1-1-2012

Phonological Complexity: The Interaction Of Three-Element Cluster And Word Lexicality On Speech Sound Disorders

Heidi Thompson

PHONOLOGICAL COMPLEXITY: THE INTERACTION OF THREE-ELEMENT CLUSTERS AND WORD LEXACILITY ON SPEECH SOUND DISORDERS

by

Heidi JoAnn Simon

Bachelor of Arts, University of North Dakota, 2010

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

May

2012

This thesis submitted by Heidi Thompson in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done, and is hereby approved.

Dr. Alycia Cummings

Dr. Manish Rami

Dr. Monika Pawlowska

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the Graduate School at the University of North Dakota and is hereby approved.

Dr. Wayne Swisher
Dean of the Graduate School

05/03/2012
Date

TITLE	Phonological Complexity: The Interaction of Three-Element Clusters and Word Lexacility on Speech Sound Disorders
Department	Communication Sciences and Disorders
Degree	Masters of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in her/his absence, by the Chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Heidi Thompson

04/01/2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

LIST OF GRAPHS

LIST OF ABBREVIATIONS USED THROUGHOUT THE THESIS

AEP .... Assessment of English Phonology

CELF-4 .... Clinical Evaluation of Language Fundamentals – 4th Edition

GFTA-2 .... Goldman Fristoe Test of Articulation – 2nd Edition

ECI .... Error Consistency Index

MOCP .... Modified Onset Cluster Probe

OCP .... Onset Cluster Probe

NW .... Non-Word

RW .... Real-Word

SSD .... Speech Sound Disorder

SSP .... Sonority Sequencing Principle

TELD – 3 .... Test of Early Language Development – 3rd Edition

ACKNOWLEDGEMENTS

I simply cannot call my thesis “complete” until I have taken a moment to express my appreciation for all of the people that have made it possible. While I, alone, set out to complete this project, it would have never came to be without the immense support that I have received in both my personal and academic life.

First and foremost, I believe that the highest word of thanks needs to be said to Dr. Alycia Cummings. Alycia has gone above and beyond her duties of a committee chair and advisor. She was always willing to take time out of her busy schedule to offer me a word of advice or simply lend a listening ear. No problem that arose throughout this process was too daunting or difficult for her. I so much appreciated her honest, wise, and encouraging words. Alycia allowed me to gain a different perspective of Speech-Language Pathology, and by doing so I have a new found respect for the field that I am about to enter. Thank you, Alycia.

I also need to voice my gratitude to my wonderful family, specifically my parents and my husband, Carson. They were always willing to offer words of encouragement when I needed it most. My mother and father always had the perfect timing to give me a phone call when I had the highest need for encouragement. They had confidence in me when I simply could not. The fact that they believed in me, without hesitation, allowed me to believe in myself. I love you Mom and Dad!

I cannot begin to express how thankful and blessed I am to have such a supportive husband. I never heard a word of complaint regarding the immense amount of time that I spent away from him in order to complete this thesis. In fact, he always encouraged me to keep working towards my goal. I am incredibly blessed to have a man of such high integrity for a husband. Carson, you are the love of my life and words can't begin to describe how thankful I am for you. This wouldn't have happened without you.

Last but not least, thank you to everyone that has given me an encouraging word throughout this past year. Especially Carson's family and my friends and academic peers; as well as anyone else that took the time to offer me a word of encouragement There have been so many of you that believed in me when I simply could not. Thank you so much.

ABSTRACT

Phonological complexity and word lexicality have been shown to induce phonological change in children with speech-sound disorders (SSDs); however, no prior research has used three-element cluster treatment target to directly compare real words (RW) to non-words (NW). Thus, the current research project sets out to answer the question of treatment efficacy in terms of real words and non-words using the complex three-element target, /str-/. Using a repeated measures single-subject research design, five children (5;7 to 7;7) were split into two SSD treatment groups (RW and NW) and directly compared in terms of learning during treatment, generalization from treatment, sound segments added to their phonetic and phonemic inventories, and error variability in their production patterns. In terms of RW and NW efficacy, the results seemed to vary depending on the measure used as the different treatments led to different effects. Overall, the RWs demonstrated a greater ability to inflict system-wide generalization.

INTRODUCTION

Treatment efficacy of speech sound disorders (SSDs) is a widely researched and debated topic and this paper will add more evidence to the literature. Olswang (1990) defines treatment efficacy as a broad term that addresses several questions related to treatment effectiveness, treatment effects, and treatment efficiency. Specifically she asks: 1) does the treatment work?; 2) in what ways does treatment alter behavior?; and 3) does one treatment work better than another? This research study will focus on the different phonological change that occurs in a sound system when presented with real words (RWs) versus non-words (NWs) in treatment while using a three-element cluster as a treatment target Both word complexity and word lexicality have proven to be efficacious in the treatment of SSDs, but the question still remains as to whether they are effective when used *together* in the treatment of SSDs. Before addressing treatment outcomes however, it is important to first understand phonology and speech sound development, as well as the basics of speech treatment.

Defining Phonological Terms. A child's phonology (that is, the grammar of sounds) describes how sounds function with his language. A phonological system includes an inventory of sounds and their features, and rules that specify how sounds interact with each other. The phonological system is often thought of as the basis for which the meaning of language emerges. An important part of understanding what a phonological system includes knowing its composition.

A *phone* can be defined as the smallest identifiable 'sound unit' that may or may not be language specific (e.g., raspberries to English speakers may be bilabial trills in another language). A child's *phonetic inventory* is a list of all of the sounds (i.e., phones) a child produces in his speech, including sounds that are produced accidently (i.e., not in the target word). Sounds included in a child's phonetic inventory are not limited to a child's native language, that is, items in a phonetic inventory can include sounds that are found in languages outside the child's target language.

A *phoneme*, on the other hand, is defined as a 'distinctive sound unit' of a language: 'unit' because the whole phoneme must be substituted to make a different word; 'distinctive' because changing a single phoneme can generate a word which is recognizably different to a speaker of the language (Coxhead, 2006). For example, the words "hat" and "cat" are considered to be minimal pairs because they differ only by a single phoneme, which makes them recognizably different to a speaker of English. Thus, a phoneme is the smallest unit that serves to change meaning between two words. A child's *phonemic inventory* is a more specific list of sounds that include phonemes that are used contrastively (Dinnsen, Chin, Elbert, & Powell, 1990). For example, the child knows that "sun" and "fun" are two different concepts, and uses /f/ and /s/ in a contrastive manner to display knowledge of that concept. The phonemic inventory is similar to that of the phonetic inventory because it may include phonemes found only in languages outside of the target language (Gierut, Dinnsen, and Elbert, 1987).

The set of phones that correspond to a single phoneme are called *allophones* of that phoneme. In other words, a phoneme is manifested as phones that can occur in different environments, and these phones are called allophones Although allophones do not change the meaning of the words they do have different phonetic representations/realizations of a phoneme, but the native listener will still judge all allophones to be the same underlying phoneme. For example, the [t] in "hat" and the [tʰ] in "tea" are allophones of /t/ in English. Allophones are also language specific, meaning that allophones in one language may mark two separate phonemes in another language (e.g., elongated vowels /ɑː/ vs. /ɑ/).

Phonemes are the individual sounds in a language, but alone they are meaningless. They begin to have meaning when they are combined together to create morphemes. *Morphemes* are the minimal meaningful units that comprise words, and eventually sentences. They are categorized as both lexical (e.g., "run", "go", "off") and grammatical (e.g., "-s" ("cats"), "-ing" ("going"). These morphemes have to be mentally stored somewhere, and that somewhere is referred to as the *lexicon*. The *lexicon* is often referred to as a mental dictionary, a list-like device in which each entry describes an arbitrary collection of linguistic properties that can describe a morpheme's purpose. For the purpose of this study, the lexicon will be discussed in greater detail in terms of word familiarity. Now that a strong foundation of phonological terms has been established, these terms will be used to describe the sound systems of children without causing confusion.

Children with speech sound disorders. The term speech sound disorder (SSD) is a relatively new term that has been introduced by the American Speech-Language Hearing Association (ASHA; 2009). This term has been adopted because it can be difficult to distinguish whether a child has an articulation disorder or phonological disorder, and very often children may present with both. ASHA refers to children with

SSDs as individuals who have difficulty in the area of articulation (making sounds) and phonological patterns (sound patterns). That being said, a SSD is not just a lack of control over articulators [1] but also a disruption in phonological acquisition. Thus, it is not just a disorder of speech production but also the mental representation of speech sounds. Because the disorder is rooted in both phonetics and phonology, determining an efficient, effective path of treatment has to be unique to that of the SSD.

Children with SSDs typically have normal social and behavioral skills, normal hearing, and cognitive abilities, but many of them may do not have age-appropriate receptive and expressive language skills. SSDs are present in approximately 16% of children 3 years of age, (Campbell, Dollaghan, Rockette, Paradise, Feldman, Shriberg, Sabo, & Kurs-Lasky, 2003) and 3.8% of these children continue to present with a SSD at 6 years of age (Shriberg, Tomblin, & McSweeney, 1999). Language impairment often co-occurs with SSD, with an estimated 6% - 21% comorbidity rate for receptive language deficits and 38-62% for children with expressive language disorders (Shriberg & Austin, 1998). Shriberg & Austin (1998) also found that more than half of the children diagnosed with SSDs encounter later academic difficulties in language, reading, and spelling. Because of the high incidence of SSDs in children and the lifelong implications that they have both academically and socially, an effective course of treatment is absolutely necessary.

Word lexicality. While the effects complexity may have on phonological systems have been a great research interest, the effects of word lexicality on the phonological system is also gaining momentum in the research word. In the present research study,

[1] Articulators are the organs or mechanisms of speech, such as the lips, tongue, teeth, hard palate, and soft palate.

word lexicality was manipulated in terms of familiarity. Specifically, comparing the effects that real words (RWs) and non-words (NWs) in terms of creating change within phonological systems of children with SSDs. Word lexicality has many formal properties, but the one property that this study was interested in was word familiarity.

Word familiarity refers to how often a word occurs in a language. A word can be very common in both verbal and written communication, or it can occur very infrequently. Thus, a word that is familiar can be described as a word that occurs frequently in a language, while a word that is not familiar is recognized as a word that does not occur often in a language (Gierut, 2001). Research has shown that the word familiarity has an implication in phonological change in children with phonological delays (Girolametto, Pearce, & Weitzman, 1997; Gierut, Morrisette, & Champion, 1999; Morrisette & Gierut, 2002). These studies observed that low-familiarity words induce greater phonological change in terms of generalization to untreated sounds.

Word familiarity in treatment words has also been manipulated by using real words (RW) and non-words (NW). While the research on this topic is limited, there are a handful of studies that have directly compared the amount of phonological change elicited by RWs and NWs in children with SSDs (e.g., Gierut, Morrisette, & Ziemer, 2010; Cummings & Barlow, 2011). These studies have shown greater efficacious evidence toward using NWs to treat children with SSDs. Word familiarity, specifically comparing RWs and NWs, will be discussed in depth in the following chapter.

Traditional Treatment. The traditional treatment approach to SSDs is based on the assumption that earlier developing, stimulable[2] sounds were easier for the child to produce and therefore they should be targeted first (Williams, 2003). When a traditional approach is used, targets that are selected are those that are stimulable for the child and early developing. Typically, this approach is very systematic, and children are moved through the various levels only once the child has produced a high level of accuracy, often 80-90% during structured activities (Paul, 2001). The treatment follows a developmental sequence of sound acquisition, beginning with sensory training or auditory discrimination, followed by production of the isolated sound, producing the sound in syllables, then in words, in structured contexts like phrases or sentences, and finally the last level is production of the sound in spontaneous speech (Bauman – Waengler, 2008). This form of treatment was the "standard" treatment approach for children with SSD's for many years and is still used today. There have been several studies that have shown that this form of treatment may not be efficacious (Van Riper & Irwin, 1958). Van Riper & Irwin's (1958) research showed that traditional approaches only results in change for that target or within that target class[3] and did not result in widespread generalization. This means that the child is not able to produce the target sound in contexts other than the context in which it was learned.

Complexity Treatment. Traditional approaches were based on the thought that if the treatment sound was an earlier developing, stimulable sound it would be easier for the child to produce that sound. The complexity treatment approach was developed to

---

[2] Stimulability refers to a child's ability to imitatively produce a sound when modeled by the clinician (Bauman-Waengler, 2008). Stimulable sounds refer to sounds that a child is able to produce while non-stimulable sounds are the sounds a child in unable to produce.
[3] The term target class refers to sounds that have shared binary features, including sonority, voicing, place of articulation, and manner of articulation (i.e., nasals, stops, fricatives, stridents).

counteract the traditional approach to treating children with SSDs. An exact definition of complexity is difficult to find, but Thompson (2007) stated that the complexity treatment approach supports training *complex* structures to promote generalized improvement of simpler, linguistically related structures. In terms of phonology, this would mean that training a complex, later developing sound would result in generalization of simpler, earlier developing sounds.

The complexity approach draws on three different theoretical perspectives: linguistics, complexity, and learnability theory. Each of these theoretical perspectives will be discussed at length in future sections of the research paper.

Children with SSDs vary in their error patterns, but despite their differences they share commonalities in their acquisition and errors of sound production (Gierut, 2007). This perspective can be addressed when selecting target sounds for treatment. It is important to base speech sound selection on the principles that govern the properties of language. Gierut (2004) found that in contrast to the traditional approach, treatment of complex sounds show system wide change[4]. More specifically, sounds follow a hierarchy of development, meaning that some sounds are earlier-developing and other sounds are later-developing. The later-developing sounds are defined as complex sounds, as they require a higher level of articulator coordination and phonological knowledge. Gierut (2007) indicates that by exposing children to the higher ordered elements, lower or secondary elements will fall into place and generalization will occur. Several studies have examined the role complexity plays in the treatment of children with SSDs (e.g., Elbert,

[4] Phonological change may be incurred at the level of the treated target phoneme or it may be widespread, affecting the overall structure and composition of a child's sound system. When this type of phonological change occurs, it is referred to as system wide change phonological change.

Dinnsen, & Powell, 1984; Gierut, 1985; Gierut, Elbert & Dinnsen, 1987; Williams, 1981). The repeated finding across many of these studies is that a more complex treatment target promotes greater learning or change on untreated related targets throughout children's phonological systems (Gierut, 2008b).

A consonant cluster is a group or sequence of consonants that appear together in a syllable without a vowel between them. In the English language alone, there are 29 word-initial consonant clusters (Gierut, 1998). These clusters can be considered two-element or three-element clusters, based on the number of consonants that appear together without vowel separation. Word-initial consonant clusters are generally considered to be among the last sound sequences that are mastered (Barton, Miller, & Macken, 1980). Clusters, as well as single sounds, have a hierarchical order. Studies of acquisition and error patterns of three-element clusters of children acquiring English as a first language (Smit, 1993; Smit, Hand, Freilinger, Bernthal, & Bird, 1990) show that, typically, three-element clusters are among the last sound sequences to be acquired. Studies of normal phonological development suggested that children typically master consonant clusters that consist of stop + liquid segments (i.e., /bl/) before they master fricative +liquid segments (i.e., /sl/) (Ingram, 1976; Smith, 1973; Templin, 1957; Wellman, Case, Mengert, & Bradfury, 1931). Three-element clusters can be described as a single unit that consists of three sounds or "elements" For example, the cluster /str/ has three-elements: the /s/, the /t/, and the /r/. Because three-element clusters have so many components they are considered to be especially complex. Based on the complexity theory, targeting a three-element cluster may result in widespread generalization across phonological systems.

Thesis layout. Now that the general ideas of phonology, the complexity theory, and word familiarity have been established, the reader should be better prepared for the forthcoming thesis. The purpose of this research study was to examine the interaction of three-element clusters and word familiarity in the treatment of children with SSDs. Chapter One presents an in-depth overview of complexity theory, word familiarity, and previous research studies that have used three-element clusters to treat children with SSDs. Chapter Two describes the children chosen to participate in this research study, providing information about the children's eligibility requirements, phonological characteristics, and other pertinent information that may have influenced the study's outcomes. Chapter Three discusses in detail the treatment study developed for the children with SSDs. This study was developed to examine the interaction that three-element cluster have on treatment outcomes of children with a speech sound disorder, with a secondary comparison of the role that RWs and NWs play in phonological change. Chapter Four discusses the results the treatment study had on the phonological systems of the participants, and specifically compares the phonological change of the RW and NW groups. Chapter Five completes the thesis by discussing theoretical and clinical implications for this work, as well as laying out a plan for future research to address questions either unanswered or unearthed by the current research study.

CHAPTER ONE

REVIEW OF LITERATURE

1.1 Connectionist model: The Spreading-Activation Model.

In order to provide efficient and effective treatment to children with speech sound disorders, it is necessary to understand how they perceive and produce speech sounds, and to identify what factors can lead to positive phonological change. Connectionist models are commonly used to stimulate cognitive tasks (Rumelhart, McClelland, & the PDP research group, 1986), which make them an ideal tool to use when examining and describing speech and language processing in processing in young children.

Connectionist models can simply be described as programs that specify the activity and layout of many simple processing units arranged in a network (Baker, Croot, McLeod, and Paul, 2001). One prominent connectionist account detailing the relationship between phonological, lexical, and semantic representations is the spreading-activation model (e.g., 1986; Dell & O' Seaghdha, 1991). This model distinguishes between semantic units, lexical (i.e., word) units, and phonological segment units, all of which are organized into a single network. The different levels of units are connected and allow for a bidirectional spread of activation, with an activated lemma/word unit activation *both* phonological units (i.e., top-down processing) and semantic (i.e., bottom-up processing). Dell and O'Seaghdha's (1991) model was used to stimulate naming: The semantic feature layer was the input layer as naming is assumed to begin with semantic processing and the phonological segment layer was the output layer because the model's task was to choose phonological segments representing the form heard in human speech.

While Dell & O' Seaghdha (1991) did not stimulate the reverse pattern of spreading activation necessary for word perception/recognition, it is hypothesized that it would work in the same manner (Baker et al., 2001). Specifically, the phonological segments corresponding to an auditory word form would be activated first, and the network would choose the appropriate semantic features necessary for "recognition". Since this is an interactive model, the activation spreads continuously and bidirectionally, with phonological units becoming active during lemma/word access and semantic units becoming active during phonological access. As a result, word selection is influenced by phonological information, and phonological segment selection is affected by semantic and other nonphonological information (Dell & O'Seaghdha, 1991). Importantly, the spreading-activation model (e.g., Dell, 1986; Dell & O'Seaghdha, 1991) does not assume that a single phonological representation is used in both perception and production of words. Instead, since the model is interactive in both top-down and bottom-up processing, what phonological segments are selected for word production may not necessarily be the same phonological segments identified during word perception and recognition.

For the purpose of this research study, time must be spent describing both top-down and bottom-up processing. By doing so, this will hopefully shed some light on to the current research question at hand.

1.2. Top-Down Processing.

The activation levels of the individual phonological segments may vary due to their accessibility in the network. In terms of top-down processing, a child may have a strong/intact semantic representation that activates a lexical item; however, some of the phonological segments comprising the [adult] lexical representation[5] may not have any, or enough activation, in order to be selected for production. Why exactly the sound production of children with phonological disorders is impaired is not known. It is possible that the phonological segments of the lexical representation are difficult to access due to lexical properties of the word, the child has not yet mastered certain phonetic features of the sound, and/or the child has atypical phonological rules/constraints that dictate what sounds can be used in what word positions, and so forth. In any care, the child does not select certain phonetic segments and his sound production does not match his underlying lexical representation.

1.3 Lexical Representation: Two-Representation Model of Spoken Word Processing.

One formal property that has emerged as a relevant predictor of spoken word processing is word familiarity. Recall, word familiarity refers to the number of times a given word occurs in a language (Storkel & Morrisette, 2002). Words can be divided into categories of high versus low familiarity items. High familiarity words are words that occur frequently in a language, while low familiarity words are words that do not occur frequently within a language.

---

[5] A lexical representation is defined as how words are represented not only in terms of their semantics, but also in terms of their formal properties (Behrens, 2000). Lexical representation has become an added variable known to enhance phonological treatment efficacy (Gierut, 2001).

Word lexicality and familiarity may influence the perception and production of words; this may be accounted for by the two-representation model of word processing (Storkel & Morrisette, 2002). The two-representation model of word processing (Storkel & Morrisette, 2002) is also a connectionist model, although it is more specific and restricted. The model specifically addresses the phonological/lexical interaction in children with speech sound disorders. Within the two-representation model, the lexical representation corresponds to a word as a whole unit, while the phonological representation corresponds to the individual sounds or sound sequences that make up the word. In this instance, hearing or thinking about a word produced external activation to a lexical representation (Storkel & Morrisette, 2002). For example, the word "sit", the lexical representation is /sɪt/, while the phonological representations are /s/, /ɪ/, and /t/. The lexical representation may play a role in how words are perceived and produced, while the phonological representation may affect how a "spoken" word is processed. Storkel & Morrisette (2002), state that this lexical representation does not just "appear" but must reach a set activation threshold. Once it reached a certain activation threshold, the listener or speaker can become consciously aware of the representation. In other words, once that word becomes familiar to them they can actively select it to use in conversation or understand it's true meaning.

The level of baseline knowledge someone has of a word is referred to as the resting threshold. Storkel & Morrisette (2002) refer to this as the "initial level of activation of a representation before further external activation is accrued either by hearing the word or by thinking of the word". Word familiarity comes into play when determining the resting threshold of words, as words that are frequently heard or produced will have a higher resting threshold than words that are infrequently recognized or produced. This model has been used to provide insight into the complex interaction between the lexicon and phonology in development. Specifically, a word's lexicality has been shown to influence children's production accuracy of target sounds.

1.4. Real Words and Non-Words in Treatment of Children with SSDs: NW Benefits.

Recall the two-representation model of word processing (e.g., Gupta & MacWhinney, 1997; Luce et al., 2000; Storkel & Morrisette, 2002) predicts that the manipulation of the lexical and/or phonological characteristics associated with a word can enhance phonological awareness and subsequent phonological acquisitions. Taking these factors into account, words that are more frequently encountered or meaningful to a child may induce greater phonological change than using unfamiliar or non-words. Thus, it is possible that real words induce greater phonological change due to their familiar lexical nature.

However, the above factors may also hinder a child's treatment performance due to his previous experiences with a specific sound in specific words (i.e., /s/ in "sand" is produced at /t/ in "tand"). For example, a child may have a "frozen" phonological representation (Bryan & Howard, 1992) for a word, in which case the child's production of a sound in the "frozen" word might not change with treatment, even if his ability to produce the same sound in other words does improve. These "frozen" words could unintentionally and negatively affect the overall outcomes of treatment. To avoid potential biases associated with life experience and semantics (Leonard, Newhoff, & Masalam, 1980), phonotactically possible non-words have been used in phonological remediation instead (Gierut, Morrisette, Ziemer, 2010, Cummings & Barlow, 2011). By presenting children with entirely new sound combinations (i.e., nonsense words) more change in a child's sound system could occur because the child will not have had any prior experiences with the treatment words.

A study done by Storkel (2004) set out to examine the tasks required by children when they are presented with a novel word. Storkel (2004a) hypothesized that when a child encounters a novel word, he or she must store a representation of the referent, a representation of the phonological form, and an association between these two representations, and then the child must integrate these new representations with existing representations. Storkel examined the effects of word frequency (i.e., familiarity) and found that early-acquired words were higher frequency than later acquired words. She stated that this might be because of the complex representation tasks necessary for a child to properly "store" this novel word. Recall that Storkel and Morrisette (2002) stated that these tasks may prove to be too cognitively demanding for some children. Storkel (2004) provided evidence that children with phonological delays can effectively learn the association between a nonsense word and its lexical and semantic representations, thus suggesting that the use of non-words in treatment is warranted.

Research began to branch out from just using high familiarity and low familiarity words to directly comparing the use of real words and non-words in treatment of children with SSDs. The use of non-words (NW) in treatment has been use for a long time, but in ways different from current research. Van Riper (1978), for example, recommended that sounds be introduced as phonotactically-permissible CV, VC, VCV, or CVC sequences in the early phases of treatment, which often mean that NWs were treatment targets. Moreover, Leonard (1973) recruited four children (4-10 years), with phonological disorders to participate in one of two treatment conditions: RW or NW. All of the participants were treated with the target sound /s/ in the initial position of three words (*side, sign,* and *sire*). Importantly, the treated words were differentially affiliated with either pictures of novel referents (e.g., computer wires corresponded with "sign") or pictures of legitimate and corresponding referents (e.g., traffic sign corresponded with "sign"). Thus in this study, the NWs were defined by novelty in word meaning and not novelty in phonological form. Leonard (1973) found that children exposed to the NW referents required fewer trials to criterion and achieved greater proportions of accuracy than those who were exposed to the RW referents. These results suggested that treatment with NWs may offer an advantage for phonological learning, but the scope of the study was very limited as there were few participants, the phonological form was not manipulated, and the generalization sampling was from just three words

More recently Gierut, Morrisette, and Ziemer (2010) evaluated the use of non-words (NWs) as stimuli in treatment of children with SSDs, as compared with the use of real words (RWs). This post-hoc study was conducted on a large scale (60 participants), with production data being drawn from the Developmental Phonology Archive of the Learnability Project (Gierut et al., 2010). The Learnability Project served children aged 3 to 7 with functional phonological disorders. Prior to treatment, the participants' vocabulary skills were examined because children treated with NWs were basically presented with the task of acquiring novel words (Gierut et al., 2010); all of the participants achieved average or above average scores on both the vocabulary and intelligence assessments. The children's phonologies were compared on three dimensions including a standardized relational assessment and two independent analyses of the children's phonetic and phonemic inventories, with no statistical differences between the groups.

Generalization learning was determined by examining the pre-treatment probe data that had been collected for each child. Two post-treatment probes were conducted to determine generalization, one being immediately post-treatment and the other being conducted 2-months post treatment. In order to determine generalization, three dependent variables were measured in Gierut et al.'s (2010) research program. These dependent variables were learning during treatment, generalization of the treated sound in untreated words, and generalization of untreated sounds in untreated words. Learning during treatment was defined as the percentage of accuracy of production of the treated sound in the treatment words, as recorded throughout treatment sessions. The generalization was reported for treated and untreated singleton phonemes by comparing the participants' pre- and post-phonetic and phonemic inventories. A repeated measures analysis of variance (ANOVA) was used to evaluate the participants' phonological generalization when NWs versus RWs were the stimuli of treatment.

Overall, both NWs and RWs seemed to benefit the participants' phonological generalization learning, but in different degrees and time frames. In both groups, generalization to treated sounds was greater than untreated sounds and this remained constant over time. The main difference was in the generalization that occurred directly after treatment, with generalization following treatment of NWs exceeding that of RWs. The RW group eventually reached comparable levels of generalization as the NW group, but the improvements were only seen at the two-month post-treatment assessment and could not be decidedly based on the treatment itself (Gierut et al., 2010).

Gierut et al.'s (2010) results support the use of RWs in treatment, but they also support the relevance and validity of NWs as stimuli of treatment for SSDs. NWs induced greater generalization than the RWs, and this hints of enhanced NW efficacy. Another contribution is that treatment with the NWs demonstrated immediate generalization, affiliated with the treatment itself. The treatment with NWs also showed a clinical implication of being at least as good as or maybe even better than RWs in promoting rapid, system-wide phonological gains.

While Gierut and colleagues (2010) showed that the use of NWs as stimuli in treatment holds an important place in treatment efficacy, their study was not without limitations. For example, their data came from a range of different studies comprising different single-subject designs and were originally collected for other research purposes. This means that much of the data were meant for other research studies and the initial post-treatment data interpretation may have been swayed based on the original research purposes. Another important limitation is that of treatment intensity. Time in treatment could have been a limitation because children in the NW instruction received 15 treatment sessions while the children in the RW instruction received only 13. Therefore, it is possible that children treated with the RW stimuli may have rapidly advanced through steps of a treatment program while the participants in the NW stimuli did not.

While the research program conducted by Gierut et al. (2010) provided new information regarding the role that NWs have in treatment, it also left unanswered questions. Most importantly, why would the treatment with NWs lead to a quicker rate of generalization that that of treatment with RWs? Cummings and Barlow (2011) also compared RWs and NWs in treatment with children who had been diagnosed with SSDs.

Their methodologies including both treatment and assessment were nearly identical to that of previous research programs by Gierut (e.g., Gierut, 2001, 2007, Gierut et al., 2010). An important difference in Cummings and Barlow's (2011) research program is that four dependent variables were measured rather than the three dependent variables that were measured in Gierut et al., (2010). Both Gierut et al. (2010) and Cummings and Barlow (2011) examined the dependent variables of: learning during treatment, generalization of treated sound in untreated words, and generalization of untreated sounds in untreated words. What distinguished the two studies was that in addition to those three variables, Cummings and Barlow (2011) also examined the variability in sound productions for treated and untreated sounds. This is called the error consistency index (ECI) and it was designed to measure the overall consistency of error substitutions within a child's phonological system (Tyler, 2002; Tyler, Lewis, & Welch, 2003). In examining the participants' ECI, Cummings and Barlow (2011) hoped to provide some additional insight regarding what type of treatment elicited greater system-wide generalization.

Four participants were selected to participant in Cummings and Barlow's (2011) research program. Three of the four children were between 3;0 and 3;11 at the onset of treatment, whereas the fourth child was 6;9. The children were split into two groups (i.e., RW vs. NW) and were treated with target sound /r/. The treatment was delivered in two phases (e.g., imitation and spontaneous production) and 19 treatment sessions were provided twice weekly in 1-hour sessions for all children. Consistent with the study conducted by Gierut et al. (2010), Cummings and Barlow (2011) found that NWs might be better than RWs in affecting change in a sound during treatment. In terms of the error variability analysis, children who were treated with NWs had much larger reductions in the number of sound substitutions produced for their treatment target compared the children who were treated with RWs. This indicated that the NWs helped the children better establish a phonemic category for their treated sound, by limiting the number of sounds determined to be “acceptable” substitutions for the treated sound (Cummings & Barlow, 2011). In other words, phonological change was occurring and more adult-like phonological representations were being developed. Once again, the data presented in the study completed by Cummings & Barlow (2011) suggested that the effectiveness of NWs in treatment might be comparable to, if not better than, that of treatment with RWs.

