Age as a Factor in the Treatment of Late-Acquired Sounds

By

Breanna I. Krueger
Ph.D., Speech-Language-Hearing: Sciences and Disorders, 2017
M.A., University of Kansas, 2013
M.A. University of Kansas, 2011
B.A., University of Wyoming, 2007

Submitted to the graduate degree program in Speech-Language-Hearing: Sciences and Disorders and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Chair: Holly Storkel

____
Nancy Brady

____
Jonathan Brumberg

____
NavinViswanathan

____
Utako Minai

Date Defended: 9 May 2017

The dissertation committee for Breanna Krueger certifies that this is the approved version of the following dissertation:

Age as a Factor in the Treatment of Late-Acquired Sounds

Chair: Holly Storkel

Date Approved: 9 May 2017

Acknowledgements

Thank you to the families who participated in this study. Without their generosity and kindness, this study would not have been possible. I extend my heartfelt gratitude to my research assistants, Gretchen Pocisk and Mariah Coomes for their tireless work on this study and for their constant enthusiasm.

I also extend my appreciation and gratitude to my family, my children and my friends. Without your love and constant support, I would not have had the confidence and inspiration to get through my studies and to this point. No one does this alone, and my support system has been beyond anything I could ever have imagined.

Thank you to my committee whose insight and ceaseless curiosity pushed this project and me, as a scholar, to be the very best that the University of Kansas has to offer. Your individual insights into my initial question regarding this research have shaped this project into something of which I am so proud.

Finally, I want to extend my deep appreciation and gratitude to Dr. Storkel. Your support and wit and sharpness kept me consistent and pushed me beyond what I thought I was capable of. Your faultless faith in me, and the example you set as an academic, helped me realize that I can do this too, and I can do it well. I am so fortunate to have ended up on this path with you as my mentor.

Abstract

The current practice for treating speech sound disorders includes initiating treatment only after the expected age of acquisition has passed, according to developmental normative data. Unfortunately, for children who experience late-acquired sound errors, this practice misses a period of accelerated learning that occurs between the ages of 4-6 (Shriberg, Gruber, & Kwiatkowski, 1994). Instead, this practice causes therapists to initiate treatment for late-acquired sounds during a plateau of learning that occurs between the ages of 6-7 to over 8.5 years (Shriberg et al., 1994). Generally speaking, early intervention is thought to be the most effective in the treatment of developmental communication disorders. Therefore, the present study investigated whether the age of treatment contributed to the efficacy and efficiency of articulation therapy for late-acquired sounds, such as /ɹ/ or / θ /. A repeated, multiple baselines, single-subjects study investigated two age groups who misarticulated late-acquired sounds: a younger group (4-5) and an older group (7-8). These age groups capture each side of the developmental trajectory of speech sound development identified by Shriberg et al. (1994). Each child received a criterion-based, standardized two-phase articulation therapy protocol, and pre- and posttest measurements were taken using speech probes, standardized articulation tests, and acoustic analysis. These measurements allowed for an examination of **treatment efficacy** through the measurement of both ***subjective accuracy*** and ***objective accuracy*** as compared with the ages of the children. In addition, **treatment efficiency** was measured by calculating the ***number*** and ***duration*** of sessions for each phase, and overall. The results of this study indicated that younger children achieved a level of accuracy similar to that of older children as a result of treatment for late-acquired sounds, and younger children's slope of treatment accuracy over time was steeper than that of older children, suggesting that younger children were more accurate sooner than older children. For young children, a higher degree of naturalness was observed than for old children, as measured by acoustic analysis. These findings serve as an initial probe to challenge the current method of speech therapy practice for late-acquired sounds.

Table of Contents

Introduction

Late-acquired sounds are known among speech-language pathologists for being difficult to treat. Late-acquired sounds, or the "late eight" (Shriberg, 1993) include /ʃ, ʒ s, θ, ð, ɹ, z, l/, and are typically expected to develop between the ages of 5 and 8, according to normative data of typically developing children (Smit, Hand, Freilinger, Bernthal, & Bird, 1990). These sounds are difficult to acquire, and difficult to treat, because they are both motorically complex, and are more marked, in terms of conceptual representations in the phonological system.

Motor Complexity

First, these sounds require more precise articulation than early or mid-acquired sounds. Each of the late-acquired sounds are either fricatives /ʃ, ʒ s, z, θ, ð/ or liquids / ɹ, l/ rather than stops /p,t,k,b,d,g/. Fricatives require the speaker to create an incomplete closure in a precise location in the oral cavity and pass air through this semi-closure to create sound through sustained friction (Ladefoged & Maddieson, 1998; Shadle, 1990; Stevens, 1971). Liquids, or approximants, require the speaker to form the tongue into a shape that allows for the redirection of the voiced airflow either laterally, as in the case of /l/, or rhotically, in the case of /ɹ/ (Ladefoged & Maddieson, 1998). Furthermore, some sounds are accurately produced in multiple ways depending on the coarticulatory context (e.g. "bunched" vs. "retroflex" /ɹ/) (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007). In contrast, early acquired sounds are stops, nasals, and glides, which require either complete closure, followed by a full release (as in stops) or a more continuous, less obstructed airflow with a lax velar or lingual closure (as in nasals and glides) (Ladefoged & Maddieson, 1998). In short, the production of late-acquired sounds is complex, and requires a level of finesse that is difficult to acquire (Bernthal, Bankson, & Flipsen Jr, 2016). These sounds may be produced with substitutions (e.g. /w/ for /ɹ/) or with distortions (e.g. nonrhoticized /ɹ/, lateralized or dental /s/). The complexity involved in production, and the wide array of error patterns provides some explanation for why late-eight sounds are typically acquired last in children's development, and explains why they can be difficult to teach and shape in speech-language therapy. Clinicians often struggle with providing placement cues to children, as they are difficult to describe in a simplistic way for children to understand.

Conceptual Difficulty

In addition to motoric complexity, late-acquired sounds are conceptually difficult in terms of acquisition and treatment, because these sounds are more marked in the phonological system (Dinnsen & Gierut, 2008). Phonological markedness is a term used to describe low frequency sounds and sound patterns across the world's languages. In other words, these phonemes and patterns are less preferable to other phonemes and sound patterns, according to what is observed in the majority of the world's recorded phonological systems (Chomsky & Halle, 1968). In this view, the language and phonology of children is considered to be a system equivalent to other languages of the world. Sounds that occur frequently in languages of the world are considered unmarked and conceptually or linguistically easier to acquire, and sounds that occur infrequently in languages of the world are considered to be marked and conceptually or linguistically more difficult to acquire (Kager, 1999). Child phonological acquisition, in general, mirrors this pattern, where children acquire unmarked sounds or sound sequences before marked sounds or sound sequences relative to their language (Elbert, Dinnsen, & Powell, 1984).

Perceptual Salience

The difficulty of acquiring marked sounds may be due, not only to the motoric and conceptual reasons described above, but also due to the perceptual salience of the sounds, and the flexibility of the sounds to combine with other phonemes in the sound system (Blumstein &

Stevens, 1979; Stevens & Keyser, 1989). In other words, the perceptual cues associated with early-acquired sounds are more salient than late-acquired sounds due to the difference in the type of acoustic information each creates. Stop phonemes (such as /t/ or /b/) are easier to differentiate in terms of voicing distinction and place of articulation due to the categorical perception of voice-onset time cues and the phoneme offset to vowel transition for these sounds, respectively. Conversely, fricatives (such as /f/), are reliably observed just by the range of high frequency energy they elicit--timing is not a useful cue for voicing in these sounds because it can vary depending on the coarticulatory context (Jongman, Wayland, & Wong, 2000). Liquids are, acoustically, more vowel-like than other consonants in the English inventory and are therefore differentiated through the identification of the distance between the second and third formants (Espy-Wilson, 1992; Flipsen, Shriberg, Weismer, Karlsson & McSweeny, 2001; Klein, Byun, Davidson, & Grigos, 2013). These cues are acquired much like other phonological skills, and, like the production of the speech sounds, may be delayed in their development.

To summarize, late-acquired sounds are difficult to acquire because they are difficult to produce in terms of motor speech coordination, in terms of conceptualization and/or because they are linguistically more complex, and perceptually less salient. Numerous commercial products and books have been written to assist speech-language pathologists in how to treat these sounds. These include oral appliances, drill techniques, and various cueing methods (Kamhi, 2006); however, few, if any, provide any such guidance on when the sound should be treated.

Speech Sound Disorders

Some speech sound errors are expected in the typical development of children's phonological systems. Speech sound disorders (SSD), on the other hand, occur when the number or type of mistakes are atypical in relation to the child's age. The American Speech-Language

Hearing Association (2014) defines SSD as "when mistakes [of speech sound production] continue past a certain age." The "certain age" is the associated normative age of acquisition for typically developing children. Children with SSD may exhibit issues with articulation, e.g. producing speech sounds, and/or may exhibit phonological processes in which the phonological features of sounds are changed (e.g. all back stop consonants become front consonants /k/ --> /t/, while other features, such as voicing and manner, remain consistent). Regardless of the type of misarticulation pattern exhibited, the emphasis for diagnosis of SSD is on age, and whether the speech sound errors are expected or unexpected. Despite this emphasis on age, there are few guidelines about when to begin treatment for the various speech sounds outside of the developmental norms. This lack of guideline is fueled by a lack of research in the area on the influence of age in the resolution of speech sound disorders. Therefore, many clinicians follow normative data of typically developing children to guide their treatment target selection in therapy. In other words, a clinician may choose to treat the earliest acquired sounds that are out of a child's phonetic inventory, in order to make that child's phonological system more like that of his or her peers.

Treatment in Clinical Practice

The practice of using normative data for treatment of SSDs is troublesome for several reasons. First, this practice assumes that children with SSD are the same as typically developing children and should conform to typical developmental norms. The developmental approach to treatment predicts that younger children will not have sufficient foundational knowledge to acquire late-sounds, and therefore early-acquired sounds must be treated first (Rvachew & Nowak, 2001). However, previous research has shown that it is possible for children to skip these early stages of phonological development (such as acquiring all early sounds), and, in fact, they often develop the skipped early or mid sounds (and the skills required to produce them) as a result of treating later sound sequencing skills “out of order” (Gierut, 1999; Gierut & Champion, 2001; Gierut, Morrisette, Hughes, & Rowland, 1996).

