Fall 2014

Discrimination of duration differences of typically developing children and children with suspected childhood apraxia of speech

Susan B. Ingram
James Madison University

Discrimination of Duration Differences of Typically Developing Children and Children with Suspected Childhood Apraxia of Speech

Susan B. Ingram

A dissertation submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

Department of Communication Sciences and Disorders

December 2014

Dedication

With appreciation to my extraordinary daughter, Emily Harper, I dedicate this dissertation manuscript to you. For every mile traveled, for every data point entered, and for every word of encouragement spoken/texted, I am grateful. Your unwavering companionship on this sometimes grueling trek has been exemplary. I look forward to celebrating the completion of your JMU degrees ('15, '16 M) in the near future!

Acknowledgments

I extend sincere gratitude to Dr. Vicki Reed for serving as my doctoral Advisor. By my side, she persevered through seemingly endless edits, fostered my courage, and nudged me out of my comfort zone. I am grateful for her mentoring and steadfast support throughout my candidature. I also wish to offer appreciation to several graduate faculty who generously served as Dissertation Committee members and helped to advance my research: Dr. Thomas Powell at Louisiana State University and Dr. Rory DePaolis and Dr. Carol Dudding, both colleagues at JMU. Recognition, too, goes to Dr. Chris Clinard for his assistance with generating my stimuli for the experimental task. Dr. Marsha Longerbeam, my longstanding comrade, frequently provided wise counsel: "Don't worry about things that fall outside your circle of control." Appreciation is also extended to my SLP colleague, Christine Reeves, for her significant contributions to this research study. And, to Dr. Nick Bankson, thank you for first planting the notion of pursuing a doctoral degree in my mind almost twenty years ago!

Several undergraduate and graduate students enrolled in the Department of Communication Sciences and Disorders played pivotal roles in this research study. Kelly Williams ('12, '14 M) was instrumental in both the development of my experimental stimuli and data analysis. Thanks, Kelly, for supporting this project from inception to defense! Nadine Joslin ('14 M) invested many hours over the course of 18 months viewing video recordings and scoring numerous assessment measures. I appreciate your attention to detail in scoring, Nadine, and fondly recall our early discussions regarding the research process. Julie Vest ('13, '15 M) took on the challenge of entering my references in EndNote; your consistent words of encouragement, Julie, and willingness to go out of your way to finish the task were valuable. Two additional graduate students, Caroline Barlow ('14 M) and Chris Wyeth ('17 M) deserve commendation for assisting with the statistical analyses of my data.

I would also like to acknowledge the support of other students who contributed to this research study in a variety of ways: Elyse Ballard ('14), Rebecca Barron ('15 M), Emily

Grabenstetter, Kasey Gray ('13, '15 M), Christina Luckey ('16 M), and Caroline Saur ('13).

Supportive friends and colleagues significantly facilitated this research endeavor, as well. Thanks to the following individuals for providing much-needed encouragement along the way: Janis Babbitt, Erin Clinard, Robin Doyle, Teresa Drulia, Kamaile Hiatt, Peggy Kennedy, Karen Lindberg, Dr. Cindy O'Donoghue, Dr. Stacey Pavelko, Betsy Runyan, Becky Shaver, Stephanie Signorino, Andrea Weaver, Carolyn West, Cam Wymer, and Julie Yoder. My appreciation is also extended to the child participants in my research study and to the parents of these children. I am inspired by you and hope that our collective efforts will further advance our knowledge of childhood apraxia of speech.

And, lastly, I am very grateful for my supportive family. I am appreciative of my empathetic husband, Dr. Richard Ingram, for being tolerant of my late nights at JMU and for holding down the fort with Tucker in my absence. When home, our "deck time" certainly helped to counter the rigor of the research! To Emily and Tim Shifflett, I knew I could always count on you to lend a helping hand. To my siblings, Sandra and Steve Shanaberger and Don and Jo Ellen Bowman, I thank you for the encouragement you consistently provided ... year after year after year! To my parents, Donald and Doris Bowman, I am thankful for the strong educational underpinning you gave us and the special memories I have of attending JMU's Anthony-Seeger Campus School as a child. To my proud Dad, I know you watch over us and continue to celebrate the accomplishments of your children and grandchildren. To my amazing Mom, also a JMU graduate, I am appreciative of the countless ways you assisted me. Throughout this almost decade-long doctoral journey, you consistently provided ample nourishment, serenity, and a listening ear.

Table of Contents

List of Tables

List of Figures

Abstract

The speech sound disorder, childhood apraxia of speech (CAS), has perplexed clinicians and researchers for many years. The perplexity has stemmed, in part, from questions about identifying characteristics that distinguish it from other childhood speech disorders. Given the reported vowel duration deficits cited in the speech production of children with sCAS, the research for this population is deficient in assessing the ability of these children to discriminate vowel duration differences.

The present study represents an initial attempt to address duration discrimination in a systematized experimental design for a group of school-age TD children (*n* = 21) and a smaller group of school-age children diagnosed with sCAS (*n* = 11). All children were asked to judge whether pairs of non-word single syllable tokens (digital recordings of single syllable /bɑ/ varying in vowel duration only) were the same or different. Using an AX paradigm, children in the current study compared a stimulus (X), which varied across trials, with a constant standard (A). The standard A interval was the stimulus with the shortest vowel duration (208 ms) and the X interval was the comparison stimuli (i.e., vowel duration = 208 ms, 248 ms, 288 ms, 328 ms, 368 ms, 408 ms, 448 ms, or 488 ms). Fundamental frequency and amplitude measures were controlled to remain uniform.

Assessing the ability of the TD population to detect duration differences in a specific experimental paradigm was prerequisite to addressing the ability of children with sCAS to detect duration differences in the same experimental task. The results of this preliminary investigation of discrimination of vowel duration in children with sCAS suggest that further research on duration discrimination skills is warranted in this population. As a group, children with sCAS displayed poorer performance on the vowel duration discrimination experimental task, compared to a similarly-aged TD group.

Chapter 1 Suspected Childhood Apraxia of Speech: Introduction and Statement of the Problem

1.1 Issues

1.1.1 Definition and characteristics of sCAS.

Over a quarter century ago, Guyette and Diedrich (1981) identified childhood apraxia of speech (CAS) as "a label in search of a population" (p. 39). Since Guyette and Diedrich questioned the existence of this disorder, ambiguity relative to the definition of CAS has persisted. In 2007, the Ad Hoc Committee on Apraxia of Speech in Children (American Speech-Language-Hearing Association, 2007) reported that the research and clinical literature on CAS, predominantly within the past ten years, contained more than 50 definitions of the speech sound disorder. Discrepancies exist in defining CAS, and variance is present with regard to the terminology used to refer to the disorder. Davis, Jakielski, and Marquardt (1998) cited the following terms used in the literature to refer to CAS: articulatory apraxia, developmental articulatory dyspraxia, childhood verbal apraxia, developmental apraxia of speech, and developmental verbal dyspraxia. As a result of the diverse terminology, the Ad Hoc Committee recommended *childhood apraxia of speech* (*CAS*) (ASHA, 2007, p. 2) as the term for this speech sound disorder and proposed it serve as a generic term for all descriptions of apraxia of speech included in childhood, regardless of specific etiology.

Furthermore, the 2007 Ad Hoc Committee recommended that CAS be recognized as a type of childhood speech sound disorder that "warrants research and clinical attention" (ASHA, 2007, p. 40). The Committee, however, advocated the use of a provisional diagnostic label due to the absence of a gold standard for differential diagnosis. Due to the lack of clarity associated with a diagnosis of CAS, the term "suspected" (Shriberg, Aram, & Kwiatkowski, 1997a, p. 282) childhood apraxia of speech (sCAS) will be used hereafter to reflect the difficulty professionals have at present to specifically identify this speech sound disorder, except when citing specific researchers' work. Issues related to difficulty identifying CAS and the term, suspected childhood apraxia of speech, include (1) lack of differential diagnostic measures, (2) fluctuating manifestations of the disorder in children over time, and (3) controversy of praxis-only versus praxis + linguistic deficits.

Differential diagnosis of sCAS is hindered, in part, by the absence of explicit clinical markers. Characteristics of sCAS discussed in the research literature are numerous and several of the features described, such as slow speech development and unintelligible speech, are applicable to other speech sound disorders (McCabe, Rosenthal, & McLeod, 1998), thus confounding differential diagnosis of sCAS, for example, from phonological disorder (PD). To make the task of differential diagnosis of sCAS even more difficult, sCAS features are typically presented as checklists of individual characteristics that mix perceptual, motor-based, and acoustic parameters of speech. Rather than being distinct, many of the features overlap and likely refer to similar characteristics, thus leading to lack of specificity in checklists. For example, the descriptors “staccato speech” (Maassen, 2002, p. 262) and “perceptual construct of isochrony” (Shriberg, Green, Campbell, McSweeny, & Scheer, 2003, p. 588) could be considered perceptual features resulting from the motor-based descriptions of “predominant use of simple syllable shapes” (Maassen, 2002, p. 262) and “inability to maintain syllabic integrity” (Gillon & Moriarty, 2007, p. 49). Furthermore, these perceptual and/or motor-based attributes could be outcomes of speech having acoustic characteristics such as “equalization of stress” (Yoss & Darley, 1974, p. 412) or “longer vowel durations” (Peter & Stoel-Gammon, 2005, p. 77). The lack of clinical markers, the mixing of perceptual/motor-based/acoustic descriptions, and the redundancy of descriptive characteristics all contribute to the conundrum of differentially diagnosing sCAS.

To complicate the diagnostic issue, not all characteristics of CAS are observed in every child (Davis et al., 1998). In addition, researchers (Lewis, Freebairn, Hansen, Iyengar, & Taylor, 2004; Skinder, Connaghan, Strand, & Betz, 2000; and Shriberg, Campbell, Karlsson, Brown, McSweeny, & Nadler, 2003) have suggested that the presence or absence of characteristics may change over time within the same individual. Results of a study by Shriberg, Campbell et al.

(2003) revealed that the presence of a lexical stress deficit may not be apparent in children with sAOS [suspected apraxia of speech] at the time of their assessment, such that segmental and suprasegmental deficits in sAOS may reduce over time with a greater frequency than previously suggested in the literature. In describing selection of participants for research, Nijland (2009) also reported challenges in differential diagnosis of sCAS and PD due to the varied presentation of characteristics for the speech disorders at different ages.

In addition to identification of speech production characteristics of sCAS, the Ad Hoc Committee cited research that identified associated deficits in the areas of nonspeech motor performance, metalinguistic/literacy, speech perception, language, and prosody. Davis et al. (1998), too, identified concurrent deficits in receptive and expressive language skills, nonverbal oral-motor functioning, and neurological development. Consequently, the uncertainty surrounding the diagnosis of sCAS is complicated by research suggesting that the central deficit of sCAS may not be limited to a speech production (praxis) deficit. As a result, the diagnostic debate has evolved from *motor* versus *linguistic* viewpoints to *motor-only* versus *motor-linguistic* stances. Summarizing, Shriberg, Green et al. (2003) stated that most empirical research defines sCAS as a movement disorder, but alternative viewpoints entertain inclusion of substantial processing deficits. Davis et al. (1998) listed the following theoretical accounts of sCAS:

> inadequate phonemic representation (Marquardt and Sussman, 1991), a deficit in underlying representations of phonemes, syllables, and suprasegmentals, manifested as motor and speech deficits (Aram, 1984; Bernhardt [*sic*], 1992; Bernhardt and Stoel-Gammon, 1994; Velleman and Strand, 1993 [*sic*]), deficits in pre-motor organization and sequencing abilities manifested at the motor output level (Hall, 1992; Hall, Jordan and Robin, 1993; Robin, 1992), and a "motolinguistic" perspective that provides for a continuum of possible clinical symptoms – from planning to execution of oral-motor movements (Crary, 1984, 1993). (p. 26)

More recently, Ozanne (2005) suggested the underlying deficits of sCAS might be categorized into four clusters:

1) a linguistic/phonological deficit characterized by "difficulty assembling the phonological plan for the word or utterance" (p. 78);

2) a motor deficit characterized by "difficulties in the motor-programming process which arises when the correct motor programme is chosen but the wrong timing and force parameters are chosen (Schmidt and Lee, 1999)" (p.78);

3) a motor deficit characterized by "difficulty at the level of phonetic programme or plan assembly" (p. 80);

4) a prosodic/segmental deficit characterized by difficulties associated with the "lack [of] the basic building blocks of speech (Bernhardt, 1993)" (p. 80).

Disregarding the controversy over motor- versus motor-linguistic deficits, the 2007 Ad Hoc Committee proposed the following working definition, simply referring to sCAS as a "speech sound disorder":

> a neurological childhood (pediatric) speech sound disorder in which the precision and consistency of movements underlying speech are impaired in the absence of neuromuscular deficits (e.g., abnormal reflexes, abnormal tone). CAS may occur as a result of known neurological impairment, in association with complex neurobehavioral disorders of known or unknown origin, or as an idiopathic neurogenic speech sound disorder. The core impairment in planning and/or programming spatiotemporal parameters of movement sequences results in errors in speech sound production and prosody. (ASHA, 2007, p. 3)

The literature is rife with descriptions of speech production characteristics of sCAS. Among these are slow speech development, restricted phonetic or phonemic inventories, multiple sound errors, reduced percentage of consonants correct, and unintelligible speech. These features are not considered discriminating because they can be frequently associated with other speech sound disorders (ASHA, 2007; McCabe et al., 1998). There are, however, several speech characteristics that might discriminate more accurately between sCAS and other speech sound disorders because they are considered to occur less frequently in children who are not suspected to have apraxic speech (ASHA, 2007). These traits include a restricted vowel inventory, vowel errors, inconsistency in multiple productions of the same target word or syllable, increased errors as length and complexity of verbalizations increase, groping articulatory movements, atypical errors, regression in accuracy of production of sounds and words previously mastered, altered production for automatic versus volitional/imitative speech, and errors in sound sequencing (Davis et al., 1998; McCabe et al., 1998; Shriberg et al., 1997a).

Although no list of diagnostic features has been validated to differentiate sCAS from other speech sound disorders, three segmental/suprasegmental features that are "consistent with a deficit in the planning and programming of movements for speech" (ASHA, 2007, p. 3) have acquired some consensus among researchers in the field: (a) inconsistent errors on consonants and vowels on attempts of multiple productions of the same target word or syllable, (b) atypical coarticulatory transitions between sounds and syllables, and (c) inappropriate prosody, especially noted with lexical or phrasal stress. Although sCAS diagnostic features remain unresolved, the Ad Hoc Committee noted a general consensus that "syllables and prosody are affected in more profound, distinctive ways in CAS than are other aspects of speech or phonology" (p. 16).

1.1.2 Proposed clinical markers.

Identification of agreed-upon diagnostic markers would reduce ambiguity surrounding sCAS and be advantageous for researchers in developing their methodologies. Davis et al. (1998) noted that use of a consistent set of characteristics to diagnose sCAS is critical so that empirically based conclusions and inferences regarding the disorder can be generated. Furthermore, frequent misdiagnosis of sCAS by professional speech-language pathologists supports the need for the development of agreed upon and consistently used clinical diagnostic protocols to facilitate accurate diagnosis of sCAS from other speech sound disorders (Davis et al., 1998). McCauley and Strand (2008) reviewed six published, standardized tests that claimed to screen/diagnose motor speech disorders in children. Of those tests reviewed, the authors found that only one met their operational definition of adequacy for validity and none presented with adequate measures of reliability. It is not surprising that consensus cannot be met regarding diagnosis of CAS if one vital piece of assessment - assessment of oral/speech motor function, for example – has deficiencies in both measures of reliability and validity. Shriberg, Lohmeier, Strand, and Jakielski (2012) reiterated the critical need of researchers to develop and standardize diagnostic criteria for CAS that meet or exceed 90% sensitivity and specificity requirements. Unfortunately, at least three primary confounds often negatively impact research efforts to identify clinical markers - circularity in identification of sCAS participants, nondiscrete inclusionary criteria for identification of sCAS participants, and researchers' decisions regarding the control groups to which to compare sCAS participants.

Davis et al. (1998) commented on the impact of circularity for subject selection criteria in which there is a blurring of subject selection criteria with the dependent variable(s) being examined. For example, research participants are assigned to groups based on a speech production characteristic or behavior. That same characteristic or behavior, however, is also a dependent variable in the study, which results in a circular argument. A specific example is apparent in a study conducted by Odell and Shriberg (2001). Participants included children with suspected apraxia of speech who had inappropriate stress (AOSci) and adults with apraxia of speech (AOS). The inclusionary criterion for the AOSci group was use of inappropriate stress occurring in a minimum of 20% of conversational utterances. Prosody-voice patterns for the AOS group, for which inclusionary criteria did not include mandatory stress misplacement, were compared to the prosody-voice patterns of the AOSci group. Not surprisingly, the results indicated the adult speakers with AOS had "significantly fewer utterances meeting criteria for inappropriate stress" (p. 275) compared to the participants with AOSci. By requiring inappropriate stress to be an inclusionary factor in the selection process for the children in the AOSci group and not for the adult AOS group, it appears the researchers selected their child subjects on the parameter they measured as a dependent variable.

Relative to nondiscrete identification of participants with sCAS, numerous studies have reflected diagnostic ambiguity regarding the selection of subjects as a result of vague inclusionary criteria. Examples are seen in Table 1.1. Forrest (2003) reported 50 different characteristics, not to be confused with 50 definitions, within the diagnostic criteria for developmental apraxia of speech used by 75 speech-language pathologists and noted, "These results are consistent with the general ambiguity of the diagnostic criteria of DAS [Developmental Apraxia of Speech] and suggest that no single deficit is used among clinicians" (p. 376).

Table 1.1. Examples of studies reflecting nondiscrete subject selection in the sCAS research literature.

RESEARCHER(S)	***sCAS SUBJECT SELECTION***
Groenen, Maassen, Crul, & Thoonen (1996)	"The purpose was to form a homogeneous group of children whose *main problem was apraxic in nature*" [emphasis added] (p.470).
Davis, Jakielski, & Marquardt (1998)	"All children had been diagnosed with DAS or suspected DAS based on the primary characteristic(s) of a *moderate or severe phonological disorder and/or slow progress in speech therapy"* [emphasis added] (p. 28).
Odell & Shriberg (2001)	"The identification of AOSc [suspected apraxia of speech in children] for each child in the AOSc sample was based on *a definition of AOSc adopted by the clinician-researcher making the diagnosis*" [emphasis added] (p. 284).
Munson, Bjorum, & Windsor (2003)	"It is always possible that *another SLP* [speech-language pathologist] *using different behavioral criteria for sDAS may have diagnosed some of the children identified as having sDAS in the current study as having PD*" [emphasis added] (p. 200).

With regard to the third research confound, the type of control groups used in research with sCAS, the majority of participants selected for comparison have been typically developing (TD) children rather than children identified with other speech sound disorders. For example, over half (62%) of 21 randomly selected studies published between 1988 and 2005 exclusively incorporated normally or TD controls. Comparison to children with phonological deficits was absent in the majority of these studies. Clinical data identifying differences in the behavior of the sCAS population and normal controls do not facilitate the identification of potential markers to clinically differentiate the sCAS population from other speech sound disorders. If comparisons occur between sCAS and another speech sound disorder, then conclusions regarding the role of linguistic stress, for example, in a particular speech sound disorder might be clarified.

Since the late 1990s, few advances have been made in the identification of inclusionary criteria for subjects in order to differentially diagnose sCAS. In the first of a series of three papers in 1997, Shriberg et al. (1997a) commented on the differential diagnostic methodology for sCAS at that time: "…the state of the art is an approach wherein clinicians and researchers consult a number of diagnostic checklists that purport to characterize *probable* [emphasis added] features of children with suspected DAS" (p. 274). Six years later, Shriberg, Campbell et al. (2003) commented, "Until a biomarker becomes available to identify children who are true positives for apraxia of speech, proposals of provisional inclusionary criteria continue to appear, commonly termed *diagnostic checklists*" (p. 552). At the present time, the standard diagnostic methodology utilized to identify sCAS subjects is typically a checklist of behavioral characteristics, but with little guidance as to how to implement the checklist for diagnosis. There is no guide as to which of the characteristics are requisite. Continued prevalence of the use of checklists seemingly implies the profession has stalled in the identification of a gold standard for differentiating sCAS from other speech sound disorders.

Building on the research that has suggested there may be some speech characteristics that are less frequently seen in children deemed not to have sCAS, the following research literature contains a range of assertions about potential diagnostic markers to differentiate sCAS among speech sound disorders. For example, as early as 1994, Velleman and Strand suggested it is the struggle to synchronize segmental and suprasegmental features (i.e. coordination of the articulators with the respiratory/laryngeal systems) that frequently distinguishes the sCAS population from children with articulatory or phonological deficits. Relative to proposed diagnostic markers for sCAS that are more segmental than nonsegmental in nature, Burns (2011) suggested that seven speech characteristics (i.e., nasalization, glottal stops, final consonant deletion, backing, frication, oral motor imitation, and fluency) may serve to differentiate sCAS from auditory perceptual/phonological disorders. Several segmental features identified by Austin and Shriberg (1997) also show some sensitivity and specificity for a praxis disorder: groping, metathetic substitution errors, inconsistent production on repetition of matching tokens, increased sound/syllable deletions, and increased vowel/diphthong errors. Furthermore, Thoonen, Maassen, Gabreëls, and Schreuder (1999) suggested measures of MPT (maximum performance tasks) may contribute to the differential diagnosis of speech sound disorders.

In contrast to characteristics that are more segmental in nature, Davis et al. (1998) noted that speech production variability, vowel errors, and suprasegmental deficits are characteristics that may facilitate a specific diagnosis of sCAS and, thus, may serve as diagnostic markers. Shriberg, Campbell et al. (2003), too, implicated suprasegmental features, along with segmental markers, as differentially diagnosing sCAS. Three suprasegmental features that appear sensitive to and specific for a praxis versus a phonological deficit include: inconsistent stress production on syllables or words, inconsistent temporal limitations on speech and pause events, and inconsistent oral-nasal resonance issues (Shriberg, Kwiatkowski, Rasmussen, Lof, & Miller, 1992). One global suprasegmental feature, speech timing, was proposed to be constrained for children with sCAS versus a severe speech delay (Shriberg, Green et al., 2003).

The literature contains multiple proposals implicating *inappropriate stress* (Yoss & Darley, 1974; Shriberg, Aram, & Kwiatkowski, 1997b, c; Davis et al., 1998; Skinder, Strand, & Mignerey, 1999; Shriberg, Campbell et al., 2003) as a distinguishing feature for diagnosing sCAS. The following list of quotations highlights implied stress problems:

- “only linguistic domain that differentiates some children with suspected DAS from those with SD [speech delay] is inappropriate stress” (Shriberg et al., 1997b, p. 286)
- “inappropriate stress is a diagnostic marker for a subtype of DAS” (Shriberg et al.,

1997b, p. 306)

- "inappropriate stress may be an excellent candidate for an 'ideal' phenotype marker" (Shriberg et al., 1997c, p. 329)
- "major support for theoretical and clinical perspectives on stress as a marker for at least one form of DAS" (Shriberg et al., 1997c, p. 320)
- "inappropriate stress might stand out as the first candidate to serve as a diagnostic marker for DAS (McNeil, Robin, & Schmidt, 1997; Hall, Jordan, & Robin, 1993; Ozanne, 1995; Shriberg et al., 1997c; Thoonen, Maassen, Gabreëls, Schreuder, & de Swart, 1997)" (Maassen, 2002, p. 262).

1.1.3 Perceptual and acoustic assessment measures of speech production.

As seen in Table 1.2, the inclusion of acoustic measures in the literature aimed at characterizing the speech of children with sCAS is less frequent than that of perceptual measurement of production, particularly for research involving aspects of timing. Research, to date, has been predominantly based on listeners' subjective assessment of the speech production of children diagnosed with sCAS. As recently as 2002, Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat, and Schreuder (2002) reported that the majority of the research in the area of sCAS was founded on perceptual assessments of production, specifically addressing the phonemic/feature qualities of segmental speech errors.

Use of perceptual assessment measures to evaluate production is not without controversy. Nijland et al. (2002) suggested there are methodological restrictions with perceptual analyses, including the inability to detect subtle phonetic differences in articulatory movements, for example. These researchers advocated use of instrumental analyses to provide quantitative/objective data regarding speech production that extends beyond auditory-based conclusions.

In addition to *subjective* assessment measures (perceptual analyses of production) and more *objective* evaluation methods (acoustic and instrumental analyses) discussed in the sCAS literature, Terband, Maassen, Guenther, and Brumberg (2009) added a *predictive* element to

Table 1.2. Studies using perceptual and/or acoustic analyses of sCAS speech.

PERCEPTUAL	***PERCEPTUAL/ ACOUSTIC***	***ACOUSTIC***
Yoss & Darley (1974)*	Groenen, Maassen, Crul, & Thoonen (1996)	Maassen, Nijland, & van der Meulen (2001)
Pollock & Hall (1991)	Thoonen, Maassen, Wit, Gabreëls, & Schreuder (1996)	Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat, & Schreuder (2003)*
Stackhouse & Snowling (1992)	Rodriguez (1998)* (Oral motor deficits vs. sCAS specifically)	Shriberg, Campbell, Karlsson, Brown, McSweeny, & Nadler (2003)*
Marion, Sussman, & Marquardt (1993)	Skinder, Strand, & Mignerey (1999)*	Shriberg, Green, Campbell, McSweeny, & Scheer (2003)*
Shriberg, Aram, & Kwiatkowski (1997b)*	Thoonen, Maassen, Gabreëls, & Schreuder (1999)	Peter & Stoel-Gammon (2005)*
Shriberg, Aram, & Kwiatkowski (1997c)*	Alcock, Passingham, Watkins, & Vargha-Khadem (2000)*	
Thoonen, Maassen, Gabreëls, Schreuder, & de Swart (1997)	Skinder, Connaghan, Strand, & Betz (2000)*	
Davis, Jakielski, & Marquardt (1998)*	Skinder-Meredith, Stoel-Gammon, Wright, & Strand (2001)*	
Velleman & Shriberg (1999)*	Munson, Bjorum, & Windsor (2003)*	
Skinder (2000)*	Peter & Stoel-Gammon (2008)*	
Odell & Shriberg (2001)*		
Marquardt, Sussman, Snow, & Jacks (2002)		
Nijland (2009)		

** Denotes research related to timing, rhythm, stress or duration.*

sCAS research. They introduced use of computer simulations to suggest how the deviant speech production system in individuals with sCAS may be attributed to impaired feed-forward commands, thus placing increased dependence on their auditory feedback system.

One method of perceptual/acoustic measurement of production that has been cited in the sCAS literature (Rodriguez, 1998; Peter & Stoel-Gammon, 2005; Peter & Stoel-Gammon, 2008) is administration of the *Tennessee Test of Rhythm and Intonation Patterns* (*T-TRIP*; Koike & Asp, 1981). This is a non-word, imitative task in which participants are asked to repeat up to 17 rhythmic sequences and 8 intonational patterns derived from the syllable "ma." Stimuli vary in both the length (two to nine syllables) and stress/intonational pattern and can provide both perceptual and acoustic data regarding production of rhythmic/intonational patterns.

Rodriguez (1998) compared the prosodic performance of kindergarten-age children (normal, phonologically impaired, oral motor impaired) using the *T-TRIP*. Based on the results of her research, Rodriguez suggested the *T-TRIP* has the potential to be used as an instrument to identify subjects whose prosodic deficits in speech production are consistent with an oral motor speech disorder, namely sCAS.

In 2005, Peter and Stoel-Gammon compared the temporal production abilities of children diagnosed with sCAS to age-matched peers during speech (sentence imitation, non-word imitation using the *T-TRIP*, monosyllabic word generation) and music-related tasks (singing, clapped rhythm imitation, repetitive tapping). Based on the results, the researchers reported,

> Of all speech tasks in this study, the non-word imitation tasks yielded the largest and most consistent differences between participants with sCAS and their TD peers, which makes it a potential assessment tool of choice in clinical and research settings involving children with sCAS. (p. 82)

Similarly, in 2008, Peter and Stoel-Gammon replicated and expanded their 2005 study. Once again, the *T-TRIP* was administered to provide data regarding production of syllable imitation and vowel duration. Comparing the results of this non-word imitation task to clapped rhythm imitation and paced repetitive tapping tasks, data suggested the greatest difference in performance between the participant groups was the comparison of participants' vowel durational patterns to the model.

The value of non-word stimuli (simple consonant-vowel syllabic patterns) has been reported in research that has addressed assessment of children's production skills.

1.1.4 Perception deficits.

To date, the primary focus of researchers in the sCAS literature has been on the production/speech motor deficits associated with this speech sound disorder and significantly less emphasis on the perception/processing aspects. Nijland (2009) commented on the relative lack of research that has been directed toward the analysis of perceptual skills for both children presenting with CAS and for children diagnosed with PD. What research has been conducted on perceptual features of the disorder led the Ad Hoc Committee on Apraxia of Speech in Children (ASHA, 2007) to suggest, "One of the major differences among alternative definitions of CAS is whether the core problem is proposed to include input processing as well as production, and if so, whether auditory, sensory, and prosodic aspects of perception may prefigure in the deficit" (p. 3). One of the six domains of behavioral research in sCAS cited by the Ad Hoc Committee is speech perception characteristics.