1.4. Real Words and Non-Words in Treatment of Children with SSDs: RW Benefits

While the research programs described above (i.e., Gierut et al, 2010 and Cummings & Barlow, 2011) both suggest that NWs may be efficacious to treating children with speech sound disorders, other research programs have found contrary evidence, at least in terms of real words that were both high and low familiarity. Recall that real words are naturally more familiar words than non-words. NWs are can be considered to be the lowest of familiarity as they are completely novel due to their non-occurrence in English.

Using the two-representation model, Storkel and Morrisette (2002) hypothesized that the learning of non-words can be very difficult for some children. Storkel & Morrisette (2002) predicted that the learning of NWs requires an additional cognitive processing, meaning that not only do the children have to establish a phonological representation and try to produce the word’s phonological form correctly, but they also have to learn the association between the novel NW and its picture (i.e., creating a lexical representation). By applying this observation, a hypothesis can be made that the tasks demanded by novel words, or in this case NWs, can prove to be too overwhelming for some children. Therefore, when taking Storkel and Morrisette's (2002) observations into account, it can be said that speech treatment targeting RWs would benefit children more than treatment with NWs.

A study conducted by Gierut, Morrisette, and Champion (1999) examined the role of lexical variables in phonological treatment. Twelve children (3;0 to 7;4) participated in an alternative treatments designed to induce change in their productive sound items. A complex single-subject experimental design combining alternating treatments with multiple baselines was used in Gierut et al, (1999). In other words, each participant received treatment on two contrasting sounds within different experimental conditions. The experimental conditions included treatment of two sounds in both familiar words and non-familiar words. For example, Subject 1 was assigned to the experimental manipulation comparing performance following high familiarity words to that of low familiarity words. In the familiar words he was taught /s/ and in the non-familiar condition he was taught /ʃ/. When the results were compared in regards to word familiarity, treatment results showed that both high familiarity words induced significantly greater generalization learning than did treatment of infrequent words.

A study by Gierut and Morrisette (2002) was done to expand on Gierut's 1999 research study. In Gierut and Morrisette's (2002) study, eight children with functional phonological delays participants in a single-subject, staggered multiple baseline, across-subjects design. The purpose of their experiment was to examine whether word frequency (i.e., familiarity) and neighborhood density affect spoken-word production. In order to do so, the design exposed children to prior of no treatment followed by treatment. Children entered the treatment phase in succession, with the number of pretreatment baselines increasing by one as each consecutive child was enrolled. For example, Child 7 completed two baselines administered over a period of 21 days before the initial treatment. Meanwhile, Child 8 continued on in the baseline phase for an additional period of 19 days to complete a total of three baselines before beginning treatment. The children were randomly assigned to 1 of 4 treatment conditions: treatment of high-frequency, low-frequency, high-density, or low-density words. The treatment procedures were consistent with previous research programs by Gierut & Colleagues (e.g., 1985, 1996, 1999). Gierut & Morrisette's dependent variables included generalization to the treated sound in untreated words and contexts, within-class generalization to untreated sounds to untreated sounds of the same manner class as the treated sound, and across-class generalization to untreated sounds in different manner classes as the treated sound.

The results of Gierut and Morrisette's research study were consistent with Gierut et al.'s (1999) study, which was described above. In general the participants seemed to benefit most the most from using treatment stimuli that were manipulated in terms of word frequency (i.e., familiarity) than words that were manipulated in terms of word density. They once again found that high frequency words induced greater phonological gains in the child's overall sound system and that low-frequency words resulted in generalization to untreated but not treated sounds.

Thus, it seems that based on the previous research both familiar words and non-familiar words are efficacious and beneficial when treating children with speech sound disorders. Recall that real words were introduced as familiar words based on their actual occurrence in the English language and non-words are considered to be non-familiar words based on their novel occurrence in the English language.

1.5. Complexity Approach: Theoretical Rationale.

Recall that the complexity approach draws on three different theoretical perspectives: linguistic, complexity, and learnability theory. A brief overview of each of these three theories will help provide rationale for the complexity approach.

Within the linguistic framework of the complexity approach, there are two major constructs: 1) distinctive features; and 2) linguistic universals. Distinctive features underlie the concept of a phonological contrast or distinctive opposition (Gierut, 1990). For example, the words “sue” and “zoo” have different meanings because the phonemes /s/ and /z/ are distinguished by voicing. Further, just as single features can distinguish phoneme pairs, they can also be grouped into a class based on their shared features. Research has shown that not all features are equal, and that they can actually be hierarchically organized based on their level of complexity. The concept of linguistic universals plays into the complexity hierarchy in several ways.

A linguistic universal is a phonological characteristic or trait across (nearly) all languages. Linguistic universals can either be absolute in nature or a tendency across languages. For example, all languages have stops (e.g., /p/ in “pop”) so it is an absolute universal whereas liquids (e.g., /r/ in “rabbit” are a tendency because most languages have at least one of them; O’ Grady, Archibal, Aronoff, & Rees-Miller, 2005). The traits that have a “tendency” to occur across languages are considered to be marked (O’ Grady et al., 2005). Marked traits have been reported to be later developing and have been described to be more complex.

Research by Elbert, Dinnsen, & Chin, (1992) showed that children gradually acquire features to create increasingly marked phonological contrasts. Elbert et al., (1992) used a five-level distinctive feature hierarchy to account for the increasing complexity of children's phonetic inventories. Elbert et al. (1992) also found that a more marked feature at a higher level implied the existence of a less marked feature at a lower level, these are referred to as implicational laws. Examples of some of these laws include consonants imply vowels (i.e., if a child produces some consonants, then the child must also produce vowels), true clusters (e.g., all clusters except /sC-/ clusters) imply affricates (e.g., / ʤ/ and /ʧ/), affricates imply fricatives (e.g., /f/, /s/, /v/, /z/), fricatives imply stops (e.g., /p/, /t/, /d/, /g/, /k/), liquids (e.g., /l/ and /r/) imply nasals (e.g., /m/, /n/, / ŋ/) and voiced obstruents (e.g., voiced stops, fricatives, affricates) imply voiceless obstruents (e.g., voiceless stops, fricatives, affricates) (Gierut, 2007). The idea of implicational laws is important to understand because it underlies much of the complexity approach to intervention. We now delve into the concept of complexity in order to gain a broader understanding of the complexity treatment approach.

1.6. Complexity from Different Perspectives.

Though complexity has been used in different disciplines, it is still remains difficult to pinpoint a comprehensive definition of complexity (Gierut, 2007). Rescher (1998) made an important contribution to this pressing issue, by defining three "modes" or ways of conceptualizing complexity from epistemic, ontological, and functional perspectives.

The epistemic perspective view of complexity is the most elementary because it simply refers to only the description of a system. Complexity at this level begins with the assessment process where a child's phonological system is analyzed and then described (Gierut, 2007). For example, in the present study, children's phonological systems were analyzed by collecting their phonemic and phonetic inventories, percent consonants correct (PCC), and their sound substitution patterns. All of these assessment measures were obtained in order to gain an accurate description of their disordered sound systems. This description, in turn, determined whether or not the children were suited for the study. Thus, complexity from an epistemic perspective allows clinicians to determine treatment goals and to gain further insight into the general outline of treatment.

Complexity from an ontological perspective refers to the constituent elements of a system and their hierarchical organization (Gierut, 2007). In terms of speech and language development, hierarchical organization is key because all domains of language are made up of elements in a hierarchical fashion. This perspective of complexity is especially key in terms of phonological learning and generalization. The idea is that if the "controlling element" is mastered, it will unlock all of the elements in the system (Rescher, 1998). Clinically, this can be used to select target sounds. Gierut (2007) states that through clinical application of the ontological perspective, we can expect that by exposing a child to higher ordered elements, other related lower ordered elements will fall into place and generalization will occur. For example, based on previous research findings (Gierut & Champion, 2001) targeting true clusters (e.g., /tr/ and /bl/) will imply affricates (e.g., /ʧ/ and /ʤ/.

Sounds can be labeled high ordered or low ordered based on their markedness and distinctive features. Gierut (2007) gives the example of a unidirectional relationship of phonological properties. If a child's grammar has the phonological property X, then that child's grammar will also have to have property Y, but not vice versa. X is the governing variable, and it is therefore a higher order and more complex than Y. Gierut (2007) goes on to refer to X as the *marked* property and *Y* the unmarked. The co-occurrence relationship between X and Y is described as markedness. Using sounds in the English language as an example, the sound class of fricatives (e.g., /s/ in "see" and /z/ in "zoo") is unmarked in comparison to the sound class of affricates (e.g., /ʧ/ or "ch" in "chew"). However, when examining the relationship between fricatives and stops, the class of fricatives is marked in relation to stops. This is not because of the individual features such as anterior or stridency[6] that make fricatives marked, but rather the combination of all the features in relation to the combination that make up the class of stops. Therefore, the ideas of high- and low-order elements are based on a sound's markedness and distinctive features.

We can now discuss learnability theory, and the role it has to play in the justification of the complexity approach. Learnability theory strives to formulate the possible ways that language can be learned from the input of the surrounding speech community (Gierut, 2007). Basically, what this implies is that what a child knows about a phonological system is a subset of what an accomplished speaker of the language knows. So, for a child that is developing language he/she must learn more and more about the phonological system and the complexities of the system so they can one day have a phonological system that matches an accomplished speaker. Gierut (2007) uses this basic

[6] Anterior sounds are those that are produced in the front of the mouth (e.g., /p/, /b/, /m/) and stridency can be described as speech sounds produced by forcing air through a constricted passage (e.g., /s/, /z/, /ʃ/).

tenet of learnability to suggest that when the input is beyond or outside a child's existing knowledge (i.e., complex input), the greater the phonological learning.

Recall that sounds can be organized on a hierarchy scale, ranging from the least complex to the most complex. The following figure represents the development of sounds in children, with the least complex sounds are called the "Early 8" and the most complex sounds being the "Late 8".

Early 8 Sounds	/m/	/b/	/j/	/n/	/d/	/p/	/h/	/w/
Middle 8 Sounds	/t/	/k/	/g/	/f/	/v/	/ʧ/	/ʤ/	/ŋ/
Late 8 Sounds	/ʃ/	/s/	/ɵ/	/ð/	/r/	/l/	/z/	/ʒ/

Figure 1.1. Typical speech sound development in children. Adapted from "Shriberg's Order of Speech-Sound Acquisition" (1992).

Complex target characteristics continue to be of great interest in determining treatment efficacy, with a subset of these research studies examining how consonant clusters help reorganize the phonological systems of children with SSDs.

1.7 Consonant Clusters and Generalization.

Powell and Elbert (1984) investigated whether later developing fricative + liquid clusters would exhibit generalization patterns that differed from those taught to produce earlier-developing stop + liquid clusters. Six children with speech sound disorders (4;4 to 6;3) were selected to participate in this study. Using a multiple-baseline-across-subjects design, the participants were matched in pairs based on their age and predominant cluster error pattern during baseline. One group of children were taught to produce the earlier developing stop + liquid clusters (i.e., /tr/, /bl, /gr/ while the other group of children were taught to produce the later developing fricative + liquid clusters (i.e., /fr/, /sl/, and /ʃr/.

The participants exhibited different learning patterns, regardless of which cluster was targeted during treatment. Powell & Elbert (1984) found that generalization occurred to both the treated and untreated categories. This finding meant that there was an incidence of system-wide generalization, which laid the foundation for future studies concerning the interaction that consonant clusters may have in treatment.

Powell & Elbert (1984) studied the role that earlier and later developing clusters may have on generalization. Having found that system wide generalization occurred after using clusters as treatment targets, the question arose as to whether three-element clusters would yield the same results. In order to discuss the interaction that three-element clusters may have on the treatment of speech sound disorders, it is important to first understand their representational structure and children's acquisition of three-element clusters.

1.8 Interaction between consonant clusters and sonority.

The syllable is thought to consist of two primary constituents: the onset and the rhyme. The rhyme can be further broken down into the nucleus and the coda. The onset makes up the initial consonant sounds of the syllable (i.e., all consonants occurring pre-vocalically in a syllable). The nucleus is composed of the vowel, as well as some syllabic consonants (e.g., /m/, /n/, and /l/). The coda contains the consonant sounds that follow the vowel within the syllable (i.e., the consonants occurring post-vocalically in a syllable). All syllables must have a nucleus; however, the presence of an onset or coda is optional as in CV (consonant-vowel) or VC (vowel-consonant) syllable structures (Gierut, 1999). For example, in the word "cat" /k/ is the onset, /æ/ is the nucleus, and /t/ is the coda.

Syllables can be divided into the categories of “simple” and the “complex”. An example of a simple syllable would be a CV (consonant-vowel) word such as “to” where the onset and rhyme constituents contain a single segment each (Barlow, 2004). Barlow (2004) stated that syllables become more complex when the onset constituent branches into two or more constituents. An example of this is when “to” becomes “true” which contains two consonants or branches (/t/ and /r/) in the onset In terms of syllable complexity, Barlow (2004) notes that the relationship between branching and non-branching onsets is crucial in describing the relationship between clusters and singletons. For example, if a child’s sound system has the cluster /tr/ it can be implied that the child also has the singletons /t/ and /r/ as well.

Sounds cannot just be put together in any order; they have a governing system that allows permissible sequences. Understanding these permissible sequences is imperative when analyzing a child’s sound system. The Sonority Sequencing Principle (SSP) is the principle by which syllables are universally governed (Clements, 1990); it is a language-universal (i.e., across all of the world’s languages) principle that governs the permissible sequences of consonants within syllables. At its most basic level, sonority refers to the resonance property (loudness) that a speech sound possesses as compared to other sounds with the same length, stress, and pitch. The sonority hierarchy places consonants along a continuum by manner class. Using a rank-ordered system, the consonants are given a numerical ranking based on their sonority values (Gierut, 2001). Steriade (1990) placed voiceless stops (e.g. /p, t, k/ at the least sonorous end of the sonority hierarchy with vowels on the most sonorous end, as in Figure 1.2.

-Voi Stops	+Voi Stops	-Voi Fricatives	+Voi Fricatives	Nasals	Liquids	Glides	Vowels
7	6	5	4	3	2	1	0

Least sonorous ⟶ Most sonorous

Figure 1.2. Sonority Sequencing Hierarchy adapted from "Beacons of Best Practice: Sound Articulation Principles." By Jennifer Taps, 2009. Communication Sciences and Disorders Colloquium at the University of North Dakota, Grand Forks, ND. Note the "Voi" refers to "voice" (e.g., +Voi = Voiced and –Voi = Voiceless.

Using the language-universal SSP, the organization of word-initial clusters in terms of complexity is possible. The SSP states that the onsets of syllables maximally rise in sonority to the nucleus, and codas fall in sonority from the nucleus. This means that the onset constituents in a syllable must differ such that the first segment is less sonorous than the second segment. Most English clusters follow this rule, such as /tr/; however, there are a few select clusters in English violate sonority sequencing: the /s/ + stop clusters such as /sp/, /st/ and /sk/. While onset clusters should rise in sonority, /s/ + stop clusters fall in sonority.[7] The term "adjunct" has been used to differentiate the /s/ + stop clusters that violate the SSP from all other true clusters (Gierut, 2001). For the true cluster /tr/, the singletons /t/ and /r/ each occupy a timing slot and they have a branching structure at the level of the onset In comparison, for the adjunct /st/, there is no branching onset because /s/ in immediately dominated by the syllable node. Barlow (2004) explains that given the unusual patterning of /s/ clusters, it can be assumed that the /s/ is outside the onset constituent and thus is viewed as an "adjunct" to the syllable.

1.9 True Clusters vs. Adjunct Clusters.

[7] Some treatment approaches (e.g., the Cycles Approach), and many clinicians, specifically target s+stop clusters, possible to the detriment of the client.

Treatment research has been conducted to determine the differing structural representations for true versus adjunct clusters. For example, Gierut (1999) examined patterns of phonological generalization across 11 children (ages 3:2 to 7;8) providing intervention that targeted clusters reflecting different sonority difference scores[8]. For example, one participant was treated with cluster /bl-/ and another participant was treated on the cluster /kw-/. The cluster /bl-/ has a smaller sonority difference than /kw-/ because the interval between stops and liquids is less than that between stops and glides. Gierut (1999) wanted to evaluate the SSP as it applied to the acquisition of onset clusters and adjuncts by children with SSDs. The study was split into two experiments (Experiment 1 and Experiment 2), with Experiment 2 having three mini-experiments (Experiment 2A, Experiment 2B, Experiment 2C).

In Experiment 1, Gierut (1999) first tested the hypothesis that all of her participants with phonological delays followed the SSP during development. The purpose of this study was to try to gain insight into the effects that clusters had on children's phonological system. More specifically, Gierut wanted to examine whether the phonological characteristics (i.e., sonority) of clusters made an impact on the child's phonological system. For example, the occurrence and use of true cluster (i.e., any cluster other than /sC-/ clusters) implies the presence of adjunct clusters (i.e., all /sC-/ clusters), but not vice versa. For example, the occurrence and use of a /bl/ (i.e., a true cluster) would imply the presence of /st/ (i.e., an adjunct cluster), but the occurrence and use of /st/ would not imply the presence of /bl/. Gierut's (1999) experiment was unique in the

[8] Refer to Appendix 6 for a complete list of all initial English clusters and their sonority differences.

fact that it did not directly emphasize learning in the manipulation of clusters in the clinical treatment.

Six children (3;2 to 7;8) with functional phonological delays were pseudo-randomly assigned to one of two experimental conditions. In order to manipulate the SSP, the markedness of the treated cluster was varied, and within each condition the degree of markedness varied incrementally. Three children were taught more marked clusters (i.e., /kl/, /kw/, /pr/), and the three other children were taught unmarked clusters (i.e., /fl/, /fl/, /bl/). Clusters /kl-/, /kw-/, and /pr-/ are considered to be marked clusters based on their sonority differences occurring between the two consonants that comprise the cluster. In other words, the smaller the sonority difference, the more marked the consonant cluster (refer to Appendix 6). Then, within each condition, the degree of (un)markedness varied incrementally in step sizes of 1, 2, or 3. Gierut (1999) found that participants who were taught unmarked clusters (e.g., /kl/) exhibited less generalization that those who were taught marked clusters (e.g., /fl/).

Gierut (1999) then evaluated whether differential learning patterns of the children with phonological disorders reflected the structure of adjunct sequences using a multiple-baseline-across-subjects design across three mini-experiments. Experiment 2A examined the learning patterns of two children who produced no onset clusters before treatment. One child received treatment with an adjunct cluster, /sp/, and the other was treated with a true cluster /bl/. Post-treatment, the child who was treated with the adjunct cluster, showed generalization only to /sC/ clusters (i.e., /st/, /sp/, /sm/, /sn/); the child who was treated with the true cluster produced several clusters which spanned in the full range of sonority differences. Given these results, it seems as though using adjunct clusters as a treatment target will result in a limited scope of generalization (i.e., only to other adjunct clusters) while targeting true clusters will result in generalization spanning the full range of sonority differences.

Experiment 2B was unique in the fact that its two participants were children who both produced only adjunct clusters (i.e., /sC-/ clusters) pre-treatment and both children were taught true clusters. One subject was taught a relatively more marked true cluster /fl/, while the other was treated with the relatively unmarked true cluster of /pr/. Cluster /fl/ was considered to be more marked based on the smaller sonority difference between the two consonants that comprised it (e.g., 3) compared to cluster /pr-/ which has a sonority difference of 5. These children demonstrated significantly less generalization than the participant presented with true clusters in Experiment 2A. Cluster acquisition in these two children in Experiment 2B was limited to only those onsets that shared a common segmental element and those were produced with less than 20% accuracy. For example, the children taught the cluster /pr-/ learned on the embedded clusters /tr-/ and /dr-. Likewise, the child taught /fl-/ generalized only to /tw-/ and /sw-/. Gierut (1999) indicated that these findings were important, as it appeared that because these children's speech sound acquisition began with the adjunct clusters, their learning patterns were restricted to within-class generalization[9]. Thus, it would seem that because the children were exposed to and produced only adjunct clusters pre-treatment, it influenced their

[9] Recall that within-class generalization refers to generalization only occurring to sounds that are within the same class of the targeted sound.

ability to gain across-class[10] generalization when other true clusters were introduced to their sound systems.

Experiment 2C presented adjunct clusters to three children who produced true clusters pre-treatment. Subject 9 was introduced to /sp/, Subject 10 to /st/, and Subject 11 to /sk/. Similar to Experiment 2A, the treatment of adjunct clusters appeared to result in limited generalization and only within-class generalization. In particular, following treatment of /sp/, Subject 9 acquired only untreated consonant + w sequences (e.g., /kw, and /sw/) and also demonstrated overgeneralization of /s/ in onset position (e.g., /bɪt/ became /sbɪt/ or /tru/ became /stru/). For subject 10, the treatment of /st/ prompted change only in untreated consonant + r sequences (e.g., /pr, tr, br, dr/). For Subject 11, treatment of /sk/ motivated accurate production of untreated /tr/ and /dr/. The findings of Experiment 2C seemed to be accordance with the previous findings of adjuncts as being unmarked in acquisition, meaning that adjunct clusters maintain a sonority structure that is inherently different from true clusters. The evidence for this conclusion was derived by consistent within-class generalization learning, complemented by instances of persistent generalization. In other words, the generalization that did develop only occurred in clusters that were within the same class as the treated cluster and the /s/ in the treated cluster began to be inappropriately added to the beginning of words.

Gierut's (1999) findings suggest that when targeting clusters in treatment, a true cluster is preferable. It appears that while adjunct clusters containing /s/ + stop consonants elicit some change, they will be less efficient in resulting in system-wide change than true clusters (e.g., all other clusters). So, while adjunct clusters containing

---

[10] Across-class generalization is the opposite as within-class, meaning that generalization is occurring to sounds that are not in the same class of the targeted sound.

/sC/ aren't actually harmful to the children, they may result in a longer and more difficult course of treatment than true clusters. While the results that can come to be expected from using /sC-/ clusters in treatment (i.e., persistent overgeneralization and within-class generalization) were comprehensive, the question still lingers as to whether this pattern would be seen in other /s/ clusters, specifically three-element /s/ clusters.

1.10 Three-element Clusters Imply Change.

Recall that three-element clusters are among the most complex type of consonantal onsets in English (Gierut & Champion, 2001). Gierut & Champion (2001) evaluated linguistic complexity as it pertained to syllable onsets in treatment of children with phonological delays, focusing on word-initial three-element consonant clusters such as /spl-/ and /skr-/. Prior to this study, there had been a gap in the literature because there had been no reports on the effects of three-element clusters on the sound systems of children with phonological delays (Gierut & Champion, 2001).

Eight children, aged 3;4 to 6;3, with phonological delays were recruited to participate in this study. All of the participants scored below the 6^{th} percentile on the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1986), had normal hearing, oral motor skills, nonverbal intelligence and age-appropriate receptive/expressive language. In addition this to criteria, participants had to display a reduced consonant inventory with a minimum of five excluded sounds as well as the absence word-initial target clusters. The participants were presented with detailed probes to sample their spontaneous productions of ambient singletons as well as two and three-element clusters. The probes were based on prior studies (Gierut, 1998; Gierut, Elbert, & Dinnsen, 1987).

The data was then analyzed to compile the children's phonemic and phonetic inventories in order to determine their phonological characteristics.

A staggered-multiple-baseline across-subjects design was used, with the participants serving as their own control. A baseline period was required to gain stability in production of onset clusters as sampled on repeated administrations of the probes given during pre-treatment assessment sessions (Gierut & Champion, 2001). Treatment involved the production of the target three-element cluster in the word initial position in 16 non-words (NWs) presented in a storybook format. Two children were treated on each cluster (/spl/, /skr/, /spr/, or /skw/). These clusters were chosen based on Gierut and Champion's (2001) observations that in determining which 3-element cluster may be most appropriate to teach, it would be valuable to consider the child's singleton phonetic inventory. Specifically, it would be beneficial if the child's singleton inventory included a subset of the consonants of the 3-element cluster. For example, if the child's inventory included the second consonant (C2) and the third consonant (C3) of the targeted cluster (e.g., /t/ and /r/ of /str-/), then /str-/ might be a preferred cluster to target for treatment. Thus, these clusters were chosen based on the children's phonetic inventories.

The treatment of the three-element clusters proceeded in two phases: imitation and spontaneous production. The imitation phase ended when a child produced the target cluster following a verbal model with 75% accuracy over two consecutive sessions, or when 7 sessions were completed, whichever came first. The spontaneous phase was considered finished when a child produced the target cluster with 90% accuracy over three consecutive sessions, or when 12 sessions were completed, whichever came first. During the treatment, two dependent variables were measured: learning during treatment and generalization from treatment of the three-element clusters. Production accuracy was determined by the child's performance on the target cluster from day-to-day treatment sessions. Generalization was determined by measuring the accuracy production of all target clusters and accuracy of production of singletons from the pre-treatment inventory.

The use of three-element clusters in treatment resulted in similar learning patterns across the participants, regardless of the specific cluster being taught. All children increased their production accuracy from 0% baseline levels to 83-92% post-treatment levels. As for generalization, there was little or no transfer of learning to other three-element clusters, whether treated or untreated. Regardless of the accuracy level at which the participants were producing the target cluster during their last treatment session, they returned to their baseline levels following treatment.

Onset clusters lie on a hierarchy (Leo & Prinz, 1997), and it was possible that generalization would occur to structures less complex, but linguistically related. Indeed, while Gierut and Champion (2001) observed no change to the three-element clusters, generalization occurred to singletons, affricates, and two-element clusters for all children. On average, five untreated sounds, which were produced with 0% accuracy at baseline, were added to each child's phonetic inventory after treatment. Production accuracy of untreated sounds ranged from 6% to 94%, with average performance being 30%. The singletons acquired were not necessarily subsets of the treated three-element cluster. For example, five children did not have affricates in their sound system before treatment. Upon completion of treatment, four of the five children were accurately producing untreated affricates (Gierut & Champion, 2001). This is a significant finding as the treatment of three-element clusters introduced an entirely new sound class in a majority of the participants' sound systems.

In general, the findings of Gierut & Champion (2001) indicate that treatment of three-element clusters resulted in both learning and generalization. The participants displayed common patterns including broad expansion of untreated singletons, acquisition of untreated affricates, and no transfer of learning to treated or untreated three-element clusters. This study showed that treatment of complex structures involving adjunct + branching onsets trigger changes in other less complex, but implicationally embedded prerequisites. Based on this study it appears that the treatment of three-element clusters may be a great benefit to children with speech sound disorders.

1.11. Summary and Statement of Questions.

Based on previous studies, both RWs and NWs truly have a place in treatment and both may actually be appropriate when treating children with SSDs. Research has also shown that using a three-element cluster as a treatment target is efficacious and has many benefits. Thus, the current research study set out to answer one main question: When children with speech sound disorders are presented with a three-element cluster, do RWs or NWs create more change in their phonological systems?

CHAPTER TWO

PARTICIPANTS

Five children (2 female; ages 5; 7 to 7; 7), with functional speech sound disorders (SSDs) were recruited to participate in this research study. These children were recruited to participate in the study through a variety of means including referrals from the University of North Dakota Speech-Language Hearing Clinic, community out-reach programs, public announcements to area schools, child-care centers, and churches, as well as referral from the clinician's supervisor and other speech-language pathologists in the Grand Forks Public School system.

When a possible participant was referred to this research study, his or her parents were contacted via a phone interview by the clinician and her supervisor. This phone interview served as a screener to determine if the child would be eligible to participate. Children who were reported to have trouble producing speech sounds, had normal hearing, had typical language development, and appeared to be functioning at normal developmental levels were eligible to participate in further screening measures.

With written permission (Institutional Review Board consent forms approved by the University of North Dakota) from the child's parents, the clinician completed a formal battery of screening measures in order to determine if the children were eligible for this research study (see below for more details). If these criteria were not met, the child was not eligible to participate in this research study. Approximately seven children were screened and considered for this research study, and six were selected to participate. Of these six children selected to participate, five were able to meet the time requirements of the research study. If the child did meet the research study criteria, a meeting was scheduled to discuss the research study, which included the purpose, procedures, risks and benefits to the child, incentives, and time commitments that were required in participating in this research study. The five participants completed the research study in 6 to 8 weeks.

The following criteria were set for all participants in the research study. Ideally, all participants had to met all of the following criteria in order to be included in the research study.

1) All children were monolingual and had been raised in a monolingual home. This criterion was set in order to ensure that there were no language-specific differences.
2) All children had to display hearing within normal limits as determined by a standard audiometric screening (American National Standard Institute, 1991), in order to control for differences in sound perception.
3) All children were determined to not have apparent organic anomalies of the speech mechanism, as determined by an oral motor examination protocol developed by Robbins and Klee (1987). This controlled for motoric involvement with speech difficulties.
4) All children had normal development as determined by a case history interview with the parents. This controlled for general cognitive or motoric involvement with speech difficulties.

5) All children had to have a minimum percentile rank of 15% on a standardized language assessment. The Clinical Evaluation of Language Fundamentals – 4th Edition (CELF-4; Semel, Wiig, & Secord, 2004) and/or the Test of Early Language Development – 3rd Edition (TELD: 3; Hersko, Reid, and Hammel, 1999) were the standardized language assessments used in this research study. The participants had to complete one test or another, but not both, to be enrolled in this research study. This controlled for other linguistic involvement with speech difficulties (see below for more details).