Another issue with applying normative data to children with SSDs is centered around the nature of children as learners. When a young child begins speech therapy, the goal of treatment is to improve the child's intelligibility to increase the success of their communication. In a longitudinal study investigating the developmental sequence, rates and error patterns of children with SSD, Shriberg et al. (1994) found that children at ages 4-6 are in a period of rapid phonological growth. Children with SSD at older ages, on the other hand (age 6-7 and 8.5+) appeared to be in a learning plateau for new sounds. Younger children learned sounds faster in the 4-6 age range. Clinicians may observe this acceleration as well. However, once the child "sounds like" his or her peers, they may be released from therapy, because, under the definition of what a SSD is, the child no longer qualifies for therapy as having a disorder. The result is that children are released from therapy before all speech sounds are learned, often resulting in the child requiring re-entry into therapy at a later age to acquire the previously abandoned late-acquired sounds. This re-entry often occurs at the point at which Shriberg et al. (1994) observed a learning plateau: ages 6-7 and 8.5 years old.

The practice of waiting for naturalistic acquisition of phonemes is a problem for children who have errors in late-acquired sounds, such as /ɹ/ or /Ɵ/ for many reasons. One issue with waiting for therapy for late-acquired sounds relates to the broad impacts that residual speech sound errors have across other communicative domains. In terms of language development, typically developing children at ages 7 or 8 have acquired a large proportion of their lexicon. This is significant because there is a large subset of words within the vocabulary that would be misarticulated. In other words, once the child at age 7 or 8 resolves their misarticulation of a late-acquired sound, they must generalize this new motor pattern to many more words than a child who is earlier in the vocabulary development. Furthermore, as a result of entrenched motor patterns, the child may have a residual speech sound error (Lewis & Freebairn, 1992; Preston & Edwards, 2007; Preston et al., 2014). Children who have long-term residual speech sound errors have lifelong difficulties related to literacy, specifically with phonological awareness and phonological processing, that are sustained and detectable even into adulthood (Lewis & Freebairn, 1992; Preston & Edwards, 2007).

Additionally, waiting to provide therapy for late-acquired sounds may increase the likelihood that the child will experience negative social impacts. Children who have SSD often experience stigma, even at very early ages. Children who misarticulate are considered less desirable as playmates above all other differences, such as age, gender, ESL status (Rice, Hadley, & Alexander, 1993). Rice et al. (1993) found that in preschool classrooms, children who have SSD initiate with peers less, and, as a result experience reduced social contact with peers. Furthermore, Krueger, Storkel and Minai (in Review) found that preschool children identify words containing misarticulations more slowly than canonical productions, which could also impact the peer groups' ability to understand and communicate further with a child who misarticulates. Older children who misarticulate may be subjected to bullying from peers, and if not resolved, may experience issues with employment, higher education and social interaction as they age into adulthood.

Perhaps the most damning result of waiting for naturalistic acquisition of late-acquired speech sounds is that doing so completely misses a potential period for accelerated phonological learning that occurs during the preschool years (ages 4-6) (Shriberg et al., 1994). In other words, this is a stage at which research has shown that children are primed to learn new sounds and alter their sound systems. Previous studies have not explored age as a factor in the success of therapy for late-acquired sounds at early ages. It is likely, however, that the developmental advantage for learning sounds at earlier ages (e.g. 4-6), may improve the outcomes for children who struggle with learning late-acquired sounds by curtailing the effects of residual speech sound errors observed in older children. Therefore, it is possible that therapy should occur during younger ages for the most complex (marked) or most late-acquired sounds when all other points are equal, such as percentile on standardized measures of articulation. This would allow for children to experience the greatest amount of growth over the course of their therapy. Furthermore, by placing the location of this treatment between the ages of 4-6, when accelerated learning occurs, there may be the amplest point of success for these children, and fewer sessions will be needed overall. If accelerated learning is a faculty that children in this age group have, then taking advantage of it by closing the gap as much as possible is intuitive and should be considered.

Phonological learning and treatment

The current practice for treating speech sound errors typically involves initiating the treatment of speech sounds at a point when the speech errors are not consistent with what is expected according to developmental normative data of typically developing children. Therefore, if a child is producing errors on many sounds at the age of 4, clinicians will provide speech therapy for the early-acquired sounds /k/ or /f/, rather than treating later-acquired sounds such as /s/ or /Ɵ/. In fact, the American Speech-Language and Hearing Association identifies a speech sound disorder as, "...when mistakes continue past a certain age. Every sound has a different range of ages when the child should make the sound correctly," and concludes that, "By the age of 8, children should be able to produce all sounds in English correctly" (American Speech-

Language Hearing Association, 2008). The assumption underlying this statement implies that, in clinical practice, therapists should wait until the expected age of acquisition for speech sounds before initiating therapy on these sounds. Indeed, waiting may be appropriate if a child is only producing errors on /ɹ/--because this error may be developmental in nature, rather than the result of a speech sound disorder. However, in children who produce multiple errors inconsistent with developmental normative data and, therefore, qualify for speech therapy at younger ages, this statement could imply that therapists should treat the age appropriate sounds, and then wait until the expected age of acquisition of late sounds before treating them. The practice of waiting in this manner incorrectly treats children with SSD as typically developing children by holding them to the normative data of typically developing peers, and assumes that, like typically developing peers, late sounds will "come in on their own." In reality, children with SSD have difficulty learning sounds—this is the cause of the delay of acquisition (American Speech-Language-Hearing Association, 2014). Therefore, it is unrealistic to assume that they will readily acquire late-sounds--again, the most difficult in the English phonology--without assistance from a speech-language pathologist.

Overall, previous research focused on which sounds to treat and how to treat them (Dean, Howell, Waters, & Reid, 1995; Gierut, 1998; Hesketh, Adams, Nightingale, & Hall, 2000; Rvachew, Rafaat, & Martin, 1999; Tyler, Edwards, & Saxman, 1987; Weiner, 1981) rather than when to treat them. In terms of which sounds to treat, there is controversy regarding whether developmental appropriate (i.e., early acquired) sounds or more advanced and complex (i.e., late-acquired) sounds should be prioritized in treatment.

On the one hand, research has shown that treating late sounds provides children with broader generalization to untreated sounds than treating early acquired sounds (Gierut et al., 1996). This method is known as the complexity approach to treatment target selection, where treating complex sounds leads to "filling in" of untreated, less complex sounds. Gierut et al. (1996) conducted two studies to investigate the influence of late-acquired and early-acquired sounds on treatment effectiveness. The first study was a within subjects, alternating treatment design in which children received treatment for an age-expected, late-acquired sound, and an early-acquired sound. The second study was a between subjects design, in which one group were treated with an early-acquired sound, and another group were treated with a late-acquired sound (relative to the child's chronological age) (Gierut et al., 1996). Children were provided with a criterion-based treatment where an established criterion of accuracy or a maximum number of sessions was required before moving on to the next phase of treatment (Gierut et al., 1996). The results of these studies showed that children who received treatment for late-acquired sounds responded to treatment at a more rapid rate than those receiving treatment for early-acquired sounds (Gierut et al., 1996). Additionally, treatment for late-acquired sounds led to broader change across the sound system (to untreated sounds) than treatment for early-acquired sounds (Gierut et al., 1996). The results of these studies suggest that “skipping” developmental phases is not only possible, but leads to the ultimate goal of accurate speech sound production faster by introducing broader change through complex treatment target selection.

On the other hand, proponents of the developmental approach advocate for meeting each developmental milestone to ensure that children have an adequate scaffold upon which to build their phonological learning (Rvachew & Nowak, 2001). *The effect of target-selection strategy on phonological learning* treated two groups of children on either early-acquired sounds or late-acquired sounds in two phases. Each block provided one session of treatment per week for six weeks, followed by an assessment and another six sessions of treatment (Rvachew & Nowak, 2001). Treatment targets were selected along two dimensions: phonological knowledge and age of acquisition. Targets were selected to be either most phonological knowledge/early acquired (ME), or least phonological knowledge/late-acquired (LL) (Rvachew & Nowak, 2001). The authors of this study suggested that children with ME sounds completed more steps in a treatment hierarchy than the LL group, but change in accuracy of the treated sounds in untreated words and overall accuracy of the sound system did not differ between the groups. These results contrast with those found by Gierut et al. (1996).

In a letter to the editor in response to this contrast of findings, Morrisette and Gierut (2003) discussed the findings of Rvachew & Nowak (2001) in the context of their own findings. First, Morrisette and Gierut (2003) suggest that the findings of Rvachew & Nowak (2001) are not necessarily in contrast to their own findings because of the nature of the treatment targets selected. Recall that Rvachew & Nowak (2001) incorporated phonological knowledge into their treatment target selection process. Gierut et al. (1996) selected early and late treatment targets based on children's least productive phonological knowledge. To put this into Rvachew & Nowak's (2001) framework of treatment target selection, Gierut et al.'s (1996) treatment targets were least phonological knowledge/early-acquired and least phonological knowledge/late-acquired. Morrisette and Gierut (2003) argued that the results of Rvachew & Nowak (2001) provide evidence for the optimal condition (ME) to move children through treatment steps at a rapid rate. If this is the goal of treatment, then ME targets would be a logical selection. However, if broader generalization to untreated sounds is the goal of therapy, Gierut et al.'s (1996) results conclude that least phonological knowledge/late-acquired is the more appropriate target. In short, neither study necessarily negates the results of the other, but the goal of therapy should be considered to determine what type of treatment target is selected. Regardless of the demonstrated efficacy of a non-traditional approach to treatment target selection, a survey of speech-language pathologists working with children aged 3-6 found that therapists typically use a traditional approach to treatment (Brumbaugh & Smit, 2013).

Purpose of the Present Study

The purpose of the present study was to investigate age as a factor in the treatment of late-acquired sounds by providing speech therapy to younger children aged 4-5 and older children aged 7-8. The present study intended to address the conflict between the two approaches by using a different tactic altogether--examining the role of age in the treatment of these sounds. In other words, the two approaches were pitted against one another on a dimension that has not yet been considered in clinical research: *when* to treat late-acquired sounds.

The developmental approach predicts that younger children will not have a foundation for learning late-acquired sounds, and therefore late-acquired sounds would be unsuccessful treatment targets for this group. Under the developmental approach to treatment target selection, young children who are taught late-acquired sounds should require more sessions and longer sessions to achieve accuracy, and, ultimately, they should not achieve a high level of accuracy because their system is not ready to learn these complex, later developing sounds. In this view, early-acquired sounds must be taught first to set the stage for learning relatively more complex late-acquired sounds (Rvachew & Nowak, 2001).

In contrast, the complexity approach predicts that younger children will learn late-acquired sounds at least as well as older children, because foundational knowledge is not a necessary prerequisite to learn complex late-acquired sounds in the context of a supportive clinical treatment (Gierut et al., 1996). The complexity approach is centered around the idea that teaching more difficult items (e.g. consonant clusters, late sounds, etc) promotes the development of less difficult items (e.g. singletons). Thus, it is possible that complex sounds could be treated effectively at any age, and, in fact, may lead to broader generalization to untreated sounds than the developmental approach—particularly in younger children who tend to have more errors. Under the complexity view of treatment target selection, young children should learn late-acquired sounds in fewer sessions and shorter sessions to achieve accuracy, and will experience a higher level of accuracy and generalization because their sound system is primed for learning new sounds.

