5-2018

The Effect of Dose Frequency on Treatment Efficacy for Children with Speech Sound Disorders

Kristen Marie Giesbrecht

Follow this and additional works at: https://commons.und.edu/theses

Part of the Communication Commons

Recommended Citation

https://commons.und.edu/theses/923

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact zeineb.yousif@library.und.edu.
THE EFFECT OF DOSE FREQUENCY ON TREATMENT EFFICACY FOR CHILDREN WITH SPEECH SOUND DISORDERS

by

Kristen Marie Giesbrecht
Bachelor of Arts, University of North Dakota, 2016

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
2018
This thesis, submitted by Kristen Giesbrecht 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.

[Signatures]

Dr. Sarah Robinson
Dr. Alycia Cummings
Dr. John Madden

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

[Signature]
Dr. Grant McGimpsey
Dean of the School of Graduate Studies

April 25, 2019
Date
PERMISSION

Title The Effect of Dose Frequency on Treatment Efficacy for Children with Speech Sound Disorders

Department Communication Sciences and Disorders

Degree Master 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 absence, by the Chairperson of the department or the Dean of the School of Graduate Studies. 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.

Kristen Giesbrecht
April 18, 2018
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ vi

LIST OF TABLES .......................................................................................................... viii

ACKNOWLEDGMENTS ............................................................................................... ix

ABSTRACT ..................................................................................................................... x

CHAPTER

  I.  INTRODUCTION ................................................................................................. 1

  II. METHOD ............................................................................................................ 27

  III. RESULTS .......................................................................................................... 42

  IV. DISCUSSION .................................................................................................... 55

REFERENCES .............................................................................................................. 67
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Academic Vocabulary (AV) words used in treatment to target /ɹ/ and /l/ with their associated picture stimuli and incidental learning definitions</td>
<td>31</td>
</tr>
<tr>
<td>2.</td>
<td>Learning curves displaying the daily treatment data for study participants involved in speech treatment targeting /l/. Data points on the graph indicate production accuracies of the targeted treatment phoneme in the word-initial position of the child’s treatment words across all twenty treatment sessions. The curve of the child in the 2x/week condition is marked in gray, while the curve for the child in the 4x/week condition is black</td>
<td>43</td>
</tr>
<tr>
<td>3.</td>
<td>Learning curves displaying the daily treatment data for study participants involved in speech treatment targeting /ɹ/. Data points on the graph indicate production accuracies of the targeted treatment phoneme in the word-initial position of the child’s treatment words across all twenty treatment sessions. The curve of the child in the 2x/week condition is marked in gray, while the curve for the child in the 4x/week condition is black</td>
<td>44</td>
</tr>
<tr>
<td>4.</td>
<td>Individual children’s generalization of the treatment sound in untreated words. The accuracy values represent each children’s ability to produce their treatment sound in the speech probes administered pre- and post-treatment. A 10% pre-to-post-treatment change was the defined qualitative measure of treatment success. One child in the 2x/week condition (Child 2) and both children in the 4x/week condition (Child 3 and Child 4) met the 10% gain criterion</td>
<td>45</td>
</tr>
<tr>
<td>5.</td>
<td>Individual children’s percent of consonants correct (PCC) scores. The white bar in each column represents children’s ability to produce all consonants in speech probes administered prior to beginning treatment. The gray bar represents children’s production of all consonants in the same probes after treatment ended. The black bar represents the PCC change from pre- to post-treatment. A 10% pre-to-post-treatment change was the defined qualitative measure of treatment success. Only one child from the 4x/week condition met the 10% gain criterion</td>
<td>47</td>
</tr>
</tbody>
</table>
6. Individual children’s treatment sound effect sizes. A small effect size was observed for Child 1, a medium effect size for Child 2 and Child 4, and a large effect size for Child 3 ................................................................. 48

7. Individual children’s PCC effect sizes. All participants in both intensity conditions displayed a medium effect size ....................................................................................................... 49

8. Individual children’s learning rates of their treated sound in untreated words (grey bars) and percent consonants correct (PCC) scores (black bars). Learning rates are represented in terms of the number of trials needed for each child to make a 1% gain in production accuracy ............................. 52
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participant characteristics of the two children treated in the 2x/week condition and the two children treated in the 4x/week condition</td>
<td>28</td>
</tr>
<tr>
<td>2. Treatment target word properties (i.e., density, age of acquisition, word frequency)</td>
<td>30</td>
</tr>
<tr>
<td>3. Intervention intensity variables for the present study's treatment program</td>
<td>32</td>
</tr>
<tr>
<td>4. Treatment dose (i.e., number of trials) across all sessions for each participant</td>
<td>33</td>
</tr>
<tr>
<td>5. Treatment sound production accuracy and standard mean difference (d) computations. Data includes the treatment sound production accuracy scores for pre-treatment (M_A) and post-treatment assessments (M_B), and individual d values for both children in either treatment condition. The qualitative generalization or percentage accuracy change following treatment is determined by subtracting the M_A from the M_B (M_B–M_A). An average value of M_B–M_A and d for the 2x/week and 4x/week groups has been calculated to improve comparisons across conditions</td>
<td>45</td>
</tr>
<tr>
<td>6. Percent Consonants Correct (PCC) and standard mean difference (d) computations. Data includes the PCC accuracy scores for pre-treatment (M_A) and post-treatment assessments (M_B), and individual d values for both children in either treatment condition. The qualitative generalization or percentage accuracy change following treatment is determined by subtracting the M_A from the M_B (M_B–M_A). An average value of M_B–M_A and d for the 2x/week and 4x/week groups has been calculated to improve comparisons across conditions</td>
<td>46</td>
</tr>
<tr>
<td>7. Learning rates of treated sound accuracy change</td>
<td>50</td>
</tr>
<tr>
<td>8. Learning rates from PCC change</td>
<td>51</td>
</tr>
<tr>
<td>9. Pre- and post-treatment phonemic inventories</td>
<td>53</td>
</tr>
<tr>
<td>10. Pre- and post-treatment cluster inventories</td>
<td>54</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I wish to express my sincere appreciation to the members of my advisory committee for their guidance and support during my time in the Master’s program at the University of North Dakota.
ABSTRACT

Speech-language pathologists (SLPs) are urged to make evidence-based treatment decisions, but it is challenging to determine the appropriate intervention intensity for children with speech sound disorders (SSD) due to limited published information. This study is a single subject, multiple baseline design that compares the phonological changes of four preschool children (4;0 to 4;9) who received therapy either twice a week or four times a week for a total of twenty, 50-minute sessions. Each child’s production accuracy of treated sounds and overall percent consonants correct (PCC) values were used to qualitatively and quantitatively measure generalization via raw scores, d scores, and learning rate scores. Daily treatment data and phonemic/cluster inventories were also considered when comparing the children amongst different intensity conditions. The two children treated with a higher dose frequency (4x/week) demonstrated greater levels of phonological change due to treatment than the two children treated with a lower dose frequency (2x/week). This trend was observed across various measurement metrics, including PCC, treatment sound accuracy in untreated words, and changes to phonemic and cluster inventories. An advantage for treatment of a high dose frequency was observed even though the learning curves, which represented sound learning during productions of treatment target words, displayed similar production accuracies between conditions. Thus, traditional treatment provided on an intensive schedule was more efficacious than treatment on a less intensive schedule when considering the system-wide
phonological changes following treatment. In other words, an intensive therapy schedule elicited faster and greater phonological change. With a finding such as this and a growing interest for researchers to study the differences in intervention intensity variables, SLPs should begin to reference the current literature when making decisions regarding treatment scheduling for children with SSD.
CHAPTER I

INTRODUCTION

Speech sound disorder (SSD) is an umbrella term that is used to describe a person’s difficulty with perception, motor production, and/or the phonological representation of speech sounds and speech segments (ASHA, n.d.). The disorder typically impacts a person’s intelligibility, or ability to be understood by others. In some cases, the cause of the disorder can be identified, but in many cases, it remains unknown. Individuals diagnosed with functional SSD have no known cause for their communication breakdowns, as they present with normal hearing, intelligence, and social, emotional, and behavioral skills. In other words, their inability to articulate speech sounds is not caused by cognitive, sensory, motor, structural, or affective issues (Bernthal, Bankson, & Flipsen, 2012; Shriberg, 2003).

SSD are known to impact either the form or function of speech sounds within a language. Impaired speech production which affects the form of speech sounds is traditionally referred to as an articulation disorder. An articulation disorder reflects a child’s inability to articulate speech sounds, often involving a motoric component (Dinnsen, 1984; Hoffman, 1989; Stoel-Gammon, 1985). In addition, a SSD that affects the way in which speech sounds function within a language is often referred to as a phonological disorder. This term describes an impairment in the way speech sound information is stored and represented in the mental lexicon or is accessed and retrieved
cognitively (Bernthal et al., 2012). Distinguishing between articulation and phonological disorders can often be difficult because speech-language pathologists (SLPs) cannot always identify the underlying cause of a child’s speech impairment; thus, the term SSD has been coined to describe a range of speech sound production impairments in children. To improve the clarity of referenced literature in this paper, the label SSD will consistently be used when referring to children identified as having articulation disorders and/or phonological impairments.

Determining the prevalence of SSD has also been difficult due to the inconsistent definitions of the disorder and the limitations in generalizing statistical information from specific samples to the population as a whole. In general, it should be emphasized that diagnoses of SSD are most common in preschool children. Campbell and colleagues (2003) presented data to suggest that 15.2% of 3-year-olds have speech sound disorders. Another study of SSD in preschool children identified a lower prevalence estimate of only 3.4% at 4 years of age (Eadie et al., 2015). By the time children are 6-years-old, however, the prevalence of SSD has been recorded as approximately 3.8% (Shriberg, Tomblin, & McSweeny, 1999). Prevalence estimates tend to vary considerably depending on the literature source that is cited. A systematic review reported prevalence estimates of SSD in children ages 5 to 7 years ranging from 2% to 25% (Law, Boyle, Harris, Harkness, & Nye, 2000). Though variation exists among prevalence estimates at particular ages, there is an agreement on an overall trend of decreasing prevalence as a child ages. This trend highlights the likelihood of treating and resolving a SSD in the preschool years.

Signs of SSD can be identified early in a child’s development of speech. Children
with SSD are often diagnosed and involved in a treatment program before they reach school age. There are several evidence-based intervention approaches that have been found to be effective in improving the speech sound production in children with SSD (Baker & McLeod, 2011a; Gierut, 1998; Kamhi, 2006; Law, Garrett, & Nye, 2004; Nelson, Nygren, Walker, & Panoscha, 2006; Williams, McLeod, McCauley, Warren, & Fey, 2010). Selection of a particular approach often requires consideration of a number of factors, including the child’s age, errors, severity, and overall intelligibility (Williams et al., 2010). Some of the well-known, evidence-based approaches include traditional intervention (e.g., van Riper & Emerick, 1984), cycles (e.g., Hodson & Paden, 1991), minimal pairs (e.g., Gierut, 1992; Weiner, 1981), multiple oppositions (Williams, 2000), Metaphon (Dean, Howell, Waters, & Reid, 1995), and core vocabulary (Dodd, Holm, Crosbie, & McIntosh, 2006).

One of the most prominent differences among intervention approaches is in regard to the number of sounds targeted during an intervention period. Contrast therapy, such as minimal pairs and multiple oppositions, utilizes multiple word pairs that differ by single sounds in order to elicit specific sound productions. Within this approach, children may have up to four different sound targets in which they practice each session. The traditional approach, however, targets a single sound throughout an intervention period by shaping its correct production and improving its accuracy through a linguistic complexity hierarchy that begins in isolation, then moves to syllables, then to words, then to phrases, and finally to sentences. It is also important to note that a child with a single sound target would have many more opportunities to produce their target sound than a child with 3-4 targets; thus, providing speech treatment via a traditional approach could
result in deeper training of a particular sound in comparison to treating sounds through a contrast therapy approach. Moreover, the traditional approach has been commonly cited in the literature on SSD treatment (Cummings & Barlow, 2011; Gierut & Morrisette, 2010; Gierut, Morrisette, Hughes, & Rowland, 1996; Morrisette & Gierut, 2002) and provides a practical and concrete method for tracking productions and daily progress.