Based on the more limited literature describing the perceptual skills of children diagnosed with sCAS, the majority of research to date has addressed their segmental/phonological skills and significantly less attention has been devoted to investigating the perceptual abilities with suprasegmental features. In fact, of the 11 identified studies addressing different aspects of children's perceptual abilities, seven looked at segmental/phonological features. Bridgeman and Snowling (1988) reported dyspraxic participants had difficulty discriminating sound sequences in nonsense words but not real words. In 1992, Stackhouse and Snowling identified auditory processing deficits characterized by weak phonological representations in their developmental verbal dyspraxia case studies. The research of Marion et al. (1993) revealed significant deficits in rhyme generation, as well as rhyme detection, in their apraxic participants; results were suggestive of a diminished phonemic representation system. Groenen et al. (1996) found their subjects with DAS presented with impaired ability to discriminate /b/ versus /d/ and Maassen, Groenen, and Crul (2003) found impaired ability to perceive vowels, in both identification and discrimination tasks, in their apraxic subjects. Deficits in both higher-level perception tasks (rhyming task and categorical classification task) and lower-level perception tasks (non-word and categorical discrimination tasks) were cited for CAS participants by Nijland (2009). And, lastly, Shriberg et al. (2012) cited speech processing deficits related to auditory-perceptual encoding of phonological representations, memory, and transcoding as "core features of CAS in both idiopathic and neurogenetic contexts" (p. 477).

The remaining four of the identified perceptual studies addressed suprasegmental and rhythmic (syllabic) parameters. In 2000, Alcock et al. identified deficits in affected KE family members related to the discrimination of rhythmic patterns, thus impacting their timing abilities. Based on the research of Maassen et al. (2001), children with DAS showed restrictions in syllabic programming, in addition to deficits in motor programming. Marquardt et al. (2002) documented severe deficits in the ability of their DAS participants to identify the number of syllables in words and to judge intrasyllabic sound positions within consonant-vowel-consonant (CVC) words. Subsequently, these researchers proposed that children with DAS have an "apparent breakdown in the ability to perceive 'syllableness' and to access and compare syllable representations with regard to position and structure" (p. 31). Additionally, in 2003, Nijland et al. cited deficits in syllable planning in children with DAS and suggested that differences between this impaired group and NS (normally-speaking) children are related to how the two groups "process prosodic aspects" (p. 15).

1.2 Suprasegmental Features Associated with sCAS

According to Lehiste (1996), the term *suprasegmental features* includes duration/quantity, tone/intonation, and stress/emphasis, and assessment of these features involves comparing their values based on segments (phonemes) within a sequence, not based on an individual segment/phoneme. The acoustic measures of duration, fundamental frequency, and amplitude along with the corresponding perceptual measures of length, pitch, and loudness define suprasegmental features. Historically, the term *prosody* referred to suprasegmental qualities

(Lehiste, 1996) but was replaced by linguists with the term *suprasegmental* to highlight the difference between prosodic and segmental features. Suprasegmental errors are commonly found in the communication profile of individuals presenting with sCAS (see Table 1.3), more commonly in reference to production than perception. A debate exists whether suprasegmental deficits exhibited by children with sCAS significantly affect their syllable/segmental production or if the latter is primarily responsible for the prosodic insufficiencies (ASHA, 2007). Davis et al. (1998) questioned whether suprasegmental issues reflect deficits in higher-order programming or, perhaps, serve as a compensatory strategy for children with sCAS.

Two of the three primary sCAS characteristics identified by ASHA (2007) relate to abnormal suprasegmental patterns: deviant co-articulatory transitioning between sounds/syllables and impaired lexical and phrasal stress. The use of the generic term "suprasegmental" to refer to more specific subordinate characteristics, such as lexical stress, found in the sCAS literature may result in confusion regarding discrete elements. Therefore, the presentation of "suprasegmental features" that follows advances from nonspecific terminology such as *suprasegmental/prosodic* and *rhythmic* to more specific vocabulary such as *timing* and, ultimately, to more precise terminology including *linguistic stress* and *duration*.

1.2.1 Rhythm.

Velleman and Strand (1994) suggested prosody can be "functionally characterized as the ability to maintain the *rhythm* [emphasis added] and intonation of speech over time" (p. 127). Furthermore, MacNeilage and Davis (1990) proposed that, at the suprasegmental/prosodic level, the "*syllable* serves as the unit of *rhythmic organization* [emphasis added] …" (p. 56). Henry (1990) suggested that Shields (1981) may have considered organization to be the function of *rhythm* in speech production because it operates as a timing system. Table 1.4 presents a summary of research findings related to errors associated with rhythm and syllabic structure in sCAS.

Table 1.3. Suprasegmental/prosodic deficits associated with sCAS.

RESEARCHER(S)	***RESULTS RELATED TO PRODUCTION***	***RESULTS RELATED TO PERCEPTION***
Rosenbek & Wertz (1972)	"*Disturbed prosody* [emphasis added] is seen most frequently on imitative speech tasks…" (p. 30).	
Yoss & Darley (1974)	"*Prosodic features may be altered* [emphasis added] …" (p. 412).	
Velleman & Strand (1994)	"Difficulty with dynamic organization is not only a problem in the articulation of segments and syllables but is also *apparent at the suprasegmental level* [emphasis added] …" (p. 126).	
Davis, Jakielski, & Marquardt (1998)	"Characteristics that led to the specific diagnosis of DAS included variability of productions, vowel errors, and *suprasegmental variability*" [emphasis added] (p. 41).	
Skinder (2000)	"These findings indicated a segmental-*suprasegmental* [emphasis added] relationship does exist" (p. 2).	
Shriberg, Campbell, Karlsson, Brown, McSweeny, & Nadler (2003)	"…they [sAOS] reflect the *prosodic consequences* [emphasis added] of a praxis deficit in speech motor control" (p. 549).	
Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat, & Schreuder (2003)		"… this difference between NS children and children with DAS is a difference in the way these two groups *process prosodic aspects*" [emphasis added] (p. 15).

1.2.2 Timing.

According to Paul, Bianchi, Augustyn, Klin, and Volkmar (2008), one component of prosody is the rhythm and *timing* patterns that comprise the phrasing of an utterance. In the sCAS literature, *timing* is a rather broad term that describes one suprasegmental feature: quantity/duration. As shown in Table 1.5, several researchers have cited timing deficits in

Table 1.4. Rhythmic and syllabic deficits associated with sCAS.

RESEARCHER(S)	***RESULTS RELATED TO PRODUCTION***	***RESULTS RELATED TO PERCEPTION***
Davis, Jakielski, & Marquardt (1998)	Diadochokinesis: "Repetitions were produced 'haltingly'and *without rhythm* [emphasis added]" (p. 38).	
Alcock, Passingham, Watkins, & Vargha-Khadem (2000)	Rhythm production (manual and vocal): "…affected family members *reproduced fewer rhythms* [emphasis added] than the controls…" (p. 38).	Rhythm perception (vocal): "The affected family members *discriminated fewer rhythms* [emphasis added] than the control group…" (p. 38).
Maassen, Nijland, & van der Meulen (2001)		Two-word utterances: "The conclusion is that children with DAS show evidence for *deficient syllabic programming* [emphasis added] as well as deficient motor planning" (p. 149).
Marquardt, Sussman, Snow, & Jacks (2002)		Syllable awareness, intrasyllabic position, intrasyllabic structure: "Results suggest that DAS children demonstrate an apparent *breakdown in the ability to perceive 'syllableness'* [emphasis added]…" (p. 31). "Intonation contour, stress, and the entire *rhythmic organization of speech output would be expected to suffer* [emphasis added] if an impoverished syllabic structure existed" (p. 44).
Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat, & Schreuder (2003)		Disyllabic utterances within a carrier phrase: "…the present study provides indications for a *problem in the planning of syllables* [emphasis added] in speech production of children with DAS…" (p. 21).

Peter & Stoel-Gammon (2005)	Singing: "…participants with sCAS had *greater difficulty than their TD peers in producing a coherent rhythmic structure* [emphasis added] based on underlying durational relationships" (p. 82).	
Gillon & Moriarty (2007)	"Speech characteristics may include…*inability to maintain syllabic integrity* [emphasis added]…" (p. 49).	
Peter & Stoel-Gammon (2008)	Nonword imitation, clapped rhythm imitation, paced repetitive tapping: "Results suggest a central timing deficit, expressed in both the oral and the limb modality, and observable in two different types of timing measures, *overall rhythmic structures* [emphasis added] and small-scale durations" (p. 171).	

individuals diagnosed with sCAS.

1.2.3 Linguistic stress.

Based on their research of movement timing, Peter and Stoel-Gammon (2005) suggested that a timing variable, rather than other vocal parameters such as intensity and pitch, may serve as the most significant contributor to lexical *stress* deficits identified in children with sCAS. According to Shriberg et al. (1997b), "Inappropriate stress meets construct validity criteria as being wholly consistent with the clinical percept of DAS as a term used for children who 'sound different'" (p. 306). Numerous accounts of linguistic stress deficits are found in the sCAS literature, including stressing and destressing syllables at word level (lexical) and stressing and destressing syllables and words according to their morphological/syntactic use at phrase/sentence level (phrasal/sentential), as summarized in Table 1.6. The research has wholly focused on production aspects of stress as shown by the empty cells in the perception column.

Table 1.5. Timing deficits associated with sCAS.

RESEARCHER(S)	***RESULTS RELATED TO PRODUCTION***	***RESULTS RELATED TO PERCEPTION***
Crary (1984)	"…DVD [developmental verbal dyspraxia] is a motor-linguistic disorder of the developing phonological system with the underlying etiology being *deficits in spatial-temporal control* [emphasis added] of the speech mechanism" (p. 80).	
Velleman & Strand (1994)	"*Inconsistent timing* [emphasis added] and control of nasality and prosody" (p. 113).	
Alcock, Passingham, Watkins, & Vargha-Khadem (2000)	Rhythm production: "…*timing abilities* are *impaired* [emphasis added]" (p. 34).	Rhythm perception: "…*timing abilities* are *impaired* [emphasis added]" (p. 34).
Shriberg, Green, Campbell, McSweeny, & Scheer (2003)	Speech production: *"…*a *constraint in speech timing* [emphasis added] is a core feature of the praxis disorder that defines a developmental form of apraxia of speech" (p. 575). "Terms such as isochrony, syllable segregation, scanning speech and staccato-like rhythmic quality have been used to characterize the *temporal regularity* [emphasis added]…" (p. 575).	
Peter & Stoel-Gammon (2005)	Speech and music production: "…participants with sCAS showed greater *temporal inaccuracies* [emphasis added]…" (p. 84).	
Peter & Stoel-Gammon (2008)	Nonspeech and rhythm production: "Results suggest a *central timing deficit* [emphasis added] expressed in both the oral and the limb modality…" (p. 171).	

Table 1.6. Linguistic stress deficits associated with sCAS.

RESEARCHER(S)	***RESULTS RELATED TO PRODUCTION***	***RESULTS RELATED TO PERCEPTION***
Yoss & Darley (1974)	Spontaneous contextual speech: "A measured effect was present even in contextual speech, with a tendency toward *equalization of stress* [emphasis added]" (p. 412).	
Shriberg, Aram, & Kwiatkowski (1997b)	Conversation: "Results suggest that the only linguistic domain that differentiates some children with suspected DAS from those with SD is *inappropriate stress* [emphasis added]" (p. 286).	
Shriberg, Aram, & Kwiatkowski (1997c)	Conversation: "Summed across the three studies, 52% of 48 eligible samples from 53 children with suspected DAS had *inappropriate stress* [emphasis added], compared to 10% of 71 eligible samples from 73 age-matched children with speech delay of unknown origin" (p. 313).	
Davis, Jakielski, & Marquardt (1998)	Speech production: "…exhibited suprasegmental errors that involved vocal quality, *syllable and word stress* [emphasis added]…" (p. 41).	
Rodriguez (1998) (Oral motor deficits vs. sCAS specifically)	Nonword repetition /mɑ/: "In general, the results of the current study support Shriberg et al.'s (1997b, 1997c) hypothesis that *inappropriate stress* [emphasis added] is a diagnostic marker for oral-motor speech disorders, such as DVD" (p. 114); "…oral-motor subjects tended to *delete syllables and to convert iambic stress into trochaic* [emphasis added]" (p. 102); "... *contrast between stressed and unstressed syllables did not exceed the JNDs* [emphasis added] for one or more parameters" (p. 103).	

Skinder, Strand, & Mignerey (1999)	Bisyllabic/multisyllabic words: "Listeners perceived the control subjects to more accurately mark *syllabic stress* [emphasis added] than DAOS [developmental apraxia of speech] subjects…" (p. 133).	
Velleman & Shriberg (1999)	Conversation: "… *syllable omissions* [emphasis added] persisted to much later ages in the SD-DAS [suspected developmental apraxia of speech] subjects, especially those children previously identified as having *inappropriate phrasal stress* [emphasis added]" (p. 1444).	
Skinder (2000)	Nonsense/real words: "Effects of phonetic complexity and stress pattern on accurate *lexical stress production* [emphasis added] were significant for nonsense words…" (p. 2).	
Odell & Shriberg (2001)	Conversation: "Although *stressing of typically unstressed vowels* [emphasis added] is the primary prosodic (i.e., stress) behaviour that defines AOSci [children with suspected apraxia of speech and inappropriate stress]…" (p. 301).	
Maassen (2002)	"There is a tendency to *neutralize vowels* [emphasis added] and a tendency *not* to neutralize, namely to *stress unstressed syllables* [emphasis added], resulting in staccato speech…" (p. 262).	
Shriberg, Campbell, Karlsson, Brown, McSweeny, & Nadler (2003)	Bisyllabic word imitation: "The primary diagnostic-marker finding of this study is that a lexical stress task and a composite lexical stress ratio derived from three acoustic features are sensitive to *stress differences* [emphasis added] in children with sAOS" (p. 566).	

Peter & Stoel-Gammon (2005)	Nonword imitation: "…participants with sCAS produced a less accurate syllable count than their TD peers, *omitting weak syllables that did not fit a simple trochaic (i.e. strong-weak) syllable template*" [emphasis added] (p. 82).	
Gillon & Moriarty (2007)	"Speech characteristics may include…*inappropriate stress* [emphasis added] and intonation patterns" (p. 49).	

Summarizing the research in the sCAS literature regarding linguistic stress, children identified with this disorder may display stress deficits. Shriberg et al. (1997c) inferred from their data that approximately half of the children referred with sCAS may present with inappropriate stress. Errors may be characterized by equalization of stress, conversion of iambic stress into trochaic, reduction of contrast between stressed and unstressed syllables, neutralization of vowels, production of stress on unstressed syllables or, simply, deletion of syllables. However, Shriberg, Campbell et al. (2003) cautioned that "not all children with sAOS have a stress deficit *at the time they are assessed*" (p. 569) because it is possible suprasegmental production deficits in sCAS may subside over time.

Proposed theoretical justifications for the occurrence of stress deficits range from an underlying linguistic deficit to a motor-skill deficit. Shriberg et al. (1997c) suggested suprasegmental deficits may originate at an underlying phonological representational level since segmental and suprasegmental development are closely linked. These researchers noted that placing stress deficits at the representational level is compatible with the fact that children with sCAS may have difficulty executing volitional, simple sequencing tasks (e.g., diadochokinesis) requiring "reliable stress assignment for each syllable" (p. 326) due to a "prelexical deficit in the ability to format stress assignment for syllables (cf. Bastiaanse, Gilbers, & van der Linde, 1994)" (p. 326). In addition, based on the research of Marion et al. (1993) on rhyming deficits in children with sCAS, a stress deficit may be manifested both receptively and expressively, further suggesting that a stress deficit may be at the underlying linguistic representational level.

Odell and Shriberg (2001) have also suggested that the stress deficits associated with children with sCAS may be linked to their reduced ability to self-correct and may be attributed to a processing deficit characterized by “unstable or fuzzy underlying representations” (p. 301). The stress deficits identified in the participants in their study may suggest delays in the acquisition of stress rules in English: Auditory/auditory-temporal processing deficits could negatively affect decoding skills and accurate representation of rhythmic patterns, thus impeding encoding skills and subsequent self-monitoring of appropriate stress patterns. Perhaps self-correction is absent because there is no discrepancy between the underlying linguistic representation of the stress pattern and the inaccurate output, thus “indicating a deficit at the highest levels of the representational aspects of stress” (p. 300).

Conversely, stress errors in children with sCAS may be attributed to a speech motor deficit. When describing the suprasegmental speech characteristics of children with sCAS, Velleman and Strand (1994) proposed the children’s tendency to prolong vowels may serve as a compensatory behavior, as they are “‘buying time’ to organize the coordination of the next series of movements” (p. 127). Skinder et al. (2000) noted a relationship between segmental accuracy of trochaic and iambic tokens and accuracy of lexical stress. In other words, “The lack of an established motor plan could adversely affect stress patterns in the speech produced by children with DAS” (p. 282).

Not all researchers have documented significant differences in stress production between sCAS and control groups. Although research by Munson et al. (2003) identified perceptual differences in the stress patterns of nonwords produced by children with sCAS versus PD as judged by trained listeners, acoustic analyses did not support any differences in linguistic stress production between the two groups. It should be noted, however, that the trochaic and iambic tokens utilized in the research were limited to 2-syllable nonwords and participants produced them in isolation, thus the final syllables were subject to phrase-final lengthening (Lehiste, 1996). Similarly, Velleman and Shriberg (1999, p. 1459) cited no "quantitative differences in lexical stress error patterns" among speech delayed versus sCAS participants using metrical analysis. Nevertheless, the researchers noted that the conversational speech transcripts that were analyzed included primarily trochaic words and few words were multisyllabic. Consequently, the need to expand the research of lexical stress errors using elicited word productions and acoustic analyses was emphasized.

1.2.4 Duration.

Table 1.7 lists the research that has focused on one component of linguistic stress in the sCAS population, *duration*. Perhaps duration, rather than the other acoustic elements of stress (e.g., fundamental frequency, intensity), has been researched because of the infrequency with which inappropriate loudness and pitch is noted in the production of the speech-disordered population (Shriberg et al., 1997b). In fact, only one of the 14 children with sCAS in the Shriberg et al. (1997b) study presented with incorrect loudness and pitch. Results of research by Alcock et al. (2000) also identified accurate perception and production of pitch during both speech and music tasks in affected KE family members; however, deficits were noted in the perception and production of rhythm during the same tasks. As a result, Peter and Stoel-Gammon (2005) suggested that the research of Alcock et al. "validates an approach to investigating prosodic errors in children with sCAS that focuses on timing rather than on other acoustic correlates of stress, such as intensity and intonation" (p. 68) and hypothesized that timing control may contribute significantly to the deficits noted in lexical stress production in children with sCAS. Support for assessing duration is also derived from the work of Sluijter and van Heuven (1996), in which they found "duration proved the most reliable correlate of stress" (p. 2471). Rodriguez (1998) reiterated the importance of the duration component:

> Generally, duration was the most frequently used acoustic parameter, followed by F0; duration was used in 90% of all correct responses to mark stress, while F0 was used in

73% of all correct responses. Overall, amplitude was only used in 57% of correct responses. (p. 87)

When assessing the role of each parameter (i.e., duration, fundamental frequency, and amplitude) and not a combination of parameters to denote stress, Rodriguez cited duration as the single indicator of stress patterns 21% of the time (13 of 63 correct items). Conversely, fundamental frequency only (3%) and amplitude only (2%) were seldom used.

Despite the literature emphasizing the importance of duration in sCAS, the research listed in Table 1.7 is focused, to date, only on duration production to the exclusion of research on duration perception, as noted by the empty cells in the perception column.

1.3 Suprasegmental Features Associated with Typical Development

In order to investigate the perceptual abilities of sCAS children to determine differences in vowel duration, a standard against which to compare their performance is required. Therefore, information related to the suprasegmental abilities of typically developing children is relevant.

Prosody is an intricate activity that involves acoustic adjustment of pitch, loudness, and length to relay linguistic information and emotion (Boutsen & Christman, 2002). With regard to both linguistic and affective prosodic development, Doherty, Fitzsimons, Asenbauer, and Staunton (1999) reported a somewhat leisurely acquisition period for perceptual skills. Age-related improvement was noted for TD children up to 8.5 years of age. Although varied prosodic patterns are observed early in production of canonical babbling, Goffman (1999) reported that refinement continues beyond the age of four to six years for production of prosody. Rodriguez (1998), too, suggested children should be able to manipulate prosody in spontaneous speech, as well as imitatively during experimental tasks, by approximately four to six years of age.

1.3.1 Rhythm.

Davis and MacNeilage (1995) identified rhythmicity as the most distinguishing characteristic of babbling to the listener, attributed to the occurrence of uniform timing. They cited that “patterns resulting from consonant-like closing phases of low energy and short duration

Table 1.7. Duration deficits associated with sCAS.

RESEARCHER(S)	***RESULTS RELATED TO PRODUCTION***	***RESULTS RELATED TO PERCEPTION***
Velleman & Strand (1994)	Children with DVD may be "heard to *prolong vowels* [emphasis added] because by prolonging the steady state of the vowel they are 'buying time' to organize the coordination of the next series of movements" (p. 127).	
Maassen, Nijland, & van der Meulen (2001)	Two-word utterances in a carrier phrase: "Children with DAS *did not produce any systematic durational pattern*" [emphasis added] (p. 148) based on location of syllable segment, as did the TD group.	
Nijland, Maassen, van der Meulen, Gabreëls, Kraaimaat, & Schreuder (2003)	Disyllabic utterances within a carrier phrase: "…children with DAS have *significantly longer total and segment durations*…" [emphasis added] (p. 15); "…children with DAS *show no effect of syllable structure on the duration* [emphasis added] …" (p. 21).	
Shriberg, Green, Campbell, McSweeny, & Scheer (2003)	Conversation: "…the children with sAOS had proportionally *more variation in the duration* [emphasis added] of pause events and/or *less variation in the duration* [emphasis added] of speech events" (p. 575).	
Peter & Stoel-Gammon (2005)	Sentence imitation: *Vowel duration was less accurate* [emphasis added] for sCAS vs. TD; Nonword imitation: *"Substantial"* (p. 77) *overlap for stressed vs. unstressed vowel duration ranges* [emphasis added] for sCAS; *"longer vowel durations in general"* [emphasis added] (p. 77) for sCAS; Word generation: *"Greater nucleus durations"* [emphasis added] (p. 78) for sCAS.	

Peter & Stoel-Gammon (2008)	Nonword imitation: "In terms of effect size, defined as Cohen's d, *the difference between the two groups was greatest for Non-word Imitation: Adult:Child Correlation Coefficient of Vowel Durations*, [emphasis added] where Cohens's d was 1.46…" (p. 188).	

alternating with vowel-like phases of higher energy and longer duration produce the impression of a regularly timed sequence of consonant-vowel syllables" (p. 1199). In fact, MacNeilage and Davis (1989) suggested even the uniformity of infants' early mandibular movements (frame) result in the perception of syllable-like productions to the listener. Similarly, Lindblom (1983) proposed that the average syllable duration is not random but that the length is determined significantly by biomechanical and physiological aspects. Allen and Hawkins (1980) noted that the speech of very young children is comprised primarily of unreduced (strong) syllables, thus generating a syllable-timed rhythm. By the age of four or five years, with the addition of reduced (light) syllable nuclei in function words and multisyllabic words, the rhythm becomes more adult-like.

Languages, categorized by the linguistic unit (isochrony) that describes the perceived durational regularity of a language, may fall within one of three classes: stress-timed (linguistic unit equals the time between two subsequent stressed syllables), syllable-timed (linguistic unit equals the duration of somewhat equivalent syllables), and mora-timed (linguistic unit equals syllable weight). Dauer (1983) proposed an alternate rationale for rhythmic differences among languages. She offered that the rhythmic variations in languages may be more related to aspects of their phonology, phonetics, lexicon, and syntax than "any attempt on the part of the speaker to equalize interstress or intersyllable intervals" (p. 55). The work of Lehiste (1977) suggested that a tendency toward isochrony (near-equal durational measures) not only exists in speech production, but in perception, as well.

1.3.2 Timing.

With regard to speech production, Stevens (1998) described two types of articulatory movement associated with timing: unidirectional (progression from one articulatory position to another) and cyclic (progression from one position to another and return to the initial articulatory position). He reported individuals present the fastest intervals of change for either unidirectional or cyclic movements associated with (a) lip or tongue movement during production of stop consonants, (b) jaw elevation/depression, and (c) vocal fold abduction/adduction. More delayed periods of adjustment were reported for tongue movements relating to vowel production, velar elevation/depression, and subglottal pressure variation. In order to facilitate the speed of articulatory movement and promote the ability to produce two or three words per second (Levelt, Roelofs, & Meyer, 1999), Levelt et al. (1999) theorized that speakers have "access to a repository of gestural scores for the frequently used syllables of the language [syllabary]" (p. 5). These rote, stored templates help alleviate the need to re-compute the motor plan, thus facilitating the production process. Relative to perception of timing parameters, Morrongiello and Trehub (1987, p. 413) noted, "The processing of temporal information in audition is of fundamental importance for sound localization, rhythm perception, speech discrimination, and the detection of signals in noise."

1.3.3 Linguistic stress.

Goffman (1999) suggested that trochaic (strong-weak) prosodic forms have their origin in canonical babbling and, consequently, develop earlier than iambic (weak-strong) forms. Further, she noted that "trochees rely on a less specified rhythmic structure, whereas iambs require a purposeful and controlled movement frame" (p. 1515). With regard to production of unstressed (weak) syllables, Echols (1993) suggested that these are especially "fragile" (p. 289) in children across early language development and prone to omission or simplification, as sometimes evidenced by reduplications. Goffman (1999) provided rationale for this occurrence:

1) unstressed syllables in initial word positions are less perceptually salient (Echols, 1993);

2) trochaic versus iambic forms may be easier to produce; and 3) a movement-related bias may exist toward production of unmodulated rhythms (trochaic) versus highly modulated (iambic) rhythmic forms.

Vihman, DePaolis, and Davis (1998) assessed stress placement in the early word learning of 13- to 20-month-olds. Based on their production of disyllabic words, English-speaking infants displayed trochaic stress patterns only slightly more frequently than iambic patterns. The research of Davis, MacNeilage, Matyear, and Powell (2000) indicated the vocal output of prelinguistic infants includes both asymmetrically and uniformly stressed patterns, with approximately equal frequencies of occurrence. Although babbling infants can alter the pitch, loudness, and length of their vocalizations, variation noted in these prosodic parameters does not suggest infants have intentionality or control of these parameters necessary for word production, which requires stress to be placed in assigned locations. Relative to production of rhythmic patterns, Davis et al. (2000) reported no trochaic stress bias in prelinguistic infants. DePaolis, Vihman, and Kunnari (2008) reported that the acquisition of stress in infants relies on linguistic knowledge. Perceptually, however, Jusczyk, Cutler, and Redanz (1993) reported American infants display a preference for the trochaic stress pattern by nine months of age.

In their study of children ages 18 to 30 months, Kehoe, Stoel-Gammon, and Buder (1995) found that this population exhibited control of the acoustic correlates of stress in the majority of their familiar word productions, although adult-like competency was not achieved. Allen and Hawkins (1980) reported variability in three- to five-year-old children to appropriately reduce syllables from their full form in early words to shortened forms found in more adult-like speech, while Goffman (1999) found four-year-olds competent and adult-like in their ability to produce weak-strong patterns. Smith and Robb (2006) reported that knowledge is limited regarding the acoustic characteristics of stress production as children progress from preschool to school age. Given that research relative to linguistic stress in the typically developing preschool population is restricted and results are variable, the need for additional research is warranted.

1.3.4 Duration.

Based on the research of Pollock, Brammer, and Hageman (1993), duration may be a relevant acoustic correlate early in a child's speech development. Addressing two-, three-, and four-year-olds, these researchers found that all age groups "demonstrated the ability to use vowel duration to differentiate stressed from unstressed syllables, although there was continued improvement with age in the ability to reduce duration for unstressed syllables" (p. 183). Although Lee, Potamianos, and Narayanan (1999) reported five-year-olds have the ability to regulate vowel duration in their speech, the research of Smith and Robb (2006) suggested children approximately six years of age still struggle to use duration to mark stress.

Although durational differences in production have been reported in children as young as 10 to 18 months (DePaolis et al., 2008; Allen & Hawkins, 1980), development of durational patterns spans years. Vowels evidence "mature patterns of relative and absolute duration quite early, by around 3 years for heavy syllable nuclei and by 4 or 5 years for light nuclei" (Allen & Hawkins, 1980, p. 236), while consonant durations require considerably more time to stabilize and may not be secured by children until age ten or beyond.