The present study expanded the findings of previous research in late vs. early acquired sounds by exploring the boundaries of children's learning of late-acquired sounds within the context of the child's age. Previous research varied age of acquisition according to the child's age, or within the limits of expected sounds for the child's age (Gierut et al., 1996; Rvachew & Nowak, 2001). However, these researchers did not explore the impact that learning late-acquired sounds during a period of accelerated or plateaued learning may have on the success in treatment. Previous research relied solely on the acquisition of treated and untreated sounds to determine treatment efficacy; however, these points alone exclude another point of interest that may inform clinical judgment on whether treating late-acquired sounds is effective. The present study examined treatment efficacy as a function of children's age. In other words, not only were children's accurate productions measured, but the amount of change over time was measured and analyzed to give a measure of treatment efficiency. After all, previous research has shown that children can learn late-acquired sounds (Gierut et al., 1996). What the present study seeks to know is whether younger children did it faster or better. The former was addressed by measuring number of sessions. The latter was addressed by confirming clinician perceptual judgment with acoustic data.

Research Questions

The results of the present study address the following questions:

1) Do younger children acquire late-acquired sounds better than older children in terms of perceptual and acoustic accuracy?
2) Do younger children learn sounds faster (in fewer sessions) and with less support (in shorter sessions) than older children?

The first research questions addressed treatment efficacy, that is, whether treatment for late-acquired sounds was more successful for younger children than older children. Based on the previous discussion regarding the findings of Shriberg et al. (1994), it is predicted that younger children will learn late-acquired sounds more accurately than older children due to their age-associated readiness to change their phonological systems. The second research question addresses treatment efficiency, that is, how efficient treatment was for each age group in terms of number of sessions required, and the average duration of sessions overall. It is predicted that younger children require fewer sessions than older children because they are in a period of accelerated learning, but also because they have fewer entrenched motor speech patterns, as a result of being younger, and fewer incorrect abstract representations for words with their respective late-acquired speech sound errors as a result of knowing fewer words.

Methods

Participants

In order to qualify for the study, participants were required to meet inclusionary criteria. In order to be admitted to the treatment portion of the study, participants were required to score within normal limits on a series of standardized assessments. First, the Listening Index on the

Test of Language Development (Primary), 4th Edition (Newcomer & Hammill, 2002) tested children's receptive language abilities through a combination of three subtests that investigated children's understanding of vocabulary and grammatical aspects of language. Children were required to score within normal limits on this index. Second, children were required to pass a hearing screening at 20dB for 1000, 2000 and 4000 Hz (American Speech-Language-Hearing Association, 1996). Third, participants were required to score within normal limits on a test of nonverbal intelligence as measured by the nonverbal index of the RIAS (Reynolds & Kamphaus, 2003). This test required children to point to an "odd item out" from an array of 6 items, and to indicate what piece or part was missing from typical real-world objects (e.g. a cow missing an ear). Participants' errors were required to be rated as "linguistic" in nature (not motoric) as measured by the Kaufman Speech Praxis Test (Kaufman, 1995). This test required children to demonstrate maintenance of their errors throughout a variety of tasks, such as maintaining the same speech error as the context increased in complexity from syllable (e.g. /ka/) to single syllable word (e.g. "cup") to multisyllabic words (e.g. "cupboard"). Maintenance of these same errors regardless of phonological complexity of the task suggested that children's errors were "linguistic" rather than motorically based (Note: Children with motor speech disorders are expected to be accurate at lower levels of phonological complexity and less accurate as complexity increases).

In terms of inclusionary criteria for articulation, participants were required to score below the 10th percentile on the Goldman-Fristoe Test of Articulation-2 (GFTA-2) (Goldman, 1986). This test examined children's production of sounds in words in a single vowel context. Then these results are compared to age and gender matched normative data. Since this assessment is age and gender-normed, the total number of sounds in error varied from child to child due to age differences. Participants were also required to have either /ɹ/ or /Ɵ/ in error, as these were the late-acquired sounds targeted for treatment. Since the GFTA-2 only elicits sounds in one vowel context for each position, a more in-depth analysis of children's phonological inventory was conducted. The Phonological Knowledge Protocol was administered (Dinnsen & Gierut, 2008). This speech probe is made up of picturable words, that allows the examiner to elicit each sound in 5 vowel contexts for initial position, 7 vowel contexts for medial position, and 5 vowel contexts for final position for a total of 17 productions for each sound. Children were required to produce /ɹ/ and /or /Ɵ/ with an accuracy of 0-7% (i.e., 0-1 correct productions out of 17 items).

Fifteen monolingual children were recruited through word-of-mouth contact and social media announcements (e.g., Facebook). Participants were divided into two age groups: 4 years, 0 months to 5 years, 11 months (the "young" group) and 6 years, 11 months and 8 years, 11 months (the "old" group). These age groups were selected because the younger group (4-5) is at an age of accelerated phonological learning, while the older group (7-8) are in the midst of a learning plateau (Shriberg et al., 1994). Two participants withdrew due to scheduling issues, 2 were excluded for accurately producing the target treatment sound on an initial probe, 4 were excluded for scoring above the 10th percentile on the articulation test (GFTA-2), and 1 was excluded for scoring below the 10th percentile on the receptive language test (TOLD-P4). Therefore, 6 participants received treatment. Of these, two were male, and four were female. All children were reported to be white, nonhispanic. Each subject's demographic and testing data is displayed in Table 1 below:

Table 1: Participant's demographic data, scores on inclusionary tests, baseline condition, and treatment target.

Subject	Age	Treatment Target	Baseline Condition	GFTA-2 %ile rank (Standard score)	RIAS %ile rank (NIX*)	TOLD-P4 RLI %ile rank (Index

						score)
Old 1	8 y 10 m	/ɹ/	3	2 (79)	47 (99)	35 (94)
Old 2	8 y 1 m	/ɹ/	4	1 (50)	70 (108)	70 (105)
Old 3	6 y 11 m	/θ/	5	4 (77)	95(124)	50 (100)
Young 1	4 y 11 m	/ɹ/	3	8 (79)	77 (111)	55 (102)
Young 2	4 y 10 m	/ɹ/	4	5 (68)	34 (94)	77 (111)
Young 3	4 y 0 m	/θ/	5	9 (65)	77 (111)	21 (88)

*Nonverbal Intelligence Index Score

Single Subjects Design

The present study was a single-subjects design. This methodology was chosen because it allowed for a more intensive and comprehensive examination of within-subjects effects to determine whether a relationship exists between age and the variables of treatment efficacy and treatment efficiency.

A total of six participants who met inclusionary criteria were enrolled: three in each age group. The first participant served as the test child. The next enrolled participant served as a within-group control. These participants were treated with the same sound (e.g. /ɹ/). The third participant served as a measure of systematic replication of the treatment within the same age group and was treated with an additional late-acquired sound (/θ/). These treatment sounds were selected because /r/ is a sound that is frequently in error among both young and old children and typically produced with a substitution, /w/, rather than a distortion. The /θ/ sound was selected because it is also typically produced with a speech sound substitution and is in a different manner and placement class than /r/. With this design, the external validity of the design was increased by providing a within-group replication of the same treatment for the same sound (/ɹ/) to demonstrate that different children show similar outcomes, and a systematic replication of the same treatment for a different sound (/θ/) to demonstrate that treatment of different late-acquired sounds produce similar outcomes.

The study employed a repeated multiple baselines design. This methodology ensures that children's progress in treatment is a result of the therapy they received, and not due to maturational effects. As each participant was enrolled, the number of baseline measurements increased by one to ensure that treatment, not concentrated exposure to target sound probes was the cause of treatment (see Table 2). For example, the first child in each group was enrolled in "Baseline Condition 3." The child's first session was a full probe of his or her phonetic and phonological inventories from the PKP, which elicits all sounds in 17 vowel contexts and positions: 5 initial, 7 medial, 5 final. If the child's accuracy for their treatment target (e.g. /ɹ/ or /Ɵ/) on this probe was 0%-7%, then the next two sessions probed these sounds using the 17 vowel contexts and positions from the PKP. Accuracy on these probes was required to be 0%-7% for the treatment targets before treatment was initiated.

Table 2: Treatment design by baseline condition and session

Baseline Condition	**Session 1**	**Session 2**	**Session 3**	**Session 4**	**Session 5**	**Session 6**	**Session 7**
3 (treatment /ɹ/)	Full Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment	Treatment	Treatment
4 (treatment /ɹ/)	Full Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment	Treatment

5 (treatment /Ɵ/	Full Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment Sound Probe	Treatment

Materials

Hardware and software

Children received treatment two to three times per week in their school setting (n=5) or in their home (n=1). Each session was video and audio recorded using a JVC handycam and a Zoom tabletop audio recorder, respectively. The treatment story was presented on a Microsoft Surface Pro 3 tablet using Microsoft Powerpoint software. The treatment pictures were an adapted story by Tomie de Paola, entitled *Bill and Pete* (DePaola, 1996), as used in Gierut et al. (2010). The names of items and verbs throughout the story were substituted with the treatment nonwords.

Nonword Stimuli

Nonwords were developed from a nonword database (Storkel, 2013) that provided information about the phonotactic probability and neighborhood density of nonwords. These factors are relevant because research has shown that children are more accurate in learning low probability sound sequences than common sequences, and that sparse neighborhood density yields more accurate word learning than dense, although this effect improved in retention (Storkel & Lee, 2011). Furthermore, children have greater difficulty repeating nonwords with low probability sequences than high probability sequences (Munson, Edwards, & Beckman, 2005). Therefore, these factors have may influence how children acquire and produce the nonwords to be used in treatment.

Table 3: Items in the story and the nonwords they replaced in each condition.

