2012-08-03

An Acoustic Analysis of Elements of Contrastive Stress Produced by 8 to 10-Year-Old Children

Nicole Michelle Clover
Brigham Young University - Provo

An Acoustic Analysis of Elements of Contrastive Stress

Produced by 8 to 10-Year-Old Children

Nicole M. Clover

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Shawn L. Nissen, Chair
Christopher Dromey
Ron W. Channell

Department of Communication Disorders

Brigham Young University

December 2012

ABSTRACT

An Acoustic Analysis of Elements of Contrastive Stress
Produced by 8 to 10-Year-Old Children

Nicole M. Clover
Department of Communication Disorders
Master of Science

Contrastive stress is an aspect of communication that can be used to highlight information, de-accent redundant information, and create distinctions between new and previously-provided information. Previous research has documented that adult speakers use relative changes in their vocal intensity, fundamental frequency (F0), and duration to mark contrastive stress in a sentence. However, less is understood about how and when children mark contrastive stress in their communication, thus the current study aims to examine a number of acoustic elements of contrastive stress in 8 to 10-year-old children. Speech samples were elicited from 20 children and analyzed to determine if the acoustic parameters of F0, intensity, and duration varied as a function of the speaking condition, speaker gender, or grammatical unit. Results of the experiment suggest that when comparing the baseline speaking condition to the speaking condition eliciting contrastive stress, significant differences were only found for the acoustic measure of mean intensity. Additionally, gender-related differences in contrastive stress were found only for the dependent measure of F0 slope, with a greater F0 slope exhibited by female speakers. All grammatical units were significantly different from one another across a number of variables, with significant interactions between baseline and target conditions and grammatical unit being analyzed. As indicated in previous research, the findings of the present study may indicate that children under 10 years of age may not have developed contrastive stress in an adult-like manner. Results may also be due to individual speaker differences, the complex nature of prosody, or measurement methodology.

Keywords: Prosody, contrastive stress, acoustic analysis, children, prosodic development

ACKNOWLEDGMENTS

I would like to thank my thesis chair, Dr. Nissen, for his continued support and guidance and for the hours spent reviewing many drafts of my thesis. I could not have completed this process without all of your help and guidance and appreciate all of your time and effort on my behalf. I would also like to thank Dr. Channell and Dr. Dromey for your encouragement, time, and insights.

I would also like to thank my family for their support and love both financially and emotionally throughout my undergraduate and graduate degrees. Your support has made a world of difference as I have worked through each part of my graduate degree and especially as I have completed each section of my thesis. My appreciation also goes out to the members of my cohort, who helped me through graduate school with their friendship and collaboration. You provided me with much-needed support in many ways.

I am extremely grateful for my husband and best friend, Richard. Thank you for dealing with me during times of stress and also celebrating my successes. I appreciate you marrying me in the midst of chaos. Your perspective and love have been invaluable to me and have helped me throughout the last two years.

Table of Contents

List of Tables

List of Figures

Description of Structure and Content

The body of this thesis is written as a manuscript suitable for submission to a peer-reviewed journal in speech-language pathology. An annotated bibliography is presented following the reference section.

Introduction

Speech communication involves the production and perception of a variety of different sound segments that are combined to form words and sentences. However, speech also has suprasegmental aspects, such as prosody, which extend across multiple sounds and words in a sentence. Prosody serves a number of linguistic purposes, including signaling the sentence type, creating linguistic boundaries, clarifying lexical ambiguities, and marking syllabic and contrastive stress (Dromey, 2010; Panagos & Prelock, 1997; Patel & Brayton, 2009). The nature of contrastive stress can be highly complex depending on the acoustic mechanism used to mark the emphasized segment and the stage of development of the speaker.

Contrastive stress can be used to emphasize or highlight one component within a sentence, or a sentence within a series of sentences (Bolinger, 1978). For example, in the sentence *I sent for them to come to the **meeting**, not to the **concert***, the words *meeting* and *concert* are set apart from the remainder of the sentence using contrastive stress. Marking the comment or psychological predicate of an utterance, or in other words creating a distinction between new versus previously-provided information, can also be accomplished using contrastive stress (Hornby & Hass, 1970). Contrastive stress can be used by a speaker to de-accent redundant information and emphasize a word to repair breakdowns and misunderstandings in a conversation as well (Dromey, 2010).