In terms of perception of durational differences, a literature search identified only four research studies addressing duration discrimination in typically developing children. However, none of these focused on discrimination of duration differences in phonemes. Morrongiello and Trehub (1987) addressed perception of both silence and signal duration in six-month-old infants, five and one-half year-old children, and adults. Participants were presented with a sequence of 18 white-noise bursts and asked to discriminate a change in duration of the middle six bursts or intervals of silence. Results revealed infants discriminated duration changes of 20 ms, preschool children discriminated duration changes of 15 ms, and adults discriminated variations as small as 10 ms. While citing the lack of research on children's development of duration discrimination, the researchers noted that estimates of duration discrimination thresholds for adults

> vary across studies as a function of psychophysical test procedure, stimulus context, and duration of the standard interval, but range from approximately 10 to 25 ms for standard intervals between 150 and 400 ms (e.g., Abel, 1972; Chistovich, 1959; Creelman, 1962; Henry, 1948). (p. 415)

A second study (Elfenbein, Small, & Davis, 1993) looked at developmental patterns of duration discrimination. Participants included 40 children, ages four to ten years, and ten adults. Difference limens for duration were obtained using a three-interval, forced-choice paradigm using a 350-ms noise burst as the standard stimulus. The duration of the standard stimulus was selected to mirror the length of an average syllable. Participants were asked to identify the variable stimulus from the two presentations of the standard stimulus in each trial. The differences in milliseconds between the standard and variable stimuli required for ≥70% correct discrimination were obtained. Results suggested improved ability to detect duration differences as age increased from four-, to six-, to eight years of age. Only the performance of the ten-year-old participants approximated that of the adults in a consistent manner.

Based on the results of their study, the researchers reflected on the factors contributing to the age-related improvement evidenced. They suggested that

> some factor or set of factors affecting discrimination of duration in a three-item forced-choice paradigm change markedly as a child matures from 4 to 8 years. Further, this factor or set of factors gradually approximates adult form as the child matures from 8 to 10 years. Candidates for this list of factors include comprehension of instructions, motivation, attention to the task, learning, maturation of the auditory system, and auditory memory. (p. 847)

Conclusions of the study suggested comprehension, motivation, attention, and task learning may have contributed to the obtained results, but were not the only influences. Auditory memory may have impacted the results observed between and within groups, given the three-interval forced-choice paradigm selected for the study.

A third study, Jensen and Neff (1993), also addressed developmental patterns of duration discrimination. Participants included 41 children, ages four to six years, and nine adults. Difference limens for duration were obtained using a three-interval, forced-choice paradigm using a 400-ms tone as the standard stimulus. Participants were asked to identify the variable stimulus from the two presentations of the standard stimulus in each trial; the approximate 70% correct response level was identified. Results described duration discrimination as "relatively poor" in many four- and five-year-old children, although the ability to detect duration differences significantly improved as age increased from four, to five, to six years of age. The performance of many six-year-olds, however, still did not approximate adult performance. Significant individual differences within all age groups were also reported. The researchers cited memory issues as a possible influencing factor for their obtained results, as participants were required to listen to the entire length of three tones (two standard and one variable) before identifying which tone was different.

A look at children's ability to discriminate duration is also found in the research of Himpel et al. (2009). Participants included 40 individuals, six to 18 years of age. Pairs of auditory stimuli, a standard interval and a variable interval, were presented to participants and they were instructed to identify which of the two intervals was longer. Two different base (standard) measures were assessed and mean 75% thresholds (the difference in duration between the standard and comparison intervals necessary to achieve 75% correct responses) were determined. Results for the task with a base duration of 50 ms revealed a mean 75% threshold of 16.6 ms (S.D. = ± 6.3 ms). For the task with a base duration of 1000 ms, the mean 75% threshold was 215.1 ms (S.D. = ± 83.6 ms).

In contrast to the four previously described studies including children, Huggins' (1972) research focused on adults. He examined perception of adults for just noticeable difference (JND) for segment duration. Greater increased sensitivity to duration differences of the vowel /ɔ/ than to changes in the duration of four consonants was found. Further, Huggins suggested that changes in phoneme duration may also affect changes in perceived stress and perceived rhythm. Kawai and Carrell (2012) addressed adults' ability to detect durational differences for the phoneme /s/ in both words and sentences. Results revealed the JND for /s/ averaged 13.3 ms at word level and 9 ms in sentences. Furthermore, a gender difference was also reported for JND, with the performance of male adults exceeding that of the female participants. Comments by the researchers included,

> Few research studies have systematically investigated human capacities for identification and discrimination of phoneme duration differences. Digital technologies have made it possible to manipulate the duration of phonemes while preserving high sound quality. Accurate estimation of human thresholds for durations of phonemes is important to understand precisely how humans process sound signals which do not have different meanings in the auditory system… (p. 191)

Ohde and German (2011) shared similar concerns about children: "Based on existing studies, it appears that very little research has been conducted on vowel perception in children" (p. 1630).

1.4 Summary

To date, differential diagnosis of sCAS from other speech sound disorders is hampered by the lack of explicit clinical markers. Nonetheless, the Ad Hoc Committee on Apraxia of Speech in Children (ASHA, 2007) cited deficits associated with "syllables" and "prosody" (p. 16) as significantly altered in the sCAS population. In fact, two of the three primary sCAS characteristics identified by the Committee relate to abnormal prosody/suprasegmental patterns. The body of research suggests *perceptual* analysis of suprasegmental parameters, including syllabic processing and vowel duration discrimination, is a likely candidate for study in the sCAS population.

Within the broad category of suprasegmental deficits in the sCAS population, the literature contains multiple studies citing children's production of inappropriate stress as a discriminating feature of sCAS. The importance of duration, in particular, as an acoustic correlate of linguistic stress (Sluijter and van Heuven, 1996; Alcock et al., 2000) is confirmed by the results of a study by Peter and Stoel-Gammon (2008). Their data suggested the greatest difference in performance between the TD and sCAS participant groups was noted in the contrast of participants' production of vowel durational patterns.

Compared to the frequency of studies in the literature focused on production skills in the sCAS population, the research related to perceptual/processing issues is more limited. And, of the existing perceptual research, the majority is dedicated to segmental, not suprasegmental, features. However, within the small body of available literature on suprasegmental qualities in the sCAS population, research has identified deficits in processing syllables (Maassen et al., 2001; Marquardt et al., 2002; Nijland et al., 2003) and discriminating rhythmic patterns (Alcock et al., 2000).

For the TD population, the amount of available research regarding their perception of durational differences is also limited. Four research studies in the literature dealing with children, identified in the present study, have looked at duration discrimination of either noise bursts or tones in children ranging in age from infants to 18 years of age. However, data regarding discrimination of duration differences in phonemes in TD children is absent in the literature. In light of the limited research regarding discrimination of duration differences, a study focused on duration discrimination in the TD population, especially at the syllable level, is warranted as part of examining duration discrimination of children with sCAS in order to explore potential differences in their abilities.

1.5 The Problem and Hypotheses

Research on the ability of children with sCAS to detect differences in vowel duration is needed to more clearly define the nature of this not-well-understood speech sound disorder and to differentiate it from other sound disorders. However, in the absence of comprehensive literature on the development of duration discrimination in TD children, research with an impaired population would be hampered without attention to the ability of TD children to detect durational differences, at least for performances on a specific task designed to examine duration discrimination. This research includes initial exploration of the TD children's performance on a task designed to assess duration discrimination in order to obtain a standard for comparison. If TD children show expected patterns of behavior, a preliminary look at a sample of children with sCAS would be justified to see if their patterns of performance appear similar to or different from those of the TD children on the same task.

Therefore, this research focuses on young, school-age children's discrimination of vowel durational differences in non-words (acoustically controlled CV syllables). After ensuring that TD children, age 5 through 8 years, can successfully navigate the experimental task designed for this research, I hypothesize that: 1) the children's performances at the extreme duration differences (i.e., pairs that are of the same duration, pairs that are toward the maximum duration differences) will demonstrate mostly correct responses, 2) as duration differences increase from the smallest difference in duration through the intermediate duration differences, the children's correct response rates will show improved performances as the duration differences increase, but variances will become large, and 3) the performances of the younger TD children will differ from the performances of the older TD children. Secondly, I hypothesize that the performances of a small sample of 5- and 6-year-old children with sCAS, compared to a subset of the TD children who are 5- and 6 years old, will show different patterns of performances on the duration-difference experimental task than those of their TD similarly-aged peers.

Chapter 2 Methodology

2.1 Pilot Study

A small pilot study was conducted on a group of TD 5- to 8-year-olds (*n* = 4) to determine if they were able to complete a non-word listening task. Acoustic stimuli, three- to five syllables in length, were created using naturally-produced speech (productions of /bɑ/) which were then digitally altered in order to assess children's ability to determine longer versus shorter vowel durations. Pairs of multi-syllabic non-words, consisting of /bɑ/ and differing only by vowel length, were used to assess the children's ability to determine if the pairs when heard were the same or different. Results of this pilot study indicated children, as young as five years of age, were able to successfully perform the task. The children's auditory memory skills, attention span, and ability to attend to a computer-based listening task all appeared adequate for task completion. Consequently, the investigator was confident that the children, ages 5;0 to 8;11, enrolled in the main study were likely to be able to complete a similar non-word listening task comprised of pairs of only single-syllable stimuli, given the performances of the children enrolled in the pilot study utilizing multisyllabic stimuli.

2.2 Main Study

The first research question exclusively addressed the performances of the TD children, ages 5;0 to 8;11. The second research question addressed performances of several 5- and 6-year-old children with sCAS compared to a subset of the TD children who were similarly aged.

2.2.1 Participants.

Participants in this research study included the following: 1) the investigator; 2) five Ph.D. faculty (Clinard, DePaolis, van Dorn, Reed, Ludlow) to assist in the development of the experimental stimuli; 3) several students to assist in the development of the experimental stimuli and the computer program to deliver the stimuli; 4) two graduate research assistants to assist in scoring children's assessment data; 5) several graduate students over the course of data collection to assist in administering assessment instruments and recording children's performance; 6) one speech-language pathologist (SLP) to assist in scoring children's assessment data; 7) the same SLP to confirm the group status (sCAS or TD) for randomly selected children; 8) two graduate research assistants to assist in data entry; 9) four statisticians to advise regarding data analyses; 10) two undergraduate students to verify all data entry; and 11) 32 children (and their parents) to serve as eligible participants. This study was approved by the Institutional Review Board of James Madison University.

To address the first hypothesis, research children were 21 TD males, ages 5;0 (years; months) to 8;11. To test the second hypothesis, a subset of the TD children who were between the ages of 5;0 and 6;11, along with 11 additional male children between 5;0 and 6;11 with a diagnosis of sCAS participated in the study. The *Diagnostic Evaluation of Articulation and Phonology* (DEAP; Dodd, Hua, Crosbie, Holm, & Ozanne, 2006) advises that if multiple testing sessions occur, a child's age at initiation of testing is the age used for scoring and evaluation. Therefore, children were classified by their age of entry into the study, even if their chronological age might have advanced by a month during the course of data collection. Table 2.1 shows relevant demographic information.

Table 2.1. Demographic information of the children.

Children	*N*	Mean Age	Age Range
TD Children	21 males	6;7 (yrs.; mos.)	5;0 – 8;8 (yrs.; mos.)
5-year-olds	7	5;6	5;0 – 5;10
6-year-olds	7	6;5	6;0 – 6;11
7/8-year-olds	7	7;9	7;2 – 8;8
7 yrs	4	7;4	7;2 – 7;6
8 yrs	3	8;5	8;1 – 8;8
sCAS Children	11 males	5;7 (yrs.; mos.)	5;0 – 6;10 (yrs.; mos.)
5-year-olds	9	5;4	5;0 – 5;10
6-year-olds	2	6;8	6;5 – 6;10

The chronological age range of the TD children in the present study (5 through 8 years of age) was selected because young children's suprasegmental maturation has been reported to develop during these ages but to do so relatively slowly. Goffman (1999) reported the acquisition of prosody in production tasks is still developing in four- to six-year-old children. Although her research participants were able to produce perceivable distinctions between iamb and trochee word forms, they reportedly lacked maturity in the extent of the difference they were able to achieve in varying the syllables of the word forms, especially trochees. The research of Rodriguez (1998) suggested kindergarten children are able to imitate prosody (acoustic correlates of stress) in non-word experimental contexts.

Although indications in the production research suggest that children as young as five years can manipulate prosody, the literature is less conclusive relative to development of the perception of prosody. Doherty et al. (1999) described a relatively "slow developmental time course for prosody" (p. 225) for typically developing children in their perception of linguistic and affective prosody. More specifically, Jensen and Neff (1993) indicated that the discrimination abilities for frequency and duration of many typically developing four- and five-year-olds in their study were "relatively poor" (p. 106), while the ability to discriminate intensity was adult-like by age five. Significant improvement in the ability to detect duration differences was noted by Jensen and Neff as participant age increased from four to six years of age, although considerable individual differences were cited within all age groups. Elfenbein et al. (1993), too, reported improved ability of their child participants to detect durational differences as age increased from four to ten years. Only the performance of the ten-year-old participants consistently approached that of the adult participants.

The chronological age range of the sCAS children in the present study (5 through 6 years of age) was based, in part, on the research of Shriberg, Green et al. (2003) who stated, "Failure to normalize speech delay by 6 years of age is a frequently cited diagnostic characteristic of [sCAS]" (p. 580). Additionally, the 5- through 6-year age range is a significant period of phonological development and is one of the common age populations SLPs have on their caseload in school settings. The upper age of sCAS children in the current study was capped at

6;11 years to increase opportunities for assessing children before, as Shriberg et al. (1997b) stated, "Both advanced age and more extensive intervention experience militate against finding persisting stress involvement" (p. 308).

The decision to include only males in this study was influenced by 1) greater prevalence of communication disorders found in the male population (Robb, 2010) and 2) the reported gender differences for production of vowel duration in males and females. Females generally produce longer vowel durations than males (Simpson & Ericsdotter, 2003; Neel, 2008). Clopper, Pisoni, and de Jong (2005) also found significant gender differences in duration for the following five vowels: /ɪ/, /ɛ/, /ɑ/, /ʌ/ and /ʊ/, with females producing longer vowels than males. Additionally, Kawai and Carrell (2012) reported gender differences for perceptual measures of just noticeable difference (JND) for duration of the phoneme /s/. Thus, using only males minimized the possibility of confounding results because of performance difference related to gender.

2.2.2 Recruitment, selection procedures, and overview of assessment process.

To recruit the TD children, the investigator distributed flyers to colleagues and SLPs employed in public schools and requested they forward the study information to parents of males within the desired chronological age range (5;0 – 8;11). To recruit the children with sCAS, SLPs employed in public schools, private clinics, and university speech-language clinics were asked to forward the study information to parents of males whom they believed presented with characteristics consistent with their clinical opinion of sCAS. Additionally, information regarding the research study was posted on a subject recruitment website, sponsored by CASANA (Childhood Apraxia of Speech Association of North America), a site specific to helping researchers recruit children with CAS.

Once referred, all children participated in a comprehensive assessment battery administered individually by the investigator and graduate student clinicians enrolled in the

Department of Communication Sciences and Disorders at James Madison University to determine eligibility for inclusion in this study. Participating graduate student clinicians were trained by the investigator on assessment administration and data recording. Assessment components were explained and demonstrated to the graduate student clinicians by the investigator prior to initiating the research study and written instructions regarding administration of all assessment components were in view of the graduate student clinicians during each assessment. Prior to initiating the research study, the investigator co-administered the assessment protocol with the graduate student clinicians for a case study child, who was not one of the research children. Feedback was provided on both methods of elicitation and data collection. For all subsequent evaluations, the investigator was in attendance for the entirety of the sessions and able to supervise all assessment procedures.

Prior to initiating the assessment protocol, parents of children were given an Informed Consent form to sign and a Case History Questionnaire (Appendix A) to complete. The questionnaire was distributed to parents to elicit responses to questions the investigator used for selection criteria of children described later. Each child was also asked to provide a Verbal Assent (age 5;0 – 6;11) or sign a Child Assent (age 7;0 – 8;11) form before testing began.

A comprehensive assessment battery was used to help determine eligibility for inclusion in the study. After passing a hearing screening, children were administered standardized articulation and phonology assessment measures (Articulation Single-Word Production and Phonology Single-Word Production subtests) using the DEAP (Dodd et al., 2006). Both subtests possess adequate sensitivity and specificity (Dodd et al., 2006). The DEAP assessment tool also provided opportunities to assess consistency of production at word level (Word Inconsistency Assessment subtest) and screen oral motor skills (Oral Motor Screen subtest). The DEAP, as a diagnostic tool for sCAS, has been used in recent research studies (Preston, Brick, & Landi, 2013; Dale & Hayden, 2013). Administration of two subtests of the *Clinical Evaluation of Language Fundamentals-4* (CELF-4; Semel, Wiig, & Secord, 2003) also provided assessment data. The

Number Repetition subtest offered information relative to auditory/verbal working memory, and the Recalling Sentences subtest provided stimuli that were used to determine the presence/absence of oral groping behaviors/silent posturing observed during attempts to imitate words and phrases in longer strings of words. Children's responses to three subtests of the *Kaufman Speech Praxis Test for Children* (KSPT; Kaufman, 1995) were phonetically transcribed to determine the frequency of occurrences of simple syllable shapes (Complex Bisyllabics subtest and Polysyllabic Synthesis/Sequencing subtest) and the number of errors on longer units of speech output (Length and Complexity subtest). Administration of Maximum Performance Tasks (Thoonen et al., 1999) provided information regarding absent/shortened frication (Maximum Fricative Duration) and diadochokinetic rate (Maximum Repetition Rate of trisyllabic sequences). Finally, a wordless picture book (*Frog, Where are You?*; Mayer, 1969) was used to elicit a connected speech sample that provided assessment information regarding production of suprasegmental features.

Following verification of selection criteria via the Case History Questionnaire and a qualifying score on the DEAP Phonology Single-Word Production subtest (see below), eligible children were asked to return for another 60-minute session. Children meeting eligibility requirements returned within 1 to 30 days subsequent to the initial assessment session, with the exception of three children affected by inclement weather, holiday observances, and illness. The second session for these children occurred 35 to 66 days following the first session. Subsequent to completion of all components of the assessment battery at the beginning of session two, children then engaged in the experimental task, a perceptual task during which the children listened to the presentation of pairs of non-word tokens (digital recordings of single syllable /bɑ/ varying in vowel duration only) and indicated if the pairs of stimuli were the *same* or *different*.

The two assessment sessions took place at a quiet location, convenient for the parents. All children were tested individually and each session was videotaped.

2.2.3 Inclusionary criteria.

Children's selection criteria were carefully identified. This was particularly important in the case of the sCAS children to be included in the study, given the reported variable nature of sCAS characteristics seen in the same children over time (Shriberg, Campbell et al., 2003; Lewis et al., 2004). All children were required to meet the following selection criteria (see Appendix B): Reside in a monolingual, English-speaking home; pass a hearing screening; exhibit non-stuttered speech; and present with no history of hearing loss, acquired speech disorder, neurological disorder/syndrome, or significant history of otitis media with effusion (defined as ≥ six episodes during the first three years of life) (Nittrouer, Lowenstein, & Tarr, 2013). No child was included in this study if he was suspected of having intellectual deficits by parents or teachers and if he had been referred for psychological testing by school personnel because of concerns regarding cognitive skills. Observations of the children by the graduate clinicians and researcher confirmed no concerns relative to level of cognitive function.

The Number Repetition subtest of the CELF-4 was administered to all children to determine if their auditory memory was sufficient to deal with the experimental task so as not to confound results because of reduced auditory memory abilities. Factors such as attention and memory skills might influence duration discrimination (Rammsayer & Brandler, 2004). Due to the possibility that speech perception and/or short-term memory deficits exist in children with developmental speech disorders (Kenney, Barac-Cikoja, Finnegan, Jeffries, & Ludlow, 2006) and speech processing deficits related to auditory-perceptual encoding, memory, and transcoding can occur in children diagnosed with sCAS (Shriberg et al., 2012), documentation of all of the children's ability to remember auditory stimuli of sufficient length to undertake the experimental task was requisite. For the experimental perceptual task in this study, the longest stimuli pair participants heard was 1580 ms in length (standard stimulus = 250 ms + 800 ms pause + a comparison stimulus not exceeding 530 ms). Therefore, if children demonstrated the ability to accurately repeat both of the two-digit test items on the CELF-4 Number Repetition subtest (presented at the prescribed rate of one digit per second, for a total length of 2000 ms), it was assumed memory skills were adequate for the perceptual task of this study. All children in this study correctly repeated both two-digit stimuli of the CELF-4 Number Repetition subtest.

The children in the TD group were required to have a scaled score of ≥8 on the Phonology Single-Word Production subtest of the DEAP and present with no more than one sCAS characteristic listed in Table 2.2. In addition to being between 5;0 and 6;11, children in the sCAS group were required to have a scaled score of ≤5 on the Phonology Single-Word Production subtest and present with ≥5 (≥45%) of 11 sCAS characteristics (Davis et al., 1998; ASHA, 2007; Burns, 2011), per Table 2.2. In addition to the relevant sCAS characteristics, Table 2.2 presents the assessment procedure for each characteristic and the criterion/criteria used in this study. Of the 42 children initially assessed for the study, 32 were found eligible for inclusion in this study. Nine children did not meet inclusionary criteria and one child withdrew from the study.

The original 11-characteristic diagnostic checklist by Davis et al. (1998) (see Appendix C) has been utilized as a participant-selection tool in numerous research studies (Peter & Stoel-Gammon, 2008; Peter & Stoel-Gammon, 2005; Munson et al., 2003; Skinder-Meredith et al., 2001; and Skinder et al., 1999). Similar subject selection criteria have been reported by Forrest (2003), Shriberg, Campbell et al. (2003), and Davis and Velleman (2000).

The literature does not provide specific quantification for many of the criteria for participant grouping with regard to lists of sCAS characteristics. Burns (2011) reported, "To keep the clinical data simple enough to be applied in clinical settings, I simply noted whether a characteristic was present (+) during the [spontaneous speech] sample or whether specific problems were evident on the oral motor and word and non-word repetition tasks" (p. 21). Ozanne (1995), too, implemented a binary rating system (present or absent) for each of 18 sCAS characteristics in her research. Similarly, Peter and Stoel-Gammon (2008) determined the presence/absence of 11 common sCAS features for participants in their study via clinical

Table 2.2. sCAS characteristics, assessment procedure, and criteria/criterion.

sCAS Characteristics	Tool	Analysis
1. Frequent omission errors (Davis et al., 1998; Burns, 2011).	*Diagnostic Evaluation of Articulation and Phonology* (DEAP, Dodd, Hua, Crosbie, Holm, & Ozanne, 2006).	Inspection of phonetic transcriptions of DEAP Phonology Single-Word Production subtest.
	Criteria/Criterion = Investigator **sCAS**: ≥30% occurrence for one of the following or ≥15% for two or more of the following: Cluster Reduction, Weak Syllable Deletion, Final Consonant Deletion, Initial Consonant Deletion, and Medial Consonant Deletion **TD**: <10% occurrence for one of the following or <5% for two of the following: Cluster Reduction, Weak Syllable Deletion, Final Consonant Deletion, Initial Consonant Deletion, and Medial Consonant Deletion.	
2. High incidence of vowel errors (Davis et al., 1998; ASHA, 2007; Burns, 2011).	DEAP (Dodd et al., 2006).	Inspection of phonetic transcriptions of DEAP Articulation Single-Word Production subtest.
	Criteria/Criterion = DEAP **sCAS**: PVC score = ≤85% **TD**: PVC score = >85%.	
3. Inconsistent articulation errors (Davis et al., 1998; ASHA, 2007).	DEAP (Dodd et al., 2006).	Results of DEAP Word Inconsistency Assessment subtest.
	Criteria/Criterion = DEAP **sCAS**: ≥40% score **TD**: <40% score.	
4. Suprasegmental errors, especially lengthened and disrupted coarticulatory transitions between sounds and syllables and lexical stress errors (Davis et al., 1998; ASHA, 2007; Burns, 2011).	Connected Speech (*Frog, Where are You?*; Mayer, 1969).	Perceptual rating of speech production within a narrative or conversational sample.
	Criteria/Criterion = Investigator and SLP **sCAS**: Presence of suprasegmental errors **TD**: Absence of suprasegmental errors.	

5. Groping behaviors/silent posturing observed during attempts to imitate words/phrases (Davis et al., 1998).	*Clinical Evaluation of Language Fundamentals-4* (CELF-4; Semel, Wiig, & Secord, 2003).	Videotaped observation of groping behaviors/silent posturing (Hall, 2007) during CELF-4 Recalling Sentences subtest (initial 10 items) (Peter & Stoel-Gammon, 2008).
	Criteria/Criterion = Investigator and SLP **sCAS**: Presence of groping behaviors/silent posturing **TD**: Absence of groping behaviors/silent posturing.	
6. Predominant use of simple syllable shapes, especially CV, V, CVC (Davis et al., 1998).	*Kaufman Speech Praxis Test for Children* (KSPT; Kaufman, 1995).	Inspection of phonetic transcriptions of KSPT Complex Bisyllabics subtest and Polysyllabic Synthesis/ Sequencing subtest.
	Criteria/Criterion = Investigator **sCAS**: Syllable reduction = ≥15% (≥5 syllables reduced out of 31) **TD**: Syllable reduction = <5% (≤1 syllable reduced out of 31).	
7. Limited consonant and vowel repertoire (Davis et al., 1998).	DEAP (Dodd et al., 2006).	Inspection of phonetic transcriptions of DEAP Articulation Single-Word Production subtest.
	Criteria/Criterion = DEAP **sCAS**: PPC score = ≤85% (≥1 *SD* below mean) **TD**: PPC score = >85% (<1 *SD* below mean).	
8. Increased errors on longer units of speech output (Davis et al., 1998).	KSPT (Kaufman, 1995).	Inspection of phonetic transcriptions of KSPT Length and Complexity subtest.
	Criteria/Criterion = Investigator **sCAS**: Error = ≥40% (≥2 in error out of 5 items) **TD**: Error = ≤20% (≤1 in error out of 5 items).	
9. Impaired volitional oral movements (Davis et al., 1998).	DEAP (Dodd et al., 2006).	Observation of DEAP Oral Motor Screen subtest.
	Criteria/Criterion = DEAP **sCAS**: Does not meet criterion score based on C.A. **TD**: Meets criterion score based on C.A.	
10. Absent or shortened frication/affrication (Burns, 2011).	DEAP (Dodd et al., 2006) or Maximum Performance Tasks (Thoonen, Maassen, Gabreëls, & Schreuder, 1999).	Inspection of phonetic transcriptions of DEAP Phonology Single-Word Production subtest for absence of fricatives and/or affricates or Maximum Fricative Duration (MFD) = grand mean of /f/, /s/, and /z/.

	Criteria/Criterion = Investigator; modified Thoonen et al. (1999) to accommodate Dutch versus English language **sCAS**: ≥25% occurrence of fricative/affricate errors (≥11 in error out of 44 items) **TD**: <10% occurrence of fricative/affricate errors (<4 in error out of 44 items) or **sCAS**: MFD = ≤4 seconds **TD**: MFD = >4.5 seconds.	
11. Reduced diadochokinetic rate (Davis et al., 1998).	Maximum Performance Tasks (Thoonen et al., 1999).	Maximum Repetition Rate of trisyllabic sequences (MRRtri) = /pʌtəkə/.
	Criteria/Criterion = Investigator; modified Thoonen et al. (1999) to accommodate Dutch versus English language **sCAS**: MRRtri = incorrect sequence or ≤2.0 syllables/second **TD**: MRRtri = correct sequence and >3.0 syllables/second.	
Note. PVC = percentage of vowels correct; CV = consonant-vowel; PPC = percentage of phonemes correct; C.A. = chronological age.		

judgment. These authors conceded, however, that

> Although these features are widely used, quantitative guidelines to indicate presence or absence of a given feature are provided for only a subset of these features; e.g. limited consonant and vowel inventory is described as less complete than the inventory of phonemes produced in adult English, while frequent omission errors are not further quantified. Furthermore, there are no published criteria on how many characteristics are required for a CAS diagnosis. (p. 177)

Therefore, the investigator used the literature, guidelines from published assessment instruments, clinical practices, and/or common empirical practices to determine specific criteria/criterion for evaluating the presence/absence for each of the 11 characteristics.

Based on results of the comprehensive assessment battery, the presence/absence of the 11 characteristics for each child (Appendix D) was determined by the investigator, two graduate research assistants (RA1 and RA2), and an experienced SLP using the criteria presented in Table 2.2. In the absence of more precise guidelines from the literature as to how many at what degree or which of 11 characteristics needed to be present for a diagnosis of sCAS, the investigator determined that for a child to be identified as having sCAS for this study, the child needed to present with approximately 50% (or in this case five) of the 11 characteristics, per the specified criterion/criteria determined for the characteristics.