Target	/ɹ/ nonword	/θ/ nonword
Baby Crocodile Name	/ɹib/	/θɪp/
Toothbrush name	/ɹɔɪm/	/θʌp/
Cage	/ɹɛb/	/θɪm/
Kidnap	/ɹʊp/	/θeɪp/
Suitcases	/ɹad/	/θoʊn/
Pick	/ɹaʊn/	/θaʊm/
Climb	/ɹʌp/	/θæm/
Jump	/ɹʊd/	/θoʊb/

Eight CVC nonwords for each treatment sound (/ɹ/ and /θ/) were selected from the database, for a total of 16 (see Table 3). The final consonant of these words were controlled to be early-acquired sounds /p,b,m,n,t,d/ to increase the likelihood that children had acquired these sounds, so that treatment was focused on the treatment acquisition of the initial sound only. The phonotactic probability of each nonword was assessed by examining the biphone probability of the CV sequence, and the biphone probability of the VC sequence for each word. For /ɹ/ each possible combination was replicated twice (high/high, high/low, low/high, low/low). The phonological neighborhood density distribution of these words was 6 dense, 2 sparse. This is due to the relative commonality of /ɹ/ in multiple vowel contexts and the restriction of the final consonant as early-acquired. For /θ/, every possible biphone probability combination could not be achieved without limiting the vowel type. According to the database, /θɪ/ is the only CV onset that is high probability. Therefore, there are only two words that are high/high, and high/low sequences. The distribution of biphone probabilities for these words is 2 high/high, 4 low/high, and 2 low/low (see Appendix C). This allowed for diversity among the vowels and control for the final consonant as early-acquired. The neighborhood density distribution for these words is 4 dense, 4 sparse. All 16 words were balanced for adequate diversity in terms of phonological features (voicing, place, vowel height, etc.) among the vowels and final consonants, while maintaining the aforementioned controls. The differences among these sets is not of concern in terms of experimental control because the grouping variable of interest is age, and both age groups received each set of words. Furthermore, since the /θ/ treatment condition is a systematic replication, differentiation among the words is beneficial.

Treatment Procedures

Once baseline was established and children met inclusionary criteria, the treatment protocol was initiated. The treatment portion of the study was conducted in two phases (see Table 4).

Table 4: Treatment words and phase procedures.

	Phase 1	Phase 2
Stimuli	8 nonwords	8 nonwords
Dose	10 imitated repetitions per item per session	10 spontaneous elicitations per item per session

Dose Frequency/ Duration of Phase	14 sessions or 75% accuracy across 2 consecutive sessions	24 sessions or 90% accuracy across 3 consecutive sessions

The first segment was an imitation phase. In this phase, children were read the treatment story containing the targeted nonwords (see Table 3). The story served to familiarize the participants with the nonwords used in treatment. After the story was read, each child was shown each treatment word on a Powerpoint slide. The examiner presented one slide and provided an imitative prompt (e.g. "Say [treatment word]"). The child then imitated the word, and the examiner provided relevant feedback. For example, if the child obtained a correct production, the examiner provided relevant positive feedback such as "Good job! You raised your tongue up!" or "You got it! You stuck your tongue out and blew!" If the child did not get the word correct, the examiner said what articulatory gesture was incorrect and provided articulatory instruction. The child would then have a second attempt at the word, and would again receive instructional feedback, but only first attempts were scored and counted to evaluate treatment progress. The slides contained each of the 8 treatment words. The examiner cycled through the slides 10 times per session, so the child produced a total of 80 scored trials each session. After the child produced 10 scored trials of each word, the session ended. The examiner judged each initial production attempt on-line during the session, and a second judge listened to audio files and independently scored the sessions to verify the examiner's scoring (see Appendix A). Once the child reached 75% accuracy across two consecutive sessions, or 14 total sessions, then the imitation phase was terminated, and the child moved to the spontaneous production phase (see Table 4). Throughout the treatment, treatment sounds were probed for generalization to real words using 5 probe words taken from the PKP: 2 in initial position, 1 in medial position, 2 in final position. These probes were administered at the end of every third session (for treatment schedule, see Appendix B) The accuracy on these probes allowed for the incremental tracking of children's acquisition of the treated sounds. The full PKP was administered upon completion of the imitation phase to examine the changes to treatment targets that occurred as a result of this first phase of treatment.

The spontaneous production phase also included reading the story and single picture naming practice via Powerpoint slides. However, rather than providing a verbal model for the child to imitate, children were asked "What's this?" by the examiner. If the child said "I don't know" or a similar response, the examiner responded by providing the child with a cloze option, such as "This is the baby crocodile named ....". If the child failed to respond with the treatment word, the examiner elicited a delayed imitation by giving the child choices, "Is it a [treatment word] or a table?" If the child still did not respond correctly, direct imitation was used "Say [treatment word]." These responses were scored. Then the child continued to the next trial. Feedback was provided in the same way as in the imitation phase for treatment. Criterion for the spontaneous treatment phase was 90% accuracy across three consecutive sessions or 24 sessions maximum (see Table 4).

After the completion of treatment, children underwent immediate posttesting at their next regularly scheduled session. The post testing was a repeated administration of the full PKP. A repeated posttest occurred 2 weeks after treatment ceased to determine if any observed growth was maintained.

Reliability and Fidelity

In terms of reliability, children's accuracy on treatment target real words, treatment accuracy on treated words, and procedural reliability were measured. Treatment target words (/ɹ/

or /θ/) from the PKP at pretest, phase-shift and posttests were transcribed from audio recordings of the sessions by the author and a trained second scorer using International Phonetic Association transcription conventions. Children's productions of treatment words and real word probes in the therapy sessions were rated on-line by the session administrator, and 30% of sessions were scored by a second rater. Any disagreements were discussed and the raters listened to the token again to reach consensus. If consensus was not reached, a third judge was consulted to provide a final answer. In addition, 30% of treatment videos were checked for fidelity to the prescribed procedures for each session. This fidelity measure was taken on a scoresheet that mirrored the scripts used in the therapy sessions. Reliability judges evaluated the number of correct and incorrect responses by the child, the overall percentage of accuracy, and duration of the session. For acoustic measures, a second scorer, trained in acoustic measurement, took a second measurement of target productions for pre- and posttest measures and 30% of treatment sessions.

For scoring reliability, session scores were rated by a second trained rater as 93% in accordance with the examiner's scores across 30% of sessions for each child. On the 7% of sessions where agreement was not reached, sessions were listened to again to obtain consensus on scoring. For treatment fidelity to the treatment procedures, the rater found 100% fidelity to prescribed treatment procedures in terms of providing the correct script for treatment, the correct picture stimuli and the correct number of exposures to treatment nonwords.

Outcome Measures

Treatment Effectiveness

In order to measure treatment effectiveness, we measured accuracy in two ways: accuracy on untreated real words, and accuracy on treatment words. In the present study, both *p*-values and effect sizes are considered in the discussion because it is possible that some analyses yield nonsignificant results, but have large effect sizes. In this case, a large effect size, but lack of statistical significance suggested that the statistical power was not sufficient enough to produce a significant result, but, with more subjects, a significant result may be found. Therefore, these results are interpreted as promising avenues for future research.

Accuracy of Untreated Real Words. For accuracy on untreated real words, each word containing the treatment target was scored as a proportion correct at pretest PKP and immediate posttest PKP. Post-treatment accuracy was compared to pre-treatment accuracy using the Wilcoxon Ranked Signs test, and a Pearson *r* effect size estimate (Cohen's conventions for *r*: 0.1= "small", 0.3= "medium", 0.5 = "large") to determine whether significant learning occurred during treatment. This was done for all children together and then for each group separately to determine whether both groups made significant change during treatment. Remaining measures and analyses focused more on differences between the Young and Old group. The rate of change in terms of generalization to untreated sounds was found by plotting accuracy on PKP probes at pretest, phase-shift and immediate posttest over time in days to determine the slope of treatment generalization to untreated real words. The slope values for each child were compared using a Mann-Whitney Test, and Pearson *r* effect size estimate to determine whether there were differences between the Young and Old group in terms of accuracy over time.

Acoustic measures were taken for each child's correct production of untreated real words to determine whether participants' accurate productions were acoustically within the expected range for a natural production of the targeted treatment sound. Acoustic measures were taken by using Praat acoustic analysis software package. For /ɹ/, the difference between the second and third formants was obtained by querying the formant listing overall and selecting the lowest point of F3 for measurement of F2 and F3 in accordance with (Klein et al., 2013). In general, this difference should be smaller than what is expected for the typical substitute for /ɹ/, which is /w/ because lip rounding lowers the F2 formant significantly. The expected F3-F2 value for /ɹ/ was between 500 and 1500 Hz (Klein et al., 2013). In terms of /θ/, children's productions were compared by measuring the relative amplitude of F5 and obtaining the spectral peak location as described in (Hedrick & Ohde, 1993; Jongman et al., 2000). The center of the fricative was extracted using a 40 ms Hamming window around the center point to a Fast Fourier Transform (FFT) spectrum and linear predictive coding spectrum (LPC). The amplitude at F5 for the fricative was obtained from the FFT and LPC. Then, the same procedure was used to extract F5 from the onset of the following vowel. This amplitude was subtracted from the amplitude of the fricative to obtain the relative amplitude of F5 at the fricative. The spectral peak of the fricative was obtained by filtering the extracted 40ms fricative with a 80 Hz pre-emphasis filter and sending it to an long-term average spectrum (LTAS) with a 250 Hz bandwidth. Then the peak amplitude value was queried from the software. The expected value for /θ/ is ~ -12.5 dB relative amplitude of F5, and >7000 Hz for spectral peak location (Jongman et al., 2000).

Accuracy on Treatment Words. For accuracy on treated words, accuracy was plotted over time and a line of best fit was obtained to acquire a slope value for each child. Slopes for all treatment sessions combined, for the imitation phase, and for the spontaneous phase were analyzed and compared between groups using a Mann-Whitney U test, and Pearson's *r* for an effect size estimate. Cohen's conventions for *r* were, again, used to compare the effect size between groups.