2.1 Standardized Language Test Protocols. In order to assess the children's expressive and receptive language skills the CELF-4 (Semel, Secord, & Wiig, 2004) and/or the TELD-3 (Hersko, Reid, and Hammel, 1999) were administered to four of the participants. Four of the CELF-4 subtests were administered to gather a core language score. The core language score is a measure of general language ability that quantifies a student's overall language performance. It is derived by summing the scaled score from the subtests that best discriminate typical language performance (Pearson Education Inc., 2008): concepts and following directions, recalling sentences, word structure, and formulating sentences. The TELD-3 was another test that was used to assess the children's expressive and receptive skills. The TELD-3 is comprised of two subtests: the Receptive Language Subtest and the Expressive Language subtest. These two subtests are combined to form an overall Spoken Language score. All of the children in this study showed expressive and receptive language difficulties. However, three of the children were able to achieve a core language index score higher than 15%, which is consistent with the requirements of this research study.

As indicated above, two of the participants performed below the 15th percentile on the above-mentioned measure. Participant SSD16_HJ received a sub-test scaled score lower than the 20 on the CELF-4 (1st percentile). This participant also had the lowest phonetic and phonemic inventory. Participant SSD34_TP also received a percentile rank of less than 15% on the CELF-4. It should be noted though, that this participant was very shy and was very reluctant to participate in the language assessment task. Throughout the treatment study as TP began to open up, his language skills did not reflect his low score on the standardized assessment. It has been reported that up to approximately 60% of children with SSDs also have some other type of language disorder (Shriberg & Austin, 1998). Because of this, having children in the sample with a possible concomitant language delay fits the overall profile of children with SSDs.

In addition to minimum entry criteria, specific characteristics were required of the children's phonological system.

1) Following Gierut et al. (2001), all children that participated in had to obtain a maximum standard score of 80 on the Goldman-Fristoe Test of Articulation (Goldman and Fristoe, 2000).

2.2 Goldman-Fristoe Test of Articulation.

The *Goldman-Fristoe Test of Articulation – 2* (GFTA-2; Goldman & Fristoe, 2000) was administered to assess the production of the participants' speech sounds. The GFTA-2 (Goldman & Fristoe, 2000) provided information about the child's articulation ability by sampling spontaneous production of sounds in specific words. The sounds-in-words sub-section of the GFTA-2 (Goldman & Fristoe, 2000), which is norm-referenced based on age and sex, was administered by showing the child a picture and having him spontaneously name it.. Fifty-three items were presented to the participants to name. The child's responses were narrowly transcribed using the International Phonetic Alphabet (IPA) and then compared to an adult production (i.e., a relational analysis) in order to assess the accuracy of sound productions.

2) Following Gierut et al. (2001), all children had to have a reduced consonant inventory, excluding a minimum of four target English sounds from their respective phonemic inventories and a minimum of two target English sounds from their respective phonetic inventories. Sounds excluded from the phonemic and phonetic inventory were identified based on extensive phonological probe measures (adapted from Gierut, 1985) using the *Assessment of English Phonology* (AEP; Barlow, 2003; Appendix 1).

2.3 Assessment of English Phonology.

The AEP (Barlow, 2003) sampled all English sound in five different exemplars in each word position (when appropriate: initial, intervocalic, and final) in both mono- and poly-morphemic forms. All phonemes in every word were examined for accuracy relative to how they were produced in ambient adult language. The probe was used to elicit English words, diminutives (i.e., "soap" and "soapi"), and progressives (i.e., "laugh" and "laughing"). Each participant's spontaneous word productions were elicited using an electronic picture-naming task and were digitally recorded. Highly trained transcribers used the IPA to narrowly transcribe all speech samples. Once the samples were transcribed, they were used to determine the participants' context-free phonetic inventory, phonetic inventory by context, and their phonemic inventory.

The context-free phonetic inventory was determined by observing each of the participant's production of the target words. When each consonant sound occurred, it was tracked on a spreadsheet developed to organize each phoneme occurrence. For this analysis, each phoneme with diacritics was treated as distinct from those without; that is, each allophone was individually tracked. For example, a participant may have had the phoneme /p/ and its aspirated allophone, [p^h], in his phonetic inventory. These productions were treated as two separate sounds for the context-free phonetic inventory. When a sound occurred at least twice, it was determined to be a part of the context-free phonetic inventory, regardless of context.

In order for a sound to be included in the phonetic inventory by context, the desired sound had to be produced twice in all appropriate word-position contexts (i.e., initial position, final position, and intervocalic medial position). If a person wants to be more conservative while assessing a child's phonetic inventory, the desired sound needed to be produced at least twice in all appropriate word positions. The sound-in-context analysis helps identify when and where children may have a rule or constraint against a certain sound in a certain word position. Both forms of the phonetic inventories (i.e., context-free and by-context) were used to help show evidence of system-wide generalization. Table 2.1 represents the participants' pre- and post-treatment context-free and by-context phonetic inventories.

Table 2.1. Participants' pre-treatment phonetic inventories. The context-free phonetic inventories are shown in black while the phonetic inventories by context are shown in red.

Child	Condition	GFTA	Phonetic Inventory Pre- and Post- Treatment																							
PD16_HJ	Nonsense	2%	pre	m	n	ŋ	p	b	t	d	k	g	f	v			s	z	ʃ	ʧ	ʤ	w	j	h	l	r
			pre	m	n	ŋ	p	b	t	d		g	f									w	j		l	
PD22_RM	Real	>1%	pre	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s̪[11]					w	j	h		
			pre	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s̪					w	j	h		
PD34_TP	Real	5%	pre	m	n	ŋ	p	b	t	d	k	g	f	v	θ		s		ʃ	ʧ	ʤ	w	j	h	l	r
			pre	m	n		p	b	t	d	k	g	f	v			s					w	j	h	l	
PD39_AW	Nonsense	7%	pre	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	ʧ	ʤ	w	j	h	l	
			pre	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	ʧ	ʤ	w	j	h		
PD42_BS	Nonsense	8%	pre	m	n		p	b	t	d	k	g	f	v			s	z	ʃ	ʧ	ʤ	w	j	h	l	r
			pre	m	n		p	b	t	d	k	g	f									w	j	h	l	

[11] Note that /s̪/ was shown in RM's phonetic inventories regardless of accuracy. This is consistent with the method derived from Gierut (1985) in that each phoneme with diacritics was treated as distinct from those without; that is, each allophone is individually tracked.

The guidelines outlined by Dinnsen, Chin, Elbert, and Powell (1990) were used to determine the participants' phonemic inventories. To determine the phonetic inventory, the clinician first referred to the child's context-free phonetic inventory of consonants. In order for a consonant to be in both the phonemic and phonetic inventory, the sound must be contained within two minimal pairs (Table 2.2). For example, /m/ was in PD42's phonemic inventory because it occurred in two minimal pairs, /mɑɪ, pɑɪ/ and /mɑp, tɑp/. It is important to note that a sound might occur in a child's phonetic inventory but not the phonemic inventory because the child might not have produced any minimal pairs for a given sound. Phonetic inventories will nearly always be larger than phonemic inventories, especially with children who have SSDs.

Table 2.2: Participants' Pre-Treatment Phonemic Inventories

Child	**Condition**		**Phonemic Inventory Pre- and Post- Treatment**																				
PD16_HJ	Nonsense	Pre	m	n	ŋ	p	b	t	d	k	g	f				s		ʧ		w	j	h	l
PD22_RM	Real	Pre	m	n		p	b	t	d	k	g	f		ɵ	ð					w	j	h	
PD34_TP	Real	Pre	m	n		p	b	t	d	k	g	f	v			s			ʤ	w	j	h	
PD39_AW	Nonsense	Pre	m	n	ŋ	p	b	t	d	k	g		v			s	ʃ	ʧ	ʤ	w			
PD42_BS	Nonsense	Pre		n		p	b	t	d	k	g	f								w	j	h	

2.4 Modified Onset Cluster Probe.

In addition to the AEP (Barlow, 2003), a second probe, the modified Onset Cluster Probe (MOCP) was elicited in order to sample the participants' spontaneous productions of 2- and 3-element clusters (Appendix 2). The MOCP was based on Gierut's (1998) Onset Cluster Probe (OCP); minimal changes were made to Gierut's OCP in order to better fit this study. Specifically, the real words used as treatment targets in the present study were originally included in OCP (Gierut, 1998). In order to ensure that the post-treatment changes were due to system-wide change rather than familiarity, the original OCP /str-/ words were discarded and replaced with different words beginning with the treatment target three-element cluster, /str/. The 146-item MOCP sampled all target English two- and three-element clusters in word-initial position in a minimum of five different exemplars. To follow Gierut's (1998) work, the words of the MOCP probe were independently randomized and assembled in an electronic slideshow format so that the children could spontaneously name each picture. The children's responses were digitally recorded and then narrowly transcribed using IPA by highly trained transcribers.

From the MOCP, a cluster substitution analysis was performed in order to determine the participants' cluster substitution patterns. For each target word, the participants' exact production of the word was recorded and transcribed using the IPA to track any cluster distortions or substitutions. After the transcription was finished, each production of the target words was analyzed for variability. These variations of the target cluster were noted and a substitution percentage for the target cluster was obtained, as well as an overall substitution percentage for all other untargeted clusters.

The following information supplies a detailed profile for each of the 5 participants, including their phonetic and phonemic inventories, cluster substitutions, their results on the CELF – 4 (Semel et al., 2004), and other factors that may have impacted the effectiveness of the treatment.

2.5. SSD34_TP Profile.

TP, a five-year, eight-month male, was assessed over three meetings, in which several formal and informal assessment measures were administered to him in order to determine his eligibility to participant in this study. He had received speech-language services for approximately one-year prior to his enrollment in this study. He was the youngest child of a family of five brothers, and had been previously enrolled in a Head Start program for two years upon entry into kindergarten. He was a compliant child, who was willing to participate in all of the assessment measures with minimal redirection.

2.5.1. Speech.

TP was administered the GFTA-2 (Goldman and Fristoe, 2000). Of the possible 77 target consonant sounds, he produced 32 sounds in error. The table below lists the sounds produced in error. The word positions in which they were produced incorrectly are marked with an “X”.

Table 2.3. SSD34_TP's sounds produced in errors during the GFTA-2.

Sound symbol	Letter and example	Word Initial	Word Medial	Word Final
/ɵ/	"th" in "thing"	X	X	X
/ð/	"th" in "these"	X		NA
/s/	"s" in "sun"	X		
/z/	"z" in "zipper"		X	X
/ʃ/	"sh" in "shoe"		X	X
/ʧ/	"ch" in "cheese"	X	X	X
/ʤ/	"j" in "jack"	X	X	X
/r/	"r" in "ring"	X	X	
/l/	"l" in "look"			X

The following sound blends were also produced in error in word initial position: "bl", "br", "dr", "fl", "gl", "gr", "kl", "kr", "kw", "sl", "sw" and "tr". His raw score converted to a standard score of 66, which placed him in the 5th percentile for children his age. These results complied with the research study's minimum criteria, which stated that the participants should have a maximum standard score of 80.

2.5.2. Language.

TP was administered the CELF-4 (Semel et al., 2004) in order to assess his receptive and expressive language development. He was compliant throughout the assessment, but he was apprehensive to interact with the clinician. He required several communication invitations before responding to the clinician. Thus, it is possible that his score may reflect his introverted personality upon first meeting with the clinician.

The participant's raw scores (i.e., the number of items he got correct), scaled scores (on a scale from 1-19), and percentile scores (50th percentile is typical) for each of the subtests are presented below.

Table 2.4. SSD34_TP's raw scores, scaled scores, and percentile scores for the standardized language measure.

Subtest	Raw Score	Scaled Score	Percentile Score
Concepts & Following Directions	20	8	25
Word Structure	10	5	5
Recalling Sentences	23	7	16
Formulated Sentences	0	1	<1

The sum of the subtest scaled scores was 21, which converted to a standard score of 72, placing the participant in the 3rd percentile for children his age. While the score earned on the CELF-4 (Semel et al., 2004) suggests that he had a moderately-severe receptive and expressive language impairment, it was felt that TP's reluctance to interact with the clinician during the initial assessment meeting may have directly influenced his results on the standardized language measure. Thus, he was still included in this research study.

2.6. SSD42_BS Profile.

BS, a 5-year, 10-month old male was assessed over three meetings, in which several formal and informal assessment measures were administered to him in order to determine his eligibility to participant in this study. When he enrolled in this research study. BS had received approximately 12 weeks of speech-language services prior to enrolling in this study. He was the oldest child in his family, with two younger brothers aged 3-years and 3-months. BS attended the Grand Forks Head Start study for a year prior to this study. Upon the first baseline session it was evident that he was a busy child and he did not show much enthusiasm for attending the speech sessions.

2.6.1. Speech. BS was administered the GFTA-2 (Goldman and Fristoe, 2000). The word positions in which they were produced incorrectly are marked with an "X".

Fifty-three items were presented to BS to label. Of the possible 77 target consonant sounds, he produced 27 sounds in error.

Table 2.5. SSD42_BS's sounds produced in errors during the GFTA-2.

Sound symbol	Letter and example	Word Initial	Word Medial	Word Final
/d/	"d" in "dog"			X
/ɵ/	"th" in "thing"	X	X	X
/s/	"s" in "sun"		X	X
/z/	"z" in "zipper"			X
/ʃ/	"sh" in "shoe"		X	X
/r/	"r" in "ring"	X		
/l/	"l" in "look"	X		X

The following sound blends were also produced in error in word initial position: "bl", “dr”, "fl", “fr”, "gl", “gr”, "kl", “kr”, "kw", “pl”, "sl", “sp”, “st”, "sw", and "tr". His raw score converted to a standard score of 72, which placed BS in the 8th percentile for children his age. These results were consistent with the research studies requirements for eligibility.

2.6.2. Language. BS was administered the CELF-4 (Secord et al., 2004) in order to assess his receptive and expressive language development. The participant required several instances of redirection and promise of reward in order to complete this assessment measure. Thus, it is possible that his score may reflect his inability to attend to the test stimuli and clinician's directions. The first four subtests of the CELF-4 (Secord et al., 2004) were administered to BS to attain a core language index score. The following table represents his performance in the four subtests.

Table 2.6. SSD42_BS's raw scores, scaled scores, and percentile scores on the standardized language measure.

Subtest	Raw Score	Scaled Score	Percentile Score
Concepts & Following Directions	28	11	63
Word Structure	17	8	25
Recalling Sentences	28	8	25
Formulated Sentences	2	3	1

Recall that the four individual subtest scaled scores are combined to form the core language score. The sum of the subtest scaled scores was 30, which converted to a standard score of 85, placing BS in the 16th percentile for children his age. This assessment suggests that he had receptive and expressive language skills that were at the low end of the normal range. It is important to note that this score fell within the required criterion to meet for expressive and receptive language established by previous research studies by Gierut (2001, 2007).

2.7. SSD39_AW Profile.

AW, a seven-year, one-month old female, was assessed over three meetings, in which several formal and informal assessment measures were administered to her in order to determine her eligibility to participant in this study. When enrolled in this research study, AW had received two-years of speech-language therapy. She had two older sisters and two younger brothers; both of her older sisters had received speech therapy in the area of speech-sound production. While she had received extensive speech therapy she seemed to be indifferent to her speech-sound production compared to her peers. She also demonstrated frustration during stimulability probes, simply saying this it was "too hard" for her. She seemed to be unaware as to what she had previously worked on with her school clinician, saying that they just got to play a lot of games.

2.7.1. Speech. The GFTA-2 (Goldman & Fristoe, 2000) was administered to formally assess AW's production of speech sounds. Fifty-three items were presented to label. Of the possible 77 target consonant sounds, she produced 12 sounds in error. The table below lists the sounds produced in error; the word positions in which they were incorrect are marked with an "X".

Table 2.7. SSD39_AW's sounds produced in errors during the GFTA-2.

Sound symbol	Letter and example	Word Initial	Word Medial	Word Final
/z/	"z" in "zipper"			X
/ʃ/	"sh" in "shoe"	X		
/r/	"r" in "ring"	X	X	X

The following sound blends were also produced in error in word initial position: "br", "dr", "fl", "fr", "gr", "kr", and "tr". Her raw score converted to a standard score of 86, which placed AW in the 7th percentile for children her age. The results of AW's GFTA converted to a standard score of 86, which actually is above this research study's maximum standard score of 80. Regardless, AW's percentile rank was well within the confines of the requirements so she still qualified to participate in the treatment study.[12]

2.7.2. Language. AW was administered the CELF-4 (Secord et al., 2004) in order to assess her receptive and expressive language development. AW completed the assessment measure without redirection, and was attentive to the clinician's directions. The first four subtests of the CELF-4 (Secord et al., 2004) were administered to her in order to attain a core language index score. The following table represents AW's performance in the four subtests.

[12] This is because the GFTA (Goldman & Fristoe, 2000) does not actually follow a bell curve. Thus, older children with a higher standard score may actually have a low percentile score.

Table 2.8. SSD39_AW's raw scores, scaled scores, and percentile scores on the standardized language assessment.

Subtests	Raw Score	Scaled Score	Percentile Score
Concepts & Following Directions	31	7	16
Word Structure	24	8	36
Recalling Sentences	30	19	99
Formulated Sentences	6	1	<1

Recall that the four individual subtest scaled scores are combined to form the core language score. The sum of the subtest scaled scores was 35, which converted to a standard score of 93, placing AW in the 32^{nd} percentile for children her age. This assessment suggested that AW had receptive and expressive language skills that are typical of children her age, and her score fell above the minimum score requirement of this study.

2.7. SSD22_RM Profile.

RM, a seven-year, seven-month old female, was assessed over three meetings, in which several formal and informal assessment measures were administered to her in order to determine her eligibility to participant in this study. When enrolled in this research study, RM had received three-years of speech-language therapy. She had also participated in two prior speech treatment research protocols enveloped within the greater speech treatment study that the present research protocol was also housed. RM had one older brother and one younger brother upon entry to this research study. While she had received extensive speech therapy she still exhibited several speech sound errors, and was not stimulable for many sounds including most affricates and strident sounds (e.g. /s, ʃ, ʧ, ʤ/). RM was agreeable during speech-language sessions, but was not motivated to produce the target sounds and had made very minimal progress.

2.7.1. Speech. The GFTA-2 (Goldman & Fristoe, 2000) was administered to formally assess RM's production of speech sounds. Fifty-three items were presented to label. Of the possible 77 target consonant sounds, she produced 26 sounds in error. The table below lists the sounds produced in error; the word positions in which they were incorrect are marked with an "X".

Table 2.9. SSD22_RM's sounds produced in errors during the GFTA-2.

Sound symbol	Letter and example	Word Initial	Word Medial	Word Final
/ɵ/	"th" in "thing"		X	X
/s/	"s" in "sun"	X	X	X
/z/	"z" in "zipper"	X	X	X
/ʃ/	"sh" in "shoe"	X	X	X
/ʧ/	"ch" in "cheese"	X	X	X
/ʤ/	"j" in "jack"		X	X
/r/	"r" in "ring"	X		
/l/	"l" in "look"	X	X	X

The following sound blends were also produced in error in word initial position: "bl", "br", "dr", "fr", "gr", "kl", "kr", "pl", "sl" and "tr". Her raw score converted to a standard score of 58, which placed RM in the 1st percentile for children her age. These results complied with the research study's minimum criteria, which stated that the participants should have a maximum standard score of 80.

2.7.2. Language. RM was administered the TELD-3 (Hresko et al., 1999) in order to assess her receptive and expressive language development. RM completed the assessment measure without redirection, and was attentive to the clinician's directions. The expressive and receptive language subtests of the TELD- 3 (Hresko et al., 1999) were administered to her in order to attain a speech language quotient. The following table represents RM's performance in the two subtests.

Table 2.10. SSD39_RM's raw scores, scaled scores, and percentile scores on the standardized language assessment.

Subtest	Raw Score	Scaled Score	Percentile Score
Receptive Language	28	93	32
Expressive Language	34	102	55

Recall that the two individual subtest scaled scores are combined to form the spoken language quotient core language score. The sum of the subtest scaled scores was 195, which converted to a standard score of 97, placing RM in the 42nd percentile for children her age. This assessment suggests that RM had receptive and expressive language skills that are typical of children her age, and her score fell above the minimum score requirement of this study.

2.6. SSD16_HJ Profile.

HJ, a 5-year, 7-month old male was assessed over three meetings, in which several formal and informal assessment measures were administered to him in order to determine his eligibility to participant in this study. When he enrolled in this research study, he had also participated in four prior speech treatment research protocols enveloped within the greater speech treatment study that the present research protocol was also housed. He had an older sister and was attending a half-day preschool study through the local school district two days a week. HJ was a busy child who enjoyed interacting with his clinician. Due to his mother's busy schedule, his speech-language services were conducted at his day care provider's home rather than at the University of North Dakota Speech-Language Hearing Clinic.

2.6.1. Speech. HJ was administered the GFTA-2 (Goldman and Fristoe, 2000). The word positions in which they were produced incorrectly are marked with an "X". Fifty-three items were presented to HJ to label. Of the possible 77 target consonant sounds, he produced 40 sounds in error.

Table 2.11. SSD16_HJ's sounds produced in errors during the GFTA-2.

Sound symbol	**Letter and example**	**Word Initial**	**Word Medial**	**Word Final**
/ɵ/	"th" in "thing"	X	X	X
/s/	"s" in "sun"	X	X	
/z/	"z" in "zipper"	X	X	X
/ʃ/	"sh" in "shoe"	X		
/ʧ/	"ch" in "cheese"	X	X	
/ʤ/	"j" in "jack"	X	X	
/k/	"k" in "king"	X		
/l/	"l" in "look"			X
/f/	"f" in "fun"	X	X	
/v/	"v" in "van"	X		X
/t/	"t" in "tub"	X	X	

The following sound blends were also produced in error in word initial position: "bl", "dr", "fl", "fr", "gl", "gr", "kl", "kr", "kw", "pl", "sl", "sp", "st", "sw", and "tr". His raw score converted to a standard score of 56, which placed HJ in the 2nd percentile for children his age. These results were consistent with the research studies requirements for eligibility.

2.6.2. Language. HJ was administered the CELF-4 (Secord et al., 2004) in order to assess his receptive and expressive language development. The participant required several instances of redirection and promise of reward in order to complete this assessment measure. Thus, it is possible that his score may reflect his inability to attend to the test stimuli and clinician's directions. The first four subtests of the CELF-4 (Secord et al., 2004) were administered to HJ to attain a core language index score. The following table represents his performance in the four subtests.

Table 2.12. SSD39_HJ's raw scores, scaled scores, and percentile scores on the standardized language assessment.

Subtest	Raw Score	Scaled Score	Percentile Score
Concepts & Following Directions	1	1	1
Word Structure	5	2	2
Recalling Sentences	5	2	2
Formulated Sentences	0	1	1

Recall that the four individual subtest scaled scores are combined to form the core language score. The sum of the subtest scaled scores was 6, which converted to a standard score of 44, placing HJ in the 1st percentile for children his age. This assessment suggests that he had receptive and expressive language skills that were below average compared to that of his peers. These findings are consistent with Shriberg and Austin's (1998) findings in that approximately 60% of children with SSDs also have some type of language disorder. Thus, having one child in the sample with a possible concomitant language delay fits the overall profile of children with SSDs.

CHAPTER THREE

METHODS

The simplest way to test the effectiveness of real (RW) and non-words (NW) in treatment is to directly compare the two. The present treated study followed the single-subject multiple baseline design methodology of previous treatment studies of children with speech sound disorders (Gierut et al, 2010; Gierut, 1999, 2001, 2007; Cummings & Barlow, 2011). Specifically, children with speech sound disorders were trained on a target cluster presented in RW or NW within a storybook context. The target words were presented in a storybook in order to establish lexical and semantic representations. The target words were chosen based on their ability to fit into the storybook context as well as their level of child "friendliness". The specific goal of this experiment was to test the effectiveness of RW and NW in treatment when presented within a complex three-element cluster. The treatment methodology was selected in order to gain the highest level of insight of this specific research goal as well as provide an effective and efficient means of treatment for the participants.

3.1 Treatment Design.

A single-subject multiple baseline design was used in this treatment study. The decision to use this treatment design was based on studies that have shown it to be useful in the study of treatment of communicative disorders (Connell & Thompson, 1986; Kearns, 1986; McReynolds & Kearns, 1983; McReynolds & Thompson, 1986). Single-subject designs are particularly useful for evaluating the relationships between independent and dependent variables, and for evaluating inter- and intra-subject variability, which is common among individuals with communicative disorders (McReynolds & Thompson, 1986). This design also allows for replication across subjects within a given treatment condition.

Following the procedures for this design, the participants with speech sound disorders (SSD) were randomly assigned to two treatment groups: real word (RW) or non-word (NW). These groups were evaluated independently and in combination. Every child was evaluated in a baseline period in which no treatment was provided. Each child first completed a full baseline assessment, which consisted of the Assessment of English Phonology (AEP; Barlow, 2003) and the modified Onset Cluster Probe (MOCP), which was based on a probe described in Gierut (1998). Two of the participants completed an additional baseline session prior to beginning treatment. This baseline assessment was developed by choosing 18 words from the MOCP that contained properties of the treatment target, /str-/; specifically, the selected words began with /st-/, /tr-/, and /str-/. The assumption behind the multiple baseline design is that each child's sound system will remain stable during the baseline phase and once treatment begins, any change or improvements in the child's performance can be directly attributed to the treatment itself (Gierut et al., 1999).

3.2 Stimuli.

Both types of treatment stimuli, real words (RW) and non-words (NW) were presented to the children in a storybook format. The story was selected based on the presence of a /str-/ word (e.g., stripe) in its title and because it was relatively easy to manipulate to include more /str-/ words. The selection of the RW stimuli was based on choosing the most appropriate child and story-friendly words as possible. This means that the words were selected in order to be cohesive with the story selected for this research study. While it would have been ideal to choose the RW treatment words based on word frequency, the amount of /str/ words in the English language are quite limited. According to Merriam Webster Dictionary (2011) the English language only has 630 words that begin with /str/ compared to 2601 words that begin with /st/. Because of the limited amount of words that begin with /str/, words were chosen based on their ability to fit into the story as well as being child friendly.

In the NW condition, treatment words were assigned lexical meaning via a story-retelling format, as described in various studies by Gierut and colleagues (Gierut, 1990, 1991, 1992; Gierut et al. 1999; Gierut & Neumann, 1992; see Appendices 3 and 4 for the story and stimuli used in the research study). Each treatment word was correlated with a picture from the story to establish meaning. These pictures were used to elicit the treatment words throughout every treatment session.

Every child's treatment program targeted a single three-element cluster, /str-/. Recall that Gierut and Champion (2001) suggested that in determining which 3-element cluster may be most appropriate to teach, it would be valuable to consider the child's singleton phonetic inventory. Specifically, it would be beneficial if the child's singleton inventory included a subset of the consonants of the 3-element cluster. For example, if the child's inventory included the second consonant (C2) and the third consonant (C3) of the targeted cluster (e.g., /t/ and /r/ of /str-/), then /str-/ might be a preferred cluster to target for treatment.

In the present study, all of the children's phonetic inventories had at least two of the consonants in the 3-element cluster /str-/, thus the cluster /str-/ was chosen for all children. While they all had two consonants in the selected three-element cluster /str-/, two of the participants did not have the C3 component of the cluster (i.e., /r/). These participants were still selected to participate in the study, but not having the C3 component may have negatively affected their results. This will be discussed later in the results section of this paper.

Using /str-/ as the treatment target, two children were randomly assigned to the real word treatment condition and the other three children were assigned to the non-word treatment condition, in which they were presented with five target words (RW or NW) in a storybook format. The treated three-element cluster was always presented in the word-initial position of all words. While the target cluster was the same for the two groups, the treatment targets differed in word familiarity. Recall, word familiarity is the familiarity in which a word occurs in a language. While this present research program did not formally assess the word familiarity of the RW treatment targets, they were considered "high familiarity" when directly compared to the NW, which had a non-occurrence in the English language.

Each of the NW stimuli was matched to a real word in terms of word category and phonetic word construction. For example, the real word, strawberry (/strɑbɛri/) had a matched non-word of "streeburaz" (/striburæz/). Primarily changing the vowel in each syllable of each real word created the non-words. This allowed for the words to be matched in terms of both syllable length and phonemic [consonant] construction.

The first treatment session consisted of several activities that ensured that the children understood the meaning of their selected treatment targets (RW or NW). These activities connected each word to their meaning using picture activities, talking about each of the words in order to familiarize the child with the word meaning, and by having the child re-tell the story using the target words.

3.3 Treatment Procedure

Treatment was delivered in two phases: imitation and spontaneous production. This treatment protocol was chosen in order to remain consistent with treatment protocols that have been used in previous studies by Gierut and colleagues (e.g., 1999, 2001, 2007, 2010). Treatment was provided twice weekly for 1-hour sessions, for a maximum of 12 treatment sessions. Overall, the children attended 12-14 one-hour sessions (e.g., baseline probes, treatment, treatment probes, post-treatment probes) with the minimum number of treatment sessions attended being 10 treatment sessions and the maximum number of attended sessions being 12 over the course of six to eight weeks.

3.3.1 Imitation Treatment Phase. During the imitation phase of treatment, each child repeated the clinician’s verbal model until either achieving a pre-established performance-based or time-based criterion, whichever came first. Specifically, imitation continued until a child maintained 75% accurate production of the treated phoneme over two consecutive sessions (i.e., performance-based criterion) or until five consecutive sessions were completed (i.e., time-based criterion). SSD34_TP and SSD42_BS achieved the 75% accuracy criterion, and SSD16_HJ, SSD22_RM, and SSD39_AW completed all five of the imitation treatment sessions.