**Service delivery decisions**

In addition to determining the type of speech treatment that is optimal for children with SSD, speech-language pathologists (SLPs) must also consider service delivery variables that may have an impact on treatment outcomes, such as the intervention intensity variables. Intervention intensity has become an important issue in the field of speech-language pathology. As such, researchers have started to become more thorough in their descriptions of treatment methods. The term “intervention intensity” has been described as a way of defining both the quality and quantity of learning experiences within and across sessions. It is more complex than merely quantifying the occurrence of treatment sessions over time. Instead, intervention intensity involves numerous components so that the exact number of teaching opportunities or episodes during a treatment program can be calculated.

In order to provide the most efficient and effective speech-language services, SLPs must understand how to administer the ideal amount of intervention, including how often to schedule therapy sessions and for how long. Treatment scheduling is often influenced by factors such as the client’s age, attention span, and severity of the disorder (Bernthal et al., 2012). Unfortunately, there are often other external factors contributing to decisions of intensity. Research has shown that SLPs with heavier caseloads offer
significantly less frequent and shorter treatment sessions to their clients with SSD (Brandel & Loeb, 2011; To, Law, & Cheung, 2012). Interestingly, a survey on program delivery revealed that students diagnosed with articulation disorders received the least intensive program of 1x/week for 20-30 minutes more often than students with any other disorder (e.g., pervasive developmental disorder, pragmatic disorder, specific language impairment, developmental disability) (Brandel & Loeb, 2011). Though the same trend was not identified across disorder types for children diagnosed with moderate and severe disabilities, it is interesting to consider that a diagnosis of mild SSD may actually reduce the intensity of provided treatment when making decisions regarding program scheduling. It has also been found that SLPs with larger caseloads provide more group interventions than SLPs with smaller caseloads (Dowden et al., 2006). This likely reduces the number of teaching opportunities a child receives as compared to if he or she was seen for the same amount of time in a one-on-one therapy session. These findings display the ongoing challenge that SLPs face when scheduling their clients amongst a busy schedule. In some cases, it may be true that the clinician only has the availability to provide services to a child once a week even though their clinical judgement would recommend sessions twice a week. In addition, clinicians may also be faced with challenges regarding the number of clients on waiting lists needing services, the number of SLPs available to provide instructional services, government policies regarding access and provision of speech-language pathology services, and the limit that insurance companies place on the provision of speech-language pathology services (Baker, 2012).

Similarly, parental expectations and/or resources may also play a role in determining the intensity variable. Some children may be unable to attend an ideal
frequency of intervention due to financial resources, distance from SLP services, lack of support from family, or misguided knowledge of the condition or intensity required to treat the condition (Baker, 2012; Bernthal et al., 2012). In some cases, it may also be the parent that “burns out” from frequent attendance at therapy sessions; thus, the overall child in the context of their family must be considered when determining the intensity of a treatment program.

In an ideal world, the important decisions surrounding intensity would be exclusively driven by evidence in the literature. Unfortunately, the current challenge in achieving this standard rests deeper than the external client- and clinician-related variables. In order to make these evidence-based decisions, we would first need a sound basis as to what the optimal treatment intensity would be for a particular intervention with SSD. There has not been much research to determine the intensity variables by which treatment programs should be administered, though. This means that there is little to no evidence for determining the optimal intervention intensity for children with SSD. As a result, SLPs consistently have a difficult time recommending the optimal treatment intensity following the child’s initial qualification for services.

As mentioned above, intervention intensity is an important way in which the outcomes of speech treatment should be examined and explained. Within a professional field, it is imperative that interventions are successful and worthwhile. It is possible that interventions may be administered in too high of a dose or offered too frequently without added benefit (McGinty, Breit-Smith, Fan, Justice, & Kaderavek, 2011). It is also possible that interventions are too infrequent and not as effective as they could be, or worse yet, equivalent to no intervention at all (Glogowska, Roulstone, Enderby, & Peters,
In order to improve the evidence base surrounding optimal intervention intensity for SSD, research needs to focus on defining the intensity variables, clearly reporting them in studies, and designing studies that make explicit comparisons between intensities.

**Intervention intensity variables**

Warren, Fey, & Yoder (2007) initiated the movement of improving clarity and completeness of reporting intensity variables in the literature by providing operational definitions for several components including: dose form, dose, session duration, dose frequency, total intervention duration, and cumulative intervention intensity.

*Dose form* refers to the context of activities and interactions within any given therapy session. It describes the type of task or activity in which teaching episodes are delivered. This variable exists on a continuum between a child-centered, play-based approach at one end and a clinician-directed, drill-based approach on the other end.

Child-centered, naturalistic play includes activities such as playing with dolls or cars while teaching through recasts or modeling of correct productions. A clinician-directed approach typically involves picture naming while incorporating games or crafts as a reinforcing activity. Typically, treatment of SSD for preschool children follows a drill-play approach in order to engage the child while providing the most opportunities for practice of their treatment target sound. In this approach, play is used as the motivating event prior to a child’s treatment trial as well as a consequent reinforcer following a response. Multiple treatment trials can be elicited using this approach because practice becomes less taxing when it is combined with exciting events, such as taking turns in a game or making progress on a craft.
**Dose** refers to the number of properly administered teaching episodes per session (e.g., 100 trials per session). When dose has been reported in the literature involving speech treatment, it is often referred to as the average number of trials a child produced in a session. In a current review of the literature, only 30 of 146 studies (i.e., 21%) provided quantitative information regarding dose (Baker & Williams, 2011). Of these studies, dose ranged between 14 and 240 trials per session, but was most commonly reported as 100 trials per session. Though it is the target dose that is often indicated, it is important to note that if time runs out before all treatment trials are elicited, the remaining number of trials are often abandoned. This means that simple multiplication using the target dose does not necessarily accurately represent the total number of trials in a treatment program since the number of trials each session is not always consistent due to the limitations of time. Interestingly, some particular speech interventions have actually determined a recommended dose for each treatment session. For example, multiple oppositions intervention suggests a minimum of 60 responses during focused practice and 20 responses during naturalistic activities within a 30-minute individual session (Williams, 2012). Other intervention approaches do not necessarily list the number of trials required for each session but it is expected that the clinician provide as many practice opportunities as possible.

**Session duration** identifies the length of each session (e.g., 30-minute sessions). Previous research of treatment for children with SSD has defined this variable within the range of 15 and 270 minutes. The most common session durations identified in the research conducted in clinical settings included 30, 45, and 60 minutes (Baker & Williams, 2011). In a survey of program intensity in the schools, students of all ages most
commonly received intervention with a session length between 20-30 minutes (Brandel & Loeb, 2011). The difference of session duration across environments is likely due to the SLP’s inability to pull children out of class for lengthy periods of time.

**Dose frequency** indicates the number of sessions provided per time unit (e.g., twice a week). Surveys of school-based SLPs have indicated that the majority of children with mild to moderate speech or language impairment received two sessions per week (Mullen & Schooling, 2010), while children with SSD in research settings were seen two or three times per week (Baker & McLeod, 2011a). In a review of the current literature, dose frequency varied from 1x/month to as frequent as 5x/week (Baker & Williams, 2011).

**Total intervention duration** is the interval time for which an intervention is provided (e.g., 20 weeks). In other words, this variable describes the total period of time a child receives intervention. In a review of 134 intervention studies, only 10 studies reported this variable and it ranged from 3-46 months (Baker & McLeod, 2011a). In school settings, children who are provided speech and/or language services through an Individualized Education Program (IEP) are often expected to receive services through the entire school year. This means that a total intervention duration of 9 months is explicitly defined and legally mandated by the IEP. In such a setting, the clinical decision to alternate intensities at different times of the year becomes difficult, if not impossible, because the treatment program is prescribed for 9 months. In addition, determination of total intervention duration is complicated in other therapeutic settings when a block
schedule\(^1\) is determined rather than a continuous schedule\(^2\). For example, if therapy is provided over several months but there are scheduled breaks between intervening blocks, the total intervention duration may need to be adjusted so that it includes only the number of weeks that the child receives direct intervention. In other cases, it becomes difficult for clinicians to explicitly define the intervention duration for treatment programs with an open duration (i.e., a specific performance criterion that must be met before treatment is completed) rather than a time-based criterion.

**Cumulative intervention intensity** is the product of “dose \(\times\) dose frequency \(\times\) total intervention duration” which results in a numerical measure of intensity (e.g., 100 trials \(\times\) 3x/week \(\times\) 10 weeks = 3000 trials). In a review of the current literature, only 11 of 146 studies (i.e., 7.5%) provided sufficient data to calculate the cumulative intervention intensity value (Baker & Williams, 2011). Cumulative intervention intensity is arguably the most important variable to describe for an intervention program as it defines the total number of teaching episodes during a child’s treatment. In some cases, while many of the intensity variables appear consistent for clients receiving SSD treatment, the total number of productions may be substantially different.

For example, consider the case of two different children receiving therapy twice a week for 30-minute sessions for a total duration of 20 weeks. If, by chance, child 1 is able

\(^1\) A block schedule is when treatment is provided for a specified period of time and then followed by a period of time devoted to indirect services or a break from treatment completely. For example, a child may receive 10 weeks of direct intervention followed by a 10-week break from services before starting up with direct intervention again.

\(^2\) A continuous schedule means that the child is provided treatment on a consistent schedule, such as twice a week, without having any scheduled breaks from treatment.
to produce an average of 100 productions each session while child 2 is less cooperative and only able to complete 50 target productions each session, the cumulative intervention intensity would differ dramatically between the two children. Though their intervention schedule looks identical in the initial description, ignoring the information of dose in each session makes it impossible to calculate their cumulative intervention intensity. If learning is based solely on the number of opportunities for sound productions during treatment, SLPs may need to redirect their focus to the total trial attempts rather than the number of minutes each child attends therapy per week.

Following a review of the literature, it was noted that several intervention intensity variables have been reported in the literature, but they have not often been explicitly compared as a treatment variable (Baker & McLeod, 2011a). In other words, the relationship between the intervention intensity variables and treatment outcomes remains unclear. Studies in which explicitly compare the difference of intervention intensity variables are necessary in order to make evidence-based decisions regarding treatment scheduling.

**Intensity as a variable for learning**

The investigation of treatment intensity is a current and relevant topic when studying how individuals learn new skills. Comparisons of different intensities have been documented for the learning of non-speech motor skills, motor-based speech skills, and language skills.

**Intensity differences for learning non-speech motor skills.** Research involving the study of motor learning not specific to speech production has explored the implications of various practice conditions on a person’s learning (Schmidt & Lee, 2011).
Several principles of motor learning have been described following research which aimed to understand how the motor system learns (Maas et al., 2008). One of these variables, named practice distribution, is comparable to intervention intensity in that it describes how often a person has opportunities for practice. Practice distribution is defined by how a given (fixed) amount of practice is distributed over time. That is, practice distribution can be identified as either massed or distributed practice. Massed practice is when a given number of trials or sessions is administered in a short period of time, while distributed practice indicates the practice of a given number of trials or sessions over a longer period of time.

There is evidence from the study of motor learning in non-speech tasks (e.g., keyboard entry task) to suggest that distributed practice (i.e., more time between practice trials or sessions) results in greater learning than massed practice (i.e., less time between trials or sessions) (Baddeley & Longman, 1978; Maas et al., 2008). The studies have shown that distributing practice over a longer period of time facilitates both immediate performance and retention for various motor tasks. This pattern may be explained through the consideration that distributed practice increases opportunity for memory-consolidation processes (Robertson, Pascual-Leone, & Miall, 2004). Evidence has also revealed that too much practice of a single skill (i.e., massed practice) leads to context-dependent learning (Fischman & Lim, 1991). This means a person has difficulty transferring the motor skill being learned to different contexts. In general, massed practice facilitates faster acquisition of a motor skill during practice opportunities but runs the risk of not actually establishing motor learning of the skill (Caruso & Strand, 1999). In other words, massed practice results in quick development of motor skills in the
environment by which it is learned with poor generalization to other contexts. Distributed practice, on the other hand, takes longer to acquire the motor skill but achieves better motor learning because long-term retention and transfer of skills to other contexts are more likely (Caruso & Strand, 1999). Thus, when teaching non-speech, motor-based skills, it is preferred that a distributed schedule of variable practice be implemented.

