Spring 2019

Segmentation and Intonation in Childhood Apraxia of Speech

Rachel Platt
University of New Hampshire, Durham

SEGMENTATION AND INTONATION IN CHILDHOOD APRAXIA OF SPEECH

BY

RACHEL TAYLOR PLATT
B.S. Communication Sciences and Disorders, University of New Hampshire, 2017

MASTERS THESIS

Submitted to the University of New Hampshire
In Partial Fulfillment of
the Requirements for the Degree of

Master of Science
in
Communication Sciences and Disorders

May 6, 2019

This thesis has been examined and approved in partial fulfillment of the requirements of the degree of Master of Science in Communication Sciences and Disorders by:

Thesis Director, Jill C. Thorson, PhD, Assistant Professor, Communication, Acquisition and Translational Studies Laboratory (CAT Lab), Director, University of New Hampshire

Donald A. Robin, PhD, Professor and Chair, Department of Communication Sciences and Disorders, University of New Hampshire

Kirrie J. Ballard, PhD, Professor, The University of Sydney, Australia

On April 19, 2019

Original approval signatures are on file with the University of New Hampshire Graduate School.

TABLE OF CONTENTS

TITLE **PAGE**

CHAPTER **PAGE**

ACKNOWLEDGEMENTS

Thank you to my advisor Dr. Jill Thorson for your support and positive motivation throughout my project. Thank you to all my committee members Dr. Thorson, Dr. Robin, and Dr. Ballard. Special thanks to the participants and their families who were part of the study. Thank you also to Lauren Gabriel and Mikaela Coombs for their help with data coding and analysis, I would not have finished without you! Thank you to the members of the CAT Lab at UNH. Thanks to Hilary Miller for recruiting CAS participants through TEMPO at UNH. And last, thank you to my mom, two sisters, and Adam for all your support in me along the way.

DEDICATION

To my mom for all your love and all have taught me.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

SEGMENTATION AND INTONATION IN CHILDHOOD APRAXIA OF SPEECH

by

Rachel Taylor Platt

University of New Hampshire, May 6, 2019

Childhood apraxia of speech (CAS) is a motor speech disorder that affects the programming of spatial and temporal parameters for speech patterns, characterized by sound distortions, segmented units, and deficits with lexical stress. CAS has notable increases in the length of time between speech segments and within syllables than do children with phonological impairments or who are developing typically. This segmentation may impact prosody at the lexical level. Prosody also includes declination of fundamental frequency (F0) and reset at the intonational level, impacting the intelligibility of speech production.

This study assessed segmentation and intonational effects on prosody across an entire utterance for 11 children with CAS and 10 typically-developing children (TD) aged 5-11-years-old. Acoustic analyses of real and non-word multisyllabic words, paired with a carrier phrase of 3-4 words, were conducted for average inter-segment duration between and within words (ms) and average slope of F0. Stimuli were generated from Treating Establishment of Motor Program Organization (TEMPO), which targets motor speech errors in CAS (Miller et al., 2018). The current study provides a TD comparison to the CAS group from the TEMPO study prior to treatment.

Results showed CAS participants produced significantly longer inter-segment durations between words and within words. A correlation analysis concluded a strong positive relationship between inter-segment duration and number of words in the sentence. These data found F0 declination over utterances for both groups, with no detectable difference between groups. F0 change over target words did not show notable declination differences between groups. Further correlation testing suggests that as between-word duration increases, F0 regression slope over utterances flattens. Comparing speech production patterns in CAS with TD children pre- and post-treatment will better establish treatment efficacy in improving communication of children with CAS. Further data from additional participants will better differentiate prosodic intonation between TD children and children with CAS.

1. INTRODUCTION

Childhood apraxia of speech (CAS) is a motor speech disorder, that affects the programming of spatial and temporal parameters for speech movement patterns, characterized by sound distortions, segmented units, and deficits with lexical stress (McNeil, Robin, & Schmidt, 1997). Children with CAS demonstrate speech errors when producing consonants and vowels, and experience deviations in speech prosody, stress, and timing (ASHA, 2007). CAS is differentiated from other speech sound disorders because of these characteristics (Murray, McCabe, Heard, & Ballard, 2015). Prosody and segmentation both impact the intelligibility and naturalness of speech production. The issue to consider is if the prosodic abnormalities seen in CAS are a direct result from the segmentation characteristics or if they are separate phenomena.

1.1. Childhood apraxia of speech

CAS is a motor speech disorder that develops when a child is first learning language and follows the same diagnostic characteristics of apraxia of speech (AOS), also a motor speech disorder, acquired later in life from a neuromuscular accident. AOS is characterized by increased intersegment durations, between sounds and syllables, distortions of speech sounds, and dysprosody (Seddoh, Robin, Sim, Hageman, Moon, & Folkins, 1996). Motor speech disorders, like CAS and AOS, define speech motor programming as the set of processes responsible for coding an abstract linguistic (phonological) message into a timed pattern of muscle contractions to produce the desired speech movement (Maas, Robin, Wright, & Ballard, 2008).

The diagnostic characteristic of sound and syllable segmentation has been investigated and measured for individuals with AOS (Kent & Rosenbek, 1983; Seddoh et al., 1996; Maas et al., 2008). To date, research on AOS has introduced models of motor programming to explain the programming breakdowns, which will be discussed further on with reference to the segmentation phenomena. Other theories investigating the perceptual and acoustical characteristics of AOS have also published on how feedback on word production affects the production and movement of speech patterns, explaining that disordered speech associated with AOS is due to impairments to the motor speech network (Ballard, Tourville, & Robin, 2014). The behavioral feature Ballard et al. (2014) analyzed in AOS was syllable segregation, which continues to be measured and used as a diagnostic marker of the disorder.

Murray et al. (2015) show that syllable segregation was moderately sensitive and specific in differentiating children with CAS from those with a different speech sound disorder, suggesting that syllable segregation has strong potential as a core feature of CAS. Iuzzini-Seigel and Murray (2017) found that fifteen out of twenty school-aged participants with CAS evidenced syllable segregation at least one-time during responses on the Goldman Fristoe Test of Articulation (GFTA-2) (Goldman & Fristoe, 2000), whereas only three out of ten participants with speech delay evidenced this feature. Another study used single word production to investigate how motor programming deficits impact the production of different behavioral measures of linguistic, cognitive, non-speech oral, and speech motor ability in adults with AOS. Ballard et al. (2016) found that these individuals have considerable difficulty with multisyllabic words because of the required motor programming needed to produce a long string of syllables.

An intervention program is currently undergoing pilot trials to target the treatment of multisyllabic word production in individuals with CAS and AOS. Treating Establishment of Motor Program Organization (TEMPO) is a motor speech disorder treatment designed and implemented to target segmentation, speech sound precision, and prosodic accuracy (Miller, Plante, Campbell, Ballard, & Robin, 2018). This study tests both acoustic and perceptual measures of sound distortions, segmentation, and lexical stress of a series of treated and untreated stimuli for children with CAS. Stimuli consisted of a collection of non-words using stop plosives and fricatives, which were used for baseline probes, a four-week treatment session at four one-hour sessions, and two post treatment probes. Miller et al. (2018) found TEMPO successful in improving speech sound production, reducing segmentation (intrasegment durations between syllables), and in improving lexical stress in children with CAS.

Like this treatment trial, other studies have investigated how lexical stress is influenced by other speech characteristics (Skinder, Connaghan, Strand, & Betz, 2000; Ballard, Robin, McCabe, & McDonald, J., 2010; Ballard, Djaja, Arciuli, James, & Doorn, 2012; Arciuli & Ballard, 2017). In these studies, stress has been measured according to fundamental frequency (F0) or perceived pitch over the syllables in words. However, to date, there has been no investigation on the F0 tracking across sentences by this population. Future directions of these studies note a need for observing how other levels of prosody are influenced by segmentation at the word and phrase level. Additionally, studies and pilot trails on AOS were limited by not including a comparison to neurotypical, healthy participants. These data would provide opportunities for the comparison of perceptual and acoustic features to better identify the core features of CAS and to aid in differential diagnosis.

The current study will present findings on the analytical models and measurement of syllable and word segmentation in CAS and present the related history of the prosodic hierarchy and its connection to motor speech disorders. These components are critical in understanding how children with CAS use prosody. The primary research goal is to investigate if any differences in the prosodic components of segmentation and intonation are emerging as part of a working memory buffer impairment and motor programming deficit or if the breakdown is happening after the motor programming, where resource allocation is not the issue. The primary research question is if children with AOS are planning prosody (duration and frequency) into their intonational unit or if they are realizing these units as multiple parts.

1.2. Segmentation

Seddoh et al. (1996) was the first to examine intersegment durations within words (inappropriate stoppages between syllables in words, where syllables are defined as the segment) and within words (inappropriate stoppages between words in an utterance, where words are defined as the segment) in adults with AOS. This study used a multisyllabic word production task with adults with AOS to examine the differences in their speech. Data found that participants with AOS differed from people with conduction aphasia (who made predominantly phonological errors) and healthy controls. People with AOS were found to have transition difficulty moving from one sound to the next, causing segmentation, whereas the those with aphasia, phonological errors, or no disorder did not. Results attributed a longer segment duration to be due to the integration of kinesthetic information and spaciotemporal programming. Investigations into why these lengthened intervals occurred in sentences and words has driven further research on segmentation.

Models of AOS describe two motor programming processes in play during the articulation of sentences (Wright et al., 2009). Wright et al. (2009) explained that segmentation is the result of a working memory buffer, which cannot concatenate or link together speech words or syllables. Syllable structures are then influenced by a “loading” delay for a syllable, and the number of unique syllables influences the number of times the loading process must occur. The two processes include the “INT”, responsible for resolving the demand of the programmed unit or message, and the “SEQ”, responsible for overseeing the serial order of longer sequences (Maas et al., 2008; Wright et al., 2009).

INT has been found to be able to preprogram units in advance to speech by preparing these units based on their length of time. The SEG cannot be preprogrammed because of its involvement in the sequencing of each programmed unit into the correct order for production. Wright et al. (2009) found that the INT process was placed in a higher demand when attempting to produce a multisyllabic word, similar to the higher number of errors found by Seddoh et al. (1996). This process explains why utterances that differ in number of syllables and syllable complexity result in a longer response time by the SEG to produce an articulated utterance. Breakdowns in these processes of motor programming, seen in AOS, influence the segmentation of words and syllables because of a latency response time (Mass et al., 2008).

Given the findings of segmentation, researchers are trying to find different ways to classify and measure these inappropriate stoppages in speech. For example, Shriberg et al. (2017a) introduced a pause marker (PM) as an acoustic and perceptual quantifying measure to function as a discriminatory tool to diagnose CAS from speech sound disorder (SSD). A PM is classified as an inappropriate intrasegment pause which was measured to be least 150 ms. The longer the pausing of speech, the more unnatural or segmented it will sound, changing its overall prosody.

Murray et al. (2015) further concluded the set of perceptual features that differentiate CAS from other speech disorders, which include syllable segmentation, lexical stress, and percent phonemes correct. Through the production of multisyllabic words in a connected speech sample and an oral motor exam, a set of acoustical measures has found to be 91% reliable against other diagnostic checklists for CAS and AOS. This study suggests further research is needed as there is still no defined diagnostic testing protocol clinically available with a valid set of measures to differentiate CAS from other speech and/or language disorders.

Past studies examining segmentation and lexical stress patterns of children have been conducted to investigate how these different speech patterns develop and emerge as children learn language. Children will begin to integrate more prosodic information to identify the boundaries of words and syllables in fluent speech. Arciuli and Ballard (2017) found that children discriminate strong from weak syllables during infancy, and then at two to three years of age typically-developing children are using strong-weak lexical stress patterns in their word productions. At this developmental stage, children are contrasting between both strong and weak syllables, but they are not yet adult-like in their prosodic productions until the age of 8- to 11-years-old (Arciuli & Ballard, 2017). Investigating these acoustical features in typically developing children will help understand typical speech development and guide goals for children who require intervention. This study will investigate similar findings in typically developing speech as they relate to segmentation and intonation levels of prosody and motor speech rather than lexical stress.

1.3. Intonation

Consideration for further investigation in syllable segmentation is whether words and syllables are uttered on a syllable-by-syllable basis, with lengthy intervals between syllables for the preparation of the next portion of the utterance (Ladd, 1996; Wright et al., 2009). This raises the question about how lengthened intervals between words or between syllables in words impacts the prosodic features of speech. Prosody is the tune and rhythm of speech that applies to multiple levels above the individual segments of speech. Selkirk (1980) introduced the prosodic hierarchy as a system of phonological representations that organize the stress and accent patterns related to the phonological system. This system is organized into a hierarchy of levels, which partially include lexical stress, word accents (or the accentuation level), and the intonational level. All levels of the prosodic hierarchy can be used by speakers simultaneously when producing an utterance.

At the word level in English, words contain lexical stress. This level has been investigated and marked as a diagnostic difference between AOS and people who are neurotypical. An example of lexical stress is the alternation between strong-weak and weak-strong stress patterns that can discriminate between certain noun or verb pairs in English (récord=noun vs. recórd=verb) (Ladd, 1996). Accentuation is how a unit (e.g., word, phrase) is accented to change the focus of the sentence. Typically, this unit exhibits changes in F0, intensity, duration, and segmental quality (Shattuck-Hufnagel & Turk, 1996). Last, intonation affects how these acoustic measures change over the entire utterance, a focus of the current study. As the intonation of a declarative utterance falls naturally, the F0 will fall with each word. Ladd (1996) describes the relationship of intonation and declination as an overlapping system that includes both the resetting and pausing of F0.

Declination is a phenomenon experienced due to the release of tension and contraction of the vocal folds as subglottal pressure is released out of the larynx and muscles relax (Ladd, 1996). One study found that declination is automatic and under a typical speaker's control because of the ability to manipulate the tension and contraction of the vocal folds by engaging their laryngeal muscles (Ladd, 1996). Connell and Ladd (1990) collected F0 values of high tone syllables to measure the decline down to the final syllable of a word. They analyzed F0 values at the end of the first syllable, antepenultimate, penultimate, and final syllable of four-syllable words to measure the downward trends. The natural control of escaping air from the larynx makes an observable downward fall in F0 in typical speakers. This study only used four syllable words, and further research is needed to observe this effect over longer multi-word utterances.

Disruption to the natural fall of F0 declination has been investigated in non-fluent Broca's aphasia, a language disorder known to cooccur with AOS (Seddoh et al., 1996; Marotta, Massimiliano, and Paolo, 2008). Marotta et al. (2008) used longer utterances made up of declarative, WH- and yes/no questions, and found that speech production was disrupted because of longer pausing within the utterances and the abnormal lengthening of syllables. Intonation across these utterances also had frequent F0 resetting from the unnatural, lengthened pauses in the natural F0 declination.