While many of the treatment activities varied based on what each child enjoyed most, the imitation phase of treatment had some constants across the sessions for all of the children. Each imitation session began with the clinician reading to the child his selected treatment story. When the clinician finished her reading, the child was asked to re-tell the story back to the clinician. Some children were very elaborate in their story re-tell; other children provided very few details. The goal of this story re-tell was for the children to produce some of the target words during their re-tell. If the child did not do so, the clinician prompted the child by showing them pictures in the story and asking them questions to elicit a target word.

Once the story was completed, the next standard activity of the imitation phase was some "sound" time. This activity centered on providing direct placement and sound-shaping therapy by giving verbal, tactile, and physical cues to each child. None of the children knew the correct way to articulate all three of the sounds in the three-element cluster (i.e., where to put their tongues, how to shape their mouths, how to control their voicing and air stream, etc.); thus, the clinician and child sat in front of a mirror and practiced making each sound in isolation or in syllables. While this may seem counterintuitive to the treatment goal, it was important to establish accuracy on each phoneme before producing them in a connected cluster.

The phonemes that required the most "sound" time were /s/ and /r/. For children working on /r/, cues such as "use your growling sound", "show me all of your teeth", "keep your tongue tight" were commonly used to elicit the target sound in isolation. A tongue depressor was also used to help monitor tongue placement by placing chocolate pudding on the roof of their mouth to elicit proper tongue placement. To shape the /s/

phoneme, tactile cues proved to be beneficial for most of the children. The main cue that was used was the clinician running her hand along the child's extended arm while producing the /s/ sound or the "snake" sound, as it was referred to during treatment. Extended shaping techniques consisted of concentrating on the correct placement of the /s/ phoneme. For example, telling the children to keep their tongue behind their teeth and smile.

Another technique that was used during the imitation phases was segmenting the three-element cluster (and the rest of the word) into various groupings of separate and combination sounds. Many of the children benefitted from this segmenting the cluster phonemes into separate sounds. This was typically done by adding a vowel schwa (/ə/) in between the /t/ and /r/ or between the /s/ and the /t/. For example, instead of the child saying "straw" he said "stuh-rah"; essentially, an extra syllable was added to a word so that the children could spend more time producing each sound in the word. This was often referred to as using their "broken" sound or "elevator" sound. When it was time to produce their target words without segmentation, the children were told to "connect" their sounds and produce a "rollercoaster" sound.

3.3.2 Spontaneous Production Treatment Phase. During the spontaneous production phase of the treatment, each child produced the treated cluster without a model. In other words, the target words were spontaneously elicited by having the children name pictures, label objects, and so on. The phase of treatment continued until the child maintained either a performance-based criterion of 90% accuracy production of the treated phoneme over three consecutive sessions, or a time-based criterion of seven sessions, whichever came first. None of the children achieved 90% accuracy for three consecutive sessions; thus, they all completed all seven of the spontaneous production treatment sessions.

The structure of the spontaneous production treatment was much more flexible than that of the imitation phase of treatment. The spontaneous production sessions typically began with some “sound” time. The amount of time spent on this activity varied greatly from child to child. Typically, when the child could produce the target sound correctly in 5 consecutive productions they were then able to move on to the designed activity for the day. Drill and drill-play activities were primarily used to elicit treatment word production. Every week the clinician chose a different theme to base her treatment activities around. The following themes were used during treatment sessions: space, pirates, zoo animals, creepy crawlies, aliens, vehicles, and farm animals.

3.4 Speech Probes.

To establish the specific phonological characteristics of each child, detailed speech probes were administered following methods of previous studies (e.g., Gierut, 1998; Gierut & Champion, 2001; Gierut et al., 1987). The probes used included the Treatment Probe, the AEP (Barlow, 2003), and the MOCP. The *Treatment Probe* measured learning as it occurred during treatment sessions, while the *AEP* (Barlow, 2003) and *MOCP* were administered to evaluate changes in the child’s sound system that had occurred based on the treatment.

3.4.1. Treatment Probe. The treatment probe specifically targeted each child’s production of the five RW or NW /str-/ words specifically targeted in treatment. In any given treatment session, each child produced the five treatment words between 37 and 189 times (both in imitation and spontaneous productions). Each child’s production of each treatment word was noted, with the clinician using a +/- accuracy rating scale for each of the phonemes that made up the cluster. A rating was given to *each* of the three phonemes in the cluster, and at the end of the sessions a percentage of accuracy was calculated to measure learning across treatment sessions.

3.4.2. Assessment of English Phonology. Using the *Assessment of English Phonology* protocol (AEP; Barlow, 2003; Appendix 1), all English phonemes were assessed in five different exemplars in each viable word position (when appropriate: initial, intervocalic, and final) in both mono- and ploy-morphemic forms. All phonemes in every word were examined for accuracy relative to how they were produced in ambient adult language. This probe was administered two times during the treatment process: (1) initial baseline session and (2) post-treatment session.

3.4.3 Modified Onset Cluster Probe. Recall that the *Onset Cluster Probe (OCP*), generated by Gierut (1998), was created to target all English two- and three-element clusters in word-initial position. For the purpose of this study, Gierut's OCP was modified into the MOCP so that it did not contain the targeted treatment words, as well as the increase the number of /str-/ clusters in the probe. The purpose of this treatment probe was to determine the participants' substitution patterns for all consonant clusters. This probe was administered twice during the treatment process: (1) initial baseline session and (2) post-treatment session.

3.4.4 Baseline Probe of Treated Cluster. The baseline probe was a phonological analysis of the English two- and three-element clusters that encompassed the /str-/ treatment cluster using a subset of items of the MOCP protocol. This was meant to be an abbreviated assessment of 18 words that specifically targeted each child's production of the two- and three-element clusters that encompassed the treatment cluster. Specifically, the target clusters consisted of words beginning with the following clusters: /tr-/, /st-/, and /str-/. The target clusters were judged to be correct if produced accurately as in the ambient adult language. This probe was administered once during the treatment program to all of the children, during the phase change from imitation to spontaneous production. It was also administered to two children with SSDs (RM and HJ) who received an extra baseline session.

3.4.5. Speech Sample Recording, Transcription, and Reliability. Every assessment and treatment session was digitally recorded with a Sony Digital Voice Recorder. Each audio sample was then digitally transferred to a PC laptop for transcription using a speech recording computer program, Adobe Audition®. All speech samples were backed-up on external hard drives and stored in a separate location. Highly-trained transcribers used the International Phonetic Alphabet (IPA) to narrowly transcribe all speech samples. Based on these transcriptions, the data were organized for phonological analyses according to target sound and word position. Approximately 90% of the speech samples were reliability-checked by a second transcriber; both transcribers agreed at least 85% of the time on the speech sample transcription. If this threshold was not reached, the speech sample in question was re-transcribed until two transcribers reached the designated threshold. Overall, the transcriber reliability was 90%.

3.4.6. Analysis and Interpretation of the Data. Given that all of the treatment results were calculated on an individual subject basis (i.e., large individual variation), *qualitative* data analysis procedures were used (Attanasio, 1994; Kazdin, 1976; McReynolds & Kearns, 1983). Specifically, analyses involved determining each child's phonetic and phonemic inventories. As previously described, the context-free phonetic status of a sound was established by following the criterion of at least two unique productions of a sound (in any word position), regardless of whether or not they were considered "accurate" by adult standards (Gierut, Simmerman, & Neumann, 1994). A phonetic inventory by context was also obtained by following the criterion of at least two unique productions of a sound (in every word position), regardless of whether or not they were considered "accurate" by adult standard (Gierut et al., 1994). Phonemic status of a sound was established following the criterion of two unique sets of minimal pairs (e.g., rub-rub), again regardless of whether they were correct relative to adult production (Gierut et al., 1994). The phonetic and phonemic inventories of the children who participated in this research program are displayed in Table 1.1, 1.2, and 1.3.

In addition to the analysis of phonemic and phonetic inventories, the children's substitution patterns were also analyzed. The analysis of substitution patterns were done based on the adult production of the MOCP target words. The participants' productions of the MOCP target words were recorded and transcribed to examine any substitution errors. The participants' substitution variability was calculated pre-treatment and post-treatment MOCP sessions

Based on the speech sample transcriptions, accuracy percentage scores for untreated sounds and the overall percent consonants correct were also calculated for each administration of the AEP (Barlow, 2003) using the Logical International Phonetic Programs 2.02 (LIPP; Oller & Delgado, 1999) PC computer transcription program (Intelligent Hearing Systems, 2000). Specifically, consonant and vowel sound was identified as being correct or incorrect according to typical adult language. All sounds produced with less than 50% accuracy pre-treatment were included in the analyses, and percentage accuracy scores for both the pre-treatment and 2-week-post treatment AEP probes were calculated. Sounds were classified as being generalized if they were produced with less than 50% accuracy pre-treatment, and greater than 50% accuracy on the post-treatment AEP, provided that the amount of change was at least 10% as per the assumptions of single-subject designs (e.g., McReynolds and Kearns, 1983).

CHAPTER 4

RESULTS

4.1. Introduction

The results of teaching the three-element cluster /str/ to children with speech-sound disorders (SSDs) will be discussed in terms of learning during treatment, generalization from treatment, system-wide phonological change, and error consistency. While examining the generalization from treatment is arguably more important as it reflects the overall interaction of treatment on a child's phonological system, examining the learning that occurs during treatment is essential for determining whether or not the treatment worked. Each participant's treatment results are discussed in detail in the following sections. In order to fully understand the results, it is necessary to first describe the data that was collected and evaluated.

The treatment learning curve for each child shows the production accuracy of the targeted treatment cluster in word-initial position of the five selected treatment words during the treatment sessions. All children completed two or three baseline sessions and 10 to 12 treatment sessions. The treatment learning curve for each child was calculated by tracking the productions of the treatment cluster. Every production of the cluster was tracked during each treatment session and assessed in terms of the production accuracy of each element of the cluster (e.g., /s/, /t /, and / r/), as well as the overall production accuracy of the entire cluster (/str-/). For example, if the target word was "straw" and it was produced "stwaw" the participant would receive credit for accurately producing the first two elements of the target cluster (/s/ and /t/) while not receiving credit for producing the /r/ incorrectly, and the entire cluster production would also be marked as incorrect. An example of how the participants' productions of the target words were tracked through the treatment sessions is shown in Appendix 5. The learning curves are described in terms of the shape of the curves (e.g., flat, monotonic, or nonmonotonic; Gierut & Champion, 2001). A monotonic curve meant that the participant had continuous improvement in speech production accuracy throughout the duration of treatment while nonmonotonic curves meant that the children's production varied across treatment sessions (i.e., some sessions elicited higher accuracy levels than others). If a curve was described as being flat, it would indicate that the participant did not show change in production accuracy (either increases or decreases) across treatment sessions.

In order to examine the overall effect that the treatment had on each participant's phonological system, which ultimately impacted their overall intelligibility (i.e., generalization), all changes in their phonological repertoires are reported in terms of individual sound change. It is important to note that this measure was different from an overall Percent Consonants Correct (PCC) measurement, which would take into account the percentage of *all* the sounds that were produced as correct. Using the LIPP program (Oller & Delgado, 1999), the percentage of accuracy for *each* individual sound in the pre-treatment AEP was determined. Sounds produced with 50% accuracy or greater were identified as generalized and established in a child's phonology; thus, they were eliminated from further analysis. Alternatively, the sounds that were not produced with 50% accuracy or greater were monitored and any change in accuracy was reported. Sounds that are produced with more than 50% accuracy are often given lowest priority when identifying treatment goals (e.g., Fey, 1986; Paul, 2007), as they are assumed to be sounds that will continue to improve without treatment. These same clinical guidelines were used to define generalization in the present study. Since every child had a different set of sounds produced in error, some children had more singleton sounds included in this individual sound change analysis.

Generalization was also reported for the treated cluster, /str/, in untreated words and the two other clusters contained within treatment cluster /str/ (i.e., /st/ and /tr/). The clusters /st/ and /tr/ were monitored for generalization in order to determine if the treated cluster /str-/ produced change in untreated clusters specifically those that are contained within the treatment cluster. In order to measure generalization of these three selected clusters, all participants were given four baseline probes (e.g., Pre-Treatment MOCP, Baseline AEP probe 1, Post-Imitation MOCP, and Post-Treatment MOCP). Moreover, two children received a second baseline consisting of selected words from the AEP that contained elements of the target cluster /str/. The second Baseline AEP probe consisted of five words that contained the elements of the target cluster (i.e., all /st/, /tr/, and /str/ words). This probe was limited to five words based on the occurrence of the initial cluster words in the probe. The productions of /str-/, /st/, and /tr/ were measured and compared over the course of treatment. The overall accuracy values for the treated cluster /str/ in untreated words and the two untreated clusters are also described in terms of the shape of their generalization curves (e.g., monotonic or nonmonotonic).

Along with examining the learning and generalization that occurred due to the treatment process, it is also necessary to examine the individual sounds that were added to the participants' phonemic and phonetic inventories. While the sounds added to a child's inventories should correspond to the overall generalization patterns, it is likely that the child needed to add a sound to his context-free phonetic inventory (i.e., by producing it twice in any word position, regardless of whether or not it corresponded to the adult model) and phonemic inventory (i.e., by producing it contrastively within two separate minimal pairs) before it could be generalized (i.e., achieving greater than 50% accuracy). A sound could be added to the phonetic and phonemic inventories *and* be generalized, but it could not be generalized without being added to the context-free phonetic inventory at the very least. Tables 4.1, and 4.2 outline the sounds added to the participants' phonetic and phonemic inventories.

Table 4.1. Post-Treatment context-free phonetic inventories and phonetic inventories by context of the participants in the treatment study. The sounds added to the participants' phonetic inventories are marked by being **bold.** Recall that the phonetic inventories by context are red, while the context-free phonetic inventories are black.

Child	**Condition**	**Phonetic Inventory Post-Treatment**																							
PD16_HJ	Nonsense	post	m	n	ŋ	p	b	t	d	**k**	g	f				**s**	**z**			**dʒ**	w	j		l	r
		post	m	n	ŋ	p	b	t	d	k	g	f	v		**ð**	s	z	ʃ	tʃ	dʒ	w	j	h	l	r
PD22_RM	Real	post	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s̪	**s**				w	j	h	**l**	
		post	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s̪	**s**				w	j	h	l	
PD34_TP	Real	post	m	n	**ŋ**	p	b	t	d	k	g	f	v			s			**tʃ**	**dʒ**		j	h	l	**r**
		post	m	n	ŋ	p	b	t	d	k	g	f	v	θ	**ð**	s	**z**	ʃ	tʃ	dʒ	w	j	h	l	r
PD39_AW	Nonsense	post	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	tʃ	dʒ	w	j	h	**l**	
		post	m	n	ŋ	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	tʃ	dʒ	w	j	h	l	
PD42_BS	Nonsense	post	m	n	**ŋ**	p	b	t	d	k	g	f	v			s	z	ʃ	tʃ	dʒ	w	j	h	l	r
		post	m	n	**ŋ**	p	b	t	d	k	g	f	**v**			s			**tʃ**	**dʒ**	w	j	h	l	**r**

Table 4.2. Phonemic Inventories of participants in the treatment study. Note that these inventories may not be all-inclusive, meaning that a child's phonemic inventory could be missing a phoneme based on a sampling error. For example, it could be potentially was due to the fact that two "good" minimal pairs were unavailable in the recorded sample.

Child	Condition	Phonemic Inventory Pre- and Post- Treatment																							
PD16_HJ	Nonsense	Pre	m	n	ŋ	p	b	t	d	k	g	f				s			ʧ		w	j	h	l	
		Post	m	n		p	b	t	d	k	g	f				s	z		ʧ	ʤ	w	j	h	l	r
PD22_RM	Real	Pre	m	n		p	b	t	d	k	g	f		ɵ	ð						w	j	h		
		Post	m	n		p	b	t	d	k	g	f	v	ɵ	ð	s	z				w	j	h	l	r
PD34_TP	Real	Pre	m	n		p	b	t	d	k	g	f	v			s				ʤ	w	j	h		
		Post	m	n		p	b	t	d	k	g	f			ð	s		ʃ	ʧ		w	j	h		r
PD39_AW	Nonsense	Pre	m	n	ŋ	p	b	t	d	k			v			s		ʃ	ʧ	ʤ	w		h		
		Post	m	n	ŋ	p	b	t	d	k	g		v		ð	s		ʃ	ʧ	ʤ	w		h		
PD42_BS	Nonsense	Pre	m	n		p	b	t	d	k	g	f							ʧ		w	j	h		
		Post	m	n	ŋ	p	b	t	d	k	g	f				s		ʃ	ʧ	ʤ	w	j	h		r

Sound added to phonemic inventory during Post-Treatment AEP probe.

The relative consistency of error substitutions is also of interest in this study as it may provide additional information to the changes that have occurred in a child's phonological system due to treatment. The Error Consistency Index (ECI) was computed for each child based on their production of the target clusters in the MOCP. This analysis was designed to measure the overall consistency of error substitutions within a child's phonological system (Tyler, 2002; Tyler, Lewis & Welch, 2003). In the present study, the ECI was a raw number that was calculated by summing the total number of different substitutions that each participant made for each word-initial cluster (i.e., 20 clusters) produced during both the pre-treatment and post-treatment MOCPs. The treatment target cluster, /str-/, was also examined more closely by examining how many substitutions each participant made during the pre-treatment and post treatment MOCP sessions. It is important to note that correct productions of the target clusters were included in the ECI raw number and analyses, thus the ideal ECI number was 1.0. The following table represents the consistency of error substitutions of the participants.

Table 4.3. Summary of sound generalization for all subjects. See text for more information.

		Error Consistency Index[A]					
Subject	**Condition**	**Pre-Tx Treatment Cluster[B]**	**Post-Tx Treatment Cluster[B]**	**Treatment Sound Proportion Change (Negative Values)**	**Pre-Tx All Clusters***	**Post-Tx All Clusters***	**All Sounds Proportion Change (Negative Values)**
SSD39_AW	Nonsense	1	1	0	1.11	1.11	0
SSD16_HJ	Nonsense	5	3	0.4	2.55	2.29	0.101960784
SSD42_BS	Nonsense	5	3	0.4	2.22	1.88	0.153153153
SSD34_TP	Real	5	2	0.6	2.32	1.64	0.293103448
SSD22_RM	Real	2	3	-0.5	1.47	1.42	0.034013605
Average							
	NW	3.66	2.33	0.26	1.96	1.76	0.085037979
	RW	3	1.5	0.1	1.85	1.51	0.163558527

[A]*The Error Consistency Index does include the correct sound productions, if produced by each child. The ideal Error Consistency Index number is 1.0, meaning that only one sound exemplar was produced for each sound (PD1 and PD15 produced the target sound, while PD8 produced one substituted sound).*

[B]*When a cluster is targeted in treatment, the Error Consistency Index refers to the number of substitutions for the first sound in the cluster.*

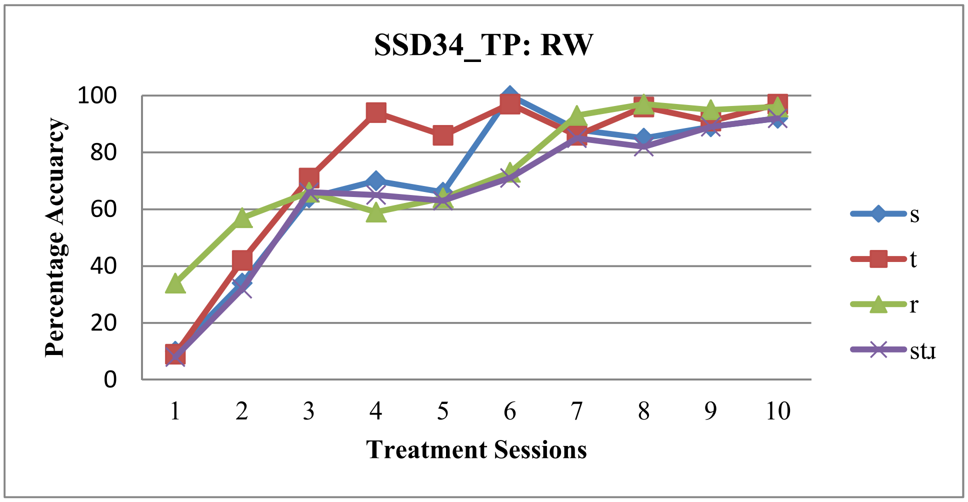
** The number of sounds varied depending on how many were produced with less than 50% accuracy during the pre-treatment probe sessions*

In the following sections, each participant's results will be presented individually, addressing each of the following areas: learning during treatment, generalization from treatment, sound segments added to their phonetic and phonemic inventories, and error variability in their production patterns. Each child's results will be followed an individualized discussion to describe factors which may or may have not influenced the participant's success with the treatment program. An overall discussion taking into account the results across children is presented in the end of this chapter.

4.2. SSD34_TP: Real-Word Treatment

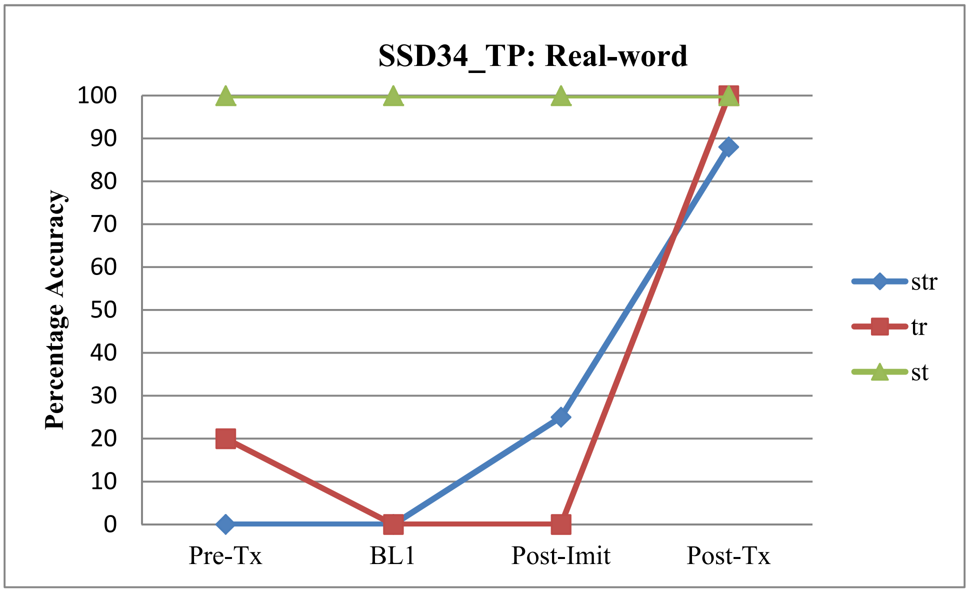
4.2.1. Learning during Treatment. The learning curve for TP, who was treated using RW treatment stimuli, is displayed in Graph 4.1. When TP entered this study, he was producing the treated cluster, /str/ with less than 5% accuracy. However, he started to show marked improvement in the production of the treated cluster almost immediately after the initiation of the treatment. The first few treatment sessions were spent shaping the /s/ and /r/ constituents of the treated cluster. He could produce each element of the cluster individually, but once combined in cluster form, he would substitute the phone /ʃ/ for the first two elements of the target cluster (i.e., /st/) as well as substitute /w/ for the last component of the target cluster (i.e., /r/). It was necessary to educate TP on the nature of the treated cluster in order to gain an understanding that each cluster's production should contain three sounds. His accuracy of the treated cluster hovered around 50% throughout most of the imitation phase of treatment and continued to slowly improve as the spontaneous phase progressed. TP's learning during treatment curve can be described as monotonic, as it continuously increased across treatment sessions. As compared to other children in his treatment group, TP demonstrated a quicker rate of learning with positive change occurring much earlier in the course of treatment. The change from treatment session 2 (with 32% accuracy) to treatment session 3 (with 66% accuracy) was a 32% improvement, which he was able to maintain through the remainder of treatment. TP produced at least 97 treatment words in each one-hour session, with a maximum of 198 words. His average production of the treatment words in each one-hour session was 147.

Graph 4.1. SSD34_TP's production of the target cluster, /str/, across treatment sessions.



4.2.2. Generalization from Treatment. The probe word generalization curves are for the production of the treated cluster, /str/, and the untreated clusters /st/ and /tr/. TP had a 100% pre-treatment accuracy level for cluster /st/, but less than 50% pre-treatment accuracy level for clusters /str/ and /tr/ due to his inaccurate /r/ productions (Graph 4.2). Accuracy levels are plotted over each administered probe (Pre-Treatment, Baseline 1, Baseline 2 (if applicable), Post-Imitation and Post-Treatment). This cluster probe generalization curve shows a monotonic curve for the target cluster /str/, a non-monotonic curve for the cluster /tr/, and a flat curve for the cluster /st/.

Graph 4.2: Generalization curve for SSD34_TP's, target cluster /str/ in untreated words and untreated clusters /st/ and /tr/.



Additionally, while the cluster /tr/ was initially produced with higher accuracy during the full MOCP, its accuracy decreased during the Baseline 1 probe and the Post Imitation probe. This was potentially due to the smaller sample size of the shorter Baseline 1 probe, which was a selective sample of words incorporating the three clusters on the AEP. Marked improvement on both the /str/ and /tr/ clusters occurred post-treatment; the /st/ cluster maintained its high level of accuracy. Based on this information, it can be concluded that the course of generalization for both the treated and untreated clusters was monotonic, in that generalization was consistently improving. It is important to note that the generalization probes did not use the same words that were targeted in treatment to determine generalization from treatment. This was done in order to ensure that generalization was due to actual change in the phonological system rather than learning the productions of a few specific words.

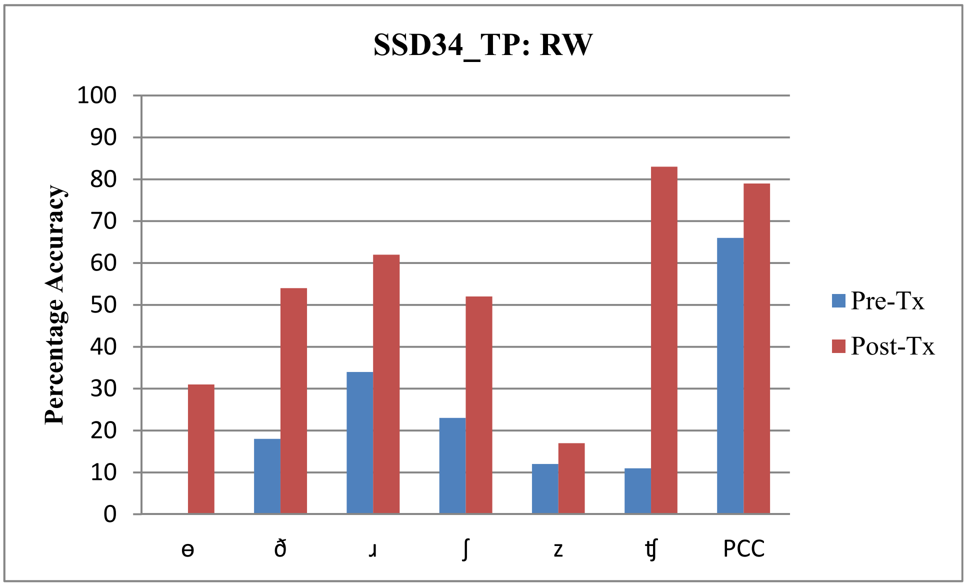
Looking more closely at the generalization of /str/, /st/, and /tr/ in untreated words containing those clusters, it is clear from Table 4.4 that the production of the untreated words reached 50% accuracy during the Post-Treatment probe. However, this may not fully represent the phonological change that was occurring. During the Post-Imitation probe (i.e., just 5 sessions into treatment), TP began to produce the /r/ phoneme with greater accuracy. His production of /r/ proved to be the main factor in whether he produced the treated cluster accurately. TP produced rhotic vowels (e.g., “or” and “ar”) with great accuracy, but his production of initial and cluster /r/ were produced incorrectly. In order to compensate for this inaccuracy, he began to produce the /r/ in the target cluster as though it was a rhotic vowel (i.e., “er” as in “sister”). Basically, he was adding an extra syllable to the word with the rhotic vowel. These productions were deemed to be inappropriate and potentially more noticeable than his pre-treatment substitution [w] for /r/. Even though the productions were in error, it was a positive step toward a correct production pattern. Table 4.4 shows TP’s exact production of the probe words to provide evidence of the very immediate phonological change due to treatment.

Table 4.4. SSD34_TP's productions of the MOCP words across treatment.

Gloss	Pre-Treatment	Post-Imitation	Post-Treatment
tree	twi	twi	tri
star	stɑr	stɑr	stɑr
train	treɪn	tweɪn	treɪn
stove	stoʊf	stoʊf	stoʊv
truck	twʌk	twʌk	trʌk
triangle	twɑɪŋgl	twɑɪeɪŋgl	trɑɪeɪŋgl
trick or treat	twɪk ɚ twit	twɪk ɚ twit	trɪk ɚ trit
street	stwit	stɚ:it	strit
string	tɚɪŋ	stɚɪŋ	ʃrɪŋ
straight	fweɪt	streɪt	streɪt
stop	stɑp	stɑp	stɑp
stamp	stæmp	stæmp	stæmp
strainer	stweɪnɚ	stweɪnɚ	streɪnɚ
stranded	stɚændɪd	stwændɪd	strændɪd
straw	swɑ	strɑ	strɑ
store	stɔr	stɔr	stɔr
strap	stwæp	stɚæp	stræp

While studying the generalization of treated and untreated clusters is of great interest and importance to understanding of how effective sound cluster treatment is, it is also important to study the broad generalization of untreated singleton sounds in order to examine the overall efficacy of the NWs and RWs in treatment. It provides information as to whether or not the treatment approach caused change to the phonological system as a whole. In order to determine broad generalization of sounds, the percentage of accuracy of singleton sounds was calculated. TP had six phonemes (/ɵ/, /ð/, /z/, /ʃ/, /ʧ/, and /r/) that were produced with less than 50% accuracy during the Pre-Treatment AEP probe. Of these six, he produced four of them with greater than 50% accuracy during the Post-Treatment AEP probe (Graph 4.3). Looking closely at the untreated sounds, all six sounds showed improvement from the beginning of treatment to the end of the treatment. Graph 4.3 represents the untreated sound change (i.e., system-wide generalization) that occurred between the Pre-Treatment AEP word probe and the Post-Treatment AEP word probe.