Treatment Efficiency

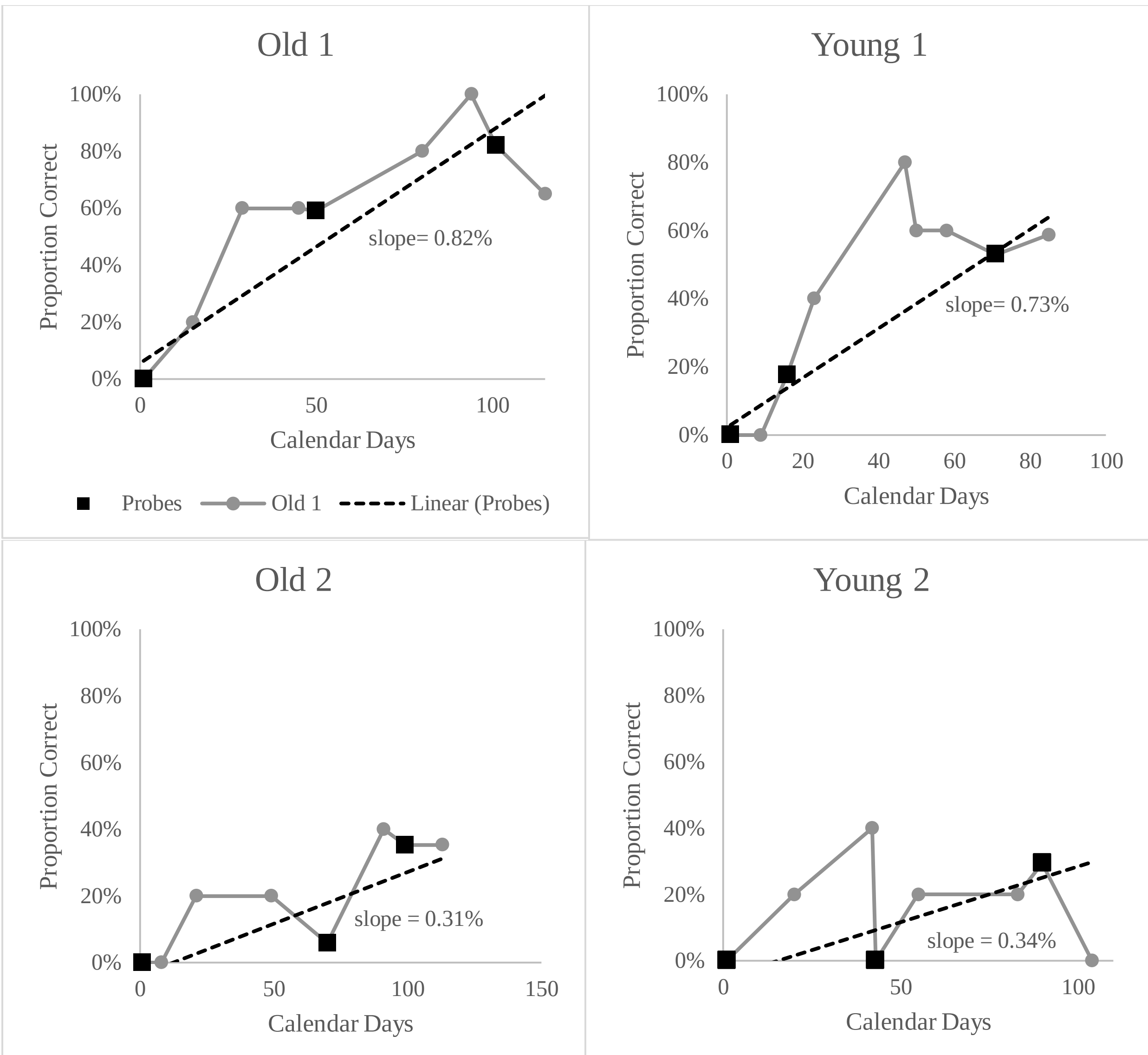
To measure treatment efficiency, the total number of sessions and duration of sessions were measured for each child. For total number of sessions, three values were obtained for each child: number of sessions overall, number of sessions for the Imitation Phase, number of sessions for the Spontaneous Phase. For length of sessions, three values were obtained for each child: mean length of all sessions, mean length of sessions in the Imitation Phase, mean length of sessions in the Spontaneous Phase. These individual values for each child were compared between groups using a Mann-Whitney test and Pearson *r* effect size estimate. Again, *p*-values and effect sizes were interpreted as indicators of differences or potential differences, respectively, between groups.

Results

Treatment Effectiveness

Accuracy of Untreated Real Words. In terms of comparing accuracy at pretest to immediate posttest, all children significantly improved their production of the treated sound in untreated real words, and the effect size was large, suggesting a lack of statistical power [Z=-2.20, *p*=0.03, *r*=0.64]. Within the Young group, children's scores were not statistically significant from pretest to posttest, but the effect size was large, suggesting a lack of statistical power [Z=-1.60, *p*=0.11, *r*=0.65]. This finding indicated that significant learning occurred as a result of treatment. Within the Old group, the children's scores were not statistically significant from pretest to posttest, but the effect size was large [Z=-1.604, *p*=0.11, *r*=0.65]. Based on the effect sizes, both young and old children appeared to show appreciable generalization of the treated sound to untreated words.

Each child's line of fit, and slope are shown in Figure 1. The gray points indicate session probe scores, while the black squares indicate accuracy on full probes at Pretest, Phase-shift and Immediate Posttest, respectively. The black dashed line is the line of best fit computed for the full probe values, and the slope is labeled below each fit line. Children's slopes were not significantly different between groups and the effect size was medium [U=3.00, Z=-0.66, p=0.51, r=0.26]. Through a visual inspection of the figure, the performance is varied. Old 1, Young 1 and Young 3 showed steep slopes, while Old 2, Old 3 and Young 2 showed flatter slopes. Overall, children's rate of generalization of the treated sound to untreated words seemed to be variable with no obvious relationship to age group.



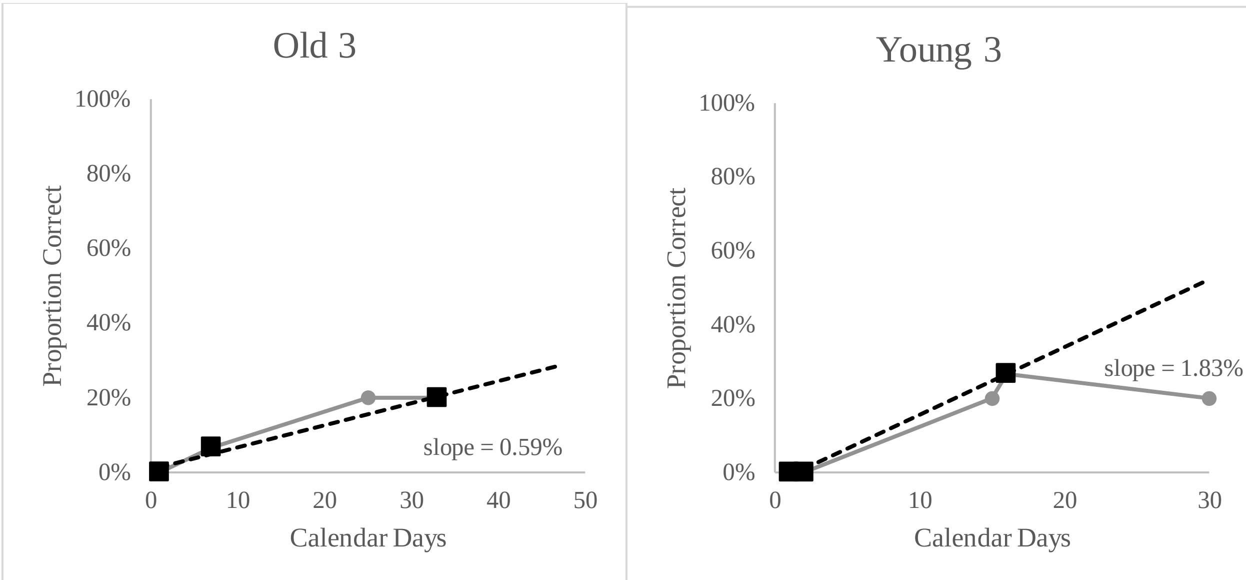


Figure 1: Figure of probe accuracy and slope by participant. "Probes" points indicate Baseline, Phase-shift and Immediate Post test results, respectively. Line points indicate children's scores on baseline and full probes together.

Acoustic Measures. Children's correct productions at immediate posttest were analyzed acoustically, and are outlined in Table 5. The (*) in the table indicates that the score was within the expected range. As seen in Table 5, all young children achieved mean acoustic values within the expected range for their productions and the majority of their productions (60%-100%) were within the expected range. Only Old 2 obtained mean acoustic values within the expected range for their treatment sound and all productions were within the expected range. Old 1 and 3 did not obtain mean acoustic values within the expected range and only a few individual productions (21-33%) were within the expected range. These results suggest that, although each of these analyzed productions were rated as perceptually accurate, only the Young group consistently obtained acoustic values that indicate consistently natural productions of the treatment sound.

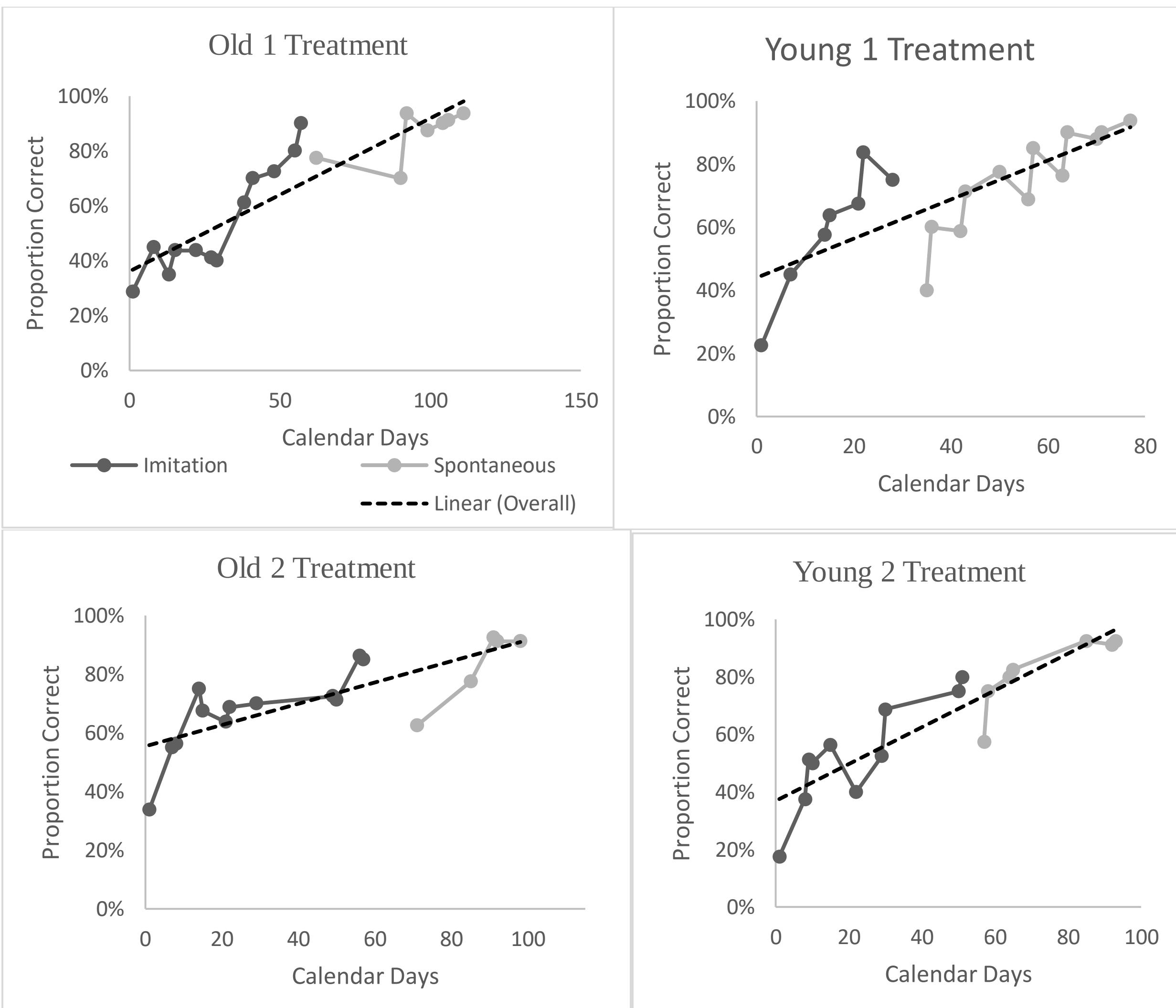
Table 5: Acoustic values of correct productions at immediate posttest by subject.