**Intensity differences for learning motor-based speech skills.** The performance difference between distributed practice and massed practice for those receiving instruction of motor-based speech tasks is still unclear. There have been several studies aimed at understanding this trend for motor-based speech disorders and the resulting evidence is not straightforward. Spielman and colleagues (2007) examined the performance of individuals with dysarthria secondary to Parkinson’s disease who received an extended 8-week Lee Silverman Voice Treatment (LSVT) program. The performance of these participants was then compared with the results of a previous study involving participants in the typical 4-week program (Ramig, Sapir, Fox, & Countryman, 2001). The total number and duration of treatment sessions remained the same for the two programs. Observations following administration of the extended treatment program were comparable to the performance of individuals in the shorter program, even after a 6-month retention test (Ramig et al., 2001). In other words, distributed practice was a viable option for individuals with dysarthria, but it did not appear to enhance learning relative to massed practice. Wohlert (2004) also compared three treatment intensities of the LSVT program for individuals with hypokinetic dysarthria and failed to reveal significant differences between group conditions. Contrary to non-speech findings which suggest improved performance on a distributed schedule, these two studies suggest there may not
actually be any differences in the re-learning of speech and/or voice behaviors for people with dysarthria receiving intervention on different intensity schedules.

Interestingly, some of the research regarding treatment of motor-speech disorders has now gone so far to say the opposite of what is known about the learning of non-speech motor tasks. That is, studies have shown that massed practice is actually better when learning motor-based speech skills. For example, children diagnosed with childhood apraxia of speech (CAS) who were treated with the motor speech treatment protocol (MSTP) showed greater gains in articulation performance on a standardized test and better functional outcomes (e.g., improved confidence, social skills/friendships, use of repair strategies) when treated twice-a-week rather than once-a-week (Namasivayam et al., 2015). Unfortunately, there was no follow-up measure to determine if these advanced skills were maintained over a period of time. In another study, speech therapy for children with cleft palate and other velopharyngeal disorders found that children who received an intensive treatment course of three treatment sessions a day for 6 weeks performed significantly better on articulation measures at various times of measurement than a control group who received treatment once-a-week (Albery & Enderby, 1984). Notably, children who received the six-week intensive course maintained their skills and continued to show an advantage over the conventional therapy group even two years after the course had ended.

Intensity differences for learning language. Intensity effects on various language treatment outcomes have also been examined. Some studies have found that more intensive treatment or teaching sessions are more efficient treatment models for patients with global aphasia (Denes, Perazzolo, Piani, & Piccione, 1996), for young
communicators with goals to improve expressive language (Barratt, Littlejohns, & Thompson, 1992), and for early school-age children receiving literacy instruction (National Reading Panel, 2000). Alternatively, it has been shown that semantic and morphological learning in preschool learners does not actually change when provided more intensive treatment compared to less intensive instruction in terms of dose frequency, or number of sessions per week (Bellon-Harn, 2012). In other words, no differences have been found between concentrated and distributed schedules for semantic and morphological learning, suggesting that both of these scheduling frameworks are equally successful. Yet still, research concerning word learning in children with specific language impairment (SLI) has described a trend in which performance improves as intensity increases until a particular intensity is reached and the amount of learning plateaus (Storkel et al., 2017). Through consideration of the research concerning intensity differences throughout the field of speech-language pathology, a statement may be made that optimal treatment intensity varies and is likely specific to factors such as the presenting disorder, the skills being taught, and the intervention approach that is applied.

**Intensity differences in SSD treatment**

Though much evidence has already been made available regarding intensity differences for individuals with other communication disorders, there is currently little systematic evidence to support one intensity method over the other in the treatment of children with SSD (Maas et al., 2008). While SLPs are continually urged to make evidence-based treatment decisions, it is challenging to determine the appropriate intervention intensity for children with SSD because of the limited information reported in the extant literature (Baker & McLeod, 2011a; Cirrin et al., 2010). In addition, the few
studies that do address components of intervention intensity report values which are successful in producing favorable treatment changes but are inconsistent across the available studies.

For example, Bowen and Cupples (1999) noted improvements to a child’s phonological processes using Parents and Children Together (PACT) therapy\(^3\) after an intervention period involving approximately 17.5 hours over an average of 10.6 months. By contrast, Klein (1996) reported positive treatment change to a child’s speech severity score when imagery therapy\(^4\) was used for an average of 68.23 hours over 13 months.

Though it may be important to consider the additional influence that home programming and parental involvement has for those treated by the PACT program, there are still substantial differences noted between the number of hours of direct intervention required

\(^3\) PACT therapy is a broad-based phonological treatment approach which is family-centered and involves active participation of the child’s parents. Treatment is often provided on a block schedule so that the therapist provides direct intervention for a number of weeks before this role is completely distributed to the parents who are expected to complete formal practice with their child. The major components of PACT therapy include parent education, metalinguistic skills, phonetic production training, multiple exemplar training, and homework.

\(^4\) Imagery therapy involves learning new phonological rules by assigning labels and images to specific phonetic characteristics. For example, a child with an active phonological process of “stopping” may learn to classify all stop sounds as “poppies” and all fricatives as “windies.” Feedback is then provided to the child by referencing these sound class labels. That is, the SLP may correct the child’s error of “tea/sea” by saying “Tea? You said it was a windy word, but you made it with a poppy sound. Can you try it again and put in the windy sound that you said it should have?” An important component of imagery therapy is that direct models of the target words are not generally provided; instead, sound errors are addressed by indicating the sound class label that should have been produced.
to make significant speech gains. The vast differences in number of hours for these two intervention protocols showcase the variability that currently exists in the literature concerning successful treatment intensity schedules. When interventions of both high and low frequencies are effective, it is difficult to determine which method is most beneficial. The considerable variation in dose, session duration, and dose frequency that is present across studies makes it challenging to speculate what the ideal speech intervention intensity should be (Warren et al., 2007). Even though intervention intensity variables have been reported in the SSD intervention literature, what is needed is information regarding whether or not more intensive intervention produces better outcomes than less intensive intervention.

A systematic review of 10 studies investigated the effect of frequency, intensity, and/or duration on the speech and language skills of preschool children (Schooling, Venediktov, & Leech, 2010). Following calculations of weighted effect sizes (Beeson & Robey, 2006; Busk & Serlin, 1992), only seven of the 35 measured outcomes displayed a significant effect in favor of one intensity condition over the other. Of the seven clinically significant outcomes, six of the outcomes were more frequent, intensive, or lengthy than the variable to which they were being compared. For example, an intervention schedule of four 20- to 30-minute sessions/week was more favorable when compared to one 60-minute session/week when the outcome measure was the number of child responses to requests in a language sample.

In general, the review completed by Schooling and colleagues (2010) substantiates the idea that more intense intervention brings about better outcomes; however, these findings were not necessarily compelling because most of the calculated
measures (i.e., 28/35) were not significant. Additionally, the review did not allow for an explicit comparison of specific intensity variables while keeping all other variables constant. For example, the comparison of treatment outcomes for children receiving treatment once daily versus twice daily after six weeks (Whitehurst & Lonigan, 1998) did not answer the question of optimal dose frequency, or number of sessions per week, because the children receiving intervention twice a day ultimately received more intervention by the end of their treatment programs. In order to truly investigate the effects of intervention intensity, studies must be designed to explicitly compare a single intensity variable while keeping all other variables constant.

**Intervention intensity as the independent variable.** *School-aged children.* One of the first studies to make an explicit comparison between intensity schedules was completed in the Chicago Public Schools during the 1953-1954 school year (Fein, Golman, Kone, & McClintock, 1956). The public-school system, at that time, was struggling to determine how to best utilize their staff of 70 speech therapists among all the schools within their city. It became general practice for speech therapists to travel between several schools, usually making only one visit a week to a particular school. This demanding schedule raised the question of whether a therapist would achieve better results with their elementary-aged students if they visited the school twice-a-week for one semester and provided no therapy during the other semester, than if they continued with the present schedule of conducting therapy sessions once-a-week for both semesters of the school year. This question developed into a research design that explicitly compared dose frequency, or number of sessions per week, for intervention involving children with SSD. One condition involved children receiving therapy twice-a-week for only one
semester, while the other condition had children participating in therapy once-a-week for two semesters. That is, the total therapy contact hours was ultimately the same between the two conditions since total intervention duration was appropriately adjusted.

Following comparisons of articulation errors on tests administered at the end of the year (i.e., June) compared to their performance at the beginning of the year (i.e., September), there did not appear to be any important differences in progress between the 299 pairs of children involved in the experiment. It appeared that therapy provided twice-a-week for one semester had no obvious advantages over therapy provided only once-a-week for the entire school year. Notably, the authors reported that children who received twice-a-week therapy during the first semester did not lose the skills learned during the second semester when no speech intervention was provided. All findings were determined by examining the difference in number of errors on articulation tests that followed their specific treatment program. Though conclusions were drawn from quantitative data, there were no statistical analyses completed to determine if there was actually a significant difference between the two conditions. With that said, this early observation was one of the first findings to support flexibility in the decisions surrounding planning of a SLP’s therapy schedule.

In addition to the conclusions drawn from speech outcomes, the Chicago Public Schools study also explored the preferences of the forty participating clinicians. More than half of the SLPs agreed that the twice-a-week system improved student’s response to homework assignments, teacher’s awareness of the speech program, rapport within therapist-student relationships, and ease of parent contact arrangements (Fein et al., 1956). This information revealed an overall preference for the more intensive therapy
option (twice-a-week versus once-a-week) among service providers. It also justified the decision of selecting intensity schedules based on the interest and convenience of the therapist and/or corresponding school if, in fact, the two models did not differentially affect the progress of the child.

Shortly after the Chicago Public Schools study, van Hattum (1959) explored the possible differences of intensity schedules following a shift of the service delivery model in the elementary schools in Rochester, New York. Van Hattum (1959) suggested that a greater intensity of treatment may in fact elicit better treatment outcomes. This statement was made with reference to the change in service delivery models from a regular schedule to a block schedule in 1955. Instead of providing therapy regularly at a frequency of once or twice a week for the entire school year, the SLPs chose to provide block therapy to students every day of the week but for only a single trimester. Following this school district-wide change in 1955, researchers noticed an increase in the dismissal rates for children receiving speech therapy under the new model. Though several limitations exist in this non-experimental design, van Hattum’s publication (1959) was one of the first to report that speech treatment administered in a block schedule may be a more efficient way to achieving articulatory and/or phonological progress than was intermittent scheduling (van Hattum, 1959, 1969).

Several years later, a study was completed in the Crawford County School System in Ohio to explicitly examine the effects of different intensity schedules of speech therapy on articulatory progress for students in Grades 2 through 8 (Ausenheimer & Irwin, 1971). All students received 30-minute sessions at various dose frequencies, or number of sessions per week. Three group conditions involved students scheduled for
therapy twice a week, three times a week, or four times a week. This meant that total contact hours varied from 60 minutes to 120 minutes in any given week and the total number of therapy hours during the school year varied considerably. Similar to the study in the Chicago schools, no specific intervention approach was implemented across all participants since each individual student’s needs were considered when planning therapy. After analysis of students’ performance on articulation tests at eight-week intervals, no intensity schedule appeared superior to another. Even when speech scores were compared between frequency condition groups after being provided the same number of sessions and thus, having an equivalent amount of contact hours, it appeared that articulatory proficiency was consistent. This experimental result suggested that any of the three scheduling methods (i.e., twice a week, three times a week, or four times a week) could be effectively implemented.

Weston & Harber (1975) also compared the differences in effectiveness of five different intensity schedules for speech therapy in elementary-aged children (i.e., grades 1-6) when following a specific intervention method called the paired-stimuli procedure\(^5\). The seventy participants were divided into five groups: Group I attended therapy on Monday, Wednesday, and Friday (3x/week); Group II on Tuesday and Thursday

---

\(^5\) The paired-stimuli intervention approach targeted a single phoneme by practicing its production in word pairs; one of the words included the phoneme in a context that was able to be produced with a high accuracy (called the “key word”) whereas the other word in the pair contained the same target phoneme in a context that it was misarticulated in at least two out of three attempts (called the “target word”). For example, a child with misarticulations of the phoneme /k/ may have a key word of “cup” and a target word of “car.” Practice of this word pair is intended to elicit correct productions of /k/ in their target word (i.e., “car”) by comparing it to the correct articulation in their key word (i.e., “cup”).
(2x/week); Group III on Monday and Wednesday (2x/week); Group IV on Monday, Tuesday, Thursday, and Friday (4x/week); and Group V attended on Monday, Tuesday, and Friday (3x/week). Progress was measured by the length of time it took for each child to meet satisfactory production accuracy of 80% or more on their particular treated phoneme during elicitation of two 15-minute probes. Results from the study indicated that children who received therapy only 2x/week (Groups II and III) required less time to reach production criterion than children scheduled in the three other intensity conditions. This significant difference indicates that the least intensive schedule was the most effective because children learned more rapidly when treated in this intensity condition. Faster learning on a distributed practice schedule may be due in part to the additional time between therapy sessions to spontaneously implement learned behaviors in the natural context of everyday conversations.