Declination is the steepest in declaratives and often absent in yes/no interrogatives as instead they exhibit a rise in F0 at the end of the utterance (Gussenhoven, 2004). The height of the F0 fall is often determinant on the syntactical markers of the sentence, such as a statement versus a question. Declination of F0 within a sentence is believed to be a process that takes no pre-planning and can be changed online (Ladd, 1996). Accepting this theory, declination of F0 should be similar for both children with CAS and typically-developing children. This rational contradicts past findings that CAS deficits are due to impairments in the programming of motor movements and not the pre-planning of speech motor movements (Maas et al., 2008; Wright et al., 2009). Some research, however contradicts this theory, showing that speakers often anticipate the length of the utterance and will begin with a higher F0 for longer sentences (Ladd, 1996). The increased complexity of multisyllabic words and sentences that place demands on the motor system can cause breakdowns in motor programming, which is a central goal of the current study in order to provide insight into the relationship between segmentation and natural pitch declination (Maas et al., 2008).

Initial high F0 reset is a declination trend where F0 resets to a higher pitch due to an interruption in the downward fall of F0. In typical speech, F0 reset is used at the beginning of a new utterance, this is because an interruption is expected due to the nature of pausing before beginning a new utterance (Gussenhoven, 2004). Natural pausing allows the speaker to reset their F0 to have enough range to fall in the following sentence. Overall, declination and reset have no other communicative function than the ability to act as a natural indicator for the end of an utterance. In disordered speech however, pausing is also expected, but these individuals additionally present with unnatural pausing, impacting the clarity of starting and ending an utterance. The present study predicts that children with CAS may show less declination for longer sentences because of the influence of motor programming deficits, particularly in the production of complex stimuli (Duffy, 2013; Maas et al., 2008).

Other research on F0 declination measured F0 reset related to utterance lengthening for typically-developing children and children with other developmental disorders, like autism spectrum disorders (ASD). Wightman, Shattuck-Hufnagel, Ostendorf, and Price (1992) investigated the acoustical phenomena that happens at phrase boundaries during intra-segmental lengthening. Prosodic intonational phrasing was found to be restricted to the rhyme of the syllable preceding the boundary and marked by a variety of acoustic cues that involved intonation, pausing, and duration. One study used acoustic measurements of F0 reset to identify perceptual differences in speakers with high-functioning ASD from typical speakers (Peppé, Mccann, Gibbon, O'Hare, & Rutherford, 2007). The measurements showed evidence of F0 reset, suggesting an overall 'exaggerated prosody', with a wider F0 range used than their typically developing peers.

Dankoviov, Pigott, Wells, and Peppé (2004) also measured changes in prosody with the manipulation of phrase boundaries in typical speech development of young children. This study used minimal pairs of utterances, one with a compound noun, in which there is no prosodic boundary (e.g., ‘coffee-cake and tea’) and the second was a list of simple nouns separated by a prosodic boundary (e.g. ‘coffee, cake, and tea’). The present study will be investigating the connection between inter-segment pause duration between and within words and how it impacts the prosodic hierarchy of speech at the intonational phrase level. If these are both still developing prosodic characteristics, this study asks how they compare to the developing language of children with a motor speech disorder like CAS.

1.4. Summary

This introduction has presented current characteristics and findings related to the perceptual and acoustical measures of CAS, specifically the prosodic elements of segmentation and intonation. Studies have presented findings of the underlying processes involved in creating segments in multisyllabic word production and offered attempts to measure the duration of these inter-segments. The TEMPO pilot trial has introduced multisyllabic word stimuli that provide an opportunity for a comparison of segmentation for both children with CAS and typically developing children. Using these words in structured sentences provides an opportunity to observe how prosody is impacted for inter-segment duration between and within words. This introduction also presented past research on the intonational prosodic features of declination and F0 reset in typical speech. It also presented how F0 can change in disordered speech due to lengthened and segmented speech produced by individuals with Broca’s aphasia and ASD. Finally, the introduction set up questions for how segmentation can have possible impacts on F0 changes over an entire utterance. Understanding the connections between these two characteristics can lead to further understanding about online motor programming and pre-planning of speech movements.

Comparing speech production patterns in English-speaking children with CAS with typically developing (TD) children will produce a TD comparison to the CAS group from the TEMPO study prior to treatment. Acoustical measures on segmentation of inter-segment durations between and within words are predicted to demonstrate the key features of the disorder, with an analysis of these measures in age-matched typical controls necessary to better measure treatment effects at baseline. The sentence stimuli from the TEMPO study provides data on both segmentation and prosody for the two groups. The current study will provide data as to what TD children do when prompted to repeat sentences with real and non-words in utterances and allow a comparison of typical child speech production to the speech of children with CAS using the same set of stimuli.

1.5. Specific Aims

There are two specific aims for this study. The goal of aim 1 is to compare segmentation of inter-segment durations between word and within word between children with CAS and TD children. Previous work has confirmed findings of increased inter-segment duration between syllables of multisyllabic words produced by children with CAS. Using the stimuli from the TEMPO pilot treatment study, productions of multisyllabic words used in sentences are analyzed, offering a direct comparison of segmentation in these two populations. It is hypothesized that segmentation will be a strong differentiating variable between groups. Significant differences are expected in inter-segmentation duration both between and within words between children with CAS and TD children. The primary hypothesis for specific aim 1 (Hypothesis 1) is that children with CAS present with significantly longer inter-segment durations between and within words compared to TD children.

The goal of aim 2 is to examine the interactions between segmentation and change in F0 for children with CAS and TD children. Previous studies have observed F0 declination with different populations of various disorders, but this has not yet been investigated for children with CAS. It is hypothesized that there would be less of a degree of F0 declination for the children with CAS than for the TD children, particularly in the case for children who exhibit increased inter-segmentation durations. Additionally, it is also hypothesized that there may be F0 reset for children with CAS, due to the lengthening of inter-segment durations between and within words. The primary hypothesis for specific aim 2 (Hypothesis 2) is that longer inter-segment durations, expressed by children with CAS will influence F0 declination more than for TD children.

2. METHODS

2.1. Participants

Participants were 10 typically developing (TD) children ages 5;10-11;00 (*M* = 8;11) and 11 children with a diagnosis of CAS ages 5;10-8;00 (*M* = 7;06). All children were native speakers of English and had no (other than CAS) developmental, neurological, genetic or speech disorders. Both CAS and TD groups passed a hearing screening of tympanometry, otoscopy, and pure tones. All participants were screened using an Oral and Motor Speech Examination testing their oral symmetry, musculature, strength, and vocal qualities (Duffy, 2005). Last, the *Clinical Evaluation of Language Fundamentals-Fifth Edition* (CELF-5) was administered to calculate a Core Language Score (CLS) for each participant demonstrating their general language abilities (Wiig, Semel, & Secord, 2013). TD participants scored an average CLS of 115 and the CAS participants scored an average CLS of 86. Table 1 below includes participant information by participant.

Table 1. Participant information.

Subject	Age	Sex	Group	Hearing Screening	Oral and Motor Speech Exam	CELF – 5 Core Language Score
CAS01	06;05	F	CAS	Pass	Inadequate	102
CAS02	06;10	M	CAS	Pass	Inadequate	70
CAS03	05;90	M	CAS	Pass	Inadequate	96
CAS04	07;07	M	CAS	Pass	Inadequate	66
CAS05	08;03	F	CAS	Pass	Inadequate	87
CAS06	08;00	M	CAS	Pass	Inadequate	81
CAS07	08;00	M	CAS	Pass	Inadequate	80
CAS08	06;11	M	CAS	Pass	Inadequate	87
CAS09	06;04	M	CAS	Pass	Inadequate	102
CAS10	07;04	M	CAS	Pass	Inadequate	86
CAS11	07;00	M	CAS	Pass	Inadequate	93
TD01	09;01	F	TD	Pass	Adequate	110
TD02	07;07	M	TD	Pass	Adequate	123
TD03	10;01	F	TD	Pass	Adequate	111
TD04	07;04	F	TD	Pass	Adequate	104
TD05	11;00	M	TD	Pass	Adequate	113
TD06	08;07	M	TD	Pass	Adequate	111
TD07	09;11	M	TD	Pass	Adequate	107
TD08	09;11	M	TD	Pass	Adequate	117
TD09	09;11	M	TD	Pass	Adequate	127
TD10	05;10	F	TD	Pass	Adequate	134

Note: Inadequate = resulted in a diagnosis of CAS for the TEMPO study.

TD participants are recruited via formal recruitment letters and recruitment flyers. Letters were emailed to and/or posted at schools, day cares, libraries, community centers, summer camps, and children’s museums. Flyers were posted in the University of New Hampshire Speech-Language-Hearing Center and around campus, as well as any other above locations to reach families. All TD participants fit the above criteria by reporting no history of any speech or language disorder and having a normal Motor Speech and Oral Mechanism Examination.

CAS participants were diagnosed using the Motor Speech Examination (Duffy, 2005) based on three features of CAS: distortions, segmentation, and unequal syllable stress.

Participants were recruited from the Treatment to Establish Motor Program Organization (TEMPO) clinical pilot treatment program (Miller et al., 2018).

2.2. Research design

This study is a between-subjects comparison between two groups for the three dependent variables of inter-segment durations between words (ms), within words (ms), and fundamental frequency (F0) in Hz. The first between-subjects group is made up of the TD participants, who are referred to as the TEMPO control group (or TD##). The second group is made up of the CAS participants, who are referred to as the CAS group (or CAS##). The CAS participants are a combination of individuals who participated in TEMPO and who either received immediate treatment or received waitlisted treatment (Miller et al., 2018). Baselines for CAS participants were all collected at the same time before intervention or outside therapy was introduced.

2.3. Procedure

2.3.1. Speech stimuli

Stimuli for the TD participants included 120 sentences made up of 10 different carrier phrases and 120 different stimuli words (shown in Tables 2 and 3). Words are a combination of real and non-words made up of three or four syllables in length. The consonants in the non-words are either all stop plosives or fricatives (e.g. BIgatu, giBUta, TAgibu, fuSIsha). These words were generated from the first baseline word list used in the TEMPO study (Miller et al., 2018). Both groups of participants completed this word list within one two-hour long session. Tables 2 and 3 list examples of the stimuli used as well as the carrier phrases that were subsequently paired with target words.

Table 2. Stimuli word types

Stimuli type	Number of items	Number of syllables	Examples
Non-word Stop plosives	30	3	BIgatu, giBUta, TAgibu, tiGUba
Non-word Stop plosives	30	4	BIgutaba, GAbatigu, tiGAtabu, gaTUbagi
Non-word Fricative	30	3	SHUziva, siTHAvu, fiTHUza, FAthisu
Real word	30	3	Cucumber, spaghetti, octopus, hamburger
Total	120		

Table 3. Carrier phrases

Phrase	Number of items	Number of words	Opportunities for segmentation	Sentence type
I saw a	12	3	3	Declarative
Can you find my	12	4	4	Interrogative
He bought a	12	3	3	Declarative
Here's the	12	2	2	Declarative
She has a	12	3	3	Declarative
There's my	12	2	2	Declarative
It's going to	12	3	3	Declarative
It's a red	12	3	3	Declarative
I want a	12	3	3	Declarative
I went to the	12	4	4	Declarative

2.3.2. Measurements

The independent variable in this study is the between-subjects grouping variable of TD and CAS. The dependent variables are the acoustic measurements of inter-segment duration and change in F0 between words and syllable to generate a linear regression slope over the utterance and target word. Inter-segment duration between and within words was measured by calculating the overall average duration in milliseconds (ms) used by the child.

Specifically, *between word duration* (addressing Hypothesis 1) was calculated for each inter-segment between words 1, 2, 3, 4, and 5 for TD and CAS participants. The average duration was then calculated for each participant over all their utterances. Opportunities for segmentation in the utterance ranged from two to four, depending on the number of words in the utterance. *Within word duration* (addressing Hypothesis 1) was calculated for each inter-segment duration within words between syllables 1, 2, and 3 for TD and CAS participants. The average duration was then calculated for each participant over all their target words. Opportunities for segmentation at the syllable level ranged from two to three, unless additional syllables were added to the target word production by the child. Please note that two out of the eleven CAS participants (CAS03 and CAS05) only produced target words of three syllables. The remaining nine participants produced target words of four syllables.

Change in F0 was calculated by extracting the average F0 over each word and syllable in the utterance and target word. Utterance level F0, or *F0 Intonation,* (addressing Hypothesis 2) was calculated by taking the average F0 for each word boundary using a pitch range from 50 HZ to 500 Hz. These data were plotted in order to observe the degree of F0 declination between each word in the utterance, and a slope of linear regression was calculated over the entire utterance. This calculation only included declarative sentences produced by the participant. A second calculation found the change in F0 from word to word by calculating the change in F0 from one word to the next in order to observe any evidence of F0 reset. *F0 over target word* (addressing Hypothesis 2) was calculated by taking the average F0 over each syllable. These were then plotted in order to observe the degree of F0 declination from syllable to syllable and a slope of linear regression was calculated over the entire target word. This calculation only included target words with initial syllable stress. Finally, the change in F0 from one syllable to the next was calculated to observe evidence of pitch reset within the target word.

2.3.3. Acoustical analysis

Initial acoustical analyses were run using Praat to transcribe and mark each of the dependent variables of inter-segment duration between and within words and F0 (Boersma &

Weenink, 2017). Acoustical coding was completed by trained graduate researchers with inter-rater reliability checked. After discussing any discrepencies, agreeability and reached 100%. Each inter-segment duration between words was labeled with a “W1” for the duration between words one and two, a “W2” between words two and three, a “W3” for the space between words three and four, and so on (W# indicating the space between words). Additionally, for opportunities between words where there was no segmentation noted, a “W#0” was added to indicate a 0 ms duration with no evidence of segmentation. Inter-segment duration within words was labeled with a “S1” for duration between syllables one and two, a “S2” for space between syllables two and three, and for 4-syllable stimuli a “S3” for between syllables three and four (S# indicating the duration between syllables). Like between words, any instances where the participant did not use segmentation and produced the word as a coarticulated unit, a “S#0” was used to indicate no evidence of segmentation between syllables. See Figure 1 for all acoustical coding features identified in Praat.