Subject /ɹ/	F3-F2	Subject /θ/	Relative Amplitude (dB)	Spectral Peak Location (Hz)
Old 1	*n*=14 *M*=1636.14 Hz Range=798-2386 % WER= 21.42%	**Old 3**	*n*=3 *M*=-1.53 dB Range= -13.2-8.6 %WER=33.33%	*n*=3 *M*=3625 Hz Range=375-10125 %WER=33.33%
Old 2	*n*=5 *M*=1282* Range=1236-1510 Hz %WER=100%			
Young 1	*n*=8 *M*=903.13 Hz* Range=383-1691 Hz %WER= 87.5%	**Young 3**	*n*=3 *M*=3.63 dB Range= 0-9 dB %WER=0%	*n*=3 *M*=10500 Hz* Range=9375-11875 Hz %WER=100%
Young 2	*n*=5 *M*=1317 Hz* Range=838-1727 Hz %WER= 60%			

*Within expected range

Note: Expected values for /ɹ/ F3-F2 is between 500 and 1500 Hz based on (Klein et al., 2013). Expected values for the /θ/ relative amplitude are -12.5 dB and spectral peak location is expected to be >7000 Hz (Jongman et al., 2000). Each child's number of tokens measured (n), mean value (M), range, and proportion of productions within the expected range (%WER) are noted for each subject .

Accuracy on Treatment Words. Treatment charts for each child are shown in Figure 2, where the dark points are scores in the Imitation Phase, light points are scores in the Spontaneous Phase, and the black dashed line is the slope for the line of best fit. Across both phases combined, as seen in Figure 2, the slope for the Young group (*M*=1.07) was steeper than that of the Old group (*M*=0.37), suggesting that younger children had steeper learning of the treated sound in treated words than older children. The difference of the slopes was marginally significant, and the effect size was large [U<0.001, Z=-1.96, *p* = 0.05, *r*=0.8].



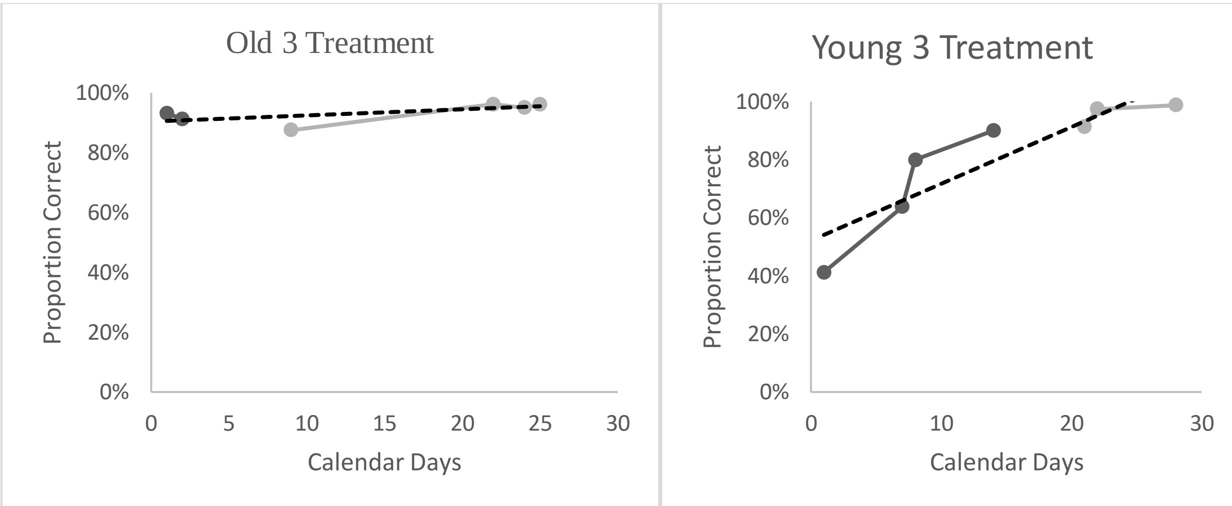


Figure 2: Treatment accuracy by subject. Data points are actual scores by participants, dashed lines are trend lines for each phase and across both phases.

In the Imitation Phase, Young children's slopes were numerically steeper (*M*=2.27) (see Table 6) than those of the Old group (*M*=-0.15). This was not a statistically significant difference, but the effect size was large, suggesting the analysis was underpowered [U=1.00, Z=-1.53, *p*=0.13, *r*=0.62]. Descriptively, Young children demonstrated steeper learning in treatment than the Old children.

In the Spontaneous Phase, children's slopes were varied. Children in the Old group had a slope of 0.82 on average, and the Young group had a slope of 0.70 on average. No significant difference between groups was observed and the effect size was medium [U=3.00, Z=-0.65, *p*=0.51, *r* = 0.26]. These values (Table 6) suggest that children's performance in treatment was varied in the Spontaneous Phase, where Old 1, Old 3 had flatter slopes than the other children, Young 2 and Young 3 had steeper slopes, and Old 2 and Young 1 had the steepest slopes of all. Taken together, younger children showed steeper learning than older children over the course of treatment and this was more pronounced in the Imitation than the Spontaneous Phase.

Table 6: Slope of each participant's treatment effectiveness scores over time.

Participant	Slope Overall Treatment Words*	Slope Imitation Phase	Slope Spontaneous Phase
Old 1	0.56	0.99	0.35
Old 2	0.36	0.55	1.21
Old 3	0.20	-2	0.55
Mean	0.37	-0.15	0.70
Young 1	0.62	2.06	1.01
Young 2	0.64	0.92	0.67
Young 3	1.95	3.82	0.77
Mean	1.07	2.27	0.82

*Significant difference between groups

Treatment Efficiency

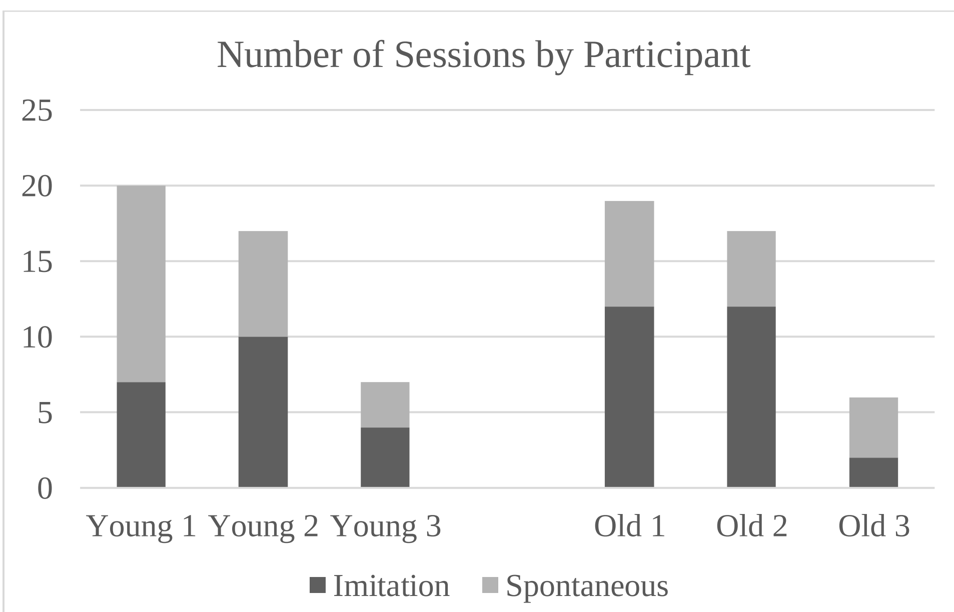


Figure 3: Number of treatment sessions required for each phase and for each subject.

Number of Sessions. Individual number of sessions are in Figure 3 above. Overall, younger children did not require fewer sessions than older children, and the effect size was small [U=3.500 Z=-0.44, *p*=0.66, *r*=0.18]. Overall, the Young group required 14.66 sessions on average, whereas the Old group required 14 sessions on average.

Within the imitation treatment phase, children in the Young group required 7 sessions on average, whereas the Old group required 8.67 sessions on average. This difference was not significantly different between groups and the effect size was small [U=3.000, Z=-0.66, *p*=0.51, *r*=0.27].

Within the spontaneous phase, children in the Young group required 7.67 sessions to meet criterion, on average, whereas the Old group required 5.33 sessions, on average. The difference between groups was not statistically significant and the effect size was small [U=3.500, Z=-0.443, *p*=0.66, *r*=0.18]. Taken together, young and old children seemed to complete the treatment program in the same number of sessions.

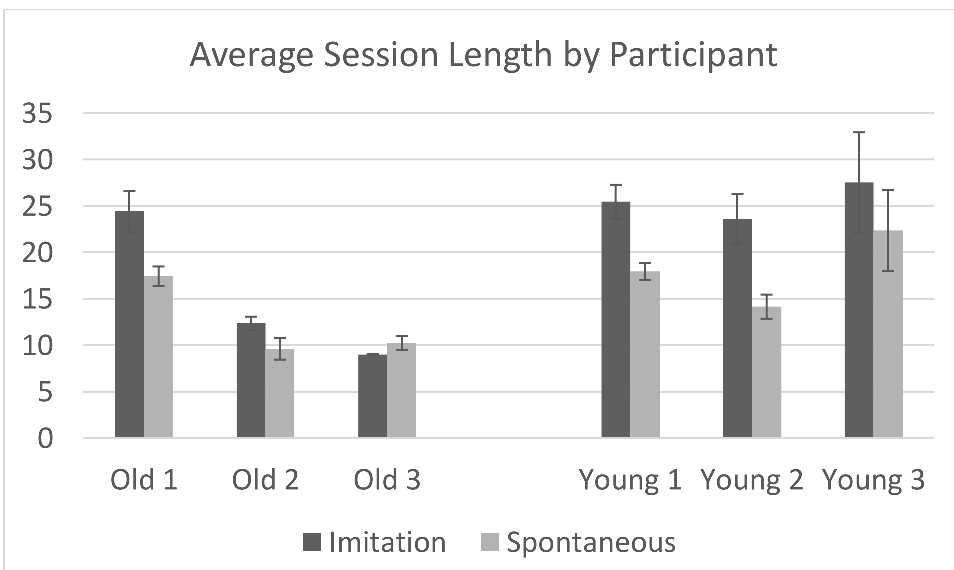


Figure 4: Mean number of minutes required for each subject for each phase.