*Preschool-aged children.* Page, Pertile, Torresi, and Hudson (1994) continued to explore which treatment intensity was more advantageous by studying the effect of dose frequency in a treatment program targeting phonological processes for phonologically-delayed preschool children. SLPs provided group therapy to children at various dose frequencies, or number of sessions per week. Half of the participants attended therapy sessions 1x/week while the other half attended therapy 3x/week. The total intervention duration was adjusted so that all participants received a total of six 1-hr treatment sessions (i.e., 6 total hours of intervention). This meant that treatment blocks of either 2 weeks or 6 weeks were applied in order to keep the total number of treatment hours consistent across conditions. All participants received pre- and post-probes consisting of single-words that were designed to identify the occurrence of three phonological
processes: final consonant deletion, cluster reduction, and velar fronting. The child’s responses were simply scored as correct or incorrect in regard to the process expectations and a percent correct score was calculated and compared for pre- and post-assessments. In other words, a child’s score was not affected by incorrect sound productions as long as the process was absent (i.e., substituting /t/ for /s/ in “house” would be marked correct if the purpose of the target word was to eliminate final consonant deletion). Comparable gains in reduction of phonological processes were observed regardless of the intensity of intervention they received (Page et al., 1994). Thus, both intensities of treatment were found to be effective forms of service delivery since neither was superior in reducing phonological processes. It may be important to note that several other possible speech production accuracy measures, such as daily treatment data, Percent Consonants Correct (PCC) scores, and intelligibility, were not considered in this study. Consequently, it is possible that differences between the two intensity conditions were actually overlooked due to the nature of this research design. It is also possible that the total intervention time of six hours was not enough to elicit substantial change to the children’s phonological systems; thus, a difference between groups was not identified due to the minimal amount of treatment provided.

Recently, Allen (2013) compared dose frequency in a multiple oppositions intervention protocol provided to preschool-aged children with SSD. As described earlier, this approach is based on the contrastive model of speech therapy and targets multiple phonemes during the same treatment session. The multiple oppositions approach differs from traditional treatment in terms of number of target sounds and method for teaching those sounds. Traditional treatment focuses on the shaping of a single phoneme, while
multiple oppositions selects up to four phoneme targets from a child’s phoneme collapse and teaches these sounds by contrasting the semantic differences when the different sounds are produced. Allen (2013) measured treatment change via the differences in Percent Consonants Correct (PCC) scores pre- and post-treatment, which was calculated using their productions on the Goldman-Fristoe Test of Articulation – 2nd Edition (GFTA-2) Sounds-in-Words subtest. Allen (2013) reported that children participating in therapy 3x/week had significantly greater improvements to their PCC scores than children seen only 1x/week after a total of 24 30-minute sessions were completed.

Such a finding contrasted with earlier research on school-aged children which suggested a lower intensity of treatment was more advantageous for sound learning (Weston & Harber, 1975). The outcomes of Allen’s study (2013) also contrasted with findings of studies on school-aged children which found no difference between intensity conditions (Ausenheimer & Irwin, 1971; Fein et al., 1956). Thus, the conflicting results between intensity studies may mean that the preferred dose frequency differs depending on the age of a child. As a result, Allen’s study (2013) was an important first step in determining what optimal intervention intensity is for preschool children with SSD.

Though Allen (2013) identified that greater intensity of treatment was the ideal condition for preschool-aged children with SSD, it is still an important and current question to investigate across the various intervention approaches. Due to the novelty of intensity research for preschool-aged children with SSD, Allen’s (2013) findings can only be associated with the multiple oppositions treatment approach at this time. Therefore, treatment intensity must be examined and explored for several other intervention approaches before determining a generalized treatment recommendation.
It is imperative that SLPs continue to recognize the importance of evidence-based decisions surrounding treatment scheduling. As professionals, SLPs must avoid inaccurate implementation of treatment intensity, be it too high or too low, as that could lead to ineffective treatment outcomes (Baker, 2012). Though Allen’s findings (2013) suggest intensive therapy, it is critical to recognize that excessive amounts of treatment may not always result in additional gains. Clinicians would be wasting their time if they chose to increase frequency of sessions for all children on their caseload when the same results could be achieved with a less intensive treatment schedule. In addition, there is not enough supporting evidence in the current literature to clearly state that an intensive schedule is superior to a less intensive schedule for all SSD interventions with preschool-aged children. The major barrier in this line of research is that intervention often requires highly individualized methods for each client and for each speech disorder (Enderby, 2012). This makes it difficult to definitively identify the intensity schedule or the amount of therapy that yields the most effective results. If optimal treatment intensity is specific to intervention approaches (Kaipa & Peterson, 2016), then it is critical for continued research to examine treatment intensities across all of the evidence-based approaches for SSD treatment in children.

**Purpose of the present study**

The purpose of the present study is to evaluate how treatment dose frequency impacts effective and efficient changes in children’s sound systems when completing traditional speech treatment. Specifically, it questions whether traditional speech treatment involving complex sounds is more effective for preschool children when it is delivered four times a week as compared to twice a week, while holding the total number
of clinical contact hours constant. The decision to hold total hours of intervention constant is an attempt to control for overall cumulative intervention intensity. Addressing the question of optimal dose frequency will add to the emerging literature on intervention intensity that enables SLPs to make informed, evidence-based decisions regarding treatment scheduling.

If the results align with previous research involving multiple oppositions treatment for preschool children (Allen, 2013), it is hypothesized that an advantage in sound learning for the children receiving the intensive 4x/week schedule would be observed. This result would suggest there is a common trend of higher intensity programming being advantageous over lower intensity programming for preschool-aged children with SSD receiving speech interventions of any type. If, however, the 2x/week schedule leads to better speech outcomes, it could be that intervention intensity is influenced by the intervention approach. Such a result would indicate that each SSD treatment approach would require its own optimal intensity schedule.
CHAPTER II

METHOD

Participants

Four male children (aged 4;0 to 4;9 years) diagnosed with SSD participated in this study. All children involved in the study were referred for therapy within the Northern Health Region in The Pas, Manitoba. The children were assigned to treatment programs of two different frequency conditions: 2x/week for 10 weeks (Children 1 & 2) or 4x/week for 5 weeks (Children 3 & 4). In other words, all participants received a treatment program involving 20 hours of intervention with the intervention intensity variable of dose frequency varying across conditions.

Formal speech and language tests were completed to determine each child’s study eligibility. Each child met all of the following criteria: (a) no prior speech or language intervention; (b) residence in a monolingual English-speaking household; (c) typical speech structures and functions as measured by an oral-peripheral mechanism exam (OPME); (d) percentile score at or below 8 on the Goldman-Fristoe Test of Articulation, 3rd edition (GFTA-3; Goldman & Fristoe, 2015); (e) normal hearing as measured by a hearing screening at 20 dB HL for each ear at 1000, 2000, and 4000 Hz; (f) expressive and receptive language skills within normal limits as assessed by the Clinical Evaluation of Language Fundamentals - Preschool - Second Edition (CELF-P2; Semel, Wiig, & Secord, 2004). Table 1 provides specific details of all participants and the conditions to
which they belong. Notice that variations in age are controlled for when participants are matched with the child in the corresponding intensity condition receiving treatment for the same treatment sound (i.e., children with the treatment sound /ɹ/ are slightly older than the children with the treatment sound /l/).

Table 1. Participant characteristics of the two children treated in the 2x/week condition and the two children treated in the 4x/week condition.

<table>
<thead>
<tr>
<th>Participant Inclusionary Criterion</th>
<th>2x/week Group</th>
<th>4x/week Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Child 1</td>
<td>Child 2</td>
</tr>
<tr>
<td>Gender</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Age at Start of Treatment</td>
<td>4;9</td>
<td>4;4</td>
</tr>
<tr>
<td>Treatment Sound</td>
<td>/ɹ/</td>
<td>/ɻ/</td>
</tr>
<tr>
<td>GFTA-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Score</td>
<td>67</td>
<td>72</td>
</tr>
<tr>
<td>Percentile Rank</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>CELF-P2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Score</td>
<td>112</td>
<td>96</td>
</tr>
<tr>
<td>Percentile Rank</td>
<td>79</td>
<td>39</td>
</tr>
<tr>
<td>Phonemes Absent from Phonemic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inventory (Pre-Tx)</td>
<td>/g, ɵ, δ, ʒ, h, l, ɹ/</td>
<td>/g, ɵ, δ, ʃ, ʒ, ʃ, ɹ/</td>
</tr>
<tr>
<td></td>
<td>/v, ɻ, ʃ, ʒ, ɻ, ɹ/</td>
<td>/v, ɻ, ʃ, ʒ, ɻ, ɹ/</td>
</tr>
</tbody>
</table>

**Stimuli**

Treatment sessions consisted of traditional speech treatment using words that began with complex sounds, either /ɹ/ or /ɻ/ (Gierut, 2001, 2007). All treatment sound targets were phonemes that were absent from the child’s phonemic inventory prior to treatment. Each child’s treatment sound target was selected using the complexity approach. The basis for selecting a complex sound target was that teaching more complex, linguistically marked phonological elements not in a child's phonological
system has the potential to generalize to the learning of other sounds (Gierut, 2001, 2007; Gierut et al., 1996).

All sound targets were elicited word-initially during treatment through productions of five academic vocabulary (AV) words. An AV word is a word that frequently occurs across a variety of academic texts and literature (Coxhead, 2000; Nagy & Townsend, 2012; Nation, 2001). Though AV words are common in written text, they occur infrequently in discourse especially in the speech of, and to, young children. The AV words provided novel phonological forms while also promoting learning of real words to populations at risk for vocabulary deficits (Leonard, Schwartz, Morris, & Chapman, 1981; Schwartz & Leonard, 1982). The AV words in this study were selected from the High-Incidence Academic Vocabulary word list (Coxhead, 2000). Treatment words included “route,” “random,” “remove,” “region,” and “rigid” for /ɹ/ treatment, and “locate,” “labor,” “layer,” “logical,” and “lecture” for /l/ treatment. A variety of lexical and phonological variables were measured for all treatment target words and have been recorded below (Table 2).

Treatment target words were first introduced to the child through a story book format. The story books were written and illustrated by certified SLP and clinician of the present study, Janet Hallgrimson, and can be accessed at the following website: http://www.bblab.org/wordtypes.php. Color picture cards were created from the illustrations in each treatment story and were introduced to the children in conjunction with their five treatment words. These cards were used during treatment to elicit the child's productions. Figure 1 provides the list of treatment target words with their corresponding picture book illustrations and quick incidental learning (Oetting, Rice, &
Swank, 1995) definitions. The definition was provided to all children during all book readings to promote learning of the academic vocabulary words. The quick definitions were used to help the children process and apply some of the semantic and syntactic characteristics of the AV words.

Table 2. Treatment target word properties (i.e., density, age of acquisition, word frequency).

<table>
<thead>
<tr>
<th>Treatment Target Word</th>
<th>Density</th>
<th>Age of Acquisition</th>
<th>Word Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>/r/ Words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>route</td>
<td>5</td>
<td>11.40</td>
<td>3.0374</td>
</tr>
<tr>
<td>random</td>
<td>1</td>
<td>9.33</td>
<td>2.7185</td>
</tr>
<tr>
<td>remove</td>
<td>4</td>
<td>5.67</td>
<td>3.0354</td>
</tr>
<tr>
<td>region</td>
<td>2</td>
<td>9.14</td>
<td>2.4099</td>
</tr>
<tr>
<td>rigid</td>
<td>1</td>
<td>9.16</td>
<td>1.9823</td>
</tr>
<tr>
<td>/l/ Words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>locate</td>
<td>1</td>
<td>7.95</td>
<td>2.7202</td>
</tr>
<tr>
<td>labor</td>
<td>1</td>
<td>8.28</td>
<td>2.7938</td>
</tr>
<tr>
<td>layer</td>
<td>8</td>
<td>8.28</td>
<td>2.2788</td>
</tr>
<tr>
<td>logical</td>
<td>0</td>
<td>9.83</td>
<td>2.5775</td>
</tr>
<tr>
<td>lecture</td>
<td>3</td>
<td>9.50</td>
<td>2.7284</td>
</tr>
</tbody>
</table>

Notes:

Density: Phonological neighborhood density was calculated using the CLEARPOND online corpus (Marian, Bartolotti, Chabal, & Shook, 2012). Density scores <10 are considered low density; thus, all treatment target words used in this study were low density.