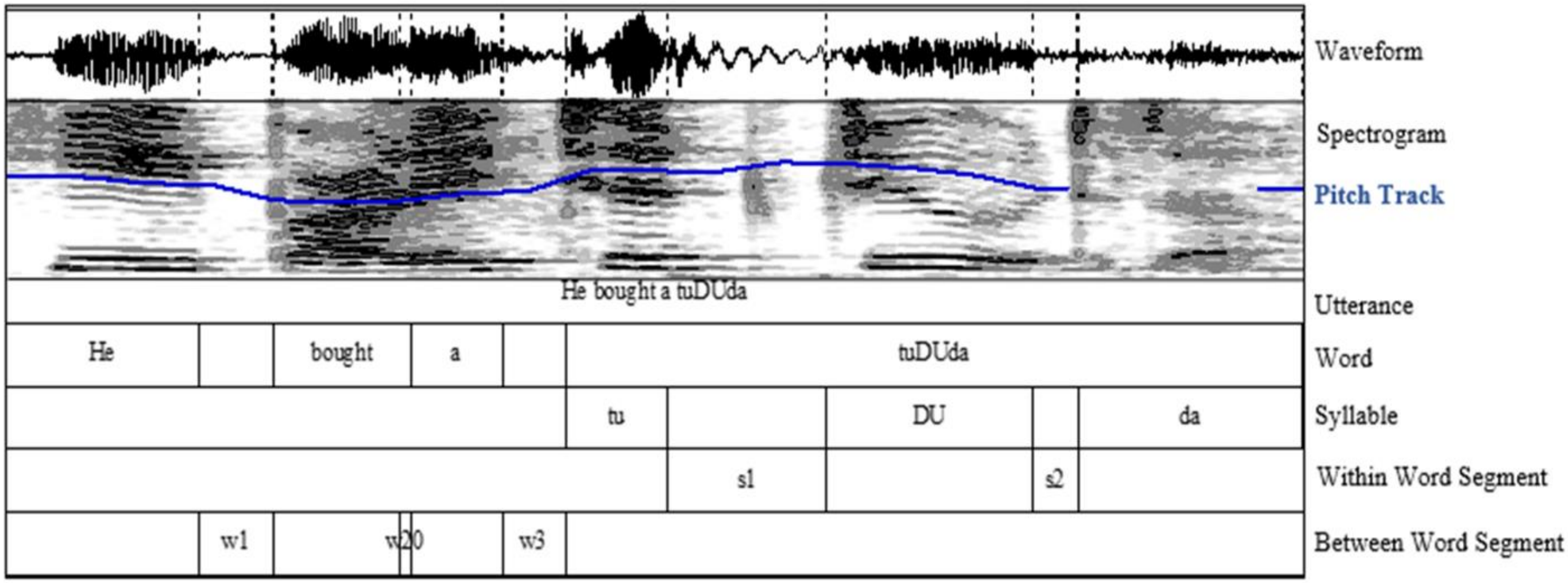

Figure 1. Praat coding example. Tier names indicated on the right reflect the type of data coded over the utterance. Interval barriers mark the start and end of each unit (utterance, word, syllable, or inter-segment). The blue line over the spectrogram indicates the F0/pitch track between the range of 50 and 500 Hz.

Each dependent variable was coded acoustically by marking the start and end of the temporal boundaries of words in the utterance and syllables in the final multisyllabic word. The start and end of segments was defined as the time from the last glottal pulse (using the spectrogram to find the end of energy extending through F1 and F2) to the onset of plosive burst in the following syllable. Average F0 was collected over each word and syllable.

Each between and within word inter-segment duration was extracted into an Excel document with their corresponding labels. The Excel spreadsheet contained the inter-segment location between words (W#) or within words (S#) and its duration in ms for the inter-segment duration measurements. Next, word and syllable boundaries were extracted with the corresponding labels of the word or syllable produced along with the average F0 in Hz for the calculation of F0 declination measurements.

2.3.4. Data analysis

For the collected speech samples, data were excluded if there were any effects that resulted in an unreliable or absent F0 track. There was an average of 13 (maximum of 99 and a minimum of 2 per participant) utterances and 32 (maximum of 42 and a minimum of 2 per participant) words or syllables that had unreliable or absent F0 tracks, resulting in the exclusion of these productions. There was a total of 211 words and 465 syllables excluded due to pitch tracking errors in Praat. There was a total of 34 utterances excluded due to poor recording quality, including the child whispering or the recording being too quiet to be picked up by the microphone, incomplete audio files, or the influence of background noise. There was a total of 54 utterances excluded from participants due to child influences, including responding with a question instead of a statement, repeating words, restarting mid-sentence, only stating one syllable of the target word, yawning, coughing, laughing, yelling, or using a silly voice. There was a total of 183 utterances excluded from participants that included the “Can you find a…” carrier phrase, due to the influence of this creating an interrogative utterance versus a declarative one. Last, there was a total of 15 utterances excluded due to the influence of a segmented utterance production modeled by the research clinician. Overall, there was a total of 374 utterances, 211 words, and 465 syllables excluded.

Following exclusions, there was a total of 4,778 inter-segment durations between words analyzed, with an average of 227 opportunities between words in the utterances among both groups of participants. There was a total of 3,572 inter-segment durations within words analyzed, with an average of 324 opportunities between syllables in words among both groups of participants. There was a total of 5,227 words analyzed for F0 measurements to calculate declination over the utterance, with an average of 248 words per participant in each group. There was a total of 3,934 syllables analyzed for F0 measurements to calculate declination over the target word, with an average of 357 syllables per participant for both groups.

The above data was organized using Microsoft Excel and prepared for statistical analysis using IBM SPSS Statistics 25 (IBM Corp., 2017). A Mann-Whitney U test was selected to compare the differences between groups for each dependent variable (Dineen & Blakesley, 1973). The independent variable in this study was whether or not the test group had CAS or not. Thus, the between-subjects groups used in the analysis are children who have CAS or not (i.e., the TD group). This statistical test follows the theory of stochastic ordering, quantifying a theory that the dependent variable may be "bigger" than another between groups. The Mann-Whitney U test also allows one to look at the two distributions of each group comparison and determine how to best interpret the results, as either a difference in distribution (if they do differ in shape) or as a difference in medians (if they have the similar distribution shapes). All data was determined to be of similar shapes and distributions and were compared for differences in medians.

3. RESULTS

3.1. Segmentation

3.1.1. Inter-segment duration between words

Inter-segment duration between words results was calculated for total average segmentation per participant and per opportunity between words within the utterance (W1 = between words one and two and so on for W2, W3, and W4). The TD group exhibited a shorter average inter-segment duration between words of 25.73 ms compared to 91.49 ms for the CAS group. Once between word duration was gathered overall, individual between word durations were calculated based on their placement in the utterance (see Table 4). Table 4 shows average, standard deviation, and standard error for TD and CAS participants for inter-segment duration results for between words.

Table 4. Inter-segment duration results between words

		Total Duration (ms)	**Avg. W1 Duration**	**Avg. W2 Duration**	**Avg. W3 Duration**	**Avg. W4 Duration**
Average (ms)	*TD*	25.73	6.56	18.37	50.38	93.38
	CAS	91.49	39.60	93.39	127.32	202.86
SD (ms)	*TD*	9.89	2.66	9.75	22.41	42.72
	CAS	48.38	22.37	75.31	60.12	101.42
Standard Error	*TD*	3.13	0.84	3.08	7.09	13.51
	CAS	14.59	6.74	22.71	18.13	30.58

A Mann-Whitney U test was run to determine if there were differences in average inter-segment duration between words between participants with and without CAS. Average between word duration across utterances was significantly shorter in the TD group (*Mdn* = 26.31) than in the CAS group (*Mdn* = 78.84), $U = 1.000$, $z = -3.803$, $p <$ inter-segment duration 0.001, using an exact sampling distribution for *U*. Measurements for inter-segmentation between words were then calculated for each segmentation opportunity between words (W1, W2, W3, and W4). Average W1 (duration between words one and two) was significantly shorter in the TD group (*Mdn* = 5.83) than in the CAS group (*Mdn* = 34.47), $U = 0.000$, z = -3.873, $p < 0.001$, respectively for W2 (*Mdn* = 15.19) than in the CAS group (*Mdn* = 73.74, $U = 2.000$, z = -3.732, $p < 0.001$), W3 (*Mdn* = 43.98) than in the CAS group (*Mdn* = 109.16), $U = 9.000$, z = -3.239, $p =$ 0.001, and for W4 (*Mdn* = 79.77) than in the CAS group (*Mdn* = 147.45), $U = 15.000$, z = -2.817, $p = 0.004$. Figure 2 shows the durations (ms) for segmentation between words.

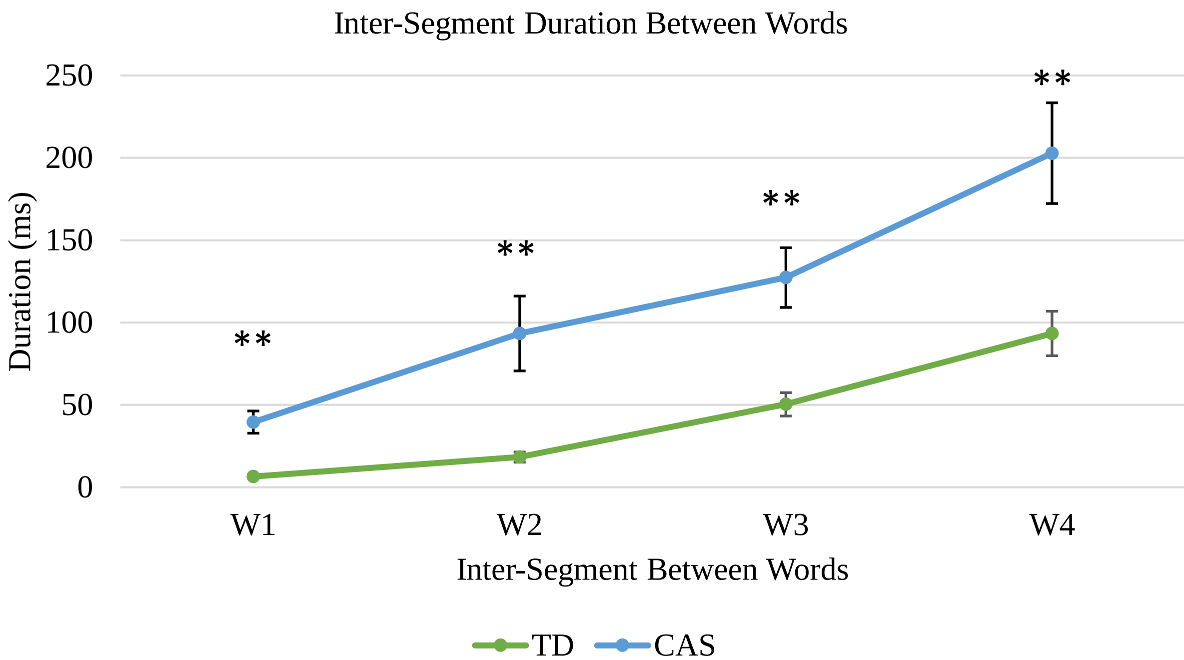

Figure 2. Inter-segment duration between words over the utterance.

Additional analysis was completed to determine a correlation between inter-segment duration between words and the number of words in the utterance. A Spearman's correlation concluded a strong positive relationship between inter-segment duration between words and number of words in the sentence ($rs(82) = 0.700$, $p < 0.001$. Overall, inter-segment duration between words was noted for both groups, with an increase in inter-segment duration between words to the end of sentence for W4 (between the carrier phrase and the final target word). Figure 3 demonstrates this strong correlation pattern.

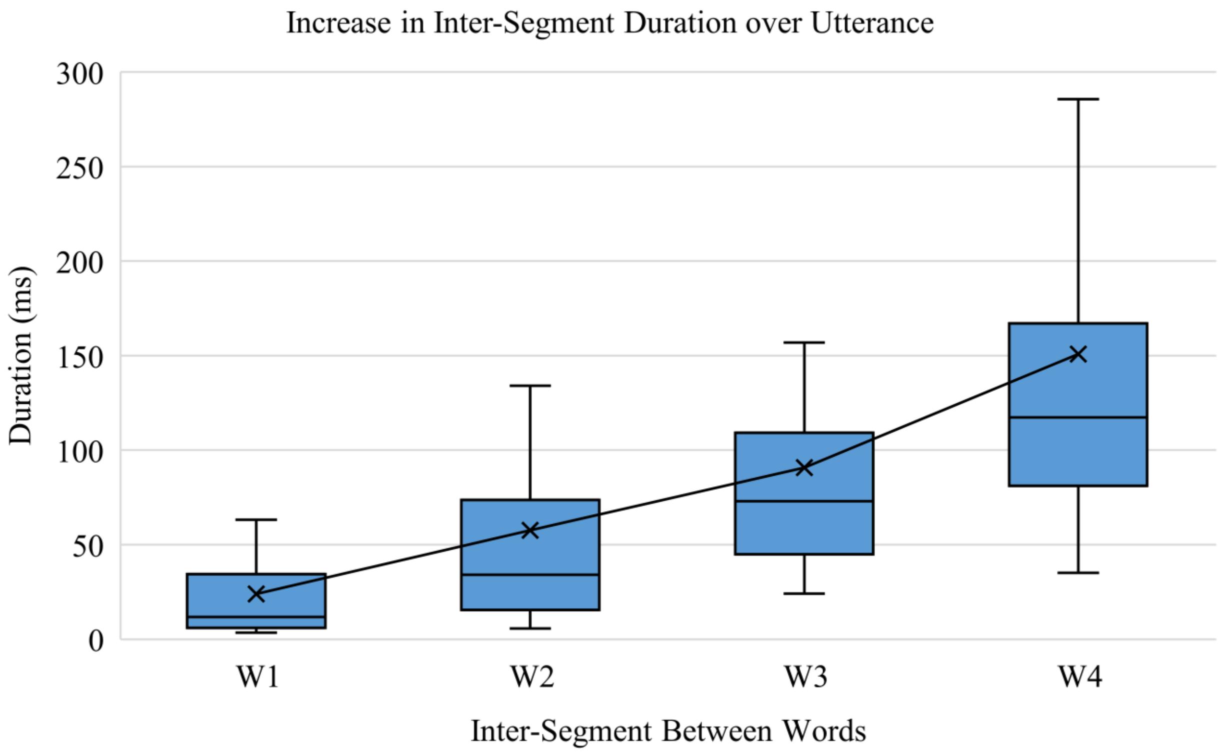

Figure 3. Correlation between inter-segment duration between words and total number of words in the utterance.

3.1.2. Inter-segment duration within words

Inter-segment duration within words results was calculated for total average segmentation per participant and per opportunity between syllables within the target word (S1 = between syllables one and two and so on for S2, and S3). Overall, the TD group exhibits a shorter average within word duration of 46.20 ms compared to 75.01 ms for the CAS group. Both groups also varied in inter-segment within word duration for each opportunity between syllables (see Table 5). Table 5 shows average, standard deviation, and standard error for TD and CAS participants. Overall, segmentation was noted for both groups.

Table 5. Inter-segment duration results within words.

		Total Duration (ms)	Avg. S1 Duration	Avg. S2 Duration	Avg. S3 Duration
Average (ms)	*TD*	46.20	42.86	37.57	57.19
	CAS	75.01	70.61	70.11	78.96
SD (ms)	*TD*	17.88	18.59	7.95	18.49
	CAS	32.21	31.42	34.28	49.49
Standard Error	*TD*	5.65	5.88	2.51	5.85
	CAS	9.71	9.47	10.34	14.92

A Mann-Whitney U test was run to determine if there were differences in average inter-segment duration within words between participants with and without CAS. Average within word duration across target words was significantly shorter in the TD group (*Mdn* = 46.08) than in the CAS group (*Mdn* = 117.26), $U = 2.000$, $z = -3.732$, $p < 0.001$, using an exact sampling distribution for *U*. Measurements for inter-segment duration within words were then calculated for each segmentation opportunity between syllables (S1, S2, and S3). Average S1 (duration between syllables one and two) was significantly shorter in the TD groups (*Mdn* = 38.10) than in the CAS group (*Mdn* = 108.29), $U = 1.000$, $z = -3.803$, $p< 0.001$, respectively for S2 (*Mdn* = 39.33) than in the CAS group (*Mdn* = 108.03), $U = 0.000$, $z = -3.873$, $p < 0.001$, and for S3 (*Mdn* = 58.63) than in the CAS group (*Mdn* = 80.85), $U = 20.000$, $z = -2.465$, $p = 0.013$. Figure 4 shows the durations (ms) for segmentation between syllables.