Length of Sessions. Overall, treatment sessions for children in the Young group were 21.82 minutes on average. Treatment sessions for children in the Old group were 13.96 minutes on average. This was not a significant difference between groups, but the effect size was large, suggesting an increase of subjects may produce a significant result [U=1.00, Z=-1.53, p=0.13, r=0.62]. In the Imitation Phase, young children required 25.51 minutes on average, and old children required 15.25 minutes on average. These differences were not significantly different between groups but the effect size was large [U=1.00, Z=-1.53, p=0.13, r=0.62]. For the Spontaneous Phase, young children required 18.13 minutes on average, while old children required 12.43 minutes. This difference was not statistically significant, but the effect size was large [U=1.00, Z=-1.53, p=0.13, r=0.62]. Each of these findings suggest that older children likely require shorter sessions in treatment than younger children according to the large effect sizes, but the lack of statistical significance is driven by a lack of statistical power rather than a lack of effect. However, this difference in time is not due to an increase of trials or an increase in feedback, because the younger children required fewer trials as shown in the treatment accuracy data. This difference is more due to redirecting the younger children to the task and requiring more extensive explanations for articulator placement than for the older children. Older children, on the other hand, had all received school-based speech therapy and required less instructive feedback than the younger children. For younger children, the experience of receiving articulatory feedback was new.

Discussion

Young children are known to make significant gains in treatment because of early intervention. However, young children are often not provided treatment for late-acquired sounds because these sounds are not expected until a later age, despite evidence that suggests younger children may be more than capable of learning new speech sounds in general. Therefore, the purpose of the present study was to determine whether age was an impacting factor in the effectiveness and efficiency of treatment for late-acquired sounds. The results of the present study showed that all children benefitted from treatment by acquiring their treatment sounds and generalizing them to untreated real words. Moreover, younger and older children achieved similar steepness of generalization of the treated sound to untreated words. However, differences in learning during treatment were apparent. Younger children showed steeper learning than older children during the full course of treatment in terms of learning treatment nonwords, and the differences between the groups was larger in the Imitation Phase than in the Spontaneous Phase. In addition, acoustic analysis of children's correct productions at posttest showed that younger children's productions were within the expected range, while older children's productions were not. While all children achieved a perceptually accurate production of their treatment sound, only the young children produced treatment sounds to a natural level. In terms of number of treatment efficiency, young and old children required a similar number of sessions to complete the treatment, but younger children's sessions were longer than older children's. Although younger children showed steeper learning in the Imitation Phase, younger children's sessions were often longer than those of older children, which suggests that more redirection and instruction was required of the clinician to keep younger children moving through treatment. Regardless of treatment session length, all participants responded to therapy by meeting treatment phase criteria before the maximum cap of sessions was reached. Younger children are capable of learning late-acquired sounds and, in fact, teaching late-acquired sounds early may support steeper learning in treatment and more natural productions of the target than waiting to teach these sounds at older ages.

Shriberg et al. (1994) found that children aged 4-6 are in a period of accelerated phonological growth. Despite this evidence, clinicians often wait to provide therapy for late-acquired sounds until the expected age of acquisition. This may be due to a belief that younger children are not ready to learn late-acquired sounds before this expected age of acquisition, and instead they focus on earlier acquired sounds for treatment. On the contrary, previous research has shown that treating more complex sounds leads to broader generalization to untreated sounds (Elbert et al., 1984; Gierut & Hulse, 2010; Powell, 1991; Powell & Elbert, 1984; Williams, 1986). The results of the current study show that not only are younger children able to acquire late-acquired sounds just as well as older children, but they show steeper learning during treatment and more natural productions than older children. Therefore, children who are aged 4-6 should be provided treatment for late-acquired sounds because they are capable of acquiring them, and do so at a steeper rate.

Treating late-acquired sounds early would allow clinicians to take advantage of an accelerated period of phonological learning as well as broad system-wide generalization that occurs when treating late-acquired sounds (Gierut et al., 1996). Taken together, these factors reduce the amount of time a child spends in therapy By reducing time spent in therapy, children with speech sound disorders may also reduce the associated impacts across other domains. For example, even typically developing children who are aged 4-6 produce errors on late-acquired sounds, so resolving late-acquired errors in children with SSD at an earlier age will not only increase intelligibility, but the social impact of speech errors is greatly reduced. The child with SSD will not produce these errors into the later elementary school years, where it is less common for peers to produce them, thus reducing the likelihood of social stigma. All kindergarten children are acquiring pre-literacy skills. Since treating late-acquired sounds produces generalization to other untreated phonemes, it is likely that the impact of speech sound disorders on pre-literacy skills will be reduced.

The results of the present study have several other implications for practicing clinicians who work with children who have speech sound disorders. First, the results provided a measure of how long sessions should be for a treatment of 80 trials for each of these age groups. Recall that older children required 14 minutes on average, whereas younger children required 21 minutes on average. Although younger children showed steeper slopes in terms of their accuracy on treatment words over time, they required more support from the examiner. First of all, younger children were all preschool-aged, and less experienced with the process of speech therapy. Only one child (Young 1) had received formal speech therapy in a school setting. The examiner in the present study often had to redirect the child back to the treatment, and had to answer questions about what the nonwords were. The older children, on the other hand, had all received speech therapy for at least one year and were all in the public-school setting. They were familiar with working at a table and were mature enough to maintain their attention to the task. Another factor that is clinically relevant from these findings is that older children were all literate. They were able to rapidly identify the difference between speech sounds because of their phonological awareness abilities. “Old 3” for example asked the examiner, “Do these all have the /θ/ sound?” and proceeded to demonstrate awareness of the accurate production of this sound. Older children in the /ɹ/ condition understood, through their literacy skills--namely sound-to-letter correspondence--that the examiner was requesting a sound different from /w/, whereas this difference, for younger children, was less meaningful because they had not yet begun to learn to read.

The hypothesis of treatment efficiency predicted that young children would progress through treatment in fewer sessions than older children due to being in a phase of accelerated learning (Shriberg et al., 1994). Although young children overall showed a steeper slope, they required as many sessions overall as the older group. In terms of the breakdown between phases, young children required fewer sessions in the imitation phase than the older children, but older children made up for this difference by progressing through the spontaneous phase faster, although these differences were not statistically significant. This was an unexpected trend, but is interesting nonetheless. This finding may be the result of younger children learning the sounds faster due to the aforementioned advantages they have due to their age. However, younger children's lengthened spontaneous phases may indicate that they are more prompt-dependent for their success than older children. This is a point that should be explored further because it has the potential to impact the type of treatment required for children in these age groups—regardless of treatment target.

Limitations

The present study was a single subjects design to explore factors that are relevant in the successful treatment of late-acquired sounds, namely age. This was conducted to ensure a full examination of children's skills at onset and to determine which factors may be relevant to prepare for a larger group design. The smaller design allowed for this extensive look, however, it limits the statistical power of the analysis. In a larger group design, the findings of this study could be strengthened with additional statistical support.

Another limitation was the variability between the subjects treated with /ɹ/ and the subjects treated with /θ/. Those who were treated with /ɹ/ remained in treatment for a longer duration due to the overall difficulty of teaching and learning that sound. Those in the /θ/

condition learned their sound at a more rapid pace, and therefore spent less time in treatment. This resulted in the /θ/ group having a lower generalization score than the /ɹ/ treatment subjects. In a follow-up procedure, a minimum number of sessions would be determined in addition to the maximum cap already in place. This would ensure that all treatment subjects receive at least a minimal number of treatment sessions to ensure that generalization could take place.

Future Directions

The above limitations would be addressed in a future direction for this study in addition to other aspects of the research design. First, opening up the treatment sounds to include all late-acquired sounds would provide a means to examine differences within age groups to examine if and how the treatment target impacts treatment effectiveness and treatment efficiency. This would provide a more comprehensive view of late-acquired sounds as a whole. In addition, expanding the criteria to include all late-acquired sounds, would allow for an examination of the differences between treatment efficacy and efficiency for distortions as well as speech sound substitutions. It is possible that teaching an adjustment of the tongue position (e.g. a frontal lisp of /s/ to an alveolar production of /s/) may yield faster results with better generalization than teaching a substitution, and these results may be related to age.

Conclusion

The present study explored the boundaries of children's learning of late-acquired sounds by testing whether young children could learn late-acquired sounds as well as or better than older children. The findings of the present study suggest that, indeed, younger children learn late-acquired sounds as well as older children, and do so at a steeper slope of learning because of their increased motor plasticity. However, in terms of treatment efficiency, or how quickly children successfully complete treatment, younger children required longer sessions due to being inexperienced with speech therapy procedures and. Older children required more sessions, but the sessions were shorter because their metalinguistic skills, such as letter-to-sound corresponded, can be harnessed to overcome the loss of plasticity in their phonological system. Therefore, it seems that although both children experienced success in treatment, there were distinct differences between the two that impacted their efficiency.

References

Adler-Bock, M., Bernhardt, B. M., Gick, B., & Bacsfalvi, P. (2007). The use of ultrasound in remediation of North American English/ɹ/in 2 adolescents. *American Journal of Speech-Language Pathology, 16*(2), 128-139.

American Speech-Language-Hearing Association. (1996). Guidelines for screening for hearing impairment—School-aged children, 5 through 18 years. ASHA Desk Reference, 4, 364-366.

American Speech-Language-Hearing Association. (2014). Speech sound disorders: Articulation and phonological processes. American Speech Language Hearing Association. Retrieved April 25, 2016, from http://www.asha.org/public/speech/disorders/SpeechSoundDisorders/

Bernthal, J. E., Bankson, N. W., & Flipsen Jr, P. (2016). *Articulation and phonological disorders: speech sound disorders in children*: Pearson.

Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America, 66*(4), 1001-1017.

Brumbaugh, K. M., & Smit, A. B. (2013). Treating children ages 3–6 who have speech sound disorder: A survey. *Language, speech, and hearing services in schools, 44*(3), 306-319.

Chomsky, N., & Halle, M. (1968). The sound pattern of English.