Results of a case study, who was not one of the research children, served to establish a training scoring protocol for the two graduate research assistants and the SLP. RA1 was trained on 1) determination of severity of speech sound production and 2) eight of the 11 characteristics (omission errors, vowel errors, inconsistency, use of simple syllable shapes, limited consonant and vowel repertoires, increased errors on longer units of speech, impaired volitional oral movements, and absent or shortened frication/affrication). RA2 was trained on one of the 11 characteristics, reduced diadochokinetic rate. After the investigator and graduate research assistants engaged in simultaneous observation and analysis of videotaped assessment results of the case study, the two graduate research assistants were determined by the investigator to be sufficiently accurate in scoring and proceeded to score the assessment performances of the first four research children (2 TD and 2 sCAS). In order to ensure continuing scoring at criterion, the graduate research assistants' scoring for these four children was compared to the scoring of the same children by the investigator. Details regarding methods for determination of inter- and intra-rater agreement are described in Appendix E. Reliability between the investigator and the two graduate research assistants for the first four children was judged to be sufficiently robust so an additional four randomly-selected children were identified for determination of inter-rater agreement. Subsequent comparison of the investigator's analyses to the two graduate research assistants' analyses for these four additional randomly-selected children (two children from each of the two diagnostic groups - sCAS and TD) indicated a similar degree of robustness (see Appendix F for inter-rater agreement for all eight children). A summary of inter- and intra-rater reliability results is presented here.

For inter-rater reliability for determining severity of speech sound production and derived percentages for three of the 11 characteristics, the mean percent agreement for RA1 and the investigator ranged from 93% to 98% for each of these characteristics. For reliability for determining categorical judgments of agreement/disagreement for four of the 11 characteristics, agreement for RA1 and the investigator ranged from 75% (agreement for 6 of 8 children) to 100% (agreement for 8 of 8 children). For determining reliability for one characteristic, absent or shortened frication/affrication, both a derived percentage (% affricate/fricative errors) and a categorical judgment of agreement/disagreement (MFD) were required for RA1 and the investigator. Mean percent agreement was 97% for the % affricate/fricative errors and 100% for MFD. Finally, for assigned categorical judgments of agreement/disagreement for one of the 11 characteristics, reduced diadochokinetic rate, agreement for RA2 and the investigator equaled 100%.

Intra-rater agreement for each of the two graduate research assistants was also obtained for the same four randomly-selected research children (two children from each of the two diagnostic groups - sCAS and TD) > 3 months following the initial scoring of the assessment results. Mean percent intra-rater agreement for RA1 ranged from 90% to 100% for derived percentages of severity and three of the 11 characteristics; percent intra-rater agreement for RA1 for categorical judgments of agreement/disagreement for four of the 11 characteristics equaled 100%, and mean percent intra-rater agreement for RA1 was 98% for the % affricate/fricative errors and 100% agreement for MFD. Finally, for assigned categorical judgments of agreement/disagreement for one of the 11 characteristics, reduced diadochokinetic rate, intra-rater agreement for RA2 equaled 100% (see Appendix G for intra-rater agreement for four randomly-selected children).

Of the 11 characteristics identified in Table 2.2, two (presence of suprasegmental errors and presence of groping behaviors/silent posturing) required video-analysis of each child's assessment by an external evaluator. An ASHA-certified and licensed SLP with 26 years of professional experience identified the presence or absence of these two communication characteristics in the children (see Appendix H for specific directions provided to the SLP). After the investigator and SLP engaged in simultaneous observation and analysis of videotaped assessment results of the case study, the SLP was determined by the investigator to be sufficiently accurate in scoring and proceeded to score from videotape observation the assessment performances of the first five (4 TD and 1 sCAS) research children, about whom the SLP had no prior knowledge (the SLP had referred and worked in her practice with some of the children). In order to ensure continuing scoring at criterion, the SLP's scoring for these five children was compared to the scoring of the same children by the investigator. For categorical judgments of presence/absence for these two characteristics, agreement for the SLP and the investigator was 80% (agreement for 4 of 5 children) for suprasegmental errors and 100% (agreement for 5 of 5 children) for groping behaviors/silent posturing. Subsequent comparison of the investigator to the SLP's analyses of six additional randomly-selected children for whom she had no prior knowledge (three children within each of the two diagnostic groups - sCAS and TD) indicated adequate accuracy of the children's scores as evidenced by 83% (agreement for 5 of 6 children) agreement for suprasegmental errors and 100% agreement (agreement for 6 of 6 children) for categorical judgments of the characteristics being present/absent (see Appendix F for inter-rater agreement for all eleven children).

Additionally, the same experienced SLP served to confirm or reject the group status (sCAS or TD) for a random number of children for whom she had no prior knowledge (8 of 32 children or 25%). Blind to the children's initial group assignment and while aware of the general area of the research but not the hypotheses, the SLP observed videotapes of the children's assessments, reviewed corresponding assessment data, and independently assigned these children to either the sCAS or TD group. Her group assignments were compared to the children's initial group placements. The frequency with which the SLP and the investigator agreed upon classification of each child was 100% (see Appendix F).

Data indicated adequate inter-rater agreement between the investigator and RA1, RA2, and the SLP for all measures of assessment. Adequate intra-rater agreement was observed for

RA1 and RA2, as well.

Appendix I presents details of the performances of each child on the assessment battery.

2.2.4 Stimuli.

Pairs of single-syllable non-words, consisting of /bɑ/ and differing only by vowel length, were used to assess the children's ability to determine if the pairs when heard were the same or different. Based on research by Huggins (1972) showing adults' increased sensitivity to vowel durational differences, as compared to changes in consonant durations, the duration of the vowel within the single syllable non-words was altered and the duration of the consonant remained constant. Vowel durational differences were also targeted due to the presence of vowel errors reported within the sCAS population (Davis et al., 1998; McCabe et al., 1998; Shriberg et al., 1997a). Use of non-words minimizes the potential effect of possible prior learning associated with real words.

As indicated, phonemes within the non-words in the present study were limited to the consonant /b/ and the vowel /ɑ/. Stimuli were derived by systematically varying the duration of the vowel /ɑ/ within a single-syllable, CV non-word (/bɑ/). The phonemes /b/ and /ɑ/ were selected, in part, based on the research of Shriberg, Lohmeier, Campbell, Dollaghan, Green, and Moore (2009) that revealed 100% of children in their study, ranging in age from 3;0 to 5;11 with speech delay of unknown origin and including those with a moderate-severe and severe speech sound disorder, had accurate production of the phonemes /b, m, n, ɑ/ in their phonetic inventory. According to these researchers, methodological confounds occur when participants are required to deal with non-words which contain target speech sounds not found in their phonetic inventory or that are typically misarticulated. Incorrect sound production can also be associated with auditory misperception (Maassen et al., 2003). The bilabial consonant /b/ was also chosen because of the negligible impact the phoneme has on the duration of an adjacent vowel (Nishi, Strange, Akahane-Yamada, Kubo, & Trent-Brown, 2008). The consonant was restricted to a single phoneme so as to not include an additional variable potentially affecting performance.

Acoustic stimuli (single-syllable productions of /bɑ/) were created using naturally-produced speech which was then digitally altered in order to assess children's ability to determine longer versus shorter vowel durations. Voice samples were recorded in a single-walled, sound-attenuating booth. A handheld microphone, placed six inches from the speaker's mouth, was used to record at a 48 kHz sampling rate. Altered speech stimuli were used in an attempt to retain some acoustic qualities of the speech signal while having the ability to manipulate other targeted acoustic characteristics of the signal (Kraus, Koch, McGee, Nicol, & Cunningham, 1999). Across tokens, vowel durations were systematically varied by lengthening or shortening the nuclei of a naturally-produced vowel using PSOLA resynthesis in Praat (Boersma &Weenink, 2011); fundamental frequency and amplitude measures were controlled to remain uniform. The fundamental frequency for /ɑ/ was systematically controlled to be 89 Hz, while the mean amplitude was 76 dB for /ɑ/. The altered stimuli were created by taking a single syllable (/bɑ/) pronounced by a male speaker of American Standard English and obtaining its designated duration via digital manipulation. A male speaker was selected since all participants were male and gender differences in vowel duration have been identified (Simpson & Ericsdotter, 2003; Clopper et al., 2005; Neel, 2008). Signal-editing functions in Praat were used to separate the consonant /b/ and the vowel /ɑ/. These syllables were then re-assembled so as to create eight single-syllable CV stimuli. The resulting stimulus items, though monotonic, were screened for naturalness and to confirm that there were no audible discontinuities in the signal. Once the investigator and one of the graduate research assistants involved in the synthesis of the stimuli were satisfied with the naturalness and continuity, the opinions of five additional listeners were obtained. Listeners included three undergraduate students and two faculty (Ludlow, Reed) in the

Department of Communication Sciences and Disorders at James Madison University. If feedback indicated the stimulus items appeared artificial or distracting, the stimuli were modified until a consensus of the listeners was reached. Subsequently, the altered non-word stimuli were digitally saved as individual WAV computer files.

For the stimuli, the duration of the consonant /b/ was 42 ms (defined by the time from the start of occlusion to the beginning of the waveform following the burst) and the duration of /ɑ/ (defined by the time interval between the onset and offset of the vowel, including the entire gesture from release/burst of the preceding consonant /b/ to the signal of decay) ranged from 208 ms to 488 ms in increments of 40 ms (eight intervals = 208 ms, 248 ms, 288 ms, 328 ms, 368 ms, 408 ms, 448 ms, and 488 ms). In addition to consensus of the five listeners, determination of the varied vowel lengths by the researcher was influenced by the duration of English syllables reported in the literature. The average duration of stressed syllables in English has been identified as 300 ms (Fant, Kruckenberg, & Nord, 1991) and 294 ms (Campbell, 1989), while the average duration of unstressed syllables in English has been identified as 140 ms (Fant et al., 1991) and 138 ms (Campbell, 1989).

Based on the work of Lehiste (1970), the length of the eight vowel durations for stimuli varied in increments of 40 ms. Lehiste noted that the JND for listeners to detect changes in the duration of a phoneme (approximately 30 to 300 ms in length) varied between 10 and 40 ms. And, since "It has been shown that young children require more acoustic information than adults to perform some speech perception tasks" (Elliott, Busse, Partridge, Rupert, & DeGraaff, 1986, p. 628), the upper limit for JND cited by Lehiste (1970) was utilized. A step-size interval of 40 ms was also used by Elfenbein et al. (1993) in their study of duration discrimination in children. Vowel length was capped at 488 ms based on research by Rammsayer and Lima (1991) that indicated cognitive factors, such as attention and memory, are more influential in temporal processing tasks where stimulus intervals exceed 500 ms. Thus, the longest stimulus was 530 ms, comprised of 42 ms /b/ and 488 ms /ɑ/; the shortest stimulus duration was 250 ms, comprised of the 42 ms /b/ and a 208 ms /ɑ/. Six additional syllable stimuli were created with 40 ms vowel duration differences between the shortest and longest syllables (i.e., 290 ms, 330 ms, 370 ms, 410 ms, 450 ms, and 490 ms).

2.2.5 Protocol.

All children were asked to judge whether pairs of these non-word single syllable tokens (digital recordings of single syllable /bɑ/ varying in vowel duration only) were the same or different. Given that in previous duration discrimination studies with children (Jensen & Neff, 1993; Elfenbein et al., 1993) auditory memory demands may have confounded results, a "same-different" task paradigm was used in order to reduce memory load (Wood, 1976). This task was congruent with the auditory memory skills demonstrated by children in their assessment task. Similar to methodology used by Himpel et al. (2009), the present study incorporated paired stimuli within a same-different task. However, based on the younger chronological ages of the children in this study, instructions were modified to have the children simply identify the two stimuli within each pair as "same" or "different", eliminating the semantic, as well as perceptual, element of "longer" that was present in the research of Himpel et al. (2009). Modeled after the works of Nittrouer, Shune, and Lowenstein (2011) and Nittrouer et al. (2013) using an AX procedure, children in the current study compared a stimulus (X), which varied across trials, with a constant standard (A). The standard A interval was the 'anchor' stimulus and always placed in the first position in each pair (Maassen et al., 2003; Kraus et al., 1999). Furthermore, the A interval was the stimulus with the shortest vowel duration (208 ms) and the X interval was the comparison stimuli (i.e., vowel duration = 208 ms, 248 ms, 288 ms, 328 ms, 368 ms, 408 ms, 448 ms, or 488 ms). Stimuli reflected the systematic 40 ms variation in vowel duration, creating pairs of syllables with values that were "not reliably discriminated to easily discriminated"

(Morrongiello & Trehub, 1987, p. 416). Each of the eight stimuli, including the standard 250 ms with the 208 ms vowel plus 42 ms /b/, was presented in the pair as a comparison stimulus (X).

Paired stimuli were presented to each child, under uniform conditions, using acoustic noise cancelling headphones (Bose QC15) at a comfortable intensity level determined by the investigator and graduate research assistants (Level 10 on Dell Latitude 13 computer). Children's observed responses during training and task items indicated they easily heard the stimuli.

Following procedures similar to that of Nittrouer et al. (2011, p. 769), the following instructions were presented to the children:

> Now you will hear a robot say two short words at a time. Decide if the words you hear are the *same* word or two *different* words. If the words are the same, point to this picture of two black squares and say "same". If the words are different, point to this picture of a black square and red circle and say "different". Listen carefully because the words cannot be repeated.

2.2.6 Training.

Several training trials were presented to confirm that all children understood the concepts of same and different (Nittrouer et al., 2011). Before using acoustic stimuli, children were shown drawings of same and different objects and asked to indicate whether the two objects on each of six cards were the same or different. Feedback was provided. Next, drawings of same and different geometric shapes were shown to children and they were asked to indicate if the two shapes on each of four cards were the same or different. This time, no feedback was provided. Finally, prior to the introduction of acoustic stimuli, children were shown a card with two black squares on one half and a red circle and a black square on the other half and asked to point to "same" and to "different". Response accuracy (100%) for identification of same versus different with visual stimuli was required prior to progressing to training with auditory stimuli.

Training continued as children listened to five digitally-recorded practice trials with the

CV syllable /bo/. For training purposes, /bo/ replaced /bɑ/ and acoustic manipulations were the same as the experimental stimuli. American English vowels vary systematically in intrinsic duration (Peterson & Lehiste, 1960) and this parameter was considered in the selection of the vowel for training purposes. Both vowels, /ɑ/ and /o/, share relatively equal intrinsic durations (Peterson & Lehiste, 1960; Pollock et al., 1993; Munson et al., 2003). The shortest training stimulus was 250 ms, comprised of the 42 ms /b/ and a 208 ms /o/, and the longest training stimulus was 530 ms, comprised of 42 ms /b/ and 488 ms /o/. Training pairs consisted of the standard (A) combined with the shortest duration /bo/ (250 ms), an intermediate duration /bo/ (370 ms), and the maximally long duration /bo/ (530 ms). Fundamental frequency and amplitude measures were controlled to remain uniform. Children received feedback from the examiner following each response (if correct, the examiner said, "That's right, those words sounded the same"; if incorrect, the examiner said, "No, those words sounded different"). If the child did not successfully perform the task (100% accuracy), the five practice items were repeated by the examiner up to five times until the participant demonstrated understanding of the testing protocol. If criterion was not achieved, the experimental task was not administered. If criterion was achieved, training then continued with the same five practice trials; however, feedback for this training was not provided following the child's responses. If the child did not successfully perform the task with 80% accuracy, the five practice items were repeated by the examiner up to five times. Again, if the child was still not able to achieve criterion, the experimental task was not administered. During training, if any child showed visible distress, all activities ceased. If criterion was achieved, the experimental task was then administered.

2.2.7 Administration of experimental stimuli.

Following successful training, each child was presented 80 pairs of non-word single syllable tokens (digital recordings of single syllable /bɑ/ varying in vowel duration only). Each of eight novel stimulus pairs was presented randomly 10 times throughout the task to determine difference duration (Nittrouer et al., 2011; Nittrouer et al., 2013). Thus, 10 trials of each pair (250-250 ms, 250-290 ms, 250-330 ms, 250-370 ms, 250-410 ms, 250-450 ms, 250-490 ms, and 250-530 ms) were presented. Repetition of stimuli by the examiner was not allowed, except for extenuating circumstances (e.g., siren sounding).

Each of the eight different stimulus pairs was presented twice within blocks comprised of 16 pairs. Thus, there were five blocks of stimulus pairs. Within each of the five blocks, the order of presentation of each pair was randomized with the exception of the first pair. To facilitate detection by the children, the initial presentation of stimuli within each block was either the standard (A) combined with the shortest duration /bɑ/ of 250 ms (same) or the maximally longest duration /bɑ/ of 530 ms. Presentation of each of the five blocks was also randomized across children. See Appendix J for a listing of the 80 stimulus pair trials.

The software Premiere (Adobe Premiere Elements 10, 2011) was used to create the stimulus pairs and computer-aided graphics. Within each trial, the two tokens comprising each stimulus pair were separated by an intra-pair interval of 800 ms. The literature shows variability in the interval length selected by researchers. For example, Nittrouer et al. (2011) used a 450 ms intra-pair interval, Elliott et al. (1986) incorporated a 500 ms intra-pair interval, and Hämäläinen et al. (2009) used a 1000 ms intra-pair interval. The length of the intra-pair interval for this study was selected based both on the researcher's attempt to minimize the demand of the interval length on children's auditory memory and on the consensus of the five listeners. As a child listened to each trial and responded as to whether the tokens were the same or different, a still graphic of a boy listening (hand cupped to his ear) was presented on a computer screen. This served as a visual cue for a child to pay attention and listen. Following this interval (approximately 8 seconds), a graphic (e.g., a frog leaping on lilypads) appeared on the computer screen to help maintain interest. This animated graphic signaled to the child both his completion of one trial (frog leaps to the subsequent lilypad) and the number of trials remaining in that block (evidenced by the number of lilypads remaining in the pond). See Appendix K for scripted instructions to children and a sample data sheet. The next trial automatically started approximately 3 seconds after the visual feedback. However, presentation of the subsequent trial could be manually paused by the administrator if the child did not initiate a response within the designated time interval or appeared not to be attending. Following the presentation of one block, children engaged in stretch breaks before the experimenter proceeded to the next block of stimuli. If a child presented with any verbal/nonverbal signs of fatigue, distress, or frustration or expressed a desire to stop during administration, the investigator stopped the task and the child and the investigator/graduate student engaged in a short break consisting of a trip to the bathroom or water fountain, or a gross motor activity like jumping jacks. Following the break, the experimental task resumed with the child's consent.

2.2.8 Scoring of responses to the experimental task.

Viewing either live or videotaped presentations, two judges independently recorded each child's responses on record sheets, in addition to the investigator's recording of responses. Judges, who were graduate student clinicians in the Department of Communication Sciences and Disorders, were aware of the purpose of the study but did not know the group status of the children. The participating graduate student clinicians were trained by the investigator relative to recording *same* and *different* responses on the data sheets provided (Appendix K). Children's responses were entered as correct identification of same or different or incorrect identification of same or different. The percent correct identification of same/different responses constituted the dependent variable data. Recording responses of a case study, who was not identified as one of the research participants, served as training for the graduate student clinicians (e.g., in the case of an immediate self-correction by a child, the last response was scored in lieu of the initial response regardless of whether it was correct or not).

2.2.9 Data preparation and handling.

Data for each child were recorded on assessment protocol sheets by the investigator, graduate student clinicians, and two graduate research assistants and, subsequently, entered into a SPSS (Version 21.0) database by a different team of two graduate research assistants. Accuracy of data entry was then ensured by two undergraduate students reviewing all data in the database. Any discrepancies in data were resolved by the investigator reviewing the videotaped data.

Chapter 3 Results

Taking into account considerations such as sample sizes, distribution of the samples' mean responses, data types for the dependent variable, and homogeneity of variances, analyses of data at this point in this line of research consist of descriptive and nonparametric inferential statistics. The percent correct responses of same or different constituted the dependent variable.

3.1 Performances of Typically Developing Children

Once the investigator determined that the TD children could perform the experimental task, the focus of research addressed three hypotheses: 1) the children's performances at the extreme duration differences (i.e., pairs that are of the same duration, pairs that are toward the maximum duration differences) will demonstrate mostly correct responses, 2) as duration differences increase from the smallest difference in duration through the intermediate duration differences, the children's correct response rates will show improved performances as the duration differences increase, but variances will become large, and 3) the performances of the younger TD children will differ from the performances of the older TD children.

3.1.1 Trials to train to task.

The investigator considered that the performances of the children on the task designed to train them to undertake the experimental task was an indicator of whether or not they could perform the task. The training task was described in Chapter 2 and consisted of two parts, with the children needing to reach criterion on the first part to progress to the second and then to reach criterion on the second part to progress to the experimental task. On both parts one and two, the children needed to reach criterion by six or fewer trials. The number of trials required to train the TD children to criteria for proceeding to the experimental task is shown in Table 3.1.

As seen in this table, all but one of the 21 children reached criterion with one or two trials on part one and only one trial on part two of the training. The one unique child required four trials to reach criterion on part one but only one trial on part two. These observations suggest that the TD children had little difficulty learning to perform the experimental task.

Table 3.1. Number of trials required to train children (TD) to criteria for experimental task.

Child	C.A.	# of training trials required out of maximum 6 to reach criterion for Training Part 1 (verbal/visual feedback provided following each stimulus pair)	# of training trials required out of maximum 6 to reach criterion for Training Part 2 (no verbal/visual feedback provided following each stimulus pair)
01	5;0	2	1
02	5;4	2	1
03	5;5	1	1
04	5;6	1	1
05	5;7	1	1
06	5;9	2	1
07	5;10	1	1
08	6;0	2	1
09	6;1	2	1
10	6;2	1	1
11	6;4	2	1
12	6;6	1	1
13	6;11	2	1
14	6;11	1	1
15	7;2	1	1
16	7;4	2	1
17	7;4	1	1
18	7;6	4	1
19	8;1	1	1
20	8;5	1	1
21	8;8	1	1
		Mode = 1 (range = 1-4)	Mode = 1
Note. C.A. = chronological age in years/months.			

3.1.2 Experimental task performances.

All statistical analyses were performed using SPSS (Version 21.0).

3.1.2.1 Descriptive statistical analyses.

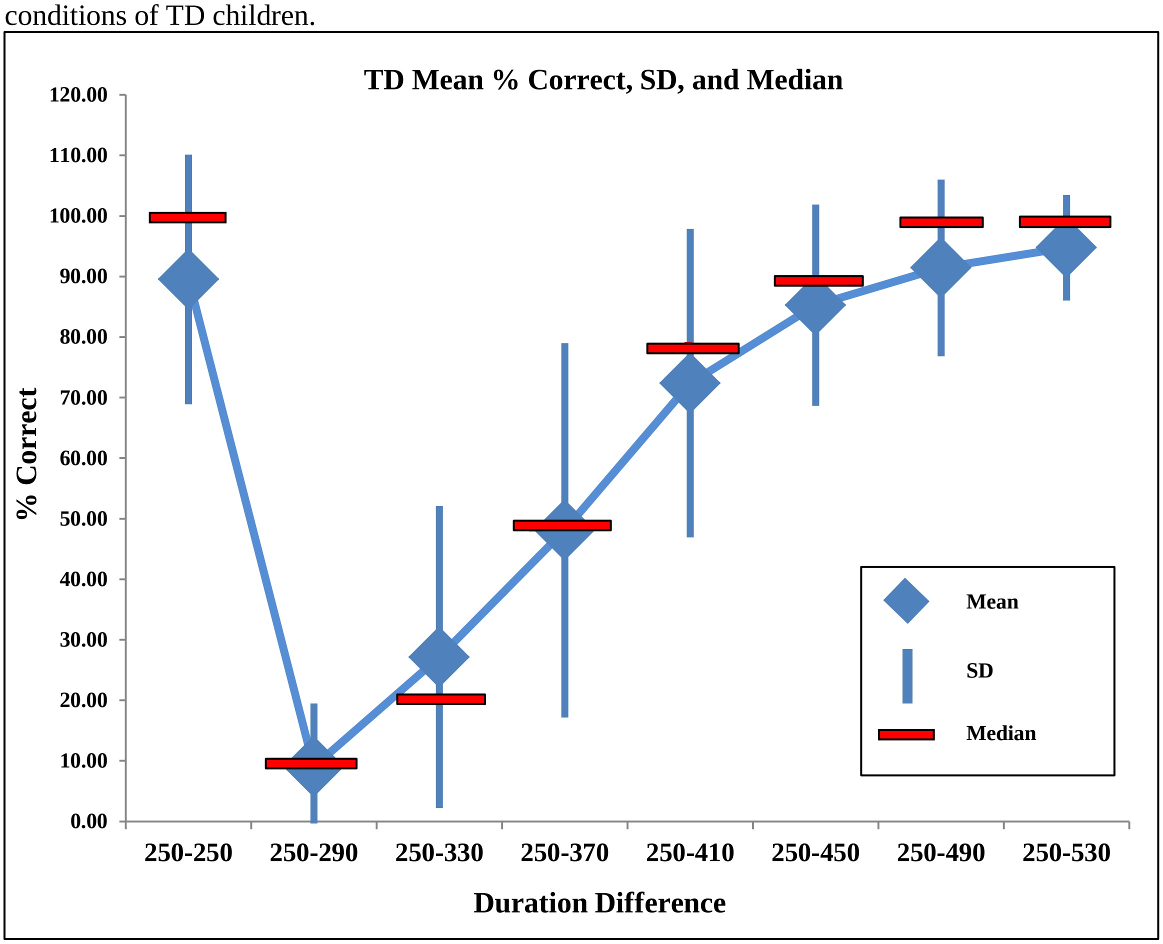
Table 3.2 shows the mean percent correct responses at each of the duration difference conditions, as well as the standard deviations (SD), medians, modes, and confidence intervals (CI) for the medians for all TD children. Figure 3.1 illustrates the mean percent correct responses and standard deviations at each duration difference, as well as the median scores.

At the 0 ms duration difference, that is, 250-250, the children achieved about a 90% correct response rate by identifying the stimulus pair as “same.” The median and mode for the 250-250 duration difference condition were each 100, with the scores ranging from 70% to 100%

Table 3.2. Mean percent correct responses, SDs, medians, modes, and CIs for 8 duration difference conditions of TD children.

	Duration Difference Condition							
	250-250	250-290	250-330	250-370	250-410	250-450	250-490	250-530
Mean	89.52 (20.61)	9.05 (10.44)	27.14 (24.93)	48.10 (30.92)	72.38 (25.48)	85.24 (16.62)	91.43 (14.59)	94.76 (8.73)
Median	100.00	10.00	20.00	50.00	80.00	90.00	100.00	100.00
Mode	100	0	10	20, 30, 50, 70, 90	90	100	100	100
CI for Median (95%)	90.00-100.00	.00-20.00	10.00-30.00	20.00-70.00	50.00-90.00	80.00-100.00	90.00-100.00	90.00-100.00

Figure 3.1. Mean percent correct responses, SDs, and medians for 8 duration difference conditions of TD children.



with the exception of one child who had a correct response rate of 10%. (This child's performance will be explained in the Discussion chapter; however, of note here, his performance was not due to difficulty determining differences with the durations.) The standard deviation was 20.61 which, upon inspection of individual scores, was influenced considerably by the result of the one child's performance.

At the greater duration differences beginning at the 250-410 duration difference condition, the children showed progressively improved response rates at the increasingly longer duration differences. The mean percent correct responses ranged from 72.38% at 250-410 to 94.76% at the 250-530 duration difference conditions. The medians and modes were consistent with the trajectory shown by the means. The size of the standard deviations decreased progressively from 250-410 to 250-530. These descriptive data support the predicted pattern of performance for the TD children with regard to the extreme duration difference conditions.

For performances at the 250-290, 250-330, and 250-370 duration difference conditions, the mean correct response rates increased progressively, ranging from 9.05% at 250-290 to 48.10% at 250-370. The standard deviations increased gradually as well, ranging from 10.44 to 30.92 across these same duration differences and suggesting increasing variability. The medians were similar to the mean scores and increased progressively from 10.00 at 250-290 to 50.00 at 250-370, indicating a 50-50 correct/incorrect response pattern at this intermediate duration difference. The modes for 250-290 and 250-330 were lower than the means or medians, at 0 and 10, respectively. As can be seen in Table 3.2, at the 250-370 duration difference condition, there were five modes, with three of the 21 children at each of the five modes. The standard deviation at the 250-370 duration difference condition was the largest of the eight duration difference conditions in the experimental task, including the same duration condition of 250-250. The TD children performed as predicted at the intermediate duration difference conditions.

Table 3.3 shows the mean percent correct responses at each of the duration difference conditions, as well as the standard deviations, medians, modes, and confidence intervals for the medians for all TD children by age group (5-, 6-, and 7/8-year olds). Figure 3.2 illustrates the mean percent correct responses by age at each duration difference condition. As evidenced by the

Table 3.3. Mean percent correct responses, SDs, medians, modes, and CIs for 8 duration difference conditions of TD children by age group.