Graph 4.3. Production accuracy of PD34_TP's percentage of consonants correct (PCC) and the untreated phonemes that had a pre-treatment production accuracy of less that 50%.



It is also important to note that his overall percent consonant correct (PCC) also improved from 66% to 79%. Thus, it is likely that using a three-element, real-word cluster caused large amounts of system-wide phonological change in TP's phonological system, as all six possible sounds improved.

4.2.3. Sounds Added to Phonetic and Phonemic Inventories. Consistent with the generalization results presented above, TP added /ʧ, ʤ, ŋ, r/ to his Post-Treatment Phonetic Inventory by Context (Table 4.1). He also added /z/ and /ð/ to his Context-Free Post-Treatment Phonetic inventory (Table 4.1). Recall that the context-free phonetic inventory is less conservative, and TP was only missing these two sounds from his pre-treatment context free phonetic inventory. Compared to his generalization of monitored singletons, his phonetic inventories reflect consistency, with all four of the sounds that were generalized appearing in either his context-free phonetic inventory (/ð/ and /r/) or his phonetic inventory by context (/ʃ/ and / ʧ/. His Phonemic Inventory was expanded to include /ð, ʃ, ʧ, r/ (Table 4.3); the sounds added to his phonemic inventory were consistent with the sounds added to his phonetic inventory.

4.2.4. Variability in Production Patterns. During the pre-treatment MOCP, TP's production of the target cluster, /str/, was highly variable with five different substitutions noted, none of which was the target cluster. During the post-treatment assessment, TP produced the target cluster with 80% accuracy; the variability in his cluster production had greatly lessened to just 2 different exemplars. Using the MOCP, 20 word-initial clusters in the English language were assessed and evaluated for variability in production patterns pre- to post-treatment. When all of the cluster substitutions were compared from pre- to post-treatment, TP's Error Consistency Index (ECI) decreased (that is, improved). For the untreated clusters produced pre-treatment, he averaged 2.32 different exemplar substitutions for each cluster. At the post-treatment MOCP, TP's ECI had decreased from 2.32 to 1.64. Refer to Table 4.3 for the larger ECI comparison.

Table 4.5. PD34_TP's treated cluster during the pre-treatment OCP sessions and his production substitution exemplars for the cluster. Note that the correct adult-like productions may also be listed.

Target Sound		**Production Exemplars**						
str	Pre	stw	fw	tɚ	sw	stɚ		
	Post						ʃr	str

4.2.5. Discussion. TP was assigned to the real-word treatment condition. Throughout his time in the treatment study, he showed phonological change for both treated and untreated sounds. During treatment, TP showed immediate learning of the target cluster, /str/, and by the fourth Imitation treatment session, his production of the /str/ treatment words was over 50% accurate. This learning curve continued to follow a monotonic course throughout the rest of the treatment sessions. All monitored clusters and phonemes (i.e., sounds produced with less than 50% accuracy during the pre-treatment AEP probe, treated cluster /str/, and untreated clusters /st/ and /tr/) demonstrated an increase in accuracy at the Post-Treatment production probes (i.e., AEP and MOCP). The generalization curve for the treated cluster, /str/, in untreated words and untreated clusters /st/ and /tr/ production was monotonic in shape, continuing to increase in accuracy over time. Nearly all untreated sounds (i.e., four of the six monitored sounds) that were initially produced with less than 50% accuracy improved to at least 50% accuracy by the Post-Treatment probe.

TP demonstrated early and consistent sound change in both his untreated and treated sounds. It is important to discuss possible reasons that may have contributed to the earlier and consistent sound change in this participant. TP was 5;8 when he entered this study, he was neither the youngest nor the oldest participant. Even so, his age could have worked against him as he had that many years to produce sounds incorrectly and for habits of "bad" speech to manifest. He was in the real-word treatment condition, so it could be argued that he had been pre-exposed to the treatment words, which may potentially be a negative factor as he had more practice producing them incorrectly. Nevertheless, TP was very receptive to the target cluster and began to make marked improvement almost immediately.

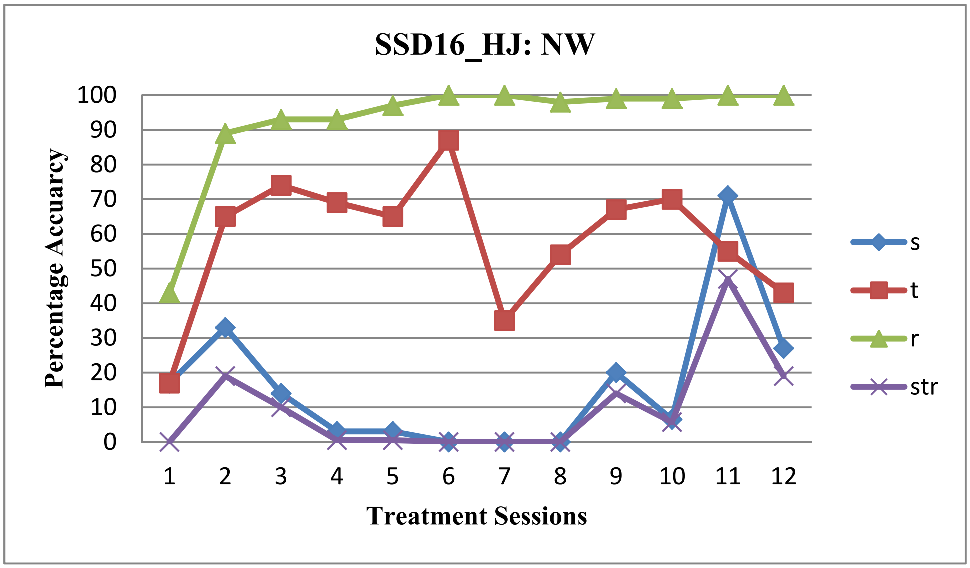
Additionally, TP was the youngest of four brothers. His mother reported that his brothers liked to pick on him for his "baby" talk. While TP did not mention this, it could have been a motivating factor for him to "sound" like all of his older brothers. He had attended one year of Head Start and a four-week kindergarten preparation study prior to his entry into this treatment study. Because of these previous academic experiences, he was accustomed to sitting still and working at a table for an extended period of time. Even though attention and self-control were not necessary for the treatment to succeed, the ability to work hard and stay on task was important for productively in terms of how many productions of the treatment words could be elicited. It was possible that regardless of the type of words (real or nonsense) chosen for the treatment, the number of productions of the treatment words could have made a difference in terms of when and how much of a sound change is observed. Compared to the other participants, TP had the highest average of total production of the treatment words (average of 147).

Anecdotally, TP seemed to really enjoy participating in the activities designed to elicit productions of the target words. He was motivated by his success and knew when he had a particularly "good speech day". He became especially adept at self-correcting himself without any prompts or cues from the clinician, often stopping himself if he produced the target inaccurately. He enjoyed reading and re-telling the story used to elicit the target words and even though he could not read the words by himself, he began to recall the story in great detail. Overall, TP was an extremely hard-working child with a respectful demeanor. His mother reported that hard-work and respect were values that they reinforced in each of their children and that their youngest seemed to be the one who took it most to heart.

4.3. SSD16_HJ: Non-Word Treatment

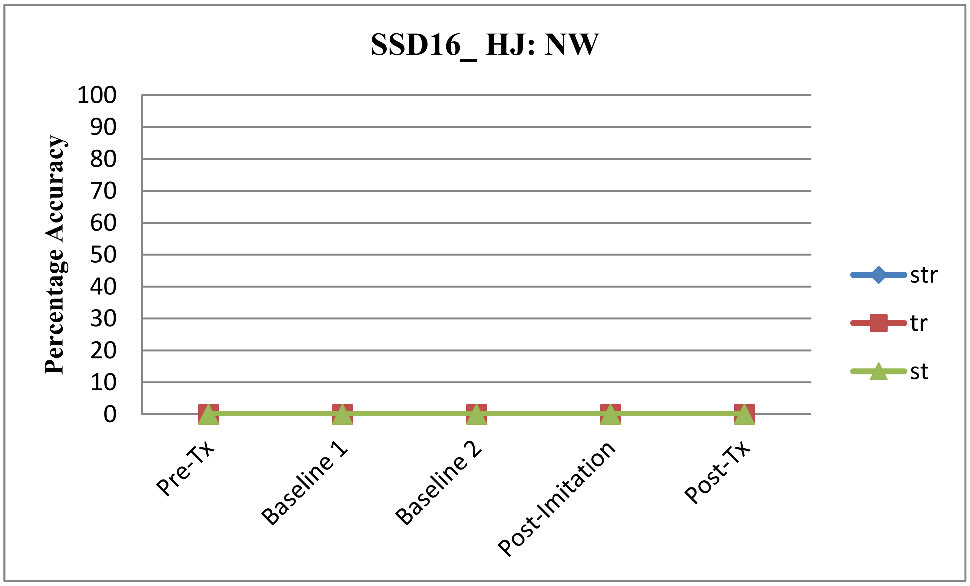
4.3.1. Learning During Treatment. The learning curve for HJ, who was treated using NW treatment stimuli, is shown in Graph 4.4. HJ showed a nonmonotonic curve in his learning during treatment, as his production accuracy greatly varied from session to session. It is important to note that although his productions were varied in the accuracy, he did show improvement. Upon entering this research study, HJ could not produce the target cluster even with the highest level of support (e.g., a combination of visual and verbal cues). He would often omit one or more of the elements or even add sounds into the cluster. For example, with the NW "strupe" he would produce the target cluster as /smɪtɚ/. The first few sessions were spent educating him on what clusters were and it was necessary to completely segment the cluster (i.e., take it "apart" and produce each sound in its own syllable). HJ could produce all three sounds in the target cluster, but not without elongation and syllable manipulation (e.g., changing the /r/ to the /ɚ/). Special attention was paid to /s/ as this was the sound that caused the most problems for HJ. When treatment ended HJ produced the target cluster with 20% accuracy when elicited through picture cards. The number of word productions produced by HJ in a session greatly varied from session to session, with the lowest being 46 and the highest number being 123. His average number of treatment word productions was 72 across all treatment sessions.

Graph 4.4. Non-monotonic Treatment Learning Curve for subject SSD16_HJ



4.3.2. Generalization from Treatment. The cluster probe generalization curves for the production the treated three-element cluster, /str/, in untreated words and the untreated clusters /st/ and /tr/ are shown in Graph 4.5. Accuracy levels are plotted over each administered probe (Pre-Treatment, Baseline 1, Baseline 2 (if applicable), Post-Imitation and Post-Treatment. HJ was one of the participants randomly selective to receive a second Baseline probe. Graph 4.5 shows a flat curve, with no variation in the production of the treated and untreated clusters across treatment. This cluster probe generalization curve shows that there was no variability and essentially no change in cluster production accuracy between the pre-treatment and post-treatment sessions. However, it is important to note that HJ did show improvement across treatment sessions. His lack of variation on the probes may be because no feedback or teaching was given through his productions of the probe words.

Graph 4.5. Cluster probe generalization curve for SSDHJ_16 for his treated cluster /str/ and untreated clusters /tr/ and /st/.



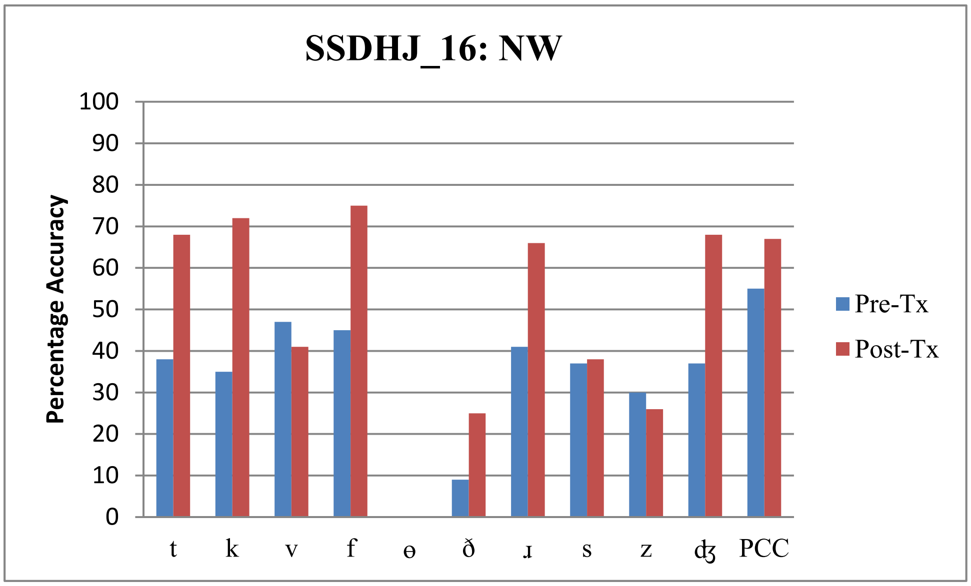
Many things could be factored into the lack of generalization that occurred within the treated and untreated cluster. HJ had been receiving speech therapy for nearly two years. His speech was characterized by inconsistent productions of sounds, meaning that words would often be produced in a different way in every production. Taking a closer look at the variability of his word productions during the probes could provide some perspective into the phonological change that occurred. Table 4.6 shows HJ's exact productions of the target words in the probes.

Table 4.6. SSD16_HJ's productions of the MOCP target words across treatment probes.

Gloss	Pre-Treatment	Post-Imitation	Post-Treatment
tree	di	di	di
star	dɑr	dɑr	dɑr
train	deɪn	deɪn	tweɪn
stove	doʊf	doʊf	doʊv
truck	gʌk	gʌk	dwʌk
triangle	gɑɪŋgl	gɑɪŋgl	dɑɪeɪŋgl
trick or treat	dɪk ɚ tit	dɪk ɚ tit	dwɪk ɚ tit
street	dwit	dwit	swrit
string	gwɪŋ	gwɪŋ	gwɪŋ
straight	gweɪt	gweɪt	tweɪt
stop	dɑp	dɑp	dɑp
stamp	dæmp	dæmp	dæmp
strainer	deɪnɚ	deɪnɚ	deɪnɚ
stranded	dwændɪd	dwændɪd	stwændɪd
straw	trɑ	trɑ	trɑ
strap	gæk	gæk	twæk
store	dor	tor	dor

While studying the generalization of /str/, /st/, /tr/ in untreated words containing those clusters, of great interest and importance to understanding of how effective sound cluster treatment is, it is also important to study the broad generalization of untreated singleton sounds in order to examine the overall efficacy of real words vs. non-words. It will allow us to know if the treatment approach caused a change to the phonological system as a whole. In order to determine broad generalization of sounds, the percentage of accuracy of single sounds was calculated. HJ had ten phonemes, /t/, /k/, /v/, /f/, /ɵ/, /ð/, /z/, /s/, /ʤ/, /r/, that were produced with less than 50% accuracy during the Pre-Treatment AEP probe. Despite the lack of improvement on the treated cluster, /str/, in untreated words, many of the untreated phonemes showed dramatic increases in production accuracy (i.e., generalization) during the Post-Treatment AEP probe. Specifically, five of the ten untreated phonemes showed generalization with HJ producing them with greater than 50% accuracy during the Post-Treatment AEP probe (Graph 4.3). Graph 4.6 represents the untreated sound change (i.e., system-wide generalization) that occurred between the Pre-Treatment AEP word probe and the Post-Treatment AEP word probe.

Graph 4.6. Production accuracy of SSD16_HJ's percent consonants correct (PCC) and the untreated phonemes that had a pre-treatment production accuracy of less than 50%.



Given that the treated cluster did not show any measurable generalization to untreated words containing the /str/ throughout the treatment period, generalization of untreated singletons is of great interest to this participant. Measureable generalization in the untreated sounds would suggest that while treatment with the cluster /str/ did not produce measureable phonological change for the treated sound, it was still an efficient and effective treatment target Indeed, large system-wide phonological change was achieved due to the treatment of the three-element cluster. These singleton sound results are very interesting because they are not consistent with the generalization of the untreated and treated clusters.

4.3.3. Phones Added to Phonetic and Phonemic Inventories. Consistent with the generalization results presented above, HJ added /k/, /z/, /ʤ/, and /r/ to his Post-Treatment Phonetic Inventory by Context (Table 4.1) and he added /ð/ to his post-treatment context-free phonetic inventory (Table 4.1). His Phonemic Inventory was expanded to include /z/, /ʤ/, and /r/ (Table 4.2). The latter three sounds added to his Phonemic Inventory are also consistent with the generalization findings. These results are quite complementary to the generalization findings and provide a more complete picture of how HJ's phonology improved due to treatment.

4.3.4. Variability in Production Patterns. During the pre-treatment MOCP, HJ's production of the target cluster, /str/, was highly variable with five different substitutions noted, none of which was the target cluster. During the post-treatment assessment, HJ was still not producing the targeted treatment cluster; however, the variability in his cluster production had lessened to three different exemplars. Using the MOCP, all word-initial clusters in the English language (i.e., 21 clusters) were assessed and evaluated for variability in production patterns. Recall that all other cluster productions in the MOCP were compared from pre- to post-treatment. When all of the cluster substitutions were compared from pre- to post-treatment, HJ's Error Consistency Index (ECI) decreased (that is, improved). For the untreated clusters produced pre-treatment, he averaged 2.55 different exemplar substitutions for each cluster. At the post-treatment MOCP, HJ's ECI had decreased from 2.55 to 2.29. This is not a drastic change but it does show there was a decrease (improvement) in production variability from pre- to post-treatment. Refer to Table 4.3 for a more comprehensive comparison of HJ's pre- and post-treatment ECI.

Table 4.7. SSD16_HJ's treated cluster during the pre-treatment MOCP sessions and his production substitution exemplars for the cluster. Note that correct adult-like productions may also be listed.

Target Sound		**Production Exemplars**						
str	Pre	dw	gw	tr	kr	g		
	Post			tr			sr	dr

4.3.5. Discussion. SSD16_HJ was assigned to the non-word treatment condition. Throughout his treatment study participation, different patterns of onset and course of phonological change were observed for the treated and untreated sounds. In terms of his learning of treatment cluster in the treatment words, he showed phonological change throughout the 12 treatment sessions. His learning curve was non-monotonic, with his production accuracy varying from session to session but it did show evidence of learning across treatment sessions. The generalization curve for the treated cluster, /str/, in untreated words and untreated clusters /st/ and /tr/ production was flat in shape, showing neither positive or negative changes. Despite the lack of generalization that occurred for the treated cluster, /str/, in untreated words and untreated clusters /st/ and /tr/, HJ did demonstrate generalization to untreated singleton sounds. Five of the ten untreated sounds that HJ produced with less than 50% accuracy during the Pre-Treatment AEP probe improved to at least 50% accuracy by the Post-Treatment AEP probe.

HJ's phonological change was one of the greatest observed in all the children participating in the research study. While his production accuracy for the treated cluster contained in the treatment words and his generalization of the treated cluster to untreated words did not reach 50% accuracy at the end of treatment, the system-wide change observed was impressive. His peculiar results of great phonological change and no generalization of the treated cluster can be attributed to multiple factors.

One of the first things to mention is the fact that HJ had completed nearly two years of speech services upon enrollment into this research study. Throughout this time HJ worked with the same clinician, and had formed a very comfortable relationship with her. He very much enjoyed telling stories and having long conversations. Because HJ also had a language impairment, the clinician was willing to allow him to "procrastinate" on saying his treatment words, as a way of inadvertently supplementing his language needs. On average, of all of the participants he produced the fewest treatment words per treatment session. This alone could be a contributing factor into the lack of generalization of the treatment cluster.

Recall that HJ received one of the lowest scores on the standardized language assessment and was diagnosed with concomitant language impairment. Because of this, HJ may have had a difficult time learning and remembering the NW treatment targets. The use of NWs with HJ may have actually been too difficult for him, as he had to concentrate on producing a complex sound and also learning a new set of words.

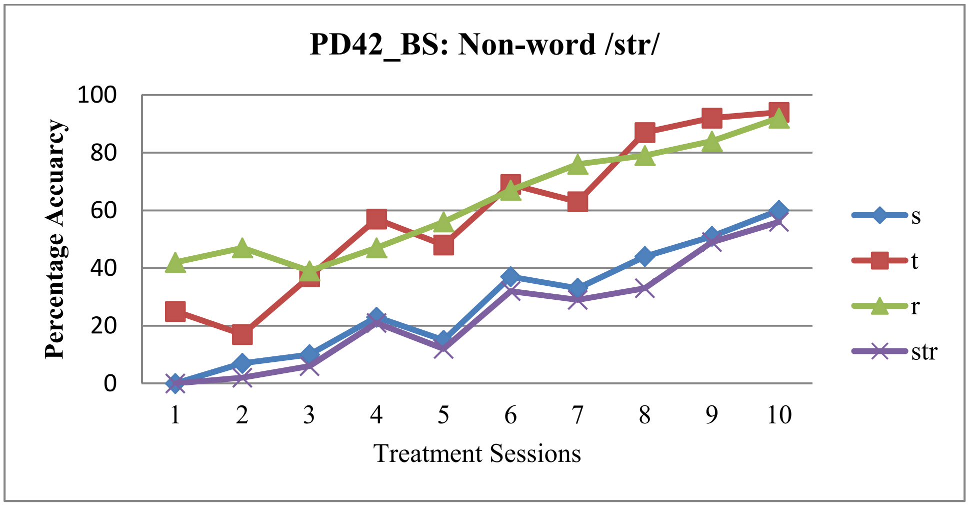
Another factor affecting HJ's performance was his ability to actually produce the cluster as a whole. A majority of the treatment sessions were spent segmenting the cluster into separate phonemes. Visual cues were a very consistent factor during the treatment sessions, as the clinician was constantly cueing him visually to make each sound of the cluster. He could produce each sound in the cluster as a lengthened unit (i.e. /s:tr:i/) but when instructed to produce the cluster faster he could not say any of the treatment words correctly. Therefore when given the probe, HJ was producing the words with the treated cluster without segmentation so he producing them incorrectly.

Considering all of these factors, it is easy to see why HJ's data yielded the results that it did. Again, HJ's phonology showed great amounts of change due to treatment with the non-word /str/ stimuli without generalizing the target cluster. From the onset, it was acknowledged that it was possible that he would not produce the target correctly by the end of the treatment, though it might make the greatest impact on his phonological system and indeed that appears to be what happened.

4.4. SSD42_BS: Non-Word Treatment

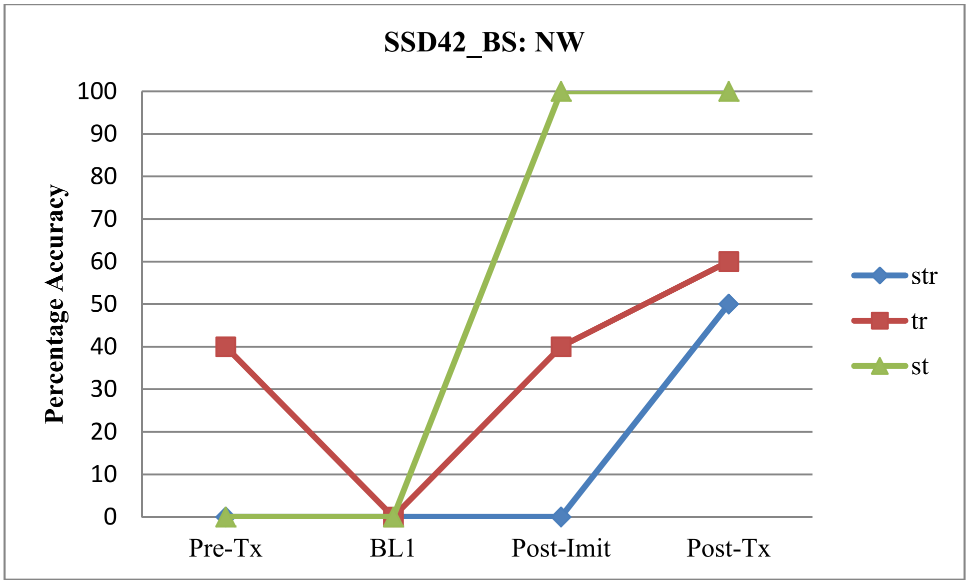
4.4.1. Learning during Treatment. The learning curve for BS, who was treated using NW treatment stimuli, is displayed in Graph 4.7. When BS entered this study, he was producing the treated cluster, /str/ with less than 5% accuracy. He began to show small improvements in the production of the treated cluster almost immediately. The first few treatment sessions were spent shaping the /s/ and /t/ constituents of the treated cluster. Similar to SSD34_TP, he could produce each element of the cluster individually, but once combined in cluster form the first two elements of the cluster, /ʃ/, was substituted for first two components of the target cluster (i.e., /st-/) and /w/ was substituted for the last element of the target cluster (e.g., /r/). It was necessary to educate BS on the nature of the treated cluster in order to gain an understanding that each cluster's production should contain three sounds. His accuracy of the treated cluster slightly fluctuated from session to session, averaging a 24% accuracy rate throughout the treatment sessions. His learning curve followed a fairly monotonic curve, though there were some subtle fluctuations in his accuracy throughout treatment. On the final day of treatment, BS produced his treatment words with 56% accuracy. BS produced at least 82 treatment words in each one-hour session, with a maximum of 135 words. Across all treatment sessions, he averaged 116 productions of the treatment words during treatment sessions.

Graph 4.7. SSD42_BS's production of the target cluster /str/ across treatment sessions.



4.4.2. Generalization from Treatment. The probe word generalization curves are for the production of the treated cluster, /str/, and the untreated clusters /st/ and tr/. BS had a 0% pre-treatment accuracy level for cluster /st/ and /str/, and a 40% pre-treatment accuracy level for cluster /tr/ (Graph 4.8). Accuracy levels are plotted over each administered probe (Pre-Treatment, Baseline 1, Baseline 2 (if applicable), Post-Imitation and Post-Treatment. This cluster probe generalization curve shows a monotonic curve for the target cluster, cluster /tr/, and cluster /st/.

Graph 4.8. Cluster probe generalization curve for SSD42_BS for his treated cluster /str/ and untreated clusters /tr/ and /st/.



Additionally, while the cluster /tr/ was initially produced with higher accuracy during the full MOCP, its accuracy decreased during the Baseline 1 probe. This was potentially due to the smaller sample size of the shorter Baseline 1 probe, which was a selective sample of words incorporating the three clusters using target words on the full AEP. Improvement on all the clusters occurred post-treatment, with the most extreme change being in untreated cluster /st/ (0% to 100%). Based on this information, it can be concluded that the course of generalization for both the treated and untreated clusters was monotonic, in that generalization was consistently improving. It is important to note that the generalization probes did not use the same words that were targeted in treatment to determine generalization during treatment. This was done in order to ensure that generalization was due to actual change in the phonological system rather than learning the productions of a few specific words.

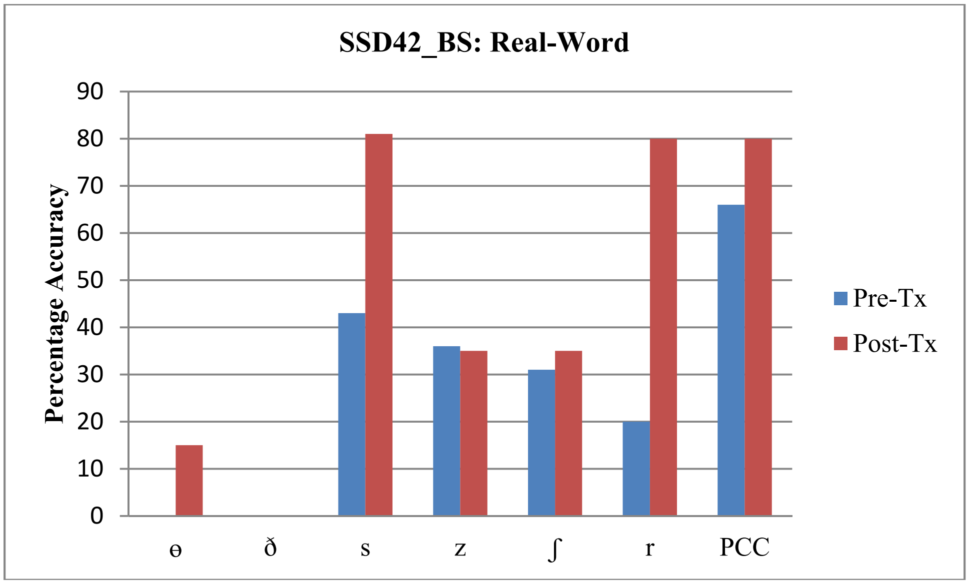
Looking more closely at the generalization of the treated cluster, /str/, /st/, and /tr/ in untreated words containing those clusters, it is clear from Graph 4.8 that the production of the target cluster in untreated words reached 50% accuracy during the Post-Treatment MOCP. However, this may not fully represent the phonological change that was occurring. During the Post-Imitation probe (i.e., just 5 sessions into treatment), BS was still producing the target cluster with 0% accuracy. His production accuracy of the target seemed to be influenced most by his tendency to substitute [ʃ] for /st/ or /s/ elements of the target BS produced the target with greater accuracy when he elongated the /s/, but these productions were deemed to be inappropriate and were just as noticeable as his substitution of [ʃ] for /st/. In fact, BS began to demonstrate some speech fluency problems when producing the treatment target words in the 7th and 8th treatment sessions, prolonging the /s/ for nearly three seconds at times. This potential stuttering issue was addressed in the 9th treatment session and did not occur in the final treatment session. It is important to note that though the prolongation of the /s/ was considered to be an error, it was a step toward a correct production pattern. Table 4.8 shows BS's exact production of the probe words to show evidence of phonological change due to treatment.