A comparison of TD and CAS participants is shown in Figure 5. This figure demonstrates the duration of each inter-segment duration for both groups for a comparison of duration between words and within words.

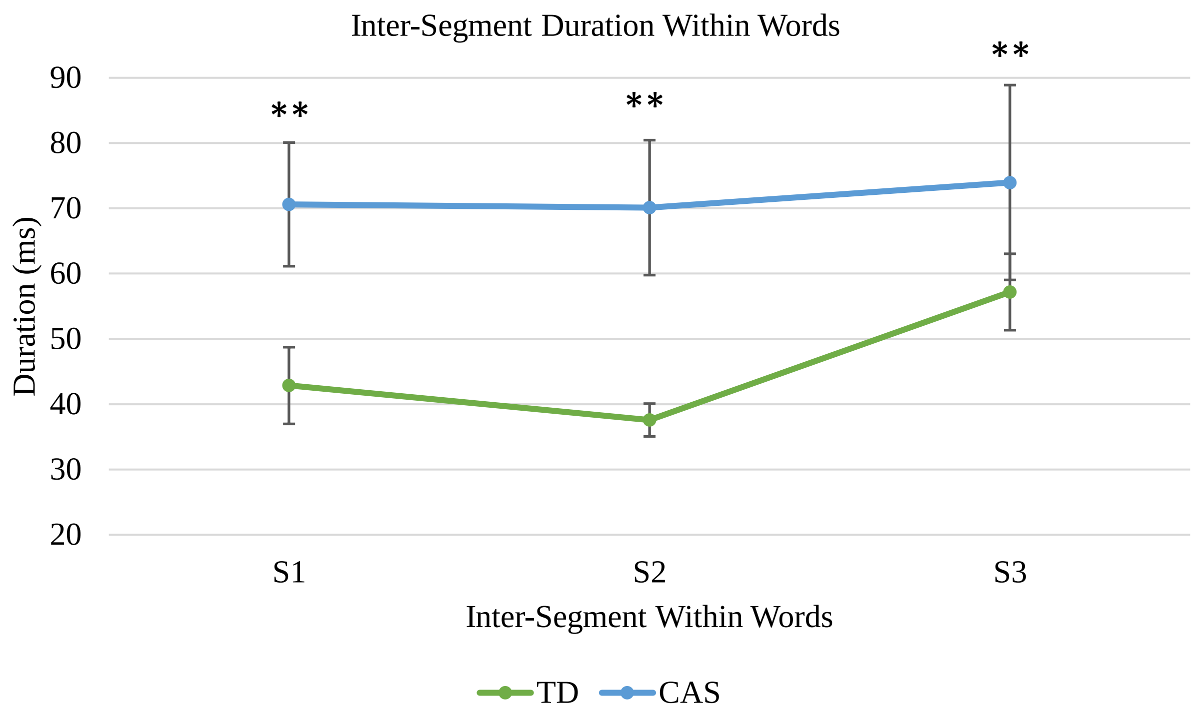

Figure 4. Inter-segment duration within words over the utterance.

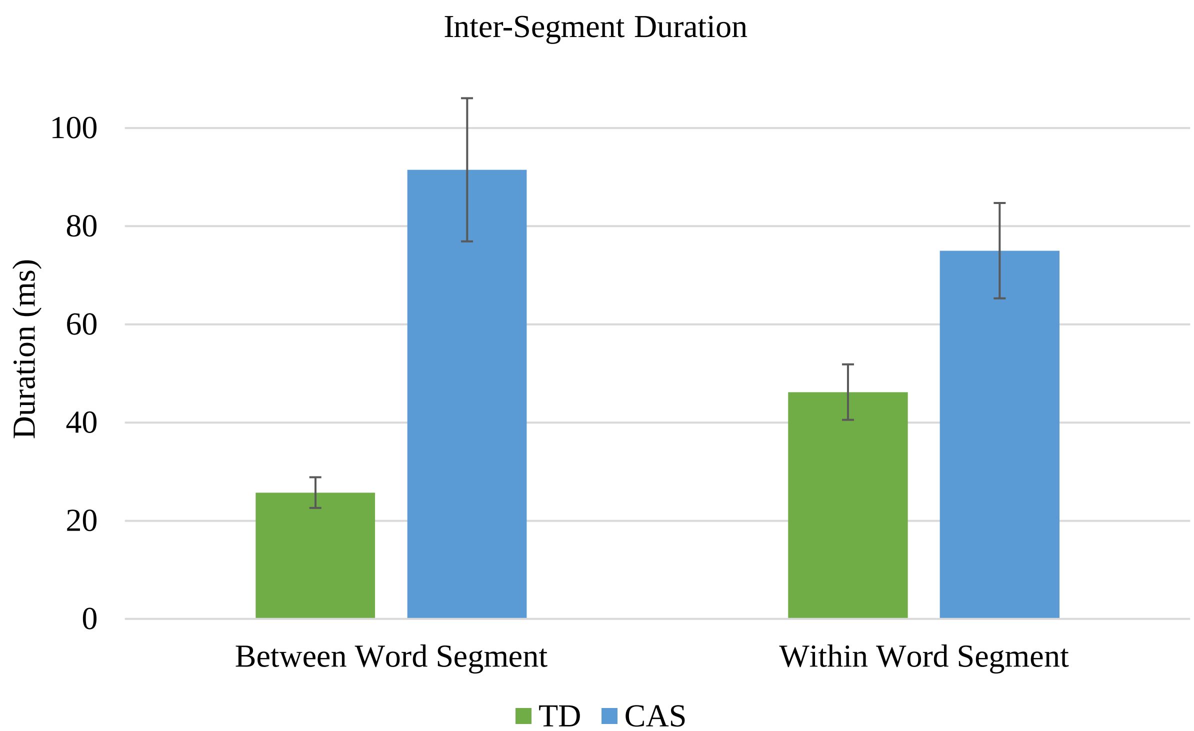

Figure 5. Comparison of inter-segment duration between and within words.

3.2. Intonation

3.2.1. F0 over utterance

Intonation was initially calculated by looking at the average change in fundamental frequency (F0) from word to word. The average F0 change from word 1 to word 2 (W1) declined by 8.87 Hz for the TD group (SD = 7.46 Hz) and 13.37 Hz for the CAS group (SD = 15.41 Hz). The average F0 change from word 2 to word 3 (W2) declined by 11.76 Hz for the TD group (SD = 6.64 Hz) and 16.27 Hz for the CAS group (SD = 16.41 Hz). The average F0 change from word 3 to word 4 (W3) declined by 8.41 Hz for the TD group (SD = 8.93 Hz) and increased by 11.50 Hz for the CAS group (SD = 13.21 Hz). Finally, the average F0 change from word 4 to word 5 (W4) declined by 1.69 Hz for the TD group (SD = 8.14 Hz) and 0.44 Hz for the CAS group (SD = 21.41 Hz).

Table 6. Change in F0 over the utterance

		W1 F0 Difference	**W2 F0 Difference**	**W3 F0 Difference**	**W4 F0 Difference**
Average	TD	-8.87	-11.76	-8.41	-1.69
(Hz)	CAS	-13.37	-16.27	+11.50	-0.44
SD	TD	7.47	6.64	8.93	8.14
(Hz)	CAS	15.42	16.41	13.21	21.42
Standard	TD	2.36	2.10	2.82	2.57
Error	CAS	4.64	4.95	3.98	6.46

A Mann-Whitney U test was run to determine if there were differences in intonation, based on average change in F0 between the words of the utterance for participants with and without CAS. Average change in F0 across the utterance was calculated for each word transition (W1, W2, W3, and W4), using an exact sampling distribution for *U*. The average F0 difference for W1 did not significantly differ between the TD group (*Mdn* = -7.19) and the CAS group (*Mdn* = -19.21), *U* = 73.000, z = 1.268, *p* = 0.223, nor for the difference for W2 (*Mdn* = -11.01) and the CAS group (*Mdn* = -16.12), *U* = 67.000, z = 0.845, *p* = 0.426. The average F0 difference for W3 was significantly different between the TD group (*Mdn* = -8.14) and the CAS group (*Mdn* = 5.31), *U* = 10.000, z = -3.169, *p* = 0.001, but did not significantly differ between the TD group (*Mdn* = -5.15) and the CAS group (*Mdn* = -5.40), *U* = 56.000, z = 0.070, *p* = 1.000 for W4. Figure 6 shows each participant's change in F0 over the utterance for participants who were typically developing and Figure 7 shows data for participants with CAS. Figure 8 shows the average change in F0 between both groups, reflecting the difference in F0 change during W3, which is the location of the transition from the penultimate word of the carrier phrase in to the target word.

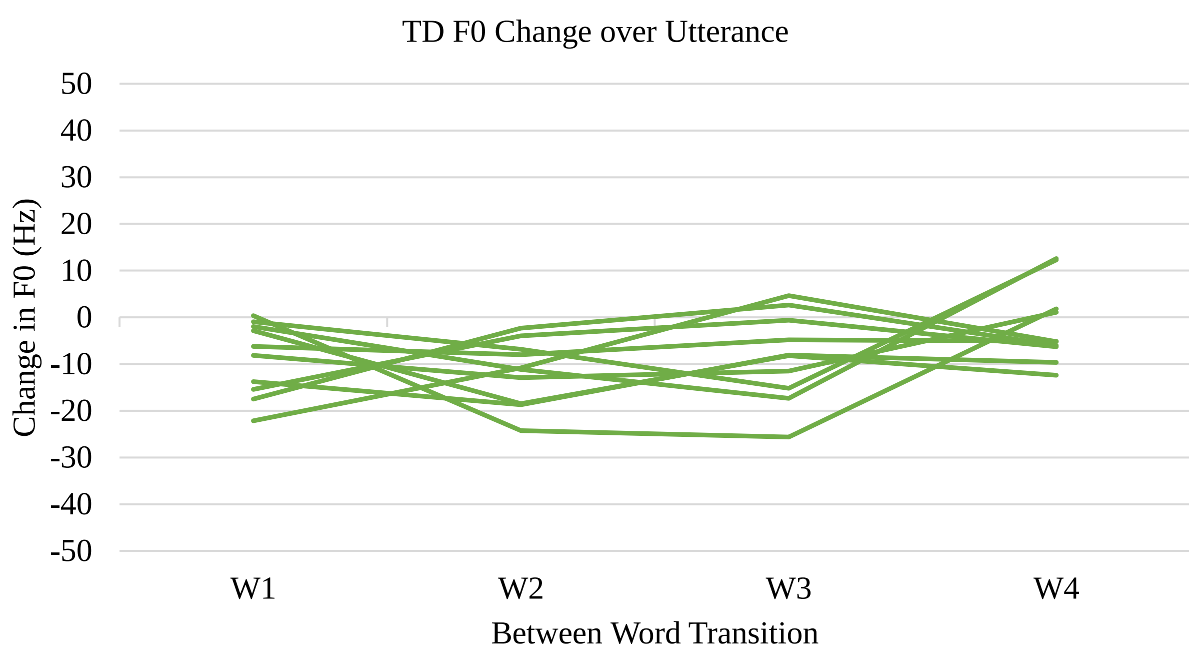

Figure 6. F0 differences between words within an utterance for TD population.

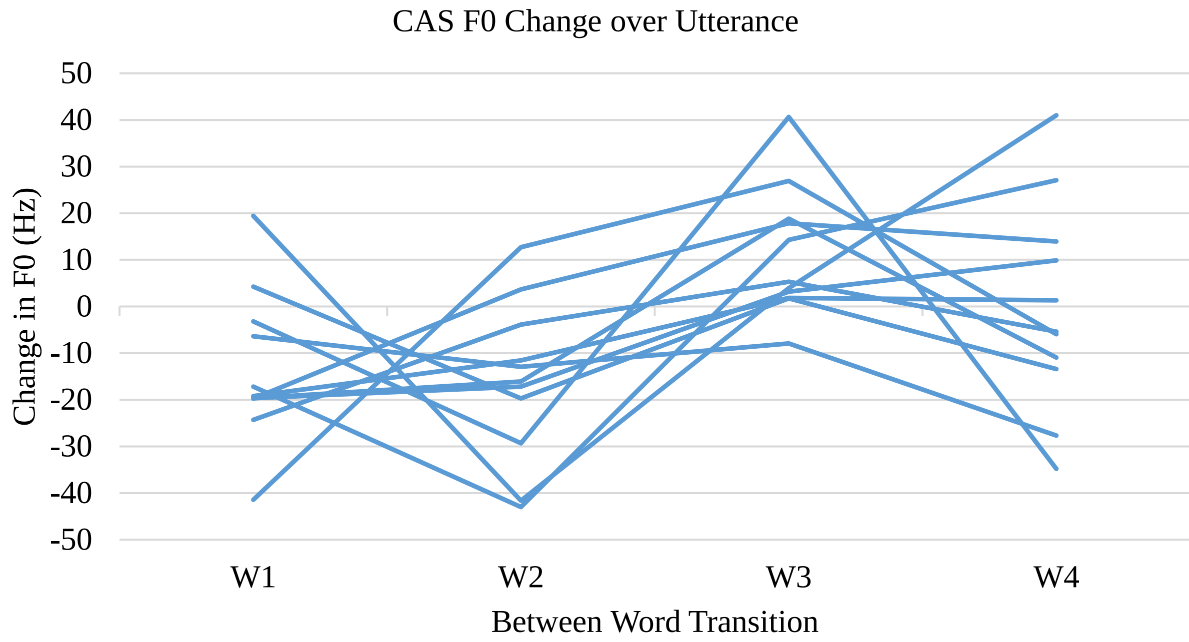

Figure 7. F0 differences between words within an utterance for CAS population.