Acoustic Mechanisms of Contrastive Stress

Contrastive stress is typically not produced using a single acoustic mechanism, but rather a combination of multiple acoustic features. Previous studies have indicated that adult speakers primarily express differences in contrastive stress through changes in relative F0, with some secondary differences in intensity and duration (Bolinger, 1978). From a speaker's perspective, the acoustic mechanisms that produce contrastive stress are the result of differences in respiratory or vocal effort, whereas the cumulative effect on the listener is often a result of relative differences in the perceptual prominence of a syllable, word, or phrase (Bauman-Waengler, 2009).

Relative changes in a speaker's pitch are one means through which contrastive stress is perceptually identified by listeners. Stressed words in an utterance are typically marked by an increase in pitch relative to unstressed words (Patel & Brayton, 2009). In addition, sentences containing phrases that are contrastively stressed are typically produced with greater overall pitch range (Ladd, 1993). Research has also indicated that, when compared to female speakers, adult males may rely less on changes in F0 to mark contrastive stress, yet use F0 changes to a greater degree when signaling syntactic and phrase-final boundaries (Cooper, Eady, & Mueller, 1985).

Another acoustic mechanism to mark contrasts in speech is an increase in the relative loudness or amplitude of a syllable, word, or phrase. Lehiste and Peterson (1959) proposed that the perception of contrastive stress is based less on the loudness of the consonants within a word or phrase and more on the listener's judgment of the physiological effort used to produce the vowels within a segment of speech, even as fundamental frequency and duration are held constant.

An increase in the duration of a syllable or word is also commonly used by speakers to mark contrastive stress (Patel & Brayton, 2009). Typically this is accomplished by increasing the length of the vowel within a syllable. A study by Patel and Grigos (2006) noted that young children around the age of 4 are especially reliant upon durational cues to mark contrasts in their own speech and comprehend it in the speech of others, while older age groups tend to use a combination of acoustic cues to make contrasts. However it is unclear if the gender of the speaker affects how the acoustic cue of duration is used to signal contrastive stress in children or adults (Beach, Katz, & Skowronski, 1996). It has also been suggested that contrastive stress may be marked to a lesser degree by differences in timing, such as speech rate, as well as the distribution and length of pauses within a sentence or conversation.

Development of Contrastive Stress in Children

Less is known about how relative changes in speaking F0, intensity, and duration are used by children to express contrastive stress. The acoustic features used to mark stress may begin to develop from acoustic elements of prosody acquired in infancy (Bauman-Waengler, 2009). Research has indicated that during the infant's canonical babbling stage, which begins around 6 months of age, the infant uses intonation, rhythm and pausing in a consistent way (Davis, MacNeilage, Matyear, & Powell, 2000), allowing them to begin building a foundation upon which they can later use these skills to mark contrasts in speech. Elements of rising and falling pitch are also observed in infancy. It was reported by Galligan in 1987 that by 17 months of age, English-speaking children used falling vs. rising tones when contrasting one-word declarative statements with utterances that were a request for assistance.

Even before 17 months of age, children have distinct rises and falls in the pitch contour, possibly indicating a sensitivity to contour direction in language (Snow, 1995), which is an essential step in the development of contrastive stress. It has also been found that adult-like intonation patterns are used by some children in the late babbling period, in which children are experimenting with sounds but have not yet produced their first words. In other children, the first meaningful words are produced before intonation patterns appear (Snow, 1995). This development of falling and rising tones is an additional step in providing the necessary background for a child to learn and use contrastive stress in discourse.

The ability to manipulate the pitch, loudness, and duration of individual syllables provides a foundation for the later development of contrastive stress within two-word and longer utterances (Bauman-Waengler, 2009). Initially, children learn to produce multi-syllabic words with one syllable having greater or lesser prominence, as in the word *daddy,* which is typically produced with stress on the first syllable. Next, they learn to generalize this ability to produce relative differences in stress to the longer utterances or phrases, often in a sequential order. This sequence is illustrated by Bauman-Waengler in the following developmental pattern of the two words *daddy eat*:

Daddy (pause) eat

Daddy (pause shortens) eat

‘Daddy ‘eat (no pause, both stressed)

‘Daddy eat (first word stressed)

Children then develop the ability to use contrastive stress to signal differences in meaning when using the same or similar words. For example, a child saying ***Daddy*** *eat* would signify that *it is daddy who is eating*, while *Daddy* ***eat*** could indicate that *Daddy needs to sit down and eat*.