Table 4.8. SSD42_BS's productions of the MOCP words in probes across treatment.

Gloss	Pre-Treatment	Post-Imitation	Post-Treatment
tree	twi	tri	twi
star	tɑr	stɑr	stɑr
train	treɪn	ʧreɪn	treɪn
stove	toʊf	stoʊv	stoʊv
truck	trʌk	trʌk	trʌk
triangle	traɪŋgl	ʃraɪeɪŋgl	ʃraɪeɪŋgl
trick or treat	ʧɪk ɚ twit	ʃrɪk ɚ ʃrit	trɪk ɚ trit
street	trit	ʃtrit	ʃtrit
string	trɪŋ	ʃtrɪŋ	strɪŋ
straight	ʃreɪt	ʃreɪt	streɪt
stop	tɑp	stɑp	stɑp
stamp	tæmp	stæmp	stæmp
strainer	treɪnɚ	ʃreɪnɚ	ʃreɪnɚ
stranded	ʃrændɪd	ʃrændɪd	ʃrændɪd
straw	ʧɑ	ʃrɑ	strɑ
strap	ʧræk	ʃtræp	stræp
store	tɔr	stɔr	stɔr

While studying the generalization of treated and untreated clusters is of great interest and importance to understanding of how effective sound cluster treatment is, it is also important to study the broad generalization of untreated singleton sounds in order to examine the overall efficacy of non-words in treatment. It will allow us to know if the treatment approach caused a change to the phonological system as a whole. In order to determine broad generalization of sounds, the percentage of accuracy of single sounds was calculated. BS produced six phonemes, /ɵ/, /ð/, /z/, /s/, /ʃ/, /r/, with less than 50% accuracy during the Pre-Treatment AEP probe. Nearly half of the untreated phonemes showed dramatic increases in production accuracy (i.e., generalization) during the Post-Treatment probe. Specifically, two of the six untreated phonemes showed generalization with BS producing them with greater than 50% accuracy during the Post-Treatment AEP probe (Graph 4.3). Graph 4.9 represents the untreated sound change (i.e., system-wide generalization) that occurred between the Pre-Treatment AEP word probe and the Post-Treatment AEP word probe.

Graph 4.9. Production accuracy of SSD42_BS's percent consonants correct (PCC) and untreated phonemes that had a pre-treatment production accuracy of less than 50%.



4.4.3. Phones Added to Phonetic and Phonemic Inventories. BS added the most sounds of all the participants into both his phonetic and phonemic inventories. Consistent with the generalization results presented above, BS added /s/ and /r/ to his Post-Treatment Phonetic Inventory by context (Table 4.1). In addition to these two sounds, he also added the sounds, / ŋ/, /v/, / ʧ/, and /ʤ/, adding six sounds into his post-treatment phonetic inventory by context (Table 4.1). This was the largest number of sounds added into a context-based phonetic inventory of all of the participants. BS only added one sound to this context-free phonetic inventory. His pre-treatment context-free phonetic inventory was only missing three sounds, so the adding of a single sound is still significant change. Recall that the phonetic inventory by context is the more conservative phonetic inventory and it is much more difficult to meet the requirements to add a sound. Because of this, adding six sounds into a context-based phonetic inventory shows significant change regardless of the amount of sounds added into the context-free phonetic inventory. BS added four sounds into his phonemic inventory as well: /ŋ/, /s/, /ʃ/, and /ʤ/ (Table 4.2). It is likely that the fact that the phoneme /l/ not being included in BS's phonemic inventory and phonetic inventory by context is a result of poor sampling, as he produced the /l/ with 80% accuracy in the post-treatment AEP probe.

4.4.4. Variability in Production Patterns. During the pre-treatment MOCP, BS's production of the target cluster, /str/, was highly variable with five different substitutions noted, none of which was the target sound. During the post-treatment assessment, BS was still not producing the targeted treatment cluster with 50% accuracy, though his variability of the target cluster had decreased from five different exemplars to three. Using the MOCP, all word-initial clusters in the English language were assessed and evaluated for variability in production patterns. Recall that all other word-initial cluster productions (i.e., 21 clusters), minus the treatment cluster /str/, in the MOCP were compared from pre- to post-treatment. When all of the cluster substitutions were compared from pre- to post-treatment, BS's Error Consistency Index (ECI) decreased (that is, improved). For the untreated clusters produced pre-treatment, he averaged 2.22 different exemplar substitutions for each cluster. At the post-treatment OCP, BS's ECI had decreased from 2.22 to 1.88. Refer to Table 4.3 for a more comprehensive comparison of pre- and post-treatment ECIs.

Table 4.9. SSD42_BS's treated cluster during the pre-treatment MOCP sessions and his production substitution exemplars for the cluster. Note that correct adult-like productions may also be listed.

Target Sound		**Production Exemplars**						
str	Pre	t	ʧ	ʧr	ʃr	sw		
	Post				ʃr		ʃtr	str

4.4.5. Discussion. PD42_BS was assigned to the non-word treatment condition. Throughout his time in the treatment study, he displayed similar monotonic patterns (consistently improving) for both treated and untreated sounds. During treatment, BS showed relatively slow learning of his treated cluster, only surpassing the 50% production accuracy during the final two treatment sessions. His learning curve followed a fairly monotonic curve, with slight variations in production accuracy throughout treatment sessions. Similar patterns of generalization were observed for the treatment cluster, /str/, and the two untreated clusters, /st/ and /tr/. BS added six sounds to his phonetic inventory by context, one sound to his context-free phonetic inventory and five sounds into his phonemic inventory, the most of any participant in the research study. He also showed a decreased ECI in both his treated cluster and untreated clusters.

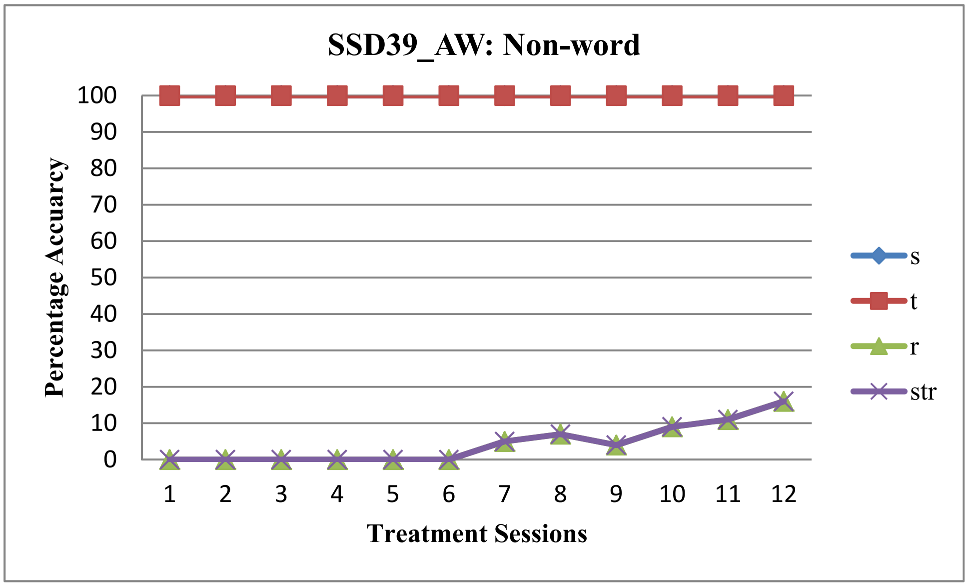
While BS demonstrated a large amount of phonological change, he was by far the most challenging child to work with during the treatment sessions. Very often, the first five to ten minutes of the session were spent coaxing BS into participating or simply setting foot into the therapy room. In order to entice him, the clinician made a visual schedule for each therapy session and allowed BS to have some control over therapy session activities. Once he was seated and ready to begin working, he could only focus on the task at hand for a very short period of time. He would often complain of being tired, saying he wished he could go back to bed. The actual working time was often cut short or very limited due to defiant behaviors. Given this information, it is quite remarkable that BS displayed so much phonological growth throughout the treatment period.

BS was 5;9 when he entered the research study and while he was not the oldest child in the research study, he was the oldest of the three boys. He had also had previous exposure speech-language concepts due to previous speech therapy at Head Start. He had received approximately three months of speech-services at Head Start where his attendance was very good due to the fact that the facility had an in house speech-language pathologist. However, his responses on the pre-treatment baseline measures were consistent, thus his phonological system was arguably stable prior to entering this treatment study. While these factors may have played a role in BS's performance it is impossible to ignore the large changes that occurred in his phonological system due to treatment. It can be said that a complex, three-element cluster treatment target presented within non-words caused widespread change in BS's phonological system.

4.5. SSD39_AW: Non-Word Treatment

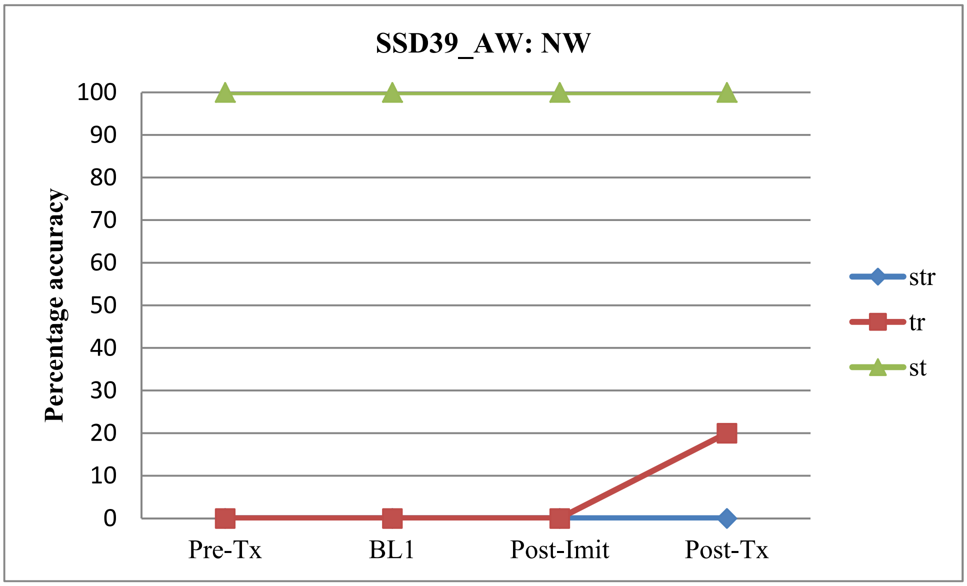
4.5.1. Learning during Treatment. The learning curve for SSD39_AW, who was treated using NW stimuli, is displayed in Graph 4.10. AW showed no productive phonological change until well into the final stages of treatment sessions. During the last two sessions, AW began to produce the target cluster in a more adult-like manner. At the end of the treatment period, AW appeared to reach a variable production accuracy of approximately 15%. During the course of treatment, she produced an average of 135 productions per 1-hour session with a minimum of 126 and a maximum of 153.

Graph 4.10. SSD39_AW's production of the target cluster /str/ across treatment sessions. Note that AW's /s/ and /t/ productions were produced with the same accuracy throughout the treatment study therefore the /t/ line is overlapping the /s/ line in this chart



4.5.2. Generalization from Treatment. The cluster probe generalization curves for the production the treated three-element cluster, /str/, in untreated words and the untreated clusters /st/ and /tr/ are shown in Graph 4.11. Accuracy levels are plotted over each administered probe (Pre-Treatment, Baseline 1, Baseline 2 (if applicable), Post-Imitation and Post-Treatment. Graph 4.11 shows a flat curve, with no change in the production of the treated and untreated clusters. However, it is important to note that AW did show improvement across treatment sessions. Her lack of variation on the probes may be because no feedback or teaching was given through his productions of the probe words. AW depended on these cues in order to correctly produce the target cluster, specifically with placement and shaping of the /r/ segment of the target cluster.

Graph 4.11. Cluster probe generalization curve for SSD39_AW, for her treated cluster /str/ and untreated clusters /tr/ and /st/.



Looking more closely at AW's generalization of the target cluster, /str/, to untreated words containing that sound, it is clear to see that generalization did not occur to the target sound and very little generalization to the untreated sound /tr/. AW began the NW treatment study with the substitution error of [w] for /r/. This error was consistent throughout the whole treatment study. Every treatment session began with 20 minutes of placement and shape "coaching" for the /r/ phoneme. In many instances a successful /r/ could not be produced as AW was not stimulable for the sound. This lack of stimulability will be discussed in length in a later section.

Looking more closely at the generalization of /str/, /st/, and /tr/ in untreated words containing elements of those clusters, it is clear from Graph 4.12 that the production of the untreated words did not reach 50% accuracy during the Post-Treatment probe. However, this may not fully represent the phonological change that was occurring. AW's production of /r/ proved to be the main factor in whether she produced the treated cluster accurately. In nearly every production of the target cluster, she substituted [w] for /r/. These productions were deemed to be inappropriate and even after a large amount of time spent shaping the /r/ phoneme, AW could not produce the sound even in isolation (i.e., separated from the rest of the target cluster). Table 4.10 shows AW's exact production of the probe words to provide evidence of the lack of generalization to both the treated cluster in untreated words and the untreated clusters.

Table 4.10. SSD39_AW's productions of the MOCP words across treatment.

Gloss	**Pre-Treatment**	**Post-Imitation**	**Post-Treatment**
tree	dwit	twi	di
star	stɑ	stɑr	dɑr
train	tweɪn	tweɪn	tweɪn
stove	stoʊv	stoʊf	doʊv
truck	twʌk	twʌk	dwʌk
triangle	twɑɪeɪŋgl	twɑɪeɪŋgl	twɑɪeɪŋgl
trick or treat	twɪk oʊ twit	twɪk oʊ twit	twɪk oʊ twit
street	stwit	stwit	stwit
string	stwi	stwɪŋ	stwrɪŋ
straight	stweɪt	stweɪt	stweɪt
stop	stɑp	stɑp	stɑp
stamp	stæmp	stæmp	stæmp
strainer	stweɪnoʊ	stweɪnoʊ	stweɪnoʊ
stranded	stwændɪd	stwændɪd	stwændɪd
straw	stwɑ	stwɑ	stwɑ
strap	stwæp	stwæp	stwæp

AW also showed very minimal improvement in her production of untreated consonant singletons (Graph 4.12). Two sounds were produced with less than 50% accuracy during the Pre-Treatment AEP probe. Of those two, neither was produced with greater than 50% accuracy during the Post-Treatment AEP probe, though they did show small levels of change. AW's generalization curve can be described as monotonic, despite the small amount of change it was there. Based on AW's data, it does not appear that the non-word treatment using /str/ stimuli generated any system-wide change in AW's phonology.

Graph 4.12. Production accuracy of SSD39_AW's percent consonants correct (PCC) and untreated phonemes that had a pre-treatment production accuracy of less than 50%.

4.5.3. Phones Added to Phonetic and Phonemic Inventories. AW only had one sound missing from her pre-treatment context-free phonetic inventory, which was /r/. It is apparent that the /r/ phoneme was not added into her post-treatment context-free phonetic inventory based on her production accuracy of /r/. Anna's pre-treatment phonetic inventory by context was missing four sounds. Of these four sounds, she was able to add one sound to her post-treatment phonetic inventory by context and that sound was /l/. Thus, adding 25% of the sounds missing from her phonetic inventory by context can be considered to be a nice accomplishment. AW's Post-Treatment Phonemic Inventory yielded a much greater phonological change, as four sounds were added: / ŋ/, /s/, /ʃ/, and /ʤ/.

4.5.4. Variability in Production Patterns. AW demonstrated limited variability during the pre-treatment and post-treatment MOCPs. During the pre-treatment baseline probes, AW produced the target cluster /str/ with one exemplar, /stw/. This exemplar remained consistent from pre- to post-treatment probes. Using the MOCP, all other word-initial clusters (i.e., 21 of them) in the English language were assessed and evaluated for variability in production patterns. Recall that all other cluster productions, minus the treatment cluster /str/, in the MOCP were compared from pre- to post-treatment. When all of the cluster substitutions were compared from pre- to post-treatment, AW's Error Consistency Index (ECI) did not change. For the untreated clusters produced pre-treatment, she averaged 1.11 different exemplar substitutions for each cluster and this amount remained the same during the Post-Treatment MOCP. These results were largely driven by Anna's substitution of /w/ for /r/, as nearly all of her cluster substitutions were reflective of this error. Refer to Table 4.3 for a more comprehensive comparison of pre- and post-treatment ECIs.

Table 4.11. SSD39_AW's treated cluster during the pre-treatment OCP sessions and her production substitution exemplars for the cluster. Note that correct adult-like productions may also be listed.

Target Sound	Production Exemplars	
str	Pre	stw
	Post	stw

4.5.5 Discussion. PD39_AW was assigned to the non-word treatment using /str/ stimuli. Throughout her treatment study, AW demonstrated similar learning and generalization patterns for both the treated and untreated sounds. During treatment, AW showed a delayed, monotonic learning curve in her production of the treated sound in the selected treatment words. The monotonic learning curve can be described as "small" as the production accuracy rate was 15%. In terms of generalization the treated cluster to untreated words, all of AW's treated and untreated clusters demonstrated a flat learning curve. AW added one sound to her phonetic inventory by context and four sounds to her phonemic inventory.

AW's lack of learning and generalization were consistent with what was observed in her everyday interactions. In general AW needed much more time to "figure out" what was being instructed during speech therapy sessions. Nearly 20 minutes was spent on shaping the phoneme /r/ in each session, and it was not uncommon for AW to not be able to produce one correct /r/ phoneme during an entire hour-long treatment session. During the speech sessions she often became very frustrated and said that it was just "too hard". Very often the clinician would speak to AW about her speech, asking if she felt that her speech was different from children her age, to which she would always reply, "yes". This was used as an encouraging factor when she became especially frustrated.

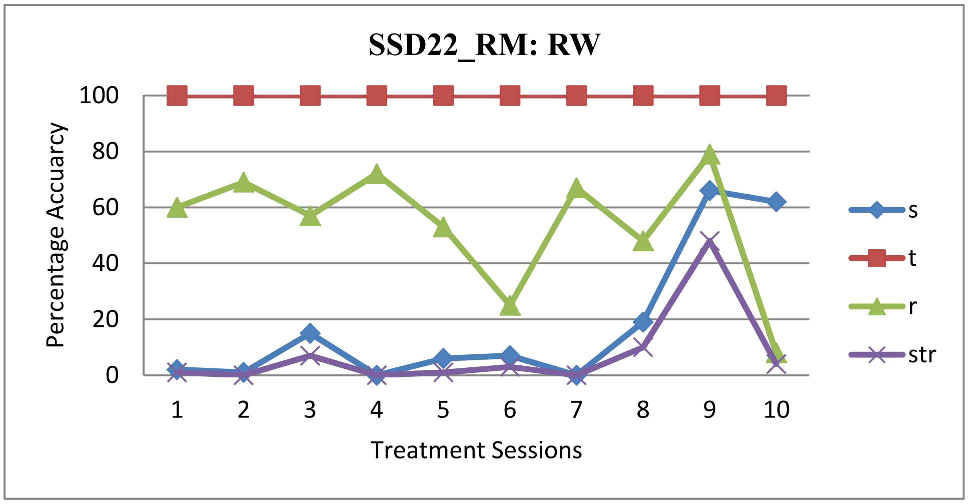
AW had been receiving speech therapy since she was in kindergarten. At 7;4 she was the one of the oldest participants in the study. In terms of age, it could be thought of years of "bad" practice. So, out of all of the participants she had the longest time of practicing incorrect speech productions. AW also a familial history of speech and language difficulties. Her mother reported that she herself had a speech impairment as a child and had received speech therapy for seven years. AW had four siblings, all of whom had speech and language difficulties to varying degrees.

As previously stated, the NW treatment condition did not appear to cause much change in AW's phonology, either within the confines of the treatment cluster and words or in a system-wide manner to other untreated sounds, clusters, and words. Anecdotally, AW did struggle to learn the word-picture mappings of the NWs. During the final treatment session she still could not produce the target words without some form of "clue" from the clinician (i.e., a carrier phrase). AW was on an Individual Education Plan (IEP) for a SSD as well as a learning disability and did receive additional resource services in school. She did not do well on the standardized language assessment and throughout the treatment demonstrated inappropriate language skills. Thus, the combination of the complex treatment target and unfamiliar words may have been too difficult for AW.

4.6. SSD22_RM: Real Word Treatment.

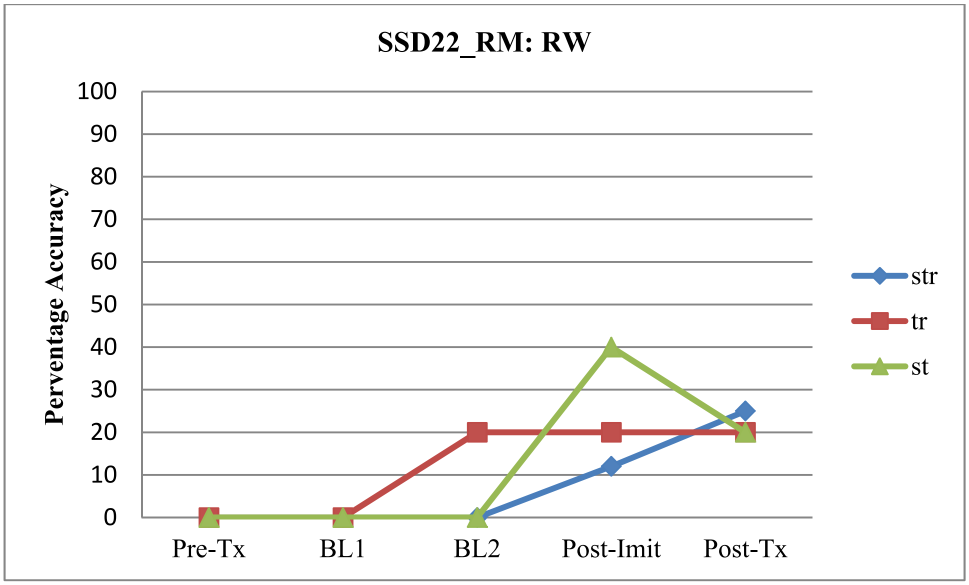
4.6.1. Learning Curve. The learning curve for SSD22_RM, who was treated using RW stimuli, is displayed in Graph 4.13. RM showed no productive phonological change until well into the final stages of treatment sessions. In the 9th session, RM had a breakthrough with the target cluster producing it with nearly 50% accuracy. This higher accuracy quickly diminished, and at the end of the treatment period, little change was shown. RM's learning curve can be described as fairly nonmonotonic, as her accuracy was sporadic throughout treatment. At the end of the treatment period, RM appeared to reach a variable production accuracy of approximately 10%. During the course of treatment, she produced an average of 126 productions per 1-hour session with a minimum of 82 and a maximum of 151.

Graph 4.13. SSD22_RM's productions of the target cluster /str/ across treatment sessions.



4.6.2. Generalization from Treatment. The cluster probe generalization curves for the production the treated three-element cluster /str/ in untreated words and the untreated clusters /st/ and /tr/ are shown in Graph 4.14. Accuracy levels are plotted over each administered probe (Pre-Treatment, Baseline 1, Baseline 2 (if applicable), Post-Imitation and Post-Treatment. Graph 4.14 shows a fairly monotonic curve, with slight improvement in the production of the treated and untreated clusters. RM did not show any change on the probes and this may be because no feedback or teaching was given through her productions of the probe words. RM depended on the clinician's cues (both verbal and visual) in order to correctly produce the target cluster, specifically with placement and shaping of the /s/ segment and to a lesser degree, the /r/ segment of the target cluster /str/.

Graph 4.14. Cluster probe generalization curve for SSD22_RM, for treated the cluster /str/ and untreated clusters /tr/ and /st/.



Looking closer at RM's generalization of/ str/, /st/, and /tr/ in and untreated word containing those clusters, there were patterns of improvement. Both the target cluster and the untreated cluster, /st/, followed a delayed monotonic generalization curve. RM's generalization to /tr/ to untreated words followed a quicker course of change, showing improvement during the BL2. This BL2 was composed of selected MOCP words and was given after the full MOCP, so there was more opportunity to produce the cluster correctly. It is common to allow a +/- 10% change during the baseline sessions, but more than that would possibly indicate that change was occurring on its own (Gierut, 2001). However, since her production of /tr/ did not improve any beyond this probe, the first measurements might have been exceptionally poor productions and her real overall average for this cluster was approximately 20%. Her accuracy of /s/ was characterized by both lateral and dental production distortions throughout the whole treatment study.

Much of the clinician's time and effort was spent on "coaching" a correct production of a /s/. Her /r/ was also compromised as it was very common for her to replace [w] for /r/. Table 4.12 shows RM's exact production of the probe words to provide evidence of the delayed monotonic generalization to both the treated cluster in untreated words and the untreated clusters.

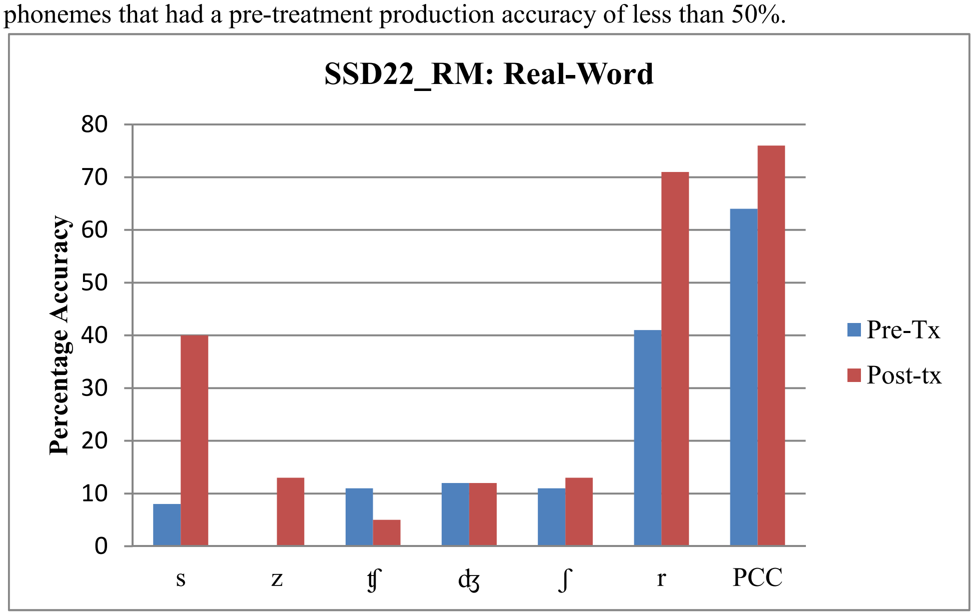
Table 4.12. SSD22_RM's productions of the MOCP words in probes across treatment.

Gloss	**Pre-Treatment**	**Post-Imitation**	**Post-Treatment**
tree	twi	twi	twi
star	*s̪tɑr	s̪tɑr	s̪tɑr
train	tweɪn	deɪn	tweɪn
stove	s̪toʊf	stoʊf	doʊv
truck	twʌk	trʌk	trʌk
triangle	twɑɪeɪŋgl	twɑɪeɪŋgl	twɑɪeɪŋgl
trick or treat	twɪk ɚ twit	twɪk ɚ twit	twɪk ɚ twit
street	s̪twit	s̪trit	s̪twit
string	s̪twɪŋ	s̪trɪŋ	strɪŋ
straight	s̪tweɪt	s̪tweɪt	s̪tweɪt
stop	s̪tɑp	stɑp	s̪tɑp
stamp	s̪tæmp	s̪tæmp	s̪tæmp
strainer	stweɪnɚ	s̪tweɪnɚ	s̪treɪnɚ
stranded	s̪twændɪd	s̪twændɪd	stwændɪd
straw	s̪trɑ	s̪trɑ	s̪trɑ
strap	stwæp	stræp	stræp
store	s̪tor	s̪tor	stor

*Note the dental marker below the /s/ segment of the target cluster. This was used so the reader could gain further insight into RM's exact productions.

While studying the generalization of treated and untreated clusters is of great interest and importance to understanding of how effective sound cluster treatment is, it is also important to study the broad generalization of untreated singleton sounds. This task is imperative to determining the efficacy of the real-word treatment approach. It provides information as to whether or not the treatment approach caused change to the phonological system as a whole. In order to determine broad generalization of sounds, the percentage of accuracy of singleton sounds was calculated. In regard to her untreated singletons, RM showed very minimal improvement (Graph 4.15). Six sounds were produced with less than 50% accuracy during the Pre-Treatment AEP probe: /s/, /z/, /ʧ/, /ʤ/, /ʃ/, and /r/. Note that RM had the /s/ phoneme in her inventories, but it was produced incorrectly with the lateral and dental distortions. By the end of treatment, her productions of /s/ improved so that nearly 50% of the time it was produced in an adult-like manner, and not in a lateral or dental manner. She could always produce a /s/, which is different from other participants who would substitute a different sound (e.g., [ʃ] for /s/). Of these six sounds, one of them showed generalization (i.e., greater than 50% accuracy on post-treatment AEP probe). This was the phoneme /r/, which was originally produced with 42% accuracy was produced with 71% accuracy post-treatment.

Graph 4.15. Production accuracy of SSD22_RM's percentage of consonants correct (PCC) and untreated phonemes that had a pre-treatment production accuracy of less than 50%.