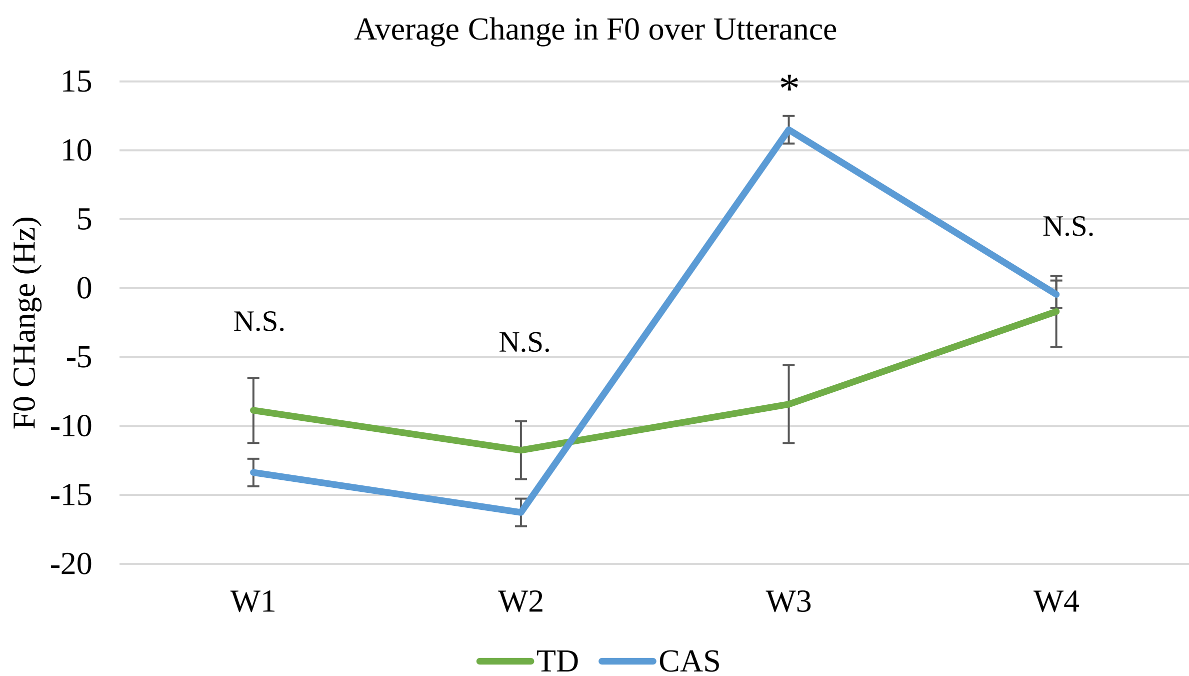

Figure 8. Average F0 differences between words within an utterance for both groups.

Additional analysis calculated the linear least squares regression line of average F0 across the utterance. The regression line was calculated using the average F0 of words one through five. The TD participants had a steeper declining average regression line with a slope of -8.16 (SD = 3.51) compared to the CAS participants who exhibited a flatter declining average regression line with a slope of -4.20 (SD = 5.14). See Table 7 for average F0 for each word, standard deviations, and standard error.

Table 7. Average F0 over word in the utterance

		Word 1	**Word 2**	**Word 3**	**Word 4**	**Word 5**
Average	*TD*	252.93	244.06	232.30	223.89	222.20
(Hz)	*CAS*	291.48	278.11	261.84	273.34	272.89
SD	*TD*	26.06	26.93	23.98	24.08	24.08
(Hz)	*CAS*	36.19	45.94	35.49	29.14	43.08
Standard	*TD*	8.24	8.52	7.58	7.80	7.61
Error	*CAS*	10.91	13.85	10.70	8.79	12.99

A Mann-Whitney U test was run to determine if there were significant differences in the slope of the linear regression lines across the utterance, based on average F0 of the words in the utterance for participants with and without CAS. Differences in average slope of the regression line was calculated using an exact sampling distribution for U. The regression lines for the utterance were not significantly different between the TD group (Mdn = -6.88) and the CAS group (Mdn = -6.17), $U = 71.000$, $z = 1.127$, $p = 0.282$.

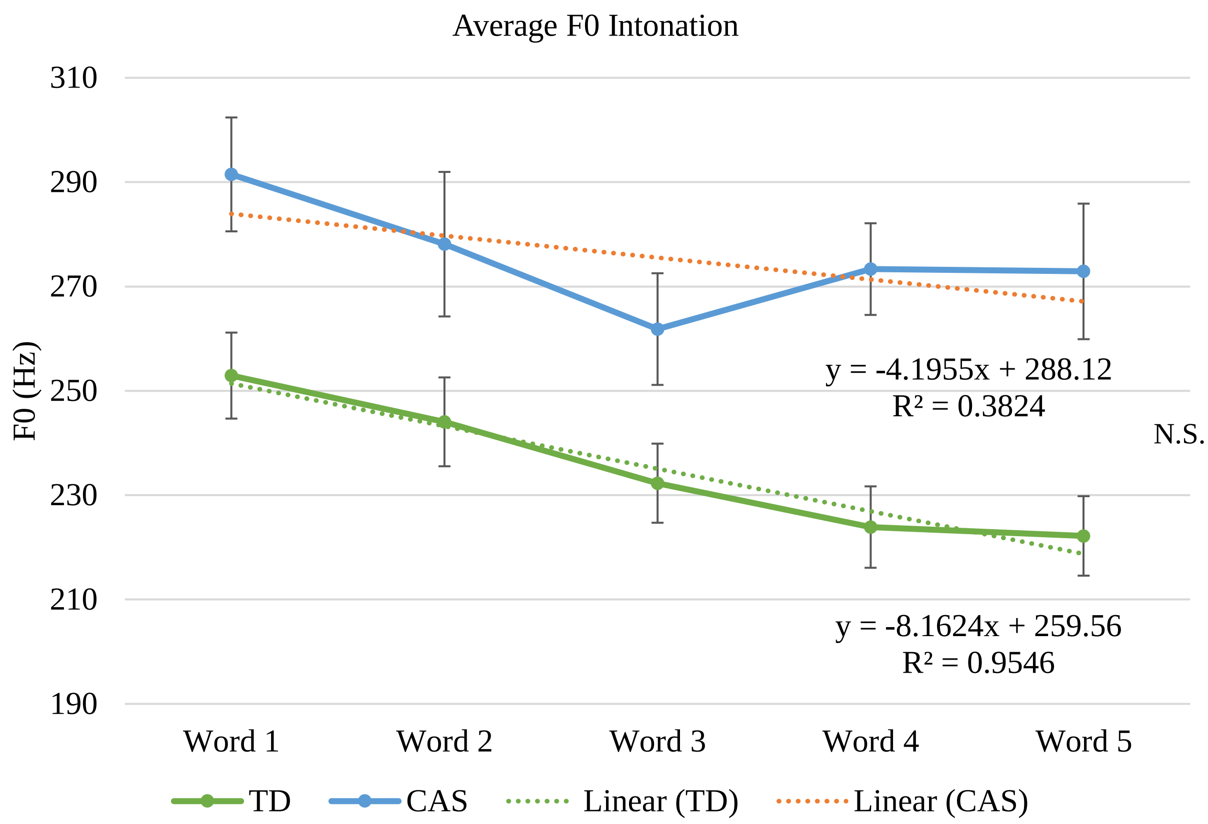

Figure 9. Average F0 intonation per word of the utterance for each group.

3. 2. 2. F0 over utterance: real- and non-words

Analyses were run to evaluate significance of utterances that use real target words versus non-real words. A regression line was calculated using the average F0 of words one through four, with the fourth word being a real word. Analysis only included four-word utterances, rather than five due to a limited number of five word utterances with real target words produced by the participants. The TD participants had a steeper declining average regression line with a slope of -9.90 (SD = 6.78) than the CAS participants who exhibited a flatter declining average regression line with a slope of -6.64 (SD = 10.86). See Table 8 for average F0 for each word, standard deviations, and standard error. The regression lines for the utterances with real words were not significantly different between the TD group (*Mdn* = -10.22) and the CAS group (*Mdn* = -9.72), *U* = 62.000, *z* = 0.483, *p* = 0.654.

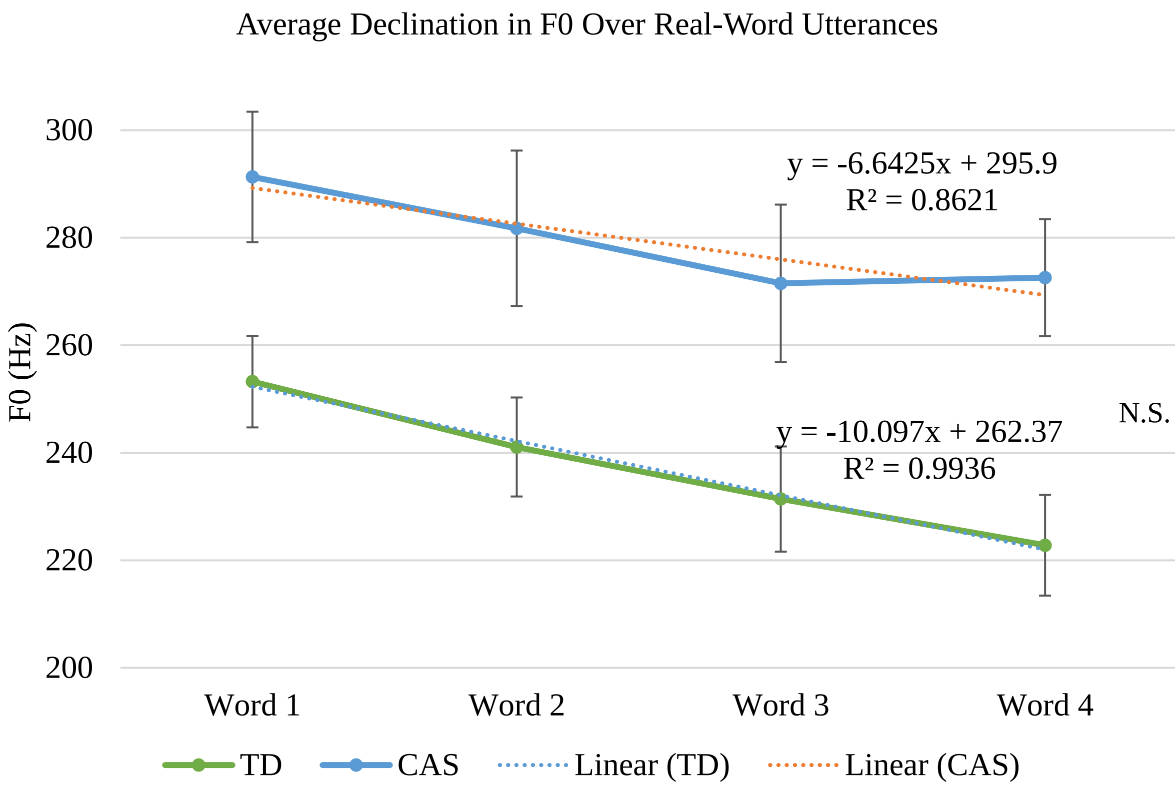

Figure 10. Average declination in F0 over real-word utterances.

Additionally, a regression line was calculated using the average F0 of words one through five, with the fifth word being a non-real word. The TD participants had a steeper declining average regression line with a slope of -8.41 (SD = 3.43) than the CAS participants who exhibited a flatter declining average regression line with a slope of -7.72 (SD = 7.83). See Table 8 for average F0 for each non-real word, standard deviation, and standard error. The regression lines for the utterances with non-words were not significantly different between the TD group (*Mdn* = -8.03) and the CAS group (*Mdn* = -7.62), *U* = 63.000, *z* = 0.563, *p* = 0.605.

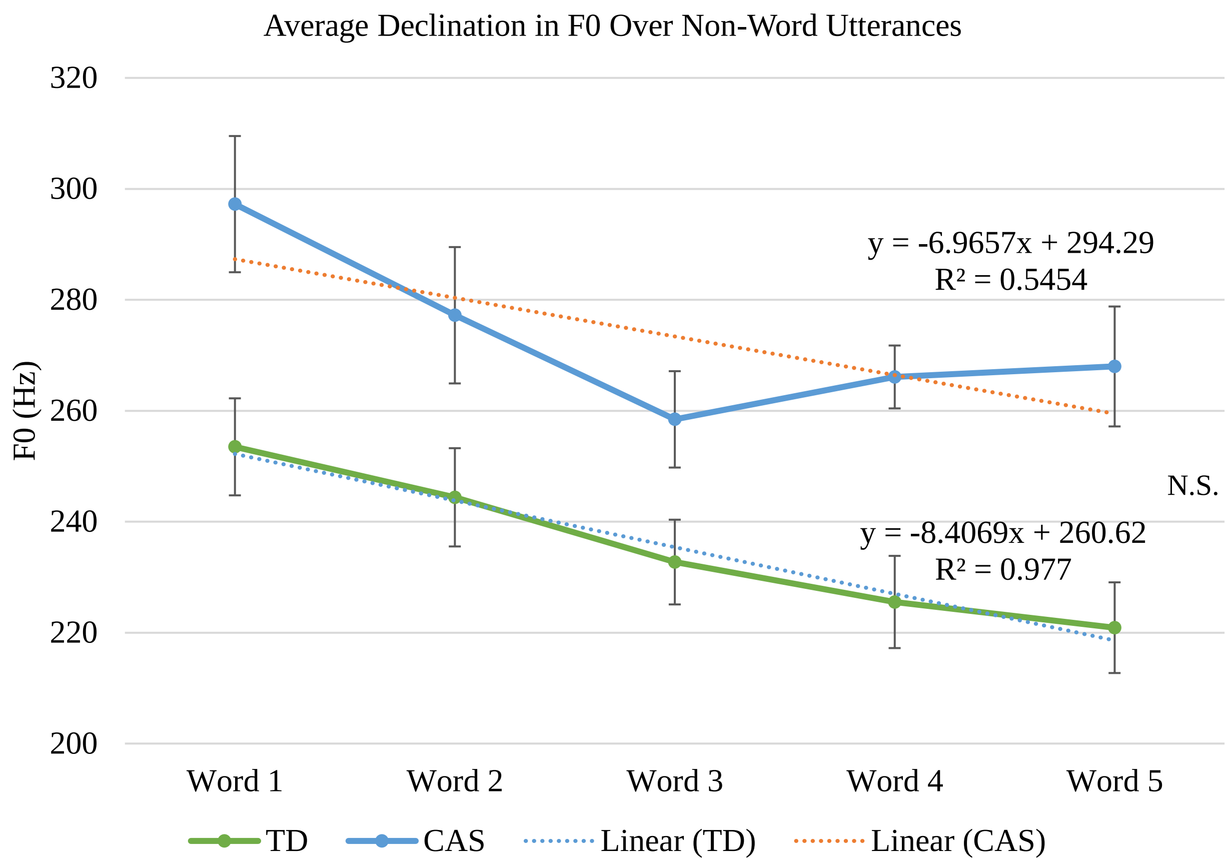

Figure 11. Average declination in F0 over non-word utterances for both groups.