Age of Acquisition: The age at which a word typically enters a child’s vocabulary was recorded based on a database of over 30,000 English words (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012).

Word Frequency: The SUBTLEX\textsubscript{US} online corpus (Brysbaert & New, 2009) was used to determine whether treatment target words were high or low in frequency. A word with a log10 frequency score of 3.0 or higher in the SUBTLEX\textsubscript{US} corpus was considered high frequency. Though most of these AV words had log10 word frequency measures below 3.0, two of the selected treatment words (i.e., “route” and “remove”) were considered high frequency.
Figure 1. The Academic Vocabulary (AV) words used in treatment to target /ɹ/ and /l/ with their associated picture stimuli and incidental learning definitions.

**Procedures**

A single subject, multiple baseline design, in which every child served as his own control, was used in this study (Dinnsen & Gierut, 2008; McReynolds & Kearns, 1983; McReynolds & Thompson, 1986). This research design has been implemented in many different types of treatment studies involving children with SSD (Cummings & Barlow,
Using this design, the four participants were divided into two groups. Parents were given the option to select one of two treatment frequency conditions for their child: 4x/week for five weeks or 2x/week for ten weeks. Both options involved a total of twenty, 50-minute individual sessions. This meant that by the end of the child’s treatment block, each participant had received individual therapy for a total of 1000 minutes. A goal of 100 productions of treatment target words was implemented in all sessions. On average, 93 productions (range: 60-112) of the child’s treatment words were produced in a single session, resulting in an average cumulative intervention intensity of 1870 productions (range: 1767-1949) across all participants. Table 3 lists the intervention intensity variables for the treatment program, while Table 4 provides the specific information regarding treatment dose across all sessions for each participant.

Table 3. Intervention intensity variables for the present study's treatment program.

<table>
<thead>
<tr>
<th>Intervention Intensity Variable</th>
<th>2x/week Condition</th>
<th>4x/week Condition</th>
<th>All Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose Form</td>
<td>Drill Play</td>
<td>Drill Play</td>
<td>Drill Play</td>
</tr>
<tr>
<td>Dose</td>
<td>~100 trials/session (Actual: 92.90)</td>
<td>~100 trials/session (Actual: 94.08)</td>
<td>~100 trials/session (Actual: 93.49)</td>
</tr>
<tr>
<td>Session Duration</td>
<td>50 minutes</td>
<td>50 minutes</td>
<td>50 minutes</td>
</tr>
<tr>
<td>Dose Frequency</td>
<td>2x/week</td>
<td>4x/week</td>
<td>-</td>
</tr>
<tr>
<td>Total Intervention Duration</td>
<td>10 weeks</td>
<td>5 weeks</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 4. Treatment dose (i.e., number of trials) across all sessions for each participant.

<table>
<thead>
<tr>
<th>Session #</th>
<th>Imitation (I) or Spontaneous (S) Productions</th>
<th>2x/week Group</th>
<th>4x/week Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Child 1</td>
<td>Child 2</td>
</tr>
<tr>
<td>1</td>
<td>I</td>
<td>82</td>
<td>85</td>
</tr>
<tr>
<td>2</td>
<td>I</td>
<td>90</td>
<td>108</td>
</tr>
<tr>
<td>3</td>
<td>I</td>
<td>91</td>
<td>99</td>
</tr>
<tr>
<td>4</td>
<td>I</td>
<td>60</td>
<td>97</td>
</tr>
<tr>
<td>5</td>
<td>I</td>
<td>62</td>
<td>102</td>
</tr>
<tr>
<td>6</td>
<td>I</td>
<td>79</td>
<td>108</td>
</tr>
<tr>
<td>7</td>
<td>I</td>
<td>80</td>
<td>110</td>
</tr>
<tr>
<td>8</td>
<td>I</td>
<td>76</td>
<td>102</td>
</tr>
<tr>
<td>9</td>
<td>I</td>
<td>89</td>
<td>104</td>
</tr>
<tr>
<td>10</td>
<td>I</td>
<td>74</td>
<td>97</td>
</tr>
<tr>
<td>11</td>
<td>S</td>
<td>102</td>
<td>101</td>
</tr>
<tr>
<td>12</td>
<td>S</td>
<td>95</td>
<td>98</td>
</tr>
<tr>
<td>13</td>
<td>S</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>14</td>
<td>S</td>
<td>100</td>
<td>92</td>
</tr>
<tr>
<td>15</td>
<td>S</td>
<td>98</td>
<td>81</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>102</td>
<td>102</td>
</tr>
<tr>
<td>17</td>
<td>S</td>
<td>97</td>
<td>80</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>95</td>
<td>90</td>
</tr>
<tr>
<td>19</td>
<td>S</td>
<td>99</td>
<td>96</td>
</tr>
<tr>
<td>20</td>
<td>S</td>
<td>101</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total # of Trials</strong></td>
<td></td>
<td>1767</td>
<td>1949</td>
</tr>
<tr>
<td><strong>Average # of Trials per Session</strong></td>
<td></td>
<td>88.4</td>
<td>97.5</td>
</tr>
</tbody>
</table>

All children completed assessments using the *Goldman-Fristoe Test of Articulation – 3rd Edition* (GFTA-3; Goldman & Fristoe, 2015) and the *Little PEEP: Shorter Protocol for the Evaluation of English Phonotactics* (PEEP; Barlow, 2012) prior to the commencement of treatment and immediately after treatment was finished. The treatment program consisted of a non-contrastive approach where a single sound target was practiced throughout all twenty sessions. The traditional approach involved
identification of the treatment sound, discriminating it from its error, shaping it through feedback until productions were correct, and then stabilizing it to improve its consistency across all speaking situations (van Riper, 1978). A licensed and certified SLP administered all of the assessment, treatment, and probing sessions for all children. To control for fidelity of treatment, all sessions were led by the same clinician who closely tracked the time spent on each treatment task throughout all treatment sessions. This information was recorded on a tracking sheet every session to ensure that the treatment program was administered in a similar manner across all participants. Consistent with procedures used previously in the literature (e.g. Cummings & Barlow, 2011; Gierut, 1992; Gierut & Morisette, 2010; Gierut et al., 1996; Gierut & Neumann, 1992; Morisette & Gierut, 2002), treatment was delivered in two phases: imitation and spontaneous production. The imitation phase encompassed the first ten sessions of treatment, followed by ten spontaneous production sessions.

**Imitation treatment phase.** During the imitation phase of treatment, each child repeated the clinician’s verbal model until a time-based criterion of 10 sessions was completed. Each imitation session began with the child verbally identifying their treatment sound (e.g., the “lollipop sound” for /l/ or the “angry dog sound” for /ɹ/). Then, the clinician read the selected treatment story, which contained the child’s five treatment target words. After completing the book, the clinician provided five to ten minutes of direct placement and sound-shaping therapy during which each child was given verbal, tactile, and physical cues to help elicit the child’s target sound. The remainder of each session involved various drill-play activities to elicit productions. In addition, children were taught to rate the correctness of their treatment target word productions on a scale of
one to five. On average, 91 responses (range: 60–112) were elicited from each child per session during the imitation phase of treatment.

**Spontaneous production treatment phase.** During the spontaneous production phase of treatment, each child produced their treatment words without a model. This phase of treatment continued for a time-based criterion of 10 consecutive sessions to ensure that all study participants received a total of 20 treatment sessions. All spontaneous production sessions began by having the child verbally identify their treatment sound. Once this task was complete, the clinician read a story to the child that did not contain the child’s treatment target words. The remainder of the session consisted of drill-play activities and tasks involving rating their own productions of treatment target words. On average, 96 responses (range: 76–105) were elicited from each child per session during the spontaneous production phase of treatment.

**Data analysis and reliability transcription**

Multiple measures were used to examine phonological change following the child’s specific treatment program. Daily treatment data was analyzed to measure learning during treatment. System-wide phonological generalization was measured by comparing the child’s productions of sounds in words during pre- and post-treatment probes. The completed probes included a standardized test protocol (i.e., *Goldman-Fristoe Test of Articulation – 3rd Edition (GFTA-3; Goldman & Fristoe, 2015)* and a 284-word list titled *Little PEEP: Shorter Protocol for the Evaluation of English Phonotactics (PEEP; Barlow, 2012)*). Specifically, several different measures were calculated both pre- and post-treatment using the child’s combined productions from these two probes: phonemic inventories (Gierut, Simmerman, & Neumann, 1994), cluster inventories
(Dinnsen, Chin, Elbert, & Powell, 1990), percent consonants correct in untreated words (PCC; Shriberg et al., 1997; Shriberg & Kwiatkowski, 1982), and production accuracy of the treatment sound in untreated words (Gierut & Morrisette, 2012a; Morrisette & Gierut, 2002). Additional measures, including the effect size $d$ (Busk & Serlin, 1992) and learning rates (Cummings & Babchishin, 2017), were also calculated to improve the comparisons between participants. The effect size $d$ is a statistical value for standard mean difference (SMD) which has been previously illustrated as a well-suited way to evaluate generalization and compare group outcomes, especially in single subject, multiple baseline designs (Gierut & Morrisette, 2011). While SMD has already been reported in previous studies of clinical phonology (Gierut & Morrisette, 2011, 2012b, 2012a, 2014), the calculation of learning rates is a newer proposed measure that controls for the effects of cumulative treatment intensity. Essentially, each child’s amount of practice during treatment (i.e., treatment trials) becomes equalized when describing a child’s performance based on the change that occurred following each trial.

**Daily treatment data.** As discussed by Gierut & Champion (2001), learning during treatment is relevant to establishing treatment effectiveness. In this study, each child’s learning during treatment is defined by their production accuracy of their treatment sound in the word-initial context of their five selected treatment words. The clinician judged production accuracy trial-by-trial during all treatment sessions. Sounds were only counted as correct if they were produced in a manner similar to that of an adult in the ambient language (i.e. prolonged sounds, segmented productions, and slightly distorted productions were judged to be incorrect); thus, this measurement provided a conservative measure of sound learning.
**Generalization.** Examining the generalization from treatment is arguably even more important than the child’s performance on treatment target sound productions, as it reflects the overall effects of treatment on the child’s entire phonological system. This change is what ultimately impacts the child’s intelligibility of spontaneous productions. Generalization is reported for treated and untreated singleton phonemes in all word positions of untreated words as a reflection of overall change in the child’s phonological system. To determine generalization of sounds, a Percent Consonants Correct (PCC) score was calculated based on the child’s GFTA-3 (Goldman & Fristoe, 2015) and PEEP (Barlow, 2012) performance measured pre- and post-treatment. These data were derived from the phonetic transcriptions of the two word probes. To ensure transcription reliability, an estimate of inter-rater agreement was obtained. Trained listeners used the International Phonetic Alphabet (IPA) to narrowly transcribe all speech samples using the PHON computer transcription and data analysis program. Using the blind transcriber function in PHON (Rose & MacWhinney, 2014), 100% of all GFTA-3 and PEEP speech samples were reliability-checked by a second transcriber. The speech samples were compared for point-by-point consonant agreement to ensure accuracy in the transcriptions of the child’s productions. An agreement threshold of at least 85% was required for each sample. Overall, the transcriber reliability was 92.63% across the 16 speech samples. Following the reliability check, transcription discrepancies were discussed until consensus was reached among the first and second transcribers.

Based on the transcriptions, the data were organized for standard descriptive phonological analysis according to target sound and word position (Dinnsen, 1984) in PHON. Two production accuracy values were calculated for each child: 1) the treatment
target sound in untreated words and 2) all consonants in untreated words in an overall percent consonants correct (PCC) analysis. Each consonant was point-by-point identified as being correct or incorrect in relation to its target phoneme. From these values, generalization data from each child were examined from both qualitative and quantitative perspectives (Gierut & Morrisette, 2012a; Morrisette & Gierut, 2002).

The qualitative description of generalization applied the accepted single subject designs learning criterion level of 10% or greater mean generalization accuracy change in treatment sounds and/or PCC scores (Elbert, Dinnsen, & Powell, 1984; Gierut & Morrisette, 2012a; McReynolds & Kearns, 1983). For each child and each treatment condition, generalization learning was examined to see if the 10% criterion was met.

**Effect Sizes.** To establish the relative magnitude of the generalization gains, an effect size measurement was included as a quantitative analysis. That is, effect size values made it possible to compare across treatment conditions to identify which is relatively most efficacious. While there are a variety of effect size measurements, the standard mean difference (SMD) was chosen as a conservative estimate of treatment effects in this single subject design (Busk & Serlin, 1992; J. M. Campbell, 2004; Olive & Smith, 2005).