Although children use contrastive stress at an early age, their production is likely not equivalent to that of an adult speaker. While children’s usage of contrastive stress may involve some of the formal properties of the adult intonational system, and their early pitch patterns reflect linguistic influences of intonation, a child’s knowledge of prosody and contrastive stress continues to develop with maturity (Snow, 1995). Cruttenden (1985) suggested that although some of the core features of contrastive stress are used both receptively and expressively by young children in early stages of development, other more subtle features of intonation, including true contrastive stress, cannot be understood or produced in adult-like ways until a child is at least 10 years of age.

In an experiment by Patel and Brayton (2009), which was performed to determine the age at which an unfamiliar listener was able to identify contrastive stress in children, adult listeners rated the production of contrastive stress in children ages 4, 7, and 11. Significant differences were found between the 4-year old age group and the 7 and 11-year-old age groups. Listeners were more successful in identifying contrastive stress placed on sentence-initial words compared with sentence-final words across all age groups, indicating that mastery of contrastive stress was not yet at an adult-like level for the majority of the children in the study. A study by Wells et al. (2004) supported these conclusions, reporting that prosodic comprehension and other aspects of intonation continue to develop between 5 and 11 years of age.

A recent study by Anita Dromey (2010) evaluated the use of contrastive stress by a group of twenty pre-adolescent children between eight and ten years of age. This study involved eliciting simple sentences from these children. Each sentence was linguistically constructed to allow the children to contrastively stress the subject, verb, or object of the sentence (e.g., *The boy is carving the pumpkin*). The resulting speech recordings were then acoustically analyzed to examine whether and how the children marked the contrastive element of the sentence through relative changes in the F0, intensity and duration of the target section of speech when compared to a baseline production of the sentences with no intended emphasis. These results were also compared to a set of data extracted from 10 adult speakers collected in a similar manner by Matthews (2010). Results of the experiment showed that speakers decreased their F0 when emphasizing the subject of the target sentences, yet increased their F0 when emphasizing verb and object parts of the sentence. No significant correlations or consistent changes were observed with intensity during the tasks, but in general children were observed to shorten the duration of words intended to be contrastively stressed. Dromey (2010) concluded that differences between the findings of this study and previous research may have been due to the methods of analysis employed in the study.

Purpose of the Study

The purpose of the current study was to analyze the data collected by Dromey (2010). Specifically, one aim of the current study was to measure the emphasis or stress of a target word relative to both a separate baseline production and the neighboring words within the same utterance. In addition, this study analyzed the intensity and F0 of a contrasted word or phrase in terms of not only mean, but also the peak and slope values. It may be that a word production with a sharp burst in intensity has a similar or greater perceptual impact on a listener compared with a production that has a sustained increase in intensity.

Method

The section below gives a brief description of how the speech samples used in this study were elicited and collected by Dromey (2010) and Matthews (2010), as well as the manner in which the data were reanalyzed in the present study. In addition, this section describes the measurement and interpretation of a supplementary set of acoustic values.

Speech Samples

The speech samples evaluated in this study were collected from 20 children between 8:0 and 9:11 years of age (M = 9.2) and 10 adults in the comparison group, who were between 21 and 28 years of age (M = 23.9). Each age group had an equal number of male and female participants, who were monolingual speakers of American English with minimal exposure to a second language (i.e., not living outside US for more than 6 months and parents/guardians with English as first language). None of the participants reported any history of a speech, language, or hearing problem. In addition, all participants passed a hearing screening with pure-tone air-conduction hearing thresholds < 25 dB HL at octave frequencies from 500 to 8000 Hz.

The speech recordings were elicited using five line-drawing pictures of simple everyday events. Each picture was designed to elicit four lexically identical sentences from each participant. One of the sentences required contrastive stress on the subject, verb, or object (which was the head word of the relevant phrase), in addition to a baseline sentence produced naturalistically but without emphatic stress placed on a certain word.

Presentation of experimental stimuli began after familiarity with the task had been demonstrated. To begin, the researcher elicited baseline productions of the ten target sentences described above. Following baseline elicitation, the researcher presented a series of five pictures, one at a time, to the participants. Upon presenting each picture, the researcher asked three questions: one each about the subject, the verb, and the object of the sentence. For example, the participants were shown a picture of a child in a swimming pool. Then, they were presented with the first question, which was intended to elicit a response about the subject of the target sentence, *Is a dog in the swimming pool?* This sentence prompted the participant to respond, *No, a* ***child*** *is in the swimming pool.* The second and third questions were constructed in a similar manner in order to elicit a response about the verb and the object of the sentence, respectively.