Table 8. Average F0 over real and non-words in the utterance

			Word 1	**Word 2**	**Word 3**	**Word 4**	**Word 5**
Real Words	*Average (Hz)*	*TD*	253.23	241.08	231.39	222.80	NA
		CAS	291.31	281.76	271.53	272.58	NA
	SD (Hz)	*TD*	26.93	29.12	30.92	29.63	NA
		CAS	40.24	47.95	48.55	36.11	NA
	Standard Error	*TD*	8.52	9.21	9.78	9.37	NA
		CAS	12.13	14.46	14.64	10.89	NA
Non-Real Words	*Average (Hz)*	*TD*	253.49	244.39	232.72	225.52	220.89
		CAS	297.24	277.21	258.22	266.09	267.98
	SD (Hz)	*TD*	27.65	28.01	24.16	26.28	25.84
		CAS	40.72	40.75	28.81	18.80	35.85
	Standard Error	*TD*	8.74	8.86	7.64	8.31	8.17
		CAS	12.28	12.29	8.69	5.67	10.81

3.2.3. F0 over target word

F0 declination over the target word was calculated using only the multisyllabic words with initial stress on the first syllable. F0 declination was first analyzed looking at the average change in F0 from syllable to syllable. The average F0 change from syllable 1 to syllable 2 (S1) declined by 8.11 Hz for the TD group (SD = 4.67 Hz) and increased by 0.26 Hz for the CAS group (SD = 16.10 Hz). The average F0 change from syllable 2 to syllable 3 (S2) declined by 6.92 Hz for the TD group (SD = 11.26 Hz) and 15.73 Hz for the CAS group (SD = 26.40 Hz). The average F0 change from syllable 3 to syllable 4 (S3) increased to 9.63 Hz for the TD group (SD = 25.48 Hz) and 18.64 Hz for the CAS group (SD = 63.20 Hz). Table 9 shows average change in F0 between the syllables in the target word, standard deviation, and standard error.

Table 9. Change in F0 over the target word syllable to syllable

		S1 F0 Difference	S2 F0 Difference	S3 F0 Difference
Average	*TD*	-8.86	-6.92	+9.63
(Hz)	*CAS*	+0.26	-15.73	+18.64
SD	*TD*	4.66	11.26	25.48
(Hz)	*CAS*	16.10	26.40	63.20
Standard	*TD*	1.48	3.57	8.06
Error	*CAS*	4.86	7.96	19.05

A Mann-Whitney U test was run to determine if there were differences in declination over the target word, based on change in F0 between the syllables of the target word for participants with and without CAS. Average change in F0 across the target word was calculated between each syllable (S1, S2, and S3), using an exact sampling distribution for U. Average F0 change for S1 (F0 change from syllable one to 2) was not significantly different in the TD group (*Mdn* = -8.53) than in the CAS group (*Mdn* = -2.31), $U = 35.000$, $z = -1.134$, $p = 0.280$, as well as, for S2 (*Mdn* = -4.50) than in the CAS group (*Mdn* = -24.61), $U = 69.000$, $z = 1.436$, $p = 0.165$, and for S3 (*Mdn* = 2.98) than in the CAS group (*Mdn* = -11.90), $U = 45.000$, $z = 0.444$, p

= 0.696. Figure 12 shows the change in F0 over the target word from syllable to syllable for each TD participant and figure 13 shows data for participants with CAS. Figure 14 shows the average change in F0 between both groups, reflecting the difference in F0 syllable to syllable.

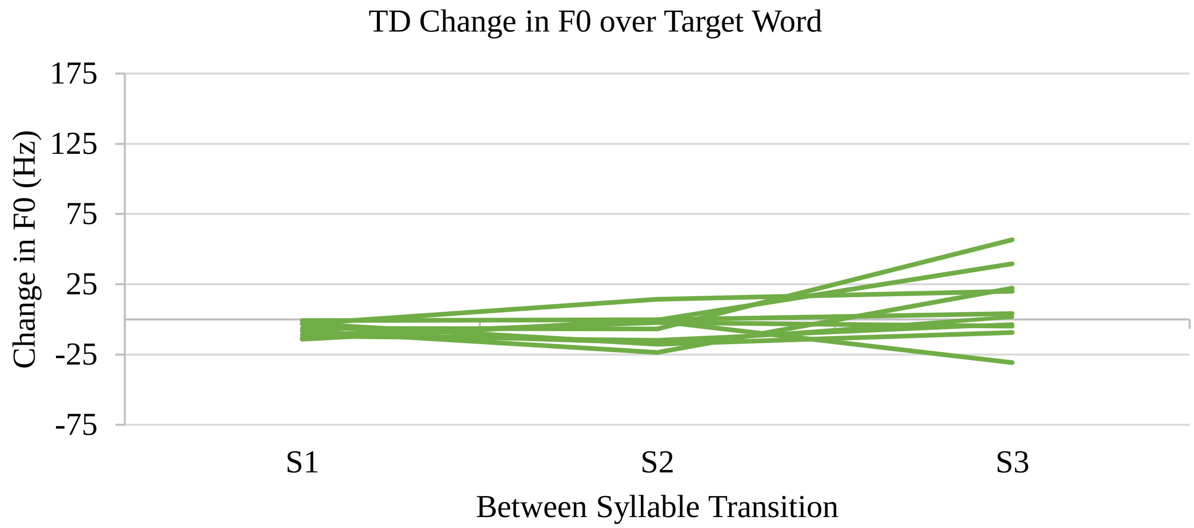

Figure 12. Participant F0 differences between syllables within a target word for TD population.

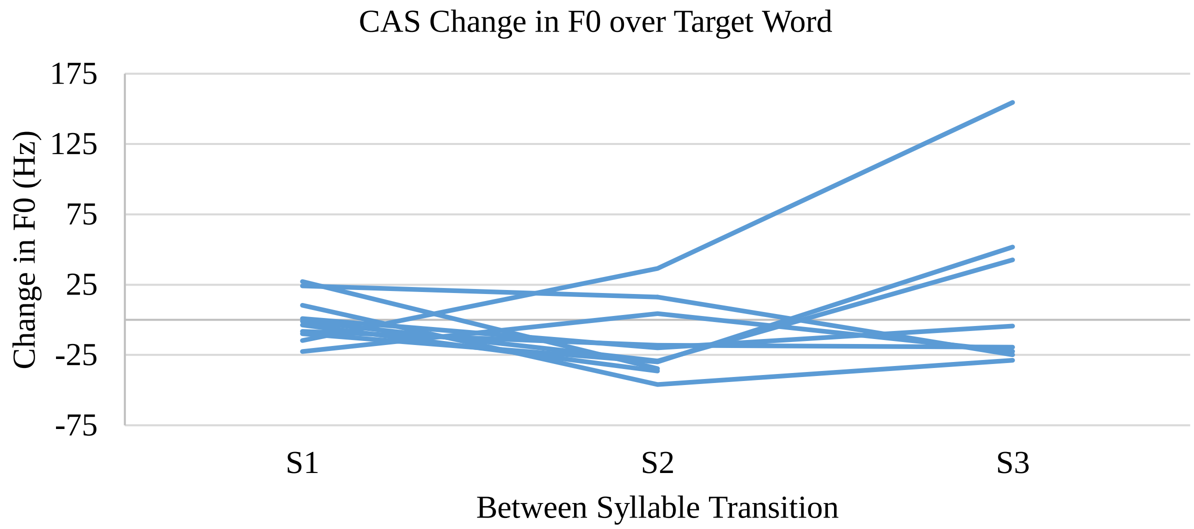

Figure 13. Participant F0 differences between syllables within a target word for CAS population.

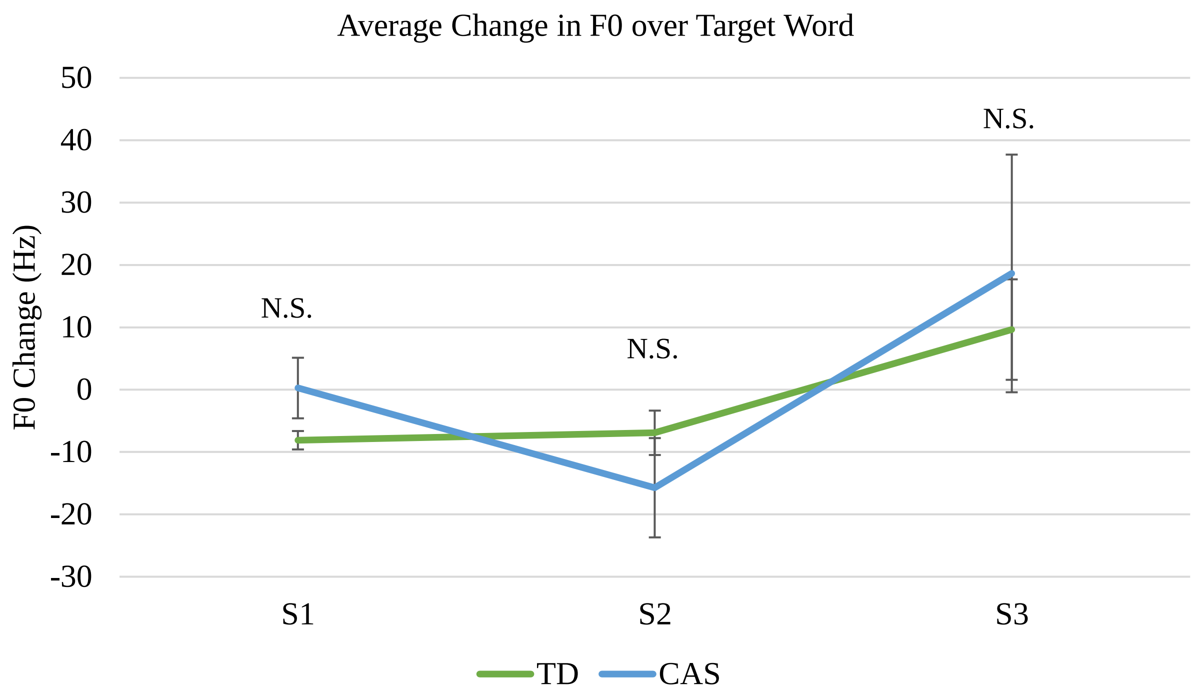

Figure 14. Average F0 differences between syllables within a target word for both groups.

An additional analysis calculated the linear least squares regression line of the average syllable F0 across the target word. The regression line was calculated using the average F0 of syllables one through four. The TD participants had a declining average regression line with a slope of -2.31 (SD = 9.58) and the CAS participants exhibited a declining average regression line with a slope of -3.78 (SD = 22.82). Table 10 shows the average F0 for each syllable of the target word, standard deviation, and standard error.

Table 10. Average F0 over the target word per syllable

		Syllable 1	**Syllable 2**	**Syllable 3**	**Syllable 4**
Average	*TD*	226.76	218.65	211.73	221.36
(Hz)	*CAS*	275.80	273.52	255.74	269.13
SD	*TD*	23.68	26.43	25.55	30.29
(Hz)	*CAS*	24.25	34.34	35.36	68.75
Standard	*TD*	7.49	8.36	8.08	9.58
Error	*CAS*	7.31	10.35	10.66	20.73

A Mann-Whitney U test was run to determine if there were differences in the slope of the linear regression lines across the target word, based on average F0 of the syllables in the target word for participants with and without CAS. The difference in average slope of the regression line was calculated using an exact sampling distribution for *U*. The regression lines for the target word were not significantly different between the TD group (*Mdn* = -6.88) and the CAS group (*Mdn* = -5.51), *U* = 78.000, *z* = 1.620, *p* = 0.114. Table 15 shows the F0 declination in both groups with average F0 of each syllable noted.

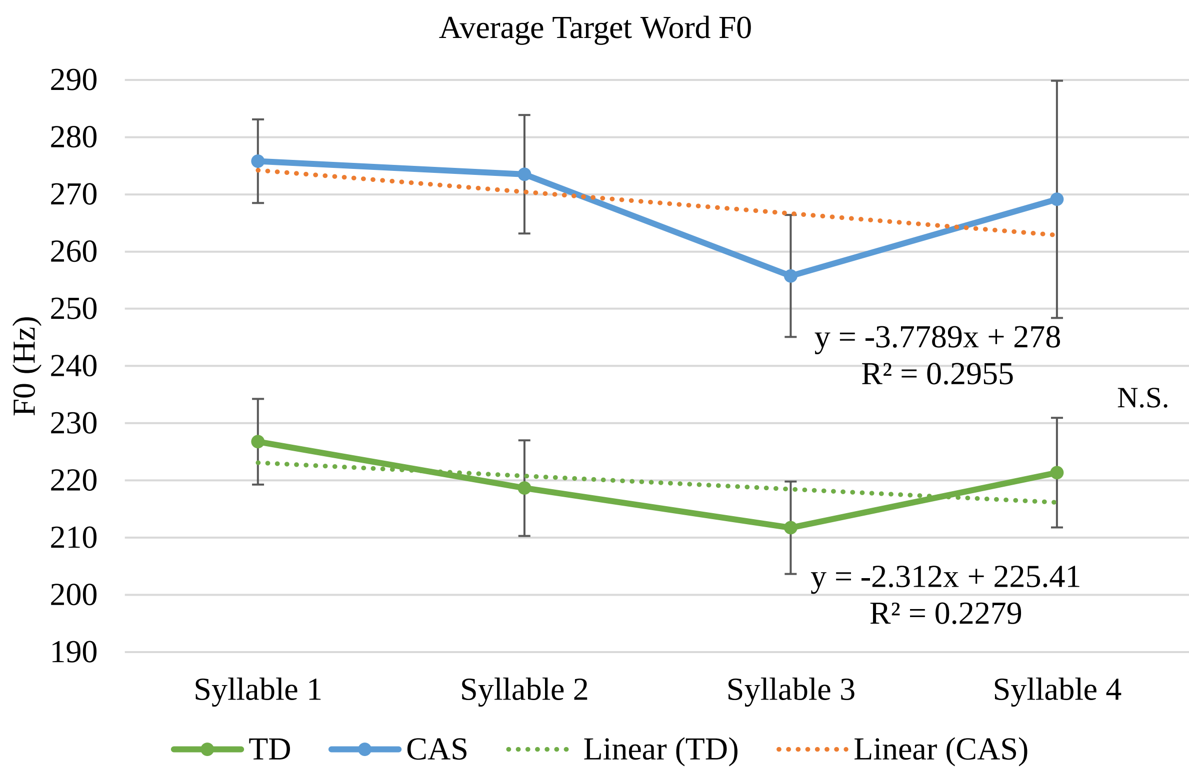

Figure 15. Average target word F0 per syllable in the word.

3.2.4. F0 over target word: real- and non-words

Analyses were run to evaluate the presence of any F0 differences between real target words versus non-real target words. A regression line was calculated using the average F0 of syllables one through three of real words, and included only words with initial stress on the first syllable. Analysis only included three-syllable words, rather than four, due to a limited number of four-syllable real target words. The TD participants had a declining average regression line with a slope of -4.78 (SD = 17.65) and the CAS participants had a declining average regression line with a slope of -16.40 (SD = 15.10). See Table 11 for average F0 for each syllable, standard deviation, and standard error. The regression lines for the utterances with real target words were not significantly different between the TD group (*Mdn* = -10.48) and the CAS group (*Mdn* = -18.82), U = 33.000, z = -1.549, p = 0.132. Table 16 shows the F0 declination for both groups over real target words.