The SMD involves the computation of two means: the mean of the baseline data ($M_A$) and the mean of the generalization data ($M_B$) (Busk & Serlin, 1992). The difference in the two means ($M_B - M_A$) creates the numerator of the equation. The standard deviation (SD) of the baseline data ($SD_A$) is the denominator of the equation, which upon division, results in the effect size $d$. The one difficulty with this equation is that children in speech treatment studies often have 0% accuracy pre-treatment and no variability in
performance across baseline sessions (leading to a SD of 0). An alternative to this equation is to pool all of the baseline data across participants to create a SD across subjects (SD_{A-pooled across Ss}) (Gierut & Morissette, 2011; Gierut, Morissette, & Dickinson, 2015). This new SD value then reflects the actual baseline variability for all of the children in a study. Thus, in the present study, the revised SMD effect size equation was used to establish the relative magnitude of generalization. For each child, SMD $d$ scores were calculated separately for the treated sound accuracy (Table 5/Figure 6) and PCC (Table 6/Figure 7) using the following formula:

$$d = \frac{(M_B - M_A)}{SD_{A-pooled across Ss}}$$

Gierut and colleagues (2015) identified mean $d$ values for small, medium, and large effect sizes as $d = 1.40$, $d = 3.61$, and $d = 10.12$, respectfully.

**Learning Rates.** Variation in the exposure and practice of treatment sounds in treatment words may impact a child’s phonological learning. In other words, the cumulative intervention intensity (Allen, 2013; Baker & McLeod, 2011b; Warren et al., 2007), or total number of productions throughout the treatment program, may be an important factor to consider when analyzing the phonological change in a child. In the case of this study, the exact number of treatment trials that occurred in each session for each child was tracked and recorded daily. Thus, cumulative intervention intensity for all participants was available and ranged between 1767 and 1949 trials. Though this difference in total productions among participants was not substantial, it was important to describe each child’s performance with this measure to eliminate the additional confounding variable of intensity. This study aimed to alter intervention intensity by
exclusively adjusting dose frequency, or number of sessions per week, between conditions; thus, variation in the other intervention intensity variables were to be kept to a minimum.

In an attempt to control for the unavoidable variations in sound production opportunities, learning rates were computed for both the treated sound accuracy change and overall PCC change. The change in each child’s production accuracy of the treated sound and PCC score were computed first. Then, the total number of treatment trials completed during the treatment program was calculated for each child. By dividing the number of trials by the overall accuracy change, a value was computed to indicate the number of trials each child needed in order to make a 1% generalization gain in either production accuracy of their treated sound or overall PCC. Each child’s learning rate per trial was also calculated, which is just another way of representing the same data. That is, the learning rate was calculated by dividing the overall production accuracy change by the number of trials completed.

**Phonemic and Cluster Inventories.** Pre- and post-treatment phonemic and word-initial cluster inventories were also calculated in *PHON*. This measure was used to determine the number of phonemes and clusters added to each child’s inventory following treatment. The AutoPATT plugin (Combiths, Amberg, & Barlow, 2016) ([https://github.com/rayamberg/AutoPATT](https://github.com/rayamberg/AutoPATT)) was used to automatically calculate these inventories. The AutoPATT identified phonemes as part of the child’s inventory if they were produced in at least two unique minimal pair sets during the assessments (Gierut et al., 1994). A word-initial cluster was identified as part of the child’s inventory if it was produced in the onset position of a syllable at least twice during the child’s speech
samples. Following the computation of pre- and post-treatment inventories, the number of phonemes and clusters added to the child’s inventory following treatment were separately calculated. The number of added phonemes excluded non-English consonants, allophones (i.e., sounds with diacritics that represent a single phoneme, such as unreleased /t/ or aspirated /k/) when the phoneme it represented was already counted, and distorted productions (such as /w/ with r coloring: /wʰ/). In regard to consonant clusters, all two- and three-element clusters produced in the post-treatment assessment probes but not in the pre-treatment probes were counted as new/added clusters regardless of their presence in English words. This criterion was considered appropriate since the number of added clusters simply described the child’s improved phonological skills in combining consonants even if the production was not correct based on the adult target.
CHAPTER III

RESULTS

The results of providing treatment at different dose frequencies are discussed in terms of daily treatment data, generalization from treatment (Gierut & Morrisette, 2012a; Gierut et al., 2015; Morrisette & Gierut, 2002; Shriberg et al., 1997; Shriberg & Kwiatkowski, 1982), effect sizes (Busk & Serlin, 1992), learning rates (Cummings & Babchishin, 2017), and phonemic/cluster inventories (Dinnsen et al., 1990; Gierut et al., 1994).

Daily treatment data

The learning curves reveal each child’s production accuracy of their targeted sound during treatment (Figures 2 & 3). Each graph displays the production accuracies of the child’s treatment target sound in their five pre-determined treatment words across all twenty sessions. Data has been plotted longitudinally for the two phases of treatment: imitation phase and spontaneous phase. Figures 2 and 3 display the daily treatment data for children treated with /l/ and /ɹ/ sounds, respectively.

Beginning with the treatment data of children treated with the /l/ phoneme (i.e., Figure 2), it can be seen that children in both conditions had a production accuracy of 0% for their treatment sound during the first four treatment sessions. By the final session of treatment, production accuracy of /l/ was 94% and 93% for Child 2 in the 2x/week and Child 4 in the 4x/week conditions, respectively. This demonstrates a significant
improvement from pre-treatment production accuracy of their treatment sound /l/.

Notably, the general trend of learning appears very similar for both children receiving treatment for /l/ as the learning curves overlapped greatly.

![Learning curves displaying the daily treatment data for study participants involved in speech treatment targeting /l/. Data points on the graph indicate production accuracies of the targeted treatment phoneme in the word-initial position of the child’s treatment words across all twenty treatment sessions. The curve of the child in the 2x/week condition is marked in gray, while the curve for the child in the 4x/week condition is black.](image)

Figure 2. Learning curves displaying the daily treatment data for study participants involved in speech treatment targeting /l/. Data points on the graph indicate production accuracies of the targeted treatment phoneme in the word-initial position of the child’s treatment words across all twenty treatment sessions. The curve of the child in the 2x/week condition is marked in gray, while the curve for the child in the 4x/week condition is black.

Treatment data of children treated with the /ɹ/ phoneme (i.e., Figure 3) shows a similar trend when comparing the performance of children in the two intensity conditions. Both children began treatment with a production accuracy of 0% for their treated sound in treatment target words. Following twenty sessions, the production accuracy of /ɹ/ improved to 78% and 89% for Child 1 in the 2x/week and Child 3 in the 4x/week conditions, respectively. The learning curve of Child 3 appears slightly more sporadic than the continuous incline displayed in the learning curve of Child 1, but in general the two learning curves display comparable trends.
Figure 3. Learning curves displaying the daily treatment data for study participants involved in speech treatment targeting /ɹ/. Data points on the graph indicate production accuracies of the targeted treatment phoneme in the word-initial position of the child’s treatment words across all twenty treatment sessions. The curve of the child in the 2x/week condition is marked in gray, while the curve for the child in the 4x/week condition is black.

Overall, the learning curve analyses for both /ɹ/ and /l/ treatment programs indicate a similar pattern of learning when producing the treatment sound in treatment words regardless of the intensity condition in which the child participated.

**Generalization: Treated sound**

When considering qualitative generalization patterns, just one of the two children in the 2x/week condition (i.e., Child 2) demonstrated a significant (McReynolds & Kearns, 1983) treatment sound production accuracy gain greater than 10% (Table 5, Figure 4). Comparatively, both children in the 4x/week condition (Child 3 & Child 4) demonstrated accuracy gains above 10%. At a group level, only the 4x/week condition demonstrated an averaged treatment sound accuracy gain greater than 10% (i.e., 19.26%); the children in the 2x/week condition improved their accuracy by 6.942%. Thus, when
considering the treatment sound generalization, the 4x/week condition displayed a clear
advantage over the 2x/week condition.

Table 5. Treatment sound production accuracy and standard mean difference ($d$) computations. Data includes the treatment sound production accuracy scores for pre-treatment ($M_A$) and post-treatment assessments ($M_B$), and individual $d$ values for both children in either treatment condition. The qualitative generalization or percentage accuracy change following treatment is determined by subtracting the $M_A$ from the $M_B$ ($M_B - M_A$). An average value of $M_B - M_A$ and $d$ for the 2x/week and 4x/week groups has been calculated to improve comparisons across conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Child</th>
<th>$M_A$</th>
<th>$M_B$</th>
<th>$M_B - M_A$</th>
<th>$M_B - M_A$ by condition</th>
<th>$SD_A^a$</th>
<th>$d$ by condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x/week</td>
<td>1</td>
<td>0.730%</td>
<td>3.676%</td>
<td>2.947%</td>
<td>6.942%</td>
<td>0.435</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.563%</td>
<td>12.500%</td>
<td>10.938%</td>
<td></td>
<td>0.980</td>
<td>6.93</td>
</tr>
<tr>
<td>4x/week</td>
<td>3</td>
<td>25.547%</td>
<td>50.000%</td>
<td>24.453%</td>
<td>19.258%</td>
<td>4.901</td>
<td>15.49</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.000%</td>
<td>14.063%</td>
<td>14.063%</td>
<td></td>
<td>0.000</td>
<td>8.91</td>
</tr>
</tbody>
</table>

Notes:

$$d = \frac{(M_B - M_A)}{SD_{A-pooled across Ss}}$$

$^a$Each child’s baseline SD is reported. The value of the pooled SDs (i.e., $SD_{A-pooled across Ss}$) = 1.579

Figure 4. Individual children’s generalization of their treatment sound in untreated words. The accuracy values represent each children’s ability to produce their treatment sound in the speech probes administered pre- and post-treatment. A 10% pre-to-post-treatment change was the defined qualitative measure of treatment success. One child in the 2x/week condition (Child 2) and both children in the 4x/week condition (Child 3 and Child 4) met the 10% gain criterion.
Generalization: All consonants

In terms of overall PCC gains, only Child 3 from the 4x/week condition demonstrated a PCC production accuracy gain greater than 10% (Table 6, Figure 5). It should be noted that the other child in the 4x/week condition (Child 4) improved his PCC accuracy by 9.786%, just missing the 10% cut-off. Overall, the averaged PCC accuracy gains were greater than 10% for the 4x/week condition; the 2x/week condition did not meet this 10% criterion.

Table 6. Percent Consonants Correct (PCC) and standard mean difference (d) computations. Data includes the PCC accuracy scores for pre-treatment (M_A) and post-treatment assessments (M_B), and individual d values for both children in either treatment condition. The qualitative generalization or percentage accuracy change following treatment is determined by subtracting the M_A from the M_B (M_B−M_A). An average value of M_B−M_A and d for the 2x/week and 4x/week groups has been calculated to improve comparisons across conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Child</th>
<th>M_A</th>
<th>M_B</th>
<th>M_B−M_A</th>
<th>M_B−M_A by condition</th>
<th>SD_A^a</th>
<th>d</th>
<th>d by condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x/week</td>
<td>1</td>
<td>61.254%</td>
<td>66.667%</td>
<td>5.413%</td>
<td>5.313%</td>
<td>1.571</td>
<td>3.79</td>
<td>3.72</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>40.664%</td>
<td>45.877%</td>
<td>5.213%</td>
<td>0.346</td>
<td>3.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4x/week</td>
<td>3</td>
<td>61.400%</td>
<td>71.822%</td>
<td>10.421%</td>
<td>10.104%</td>
<td>0.420</td>
<td>7.30</td>
<td>7.08</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>49.284%</td>
<td>59.069%</td>
<td>9.786%</td>
<td>3.374</td>
<td>6.85</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

^aEach child’s baseline SD is reported. The value of the pooled SDs (i.e., SD_A-pooled across Ss) = 1.428
Effect sizes

When considering the treatment sound effect sizes for children individually, the children in the 2x/week group had $d$ scores of 1.87 and 6.93, while children in the 4x/week group received $d$ scores of 8.91 and 15.49 (Table 5; Figure 6). Thus, both children in the 4x/week group displayed an advantage over the 2x/week participants when comparing the magnitude of the effect size for learning their treatment sound. These effect sizes were also interpreted using the criteria proposed by Gierut and colleagues (2015): small effect: $d = 1.40$, medium effect: $d = 3.61$, large effect: $d = 10.12$. One participant of the 4x/week condition (Child 3) displayed a large effect while the
other participant of the 4x/week condition (Child 4) had a medium effect size. Children from the 2x/week condition had effect sizes of medium (Child 2) and small (Child 1).