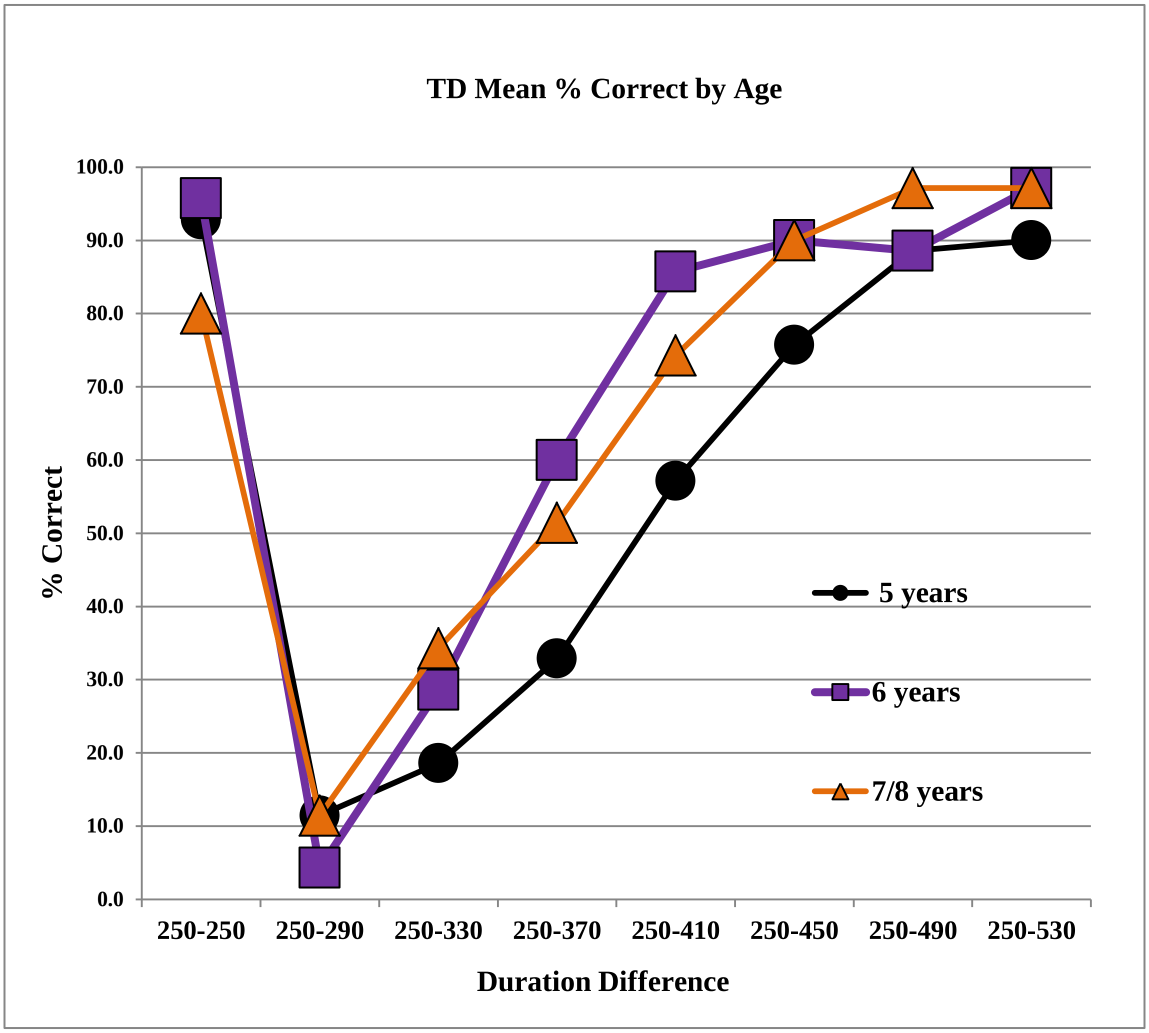
	Duration Difference Condition							
	250- 250	250- 290	250- 330	250- 370	250- 410	250- 450	250- 490	250- 530
Five-year-olds								
Mean	92.86 (11.13)	11.43 (12.15)	18.57 (14.64)	32.86 (24.30)	57.14 (22.15)	75.71 (7.87)	88.57 (12.15)	90.00 (11.55)
Median	100.00	10.00	30.00	30.00	50.00	80.00	90.00	90.00
Mode	100	0	30	20, 30	80	80	100	100
CI for Median (95%)	70.00- 100.00	.00- 30.00	.00- 30.00	.00- 70.00	30.00- 80.00	60.00- 80.00	70.00- 100.00	70.00- 100.00
Six-year-olds								
Mean	95.71	4.29	28.57	60.00	85.71	90.00	88.57	97.14
	(5.35)	(7.87)	(31.85)	(33.67)	(25.07)	(22.36)	(21.93)	(7.56)
Median	100.00	0.00	10.00	70.00	90.00	100.00	100.00	100.00
Mode	100	0	10	70	90, 100	100	100	100
CI for Median (95%)	90.00- 100.00	.00- 20.00	.00- 90.00	10.00- 100.00	30.00- 100.00	40.00- 100.00	40.00- 100.00	80.00- 100.00
Seven- Eight-year-olds								
Mean	80.00	11.43	34.29	51.43	74.29	90.00	97.14	97.14
	(33.17)	(10.69)	(26.37)	(31.85)	(23.71)	(14.14)	(4.88)	(4.88)
Median	100.00	10.00	20.00	50.00	80.00	100.00	100.00	100.00
Mode	100	10	10, 20, 50	50	90	100	100	100
CI for Median (95%)	10.00- 100.00	.00- 30.00	10.00- 80.00	.00- 90.00	30.00- 100.00	70.00- 100.00	90.00- 100.00	90.00- 100.00

plots, the three age groups share an upward trajectory as the duration difference increases up to the 250-450 duration difference condition. At that point, the three age groups start to plateau but with slightly different patterns.

3.1.2.2 Inferential statistical analyses.

To determine if there was an overall difference in percent correct performances of the children (n = 21) on the basis of the seven duration difference conditions, a Friedman F_r test statistic was computed. This test statistic is appropriate for repeated measures, that is, when samples are related. Because the question of interest was for performances related to differences

Figure 3.2. Mean percent correct responses for TD children by age for 8 duration difference conditions.



in duration, the children's performances for the 250-250 condition (0 ms difference) were not included in the analysis. The Friedman test was significant ($F_{r\,(6)} = 112.15$, $p < 0.05$). Therefore, six follow-up contrasts (250-290/250-330, 250-330/250-370, 250-370/250-410, 250-410/250-450, 250-450/250-490, 250-490/250-530) using the Wilcoxon Signed Ranks test were conducted. Because multiple tests were performed, the critical alpha level was adjusted using the Bonferroni correction procedure: .05 alpha/6 comparisons = .008. Table 3.4 shows the one-tailed results of these tests. The one-tailed direction was selected because the researcher hypothesized that the children's accuracy would improve as duration differences increased.

The results of the Wilcoxon Signed Ranks test were statistically significant at the .008 level for the first four duration difference comparisons (250-290/250-330, $T = 12.00$, $n = 21$, $p < .008$; 250-330/250-370, $T = 14.50$, $n = 21$, $p < .008$; 250-370/250-410, $T = 0.00$, $n = 21$, $p < .008$; and 250-410/250-450, $T = 4.50$, $n = 21$, $p < .008$), indicating the percent correct performances of the TD children were significantly different as the duration differences progressively increased from the 250-290 to the 250-450 duration difference condition. The effect sizes (Rosenthal, 1991) for the significant matched-pair conditions ranged from $r = .67$ (250-410/250-450) to $r = .84$ (250-370/250-410) ("$r = |z|/\sqrt{n}$" Field, 2005, p. 532). These effect sizes were large. In addition, the sums of the positive difference ranks were consistently larger than the negative difference ranks across all duration difference comparisons, showing improvement in children's ability to detect differences in duration as the length of the duration difference increased. For the 250-450/250-490 and 250-490/250-530 duration difference comparisons, the results indicated the percent correct performances of the TD children were not significantly different. However, at the 250-450/250-490 duration difference comparison, the results were significant at the .05 alpha level but not the adjusted alpha level of .008.

To determine if there was an overall difference in percent correct performances among the three age groups (5-year-olds, $n = 7$; 6-year-olds, $n = 7$; and 7/8-year-olds, $n = 7$) on the basis of the eight duration difference conditions, a Kruskal-Wallis *H*-test statistic was computed for each duration difference condition. This test statistic is appropriate for comparing three or more unrelated samples. Because the question of interest was for comparison of age-group performances, all duration difference conditions, including 250-250, were analyzed. Furthermore, because multiple tests were performed, the critical alpha level was adjusted using the Bonferroni correction procedure, resulting in an alpha = .006. The Kruskal-Wallis *H*-test was not significant at any duration difference condition, indicating that the performances of the three age groups were not statistically different at the .006 significance level (Table 3.5). Although none of the Kruskal-Wallis *H*-tests reached statistical significance at the alpha level .006, at the

Table 3.4. Results of the Wilcoxon Signed Ranks test for six follow-up contrasts for TD children (n = 21).

	Duration Difference Comparisons											
	250-290	250-330	250-330	250-370	250-370	250-410	250-410	250-450	250-450	250-490	250-490	250-530
T statistic	12.00		14.50		.00		4.50		9.00		9.00	
p	.0005*		.001*		.000*		.001*		.015		.098	
$\Sigma R+$	159.00		156.50		190.00		100.50		57.00		27.00	
$\Sigma R-$	12.00		14.50		.00		4.50		9.00		9.00	
Effect Size	0.72		0.68		0.84		0.67		---		---	
Confidence Interval (95%)	10.00-30.00		10.00-40.00		10.00-30.00		.00-20.00		.00-20.00		.00-10.00	

*Significant at .008

Table 3.5. Results of the Kruskal-Wallis *H*-test for percent correct performances for TD children by age (n = 21).

	Duration Difference Condition							
	250-250	250-290	250-330	250-370	250-410	250-450	250-490	250-530
H statistic	.438	2.321	1.105	2.577	6.209	6.110	1.949	2.977
p	.803	.313	.575	.276	.045	.047	.377	.226
df	2	2	2	2	2	2	2	2

duration differences of 250-410 and 250-450, statistical significance would have been achieved at the alpha level .05 if it had not been adjusted to account for multiple tests of significance. To explore these two trends toward statistical significance, follow-up Mann-Whitney *U*-tests for the two duration difference conditions (250-410, 250-450) were conducted.

Because multiple tests were performed using the Mann-Whitney *U*-statistic, the critical alpha level was adjusted using the Bonferroni correction procedure, resulting in an alpha = .017. Table 3.6 shows the one-tailed results of these tests. The one-tailed direction was selected because it was predicted that the older children (7/8-year-olds) would perform better than the younger children (5- and 6-year-olds). Comparing the 5- and 6-year-old groups, the results of the Mann-Whitney *U*-tests were statistically significant at the .017 level for both the 250-410 and 250-450 duration difference conditions, suggesting the 5-year-olds' performance was significantly different from that of the 6-year-olds (Table 3.3). The effect size (Rosenthal, 1991) was $r = .62$ for the 250-410 condition and $r = .63$ for the 250-450 condition ("$r = |z|/\sqrt{n}$" Field, 2005, p. 532). These effect sizes were large. The sum of ranks for the 6-year-olds was larger

Table 3.6. Results of the Mann-Whitney *U*-tests for comparisons by age (n = 7 per age group).

	250-410 Duration Difference Condition			**250-450 Duration Difference Condition**		
	5-yr-olds/ 6-yr-olds	5-yr-olds/ 7/8-yr-olds	6-yr-olds/ 7/8-yr-olds	5-yr-olds/ 6-yr-olds	5-yr-olds/ 7/8-yr-olds	6-yr-olds/ 7/8-yr-olds
U statistic	6.50	14.00	14.00	7.00	11.00	21.50
p	.009*	.105	.105	.013*	.049	.355
Effect Size	.62	---	---	.63	---	---
ΣR_1	34.50	42.00	63.00	35.00	39.00	55.50
ΣR_2	70.50	63.00	42.00	70.00	66.00	49.50

*Significant at .017

than that of the 5-year-olds, suggesting the performance of the 6-year-olds exceeded that of the 5-year-olds. No statistically significant differences were obtained for comparison of the 5- and 7/8-year-old groups and for the 6- and 7/8-year-old groups.

3.2 Performances of Children with sCAS

The second research question addressed the performances of 5- and 6-year-old children with sCAS compared to a subset of the TD children who were 5 and 6 years old. The investigator hypothesized that the sCAS children's performances would show different patterns on the duration difference task than that of their TD age peers. In the absence of any previous research known to the investigator that compared performances of young school-age children with and without sCAS on discrimination of duration differences, the investigator at this point in her line of research limited her hypothesis to predicting different patterns of performances, rather than predicting specific patterns of differences. The mean age of the TD children was 5 years; 11 months and the mean age of the children with sCAS was 5 years; 6 months. The result of an independent *t*-test, conducted to determine if there was a statistically significant difference in the ages of the two groups, was not significant at an alpha level of 0.05 ($t = 1.89$, $p = .07$).

3.2.1 Trials to train to task.

The investigator considered that the performances of the children on the task designed to train them to undertake the experimental task was an indication of whether or not they could perform the task. The number of trials required to train the children with sCAS to criteria for proceeding to the experimental task is shown in Table 3.7.

As seen in this table, two of the 11 children reached criterion with one or two trials on part one and six of the 11 children reached criterion with one or two trials on part two of the training. The performances of the children were bimodal (3,5), with a range of 1 to 6+ for part one. The mode for the number of training trials for part two was 1, also with a range of 1 to 6+. Two children, #29 and #31, did not successfully train to task and, consequently, did not proceed to the experimental task.

Table 3.7. Number of trials required to train children (sCAS) to criteria for experimental task.

Child	C.A.	# of training trials required out of maximum 6 to reach criterion for Training Part 1 (verbal/visual feedback provided following each stimulus pair)	# of training trials required out of maximum 6 to reach criterion for Training Part 2 (no verbal/visual feedback provided following each stimulus pair)
22	5;0	5	1
23	5;1	3	2
24	5;1	2	1
25	5;4	5	1
26	5;4	5	≤3*
27	5;5	6	1
28	5;6	3	3
29	5;8	>6 (Did not train)	N/A (Not administered)
30	5;10	3	3
31	6;5	6	>6 (Did not train)
32	6;10	1	1
		Mode = 3 and 5 (range = 1-6+)	Mode = 1 (range = 1-6+)

Note. C.A. = chronological age in years/months; *= researcher error.

3.2.2 Experimental task performances.

All statistical analyses were performed using SPSS (Version 21.0).

3.2.2.1 Descriptive statistical analyses.

Table 3.8 shows the mean percent correct responses at each of the duration difference conditions, as well as the standard deviations (SDs), medians, modes, and confidence intervals (CIs) for the medians for the children with sCAS who completed the experimental task (n = 9). Figure 3.3 illustrates the mean percent correct responses and standard deviations at each duration difference, as well as the median scores.

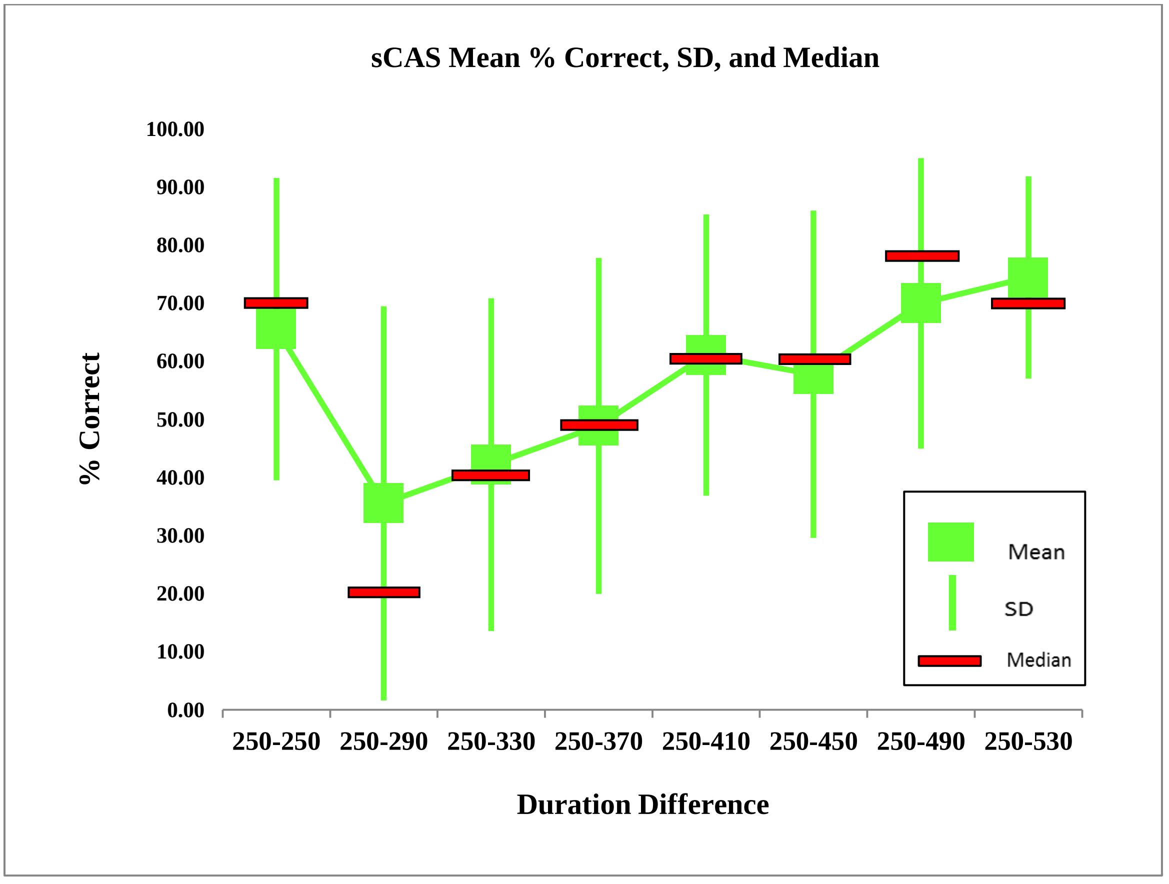
At the 0 ms duration difference, that is, 250-250, the children achieved approximately a 65% correct response rate by identifying the stimulus pair as “same.” The median for the 250-250 duration difference condition was 70.00 and the mode was split at 40 and 90, with the children’s scores ranging from 30% to 100%. The standard deviation was 26.03.

At the greater duration differences beginning at the 250-410 duration difference condition, the children did not show progressively improved response rates at the increasingly longer duration differences. In fact, their mean percent correct decreased from approximately

Table 3.8. Mean percent correct responses, SDs, medians, modes, and CIs for 8 duration difference conditions of children with sCAS.

	Duration Difference Condition							
	250-250	250-290	250-330	250-370	250-410	250-450	250-490	250-530
Mean	65.56 (26.03)	35.56 (33.95)	42.22 (28.63)	48.89 (28.92)	61.11 (24.21)	57.78 (28.19)	70.00 (25.00)	74.44 (17.40)
Median	70.00	20.00	40.00	50.00	60.00	60.00	80.00	70.00
Mode	40, 90	0, 20	10	50	30, 60, 80	80	80	70
CI for Median (95%)	40.00-90.00	.00-80.00	10.00-70.00	20.00-90.00	30.00-80.00	30.00-80.00	50.00-90.00	50.00-90.00

Figure 3.3. Mean percent correct responses, SDs, and medians for 8 duration difference conditions of children with sCAS.



61% at 250-410 to approximately 58% at the 250-450 duration difference condition. The range of mean percent correct responses was 61.11% at 250-410 to 74.44% at the 250-530 duration difference condition. The medians and modes at these greater duration differences also did not reflect an increasingly upward trajectory as evidenced by median values that were stationary at 60.00 at both the 250-410 and 250-450 duration difference conditions, rose to 80.00 at 250-490, and decreased to 70.00 at the most extreme duration difference condition of 250-530. Similarly, the mode was divided at 250-410 (30, 60, 80), remained stationary at 80 for the next two greater duration difference conditions of 250-450 and 250-490, and decreased to 70 at the largest 250-530 duration difference. The size of the standard deviations did not decrease progressively as the duration difference increased from the 250-410 condition to the 250-530 condition.

For performances at the intermediate 250-290, 250-330, and 250-370 duration difference conditions, the mean correct response rates increased progressively, ranging from 35.56% at 250-290 to 48.89% at 250-370. The medians also increased gradually from 20.00 at 250-290 to 50.00 at 250-370, indicating a 50-50 correct/incorrect response pattern at this intermediate duration difference. The modes evidenced little improvement as these intermediate duration differences increased, with bimodal values of 0 and 20 at 250-290 and 10 at the 250-330 condition. Mode values then escalated to 50 at the 250-370 duration difference condition. The large standard deviation at 250-290, 33.95, decreased only slightly at 250-330 and then showed no decline at the 250-370 duration difference condition. Of note, the standard deviation for the children with sCAS at the 250-290 condition was the largest of the eight duration difference conditions in the experimental task.

3.2.2.2 Inferential statistical analyses.

To determine if there was an overall difference in percent correct performances of the children (n =9) on the basis of the seven duration difference conditions, a Friedman F_r test statistic was computed. Because the question of interest was for performances related to differences in duration, the children's performances for the 250-250 condition (0 ms difference) were not included in the analysis. The Friedman test was significant ($F_{r\,(6)} = 22.29$, $p < 0.05$). Therefore, six follow-up contrasts (250-290/250-330, 250-330/250-370, 250-370/250-410, 250-410/250-450, 250-450/250-490, 250-490/250-530) using the Wilcoxon Signed Ranks test were conducted. Because multiple tests were performed, the critical alpha level was adjusted using the Bonferroni correction procedure: .05 alpha/6 comparisons = .008. Table 3.9 shows the two-tailed results of these tests. The two-tailed direction was selected because the investigator did not, at this point in her line of research, hypothesize beyond predicting different patterns of performance, rather than predicting the specific patterns of the differences.

The results of the Wilcoxon Signed Ranks test were not statistically significant at the .008 level for any of the six duration difference comparisons (250-290/250-330, 250-330/250-370, 250-370/250-410, 250-410/250-450, 250-450/250-490, 250-490/250-530), indicating the percent correct performances of the children with sCAS were not significantly different as the duration differences progressively increased from the 250-290 to the 250-530 duration difference conditions, although the sums of the positive difference ranks were larger than the negative difference ranks across all duration difference comparisons with the exception of one (250-410/250-450). Results did not suggest, therefore, consistent improvement in the children's ability to detect differences in duration as the length of the duration difference increased.

3.3 Comparison of a Subset of TD Children to Children with sCAS

When the investigator compared the performances of a subset (5- and 6-year-olds; n = 14) of the total TD group to a group of children with sCAS (5- and 6-year-olds; n = 11), three primary differences were observed. Distinctions were noted in areas relating to training trials; performance trajectory across duration difference conditions, including variance across duration difference conditions, confidence intervals, and >70% correct response thresholds; and default-to-different response patterns.

3.3.1 Trials to train to task.

Table 3.10 combines relevant data from Tables 3.1 and 3.7. As apparent in Table 3.10 comparing the performances of the TD 5- and 6-year-olds to those of the children with sCAS, all TD children (n = 14) were able to train to criteria with one or two trials on part one, with an equal number of children needing one and two trials (mode = 1, 2), and only one trial (mode = 1) on

Table 3.9. Results of the Wilcoxon Signed Ranks test for six follow-up contrasts for children with sCAS (n = 9).

	Duration Difference Comparisons					
	250-290 / 250-330	250-330 / 250-370	250-370 / 250-410	250-410 / 250-450	250-450 / 250-490	250-490 / 250-530
T statistic	5.00	10.00	9.50	16.00	8.00	5.50
p	.236	.256	.121	.778	.159	.588
$\Sigma R+$	16.00	26.00	35.50	16.00	28.00	9.50
$\Sigma R-$	5.00	10.00	9.50	20.00	8.00	5.50
Effect Size	---	---	---	---	---	---
Confidence Interval (95%)	.00-40.00	.00-30.00	10.00-40.00	.00-30.00	.00-50.00	.00-50.00

*Significant at .008

part two of the training. In contrast, the range of trials for part one and part two for children with sCAS (*n* = 11) was one to > six (mode = 3, 5 for part one and mode = 1 for part two). Furthermore, where all TD children in the 5- and 6-year-old groups were able to train to criteria, two of the children in the sCAS group were unable to train to criteria following six trials and therefore unable to proceed to the experimental task. Figure 3.4 illustrates the differences in the number of training trials required for the TD versus sCAS groups.

3.3.2 Experimental task performances.

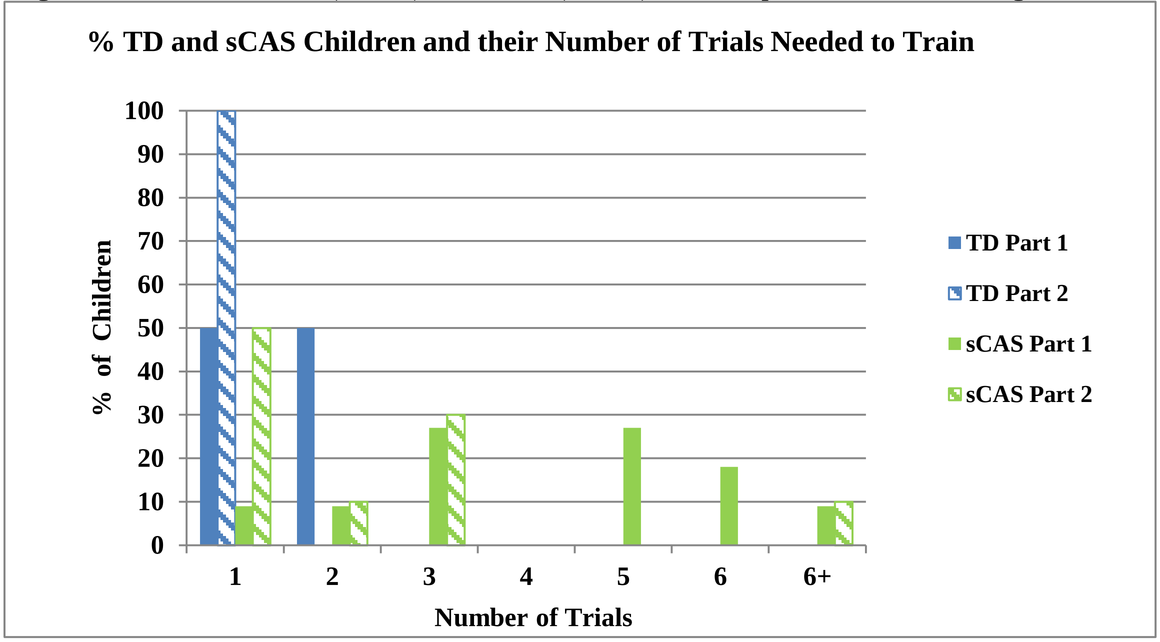
All statistical analyses were performed using SPSS (Version 21.0).

Table 3.10. Number of trials required to train children (*n* = 14 for TD; *n* = 11 for sCAS) to criteria for experimental task.

Group	# of training trials required for Training Part 1	# of training trials required for Training Part 2
TD	2	1
TD	2	1
TD	1	1
TD	1	1
TD	1	1
TD	2	1
TD	1	1
TD	2	1
TD	2	1
TD	1	1
TD	2	1
TD	1	1
TD	2	1
TD	1	1
	Mode = 1,2 (range = 1-2)	Mode = 1
sCAS	5	1
sCAS	3	2
sCAS	2	1
sCAS	5	1
sCAS	5	≤3*
sCAS	6	1
sCAS	3	3
sCAS	>6 (Did not train)	N/A (Not administered)
sCAS	3	3
sCAS	6	>6 (Did not train)
sCAS	1	1
	Mode = 3,5 (range = 1-6+)	Mode = 1 (range = 1-6+)

Note. TD = typically developing; sCAS = suspected apraxia of speech; *= researcher error.

Figure 3.4. Percent of TD (n = 14) and sCAS (n = 11) children per number of training trials.



3.3.2.1 Descriptive statistical analyses.

Table 3.11 shows the mean percent correct responses at each of the duration difference conditions, as well as the standard deviations (SDs), medians, modes, and confidence intervals (CIs) for the medians for the 5- and 6-year-old TD children (n = 14) and the children with sCAS (n = 9). Figure 3.5 illustrates the mean percent correct responses at each duration difference for both groups, as well as the median scores.

At the 0 ms duration difference, that is, 250-250, the TD children achieved approximately a 94% correct response rate by identifying the stimulus pair as “same,” while the children with sCAS achieved about 66% accuracy at that same duration condition. The median and mode for the 250-250 duration difference condition were each 100 for the TD children, with scores ranging from 90% to 100% with the exception of one child who had a correct response rate of 70%. The standard deviation was 8.52. For the children with sCAS, the median for the 250-250 duration difference condition was 70.00 and there was a bimode, 40 and 90, with the children’s scores ranging from 30% to 100% correct. The standard deviation was 26.03.