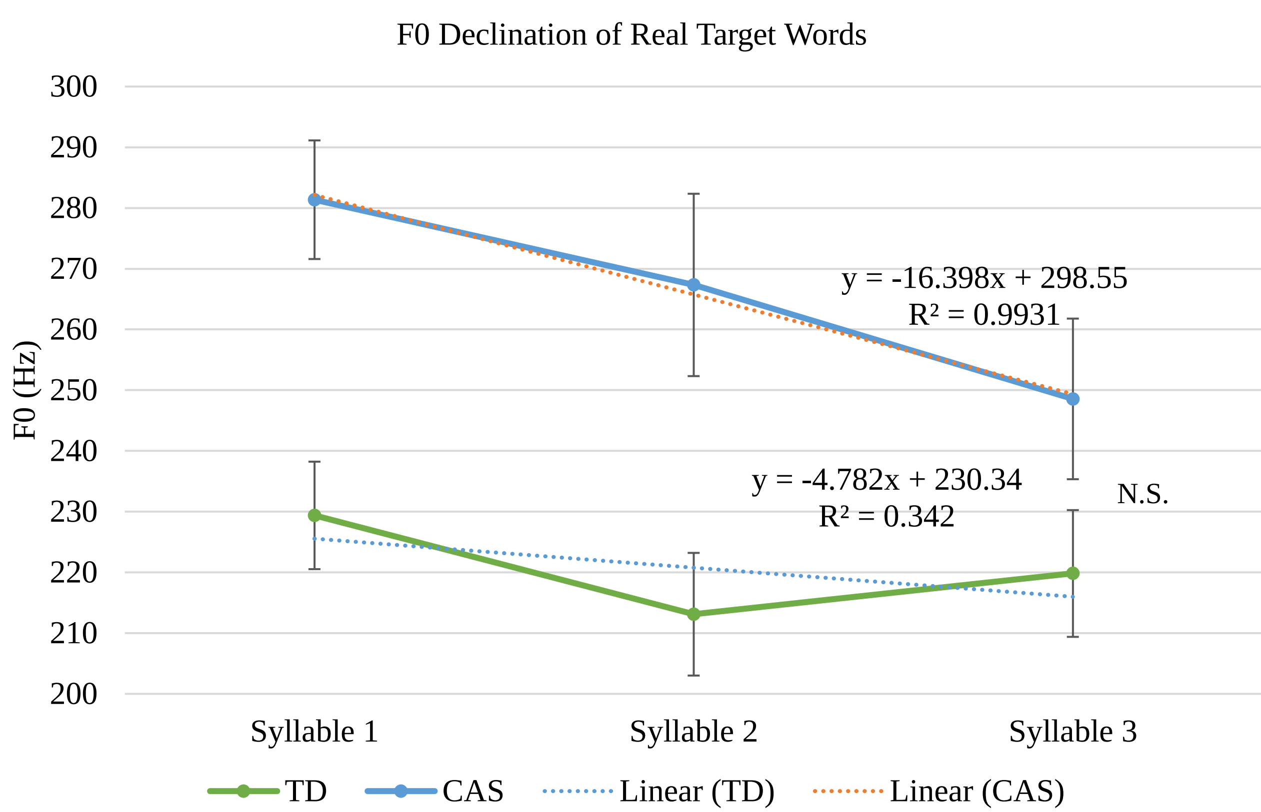

Figure 16. F0 declination of real target words for both groups.

A regression line was calculated using the average F0 of syllables one through four of all non-words with initial stress on the first syllable. The TD participants had a declining average regression line with a slope of -0.83 (SD = 10.62) and the CAS participants had a declining average regression line with a slope of -7.88 (SD = 23.53). See Table 11 for average F0 for each non-syllable, standard deviation, and standard error. The regression lines for the utterances with non-words were significantly different between the TD group (*Mdn* = -3.04) and the CAS group (*Mdn* = -17.44), *U* = 26.000, *z* = -2.04, *p* = 0.043. Table 17 shows the F0 declination for both groups over non-word target words.

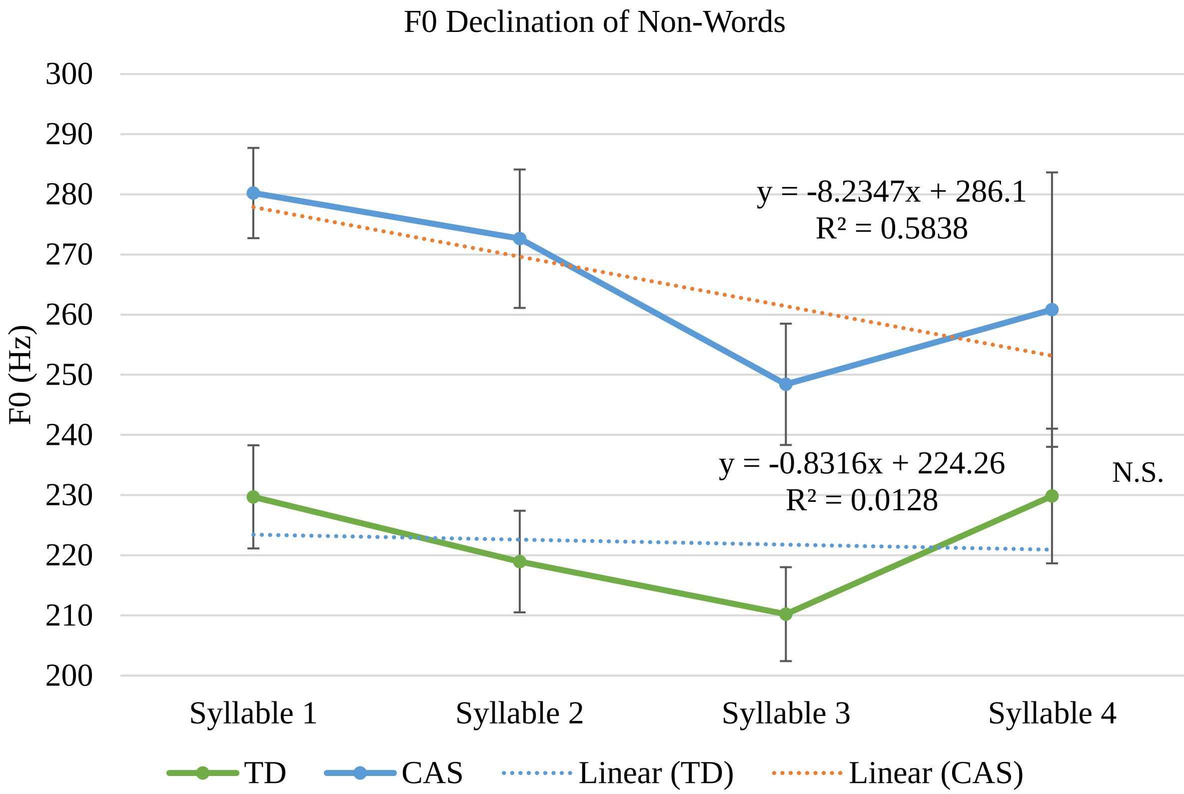

Figure 17. F0 declination of non-word target words for both groups.

Table 11. Average F0 over real and non-word target words

			Syllable 1	Syllable 2	Syllable 3	Syllable 4
Real Syllables	*Average (Hz)*	*TD*	229.38	213.11	219.82	NA
		CAS	281.36	267.34	248.57	NA
	SD (Hz)	*TD*	27.98	31.94	33.00	NA
		CAS	32.39	49.80	43.86	NA
	Standard Error	*TD*	8.85	10.10	10.44	NA
		CAS	9.77	15.01	13.22	NA
Non-Syllables	*Average (Hz)*	*TD*	229.71	218.96	210.22	229.85
		CAS	280.21	272.62	248.42	260.83
	SD (Hz)	*TD*	27.11	26.71	24.68	35.40
		CAS	24.88	38.15	33.44	75.65
	Standard Error	*TD*	8.57	8.45	7.80	11.20
		CAS	7.50	11.50	10.08	22.81

3. 3. Variable correlation

A Pearson's correlation was run to assess the relationship between inter-segment duration between-words and the linear regression slope of the utterance’s F0. All 21 participants were assessed. There was no statistically significant correlation between inter-segment duration between words and the linear regression slope, $r(19) = 0.275$, $p = 0.227$, with between word duration explaining 5.1% of the variation in the utterance’s F0 declination. Results show that as between word duration increases, the slope of linear regression of F0 flattens. Likewise, as between duration decreases, the slope of the regression falls or declines.

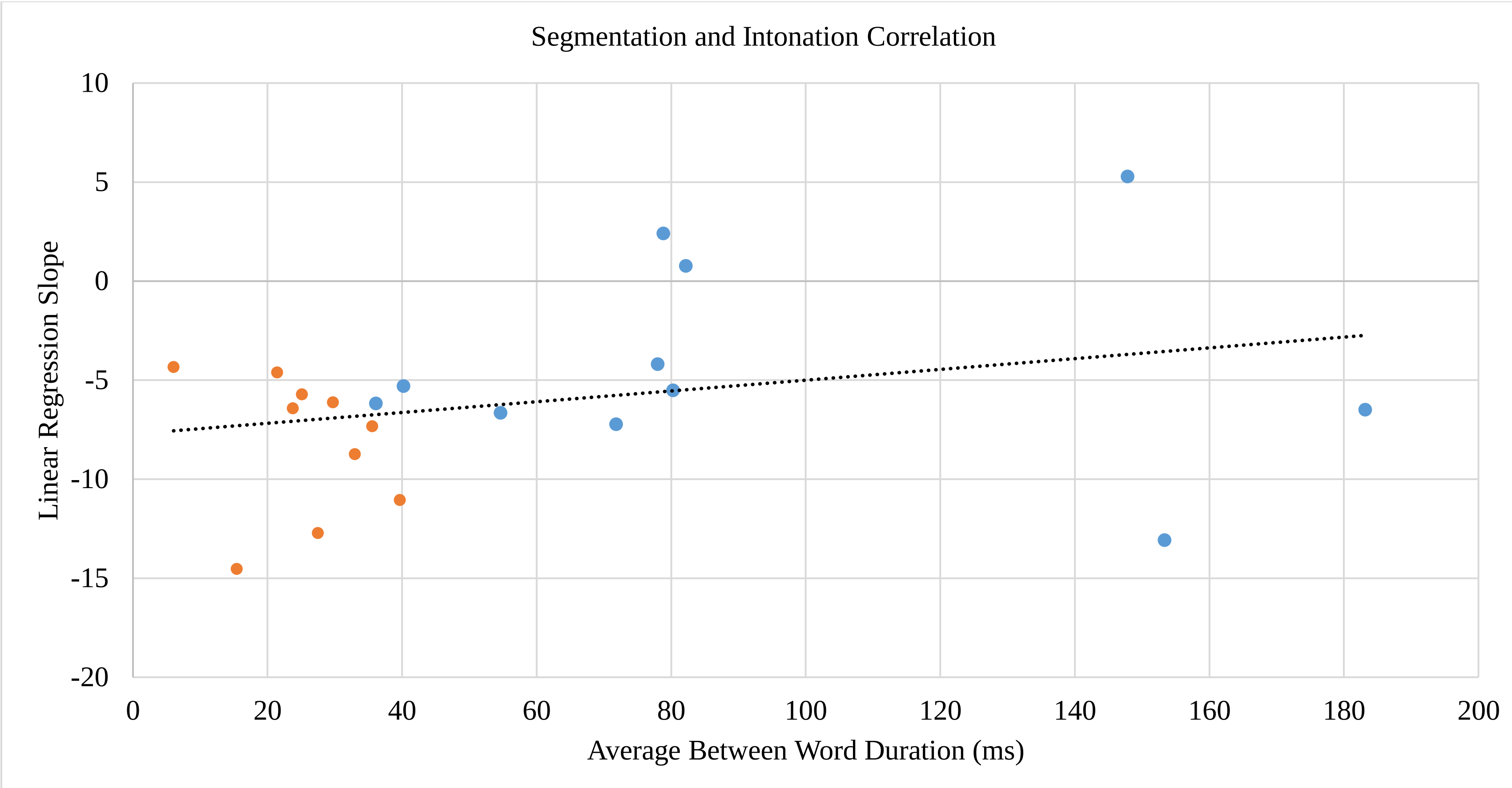

Figure 18. Pearson correlation depicting interaction between inter-segment duration between words and slope of the utterance related to F0 intonation. TD participants are represented in orange and CAS participants are represented in blue.

4. DISCUSSION

4.1. Segmentation

Segmentation is one of the three main diagnostic criteria of diagnosing and differentiating CAS (Kent & Rosenbek, 1983; Seddoh et al., 1996; Murry et al., 2015; Iuzzini-Seigel & Murray, 2017). Inter-segment duration was analyzed due to its prominence when deficits in motor programming are evident. This study asked the question about whether segmentation differences between CAS and TD children emerge as part of a working memory buffer impairment, motor programming deficit, or as a resource allocation issue due to complexity in speech production.

Past research has looked at segmentation or pauses for different motoric or linguistic reasons. The current study produced a comparison of segmentation for motor programming information between children with CAS and TD children. It provides a healthy control comparison to the CAS group from the TEMPO study prior to treatment, which provides insight into the development of further efficacy data following treatment.

The first hypothesis of this studyasking if children with CAS present with significantly longer inter-segment durations between and within words compared to TD children was accepted. Results showed a strong statistically significant difference in inter-segment duration between and within words in the CAS group. A comparison of medians suggests a longer duration in the CAS group than the TD group. Over the utterance, the CAS population had a significantly longer average inter-segment duration between words of 91.49 ms and respectively the TD group with 25.73 ms duration. Over the target word, the CAS population also showed a significantly longer duration of 75.01 ms compared with the TD group with 46.20 ms.

These findings are compatible with past findings on segmentation, justifying theories of the inabilities to concatenate speech segments into timed and coordinated programs for production. Past research has relied on using a trained clinician's perception to note inter-segment duration between and within words (Murray et al, 2015; McNeil et al., 1997; Seddoh et al., 1996), while no studies to date have gathered quantitative measures of inter-segment duration at the lexical level (between syllables) or sentence level (between words) (Kent & Rosenbek, 1983; Shriberg, 2017a et al.). The findings of this study present the first quantitative measures to the length of duration between segments of speech. This study measured segmentation at the utterance level (word to word) and lexical level (syllable to syllable), giving new insight into quantifying segmentation on a larger scale when diagnosing or observing acoustical differences between CAS and typical peers. Additionally, these data provide new insight into the organization of speech parameters relative to timing and duration for individuals with CAS.

In agreement with past findings, measurements of between and within word segmentation show that lengthening of inter-segments is likely to occur as utterances and words become more complex (Wright et al., 2009; Maas et al., 2008). The question raised was if it is possible that inter-segmental lengthening in CAS is in part due to a demand for resources when programming the motor plan into an intonational unit. Assigning the time, duration, intensity, and frequency to each of the parameters of speech units is causing the program and speech output to be disrupted.

The duration between words one and two had an average of 6.56 ms for the TD group, and the duration between words three and four (from the carrier phrase to the target word) was an average of 94.38 ms for the TD population. The radical differences in segmentation at the beginning of the utterance to the end reflect the theories of Wright et al. (2009) and Maas et al. (2008) regarding the demand for resources in motor programming. Similarly, the CAS population showed the same effect as the utterance became more linguistically complex and required more motor programming. The duration between words one and two was then an average of 39.60 ms for the CAS group and respectively, the duration between words three and four (from the carrier phrase to the target word) was an average of 202.86 ms for the CAS population.