At the group level, the effect size associated with the 4x/week group ($d = 12.20$) was larger than the effect size for the 2x/week condition ($d = 4.40$). Specifically, the 4x/week group elicited a large treatment effect, while the 2x/week condition elicited a medium treatment effect (Gierut et al., 2015). Furthermore, the 4x/week condition had a $d$ score that tripled the magnitude of the 2x/week condition, suggesting that the more intensive treatment schedule had a greater impact on children’s phonological learning of treated sounds than did the less intensive program.

![Figure 6](image.png)

Figure 6. Individual children’s treatment sound effect sizes. A small effect size was observed for Child 1, a medium effect size for Child 2 and Child 4, and a large effect size for Child 3.

Effect sizes for the overall PCC scores were also calculated as a measure to compare system-wide phonological change across all participants. Individually, children in the 2x/week group had $d$ scores of 3.65 and 3.79 while children in the 4x/week group
received \(d\) scores of 6.85 and 7.30 (Table 6; Figure 7). At the group level, the 2x/week condition elicited a \(d\) score of 3.72 and the 4x/week condition resulted in a \(d\) score of 7.08. Interpreting these PCC effect sizes based on the criteria presented by Gierut and colleagues (2015) for treated sound effect sizes, both treatment conditions elicited medium treatment effects. Given that Gierut and colleagues did not specifically address PCC \(d\) scores though, it is not known whether these effect sizes are appropriate for a PCC analysis. Continued use of effect sizes in treatment studies with PCC scores will hopefully establish some common standards. Even though interpretation using Gierut and colleague’s criteria indicate the same effect size across intensity conditions, it should be noted that the 4x/week condition had a \(d\) score that was double the magnitude of the \(d\) score for the 2x/week condition; thus, comparing effect sizes for PCC gains still demonstrates an advantage for an intensive, 4x/week treatment program.

![Figure 7. Individual children’s PCC effect sizes. All participants in both intensity conditions displayed a medium effect size.](image-url)
Learning rates

It is important, especially in a study designed to examine the effect of a single intensity variable, to control for the variability in cumulative intervention intensity across participants. Such a measure provides a fair and conservative measure of treatment outcomes for participants within the two group conditions since the amount of learning and practice for each child during treatment (i.e., treatment trial attempts) becomes equalized.

In the treated sound analysis, children in the 2x/week condition required 599.69 trials (0.0017 percentage accuracy/trial) and 178.19 trials (0.0056 percentage accuracy/trial) to achieve a 1% accuracy gain. Children in the 4x/week condition required 129.00 trials (0.0078 percentage accuracy/trial) or 79.71 trials (0.0125 percentage accuracy/trial) for a 1% accuracy gain (Table 7, Figure 8). At the group level, the 4x/week condition displayed more efficient learning than the 2x/week condition. Children in the 4x/week condition required 104.35 trials for a single percentage point accuracy gain (0.0101 percentage accuracy/trial), while children in the 2x/week condition required 388.94 trials (0.0036 percentage accuracy/trial) to make the same gain. Thus, even when the number of treatment trials was controlled, the 4x/week condition elicited more efficient change of the treatment target sound in untreated words.

Table 7. Learning rates of treated sound accuracy change.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Child</th>
<th>Treated Sound Accuracy Change (%)</th>
<th>Cumulative Intervention Intensity</th>
<th>Learning Rate (%Gain Per Trial)</th>
<th># of Trials to Make 1% Generalization Gain</th>
<th>Learning Rate by condition</th>
<th># of Trials to Make 1% Generalization Gain by condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x/week</td>
<td>1</td>
<td>2.947</td>
<td>1767</td>
<td>0.0017</td>
<td>599.69</td>
<td>0.0036</td>
<td>388.94</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>10.938</td>
<td>1949</td>
<td>0.0056</td>
<td>178.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4x/week</td>
<td>3</td>
<td>24.453</td>
<td>1949</td>
<td>0.0125</td>
<td>79.71</td>
<td>0.0101</td>
<td>104.35</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>14.063</td>
<td>1814</td>
<td>0.0078</td>
<td>129.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the PCC analysis, children in the 2x/week condition required 326.43 trials (0.00306 percentage accuracy/trial) and 373.85 trials (0.00267 percentage accuracy/trial) to achieve a 1% accuracy gain. Children in the 4x/week condition required 187.02 trials (0.00535 percentage accuracy/trial) or 185.37 trials (0.00539 percentage accuracy/trial) for a 1% accuracy gain (Table 8, Figure 8). At the group level, the 4x/week condition displayed more efficient learning than the 2x/week condition. On average, children in the 4x/week condition required 186.20 trials for a single percentage point accuracy gain (0.00537 percentage accuracy/trial), while children in the 2x/week condition required 350.14 trials (0.00287 percentage accuracy/trial) to make the same gain. Thus, similar to the treated sound analysis, the 4x/week condition was more efficient in eliciting system-wide phonological change than was the 2x/week condition.

Table 8. Learning rates from PCC change.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Child</th>
<th>PCC Change (%)</th>
<th>Cumulative Intervention Intensity</th>
<th>Learning Rate (%Gain Per Trial)</th>
<th># of Trials to Make 1% Generalization Gain</th>
<th>Learning Rate by condition</th>
<th># of Trials to Make 1% Generalization Gain by condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x/week</td>
<td>1</td>
<td>5.413</td>
<td>1767</td>
<td>0.00306</td>
<td>326.43</td>
<td>0.00287</td>
<td>350.14</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.213</td>
<td>1949</td>
<td>0.00267</td>
<td>373.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4x/week</td>
<td>3</td>
<td>10.421</td>
<td>1949</td>
<td>0.00535</td>
<td>187.02</td>
<td>0.00537</td>
<td>186.20</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9.786</td>
<td>1814</td>
<td>0.00539</td>
<td>185.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To summarize, both the treated sound learning rates and the PCC learning rates can be used as a reference to make a stronger argument for the 4x/week condition. That is, children in the 4x/week condition responded better to treatment even after controlling for the differences among total number of practice opportunities during treatment sessions.

**Phonemic and cluster inventories**

The phonemes missing from each child’s phoneme inventory pre-treatment and those added post-treatment are listed in Table 9. The number of phonemes added per child ranged from 1 to 3. For the 2x/week group, children added one or two phonemes, with a total of three phonemes added for the group. In the 4x/week group, children added either two or three phonemes to their inventory, with a total of five phonemes added for
the group. Children treated with /l/ in both intensity conditions added their treatment target sound to their inventory. None of the children treated with the /u/ phoneme added this sound to their inventory. Thus, in terms of the number of phonemes added, the 4x/week condition appeared to elicit the most widespread phonological change.

Table 9. Pre- and post-treatment phonemic inventories.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Child</th>
<th>Phonemes Missing</th>
<th>Added Phonemes</th>
<th># of Added Phonemes</th>
<th># of Added Phonemes by condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x/week</td>
<td>1</td>
<td>/g, o, ɔ, ʒ, h, l, j/</td>
<td>/o/</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>/v, ɔ, f, ʒ, ʃ, dʒ, l, j/</td>
<td>/v, l/</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4x/week</td>
<td>3</td>
<td>/g, η, v, o, ɔ, z, ʒ, h, j/</td>
<td>/ŋ, v, θ/</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>/g, η, f, ʒ, ʃ, dʒ, l, j, j/</td>
<td>/ŋ, l/</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

The consonant clusters present in each child’s inventory pre-treatment and the clusters that were added following treatment are listed in Table 10. The number of clusters added per child ranged from 1 to 7. For the 2x/week group, children added either one or two consonant clusters, with a total of three clusters added for the group. None of the three clusters added were adult-like English clusters, and none of the clusters included the child’s treatment target sound. In the 4x/week group, children added either two or seven clusters to their inventory, with a total of nine clusters added for the group. Five of these nine clusters can be found in English target words. Furthermore, two of these clusters contained the child’s treatment target sound (i.e., /gu/ and /stu/). Three-element clusters were only added to the inventory of a child treated in the 4x/week intensity condition. Notably, one of these three-element clusters included the child’s treatment
sound /\(\textit{s}\)/ (i.e., /\textit{st}1/). Thus, in terms of the number and complexity of clusters added, the 4\(x\)/week condition appeared to elicit greater phonological change in children.

Table 10. Pre- and post-treatment cluster inventories.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Child</th>
<th>Pre-Tx Clusters</th>
<th>Added Clusters</th>
<th># of Added Clusters</th>
<th># of Added Clusters by condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2(x)/week</td>
<td>1</td>
<td>/ pw, tw, kw, bw, gw, sw, fw, ʃw, f(w), d(w) /</td>
<td>/ dw /</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>none</td>
<td>/ bw, mw /</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4(x)/week</td>
<td>3</td>
<td>/ pw, tw, kw, pl, bw, dw, gw, bl, gl, sw, ʃw, fl, d(w), sm, sn, sp, st, sk, skw, f(tw) /</td>
<td>/ kl, g(ʒ), f(w), sl, st, sp, stw /</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>/ pw, sm, sn, sb, sg, sp, sk /</td>
<td>/ f(w), st /</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER IV
DISCUSSION

Using a single subject, multiple baseline design, this study explicitly compared the differences between providing treatment 2x/week versus 4x/week in four children with SSD. This study found that a traditional treatment approach with a more intensive treatment schedule resulted in greater system-wide phonological change than did a less intensive treatment schedule. Specifically, children in the 4x/week condition displayed an advantage over the children in the 2x/week condition when considering generalization of the child’s treatment sound, system-wide generalization of all consonants, effect sizes, learning rates, and added sounds to phonemic and cluster inventories.

Understanding the present study’s findings

It was determined that the children in the 4x/week condition displayed greater gains to production accuracy of their treated sound than the children in the 2x/week condition. On average, the 2x/week condition improved the accuracy of their treatment sound in untreated words by approximately 7% while the 4x/week condition improved treated sound accuracy by nearly 20%. This difference reveals that the more intensive treatment schedule elicited gains in treatment sound accuracy that were almost three times greater than the gains experienced by children on the less intensive schedule. Since much of the previous research on intensity (Allen, 2013; Ausenheimer & Irwin, 1971; Fein, Golman, Kone, & McClintock, 1956; van Hattum, 1959) has focused on a child’s
broader learning of articulatory skills rather than the learning of a specific sound, this
difference observed in the learning of the single sound being practiced during sessions is
a new and important one. The present study’s non-contrastive treatment approach which
targeted only a single sound allows for further interpretation regarding the direct learning
of sounds taught during a therapy session. Even when the child has the same number of
practice opportunities for a given sound, he is able to produce a higher accuracy of that
sound following a treatment program that is completed on an intensive schedule.

System-wide generalization of all English consonants displayed the same trend
when comparing intensity conditions. The 4x/week treatment condition displayed an
average PCC increase that was almost double the PCC of the 2x/week condition (i.e.,
10.1% versus 5.3%). This finding is consistent with Allen’s study (2013) which supports
the idea that more intensive intervention brings about greater improvements to a child’s
overall PCC score.

The effect size analyses further support the qualitative results that have already
been mentioned. Larger effect sizes were determined for both treatment sound change
and PCC change following treatment in the 4x/week group as compared to the 2x/week
group.

In addition, more phonemes and word-initial consonant clusters were added to the
inventories of children in the more intensive treatment program. Since this outcome
measure is new to the study of intensity for SSD treatment, a greater understanding of the
effects of intervention intensity is possible. Unlike the learning of non-speech motor
skills, practice on an intensive schedule displays an advantage for the child even beyond
the context of the sound being taught. This is in contrast with the finding that massed
practice leads to context-dependent learning (Fischman & Lim, 1991). In other words, an advantage is observed for children regarding the accuracy of their treatment sound productions as well as the generalization that occurs and improves accuracy of sounds not targeted during therapy.

The efficacy of the 4x/week treatment schedule is heightened when the learning rates for the treated sound and PCC score are considered. Children in the 4x/week condition responded better to treatment even after controlling for the differences among total number of practice opportunities during treatment sessions. Specifically, the children participating in the less intensive treatment condition required almost four times as many trials as the children in the more intensive treatment condition to make the same percentage accuracy gain in their treatment target sound. When considering change to PCC scores, children in the less intensive treatment required almost twice as many trials than children receiving intensive treatment before they made the same percentage accuracy gain. This difference is important to consider as it reveals the efficiency of a higher treatment dose frequency for eliciting greater change to a child’s phonological system.