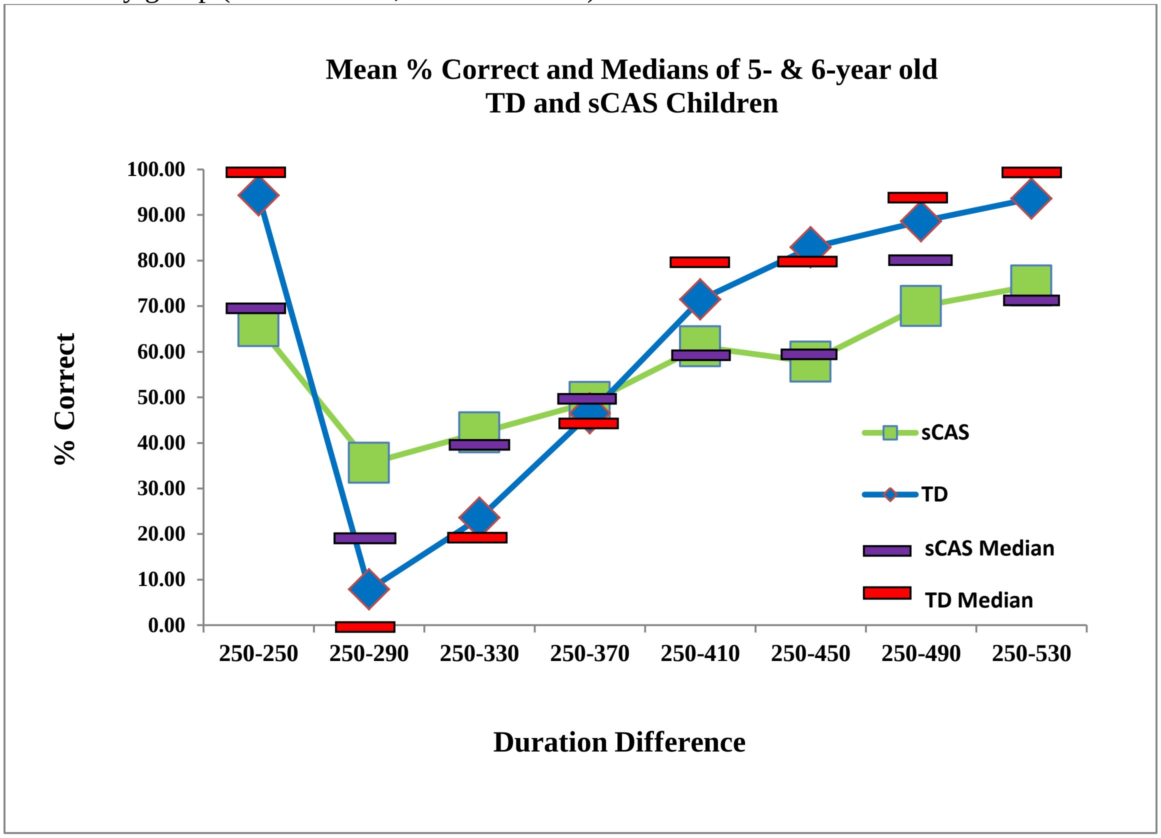
For large duration differences, beginning with the 250-410 duration difference condition,

Table 3.11. Mean percent correct responses, SDs, medians, modes, and CIs for 8 duration difference conditions of children by group (*n* = 14 for TD; *n* = 9 for sCAS).

Duration Difference Condition								
	250-250	250-290	250-330	250-370	250-410	250-450	250-490	250-530
TD Children (*n* = 14)								
Mean	94.29	7.86	23.57	46.43	71.43	82.86	88.57	93.57
	(8.52)	(10.51)	(24.37)	(31.53)	(27.13)	(17.73)	(17.03)	(10.08)
Median	100.00	0	20.00	45.00	80.00	80.00	95.00	100.00
Mode	100	0	30	20, 70	80, 90, 100	80, 100	100	100
CI for Median (95%)	90.00-100.00	.00-20.00	.00-30.00	20.00-70.00	40.00-100.00	70.00-100.00	80.00-100.00	80.00-100.00
Children with sCAS (*n* = 9)								
Mean	65.56 (26.03)	35.56 (33.95)	42.22 (28.63)	48.89 (28.92)	61.11 (24.21)	57.78 (28.19)	70.00 (25.00)	74.44 (17.40)
Median	70.00	20.00	40.00	50.00	60.00	60.00	80.00	70.00
Mode	40, 90	0, 20	10	50	30, 60, 80	80	80	70
CI for Median (95%)	40.00-90.00	.00-80.00	10.00-70.00	20.00-90.00	30.00-80.00	30.00-80.00	50.00-90.00	50.00-90.00

the TD children showed progressively improved response rates at the increasingly longer duration differences. The mean percent correct responses approximated 71% at 250-410 to about 94% at the 250-530 duration difference conditions. The medians and modes were consistent with the trajectory shown by the means, although the mode was tri-modal (80, 90, 100) at the 250-410 duration difference condition. The size of the standard deviations decreased progressively from 250-410 to 250-530. For the children with sCAS, performances did not show progressively improved response rates at the increasingly longer duration differences. In fact, their mean percent correct decreased from approximately 61% at 250-410 to approximately 58% at the 250-450 duration difference condition. The range of mean percent correct responses approximated 61% at 250-410 to about 74% at the 250-530 duration difference condition. The medians and modes at these greater duration differences also did not reflect an increasingly upward trajectory as evidenced by median values of 60.00 at both the 250-410 and 250-450 duration difference

Figure 3.5. Mean percent correct responses and medians for 8 duration difference conditions of children by group (n = 14 for TD; n = 9 for sCAS).



conditions, 80.00 at 250-490, and 70.00 at the most extreme duration difference condition of 250-530. Similarly, the mode was divided at 250-410 (30, 60, 80), remained stationary at 80 for the next two greater duration difference conditions of 250-450 and 250-490, and decreased to 70 at the largest 250-530 duration difference. The size of the standard deviations also did not decrease progressively as the duration difference increased from the 250-410 condition to the 250-530 condition but instead remained large compared to the TD children.

For performances at the intermediate 250-290, 250-330, and 250-370 duration difference conditions, the mean correct response rates for the TD children increased progressively, ranging from approximately 8% at 250-290 to 46% at 250-370. The standard deviations increased gradually as well, ranging from 10.51 to 31.53 across these same duration differences, which suggested increasing variability as the duration differences increased for these intermediate duration differences. The medians were similar to the mean scores and increased progressively from .00 at 250-290 to 45.00, an approximate 50-50 correct/incorrect response rate, at 250-370. The mode was 0 at 250-290, increased to 30 at 250-330, and split (20, 70) at the 250-370 duration difference condition. The standard deviation at the 250-370 duration difference condition was the largest of the eight duration difference conditions in the experimental task, including the same duration condition of 250-250. For performances of the children with sCAS at the intermediate 250-290, 250-330, and 250-370 duration difference conditions, the mean correct response rates increased progressively, ranging from approximately 36% at 250-290 to about 49% at 250-370. The medians also increased from 20.00 at 250-290 to 50.00 at 250-370, indicating a 50-50 correct/incorrect response pattern at this intermediate duration difference. The modes evidenced little improvement as two intermediate duration differences increased, with bimodal values of 0 and 20 at 250-290 and 10 at the 250-330 condition. Mode values then increased to 50 at the 250-370 duration difference condition. The large standard deviation at 250-290, 33.95, decreased only slightly at 250-330 and then showed no decline at the 250-370 duration difference condition. Of note, and quite unlike the TD children, the standard deviation for the children with sCAS at the 250-290 condition was the largest of the eight duration difference conditions in the experimental task.

Looking at the performance trajectories of the TD and sCAS groups across duration difference conditions, disparities are apparent (Figure 3.5). The TD children performed as expected with the large mean percent correct responses observed at the same (250-250) duration condition. At 250-290, the TD children again performed as anticipated with little percent accuracy detecting changes in the 40 ms duration difference. The intermediate duration difference conditions of 250-330 and 250-370 also reflected expected growth in performance accuracy. Finally, a regularly increasing trajectory was observed at the extreme duration difference conditions of 250-410 through 250-530.

Unlike the TD children's performance at the same duration condition, the mean percent response accuracy for the children with sCAS was approximately 30% poorer. At the 250-290 duration difference condition, the children with sCAS achieved approximately 36% accuracy detecting changes in the 40 ms duration difference. However, an elevated standard deviation (33.95) and large confidence interval for the median (.00- 80.00) were also evidenced at this duration difference condition. Group performance of the sCAS children at the intermediate conditions of 250-330 and 250-370 hovered around 50%. In fact, this 50-50 correct/incorrect response pattern continued into the extreme duration difference conditions of 250-410 and 250-450. A slight upward shift was then observed for the last two extreme duration difference conditions. In contrast to the consistent growth and upward trajectory of the TD's performance, the overall performance of the children with sCAS across the duration difference conditions appeared flattened.

In addition, a look at the performances of the TD and sCAS groups across the duration difference conditions revealed dissimilarities in 1) the variance of the mean percent correct responses, and 2) the confidence intervals of the medians. For the TD group, across all duration difference conditions, the standard deviation exceeded 24.00 (rounded) for the children's performances at only two conditions. Conversely, for the sCAS group, across all duration difference conditions, the standard deviation exceeded 24.00 (rounded) for the children's performances for all but one condition. Relative to confidence intervals at the more extreme duration difference conditions, for example, the lower bound of the confidence interval for TDs at 250-450 was 70.00 while the lower bound for the sCAS group at the same condition was 30.00. Likewise, at the last two duration discrimination conditions of 250-490 and 250-530, the lower bounds for TDs were 80.00 while the lower bounds for the sCAS group were 50.00. In general, standard deviations and confidence intervals were large across the majority of the duration difference conditions for the children with sCAS.

And, finally, the TD and sCAS groups differed on the determination of their >70% correct response thresholds. Appendix L outlines previous research related to the perceptual abilities of normal children and adults showing a variety of different threshold points but most focusing on 70% - 75% accuracy. Besides the same duration difference condition of 250-250, the TD children demonstrated evidence of first meeting the criterion of >70% accuracy at the 250-410 duration difference condition and then stabilized at 250-450 and beyond. Conversely, the children with sCAS did not show evidence of even approximating the criterion of >70% until the 250-490 duration difference condition. Here, their mean percent correct responses reached 70%, however, they did not exceed 70% until the most extreme duration difference condition of 250-530 (74.44).

3.3.2.2 Inferential statistical analyses.

In order to compare the performances of the TD children and the children with sCAS at each of the eight duration difference conditions, Mann-Whitney *U*-tests were performed. This test statistic is appropriate for comparing two samples that are unrelated. Because multiple tests were performed using the Mann-Whitney *U*-statistic, the critical alpha level was adjusted using the Bonferroni correction procedure, resulting in an alpha = .007. Table 3.12 shows the two-tailed results of these tests. A two-tailed analysis was used because, although the investigator hypothesized differences in performance between the two groups, she did not predict the specific patterns of the differences for the different duration conditions. The results of the Mann-Whitney

Table 3.12. Results of the Mann-Whitney *U*-tests for comparison of groups (n = 14 for TD; n = 9 for sCAS).

	Duration Difference Condition							
	250-250	250-290	250-330	250-370	250-410	250-450	250-490	250-530
U statistic	18.50	30.00	37.50	62.50	45.50	28.00	30.00	21.00
p	.003*	.039	.109	.975	.277	.028	.039	.007*
Effect Size	.61	---	---	---	---	---	---	0.58
ΣR_1	212.50	135.00	142.50	168.50	185.50	203.00	201.00	210.00
ΣR_2	63.50	141.00	133.50	107.50	90.50	73.00	75.00	66.00

*Significant at .007

U-tests were statistically significant at the .007 level for both the extreme 250-250 and the 250-530 duration difference conditions, suggesting the TD children’s performance was statistically significantly different from that of the children with sCAS. Moreover, the effect sizes (Rosenthal, 1991) for the significant duration difference conditions were $r = .61$ (250-250) and $r = .58$ (250-530) ("$r = |z|/\sqrt{n}$" Field, 2005, p. 532). These effect sizes were large. In addition, the sum of the ranks for the TD children was typically larger than for the children with sCAS, suggesting the performance of the TD children exceeded that of the children with sCAS. If the alpha level had not been adjusted to account for multiple tests of significance, five of the eight duration difference comparisons of TD and sCAS (250-250, 250-290, 250-450, 250-490, 250-530) would have reached statistical significance at the alpha level of .05, suggesting a trend. The statistical differences cited here between the TD and sCAS groups serve as additional evidence of dissimilar performances shown by the two groups across duration difference conditions.

3.3.2.3 Default-to-different response patterns.

The difference in percent correct response rate for the 250-250 duration difference condition and the 250-290 duration difference condition was examined with a Wilcoxon Signed Ranks test for both the TD children and children with sCAS. Because multiple tests were not performed, the critical alpha level was not adjusted and remained at .05. A one-tailed direction was selected because the researcher hypothesized that the children's accuracy would decrease from the 250-250 to 250-290 duration difference condition.

For the TD children, the result of a Wilcoxon Signed Ranks test was statistically significant for the 250-250/250-290 duration difference comparison ($T = 0.00$, $n = 14$, $p < .05$) indicating the percent correct performance was significantly better at the duration difference of 250-250 than the 250-290 duration difference condition. The effect size (Rosenthal, 1991) for this significant matched-pair condition was $r = .89$ ("$r = |z|/\sqrt{n}$" Field, 2005, p. 532). This effect size was large.

Conversely, for the children with sCAS, the result of a Wilcoxon Signed Ranks test was not statistically significant ($T = 10.00$, $n = 9$, $p > .05$), indicating the percent correct performance was not significantly different between the 250-250 and the 250-290 duration difference conditions. Results suggest the children with sCAS, but not the TD children, may have used a "default-to-different" strategy.

Chapter 4 Discussion

4.1 Summary of Results

The present study addresses duration discrimination in a systematized experimental design for a group of school-age TD children (*n* = 21) and a smaller group of school-age children diagnosed with sCAS (*n* = 11). Furthermore, this research provides preliminary support for two hypotheses formulated by the investigator: The first research question exclusively addressed the performances of the TD children. The investigator hypothesized that on the experimental task designed for this research: 1) the children's performances at the extreme duration differences (i.e., pairs that were of the same duration, pairs that were toward the maximum duration differences) would demonstrate mostly correct responses, 2) as duration differences increased from the smallest difference in duration through the intermediate duration differences, the children's correct response rates would show improved performances as the duration differences increased, but sizes of variances would become large, and 3) the performances of younger TD children would differ from the performances of older TD children.

Descriptive and inferential statistical analyses indicate the TD children could perform the experimental task and supported the predicted pattern of performance for the TD group as a whole. As noted in a previous chapter, the performance of one child in the 7/8-year-old group skewed the results at the same duration condition of 250-250. The standard deviation at this condition, upon inspection of individual scores, was influenced considerably by the hypervigilant manner that this individual child approached the task. It was obvious to the investigator, however, that he could perform the task.

The TD group's performance showed a progressively upward trajectory, as the duration differences increased from the smallest duration difference condition of 250-290 to the most extreme duration difference condition of 250-530. Performance at the intermediate duration difference conditions was as expected, with large standard deviations.

Finally, with regard to the initial research question, the three age groups (5-, 6-, and 7/8-year-olds) shared an upward trajectory of performance as the duration difference increased from the 250-290 up to the 250-450 duration difference condition. At that point, the three age groups started to diverge slightly in the pattern of their performances. However, contrary to the investigator's hypothesis, no statistically significant differences were found at any duration difference condition when the performances of the three age groups were analyzed using an adjusted conservative alpha level.

The second research question addressed the performances of 5- and 6-year-old children with sCAS compared to a subset of the TD children who were 5- and 6 years old. The investigator hypothesized that the sCAS children's performances would show different patterns on the duration difference task than that of their TD age peers.

Most of the children with sCAS, but not all, could perform the experimental task. Descriptive and inferential statistical analyses indicate the sCAS group's performance did not show a progressively upward trajectory, as the duration differences increased from the smallest duration difference condition of 250-290 to the most extreme duration difference condition of 250-530. Instead, their performance suggested a more flattened course across the experimental conditions. In addition to not showing progressively improved response rates at the increasingly longer duration differences, the children with sCAS did not progressively decrease the size of their variance as the duration difference increased from the 250-410 condition to the 250-530 condition.

When the investigator compared the performances of a subset (5- and 6-year-olds; n = 14) of the total TD group to a group of children with sCAS (5- and 6-year-olds; n = 11), three primary differences were observed. Distinctions were noted in areas relating to training trials; performance trajectory across duration difference conditions, including variances across duration difference conditions, confidence intervals, and >70% correct response thresholds; and possible default-to-different response patterns. Relative to the last difference, the observation that the sCAS group did not show a significant difference in their performance between the same duration difference of 250-250 and the next duration difference condition of 250-290, as did the TD group, may be suggestive of category formation deficits for some children in the sCAS group. Nittrouer et al. (2011) indicated, “In an AX task, the ability to correctly judge physically identical stimuli as the same is an indication of how well the listeners have formed a category for the standard (A) and are able to recognize stimuli as members of that category” (p. 775). The researchers continue to say, “If children with PPD [phonological processing deficits] actually have difficulty in forming categories, it means that they may be biased toward responding that stimuli are different whether they really recognize that difference or not” (p. 776). On the other hand, a less sophisticated explanation for over-identification of responses as “different” by many children in the sCAS group may be attributed to guessing. The more flattened pattern of performances of the sCAS children across multiple duration difference conditions that hovered in the mid-range of percent correct responses suggests either a greater guessing pattern and/or a “default-to-different” pattern.

Overall, the present study provides preliminary support for most of the investigator’s hypotheses. That is, TD children were able to perform the task and showed predictable improvement patterns, as the duration differences increase across conditions from minimally to maximally different.

Assessing the ability of the TD population to detect duration differences in a systematized experimental paradigm was prerequisite to addressing the ability of children with sCAS to detect duration differences in the same experimental task. It was imperative to first establish what TD children could do on a specific task before looking at speech disordered children’s performance on the same task, as Morrongiello and Trehub (1987) indicated that duration discrimination thresholds “vary across studies as a function of psychophysical test procedure, stimulus context, and duration of the standard interval” (p. 415). Although not the primary focus of the investigator’s study, the results of a preliminary look at a small group of children with sCAS revealed that, although this sample could perform the experimental task, they do so differently and with greater variability than that demonstrated by the TD children.

4.2 Strengths of the Research

An important strength of the present study is the significant attention the investigator placed on the sCAS selection criteria for participation in this research study. As detailed in Chapter 1, the literature does not provide specific quantification for many of the criteria for participant grouping with regard to lists of sCAS characteristics. Therefore, the investigator used the literature, guidelines from published assessment instruments, clinical practices, and/or common empirical practices to determine specific criteria/criterion for evaluating the presence/absence for each of the 11 characteristics (Table 2.2). Unlike much of the sCAS research, the present study included explicit criterion/criteria for each of the 11 characteristics that served to qualify or disqualify children as participants, leading the investigator to be as certain as is reasonably possible that the children within her groups were indeed TD or sCAS. As such, the participant selection protocol that the investigator developed can serve as a rigorous practice for her subsequent research. It can also serve as a robust protocol for other investigators to use to improve the quality of the research in the speech sound disorders area.

Looking at the obtained data in the present study for each of the 11 characteristics assessed, there were two characteristics, absent/shortened frication and reduced diadochokinetic rate, that were observed in all of the children enrolled in the sCAS group (n = 11). One-hundred percent of the sCAS group presented with absent or shortened frication/affrication, observed via either inspection of phonetic transcriptions of the DEAP Phonology Single-Word Production (Dodd et al., 2006) subtest or by determination of the child's Maximum Fricative Duration score (grand mean of prolonged productions of /f/, /s/, and /z/) (Thoonen et al., 1999). Additionally, all children with sCAS presented with a reduced diadochokinetic rate (/pʌtəkə/).

A second strength of this study is that it is the only known research addressing the ability of TD children to discriminate duration-only differences within a linguistic context, namely vowels. In fact, for young children, duration discrimination studies using white noise or tone stimuli are even limited. Furthermore, given the reported vowel duration deficits cited in the speech production of children with sCAS, the research for this population is deficient in assessing the ability of these children to discriminate vowel duration differences. In light of the limited research regarding discrimination of duration differences in the TD and speech-sound-disordered child populations, the investigator concluded that a study focused on detection of durational differences in a vowel context was warranted for both populations. Additionally, for the present study, the investigator incorporated the vowel stimulus into a CV non-word context (/bɑ/) for the experimental design. This decision was based on the usefulness of non-word stimuli (simple CV syllables) cited in both the production and perception literature. Peter and Stoel-Gammon (2005; 2008) found a non-word imitation task discerning in assessment of vowel duration in both TD and speech-impaired participants. Similarly, the research of Bridgeman and Snowling (1988) and Nijland (2009) cited the usefulness of non-word stimuli in their discrimination tasks with TD and speech-impaired participants, as well.

Four research studies in the literature, identified in the present study, have looked at duration discrimination in children ranging in age from infants to 18 years of age. Relative to the researchers' selection of a task paradigm, one of the four studies presented sequences of 18 white noise bursts (Morrongiello and Trehub, 1987), two studies employed a three-interval, forced choice task using noise bursts or tones (Elfenbein et al., 1993; Jensen and Neff, 1993), and one study used pairs of auditory stimuli (Himpel et al., 2009). Given the researchers' reported concern of auditory memory negatively impacting the performances of the children in the two studies using a three-interval forced choice task, this investigator chose a paired stimulus task. However, unlike the paired-stimuli task used by Himpel et al., a same-different task was selected for the present study and children simply identified the two stimuli within each pair as "same" or "different", eliminating the semantic element of "longer" that was present in the research of

Himpel et al. (2009). The experimental paradigm for the current study addressed what were considered weaknesses in the protocols of these few studies.

In addition to selection of vowel stimuli within non-words and a same-different task paradigm, a strength of the present study was confirmation that children had sufficient auditory memory to undertake the experimental task. Based on results of a pilot study and the children's performances on the Number Repetition subtest of the CELF-4 (Semel et al., 2003) within the present study, it was determined that the children's auditory memory was sufficient to deal with the experimental task so as not to confound results because of reduced auditory memory abilities.

4.3 Limitations of the Present Study

One consideration that must be taken into account when interpreting the results of this study is the small sample sizes of both the TD and sCAS populations. In addition to smaller sample sizes, firm interpretation of the results of the main and preliminary studies must also be viewed in light of the significant performance variability observed within the two groups. Sample sizes reflected the low prevalence of sCAS, the investigator's decision to include only males in the study, and the narrow age range identified in the selection criteria. The lower age limit (5-year-olds) was dictated, in part, by the research methodology; that is, the child had to be old enough to maintain attention throughout the 80-item experimental task. Conversely, the upper age limit was influenced by concerns expressed by Shriberg et al. (1997b) that "both advanced age and more extensive intervention experience militate against finding persisting stress involvement" (p. 308). Consequently, the upper age limit was capped at 6;11 for children in the sCAS group.

Indeed, the investigator believes that the effects of intensive habilitation were significantly noted for one child enrolled in the sCAS group in the present study. He was diagnosed with sCAS at an early age by his SLP and the diagnosis was confirmed by a neurologist; however, at the time of his initial evaluation for possible inclusion in this study, he was 5;5 years of age and had received approximately 2.5 years of intensive, private intervention.

Consequently, at the time of the evaluation, he presented with among the fewer number of the sCAS characteristics, likely due to many years of private speech habilitation. Nevertheless, his performance on the experimental task looking at duration difference discrimination was as poor as that of the other children with sCAS. This pattern of performance also suggests that duration discrimination ability may be an underlying clinical characteristic of sCAS that may not remediate even with intensive habilitation.

Secondly, relative to selection criteria for inclusion in this study, consideration may be given to the method of evaluation for one of the 11 characteristics, vowel errors. For the present study, children's production of vowel errors was evaluated using the DEAP Articulation Single-Word Production (Dodd et al., 2006) subtest. This assessment instrument evaluates vowel accuracy at the word level (26 monosyllabic words, three bisyllabic words, and one multisyllabic word). However, thought might be given to expanding the evaluation to include vowel assessment during production of additional multisyllabic words or non-word utterances, given the research of McNeil et al. (1997). These researchers, in their work with individuals with acquired apraxia of speech, cited longer vowel durations in the context of more complex words (multisyllabic words or non-word utterances), as compared to monosyllabic words.

As reported previously, results of the preliminary study involving the sCAS group led the investigator to wonder why some children within this group presented with what seemed to be a bias of default-to-different strategy on the experimental task. Therefore, it might be beneficial to consider the inclusion of additional pairs of identical stimuli, for which the correct response is "same," in subsequent research to reduce the potential effect of more "different" items triggering "different" as a default response for those children who guessed.

Finally, the investigator was not blind to the children's group assignments during data collection so possible researcher bias might exist. However, the high inter-rater agreement between the investigator and other raters in recording children's responses mitigates against the presence of any notable researcher bias in the data.

4.4 Clinical Implications and Future Directions

The results of this preliminary investigation of discrimination of vowel duration in children with sCAS suggest that further research on duration discrimination skills is warranted in this population. As a group, children with sCAS displayed poorer performance on the vowel duration discrimination experimental task, compared to a TD group. Although the analyses of the data in this present study were primarily focused on the collective group performances of the TD children and children with sCAS, future exploration of the data could probe individual performances on the experimental task. These probes could reveal participants' patterns of performance that could contribute to additional understanding of the children's strengths and weaknesses with regard to vowel duration discrimination. Scatterplot analyses, for example, might be utilized to examine variability among the performances of individual children within the two groups on the task. As the primary component of this present study, the performance of the TD children on the experimental task serves as affirmation of both the methodology and the absence of significant confounds in the task paradigm.

Prior to enrollment in the study and participation in the experimental task, school-age children were required to meet certain qualifications founded on a criterion-based assessment protocol created by the investigator, addressing 11 sCAS characteristics that she identified in the research literature. Concern over the lack of clinical markers to diagnose CAS as been expressed for some time. As a result, the Ad Hoc Committee on Apraxia of Speech in Children (American Speech-Language-Hearing Association, 2007) identified three features that appear to be associated with a praxis deficit: (a) inconsistent errors on consonants and vowels on attempts of multiple productions of the same target word or syllable, (b) atypical coarticulatory transitions between sounds and syllables, and (c) inappropriate prosody, especially noted with lexical or phrasal stress.

Surprisingly, the results of the present preliminary study did not exclusively support these features. Based on a small sample size (*n* = 11), the investigator identified two of the 11 sCAS characteristics assessed, absent/shortened frication (Thoonen et al., 1999) and reduced diadochokinetic rate, in all of the children enrolled in the sCAS group. One-hundred percent of the sCAS group presented with absent or shortened frication/affrication, observed via inspection of phonetic transcriptions at word level or by determination of the child's Maximum Fricative Duration score. Additionally, all children with sCAS presented with a reduced diadochokinetic rate. Contrastively, < 30% of the children in the sCAS group presented with inconsistent speech production errors and only approximately 80% presented with suprasegmental errors. If the trends suggestive in the present study prove accurate in future research with a larger sample size, modifications to clinical assessment protocols for diagnosis of sCAS that include absent/shortened frication (Thoonen et al., 1999) and reduced diadochokinetic rate may be warranted.

The results of the present experimental task, suggestive of a deficit in discrimination abilities of children with sCAS, are congruent with recent research by Shriberg et al. published in 2012. Based on their work using the Syllable Repetition Task (Shriberg et al., 2009), these researchers concluded that "speakers with CAS have speech processing deficits in encoding, memory, and transcoding" (p. 445). In addition to advocating for standardized assessment protocols for children with sCAS, Shriberg et al. (2012) reported, "A need in future studies, using extended controlled stimuli, is to include measures of participants' auditory-perceptual status, possibly using physiological measures…" (p. 473).

Results of this present study, relative to both the diagnostic criteria for identification of sCAS and the performance of sCAS children on a vowel duration discrimination task, may contribute to the potential development of a proposed standardized assessment protocol that would include just three tasks, the duration discrimination task used in the present study, the fricative task (Thoonen et al., 1999), and the diadochokinetic task. Future research could investigate the sensitivity and specificity of such a 3-task protocol to determine if it would accurately predict group classification of children as TD or sCAS. It might also serve as a differential diagnostic protocol for accurately classifying sCAS and PD children. Thus, future directions of this present research, following confirmation of the trends suggested in the current study within larger sCAS and TD samples, may be to extend participants to include children with phonological impairment. In turn, the translation of such results could have clinical implications in the form of extended development of evidence-based intervention for the sCAS population.

Appendices

Appendix A. Case history questionnaire completed by parent(s) of all children.

1. Is English the only language spoken in your home?	____ Yes	____ No	Explain:
2. Does your child have any history of… Stuttering?	____ Yes	____ No	Explain:
Hearing loss?	____ Yes	____ No	Explain:
Ear infections with fluid behind the eardrum?	____ Yes	____ No	Explain:
A speech disorder that he obtained since learning to speak?	____ Yes	____ No	Explain:
Nervous system disorders?	____ Yes	____ No	Explain:

3. How much did your child vocalize as a baby (particularly between the ages of 6 and 12 months)? Please rate on a scale of 1 to 5, with 1 indicating that your child seemed to vocalize rarely (or was a 'quiet' baby), and 5 indicating that your baby seemed to vocalize often (or was a 'vocal' baby) (Highman, Hennessey, Sherwood, & Leitão, 2008).