Further correlation concluded a strong positive relationship between inter-segment duration between words and number of words in the sentence. The increased phonological and linguistic complexity of longer utterances is affecting the motor programming of speech units in CAS. Theories of resource allocation during motor programming explain these data and inter-segmental lengthening (Maas et al., 2008; Wright et al., 2009). Considering these data for clinical application will have high impact on diagnosis and treatment in CAS. During the motor program, relative time and duration is being planned, assigned, and programmed for speech production.

This study investigated the segmentation in CAS and TD participants and found strong evidence to support prior suggestions of longer inter-segment durations. Future analyses also quantifying the duration of each segment (syllable or word) could potentially also yield significant differences in productions between populations. Past research has found that segmental lengthening is also a differentiating characteristic for CAS children, due to the breakdown happening in the motor program (Seddoh et al., 1996; Maas et al., 2008; Wright et al., 2009). It would be interesting to observe if children with CAS are not segmenting their utterance, but instead significantly increasing the length of their segments. This would result in a perpetual difference in speech naturalness and intelligibility. Another question to explore is if it is possible that children are doing one and not the other because of inconsistences in their motor programming. Additionally, a second question for future research is to explore how the lengthening of segments could impact the change in intonation over the utterance or target word.

4.2. Intonation

Dysprosody is described in AOS literature as deficits impacting the rhythm, stress, and intonation (McNeil et al., 1997; Maas et al., 2008; Basilakos et al., 2017) and described in the prosodic literature as inconsistent prosody when deficits arise (Ladd, 1996). Past finding into prosodic differences in CAS conclude that deficits occur because of speakers compensating for accurate articulation and view dysprosody as a result of impaired motor planning (Basilakos et al., 2017). Most findings all observe prosody at the lexical level of the prosodic hierarchy, differentiating stress patterns in weak-strong versus strong-weak stress placements (Kent & Rosenbek, 1983; Ballard et al., 2016).

This study proposed questions asking how children with CAS use prosody at the intonational phrase level and how this relates to children who are typically developing. Critically, the current study provides new insights into how both groups change their F0 over an utterance and within a multisyllabic target word. Data gathered offer new discussion as to whether children with CAS have a highly robust prosodic system of the intonational phrase or if they are experiencing breakdowns due to a resource allocation issue.

The second hypothesis in this study that proposed longer inter-segment durations expressed by children with CAS will influence F0 declination more than for TD children was partially accepted. This is because this research provides new insight into intonation and declination patterns. The longer inter-segment durations noted in the CAS group are believed to be influencing F0 declination over the utterance to a certain degree. This is evidenced by a statistically significant difference in change in F0 from the penultimate word of the carrier phrase into the multisyllabic target word. Other influences on declination include an inconsistency in F0 intonation of falling word to word in the utterance with a higher pitch reset in Hz in the CAS group than in the TD group. These findings relate to past theories and findings related to motor programming and prosody.

Partial rejection of the second hypothesis was considered since both populations demonstrated declination of F0 but did not differ statistically between one another. The TD and CAS groups both showed declination over the utterance, with the TD group measuring a linear regression of -8.16 slope and the CAS group measuring a slope of -4.20. Additionally, the overall pattern for the individual's with CAS showed high interspeaker inconsistency due to higher standard deviations than the TD group. These findings are comparable to theories that F0 declination is a highly robust characteristic of speech, which even though variable between speakers, does not require any preprogramming into the linguistic code (Kent & Rosenbek, 1983; Ladd, 1996).

Accepting this theory that programming initial high F0 at the beginning of a sentence is not necessary, and instead innate, does not however explain for the breakdown resulting in F0 reset experienced by the CAS group. The breakdown in F0 from word to word, resulting in F0 reset instead connects to other theories of resource allocation issues, causing the second hypothesis to be partially accepted. During the natural slow release of tension and subglottal pressure from the larynx as muscles relax, there is time for the preparation of the vocal folds to relax and an innate decrease in F0. These findings suggest that in declarative sentences, F0 declination may not be carried in the linguistic intent during motor programming, but instead is a natural phenomenon of the speech mechanism.

Evidence of differences in the change in F0 between words from carrier phrases into the multi-syllabic target word was also found in both groups. Participants with CAS exhibited a higher F0 on the target word than on the final and penultimate words of the carrier phrase (increase in F0 from 261.84 Hz to 272.89 Hz), while the TD group continued to decrease their F0 following the carrier phrase. That is, F0 significantly increased from the penultimate word of the carrier phrase to the target word by 11.50 Hz for CAS and declined by 8.41 Hz in the TD group. This evidence of F0 reset explains that the CAS population is potentially programming the target word as an entirely new utterance (Maas et al., 2008).

The F0 reset on the penultimate word of the carrier phrase raises questions to why the CAS participants are resetting before the phrase is over and not right at the onset of the target word. Acoustical analysis for the current study collapsed both four-word and five-word utterances for a combined analysis, impacting the location of the target word in the utterance. In the current analysis, the target word landed in the "word 4" location and others in the "word 5" location, further impacting the location of F0 reset. Future analyses should separate out four-word and five-word utterances for a more in-depth analysis into the significant differences of pitch reset for both groups following the carrier phrase and into the target word. Another aspect for further analysis includes analyzing the duration of the carrier phrase. Gathering a correlation of the length of the carrier phrase to the degree of F0 reset would provide insight into past theories of motor learning (Maas et al., 2008; Seddoh et al., 1996).

Change in F0 over the target word did not conclude notable differences in declination between both groups, as past findings have concluded with lexical stress (Arciuli & Ballard, 2017; Kent & Rosenbek, 1983). Both groups were noted to decline in the slope of linear regression of F0 at a similar rate with a slope of -2.31 for the TD group and -3.78 for the CAS group. Additionally, there was evidence of F0 reset found for both groups as they approached the final syllable of the multisyllabic target word. Although F0 reset was noted for both groups on the final syllable of the multisyllabic words, the degree of F0 reset was twice as high for the CAS population with a 2:1 ratio (9.63 Hz change for TD group and 18.64 Hz change for CAS group).

Past studies looking at prosody in CAS have only looked at the initial dyad of the first two syllables in multisyllabic words, looking at either the strong-weak or weak-strong relationship (Arciiuli & Ballard, 2017; Miller et al., 2018). This study looked at the F0 of each syllable in the multisyllabic words, with strong-weak initial stress patterns, finding evidence of pitch reset on the last syllable. Research looking further into the syllable structures of words should be analyzed for more information into the stress patterns of additional syllables.

4.3. Limitations

The limitations in the current study were noted in regard to the participants, data collection, and data analysis. The current study was limited by the small sample size. Further data from additional participants will better differentiate prosodic intonation between TD children and children with CAS. Additionally, the average age ranges of the TD participants were noted to be slightly higher than the participants with CAS, which has potential impacts on their speech production. The participants in this study were also noted to have different average CELF-5 scores, however all in the average range.

Another limitation was due to missing data that could not be acoustically collected to capture the pitch for all words and syllables in each utterance. Pitch tracking errors noted in the Methods section reflect items that needed to be excluded from analysis. The results of F0 declination over the target word did not yield significant differences, however, the data did contain a large degree of variation in data from one of the CAS participants.

4.4. Broader Impacts

While the outcomes of this study were specific to this particular sample of participants, it also contained a wide range of characteristics in agreement with past studies on CAS and typically developing children. Therefore, it is reasonable to present these findings when examining the characteristics of CAS compared to children who are typically developing. Many findings from this study produce a boarder understanding for comparing and justifying to past findings, new research opportunities, and clinical application.

Comparing speech production patterns in CAS with typically developing children pre- and post-treatment will better establish treatment efficacy in improving communication of children with CAS. For example, future research should compare this group of CAS participants post-TEMPO treatment to the TD population. This treatment program focuses on targeting multisyllabic words, rather than over sentences, in order to directly target the motor programing mechanism. Treating CAS at the word level, to focus on syllable segmentation, would then generalize word to word over utterances. Past research and intonational findings from the current study also suggest that targeting the strong-weak syllable stress control will also target treating the mechanism for correcting inconsistencies in prosody.

Additional future research should investigate within-subject variability and group variability, with the potential to review more data regarding predictions of intelligibility, naturalness, and articulatory accuracy in the CAS population. It would also be productive to further match each of the CAS participants with an apraxia severity rating for correlation analysis. If each CAS participant was assigned into a numeric category to characterize the perceptual severity of their apraxia, then further analyses could investigate how segmentation changes based on severity.

Overall, the current study provides further insight into the motor programming of words and syllables and impacts on segmentation and intonation. Acoustical coding and analysis for diagnosing and treating CAS is important in clinical application for measuring progress for quantifying segmentation and pitch, which in the past have only been measured perceptually. Future research looking into automating the coding process in real time would bring new and innovative services to the field of communication sciences and disorders.

REFERENCES

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

Arciuli, J., & Ballard, K. J. (2017). Still not adult-like: lexical stress contrastivity in word productions of eight-to eleven-year-olds. *Journal of child language, 44*(5), 1274-1288.

Ballard, K. J., Robin, D. A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 53* (5), 1227-1245. doi:10.1044/1092-4388(2010/09-0130)

Ballard, K. J., Djaja, D., Arciuli, J., James, D. G., & Doorn, J. V. (2012). Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research, 55*(6), 1822-1835. doi:10.1044/1092-4388(2012/11-0257)

Ballard, K. J., Tourville, J. A., & Robin, D. A. (2014). Behavioral, computational, and neuroimaging studies of acquired apraxia of speech. *Frontiers in Human Neuroscience, 8*. doi:10.3389/fnhum.2014.00892

Ballard, K. J., Azizi, L., Duffy, J. R., McNeil, M. R., Halaki, M., O'Dwyer, N., . . . Robin, D. A. (2016). A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia, 81*, 129-139. doi:10.1016/j.neuropsychologia.2015.12.010

Basilakos, A., Yourganov, G., Ouden, D. D., Fogerty, D., Rorden, C., Feenaughty, L., & Fridriksson, J. (2017). A multivariate analytic approach to the differential diagnosis of apraxia of speech. *Journal of Speech Language and Hearing Research, 60*(12), 3378-3392. doi:10.1044/2017_jslhr-s-16-0443

Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer [Computer program]. Version 6.0.49, retrieved 2 March 2019 from http://www.praat.org/.

Case, J., & Grigos, M. I. (2016). Articulatory control in childhood apraxia of speech in a novel word–learning task. *Journal of Speech Language and Hearing Research, 59*(6), 1253. doi:10.1044/2016_jslhr-s-14-0261

Connell, B., & Ladd, D. R. (1990). Aspects of pitch realization in Yoruba. *Phonology*, *7*(1), 1-29.

Dankoviov, J., Pigott, K., Wells, B., & Peppé, S. (2004). Temporal markers of prosodic boundaries in children's speech production. *Journal of the International Phonetic Association, 34*(1), 17-36. doi:10.1017/s0025100304001525

De Looze, C., Yanushevskaya, I., Murphy, A., O'Connor, E., and Gobl, C. (2015). Pitch declination and reset as a function of utterance duration in conversational speech data, *Interspeech*, *7*(10), 3071 – 3075.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management.* St. Louis, MO: Elsevier.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press.

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Iuzzini-Seigel, J., & Murray, E. (2017). Speech assessment in children with childhood apraxia of speech. *Perspectives of the ASHA Special Interest Groups*, *2*(2), 47. doi:10.1044/persp2.sig2.47

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. doi:10.1111/j.1467-8624.1993.tb02935.x

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, *39*(3-4), 159-207. doi:10.1006/cogp.1999.0716

Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *26*(2), 231-249.

Ladd, D. R. (1996). *Intonational Phonology* (Vol. 79). Cambridge University Press.

Marotta, G., Massimiliano, B., & Paolo, B. (2008). Prosody and Broca's aphasia: An acoustic analysis. *Study Linguistics Philological Online*, *6*(1), 79-88.

Maas, E., Robin, D., Hula, S., Freedman, S., Wulf, G., Ballard, K., & Schmidt, R. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, *17*, 277-298.

Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain and Language*, *106*(2), 107-118. doi:10.1016/j.bandl.2008.03.004

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.

Murray, E., McCabe, P., Heard, R., & Ballard, K. J. (2015). Differential diagnosis of children with suspected childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research,* 58(1), 43–60.

Merwe, A. V., & Steyn, M. (2018). Model-driven treatment of childhood apraxia of speech: Positive effects of the speech motor learning approach. *American Journal of Speech-Language Pathology, 27*(1), 37. doi:10.1044/2017_ajslp-15-0193

Miller, H., Plante, A. S., Campbell, J. L, Ballard, K. J., & Robin, D. A. (2018). Treatment of lexical stress, segmentation, and sound distortions in childhood apraxia of speech. Presented at the *Conference on Motor Speech*, Savannah, Georgia.

Peppé, S., Mccann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech Language and Hearing Research, 50*(4), 1015. doi:10.1044/1092-4388(2007/071)

Preston, J. L., Leece, M. C., Mcnamara, K., & Maas, E. (2017). Variable practice to enhance speech learning in ultrasound biofeedback treatment for childhood apraxia of speech: A single case experimental study. *American Journal of Speech-Language Pathology*, *26*(3), 840. doi:10.1044/2017_ajslp-16-0155

Seddoh, S. A., Robin, D. A., Sim, H., Hageman, C., Moon, J. B., & Folkins, J. W. (1996). Speech timing in apraxia of speech versus conduction aphasia. *Journal of Speech Language and Hearing Research, 39*(3), 590. doi:10.1044/jshr.3903.590

Selkirk, E. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry*, 11(3), 563-605. Retrieved from http://www.jstor.org.libproxy.unh.edu/stable/4178179

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*, *4*, 279–284.

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., ... Wilson, D. L. (2017a). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: I. Development and description of the pause marker (PM). *Journal of Speech, Language, and Hearing Research, 60*, S1096–S1117. doi.org/ 10.1044/2016_JSLHR-S-15-0296

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., . . . Wilson, D. L. (2017c). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: III. Theoretical coherence of the pause marker with speech processing deficits in childhood apraxia of speech. *Journal of Speech Language and Hearing Research*, *60*(4). doi:10.1044/2016_jslhr-s-15-0298

Wright, D. L., Robin, D. A., Rhee, J., Vaculin, A., Jacks, A., Guenther, F. H., & Fox, P. T. (2009). Using the self-select paradigm to delineate the nature of speech motor programming. *Journal of Speech Language and Hearing Research*, *52*(3), 755. doi:10.1044/1092-4388(2009/07-0256)

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707-1717. doi:10.1121/1.402450

Wiig, E. H., Semel, E., & Secord, W. A. (2013). Clinical Evaluation of Language Fundamentals–Fifth Edition (CELF-5). Bloomington, MN: NCS Pearson.

Yuan, J., & Liberman, M. (2014). F0 declination in English and Mandarin broadcast news speech. *Speech Communication*, *65*, 67-74. doi:10.1016/j.specom.2014.06.001