Looking more closely at RM's untreated sounds, it can be said that her RW treatment did induce some phonological change in the monitored untreated sounds (Graph 4.15). While only one of the six untreated sounds, /r/, showed greater than 50% accuracy at the Post-Treatment probe, three other sounds did show an increase in production accuracy (i.e., /s/, /z/, and /ʃ/). RM's overall PCC showed improvement, gaining 12% accuracy compared to pre- and post-treatment (i.e., 76% compared to 64%). Her production of /s/ nearly reached the 50% accuracy generalization criteria at the Post-Treatment probe, and it was added into both of her phonetic inventories. Because of RM's distortions of the /s/, her level of accuracy of this singleton had to be done by hand. RM's productions of /s/ had to be counted by hand because the LIPP software did not account for diacritics when calculating the percentages of accuracy. For example, the transcriber inserted her productions of /s/ as dentalized but the software itself only accounted for the /s/ production disregarding the dental marker. Therefore, the chance of an error occurring in calculation of accuracy is a slim possibility.

4.6.3. Phones Added to Phonetic and Phonemic Inventories. The inventory analyses provided more evidence of system-wide change in RM's phonology. She added four sounds to her phonetic inventory by context: /s/, /z/, /l/, /r/ (Table 4.1). Recall that RM had a dental or lateral /s/, but not an adult-like /s/. The dental or lateral /s/ was considered to be a separate sound than the adult-like /s/. She also added two sounds to her context-free phonetic inventory: /s/ and /l/ (Table 4.1). She added an additional five sounds to her phonemic inventory: /v/, /s/, /z/, /l/, /r/ (Table 4.2). These results are quite complementary to her sound generalization results, given that she only generalized /r/ with greater than 50% accuracy. Recall however, that her production of /s/ nearly reached the 50% accuracy mark and because it occurred in all of her inventories it is likely that generalization of /s/ actually occurred.

4.6.4.Variability in Production Patterns. RM showed limited variability during the pre-treatment and post-treatment probes. During the pre-treatment baseline probes, RM produced the target cluster /str/ with two exemplars, /stw/ and /s̪tw/. These exemplars remained consistent from pre- to post-treatment probes. Using the MOCP, all clusters in the English language were assessed and evaluated for variability in production patterns. Recall that all other cluster productions, minus the treatment cluster /str/, in the MOCP were compared from pre- to post-treatment. When all of the cluster substitutions were compared from pre- to post-treatment, RM's Error Consistency Index (ECI) demonstrated minimal change. For the untreated clusters produced pre-treatment, she averaged 1.47 different exemplar substitutions for each cluster and this amount decreased to 1.44 during the Post-Treatment MOCP. Refer to Table 4.3 for a more comprehensive comparison of pre- and post-treatment ECIs.

Table 4.13. SSD42_RM's treated cluster during the pre-treatment MOCP sessions and her production substitution exemplars for the cluster. Note that correct adult-like productions may also be listed.

Target Sound		**Production Exemplars**		
str	Pre	stw	s̪tw	
	Post	stw	s̪tw	str

4.6.5. Discussion. SSD22_RM was assigned to the real-word condition, using a complex-three-element cluster as the treatment target Throughout her time in the treatment study, similar patterns of both onset and course of phonological change were observed for both treated and untreated sounds. During treatment, RM's production of her treated sound began to show signs of learning towards the end of the treatment sessions. Her learning curve followed a fairly nonmonotonic curve throughout the treatment sessions, with her accuracy varying from session to session primarily depending on which sound the session focused on that day, /s/ or /r/. RM showed similar patterns in generalizing her treated cluster and untreated clusters, /st/ and /tr/ to untreated probe words. All three clusters showed a delayed onset in improvement, and if change occurred, it varied across assessments. Six single phonemes were monitored for generalization from the pre-treatment AEP probe to the post-treatment AEP probes; of these six phonemes, one showed generalization. Thus, minimal generalization due to real-word treatment was observed.

SSD22_RM was the oldest child in the treatment study, at age 7;7. She began this treatment study having already completed several years of speech therapy. RM's overall motivation to make progress in her speech sounds was often minimal at best; she was unwilling to monitor her own speech. Since self-monitoring is a key step in the generalization and improvement of speech, the fact that she refused to do it may be a key reason for her lack of improvement. Also, it really seemed to be the case that some days she was more focused on her /s/ and /r/ productions than other days. She could easily fix an incorrect /r/ and a lateral/dental /s/; however, she appeared to be unbothered by her distorted /s/ sounds, and only somewhat aware of when her /r/ was incorrect.

RM's speech was characterized by distortions (i.e., dental and lateral productions) of all strident and affricate sounds, as well as distortions of the /l/ and /r/ phonemes. RM's speech therapy had primarily focused on the /r/ and /s/ singleton productions for over a year, with very little progress being shown. Because of this, the cluster /str/ was an extremely complex target for her, as she had difficulty producing two of the three cluster elements correctly even in isolation. It could be argued that RM may not have been a good candidate for this research study, but her phonological system did benefit. Treatment sessions were often tedious as much of the time was spent shaping and segmenting each sound, but it is possible that the use of the most complex cluster possible for RM resulted in a greater amount of phonological change than would treatment focusing on singletons alone. These findings are consistent with the complexity theory (Gierut, 1999, 2001, 2007) that when the most complex sounds are targeted they will induce the largest amounts of phonological change.

4.7. Results: Summary

The present study compared the phonological change that occurred in five children after they participated in a treatment study using a complex treatment target (i.e., three-element cluster /str-/) in RWs and NWs. The treatment targets were then manipulated in terms of word lexicality: real words and non-words. The issue of word lexicality was addressed at the individual subject level by examining the learning (of the treated cluster /str/) that occurred during treatment, the generalization of the treated cluster in untreated words, the generalization of untreated clusters /st/ and /tr/ in untreated words, the addition of sounds to the phonetic and phonemic inventories, the increase or decrease of production accuracy across monitored singletons, as well as the consistency of the substitution errors. When taken together the data suggests that, while the outcomes of both treatment conditions were very similar, with a nearly 50/50 split towards both treatment conditions (Table 4.14). These results suggest that RWs may induce larger phonological change in some measurements while NWs may induce larger change in other measurements. This phenomenon will be discussed at length in the next section. These findings may have implications for the structure of clinical intervention, as well as future SSD treatment efficacy research. The results of using RW and NW in treatment are discussed in terms of learning during treatment, generalization from treatment, system-wide phonological change, and error consistency.

Table 4.14. Summary of group post-treatment results in terms of generalization of treated cluster and untreated singletons, learning curves for treated and untreated clusters, phonetic/phonemic inventories, and substitution variability for treated and untreated clusters.

Variable	Real Word Treatment		Non-Word Treatment
Quantitative Phonological Change			
% Accuracy: Treated	49%	Real Word > Non-Word	16%
% Accuracy: Untreated	23%	Real Word > Non-Word	12%
Qualitative Phonological Change:			
Course: Treated	Monotonic, Monotonic	Real Word > Non-Word	Flat, Flat, Monotonic
Course: Untreated	Monotonic, Monotonic	Real Word > Non-Word	Flat, Flat, Monotonic
Phonetic Inventories by Context:			
Average of Sounds Added	4	Real Word ≥ Non-Word	3.3
Average of Sounds Generalized	2.5	Non-word > Real Word	3.5
Sounds Added & Generalized*	5 of 8	Non-word ≥ Real Word	7 of 10
Context-Free Phonetic Inventories			
Average of Sounds Added	2	Real Word > Non-Word	.66
Sounds Added & Generalized*	1 of 4	Real Word > Non-Word	0 of 2
Phonemic Inventory			
Average of Sounds Added	4.5	Non-Word > Real Word	5
Error Consistency			
Decrease in Substitution Errors: Treated	26%	Non-Word > Real Word	10%
Decrease in Substitution Errors: Untreated	16%	Real Word ≥ Non-Word	14%

*"Sounds Added and Generalized" refers to the number of sounds that were both added to the phonetic inventory and produced with more than 50% accuracy on the post-treatment AEP probe.

Given the single-subject design of the treatment study, the individual results provided varying demonstrations of phonological change due to treatment, which can be difficult to generalize to a group as a whole. All of the participants demonstrated varying degrees of phonological change, with some being similar more similar than others. The following sections will also address the overall results and interpret them at a group level for a more thorough discussion of treatment effectiveness.

4.7.1. Learning during Treatment. The learning patterns of the participants suggest that the treatment was efficacious and valid, regardless of whether it targeted RWs or NWs. All five participants began to show change in their production of the treated cluster /str/ at least by the second session of the Spontaneous Production phase of treatment; however, only two of the participants surpassed 50% accuracy in the production of their treatment sounds for more than two sessions. Three of the five children demonstrated continual improvement through the end of their treatment sessions. RM's and HJ's productions did show improvements in accuracy, but their productions varied from session to session and cannot be described as "continual improvement". It is likely that productions accuracies would have continued to improve had more treatment sessions been allowed. None of the participants demonstrated a plateau in her/her productions. This indicated that through the entire course of treatment their phonological systems were changing.

Some factors must be taken into consideration while evaluating the learning during treatment. As discussed in the individual participants' results, the variation in the number of treatment token productions was potentially a factor in how fast each child's productions of the treatment target began to change. For example, TP demonstrated the fastest amount of change in the treatment cluster and he produced, on average, the highest number of treatment words in a given treatment session. Alternatively, HJ demonstrated a slow course of change and was the participant who, on average, produced the fewest number of treatment words in a given session. It is likely that the number of productions spoken may be just as important as the number of treatment sessions in determining which children will demonstrate phonological change. These findings do not suggest that the use of NWs in treatment is not efficacious, but rather it may be necessary to only use NWs with children who do not have language impairments. For example, both AW and HJ had very low scores on the standardized language assessment and they both demonstrated minimal learning throughout the treatment. On the other hand, the other participant in the NW group (BS) demonstrated a monotonic learning curve. Based on these findings, AW and HJ might have benefitted more from treatment using RWs based on their language skills.

4.7.2. Treatment Cluster Generalization to Untreated Words. Overall, there was a large amount of individual variation in terms of how much generalization occurred in the production of the treated cluster /str/ in untreated contexts. Children in both treatment groups displayed immediate, delayed, and/or no measurable change during the treatment. Treatment with RWs arguably induced more rapid gains in production of the treated sound as a whole, while two of the three participants in the NW group demonstrated no measurable change in the treatment cluster during treatment. One of the participants (TP) in the NW group displayed the most immediate measureable change during treatment. Thus, it is very difficult to claim that the treatment the children received was the factor that was responsible of the observed phonological changes displayed in both groups.

Again, any change in productions that was closer to an adult model should be considered a positive effect, but for the purposes of this analysis only immediate change can be directly attributed to treatment.

4.7.3. Generalization of Untreated Clusters and Untreated Singletons in Untreated Words. While not every child demonstrated generalization to the treatment cluster in untreated contexts, nearly all of the participants demonstrated at least some phonological change in the untreated clusters /st/ and /tr/ that made up the target cluster /str/, and with their untreated sounds that were initially produced with less than 50% accuracy. The topic of generalization to untreated clusters /st/ and /tr/ will be discussed first.

Both immediate onset and delayed onset patterns of phonological change were observed in terms of the untreated clusters. One participant (HJ) demonstrated no measurable change for the untreated clusters. All other participants demonstrated a monotonic learning for the production of the untreated clusters /st/ and /tr/. Both the RW and NW treatments appear to have the potential to cause change in untreated clusters that are contained within a given three-element cluster.

It is important to take into consideration the amount of generalization that occurred in untreated singleton sounds that were initially produced with less than 50% accuracy. Again, both groups demonstrated at least some phonological change in their untreated phonemes. Similar to the previous findings, an overarching pattern is hard to find. Children in the NW group generalized more of the untreated sounds that were initially produced with less than 50% accuracy as compared to children in the RW group. It should be noted however, that not all of the participants displayed generalization in untreated sounds. One of the participants (AW) did not generalize any of the untreated sounds that she initially produced with less than 50% accuracy. Thus, while the NW condition might have induced slightly higher rates of phonological change, the RW condition was also able to create measurable change in the participants' phonologies.

4.7.4. Sounds Added to Phonetic and Phonemic Inventories. Every participant in the study added at least one phone into his or her phonetic and phonemic inventories. When compared directly, children in the RW condition added an average of 4 sounds to their phonetic inventory by context, with a total of 9 sounds in their phonemic inventories while children in the NW condition added an average of 3.3 sounds to their phonetic inventory by context and 10 total sounds in their phonemic inventory. The context-free phonetic inventories yielded smaller results with the RW adding an average of 2 sounds and the NW adding an average of less than 1 sound to their inventories. Overall, children in the RW condition displayed a greater number of sounds added into the phonetic inventories while children in the NW condition displayed a greater number of sounds added into their phonemic inventories.

4.7.4. Variability in Production Patterns. Nearly every child in both conditions decreased the number of substitution errors for both the treatment cluster /str/, as well as the untreated clusters in the MOCP.

At the group level, children in the NW treatment condition demonstrated a larger amount of change from pre- to post-treatment in the number of error variants produced for the treatment target /str/. However, the amount of error variants for children in both conditions was very comparable (30% and 40%, respectively). This suggests that both treatments did induce some amount of change across the treated cluster. Looking at all of the untreated clusters assessed in the MOCP, the ECI scores for the two treatment conditions are also very similar, with a very small advantage given to the real word condition (Table 4.3 and 4.14).

4.8. Results: Direct Comparisons

Five children were selected to participant in the present study. Some of them had very comparable phonological systems and some did not. For effective single-subject treatment designs, it is desirable for each child to have a “match” in the other group. While this goal was not attained due to time constraints and participants dropping out of the study, there were four participants who were very evenly matched. Because of this, two separate pairs of participants will be compared directly, as they began the treatment study with very similar phonological systems and ecological factors. The following sections will discuss these four participants in pairs and directly compare their results. This will allow for a clearer picture of the treatment outcome, and will allow us to determine if using the three-element cluster /str/ was efficacious.

4.8.1. SSD42_TP and SSD39_BS: A Direct Comparison. TP and BS had very similar phonological systems upon entering this research study. They were also within months of the same age and had had very similar histories of previous speech therapy. TP participated in the RW group, while BS was chosen for the NW group.

Upon entering the study, TP had eight sounds out of his phonetic inventory and eleven sounds missing from his phonemic inventory. These sounds consisted primarily of affricates, fricatives, laterals, and strident sounds. His pre-treatment ECI score was one of the highest, having an ECI of 5 for the treatment cluster /str/ and 2.32 for all other untreated clusters. Upon entry to the study, he was stimulable for each of the sounds in the treatment cluster but could not produce the treatment cluster correctly as a whole. He often substituted [ʃ] for /st/ and [w] for /r/ when trying to produce the cluster. These initial characteristics were very comparable to BS. Upon entering the study, BS had ten sounds missing from his phonetic inventory and eleven sounds missing from his phonemic inventory. These sounds consisted primarily of affricatives, fricatives, stridents, and lateral sounds just like TP. His ECI score was very similar to TP's, being 5 for the treated cluster and 2.22 for all other untreated clusters. He was stimulable for each of the sounds in the treatment cluster, but when he tried to produce the target cluster he would often substitute /ʃ/ for /st/.

Thus, prior to treatment TP and BS had comparable phonological systems and a hypothesis was formed that their phonological systems would undergo very similar changes. Their learning curves, generalization curves, phonemic and phonetic inventories, variability consistency, and their generalization of untreated sounds are presented in Table 4.15.

Table 4.15. A direct comparison of SSD34_TP and SSD42_BS's post-treatment results in terms of generalization of treated cluster and untreated singletons, learning curves for treated and untreated clusters, phonetic/phonemic inventories, and substitution variability for treated and untreated clusters.

Variable	SSD34_TP: RW		SSD42_BS: NW
Quantitative Phonological Change			
% Accuracy: Treated	88%	Real Word > Non-Word	50%
% Accuracy: Untreated	34%	Real Word > Non-Word	20%
Qualitative Phonological Change:			
Course: Treated	Monotonic	Real Word = Non-Word	Monotonic
Course: Untreated	Monotonic	Real Word = Non-Word	Monotonic
Phonetic Inventory by Context:			
Sounds Added	6	Non-word = Real Word	6
Sounds Generalized	4	Real word > Non-Word	2
Sounds Added & Generalized	4 of 6	Real word > Non-Word	2 of 6
Context-Free Phonetic Inventory			
Sounds Added	2	Real Word > Non-Word	1
Sounds Added & Generalized	1 of 2	Real Word > Non-Word	0 of 1
Phonemic Inventory			
Sounds Added	4	Non-word > Real Word	6
Error Consistency			
Decrease in Substitution Errors: Treated	60%	Real Word > Non-Word	40%
Decrease in Substitution Errors: Untreated	29%	Real Word > Non-Word	15%

When directly compared, the post-treatment results of TP and BS are also similar. However, it appears the using RW for the treatment stimuli induced more change in TP's phonological system than NW did for BS. These findings are consistent with the overall results of the RW group showing more phonological change than the NW group. While these two participants were the "perfect pair", there were two other children in the study who were also very comparable and could quite possibly provide more information of treatment efficacy regarding this study.

4.8.2. SSD22_RM and SSD39_AW: A Direct Comparison. AW and RM were the only females in this study, as well as the oldest at 7;4 and 7;7. Their similarities did not end there, however. Both had received at least two years of speech therapy, and their speech sound improvement had plateaued. Their phonological systems were comparable, as they each one sound class missing from both their phonetic and phonemic inventory. AW was not able to produce a /r/ phoneme, regardless of the level of visual and verbal cues. RM had been working on strident sounds (e.g. /s/, /z/, /ʃ/) as well as the /r/ phoneme for nearly two years and had very minimal progress to show for it.

Both children were not stimulable for some of the sounds contained within the target cluster, and the clinicians spent a lot of time shaping their "problem" sounds (i.e., /r/ for AW and /s/ and /r/ for RM). Neither RM nor AW could self-monitor their incorrect productions and they showed little motivation or interest in improving their speech. Because of their age, they had both spent a longer time than the other participants practicing those incorrect speech patterns. By comparing their post-treatment results, more information can be gathered regarding the use of a complex target to treat SSDs in older children with few sounds in error.

Table 4.16. A direct comparison of SSD22_RM and SSD39_AW's post-treatment results in terms of generalization of treated cluster and untreated singletons, learning curves for treated and untreated clusters, phonetic/phonemic inventories, and substitution variability for treated and untreated clusters.

Variable	SSD22_RM: RW		SSD39_AW: NW
Quantitative Phonological Change			
% Accuracy: Treated	27%	Real Word > Non-Word	0%
% Accuracy: Untreated	12%	Real Word > Non-Word	1%
Qualitative Phonological Change:			
Course: Treated	Monotonic	Real Word = Non-Word	Flat
Course: Untreated	Monotonic	Real Word = Non-Word	Flat
Phonetic Inventory by Context:			
Sounds Added	4	Real word > Non-Word	1
Sounds Generalized	1	Real word > Non-Word	0
Sounds Added & Generalized	1 of 4	Real word > Non-Word	0 of 1
Context-Free Phonetic Inventory			
Sounds Added	1	Real Word > Non-Word	0
Sounds Added & Generalized	0 of 1	Real Word = Non-Word	0 of 0
Phonemic Inventory			
Sounds Added	4	Non-word > Real Word	6
Error Consistency			
Decrease in Substitution Errors: Treated	-50%	Real Word > Non-Word	0%
Decrease in Substitution Errors: Untreated	3%	Real Word > Non-Word	0%

When directly comparing these two participants, it is obvious that very minimal phonological change occurred due to treatment. These results may be indicative of how future treatment using three-element clusters should be conducted. While they both made minimal progress, RM demonstrated more phonological change than AW. These findings are consistent with the other findings in both the overall results and the results that were found directly comparing BS and TP, which were that the participants treated with RWs demonstrated more phonological change than the participants treated with NWs.

4.9. Real Word vs. Non-Word: A Final Summary.

In this treatment study, RW treatment created more change in the participants' phonologies than did NW treatment. These treatment results both extend and clarify previous findings by Gierut et al. (1999, 2002). A detailed discussion of the results and the clinical implications that they may have will be discussed at length in the following section.

CHAPTER 5

DISCUSSION

5.1. Discussion Introduction.

Recall the questions that define treatment efficacy introduced in the very first paragraph of this paper: Does the treatment work? In what ways does treatment alter behavior? Does one treatment work better than another? This research study set out to compare the treatment effectiveness of RWs and NWs in a complex three-element cluster. After close examination of the treatment results, we are able to answer the question: "Which are a more effective treatment target in children with SSDs: RWs or NWs?" The following section will discuss the current research studies results and directly compare them to past studies of similar nature. By doing this, it is hoped that an overall idea of treatment efficacy in terms of complexity and word lexicality will emerge or be strengthened. This study is not without limitations and those will be discussed as well. This study has also introduced other unanswered questions that may warrant further research in order to gain more insight. These questions and their implications of further research will also be discussed at length.

5.2. Treatment Cluster.

In examining the interactions of a three-element cluster had on the phonological system of children with SSDs, an overall idea has emerged. Nearly all of the children in the research study exhibited some sort of phonological change, be it adding sounds to their inventories, reducing their amount of substitution variability, or increasing their overall PCC. Regardless of whether the children were in the RW or NW group, change occurred. Because of this it can be said that using a three-element cluster target to treat children with SSDs is efficacious.

Recall that Gierut & Champion (2001) evaluated linguistic complexity focusing on word-initial 3-element consonant clusters. The use of 3-element clusters in treatment resulted in similar learning patterns across the participants regardless of the specific cluster being taught. All children increased their production accuracy from 0% baseline levels to 83-92% post-treatment levels. As for generalization, there was little or no transfer of learning to 3-element clusters, whether treated or untreated. Regardless of the accuracy level at which the participants were producing the target cluster during their last session, they returned to their baseline following treatment.

The findings in this research study corresponded well with the findings by Gierut & Champion (2001). Recall that every participant increased his or her production accuracy from 0% baseline levels to at least 15% post-treatment levels. As for generalization, many of the participants did not transfer the learning (generalize) the treated cluster regardless of the level that they were producing it at during the treatment sessions. Unlike, Gierut & Champion (2010) some of the participants did transfer learning and were able to generalize the treated three-element cluster following treatment. This may be because of the children's relatively stable phonological systems and their age, both of which would have allowed for generalization of the treated three-element cluster /str/.

Gierut & Champion (2001) found that the treatment of 3-element clusters did not induce generalization to like-3 element clusters, but it did trigger broader changes in other onsets, including singletons, affricates, and 2-element clusters. In that study, five untreated sounds, on average, which were produced with 0% accuracy at baseline were added to each child's inventory after treatment. On average, the participants in this study added 4.2 untreated sounds to their phonetic inventory by context after treatment. Upon completion of treatment, 3 of the 5 children in this study were accurately producing untreated affricates that were missing from their pre-treatment phonemic inventory and/or phonetic inventory by context. This is a significant finding as the study by Gierut & Champion (2001) observed similar results, with 4 of the 5 of their children accurately producing affricates. Some of the participants in this current research study did not show results of this nature, in fact they hardly showed any phonological change from treatment.

Recall that Gierut and Champion (2001) suggested that in determining which 3-element cluster may be most appropriate to teach, it would be valuable to consider the child's singleton phonetic inventory. Specifically, it would be beneficial if the child's singleton inventory included a subset of the consonants of the 3-element cluster. For example, if the child's inventory included the second consonant (C2) and the third consonant (C3) of the targeted cluster (e.g., /t/ and /r/ of /str-/), then /str-/ might be a preferred cluster to target for treatment. Recall that all but two participants (i.e., AW and RM) in the current research study had C2 and C3 of the target cluster /str-/. Specifically, the C3 cluster element, /r-/, was not in two of the participants inventories prior to beginning this treatment study. However, one of the two participants was stimulable for the sound while the other was not. This simple fact seemed to make a world of difference in their success in this treatment study.

AW did not have the C3 element of the treatment cluster /str-/ in her inventories and she was not stimulable for it. She was also one of the oldest participants in the study and she had also had two years of previous speech therapy. Because of this, her singleton inventory was very stable and was only missing one phoneme, /r/. Regardless of the previous findings of Gierut & Champion (2001) cautioning future researchers to only target a 3-element cluster if the child's singleton inventory included a subset of the clusters consonants, AW was still selected to participate in this research study. Gierut & Champion (2001) were right to caution, as the treatment program targeting the three-element cluster /str-/ induced little to no phonological change in AW.

Likewise, RM did not have the C3 element of the treatment cluster /str-/ in her inventories, but she was stimulable for it. She was the oldest participant in the study and had also received nearly three years of previous speech therapy. Along with not have the C3 element of the treatment cluster /str-/ in her inventories, the C1 element of the treatment cluster was also absent. The fact that RM was stimulable for both the C1 and C3 elements of the target cluster /str-/ may have greatly helped her in the course of this treatment program. While she did not show any learning or generalization to the treated cluster, her overall PCC did improve and both /s/ and /r/ were added to at least one of her post-treatment inventories. Thus, it seems that even though she did not have the C1 and C3 elements of the treated cluster /str-/ in her inventory, the fact that she was stimulable for them allowed for some phonological change to occur.

Overall, the findings in this research study indicate that treatment of three-element clusters resulted in both learning and generalization. In general, the participants displayed common patterns that included broad expansion of untreated singletons, acquisition of untreated affricates, and some transfer of learning to the treated 3-element cluster /str/. The findings also concur with previous research (Gierut & Champion, 2001), in that when using a three-element cluster as a treatment target, it is valuable to consider the child's singleton phonetic inventory. Specifically, it would be beneficial if the child's singleton inventory included a subset of the consonants in the three-element cluster.

5.3. Real Words vs Non-Words: A Discussion.

The issue of word lexicality was addressed at the individual subject level by examining the learning (of the treated cluster) that occurred during treatment, the generalization of the clusters /str-/, /tr-/, and /st-/ to untreated words, the addition of sounds to phonetic and phonemic inventories, the generalization of untreated singletons, as well as the consistency of the substitution errors. All of these issues have been discussed in length above in the results section, but they have yet to be compared to previous research. By comparing this research study's results to those found in previous research studies, we can hope to continue to address the questions of treatment efficacy formed at the beginning of this research study. Which were: Did the treatment work? In what ways did treatment alter behavior? Did one treatment work better than another?

5.3.1. Learning during Treatment. As discussed in the results section, the learning patterns of the participants suggested that the treated was efficacious and valid, regardless of whether it targeted real words or non-words. All five participants began to show change in their production of the treated /str/, but only two of the participants surpassed 50% accuracy in the production of the treatment cluster for more than two sessions. Three of the five participants demonstrated continual improvement through the end of the treatment sessions and it is likely that the production accuracies would have continued to improve had more treatment sessions been allowed. The individual treatments were intensive one-on-one (child-clinician) sessions, they only occurred twice a week for an hour. It is possible that had the treatments occurred more frequently (e.g., daily or even three times a week), the children would have shown even more improvement (e.g., Denes, Perazzolo, Piani, & Piccione, 1996). The children only received a maximum of 12 treatment sessions, rather than the standard 19 delivered by previous studies by Gierut and colleagues (2001, 2007, 2010) therefore just adding seven more treatment sessions could have greatly impacted their learning during treatment.

The changing learning curves show that children with SSDs are able to learn NWs and in effect treat them like RWs. Most of the children enrolled in treatment were able to learn the associations between the NWs and objects/actions presented in the short storybook context. Two of the participants in the NW group struggled with this task, but they were also the two children with the lowest scores on the standardized language assessment. They also demonstrated minimal learning throughout the treatment, which may have been directly corembedded with their inability to learn the novel words based on their co-occurring language disorders. The other participant in the NW group demonstrated an immediate monotonic learning curve, which corresponds with Storkel (2004). Recall that Storkel (2004) provided evidence that children with phonological delays can effectively learn the association between a nonsense word and its lexical and semantic representations, thus suggesting that the use of non-words in treatment is warranted.

While all of the participants demonstrated patterns of learning during the treatment phase, future studies should done to better control for the selection of participants in terms of their language abilities. Storkel & Morrisette (2002) predicted that the learning of NWs requires an additional cognitive processing, meaning that not only do the children have to establish a phonological representation and try to produce the word's phonological form correctly, but they also have to learn the association between the novel NW and its picture (i.e., creating a lexical representation). This prediction seems to be entirely true as two of the NW participants demonstrated significantly different learning patterns from that of the other participants. These two participants were also the children who had low scores on the standardized language assessment. Thus, perhaps the use of NWs was not the most efficacious course of treatment for these two participants based on their language skills alone. Future studies should better control for this potential confound.

5.3.2. Generalization of the Treated Cluster to Untreated Words. The generalization of the treated sound to untreated words is an important factor to consider in treatment effectiveness, as it would not be desirable if a child only learned to correctly produce a targeted treatment sound in the few words used in treatment. Overall, children in both treatment groups displayed immediate, delayed, and/or no measurable change during treatment. Treatment with RWs arguably induced more rapid gains in production of the treated sound as a whole, while two of the three participants in the NW group demonstrated no measurable change in the treatment cluster during treatment.

The results of the present experiment correspond somewhat with the previous findings of Gierut et al. (1999, 2002). This study, manipulated the word familiarity and word density in treatment targets; however, instead of controlling for one variable, each child in Gierut et al.'s (1999) study was provided with alternating treatments exposing them to two different conditions (e.g., treating a child on /s/ in high familiarity words and /ʃ/ on low familiarity words, with both types of words being variable in word density). Of the two children who were treated on low familiarity and high familiarity words, both children demonstrated earlier increases in their productive accuracy of their treated sound in high familiarity words, as compared to the sound in low familiarity words. Thus, there appeared to be at least some evidence that high familiarity words elicit earlier phonological changes than do low familiarity words. The present research study also alludes to the idea that high familiarity words elicit earlier phonological changes that do low familiarity words.

4.3.3 Generalization of Untreated Clusters and Singletons in Untreated Words. Nearly all of the participants demonstrated at least some phonological change in the untreated clusters /st/ and /tr/ that made up the target cluster /str/, and with their untreated sounds that were initially produced with less than 50% accuracy.