1 2 3 4 5

Rarely vocalized except for crying, etc. — Frequently vocalized

4. Did your child make 'cooing' noises, like 'ah', 'ee' (Highman et al., 2008)?

____ Yes ____ No ____ Unsure

If yes, at what age? (You can circle one particular month, or a range of months)

3 mos	4 mos	5 mos	6 mos	7 mos	8 mos	9 mos	10 mos	11 mos	12 mos	13 mos	14 mos	15 mos

Other: _____________ ____ Cannot remember

5. Did your child 'babble' as a baby, like "ba-ba", "ma-ma", "da-da-da" where the sound is repeated (Highman et al., 2008)?

____ Yes ____ No ____ Unsure

If yes, at what age? (You can circle one particular month, or a range of months)

6 mos	7 mos	8 mos	9 mos	10 mos	11 mos	12 mos	13 mos	14 mos	15 mos	16 mos	17 mos	18 mos

Other: ______________ ____ Cannot remember

6. Did your child ever produce babble where the consonant sound changed, like "ba-da", "te-da", or "be-de-ga" (Highman et al., 2008)?

____ Yes ____ No ____ Unsure

If yes, at what age? (You can circle one particular month, or a range of months)

6 mos	7 mos	8 mos	9 mos	10 mos	11 mos	12 mos	13 mos	14 mos	15 mos	16 mos	17 mos	18 mos

Other: ______________ ____ Cannot remember

7. Did your child babble as much as other children? Please state to whom you are comparing your child (e.g., brother/sister, friend's son/daughter) (Highman et al., 2008).

____ Babbled less ____ Babbled more ____ Babbled about the same

____ Unsure

Compared to ____ .

Appendix B. Children selection criteria.

CHILDREN SELECTION CRITERIA	**TD**	**sCAS**
C.A. = 5;0 to 8;11 year-old males	√	
C.A. = 5;0 to 6:11 year-old males		√
Monolingual, English-speaking home	√	√
Hearing WNL	√	√
Absence of stuttering disorder	√	√
No hx of hearing loss	√	√
No hx of frequent otitis media with effusion	√	√
No hx of acquired speech disorder	√	√
No hx of neurologic disorders /syndromes	√	√
Mod-severe speech sound deficit: *Diagnostic Evaluation of Articulation and Phonology* (DEAP, Dodd, Hua, Crosbie, Holm, & Ozanne, 2006)	Cut-off score = Scaled score of ≥8 on the Phonology Single-Word Production subtest	Cut-off score = Scaled score of ≤5 on the Phonology Single-Word Production subtest
sCAS Characteristics, per Table 2.2 (Davis, Jakielski, & Marquardt, 1998; ASHA, 2007; Burns, 2011): Performance on specific assessment tasks; gold standard clinical agreement; judgment of an independent, experienced SLP to confirm group assignment (sCAS, TD) based on videotape analyses.	≤1 characteristic	≥5 (≥45%) of 11 characteristics

Appendix C. Speech and non-speech characteristics of developmental apraxia of speech (Davis, Jakielski, & Marquardt, 1998, pp. 28-29).

Speech Characteristics
1. Limited consonant and vowel repertoire. The phonetic inventory of consonants and vowels (a count of presence of consonants and vowels, regardless of correct usage) does not include a complete inventory of phonemes produced in adult English. 2. Frequent omission errors. Errors include frequent omission of phonemes in speech output. 3. High incidence of vowel errors. The vowel inventory available is not used correctly. 4. Inconsistent articulation errors. Variability and lack of consistent patterns characterizes speech output for both consonants and vowels. Variability is not always context-dependent, but may occur in repeated productions of the same lexical item. 5. Altered suprasegmental characteristics. Suprasegmental characteristics of rate, pitch and loudness may also be inconsistent and variable, undermining intelligibility. 6. Increased errors on longer units of speech output. All types of errors as well as variability of error types increase with increasing length of utterance. 7. Significant difficulty imitating words and phrases. Groping postures or lack of willingness or ability to imitate a model. 8. Predominant use of simple syllable shapes. The child either employs simple lexical types or reduces more complex word types to CV, V, or CVC predominantly. Variability may be seen in use of syllable shapes, as well as in consonant and vowel types and errors.
Non-speech Characteristics
1. Impaired volitional oral movements. 2. Reduced expressive compared to receptive language skills. 3. Reduced diadochokinetic rates.

Appendix D. Presence/absence of 11 sCAS characteristics for all children.

ID	Group	Omission Errors	Vowel Errors	Inconsistent Errors	Suprasegmental Errors	Groping/ Silent Posturing	Simple Syllable Shapes
01	TD	-	-	-	-	-	-
02	TD	-	-	-	-	-	-
03	TD	-	-	-	-	-	-
04	TD	-	-	-	-	-	-
05	TD	-	-	-	-	-	-
06	TD	-	-	-	-	-	-
07	TD	-	-	-	-	-	-
08	TD	-	-	-	-	-	-
09	TD	-	-	-	-	-	-
10	TD	-	-	-	-	-	-
11	TD	-	-	-	-	-	-
12	TD	-	-	-	-	-	-
13	TD	-	-	-	-	-	-
14	TD	-	-	-	-	-	-
15	TD	-	-	-	-	-	-
16	TD	-	-	-	-	-	-
17	TD	-	-	-	-	-	-
18	TD	-	-	-	+	-	-
19	TD	-	-	-	-	-	-
20	TD	-	-	-	-	-	-
21	TD	-	-	-	-	-	-
22	sCAS	+	-	+	+	-	-
23	sCAS	+	-	-	+	-	+
24	sCAS	+	+	+	+	-	+
25	sCAS	+	-	-	-	-	+
26	sCAS	-	-	+	+	-	-
27	sCAS	-	-	-	+	-	-
28	sCAS	+	-	-	+	-	+
29	sCAS	+	-	-	+	-	+
30	sCAS	+	+	-	+	-	+
31	sCAS	+	-	-	+	+	-
32	sCAS	+	-	-	-	-	-

Note. TD = typically developing; sCAS = suspected apraxia of speech; + = characteristic present per criteria; - = characteristic absent per criteria.

Appendix D continued.

ID	Group	Limited Consonants and Vowels	Increased Errors on Longer Speech Units	Impaired Volitional Oral Movements	Absent or Shortened Frication/ Affrication	Reduced DDK Rate	Referral Status (- = TD; + = sCAS) (prior to study)*
01	TD	-	-	-	+	-	-
02	TD	-	-	-	-	-	-
03	TD	-	-	-	-	-	-
04	TD	-	-	-	-	-	-
05	TD	-	-	-	-	-	-
06	TD	-	-	-	-	-	-
07	TD	-	-	-	-	-	-
08	TD	-	-	-	-	-	-
09	TD	-	-	-	-	-	-
10	TD	-	-	-	-	-	-
11	TD	-	-	-	-	-	-
12	TD	-	-	-	-	-	-
13	TD	-	-	-	-	-	-
14	TD	-	-	-	-	-	-
15	TD	-	-	-	-	-	-
16	TD	-	-	-	-	-	-
17	TD	-	-	+	-	-	-
18	TD	-	-	-	-	-	-
19	TD	-	-	-	-	-	-
20	TD	-	-	-	-	-	-
21	TD	-	-	-	-	-	-
22	sCAS	+	-	+	+	+	+
23	sCAS	+	+	+	+	+	+
24	sCAS	+	+	+	+	+	+
25	sCAS	+	+	+	+	+	+
26	sCAS	+	+	+	+	+	+
27	sCAS	+	-	-	+	+	+
28	sCAS	+	-	+	+	+	+
29	sCAS	+	-	+	+	+	+
30	sCAS	+	-	+	+	+	+
31	sCAS	-	-	+	+	+	+
32	sCAS	+	-	+	+	+	+

Note. TD = typically developing; sCAS = suspected apraxia of speech; + = characteristic present per criteria; - = characteristic absent per criteria; DDK = diadochokinetic; * = status not included as one of the 11 characteristics.

Appendix E. Methods for determination of inter- and intra-rater agreement.

Classification: An ASHA-certified, licensed, experienced SLP served to confirm or reject the initial group status (sCAS or TD) for a random number of children (25%). Blind to the children's initial group assignment, although aware of the general area of the research but not the hypotheses, the SLP observed videotapes of the children's assessments, reviewed corresponding assessment data, and independently assigned these children to either the sCAS or TD group. Her group assignments were compared to the children's initial assignments.

Severity: One graduate research assistant (RA1) was trained on 1) determination of severity of speech sound production, and 2) eight of the 11 characteristics (omission errors, vowel errors, inconsistency, use of simple syllable shapes, limited consonant and vowel repertoires, increased errors on longer units of speech, impaired volitional oral movements, and absent or shortened frication/affrication). Inter-rater agreement between the investigator and RA1 was determined by obtaining the percent agreement, per phoneme, on the *Diagnostic Evaluation of Articulation and Phonology* (DEAP; Dodd, Hua, Crosbie, Holm, & Ozanne, 2006) Phonology Single-Word Production subtest. A phoneme was considered agreed upon when both raters judged the designated phoneme to be correct or to be in error. Agreement was not based on determination of the type of error (distortion, substitution, or deletion). Given that this subtest included the opportunity for production of 226 phonemes, the number of phonemes agreed upon by the two raters was divided by 226 to determine the percent agreement.

Omission errors: Inter-rater agreement between the investigator and RA1 was obtained for five different error patterns on the DEAP Phonology Single-Word Production subtest: Consonant Cluster Reduction (CCR), Weak Syllable Deletion (WSD), Final Consonant Deletion (FCD), Initial Consonant Deletion (ICD), and Medial Consonant Deletion (MCD). To obtain percent agreement for these error patterns, the number of targeted data points agreed upon by the two raters per error pattern was divided by the number of data points identified for the error pattern. A data point was considered agreed upon when both raters judged the designated data point to be present/represented or to be omitted. For CCR, there were 27 opportunities for production or data points identified by the DEAP; WSD = 24 data points; and FCD = 36 data points. The investigator and RA1 also identified the number of data points within the DEAP Phonology Single-Word Production subtest for the error patterns of ICD (29 data points) and MCD (19 data points).

Vowel errors: Inter-rater agreement between the investigator and RA1 was recorded for the Percentage of Vowels Correct (PVC) on the DEAP Articulation Single-Word Production subtest. Percent agreement was determined by the frequency with which the two raters agreed on the level of severity (mild, mild-moderate, moderate-severe, or severe). Vowel production was considered agreed upon if both raters' calculations placed the child within the same severity category.

Inconsistency: Inter-rater agreement between the investigator and RA1 was determined for the presence or absence of consistency on the DEAP Word Inconsistency subtest. Percent agreement was determined by the frequency with which the two raters agreed on the subtest's final category (meets/does not meet criterion). This subtest included 25 test items and three trials offered per item. A test item was marked *same* if all trials of the same item were similarly produced and a

Appendix E continued.

test item was marked *different* if one or more of the three trials was produced in a different manner. A word inconsistency score was obtained by dividing the total number of items marked *different* by the total number of items (25) and multiplying this number by 100. Based on criterion established by the DEAP, a score of ≥40% is suggestive of an inconsistent phonological disorder. Consistency/inconsistency was considered agreed upon if both raters categorized the child within the same category (met/did not meet criterion).

Suprasegmental errors: Inter-rater agreement between the investigator and an ASHA-certified, licensed, experienced SLP was determined for the presence or absence of suprasegmental errors observed during video-analysis of each child's connected speech sample (*Frog, Where are You?*; Mayer, 1969). Using a checklist (Appendix H), percent agreement was determined by the frequency with which the two raters agreed on the final category (presents/does not present with suprasegmental errors).

Groping behaviors/silent posturing: Inter-rater agreement between the investigator and the same ASHA-certified, licensed, experienced SLP was determined for the presence or absence of groping behaviors/silent posturing observed during video-analysis of each child's Recalling Sentences subtest (initial 10 items) of the *Clinical Evaluation of Language Fundamentals-4* (CELF-4; Semel, Wiig, & Secord, 2003). Percent agreement was determined by the frequency with which the two raters agreed on the final category (presents/does not present with groping behaviors/silent posturing) (Appendix H).

Use of simple syllable shapes: Inter-rater agreement between the investigator and RA1 was obtained for use of simple syllable shapes during administration of the *Kaufman Speech Praxis Test for Children* (KSPT; Kaufman, 1995) Complex Bisyllabics and Polysyllabic Synthesis/Sequencing subtests. For the purpose of this study, these combined assessments included 31 targeted syllables (subtest item #7 was omitted because it was a non-word) and each rater identified the frequency that syllables were reduced (contained deleted phonemes) and not reduced (contained accurate or attempted production of all phonemes, with credit given for substitution errors). A syllable was considered agreed upon when both raters judged that syllable to be reduced/not reduced. To obtain percent agreement for syllable shapes, the number of targeted syllables agreed upon by the two raters was divided by the number of total syllables (31) assessed.

Limited consonant and vowel repertoires: Inter-rater agreement between the investigator and RA1 was recorded for the Percentage of Phonemes Correct (PPC) on the DEAP Articulation Single-Word Production subtest. Percent agreement was determined by the frequency with which the two raters agreed on the level of severity (mild, mild-moderate, moderate-severe, or severe). Phoneme errors were considered agreed upon if both raters' calculations placed the child within the same severity category.

Increased errors on longer units of speech: Inter-rater agreement between the investigator and RA1 was obtained for frequency of increased errors on longer speech units during administration of the KSPT Length and Complexity subtest. This assessment included five test items

Appendix E continued.

(sequences), each containing three words of increasing complexity (monosyllabic, bisyllabic, and polysyllabic words). Each word was phonetically transcribed (broad transcription) to see if similar phoneme production was maintained. Each rater identified the frequency with which the child maintained similar phoneme production across the three words and scored *yes* or *no* for each of the five test items. To obtain percent agreement for frequency of errors, the number of test items agreed upon by the two raters was divided by 5.

Impaired volitional oral movements: Inter-rater agreement between the investigator and RA1 was determined for the presence or absence of oral motor deficits on the DEAP Oral Motor Screen subtest. The child's performance was summed across three categories, with the total oral motor score either meeting or not meeting criterion based on chronological age. Percent agreement was determined by the frequency with which the two raters agreed on the child's summed performance. Performance was considered to be agreed upon if both raters categorized the child within the same category (met/did not meet criterion).

Absent or shortened frication/affrication: Inter-rater agreement between the investigator and RA1 was obtained for production of fricatives and affricates on the DEAP Phonology Single-Word Production subtest. To obtain percent agreement for these phonemes, the number of targeted data points agreed upon by the two raters was divided by the number of opportunities for production of affricates and fricatives. The investigator and RA1 identified 44 opportunities for production of these phonemes on the subtest. A data point was considered agreed upon when both raters judged the child's production of the fricative/affricate to be accurate or inaccurate (misarticulated/omitted). Absent or shortened frication/affrication was also measured by determination of the child's Maximum Fricative Duration score. Percent agreement between the investigator and RA1 was based on comparison of the average number of seconds that a child prolonged production of the phonemes /f/, /s/, and /z/ on one breath. The calculation of RA1 was required to be within ±10% of the investigator's value to be considered agreed upon.

Reduced diadochokinetic rate: A second graduate research assistant (RA2) was trained on one of the 11 characteristics, reduced diadochokinetic rate. Inter-rater agreement between the investigator and RA2 was obtained for each child's maximum repetition rate of trisyllabic sequences (/pʌtəkə/). Percent agreement between the investigator and RA2 was based on comparison of the average number of syllables that a child could repeat per second. Each child was instructed to repeat as many repetitions as possible of the sequence /pʌtəkə/ on one breath. The investigator and RA2 simultaneously viewed each child's recording and perceptually determined the most accurate/fluent/rapid series (five repetitions of the sequence /pʌtəkə/) to analyze. Subsequently, the audio file was converted into a WAV file. Using Praat, the first and last two syllables of the selected series were excluded from analysis and the number of repeated syllables per second was calculated based on 12 syllables. The calculation of RA2 was required to be within ±10% of the investigator's value to be considered agreed upon.

For intra-rater agreement, similar procedures occurred. The comparison, however, was within raters, as opposed to across raters.

Appendix F. Inter-rater agreement: Investigator and two graduate research assistants and SLP.

Parameters	*Task*	*% Agree/Disagree*	*Mean % Agreement*
Classification	Children Selection Criteria	100%	
Severity	DEAP Phonology Single-Word subtest		96%
Omission errors	DEAP Phonology Single-Word subtest		98%
	CCR		97%
	WSD		99%
	FCD		97%
	ICD		100%
	MCD		96%
Vowel errors	DEAP Articulation Single-Word subtest (PVC)	88%	
Inconsistency	DEAP Word Inconsistency subtest	100%	
Suprasegmental errors	Connected Speech (*Frog, Where Are You?*)	82%	
Groping behaviors/Silent posturing	CELF-4 Recalling Sentences subtest	100%	
Use of simple syllable shapes	KSPT Complex Bisyllabics subtest & Polysyllabic Synthesis/Sequencing subtest		95%
Limited consonant and vowel repertoires	DEAP Articulation Single-Word subtest (PPC)	75%	
Increased errors on longer units of speech	KSPT Length and Complexity subtest		93%
Impaired volitional oral movements	DEAP Oral Motor Screen	100%	
Absent or shortened frication/affrication	DEAP Phonology Single-Word subtest (% affricate/fricative errors)		97%
	MFD	100%	
Reduced diadochokinetic rate	MRRtri	100%	

Note. DEAP = Diagnostic Evaluation of Articulation and Phonology; CCR = consonant cluster reduction; WSD = weak syllable deletion; FCD = final consonant deletion; ICD = initial consonant deletion; MCD = medial consonant deletion; PVC = Percentage of Vowels Correct; CELF-4 = Clinical Evaluation of Language Fundamentals-4; KSPT = Kaufman Speech Praxis Test for Children; PPC = Percentage of Phonemes Correct; MFD = Maximum Fricative Duration (sec); MRRtri = Maximum Repetition Rate of trisyllabic sequences (syllables/sec).

Appendix G. Intra-rater agreement: Two graduate research assistants.

Parameters	*Task*	*% Agree/Disagree*	*Mean % Agreement*
Severity	DEAP Phonology Single-Word subtest		98%
Omission errors	DEAP Phonology Single-Word subtest		100%
	CCR		99%
	WSD		100%
	FCD		100%
	ICD		100%
	MCD		99%
Vowel errors	DEAP Articulation Single-Word subtest (PVC)	100%	
Inconsistency	DEAP Word Inconsistency subtest	100%	
Use of simple syllable shapes	KSPT Complex Bisyllabics subtest & Polysyllabic Synthesis/Sequencing subtest		99%
Limited consonant and vowel repertoires	DEAP Articulation Single-Word subtest (PPC)	100%	
Increased errors on longer units of speech	KSPT Length and Complexity subtest		90%
Impaired volitional oral movements	DEAP Oral Motor Screen	100%	
Absent or shortened frication/affrication	DEAP Phonology Single-Word subtest (% affricate/fricative errors)		98%
	MFD	100%	
Reduced diadochokinetic rate	MRRtri	100%	

Note. DEAP = Diagnostic Evaluation of Articulation and Phonology; CCR = consonant cluster reduction; WSD = weak syllable deletion; FCD = final consonant deletion; ICD = initial consonant deletion; MCD = medial consonant deletion; PVC = Percentage of Vowels Correct; KSPT = Kaufman Speech Praxis Test for Children; PPC = Percentage of Phonemes Correct; MFD = Maximum Fricative Duration (sec); MRRtri = Maximum Repetition Rate of trisyllabic sequences (syllables/sec).

Appendix H. Data sheet used in scoring two of the 11 characteristics (presence/absence of suprasegmental errors and groping behaviors/silent posturing).

Suprasegmental Characteristics and Groping/Silent Posturing		
	IRB #:	
Suprasegmental Characteristics	Presence	Absence
Frog, Where are You? Part __ - _________		
Staccato speech; excessive/equal stress		
Word segregation; lengthened transitions/pausing between phonemes and syllables		
Slurred speech; limited word boundaries		
Monotone speech; minimal prosodic variation		
Appropriate suprasegmentals		
Suprasegmental Errors?	YES or NO	
Groping/ Silent Posturing	Presence	Absence
CELF-4 Sentence Repetition Part __ - _________		
Groping/Silent Posturing?	YES or NO	

Appendix I. Description of child performances on comprehensive assessment battery.

ID	C.A. (yrs.; mos.)	Group	DEAP Phon. SS	DEAP Phon. % Error CCR	DEAP Phon. % Error WSD	DEAP Phon. % Error FCD	DEAP Phon. % Error ICD	DEAP Phon. % Error MCD	DEAP Phon. % Error Affricates/ Fricatives
01	5;0	TD	8	0	0	0	0	0	0
02	5;4	TD	12	0	0	0	0	0	0
03	5;5	TD	11	0	0	0	0	0	7
04	5;6	TD	10	0	0	0	0	0	0
05	5;7	TD	9	0	0	0	0	0	5
06	5;9	TD	8	4	0	3	0	0	5
07	5;10	TD	9	0	0	0	0	0	5
08	6;0	TD	12	0	0	0	0	0	0
09	6;1	TD	12	0	0	0	0	0	0
10	6;2	TD	8	7	0	0	0	0	5
11	6;4	TD	12	0	0	0	0	0	0
12	6;6	TD	12	0	0	0	0	0	0
13	6;11	TD	8	4	0	0	0	0	7
14	6;11	TD	12	0	0	0	0	0	0
15	7;2	TD	10	0	0	0	0	0	2
16	7;4	TD	11	0	0	0	0	0	2
17	7;4	TD	9	4	0	0	0	0	0
18	7;6	TD	12	0	0	0	0	0	0
19	8;1	TD	11	4	0	3	0	0	0
20	8;5	TD	12	0	0	0	0	0	0
21	8;8	TD	8	0	0	3	0	0	2
22	5;0	sCAS	1	41	0	6	3	11	86
23	5;1	sCAS	3	22	0	28	0	11	57
24	5:1	sCAS	1	85	21	22	34	47	70
25	5;4	sCAS	1	85	0	0	7	5	93
26	5;4	sCAS	2	7	0	0	0	0	36
27	5;5	sCAS	4	11	0	3	0	0	25
28	5;6	sCAS	1	81	17	61	34	53	95
29	5;8	sCAS	1	74	13	28	3	26	59
30	5;10	sCAS	1	85	4	61	0	42	93
31	6;5	sCAS	4	30	4	0	0	0	36
32	6;10	sCAS	1	67	13	3	3	11	68

Note. ID = child identifier; C.A.= chronological age in years/months; TD = typically developing; sCAS = suspected childhood apraxia of speech; DEAP Phon.= Diagnostic Evaluation of Articulation and Phonology: Phonology Single-Word Production subtest; SS = scaled score; CCR = consonant cluster reduction; WSD = weak syllable deletion; FCD = final consonant deletion; ICD = initial consonant deletion; MCD = medial consonant deletion.

Appendix I continued.

ID	C.A. (yrs.; mos.)	Group	DEAP Artic. PVC	DEAP Artic. PPC	DEAP Oral Motor Screen	DEAP Word Inconsistency	KSPT Complex Bisyllabics & Polysyllabic Synthesis/ Sequencing % Syllable Reduction	KSPT Length & Complexity % Error
01	5;0	TD	94%	94%	+	+	0	0
02	5;4	TD	100%	100%	+	+	3	0
03	5;5	TD	100%	99%	+	+	0	0
04	5;6	TD	100%	100%	+	+	0	0
05	5;7	TD	100%	98%	+	+	0	0
06	5;9	TD	100%	94%	+	+	0	0
07	5;10	TD	94%	96%	+	+	0	20
08	6;0	TD	100%	100%	+	+	0	0
09	6;1	TD	97%	99%	+	+	0	0
10	6;2	TD	100%	98%	+	+	0	0
11	6;4	TD	100%	100%	+	+	0	0
12	6;6	TD	100%	100%	+	+	0	0
13	6;11	TD	100%	99%	+	+	0	0
14	6;11	TD	100%	100%	+	+	0	20
15	7;2	TD	100%	100%	+	+	0	0
16	7;4	TD	100%	100%	+	+	0	0
17	7;4	TD	94%	96%	-	+	3	0
18	7;6	TD	100%	100%	+	+	0	0
19	8;1	TD	100%	99%	+	+	0	0
20	8;5	TD	100%	99%	+	+	0	0
21	8;8	TD	100%	100%	+	+	0	20
22	5;0	sCAS	92%	60%	-	-	0	0
23	5;1	sCAS	96%	75%	-	+	35	80
24	5:1	sCAS	64%	51%	-	-	58	80
25	5;4	sCAS	94%	63%	-	+	16	40
26	5;4	sCAS	94%	76%	-	-	13	60
27	5;5	sCAS	94%	85%	+	+	0	20
28	5;6	sCAS	88%	53%	-	+	48	20
29	5;8	sCAS	92%	63%	-	+	48	20
30	5;10	sCAS	66%	39%	-	+	48	20
31	6;5	sCAS	94%	88%	-	+	6	0
32	6;10	sCAS	94%	58%	-	+	6	20

Note. ID = child identifier; C.A.= chronological age in years/months; TD = typically developing; sCAS = suspected childhood apraxia of speech; DEAP Artic.= Diagnostic Evaluation of Articulation and Phonology: Articulation Single-Word Production subtest; PVC = Percentage of Vowels Correct; PPC = Percentage of Phonemes Correct; - = did not meet criterion; + = met criterion; KSPT = Kaufman Speech Praxis Test for Children.

Appendix I continued.

ID	C.A. (yrs.; mos.)	Group	Connected Speech (*Frog, Where are you?*)	CELF-4 Recalling Sentences	MFD	MRRtri	CELF-4 Number Repetition (correct 2-digit repetition)
01	5;0	TD	1	-	3.6	4.58	√
02	5;4	TD	1	-	5.5	3.57	√
03	5;5	TD	1	-	7.7	4.60	√
04	5;6	TD	1	-	5.2	3.80	√
05	5;7	TD	1	-	7.7	4.44	√
06	5;9	TD	1	-	4.7	3.51	√
07	5;10	TD	1	-	6.1	4.11	√
08	6;0	TD	1	-	6.7	4.26	√
09	6;1	TD	1	-	12.3	4.51	√
10	6;2	TD	1	-	6.7	3.92	√
11	6;4	TD	1	-	7.2	5.38	√
12	6;6	TD	1	-	12.0	4.71	√
13	6;11	TD	1	-	8.3	3.39	√
14	6;11	TD	1	-	5.0	4.92	√
15	7;2	TD	1	-	8.1	5.33	√
16	7;4	TD	1	-	7.1	4.88	√
17	7;4	TD	1	-	6.2	4.82	√
18	7;6	TD	0	-	11.0	4.72	√
19	8;1	TD	1	-	6.0	4.41	√
20	8;5	TD	1	-	9.7	4.80	√
21	8;8	TD	1	-	15.9	4.26	√
22	5;0	sCAS	0	-	3.2	X	√
23	5;1	sCAS	0	-	6.1	1.78	√
24	5;1	sCAS	0	-	2.8	X	√
25	5;4	sCAS	1	-	4.4	X	√
26	5;4	sCAS	0	-	3.2	X	√
27	5;5	sCAS	0	-	6.7	X	√
28	5;6	sCAS	0	-	X	X	√
29	5;8	sCAS	0	-	2.8	X	√
30	5;10	sCAS	0	-	4.6	X	√
31	6;5	sCAS	0	+	2.7	1.97	√
32	6;10	sCAS	1	-	3.9	X	√

Note. ID = child identifier; C.A.= chronological age in years/months; TD = typically developing; sCAS = suspected childhood apraxia of speech; 0 = presence of suprasegmental errors; 1 = absence of suprasegmental errors; CELF-4 = Clinical Evaluation of Language Fundamentals-4; + = presence of groping/silent posturing; - = absence of groping/silent posturing; MFD = Maximum Fricative Duration (sec); MRRtri = Maximum Repetition Rate of trisyllabic sequences (syllables/sec); X = could not perform task; √ = met criterion.

Appendix J. Five different stimuli blocks of single-syllable non-word tokens for experimental task.

BLOCK A		BLOCK B		BLOCK C	
250 ms	530 ms	250 ms	530 ms	250 ms	250 ms
250 ms	330 ms	250 ms	290 ms	250 ms	370 ms
250 ms	490 ms	250 ms	410 ms	250 ms	410 ms
250 ms	250 ms	250 ms	490 ms	250 ms	530 ms
250 ms	370 ms	250 ms	370 ms	250 ms	450 ms
250 ms	290 ms	250 ms	250 ms	250 ms	290 ms
250 ms	450 ms	250 ms	330 ms	250 ms	490 ms
250 ms	410 ms	250 ms	450 ms	250 ms	330 ms
250 ms	490 ms	250 ms	410 ms	250 ms	530 ms
250 ms	250 ms	250 ms	330 ms	250 ms	450 ms
250 ms	410 ms	250 ms	490 ms	250 ms	410 ms
250 ms	530 ms	250 ms	290 ms	250 ms	490 ms
250 ms	330 ms	250 ms	450 ms	250 ms	250 ms
250 ms	450 ms	250 ms	370 ms	250 ms	330 ms
250 ms	370 ms	250 ms	250 ms	250 ms	290 ms
250 ms	290 ms	250 ms	530 ms	250 ms	370 ms

BLOCK D		BLOCK E	
250 ms	530 ms	250 ms	250 ms
250 ms	370 ms	250 ms	410 ms
250 ms	410 ms	250 ms	530 ms
250 ms	330 ms	250 ms	490 ms
250 ms	250 ms	250 ms	330 ms
250 ms	290 ms	250 ms	290 ms
250 ms	450 ms	250 ms	450 ms
250 ms	490 ms	250 ms	370 ms
250 ms	250 ms	250 ms	410 ms
250 ms	370 ms	250 ms	330 ms
250 ms	490 ms	250 ms	290 ms
250 ms	450 ms	250 ms	450 ms
250 ms	530 ms	250 ms	250 ms
250 ms	290 ms	250 ms	370 ms
250 ms	410 ms	250 ms	490 ms
250 ms	330 ms	250 ms	530 ms

Appendix K. Sample data sheet used to record child's "same" and "different" responses for experimental task.