One of the most interesting observations made from this study is that if only considering the children’s production accuracy of their treatment words during treatment sessions, a difference between intensity conditions was not apparent. In other words, the learning of treatment sounds in treatment words looked almost identical for children in both intensity conditions. This means that children in both conditions improved their production accuracy of their treatment sound during sessions at a similar rate. By the end of the twenty-session treatment block, each child obtained a production accuracy of their
treatment sound in treated words that was almost identical to the child treated with the same sound in the opposing intensity condition. This point needs to be addressed and emphasized, as it reveals the importance of not only relying on a child’s treatment data when discussing treatment change.

**Why was intensive treatment more effective?**

There may be several explanations for why children elicit greater change to their sound systems when participating in intensive treatment. One possible explanation is that memory processes may impair performance more greatly on a distributed schedule\(^6\) than a condensed schedule\(^7\). For example, some children may have difficulty transferring learned articulatory skills into long-term memory stores. This limitation could increase the need for re-teaching on the distributed schedule since there is more time to forget the skills between sessions. In contrast, an intensive, condensed schedule (e.g., 4x/week) may reduce the need for re-teaching because participants are more likely to remember the skills practiced in a session that more recently occurred.

Another possible advantage of intensive treatment schedules involves the relationship between the clinician and child. Barratt and colleagues (1992) proposed that children quickly build relationships with their therapist when they are seen more frequently. In other words, the therapist-client relationship did not need to be re-established every session as it often does on a more distributed schedule. Building that strong rapport early with a child may elicit greater attention, motivation, and participation from the child throughout a treatment block. In turn, the close relationship between a

---

\(^6\) Term used to describe a treatment program with a low dose frequency (i.e., 2x/week)

\(^7\) Term used to describe a treatment program with a high dose frequency (i.e., 4x/week)
clinician and a child can elicit faster and greater change in a child’s learning of speech sounds.

Other psychosocial benefits of a more intensive treatment approach may have contributed to outcomes as well. For example, conducting more frequent treatment sessions could allow clinicians to become more familiar with their clients and remember the specific strategies used to elicit successful productions in previous sessions. The clinician is more likely to remember useful cueing strategies and/or specific target words in which the client was most successful. This could make treatment more effective since the teaching episodes become more specific to the child’s needs. Furthermore, parents might be more satisfied with the shorter treatment block that corresponds with an intensive treatment schedule because it would not affect family activities for an extended duration of time. Moreover, frequent treatment sessions could also help families establish a consistent routine, leading to more reliable treatment attendance. Thus, preference for an intensive treatment schedule by all participators (i.e., child, clinician, and parent) could improve the outcome of treatment.

Caveats of the present study

The small sample size of this study limits the generalizability of its findings. Due to the limited number of participants, it is possible that group differences were exaggerated from the performance of a single participant. For example, one child from the 4x/week condition (i.e., Child #3) displayed much greater gains than all of the other participants. While this study suggests that intensive dose frequency enhanced the performance of children in the 4x/week group, it is impossible to attribute their greater gains to that variable alone. Randomization with a larger sample size is necessary in
order to determine this causal relationship. While preliminary analyses suggest 4x/week has greater benefits, further research still needs to be conducted before making generalizations regarding optimal intervention intensity for children with SSD.

Another limitation of this study is that post-treatment data was only collected immediately following the child’s 5- or 10-week treatment program. The performance of a skill during or immediately following its acquisition is a poor predictor of long-term retention and learning (Schmidt & Bjork, 1992). In other words, true learning of a skill needs to be measured after a period of time that the child has not received explicit teaching and/or controlled practice opportunities. In order to determine the implications of dose frequency on maintenance of learning, multiple measurements after treatment had been removed would have been appropriate for this study. For example, word probes could have been collected and analyzed for all participating children a few weeks after their treatment block had ended as well as several months post-treatment. Additional conclusions may have emerged with this type of follow-up data collection.

If intensive treatment displays an advantage only when measured immediately after the treatment program ends, a preference for higher intensity treatment would not be purposeful. Children must be able to retain their skills after a maintenance period in order for treatment to be considered useful. It is possible that measurement of a child’s performance following a period of treatment withdrawal could indicate opposing results in the present study. This study may have captured the phenomenon that massed practice elicits faster skill acquisition initially, but overlooked the finding that distributed practice is actually better for generalization and maintenance of these newly learned motor skills. If the learning of speech sound production skills is comparable to the learning of non-
speech motor skills (Baddeley & Longman, 1978; Maas et al., 2008), children may actually display better outcomes in maintenance of learned skills when teaching is delivered in a distributed, less-intensive format. Unfortunately, the present study does not involve measurement of speech sound abilities after a maintenance period; thus, conclusions about intensity differences on maintenance of newly learned skills cannot be made. In order to validate the benefits of one intensity schedule over another, it is important to provide evidence to show that the learned skills have been stabilized by the end of the treatment program. Treatment programs are only successful when maintenance of the learned behaviors continues even after the reinforcement of treatment has been removed.

**Future Research**

All future research in the area of SSD treatment must place a greater emphasis on explicitly defining the intensity variables used in the study’s treatment protocol. Clarity of treatment procedures will help to determine the implications of different intervention intensities on treatment outcomes. It is important to note that even when intensity information has been reported in previous studies, authors have not always used a consistent terminology to define it. Definitions for treatment intensity have varied substantially across studies, including features such as the quality and quantity of service, the number of hours, the level of participation, the proportion of adults to children during treatment, and the number of therapy sessions over a period of time (Warren et al., 2007). Thus, consistently using the intervention intensity terms proposed by Warren, Fey, & Yoder (2007) will improve the conciseness and clarity of identifying intensity variables in future research.
Additional research should create studies that replicate the present study in order to determine the reliability of this study’s results. If a high dose frequency does indeed improve the performance of preschool children receiving traditional SSD treatment, then consistency across repeated studies should be observed. Future research should follow the procedures outlined for a randomized controlled trial (RCT). Unlike the single-subject design utilized in this study, a RCT improves the external validity and generalizability of its findings to all children with SSD. A sample size equal to or greater than 30 (Howell, 2012) is encouraged in order to adequately make group comparisons between intensity conditions.

Further research could also explore the effect of dose frequency across a variety of other treatment approaches for children with SSD. If optimal treatment intensity is indeed specific to the intervention approach being applied (Kaipa & Peterson, 2016), then continued research is necessary across all evidence-based approaches for children with SSD. This means that treatment designs should compare intensity differences for the variety of approaches including traditional intervention (e.g., van Riper & Emerick, 1984), cycles (e.g., Hodson & Paden, 1991), minimal pairs (e.g., Gierut, 1992; Weiner, 1981), multiple oppositions (Williams, 2000), Metaphon (Dean et al., 1995), core vocabulary (Dodd et al., 2006), among others.

Future studies may also examine the effect of dose frequency on treatment outcomes when applying the traditional treatment approach using early-developing treatment sound targets\(^8\) rather than complex sound targets which were used in this study.

---
\(^8\) Early-developing treatment sound targets are selected by referencing studies that have determined the age of acquisition of particular English consonants (Prather & Hedrick, 2012).
It is plausible that children respond differently to frequency conditions when earlier-acquired sounds are targeted in treatment compared to targeting later-developing, more difficult sound targets.

In addition, different treatment durations, such as 30-minute sessions, may also be explored to determine the implications of providing treatment on various intensity schedules. In the school setting, 50-minute sessions are not appropriate to implement with students. This is because taking students away from their classroom work for an extended period of time could have negative implications on the student’s learning of academic skills. Furthermore, the academic schedules in the classroom are often already divided into 20- or 30-minute segments for elementary-aged students. This organization makes it simple to pull a child out of the classroom for one of these pre-determined slots. Essentially, the possibilities for treatment intensities are vast in number and exceed the two intensity conditions (i.e., 2x/week and 4x/week) that this study compares. Thus, future studies should utilize the opportunity to compare treatment outcomes for children at different intensities, including sessions of shorter duration such as 30 minutes.

It is possible that future research could shift its focus from session time to session dose when designing treatment studies. For example, this present study defined each treatment session by a time duration of 50 minutes. Many clinicians would assume that a 30-minute session would not be applicable to this study since it elicits less change than the longer, 50-minute session. It is possible, however, that a highly motivated child could

---

1975; Smit, Hand, Freilinger, Bernthal, & Bird, 1990; Templin, 1957). The English consonants have also been organized into three developmental stages: early eight: /m n j b w d p h/, middle eight: /t η k ɡ f v tj dʒ/, late eight: /ʃ z l s z θ ɵ/ (Shriberg, 1993).
produce the same number of trials in a 30-minute session as another [potentially less-cooperative] child does in a 50-minute session. This concept emphasizes the importance of defining all intensity variables for a particular treatment program. For example, it is possible that the relevance of this study could extend across session times as long as a session dose of approximately 100 trials remains consistent. Future research could examine the idea of whether or not session dose is indeed the more important variable than session duration.

**Clinical implications**

If intensive treatment continues to be supported as the preferred schedule for all children with SSD receiving any type of intervention approach, a generalized recommendation for high intensity treatment may be established. In such a case, children should be provided therapy with the most effective scheduling method, eliciting faster progress across all cases. When children are able to meet satisfactory performance for dismissal at quicker rates, the availability of services for those waiting to receive treatment will improve. In other words, more children will have an opportunity to access treatment because faster progress will elicit earlier dismissal times of those currently accessing speech services.

*Reduction in related difficulties.* In addition, improved performance following an intensive treatment block will clear up speech errors earlier in the child’s development. This means that the child will be able to catch up to their typically developing peers at a faster pace than if treated with a low-intensity, distributed schedule. There may be a variety of reasons why this would be beneficial for children. Persisting speech delays have the potential to affect learning of language and literacy when the child reaches
school-age. Several studies have indicated significant relationships between articulatory performance and reading readiness or articulatory performance and reading scores (Fitzsimons, 1958; Weaver, Furbee, & Everhart, 1960). Research has also shown that individuals with speech errors persisting through school years are much more likely to repeat a grade or receive academic tutoring during school (Felsenfeld, McGue, & Broen, 1995). Thus, avoiding lengthy periods of remediation for children with speech errors would likely decrease the chances of additional academic difficulties. Furthermore, errors that are present in a child’s speech may negatively influence their social perceptions and interactions with peers. Thus, being able to use age-appropriate speech skills could reduce the chances for bullying or negative social stigmas that are often observed in children with communication disorders (Crowe-Hall, 1991).

**Consideration of block scheduling.** SLPs may be concerned about the practicality of providing intensive treatment schedules when they must manage a large caseload. One possible suggestion to overcome this problem is to consider scheduling high-intensity treatment on a block therapy schedule rather than a continuous therapy schedule. This means that only some of the children on an SLP’s caseload will receive therapy for a given period of time while the other children on the caseload wait to access services. Once the first treatment block is complete, the clinician can cycle through to their next group of children in order to provide them with a high-intensity block of treatment as well. If the intensive treatment schedule does in fact elicit greater gains even after a maintenance period, a block treatment schedule has the potential to increase the number of children who can receive treatment. This becomes possible because dismissal from
treatment occurs earlier in a child’s development and thus, opens up space for those children waiting to access services.

In the present study, an intensive, 4x/week treatment schedule appeared to be more efficacious than a 2x/week schedule in terms of the changes made to a child’s overall phonological system following a total of twenty sessions. Given that a higher dose frequency was recently indicated as advantageous when using the multiple oppositions treatment approach with preschool children (Allen, 2013), this should not be a complete surprise to consumers of the research. More information regarding the generalizability of this finding across all intervention approaches is still necessary. At this time, optimal treatment intensity for preschool-aged children with SSD is not confirmed to be a dose frequency of 4x/week. In other words, treatment sessions for children with SSD do not necessarily need to be held at 4x/week, but the present study suggests a higher dose frequency should be strongly considered when treating preschool-aged children using the traditional treatment approach.
REFERENCES


https://doi.org/10.1044/0161-1461(2001/021)


https://doi.org/10.3109/02699200903440975

https://doi.org/10.3109/02699206.2011.601392

https://doi.org/10.1017/S0142716411000294


Movement Disorders, 16(1), 79–83. https://doi.org/10.1002/1531-8257(200101)16:1<79::AID-MDS1013>3.0.CO;2-H