In the present study, both immediate onset and delayed onset patterns of phonological change were observed in terms of the untreated clusters. Gierut and Champion (2001) found that the results of teaching word-initial three-element clusters to children with functional phonological delays resulted in generalization to two-element /s/ clusters and non-/s/ true clusters. The results of the present study seem to agree with Gierut and Champion's (2001) findings, although on a smaller scale. In order to truly agree with Gierut and Champion, analyzing only two untreated clusters is not enough. Both real RW and NW treatments appear to have the potential to cause change in untreated clusters that are contained within a given three-element cluster. This does not clarify whether NWs or RWs were more effective in terms of generalization to untreated clusters, as there was simply not enough information analyzed in this research study. There simply was not enough time or resources to analyze the generalization of all untreated clusters. Future research should address this issue by examining all untreated clusters.

The amount of generalization that occurred in untreated singleton sounds was the area where most of the analysis took place. Recall that singleton sounds that were initially produced with less than 50% accuracy were monitored across the treatment phase. Once again, an overarching pattern of generalization was hard to find. Children in the NW group generalized more of the untreated sounds that were initially produced with less than 50% accuracy, as compared to children in the RW group. It should be noted however, that not all of the participants in the NW group displayed generalization in untreated sounds. Cummings and Barlow's (2011) results also showed similar results in terms of generalization to untreated singletons. The RW and NW treatment data showed similar levels of generalization, but there appeared to be a slight advantage for using NWs in the area of generalization to untreated singletons. In Cummings and Barlow's (2011) research program, the non-word condition might have induced slightly higher rates of phonological change, but the real word condition was also able to create measurable change in the participants' phonologies.

The current research study conducted two probes in order to directly assess the generalization that had taken place based on treatment: a pre-treatment probe and a post-treatment probe. There was no probe given post-treatment after time had passed and this may have directly impacted the integrity of the results. In previous studies of the same nature (Cummings & Barlow, 2011; Gierut & Champion, 2001; Gierut, 2010) a two-week post-treatment probe was given to ensure that the results shown are not just a result of the recent treatment but, rather a result of a true change in their phonological change. The fact that this research study did not enlist this two-week post-treatment probe is a true limitation to the findings. Future research studies should be conducted to control for this.

5.3.4 Sounds Added to Phonetic and Phonemic Inventories. Recall that every participant in this study added at least one phone into their phonetic and phonetic inventories. In terms of treatment efficacy, the RW condition added more sounds to their phonetic inventory by context than the children in the NW did. Recall that the phonetic inventory by context is a more conservative measure of the phonological inventories. This information has typically not been included in previous research programs (Gierut, 2001, 2007, 2010; Cummings & Barlow, 2010) but it was included in this research study for a specific reason.

When the pre-treatment context-free phonetic inventory was gathered it was apparent that the participants had very stable phonological systems. This meant that they had a limited amount of sounds excluded from their inventories. This could have been due to a number of factors, including both their age and that their phonological systems had already underwent some form of change because of previous speech therapy. Post-treatment their context-free phonetic inventory was assessed once again and the results yielded limited amounts of change in both the RW and the NW group (i.e., an average of 2 sounds and 1 sound, respectively). This small amount of change did not correspond with the larger amounts of generalization that had occurred and thus a look at their phonetic inventory by context was warranted. This inventory analysis allowed a better look into the phonological change that was taking place. In terms of the phonemic inventories there results were very similar, with the RW condition adding a total of 9 sounds and the NW condition adding a total of 10 sounds. Thus, once again, it appears that both the RW and NW condition resulted in phonological change for the participants. It is very difficult to pinpoint which treatment condition yielded better results in terms of phonological change for the participants.

5.3.5. Variability in Production Patterns. Nearly every child in both conditions decreased the number of substitution errors for both the treatment cluster /str/, as well as the untreated clusters in the MOCP. The children who showed the most number of pre-treatment variants for the treatment cluster /str-/ were the children that showed the highest production accuracy change during treatment.

Variability in sound production may provide some additional insight regarding not only which children are more likely to show change due to treatment, but also perhaps what type of treatment elicits more generalization. Variability can be interpreted as instability in the representation of a phoneme in a child's phonology (e.g., Tyler & Lewis, 2005); however, there are varying interpretations of what variation in sound patterns actually represents. Another view of variability is that it may represent a child's inability to learn from treatment. Previous research (e.g., Dodd & Bradford, 2000; Forrest, Dinnsen, & Elbert, 1997; Forest, Elbert & Dinnsen, 2000) has found that children with variable error patterns required more treatment sessions than children with consistent substitutions. In these studies, the children with variable error patterns still did not learn to generalize the target sound to untreated words in different positions. At the single subject level, this study does not provide support for the claim that the more varied a child's phonological production of the treatment target is at the outset of treatment, the more likely it is that a delayed, or lack of, response due to treatment will be observed.

At the system-wide level, the three children who had the highest pre-treatment ECI scores (TP, BS, and HJ) were the children who showed the greatest decrease in sound variability at the Post-Treatment MOCP. In other words, the more error substitution sounds a child had to eliminate, the greater the overall system-wide reorganization. There are a few reasons as to why these contradictive results were displayed. Recall that the other two participants (AW and RM) were older than the other three participants by at least two years. They had errors that were very specific and were not stimulable for some of the individual sounds contained within the treated cluster. For example, AW was not stimulable for the phoneme /r/ and this was one of the only errors present in her phonological system. Because of this, her phonological system did not lend itself to large amounts of reorganization.

While most the treatment evidence presented here suggests that children were more successful in treatment when they produced fewer error variants, there have been studies that are consistent with the findings in this study. For example, Tyler and Lewis (2005) observed that children who had both highly variable and highly consistent phonological substitutions both followed similar linear trends in both increasing their PCC, as well as decreasing the number of different error substitutes. Thus, neither group had a delayed response or lack of response to intervention. Tyler and Lewis (2005) claimed that the multiple substitutions in children's productions reflected instability in their representation of different phonemes. This instability in turn made the entire phonological system more likely to show reorganization and change.

At the group level, children in the NW treatment condition demonstrated a larger amount of change from pre- to post-treatment in the number of error variants produced for the treatment target /str/. However, the amount of error variants for children in both conditions was very comparable (30% and 40%, respectively). This suggests that both treatments did induce some amount of change across the treated cluster. Looking at all of the untreated clusters assessed in the MOCP, the ECI scores for the two treatment conditions are also very similar, with a very small advantage given to the real word condition (Table 4.3 and 4.18). Recall that the ECI in the present study only took into account the errors in clusters, whereas prior studies (i.e., Cummings & Barlow, 2011) took into account the errors in singletons. The fact that the current research ECI only took clusters into account may have effected the overall results and it is quite possible that if the singletons were taken into account the ECI scores may have pointed in a stronger direction (i.e., NW vs. RW).

5.3.6. Real word vs. Non-Word Treatment: A summary. In this treatment efficacy study, it seems that both RWs and NWs were effective treatment targets. In this research study, RWs and NWs both created substantial amounts of change in the phonological systems of children with SSD. RWs induced greater change in some variables while NWs appeared to elicit more change in other variables. This is an important factor to consider, as this may lead us to final conclusion regarding the place that RWs and NWs have in treatment.

It is possible that the reason behind why some of the variables were better for RW and others for NWs have to do with the specific nature of the measurements. For example, when comparing learning during treatment the RW condition resulted in immediate monotonic learning curves while the NW condition was extremely variable. Recall the prediction made by Storkel & Morrisette (2002) that the learning of NWs requires an additional cognitive processing, meaning that not only do the children have to establish a phonological representation and try to produce the word's phonological form correctly, but they also have to learn the association between the novel nonsense word the its picture (i.e., creating a lexical representation). This task may have been too difficult for the children in the NW group, thus the nature of their learning curve was affected. So, not only did the NW group need to learn how to say their treatment words correctly, they also had to learn to associate the new NWs with their respective pictures.

The most compelling evidence for the effectiveness of NW treatment was observed in the error variability analysis. Specifically, children in the NW treatment condition showed a larger reduction in the number of sound substitutions produced for their treatment target Thus, consistent with previous research of Cummings and Barlow (2011) and Gierut and Morrisette (2010), this study suggests that NWs may be just as effective as RWs in the treatment of SSDs.

5.4. Study Limitations.

While indirectly replicating previous research, this study is not without its limitations. All five of the children had received previous speech therapy upon entry into this research study. This could have directly affected the treatment results, as many of the children's phonological systems were deemed to be very stable. Because of this, small amounts of phonological change were seen in some of the measurements.

Another limitation of this study is that of the odd number of participants. When this research study began, there were six participants and each participant was evenly matched by their age and phonological system. One of the participants withdrew from the research study after two treatment sessions, misaligning the RW and NW treatment groups. Because of this one of the participants did not have a match and the groups could not be compared evenly. The other four participants were matched very evenly and even showed very similar results over the course of treatment. It would be extremely beneficial to find a match for SSD16 in order to gather a more conducive comparison of RW and NW treatment. The fact that there was not a 2-week post-treatment probe does not allow us to determine if the results were an actual change occurring in the participants' phonological systems or if it was simply due to the treatment that had just occurred. Future studies should be done to address this question.

In terms of the participants, the girls in the research study yielded the least amount of phonological change due to a number of reasons. Both of them were the oldest in the research study and they also had been receiving speech therapy for many years. Because of this, their phonological systems were very stable, and they had been practicing incorrect speech patterns for a number of years. If the target selection criteria established by Gierut et al. (2010) had been followed, neither one of these children would have been selected to participate in this research study simply based on their phonetic inventories not including consonants within the target cluster /str-/. Having these two children in this research study directly impacted the studies results, especially in terms of AW. Without these two children, the results may have looked very different in terms of providing evidence of treatment efficacy towards RW and NW.

Perhaps the biggest limitation of this study pertains to the RW treatment stimuli used. The treatment stimuli words were not directly assessed for word familiarity, but rather chosen based on their ability to fit into a story as well as being "child friendly". Because of the limited amount of words beginning with /str/ in the English language, there are actually not many high familiarity words that begin with /str/. Further studies should control for this limitation by more precisely controlling for word familiarity, preferably using high familiarity RWs as a more direct comparison to the NWs.

5.5 Conclusion

This treatment study compared the efficacy of RW and NW targets containing a word-initial three-element cluster in the treatment of children with speech sound disorders. Converging evidence from analyses of data from the single-subject design revealed that both RWs and NWs could be considered to be efficient and efficacious options when working with children who have SSDs. Specifically, RW treatment led to a quicker, more monotonic rate of learning, and greater number of sounds being added to the phonetic inventories. On the other hand, NW treatment led to a decrease in the number of error variants for the treated cluster, a greater increase to the PCC, and a greater number of sounds being added to the phonemic inventories. While a strong claim cannot be made as to whether RWs or NWs are more efficacious that the other when used as treatment stimuli, this study did yield interesting findings that can be applied clinically.

Treatment suggestions have emerged from the results, meaning that there are instances when RWs would be especially beneficial and instances when NWs would be especially beneficial. These treatment suggestions can be found in Appendix 7. The treatment target itself yielded phonological change in all of the participants regardless of its treatment condition. Thus it appears that both RW and NW can be efficacious when presented within a complex target, such as the three-element cluster /str-/.

Appendix 1. Words presented in the *Assessment of English Phonology* (AEP; Barlow, 2003)

glasses	other	sock	GLOVE I	CHEESE I	Smile ing	Climb ing
ghost	rocket	jeep	CHALK	SQUARE	Splash	Stir
six	space	doll	CHALK I	SQUARE I	Splash ing	Stir ing
vacuum	ocean	game	JUICE	COMB	Blow	Rake
elephant	teacher	nail	JUICE I	COMB I	Blow ing	Rake ing
zebra	finger	hammer	CUP	ICE	Skate	Read
zoo	wagon	hanger	CUP I	ICE I	Skate ing	Read ing
brush	chair	ladder	WITCH	PEACH	Break	Drive
blue	thirsty	magician	WITCH I	PEACH I	Break ing	Drive ing
plate	queen	rabbit	HAT	DUCK	Swing	Throw
knife	twins	chain	HAT I	DUCK I	Swing ing	Throw ing
shark	them	crayons	BADGE	RING	Buzz	Snow
teeth	zipper	grapes	BADGE I	RING I	Buzz ing	Snow ing
feather	pig	slippers	NOISE	WATCH	Catch	
green	screw	jump (rop)ing	NOISE I	WATCH I	Catch ing	
vanilla	shower	light (or lamp)	OFF	DOOR	Kiss	
shovel	theirs	spoon	OFF I	DOOR I	Kiss ing	
this/that	thunder	yawn	MOTH	BRIDGE	Ride	
there	vest	princess	MOTH I	BRIDGE I	Ride ing	
yoyo	treehouse	smoke	ROBE	MOUTH	Rub	
guitar	cob	thermometer	ROBE I	MOUTH I	Rub ing	
lemon	these/them	mop	TOOTH	WISH	Shave	
yellow	sun	chain	TOOTH I	WISH I	Shave ing	
stretch	pie	SOAP	VAN	STOVE	Cough	
popcorn	drum	SOAPI	VAN I	STOVE I	Cough ing	
mother	water	FISH	WEB	CAGE	Crash	
father	toes	FISHI	WEB I	CAGE I	Crash ing	
behind	twelve	BATH	HILL	SPRING	Swim	
yes	zero	BATH I	HILL I	SPRING I	Swim ing	
you	gum	TUB	TRAIN	GOAT	Fall	
present	three	TUB I	TRAIN I	GOAT I	Fall ing	
thank you	judge	THUMB	MUD	Laugh	Sleep	
please	jelly	THUMB I	MUD I	Laugh ing	Sleep ing	
strawberry	brother	FROG	PAGE	Fly	Dig	
snail	(french)fries	FROG I	PAGE I	Fly ing	Dig ing	
beehive	skunk	CAR	DOG	Run	Rain	
chicken	noise	CAR I	DOG I	Run ing	Rain ing	
cloud	quiet	LEAF	BED	Hug	Hiss	
vase	shoes	LEAF I	BED I	Hug ing	Hiss ing	
music	zipping	GLOVE	CHEESE	Smile	Climb	

Appendix 2. Probe of Modified Onset Clusters Adapted from Gierut (1998)

pr	pretzel	**fr**	fruit	**sm**	smell	**skr**	scream
	princess		frog		smile		scratch
	pretty		french fries		small		scrub
	prize		friend		smoke		scribble
	present		front		smooth		screw
br	bread	**thr**	three	**sn**	sneee	**skw**	squeak
	brush		throw		snowman		square
	broom		thread		snake		squirrel
	bridge		throne		snail		squeeze
	brown		throat		snack		squirt
tr	tree	**shr**	shrink	**sw**	swing	**tw**	twinkle
	truck		shrub		swim		tweety
	train		shrug		sweep		twelve
	trick-or-treat		shrimp		sweater		twins
	triangle		shred		sweet		twist
dr	dress	**pl**	plate	**spr**	sprite	**kw**	quack
	drum		plane		spring		queen
	drive		plant		spray		quarter
	draw		play		sprinkler		quiet
	drink		plug				quilt
kr	crayon	**bl**	black	**spl**	splash	**Cj**	music
	cry		blocks		splinter		viewmaster
	crawl		blow		split		cute
	crack		blanket		splitting		few
	cream		blue				beauty
gr	grass	**kl**	clown	**str**	stranded		
	green		clean		straw		
	grow		clock		street		
	grapes		clothes		straight		
	grandma		cloud		string		
					strap		
					strainer		

Appendix 3: Treatment stimuli used for /str/ treatment sessions.

stroonge: strunʤ

strange: streɪnʤ

strang: streɪ

strong: strɔ

strupe: strup

stripe: straɪp

streeboorazz: striburæz

strawberries: strɑbɛriz

stree: stri

straw: strɑ

Appendix 4: Story used to present words for /str/ treatment

How the Zebras Got Their **Stripes/Stroopes**

Greetings, children! I'm Professor Linus Pinstripe and I am on the trail of a **stroonge/strange** mystery: How did the Zebra get its **stripes/strupes**? I'm here at the local watering hole on the plains of Africa with some talkative creatures that say they know the answer to this **stroonge/strange** mystery of the Zebra's stripes. "Long ago," said Giraffe, "there were two kids of zebras – the all-white kind and the all-black kind. The all-black kind was this color because all they ate was **stree/straw**. Everyone knows that if you only eat **stree/straw** you will have black fur. The black zebras were also so hot in the African sun."

"Yes," said Ostrich, "and the white zebras only ate **streeboorazz/strawberries**! Everyone knows that if you only eat that fruit you will have white fur. The white zebras weren't hot from the sun, but lions could always find them."

"One day," said Giraffe, "the zebras went to Baboon for help. He gave the black zebras **streebooras/strawberries.** And he gave the white zebras **stree/straw.** The zebras had a hard time switching their food, because they found the other food to be **stroonge/strange**. After a few days of the white zebras only eating **stree/straw** and the black zebras only eating **streebooraz/strawberries**, a very **stroone/strange** thing happened. The zebras switched colors completely! Now the white zebras were black and the black zebras were white."

"They really thought their problems were solved," exclaimed Ostrich, "but they were still having the same problems with the **strange/strang** sun and lions. So, they went back to the Baboon for help."

Those poor zebras begged the baboon for help, "Please help us Baboon, we are still so scared of the sun and lions," said the Zebras.

"I have the answer to your **stroonge/strange** problem." Baboon said, "You have to eat a mix os the **streebooraz/strawberries** and **stree/straw**. If you eat both of these foods you will have both white and black fur. This should solve your problem!" So, the zebra's followed the advice and soon they have black and white ***strupes/stripes.*** They were suddenly ***struped/striped*** creatures that could outsmart the African sun and lions. Giraffe laughed, "The zebras soon realized that their **strupes/stripes** were a great way to escape the lion and the African sun. Even though the other animals laughed at their **strupes/stripes**, they kept eating **streebooras/strawberries** and **stree/straw** to keep them safe. Now, they are proud of their **strupes/stripes**, because even though they are **stroonge/strange** it keeps them **strang/strong** and healthy."

Appendix 5. Tracking sheet used to collect treatment target word accuracy across treatment sessions.

	s
	t
	r

	5			10			15			20			25			30			35			40			
	s	**t**	**r**	**s**	**t**	**r**	**s**	**t**	**r**	**s**	**t**	**r**	**s**	**t**	**r**	**s**	**t**	**r**	**s**	**t**	**r**	**s**	**t**	**r**	
stri/straw																									s t r
stroonge/ strange																									s t r
strups/ stripes																									s t r
strang/strong																									s t r
streeboorazz/ strawberries																									s t r

Appendix 6: Initial clusters in the English language organized by Sonority Distance. Adapted from Barlow, Taps, and Storkel (2010).

SD = 6	SD = 5	SD = 4	SD = 3	SD = 2	SD = -2	/s/CC
/tw-/	/bj-/	/br-/	/fr-/	/mj-/	/sp-/	/skw-/
/kw-/	/pr-/	/dr-/	/ɵr-/	/sm-/	/st-/	/spr-/
/pj-/	/tr-/	/gr-/	/ʃr-/	/sn-/	/sk-/	/str-/
/kj-/	/kr-/	/bl-/	/fl-/			/skr-/
	/pl-/	/gl-/	/sl-/			/spl-/
	/kl-/	/fj-/	/vj-/			
		/sw-/				

Appendix 7. The table below contains treatment suggestions as to when it is appropriate to use RWs or NWs with children as treatment stimuli.

Real Words might be more appropriate if the child presents with:		**Non-Words might be more appropriate if the child presents with:**
many sounds missing from their inventories		
		a low Percent Consonant Correct (PCC)
a need to learn a single sound		
		many sound substitutions
a need to promote change across the sound system		
a need to promote change across two-element clusters		a need to promote change across two-element clusters

REFERENCES

Attanasio, J. (1994). Inferential statistics and treatment efficacy studies in communication disorders. *Journal of Speech and Hearing Research*, 37, 755-759.

Barlow, J., (2003). *The assessment of English phonology,* San Diego, CA: San Diego State University.

Barlow, J. (2004). Consonant clusters in phonological clusters in phonological acquisition: Applications to assessment and treatment. CSHA Magazine. 34: 10-13.

Barton, D., Miller, R., & Macken, M. (1980). Do children treat clusters as one unit or two? *Papers and Reports on Child Language Development,* 18, 105-137.

Bauman-Waengler, J. (2008). *Articulatory and Phonological Impairments: A Clinical Focus*. (3rd ed.). Boston, MA: Allyn & Bacon.

Campbell, T. F., Dollaghan, C.A., Rockette, H.E., Paradize, J. L., Feldman, H.M., Shriberg, L. D., Sabo, D., Kurs-Lasky, M. (2003). *Risk factors for speech delay in three-year-old children.* Child Development, 74, 346-357.

Clements, G. N., 1990. "The Role of the Sonority Cycle in Core Syllabification." In John Kingston & M. Beckman, eds., *Papers in Laboratory Phonology I*, 283-333. Cambridge: Cambridge University Press.

Connell , P. & Thompson, C. (1986). Flexibility of single-subject experimental designs, Part III: Using flexibility to design or modify experiments. *Journal of Speech and Hearing Disorders*, 51, 214-225.

Coxhead, A. (2006). *Essentials of teaching academic vocabulary.* Boston: Heinle Cengage

Cummings, A. E & Barlow, J. A. (2011). A comparison of word lexicality in the treatment of speech sound disorders. *Clinical Linguistics & Phonetics.* 25: 265-286.

Denes, G., Perazzolo, C., Piani, A., & Piccione, F. (1996). Intensive versus regular speech therapy in global aphasia: A controlled study. *Aphasiology*, *10*, 385–394.

Dinnsen, D. A., Chin, S. B., & Elbert, M. (1992). *On the lawfulness of change in phonetic inventories.* Lingua, 8, 207-222

Dinnsen ,D., Chin, S., Elbert, M., & Powell, T. (1990). Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech and Hearing Research.*

Dinnsen, D.A., & Gierut, J. A. (2008b). *Optimality theory: phonological acquisition and disorders.* London: Equinox.

Elbert, M., Dinnsen, D. A., & Powell, T. W. (1984). *On the prediction of phonologic generalization learning patterns*. Journal of Speech and Hearing Disorders, 49, 309-317.

Fey, M.E.(1986). *Language Intervention with Young Children.* San Diego, CA: College-Hill Press

Forrest, K., Dinnsen, D., & Elbert, M. (1997). Impact of substitution patterns on phonological learning by misarticulating children. *Clinical Linguistics and Phonetics*, 11, 63-76.

Forrest, K., Dinnsen, D., & Elbert, M. (2000). The effect of substitution patterns on phonological treatment outcomes. *Clinical Linguistics and Phonetics*, 14, 519-531.

Gierut, J. (1985). *On the relationship between phonological knowledge and generalization learning n misarticulating children.* Indiana University Dissertation. Dissertation Abstracts International, 46(6-B).

Gierut, J. (1990). Differential learning of phonological oppositions. *Journal of Speech and Hearing Research*, 33, 540-549.

Gierut, J. (1991). Homonymy in phonological change. *Clinical Linguistics & Phonetics*, 5, 119-137.

Gierut, J. (1992). The conditions and course of clinically induced phonological change. *Journal of Speech and Hearing Research*, 41, S85-S100.

Gierut, J. A. (1994). *Cyclicity in the acquisition of phonemic distinctions*. Lingua, 94, 1-23.

Gierut, J., (1998). Natural domains of cyclicity in phonological acquisition. *Clinical Linguistics and Phonetics*, 12, 481-499.

Gierut, J. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research*, 42, 708-726.

Gierut, J. A. (2001). *A model of lexical diffusion in phonological acquisition.* Clinical Linguistics & Phonetics, 15, 19-22.

Gierut, J. A. (2004a). *Clinical application of phonological complexity.* CSHA, 6-7, 16.

Gierut, J. (2007). *Phonological complexity and language learnability.* American Journal of Speech-Language Pathology, 16, 6-17.

Gierut, J & Champion, A.H. (2001). Syllable onsets II: Three-element clusters in phonological treatment, *Journal of Speech, Language, and Hearing Research,* 44, 889-904.

Gierut, J & Neumann, H. (1992) Teaching and Learning/T/: A nonconfound. *Clinical Linguistics and Phonetics*, 6, 191-200.

Gierut, J., Elbert, M., & Dinnsen, D., (1987). A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research*, 30, 462-479.

Gierut, J., Morrisette, M., & Champion, A. (1999). Lexical constraints in phonological acquisition. *Journal of Child Language*, 26, 261-294.

Gierut, J., Morrisette, M., & Ziemer, S. (2010) Nonwords and generalization in children with phonological disorders. *American Journal of Speech-Language Pathology*, 19, 167-177.

Gierut, J., Simmerman, C., & Neumann, H. (1994). Phonemic structures of delayed phonological systems. *Journal of Child Language*, 21(2), 291-316.

Girolametto, L., Steig-Pearce, P., & Weitzman, E. (1997). Effects of lexical intervention on the phonological of late talkers. *Journal of Speech, Language, and Hearing Research*, 40(2), 338-348.

Goldman, R. & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2.* Circle Pines, MN: American Guidance Service, Inc.

Hersko, W., Reid, D., & Hammill, D. (1999). *Test of Expressive Language Development (3rd ed.).* Austin, TX: Pro-Ed.

Kazdin, A. (1976). Statistical analyses for single-case experimental designs. In M. Hesen & D. H. Barlow (eds.), *Single-case experimental designs: Strategies for studying behavior change* (pp. 265-316). New York: Pergamon Press.

Kearns, K., (1986). Flexibility of single-subject experimental designs, Part II: Design selection and arrangement of experimental phases. *Journal of Speech and Hearing Disorders*, 51, 204-213.

Lleo, C. & Prinz, M. (1997). *Syllable structure parameters and the acquisition of affricates.* In *Focus on phonological acquisition,* M. Young-Scholten & S. J. Hannas, eds., John Benjamins, Philadelphia, pp. 143-164.

McReynolds, L. & Kearns, K. (1983). *Single-subject experimental designs in communicative disorders.* Baltimore, MD: University Park Press.

McReynolds, L. & Thompson, C. (1986). Flexibility of single-subject experimental designs, Part I: Review of the basics of single-subject designs. *Journal of Speech and Hearing Disorders,* 51, 194-203.

Morrisette, M., Gierut, J. (2002). Lexical organization and phonological change in treatment. *Journal of Speech and Hearing Research*, 17, 352-366.

O'Grady, W., Archibal, J., Aronoff, M., & Reese-Miller, J. (2005) *Contemporary Linguistics: An introduction – 9th edition.* New York: Avery Publishing Group.

Oller, D. K., & Delgado, H. K. (1999). *Logical International Phonetics Programs*. (Computer program: Windows Version). Miami: Intelligent Hearing Systems.

Olswang, L. B. (1990). *Treatment efficacy: The breadth of research.* In L. B. Olswang, C. K. Thompson, S. F. Warren & N. J. Minghetti (Eds.), Treatment efficacy research in communication disorders (pp. 99-103), Rockville, MD: ASHA Foundation.

Powell, T. & Elbert, M. (1984). *Generalization following the remediation of early- and later-developing consonant clusters.* Journal of Speech and Hearing Disorders, 49, 211-218.

Rescher, N. (1998) *Complexity: A philosophical overview.* New Brunswick: Transaction.

Robbins, J. & Klee, T. (1987). Clinical assessment of oropharyngeal motor developmental in young children. *Journal of Speech and Hearing Disorders,* 52, 271-277.

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical evaluation of language fundamentals – 4th edition.* Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Shriberg, L. (1992). *Shriberg's Order of Speech-Sound Acquisition.* New York:

Shriberg, L. & Austin, D. (1998). Comorbidity of speech-language disorder. Implications for a phenotype marker for speech delay. In R. Paul (ed.), *Exploring the Speech-Language Connection* pp. 73-117). Baltimore, MD: Brookes

Shriberg, Tomblin, & McSweeney. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research*, 42(6): 61-81.

Smit, A. B. (1993). Phonological error distributions in the Iowa-Nebraska Articulation Norms Project: Word-initial consonant clusters. *Journal of Speech and Hearing Research,* 36, 931-947.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990) The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, 55, 779-798.

Steriade, D. (1990) *Greek prosodies and the nature of syllabification*. (Doctoral dissertation, Massachusetts Institute of Technology, 1982). New York: Garland Press.

Storkel, H. (2004a). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisitions. *Applied Psycholinguistics*, 25, 201-221.

Storkel, H. & Morrisete, M. (2002) The lexicon and phonology: Interactions in language acquisitions. *Language, Speech, and Hearing Services in Schools*, 33, 24-37.

Templin, M.C., (1957). *Certain language skills in children: Their development and interrelationships.* (Institute of Child Welfare, Monograph Series 26). Minneapolis, MN: University of Minnesota Press.

Tyler, A. (2002). Language-based intervention for phonological disorders. *Seminars in Speech and Language*, 23, 69-81.

Tyler, A. & Lewis, K. E. (2005) Relationships among consistency/variability and other phonological measures over time. *Topics in Language Disorders*, 25(3), 243-253.

Tyler, A., Lewis, K. E., & Welch, C. M. (2003). Predictors of phonological change following intervention. *American Journal of Speech Language Pathology*, 12, 289-298.

Wellman, B., Case, I., Mengert, I., & Bradfury, D. (1931). *Speech sounds of young children.* University of Iowa Studies on Child Welfare, 5, 1-82.

Paul, R. (2001). *Language Disorders from Infancy through Adolescence: Assessment & Intervention.* St. Louis: Mosby.

Van Riper, C., (1978). *Speech Correction: Principles and methods* (6th edition). Englewood Cliffs, NJ: Prentice Hall.

Van Riper, C. and Irwin, J. V., (1958). *Voice and Articulation*. London: Pitman Medical (2nd Edition).

Williams, N., (2003). Phonological Memory and Rule Learning. *Language Learning*, 53, 1, 67-121.