Part 1 Training: PICTURES

(OBJECTS) Feedback Given *"You are going to see two pictures on a piece of paper. Say if they are the same or different."* *Correct response: "Good, you are right, they are [the same/different]."* *Incorrect response: "No, these pictures are [the same/different]. Let's look at another one."*				
#	**ITEMS**	**Response**	***ERROR***	**NOTES**
1.	Pizza/Pizza	**S** D	*E*	
2.	House/Mouse	S **D**	*E*	
3.	Cat/Cat	**S** D	*E*	
4.	Dog/Dog	**S** D	*E*	
5.	Doll/Ball	S **D**	*E*	
6.	Apple/Banana	S **D**	*E*	
After child responds and is given feedback for each answer to the items above, go to the next four items below, providing no feedback.				
(SHAPES) No Feedback Given				
#	**ITEMS**	**Response**	***ERROR***	**NOTES**
1.	Circle/Circle	**S** D	*E*	
2.	Triangle/Triangle	**S** D	*E*	
3.	Circle/Triangle	S **D**	*E*	
4.	Circle/Square	S **D**	*E*	
Child will subsequently be shown a card with two squares on one half and a circle and a square on the other half and asked to point to "same" and to "different", in preparation for the experimental task.				
Note. Bolded and shaded "S" and "D" (**S** and **D)** indicate which is correct response; "*E*" is circled if response is incorrect.				

Part 2 Training: NON-WORDS

(NON-WORDS) Feedback Given *"Now you are going to hear a robot say two sounds. Point to the 'same' picture if the two sounds are the same, and point to the 'different' picture if the sounds are different (different picture on the right of the child).* *Now I am going to put the headphones on you. When you are ready to begin, give me a thumbs-up sign and I will play the sounds."* *Correct response: "Good, you are right, they are [the same/different]."* *Incorrect response: "No, this one is [the same/different]. Let's listen to another."*				
#	**ITEMS**	**Response**	***ERROR***	**NOTES**
1.	bo bo*	**S** D	*E*	
2.	bo BO**	S **D**	*E*	
3.	bo Bo***	S **D**	*E*	
4.	bo bo	**S** D	*E*	
5.	bo BO	S **D**	*E*	
After child responds and is given feedback for each answer to the items above, go to next five items below, providing no feedback. **Child must get 5/5 correct in order to proceed.**				
(WORDS) No Feedback Given				
#	**ITEMS**	**Response**	***ERROR***	**NOTES**
1.	bo BO	S **D**	*E*	
2.	bo bo	**S** D	*E*	
3.	bo Bo	S **D**	*E*	
4.	bo BO	S **D**	*E*	
5.	bo bo	**S** D	*E*	
Child must get 4/5 correct in order to proceed.				
Note. * indicates same duration difference of 250 ms; **indicates maximally long duration difference of 530 ms; *** indicates intermediate duration difference of 370 ms.				

Version 1: ABCDE (frog, turtle, rabbit, monkey, frog)

Version 2: DAECB (monkey, frog, frog, rabbit, turtle)

Version 3: BCADE (turtle, rabbit, frog, monkey, frog)

Version 4: CBAED (rabbit, turtle, frog, frog, monkey)

Part 3: STIMULUS ITEMS

	Block A: Frog #1			**Block B: Turtle #2**	
#	Response	Error (/)	#	Response	Error (/)
1	S **D**	*E*	1	S **D**	*E*
2	S **D**	*E*	2	S **D**	*E*
3	S **D**	*E*	3	S **D**	*E*
4	**S** D	*E*	4	S **D**	*E*
5	S **D**	*E*	5	S **D**	*E*
6	S **D**	*E*	6	**S** D	*E*
7	S **D**	*E*	7	S **D**	*E*
8	S **D**	*E*	8	S **D**	*E*
9	S **D**	*E*	9	S **D**	*E*
10	**S** D	*E*	10	S **D**	*E*
11	S **D**	*E*	11	S **D**	*E*
12	S **D**	*E*	12	S **D**	*E*
13.	S **D**	*E*	13.	S **D**	*E*
14.	S **D**	*E*	14.	S **D**	*E*
15	S **D**	*E*	15.	**S** D	*E*
16	S **D**	*E*	16.	S **D**	*E*
Notes:			Notes:		

	Block C: Rabbit #3	
#	Response	Error (/)
1	**S** D	*E*
2	S **D**	*E*
3	S **D**	*E*
4	S **D**	*E*
5	S **D**	*E*
6	S **D**	*E*
7	S **D**	*E*
8	S **D**	*E*
9	S **D**	*E*
10	S **D**	*E*
11	S **D**	*E*
12	S **D**	*E*
13	**S** D	*E*
14	S **D**	*E*
15	S **D**	*E*
16	S **D**	*E*
Notes:		

	Block D: Monkey #4	
#	Response	Error (/)
1	S **D**	*E*
2	S **D**	*E*
3	S **D**	*E*
4	S **D**	*E*
5	**S** D	*E*
6	S **D**	*E*
7	S **D**	*E*
8	S **D**	*E*
9	**S** D	*E*
10	S **D**	*E*
11	S **D**	*E*
12	S **D**	*E*
13	S **D**	*E*
14	S **D**	*E*
15	S **D**	*E*
16	S **D**	*E*
Notes:		

	Block E: Frog #5	
#	Response	Error (/)
1	**S** D	*E*
2	S **D**	*E*
3	S **D**	*E*
4	S **D**	*E*
5	S **D**	*E*
6	S **D**	*E*
7	S **D**	*E*
8	S **D**	*E*
9	S **D**	*E*
10	S **D**	*E*
11	S **D**	*E*
12	S **D**	*E*
13	**S** D	*E*
14	S **D**	*E*
15	S **D**	*E*
16	S **D**	*E*
Notes:		

Scripts for each block
(ABCDE: frog, turtle, rabbit, monkey, frog)

1. BLOCK A - FROG

Mr. Frog needs your help to get across the pond! To help him, you will listen to two sounds and decide if they are the same or different.

If the words are the same, point to this picture of two black squares and say 'same'. If the words are different, point to this picture of a black square and red circle and say 'different'.

Listen carefully because the sounds cannot be repeated.

2. BLOCK B - TURTLE

Mr. Turtle wants to cross a river and needs your help! To help him, you will listen to two sounds and decide if they are the same or different.

Remember… If the words are the same, point to this picture of two black squares and say 'same'. If the words are different, point to this picture of a black square and red circle and say 'different'.

3. BLOCK C - RABBIT

Mr. Rabbit is hungry and needs your help! To help him, you will listen to two sounds and decide if they are the same or different.

Remember… If the words are the same, point to this picture of two black squares and say 'same'. If the words are different, point to this picture of a black square and red circle and say 'different'.

4. BLOCK D - MONKEY

Mr. Monkey is hungry and needs your help! To help him, you will listen to two sounds and decide if they are the same or different.

Remember… If the words are the same, point to this picture of two black squares and say 'same'. If the words are different, point to this picture of a black square and red circle and say 'different'.

5. BLOCK E - FROG

Mr. Frog needs your help to get across the pond! To help him, you will listen to two sounds and decide if they are the same or different.

Remember… If the words are the same, point to this picture of two black squares and say 'same'. If the words are different, point to this picture of a black square and red circle and say 'different'.

Appendix L. Definitions of response accuracy as percent correct found in the research literature.

Definition of Response Accuracy	Citation	Research
75% differential threshold	Nooteboom & Doodeman (1980)	Vowel duration (adult)
50% difference limen	Elliott, Busse, Partridge, Rupert, & DeGraaff (1986)	VOT (child/adult)
70.7% correct response point	Elfenbein, Small, & Davis (1993)	Duration discrimination (child/adult)
70.7% correct response point	Jensen & Neff (1993)	Auditory discrimination (child/adult)
69% accuracy	Kraus, Koch, McGee, Nicol, & Cunningham (1999)	Formant transition duration discrimination (child)
75% accuracy	Kato, Tsuzaki, & Sagisaka (2003)	Duration discrimination (adult)
Discriminability of 50% of the maximum discriminability value	Maassen, Groenen, & Crul (2003)	Vowel perception (child)
75% accuracy	Himpel, Banaschewski, Grüttner, Becker, Heise, Uebel, Albrecht, Rothenberger, & Rammsayer (2009)	Duration discrimination (child)
75% accuracy	Lidestam (2009)	Vowel duration discrimination (adult)
71% accuracy	Kawai & Carrell (2012)	Phoneme duration discrimination (adult)
50% points on the discrimination functions	Nittrouer, Lowenstein, & Tarr (2013)	Amplitude/spectral discrimination (child/adult)

References

Abel, S. M. (1972). Duration discrimination of noise and tone bursts. *The Journal of the Acoustical Society of America, 51*(4), 1219-1223.

Adobe Premiere Elements (Version 10.0) [Computer software]. (2011). Adobe Systems, Inc.

Alcock, K. J., Passingham, R. E., Watkins, K., & Vargha-Khadem, F. (2000). Pitch and timing abilities in inherited speech and language impairment. *Brain and Language, 75*, 34-46.

Allen, G. D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.*), Child phonology*-*volume 1*-*production* (pp. 227-256). New York: Academic Press, Inc.

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

Aram, D. M. (1984). Assessment and treatment of developmental apraxia. *Seminars in Speech and Language*, *5*(2).

Austin, D., & Shriberg, L. D. (1997). *Lifespan reference data for ten measures of articulation competence using the speech disorders classification system (SDCS)* (Phonology Project Technical Report No. 3).

Bastiaanse, R., Gilbers, D., & van der Linde, K. (1994). Sonorant sustainment: A single case study of a rare phonemic deficit in aphasia. *Clinical Linguistics & Phonetics*, *8*(1), 27-43.

Bernhardt, B., & Gilbert, J. (1992). Applying linguistic theory to speech-language pathology: The case for nonlinear phonology. *Clinical Linguistics & Phonetics*, *6*(1-2), 123-145.

Bernhardt, B., & Stoel-Gammon, C. (1994). Nonlinear phonology: Introduction and clinical application. *Journal of Speech and Hearing Research*, *37*, 123-143.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.3) [Computer program]. Retrieved from http://www.praat.org/

Boutsen, F. R., & Christman, S. S. (2002). Prosody in apraxia of speech. *Seminars in Speech and Language, 23*(4), 245-256.

Bridgeman, E., & Snowling, M. (1988). The perception of phoneme sequence: A comparison of dyspraxic and normal children. *British Journal of Disorders of Communication, 23*, 245-252.

Burns, M. S. (2011). Apraxia of speech in children and adolescents: Applications of neuroscience to differential diagnosis and intervention. *SIG 2 Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders, 21*(1), 15-32.

Campbell, W. N. (1989). Syllable level duration determination. *Eurospeech 89*, *2*, 698-701.

Chistovich, L. A. (1959). Discrimination of the time interval between two short acoustic pulses. *Soviet Physics-Acoustics*, *5*, 480-484.

Clopper, C. G., Pisoni, D. B., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America*, *118*(3), 1661-1676.

Crary, M. A. (1984). Phonological characteristics of developmental verbal dyspraxia. *Seminars in Speech and Language*, *5*, 71-83.

Crary, M. A. (1993). *Developmental motor speech disorders*. San Diego, CA: Singular Publishing Group, Inc.

Creelman, C. D. (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, *34*(5), 582-593.

Dale, P. S., & Hayden, D. A. (2013). Treating speech subsystems in childhood apraxia of speech with tactual input: The PROMPT approach. *American Journal of Speech-Language Pathology, 22*, 644-661.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51-62.

Davis, B. L., Jakielski, K. J., & Marquardt, T. P. (1998). Developmental apraxia of speech: Determiners of differential diagnosis. *Clinical Linguistics & Phonetics*, *12*(1), 25-45.

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research, 38*, 1199-1211.

Davis, B. L., MacNeilage, P. F., Matyear, C. L., & Powell, J. K. (2000). Prosodic correlates of stress in babbling: An acoustical study. *Child Development*, *71*(5), 1258-1270.

Davis, B. L., & Velleman, S. L. (2000). Differential diagnosis and treatment of developmental apraxia of speech in infants and toddlers. *Infant-Toddler Intervention: The Transdisciplinary Journal*, *10*(3), 177-192.

DePaolis, R. A., Vihman, M. M., & Kunnari, S. (2008). Prosody in production at the onset of word use: A cross-linguistic study. *Journal of Phonetics, 36*, 406-422.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic evaluation of articulation and phonology (DEAP)*. London: Psychology Corporation.

Doherty, C. P., Fitzsimons, M., Asenbauer, B., & Staunton, H. (1999). Discrimination of prosody and music by normal children. *European Journal of Neurology*, *6*(2), 221-226.

Echols, C. H. (1993). A perceptually-based model of children's earliest productions. *Cognition*, *46*, 245-296.

Elfenbein, J. L., Small, A. M., & Davis, J. M. (1993). Developmental patterns of duration discrimination. *Journal of Speech and Hearing Research, 36*, 842-849.

Elliott, L. L., Busse, L. A., Partridge, R., Rupert, J., & DeGraaff, R. (1986). Adult and child discrimination of CV syllables differing in voicing onset time. *Child Development, 57*, 628-635.

Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French, and English. *Journal of Phonetics*, *19*, 351-365.

Field, A. (2005). *Discovering statistics using SPSS*. London: Sage Publications.

Forrest, K. (2003). Diagnostic criteria of developmental apraxia of speech used by clinical speech-language pathologists. *American Journal of Speech-Language Pathology, 12*, 376-380.

Gillon, G. T., & Moriarty, B. C. (2007). Childhood apraxia of speech: Children at risk for persistent reading and spelling disorder. *Seminars in Speech and Language, 28*(1), 48-57.

Goffman, L. (1999). Prosodic influences on speech production in children with specific language impairment and speech deficits: Kinematic, acoustic, and transcription evidence. *Journal of Speech, Language, and Hearing Research, 42*, 1499-1517.

Groenen, P., Maassen, B., Crul, T., & Thoonen, G. (1996). The specific relation between perception and production errors for place of articulation in developmental apraxia of speech. *Journal of Speech and Hearing Research, 39*, 468-482.

Guyette, T. W., & Diedrich, W. M. (1981). A critical review of developmental apraxia of speech. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 5, pp. 1-49). New York: Academic Press.

Hall, P. K. (1992). At the center of controversy: Developmental apraxia. *American Journal of Speech-Language Pathology, 1*, 23-25.

Hall, P. K. (2007). Speech characteristics of developmental apraxia of speech. *Developmental apraxia of speech: Theory and clinical practice* - *2nd edition*. Austin, TX: Pro-Ed.

Hall, P. K., Jordan, L. S., & Robin, D. A. (1993). *Developmental apraxia of speech: Theory and clinical practice*. Austin, TX: Pro-Ed.

Hämäläinen, J. A., Leppänen, P. H. T., Eklund, K., Thomson, J., Richardson, U., Guttorm, T. K., ... & Lyytinen, H. (2009). Common variance in amplitude envelope perception tasks and their impact on phoneme duration perception and reading and spelling in Finnish children with reading disabilities. *Applied Psycholinguistics, 30*, 511-530.

Henry, C. E. (1990). The development of oral diadochokinesia and non-linguistic rhythmic skills in normal and speech-disordered young children. *Clinical Linguistics & Phonetics, 4*(2), 121-137.

Henry, F. M. (1948). Discrimination of the duration of a sound. *Journal of Experimental Psychology, 38*(6), 734-743.

Highman, C., Hennessey, N., Sherwood, M., & Leitão, S. (2008). Retrospective parent report of early vocal behaviours in children with suspected childhood apraxia of speech (sCAS). *Child Language Teaching and Therapy, 24*(3), 285-306.

Himpel, S., Banaschewski, T., Grüttner, A., Becker, A., Heise, A., Uebel, H., ... & Rammsayer, T. (2009). Duration discrimination in the range of milliseconds and seconds in children with ADHD and their unaffected siblings. *Psychological Medicine, 39*, 1745-1751.

Huggins, A. W. F. (1972). Just noticeable differences for segment duration in natural speech. *The Journal of the Acoustical Society of America, 51*(4), 1270-1278.

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science*, *4*(2), 104-107.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675-687.

Kato, H., Tsuzaki, M., & Sagisaka, Y. (2003). Functional differences between vowel onsets and offsets in temporal perception of speech: Local-change detection and speaking-rate discrimination. *The Journal of the Acoustical Society of America, 113*(6), 3379-3389.

Kaufman, N. R. (1995). *The Kaufman speech praxis test for children*. Detroit, MI: Wayne State University Press.

Kawai, N., & Carrell, T. D. (2012). Discrimination of differences in digitally manipulated phoneme length during speech. *Perceptual and Motor Skills, 114*(1), 189-203.

Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research, 38*, 338-350.

Kenney, M. K., Barac-Cikoja, D., Finnegan, K., Jeffries, N., & Ludlow, C. L. (2006). Speech perception and short-term memory deficits in persistent developmental speech disorder. *Brain and Language, 96*, 178-190.

Koike, K. J. M., & Asp, C. W. (1981). Tennessee test of rhythm and intonation patterns. *Journal of Speech and Hearing Disorders, 46*(1), 81-87.

Kraus, N., Koch, D. B., McGee, T. J., Nicol, T. G., & Cunningham, J. (1999). Speech-sound discrimination in school-age children: Psychophysical and neurophysiologic measures. *Journal of Speech, Language, and Hearing Research, 42*, 1042-1060.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America, 105*(3), 1455-1468.

Lehiste, I. (1970). *Suprasegmentals*. Oxford, England: MIT Press.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics, 5*, 253-263.

Lehiste, I. (1996). Suprasegmental features of speech. In N. J. Lass (Ed.), *Principles of experimental phonetics* (pp. 226-244). St. Louis: Mosby, Inc.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Iyengar, S. K., & Taylor, H. G. (2004). School-age follow-up of children with childhood apraxia of speech. *Language, Speech, and Hearing Services in Schools, 35*, 122-140.

Lidestam, B. (2009). Visual discrimination of vowel duration. *Scandinavian Journal of Psychology, 50*, 427-435.

Lindblom, B. (1983). Economy of speech gestures. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 217-245). New York: Springer-Verlag.

Maassen, B. (2002). Issues contrasting adult acquired versus developmental apraxia of speech. *Seminars in Speech and Language, 23*(4), 257-266.

Maassen, B., Groenen, P., & Crul, T. (2003). Auditory and phonetic perception of vowels in children with apraxic speech disorders. *Clinical Linguistics & Phonetics, 17*(6), 447-467.

Maassen, B., Nijland, L., & van der Meulen, S. (2001). Coarticulation within and between syllables by children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 15*(1&2), 145-150.

MacNeilage, P. F., & Davis, B. (1989). Acquisition of speech production: Frames, then content. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control (Attention and performance series)* (pp. 453-476). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

MacNeilage, P. F., & Davis, B. L. (1990). Acquisition of speech production: The achievement of segmental independence. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 55-68). Netherlands: Kluwer Academic Publishers.

Marion, M. J., Sussman, H. M., & Marquardt, T. P. (1993). The perception and production of rhyme in normal and developmentally apraxic children. *Journal of Communication Disorders, 26*, 129-160.

Marquardt, T. P., & Sussman, H. M. (1991). Developmental apraxia of speech: Theory and practice. In D. Vogel & M. Cannito (Eds.), *Treating disordered speech motor control* (pp. 341-390). Austin, TX: Pro-Ed.

Marquardt, T. P., Sussman, H. M., Snow, T., & Jacks, A. (2002). The integrity of the syllable in developmental apraxia of speech. *Journal of Communication Disorders, 35*, 31-49.

Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Books for Young Readers.

McCabe, P., Rosenthal, J. B., & McLeod, S. (1998). Features of developmental dyspraxia in the general speech-impaired population? *Clinical Linguistics & Phonetics*, *12*(2), 105-126.

McCauley, R. J., & Strand, E. A. (2008). A review of standardized tests of nonverbal oral and speech motor performance in children. *American Journal of Speech-Language Pathology*, *17*, 81-91.

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (1997). Apraxia of speech: Definition, differentiation, and treatment. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 311-344). New York: Thieme.

Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology, 44*, 413-426.

Munson, B., Bjorum, E. M., & Windsor, J. (2003). Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. *Journal of Speech, Language, and Hearing Research, 46*, 189-202.

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*, 574-585.

Nijland, L. (2009). Speech perception in children with speech output disorders. *Clinical Linguistics & Phonetics*, *23*(3), 222-239.

Nijland, L., Maassen, B., van der Meulen, S., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, *16*(6), 461-483.

Nijland, L., Maassen, B., van der Meulen, S., Gabreëls, F., Kraaimaat, F. W., & Schreuder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, *17*(1), 1-24.

Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., & Trent-Brown, S. A. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America, 124*(1), 576-588.

Nittrouer, S., Lowenstein, J. H., & Tarr, E. (2013). Amplitude rise time does not cue the /ba/-/wa/ contrast for adults or children. *Journal of Speech, Language, and Hearing Research, 56*, 427-440.

Nittrouer, S., Shune, S., & Lowenstein, J. H. (2011). What is the deficit in phonological processing deficits: Auditory sensitivity, masking, or category formation*? Journal of Experimental Child Psychology, 108,* 762-785.

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *The Journal of the Acoustical Society of America, 67*(1), 276-287.

Odell, K. H., & Shriberg, L. D. (2001). Prosody-voice characteristics of children and adults with apraxia of speech. *Clinical Linguistics & Phonetics, 15*(4), 275-307.

Ohde, R. N., & German, S. R. (2011). Formant onsets and formant transitions as developmental cues to vowel perception. *The Journal of the Acoustical Society of America, 130*(3), 1628-1642.

Ozanne, A. (1995). The search for developmental verbal dyspraxia. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorders* (pp. 91-109). London: Whurr Publishers.

Ozanne, A. (2005). Childhood apraxia of speech. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorders* (2nd ed.). London: Whurr Publishers Ltd.

Paul, R., Bianchi, N., Augustyn, A., Klin, A., & Volkmar, F. R. (2008). Production of syllable stress in speakers with autism spectrum disorders. *Research in Autism Spectrum Disorders, 2*, 110-124.

Peter, B., & Stoel-Gammon, C. (2005). Timing errors in two children with suspected childhood apraxia of speech (sCAS) during speech and music-related tasks. *Clinical Linguistics & Phonetics, 19*(2), 67-87.

Peter, B., & Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics, 22*(3), 171-198.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32*(6), 693-703.

Pollock, K. E., Brammer, D. M., & Hageman, C. F. (1993). An acoustic analysis of young children's productions of word stress. *Journal of Phonetics, 21*, 183-203.

Pollock, K. E., & Hall, P. K. (1991). An analysis of the vowel misarticulations of five children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 5*(3), 207-224.

Preston, J. L., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology, 22*, 627-643.

Rammsayer, T. H., & Brandler, S. (2004). Aspects of temporal information processing: A dimensional analysis. *Psychological Research, 69*, 115-123.

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception and Psychophysics, 50*(6), 565-574.

Robb, M. P. (2010). *INTRO A guide to communication sciences and disorders.* San Diego, CA: Plural Publishing, Inc.

Robin, D. A. (1992). Developmental apraxia of speech: Just another motor problem. *American Journal of Speech-Language Pathology, 1*, 19-22.

Rodriguez, R. H. (1998). *Acoustic and perceptual comparisons of imitative prosody in kindergartners with and without speech disorders.* (Unpublished master's thesis). University of South Florida, Tampa.

Rosenbek, J. C., & Wertz, R. T. (1972). A review of fifty cases of developmental apraxia of speech. *Journal of Speech and Hearing Disorders, 3*, 23-33.

Rosenthal, R. (1991). *Meta-analytic procedures for social research (Vol. 6)*. Newbury Park, CA: Sage Publications.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals - 4th edition (CELF-4*). London: Psychological Corporation.

Shields, J. R. (1981). A study of rhythmic abilities and language abilities in young children. *First Language, 2*(5), 131-140.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997a). Developmental apraxia of speech: I. Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research, 40*, 273-285.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997b). Developmental apraxia of speech: II. Toward a diagnostic marker. *Journal of Speech, Language, and Hearing Research, 40,* 286-312.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997c). Developmental apraxia of speech: III. A subtype marked by inappropriate stress. *Journal of Speech, Language, and Hearing Research, 40*, 313-337.

Shriberg, L. D., Campbell, T. F., Karlsson, H. B., Brown, R. L., McSweeny, J. L., & Nadler, C. J. (2003). A diagnostic marker for childhood apraxia of speech: The lexical stress ratio. *Clinical Linguistics & Phonetics, 17*(7), 549-574.

Shriberg, L. D., Green, J. R., Campbell, T. F., McSweeny, J. L., & Scheer, A. R. (2003). A diagnostic marker for childhood apraxia of speech: The coefficient of variation ratio. *Clinical Linguistics & Phonetics, 17*(7), 575-595.

Shriberg, L. D., Kwiatkowski, J., Rasmussen, C., Lof, G. L., & Miller, J. F. (1992). *The prosody-voice screening profile (PVSP): Psychometric data and reference information for children* (Phonology Project Technical Report No.1).

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The syllable repetition task (SRT). *Journal of Speech, Language, and Hearing Research, 52*, 1189-1212.

Shriberg, L. D., Lohmeier, H. L., Strand, E. A., & Jakielski, K. J. (2012). Encoding, memory, and transcoding deficits in childhood apraxia of speech. *Clinical Linguistics & Phonetics, 26*(5), 445-482.

Simpson, A. P., & Ericsdotter, C. (2003). *Sex-specific durational differences in English and Swedish*. Paper presented at the 15th International Congress of Phonetic Sciences, Barcelona.

Skinder, A. E. (2000). *The relationship of prosodic and articulatory errors produced by children with developmental apraxia of speech.* (Unpublished doctoral dissertation). University of Washington, Seattle.

Skinder, A., Connaghan, K., Strand, E., & Betz, S. (2000). Acoustic correlates of perceived lexical stress errors in children with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology, 8*(4), 279-284.

Skinder, A., Strand, E. A., & Mignerey, M. (1999). Perceptual and acoustic analysis of lexical and sentential stress in children with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology, 7*(2), 133-144.

Skinder-Meredith, A., Stoel-Gammon, C., Wright, R., & Strand, E. (2001, November). *The effects of articulatory accuracy on prosody in children with developmental apraxia of speech*. Poster session conducted at the meeting of the ASHA National Convention, New Orleans, LA.

Sluijter, A. M. C., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America, 100*(4), 2471-2485.

Smith, A. B., & Robb, M. P. (2006). The influence of utterance position on children's production of lexical stress. *Folia Phoniatrica et Logopaedica, 58*, 199-206.

Stackhouse, J., & Snowling, M. (1992). Barriers to literacy development in two cases of developmental verbal dyspraxia. *Cognitive Neuropsychology, 9*(4), 273-299.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Terband, H., Maassen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech (CAS). *Journal of Speech, Language, and Hearing Research, 52*, 1595-1609.

Thoonen, G., Maassen, B., Gabreëls, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics, 13*(1), 1-23.

Thoonen, G., Maassen, B., Gabreëls, F., Schreuder, R., & de Swart, B. (1997). Towards a standardised assessment procedure for developmental apraxia of speech. *European Journal of Disorders of Communication, 32*, 37-60.

Thoonen, G., Maassen, B., Wit, J., Gabreëls, F., & Schreuder, R. (1996). The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. *Clinical Linguistics & Phonetics, 10*(4), 311-336.

Velleman, S. L., & Shriberg, L. D (1999). Metrical analysis of the speech of children with suspected developmental apraxia of speech. *Journal of Speech, Language, and Hearing Research, 42*, 1444-1460.

Velleman, S. L., & Strand, K. (1994). Developmental verbal dyspraxia. In J. E. Bernthal & N. W. Bankson (Eds.), *Child phonology: Characteristics, assessment, and interventions with special populations* (pp.110-139). New York: Thieme Medical Pub.

Vihman, M. M., DePaolis, R. A., & Davis, B. L. (1998). Is there a "trochaic bias" in early word learning? Evidence from infant production in English and French. *Child Development*, *69*(4), 935-949.

Wood, C. C. (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *The Journal of the Acoustical Society of America, 60*(6), 1381-1389.

Yoss, K. A., & Darley, F. L. (1974). Developmental apraxia of speech in children with defective articulation. *Journal of Speech and Hearing Research, 17*(3), 399-416.