Dalston, R. M. (1975). Acoustic characteristics of English/w, r, l/spoken correctly by young children and adults. *The Journal of the Acoustical Society of America, 57*(2), 462-469.

Dean, E., Howell, J., Waters, D., & Reid, J. (1995). Metaphon: A metalinguistic approach to the treatment of phonological disorder in children. *Clinical Linguistics & Phonetics*, *9*(1), 1-19.

Dinnsen, D. A., & Gierut, J. A. (2008). Optimality theory, phonological acquisition and disorders. Oakville, CT: Equinox Publishing Ltd.

Elbert, M., Dinnsen, D. A., & Powell, T. W. (1984). On the prediction of phonologic generalization learning patterns. *Journal of Speech and Hearing Disorders, 49*(3), 309-317.

Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels/w,j,r,l/in American English. *The Journal of the Acoustical Society of America, 92*(2), 736-757.

Flipsen, P., Shriberg, L. D., Weismer, G., Karlsson, H.B., & McSweeny, J.L. (2001). Acoustic phenotypes for speech-genetics studies: Reference data for residual/ɜ/distortions. *Clinical Linguistics & Phonetics, 15*(8), 603-630.

Gierut, J. A. (1998). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech Language and Hearing Research, 41*(1), S85.

Gierut, J. A. (1999). Syllable OnsetsClusters and Adjuncts in Acquisition. *Journal of Speech, Language, and Hearing Research, 42*(3), 708-726.

Gierut, J. A., & Champion, A. H. (2001). Syllable Onsets IIThree-Element Clusters in Phonological Treatment. *Journal of Speech, Language, and Hearing Research, 44*(4), 886-904.

Gierut, J. A., & Hulse, L. E. (2010). Evidence-based practice: A matrix for predicting phonological generalization. *Clinical Linguistics & Phonetics, 24*(4-5), 323-334.

Gierut, J. A., Morrisette, M. L., Hughes, M. T., & Rowland, S. (1996). Phonological treatment efficacy and developmental norms. *Language, speech, and hearing services in schools, 27*(3), 215-230.

Hedrick, M. S., & Ohde, R. N. (1993). Effect of relative amplitude of frication on perception of place of articulation. *The Journal of the Acoustical Society of America, 94*(4), 2005-2026.

Hesketh, A., Adams, C., Nightingale, C., & Hall, R. (2000). Phonological awareness therapy and articulatory training approaches for children with phonological disorders: a comparative outcome study. *International Journal of Language & Communication Disorders, 35*(3), 337-354.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108*(3), 1252-1263.

Kager, R. (1999). Optimality theory. New York, NY : Cambridge University Press.

Kamhi, A. G. (2006). Treatment decisions for children with speech–sound disorders. *Language, speech, and hearing services in schools, 37*(4), 271-279.

Klein, H. B., Byun, T. M., Davidson, L., & Grigos, M. I. (2013). A Multidimensional Investigation of Children's/ɹ/Productions: Perceptual, Ultrasound, and Acoustic Measures. *American Journal of Speech-Language Pathology, 22*(3), 540-553.

Krueger, B.I., Storkel, H.L. & Minai, U. (Under Review). The influence of misarticulations on children's word identification. *American Journal of Speech-Language and Hearing Research*. Manuscript under review.

Ladefoged, P., & Maddieson, I. (1998). The sounds of the world's languages. *Language, 74*(2), 374-376.

Lewis, B. A., & Freebairn, L. (1992). Residual effects of preschool phonology disorders in grade school, adolescence, and adulthood. *Journal of Speech, Language, and Hearing Research, 35*(4), 819-831.

Morrisette, M. L., & Gierut, J. A. (2003). Unified Treatment Recommendations: A Response to Rvachew and Nowak (2001). *Journal of Speech, Language, and Hearing Research, 46*(2), 382-385.

Powell, T. W. (1991). Planning for Phonological GeneralizationAn Approach to Treatment Target Selection. *American Journal of Speech-Language Pathology, 1*(1), 21-27.

Powell, T. W., & Elbert, M. (1984). Generalization following the remediation of early-and later-developing consonant clusters. *Journal of Speech and Hearing Disorders, 49*(2), 211-218.

Preston, J. L., & Edwards, M. L. (2007). Phonological processing skills of adolescents with residual speech sound errors. *Language, speech, and hearing services in schools, 38*(4), 297-308.

Preston, J. L., Molfese, P. J., Mencl, W. E., Frost, S. J., Hoeft, F., Fulbright, R. K., . . . Felsenfeld, S. (2014). Structural brain differences in school-age children with residual speech sound errors. *Brain and language, 128*(1), 25-33.

Rice, M. L., Hadley, P. A., & Alexander, A. L. (1993). Social biases toward children with speech and language impairments: A correlative causal model of language limitations. *Applied Psycholinguistics, 14*, 445-445.

Rvachew, S., & Nowak, M. (2001). The effect of target-selection strategy on phonological learning. *Journal of Speech, Language, and Hearing Research, 44*(3), 610-623.

Rvachew, S., Rafaat, S., & Martin, M. (1999). Stimulability, speech perception skills, and the treatment of phonological disorders. *American Journal of Speech-Language Pathology, 8*(1), 33-43.

Shadle, C. H. (1990). Articulatory-acoustic relationships in fricative consonants *Speech production and speech modelling* (pp. 187-209): Springer.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech, Language, and Hearing Research*, *36*(1), 105-140.

Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J. (1994). Developmental Phonological Disorders III: Long-Term Speech-Sound Normalization. *Journal of Speech, Language, and Hearing Research, 37*(5), 1151-1177.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779.

Stevens, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants: Static considerations. *The Journal of the Acoustical Society of America, 50*(4B), 1180-1192.

Stevens, K. N., & Keyser, S. J. (1989). Primary features and their enhancement in consonants. *Language*, 81-106.

Storkel, H. L. (2013). A corpus of consonant–vowel–consonant real words and nonwords: Comparison of phonotactic probability, neighborhood density, and consonant age of acquisition. *Behavior Research Methods, 45*(4), 1159-1167.

Tyler, A. A., Edwards, M. L., & Saxman, J. H. (1987). Clinical application of two phonologically based treatment procedures. *Journal of Speech and Hearing Disorders, 52*(393-409).

Weiner, F. F. (1981). Treatment of phonological disability using the method of meaningful minimal contrast: two case studies. *Journal of Speech & Hearing Disorders*.

Williams, A. (1986). Generalization to singletons following cluster training. *Unpublished master's thesis, Indiana University, Bloomington*.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics bulletin*, *1*(6), 80-83.

Appendix A

NONWORD	ACCURACY: EXAMINER JUDGMENT										TOTAL (out of 10)
1.											
2.											
3.											
4.											
5.											
6.											
7.											
8.											

CORRECT NUMBER OF EXPOSURES? YES OR NO

TOTAL ACCURACY: ____________________/80 * 100 = ______________%

Appendix B

Treatment Schedule, Baseline Condition 3		
Session 1	Hearing, GFTA-2, RIAS, Full PKP Part 1 (Baseline 1)	
Session 2	CTOPP, Full PKP Part 2 (Baseline 1)	
Session 3	TOLD-P4, PPVT BASELINE PROBE 2 (Treated Sound)	
Session 4	Baseline Probe 3 PPVT, KSPT	
Session 5	Baseline Probe 4 (Treated Sound, if in 5 baseline condition)	
Session 6	Baseline Probe 4/5 (Treated Sound, depends on baseline condition) & Imitation Treatment (if baseline stable)	
Session 7	Imitation	
Session 8	Imitation	Treated Sound Probe
Session 9	Imitation	
Session 10	Imitation	
Session 11	Imitation	Treated Sound Probe
Session 12	Imitation	
Session 13	Imitation	
Session 14	Imitation	Treated Sound Probe
Session 15	Imitation	
Session 16	Imitation	
Session 17	Imitation	Treated Sound Probe
Session 18	Imitation	
Session 19	Imitation	
Session 20	Phase-Shift Probe (PKP)	
Session 21	Spontaneous	
Session 22	Spontaneous	
Session 23	Spontaneous	Treated Sound Probe
Session 24	Spontaneous	
Session 25	Spontaneous	
Session 26	Spontaneous	Treated Sound Probe
Session 27	Spontaneous	
Session 28	Spontaneous	
Session 29	Spontaneous	Treated Sound Probe
Session 30	Spontaneous	
Session 31	Spontaneous	
Session 32	Spontaneous	Treated Sound Probe
Session 33	Spontaneous	
Session 34	Spontaneous	
Session 35	Spontaneous	Treated Sound Probe
Session 36	Spontaneous	
Session 37	Spontaneous	
Session 38	Spontaneous	Treated Sound Probe
Session 39	Spontaneous	
Session 40	Spontaneous	
Session 41	Spontaneous	Treated Sound Probe
Session 42	Spontaneous	
Session 43	Spontaneous	
Session 44	Spontaneous	Treated Sound Probe
Session 45	Immediate Post-test (PKP)	
Session 46	2 week Post-test (PKP)	

Appendix C

/ɹ/						
Word	CV Biphone Probability	Code	VC Biphone Probability	Code	Neighborhood Density	Neighborhood Density Code
ɹib	0.0044	high	0.0007	low	12	high
ɹɛb	0.0085	high	0.0007	low	10	high
ɹad	0.0011	high	0.0025	high	16	high
ɹʌp	0.0026	high	0.0012	high	18	high
Mean	**0.00415**		**0.001275**			
SD	**0.003198437**		**0.00085**			
ɹɔɪm	0.0001	low	0	low	7	low
ɹʊp	0.0002	low	0.0003	low	6	low
ɹaʊn	0.0005	low	0.004	high	13	high
ɹʊd	0.0002	low	0.0013	high	13	high
Mean	0.00025		0.0014			
SD	0.000173205		0.001820256			
Overall M	**0.0022**		**0.001338**			
Overall SD	**0.002957**		**0.001316**			

/θ/						
Word	CV Biphone Probability	Code	VC Biphone Probability	Code	Neighborhood Density	Neighborhood Density Code
θɪp	0.0011	high	0.0049	high	13	high
θɪm	0.0011	high	0.0068	high	10	high
θoun	0.0002	low	0.0022	high	12	high
θæm	0.0002	low	0.0049	high	10	high
Mean	**0.00065**		**0.0047**			
SD	**0.00052**		**0.00189**			
θʌp	0.0004	low	0.0012	high	7	low
θeɪp	0	low	0.0017	high	6	low
θaʊm	0.0001	low	0	low	2	low
θoʊb	0	low	0	low	2	low
Mean	**0.000125**		**0.000725**			
SD	**0.000189**		**0.000862**			
Overall M	**0.000388**		**0.002713**			
Overall SD	**0.000458**		**0.002523**			