A total of 20 sentences were recorded from each participant in this manner. Sentences produced in the five picture elicitations were designed to be similar syntactically and of a similar length and complexity.

Recording

The participants' speech was recorded directly to a PC computer while each participant was seated in a quiet room. Speech samples were recorded using a high-quality head-set microphone (Shure 4011), positioned approximately 2.5 centimeters from the participant's mouth. The sampling rate of the recordings was 44.1 kHz with a quantization of 16 bits with Adobe Audition software. Sound files were subsequently archived to a PC computer hard drive for further analysis. All sentences recorded were high-pass filtered at 70 Hz. If inaccurate articulation, peak clipping, or an error in the recording took place, the participant was asked to repeat the test item and the stimulus was re-recorded.

Measurement of Acoustic Variables

F0 measurements. The mean F0 values for the target grammatical units (subject, verb, and object) were previously extracted by Dromey (2010). These data were extended in the current study by exacting the mean F0 of the neighboring non-target content words in each sentence, the maximum and minimum F0 for each target word (subject, verb, and object), and the maximum and minimum F0 for the neighboring non-target words. Praat acoustic analysis software (version 5.1.20; Boersma & Weenink, 2009) was used to measure the F0 values of each target word by extracting an F0 track plotted over time. The algorithm for the extraction of the F0 track relied on autocorrelation, as described in Boersma (1993). To determine the amount of relative emphasis of the target segment, a ratio of the mean F0 of the target word to the mean F0 of the other words in the sentence was calculated. This value was then compared to a similar ratio calculated for the baseline sentence (the sentence with the identical linguistic structure produced without any target emphasis). This procedure was duplicated for the F0 peak values.

Intensity measurements. The average root-mean-square (RMS) intensity values for the target grammatical units (subject, verb, and object) were previously extracted by Dromey (2010). These data were extended in the current study by exacting the mean intensity of the neighboring non-target words in each sentence, the maximum intensity for each target word (subject, verb, and object), and the maximum intensity for neighboring non-target words in each sentence. Praat acoustic analysis software (Boersma & Weenink, 2009) was used to measure the intensity values of each target word by extracting an intensity track plotted over time. The algorithm for the extraction of the intensity track is further described in Boersma (1993). To determine the amount of relative emphasis of the target segment, a ratio of the mean intensity of the target word to the mean intensity of the other words in the sentence was calculated. This value was then compared to a similar ratio calculated for the baseline sentence. This procedure was duplicated for the intensity peak values.

Duration measurements. The duration values for the individual target words (subject, verb, and object) were previously extracted by Dromey (2010). These data were extended in the current study by exacting the duration of the neighboring non-target words in each sentence. These duration measures were computed to the nearest millisecond (ms) using the Praat analysis software, utilizing both waveform and spectrographic displays for the analysis. From the measures, a ratio between the target word and the words in the same sentence was calculated. This value was then compared to a similar ratio calculated for the baseline sentence.

Reliability of the Measures

To establish reliability of the extracted acoustic measures, speech samples from 10% of the speaker productions were selected and reanalyzed by another individual. The additional sets of duration, intensity, and F0 measurements that resulted were extracted, recorded, and checked in the same manner as the original measures. Comparisons of durational measures produced correlations of 0.96, F0 measures produced correlations of 0.89, and intensity measures produced correlations of 0.99.

Results

Data in this experiment were analyzed using a repeated measures analysis of variance (ANOVA) to determine the significance of acoustic changes in speakers' productions as a function of the speaking condition (baseline versus target emphasis), grammatical unit, and the speaker's sex. As described in the previous section, the dependent measures included the ratios for duration, F0 mean, F0 peak, intensity mean, intensity peak, and F0 slope. Partial eta squared (η2) measures of effect size were computed for ANOVA results that were significant. Post hoc analyses consisted of pairwise comparisons, with Bonferroni adjustments for multiple comparisons. The descriptive statistics for male and female participants are illustrated in Tables 1 and 2.

The statistical analysis indicated a significant difference between the speaker's baseline and target productions for the dependent measure of intensity mean ratio $F(1, 18) = 14.15$, $p = .001$, $\eta^2 = .44$. The intensity mean ratio for the baseline sentences was 1.00, while the same ratio for the target sentences averaged 1.03. Although the difference in intensity mean ratios appears small, the difference between the intensity values of the target word in the baseline sentence versus the target sentence was 1.4 dB, .95 dB, and 1.0 dB for the subject, verb, and object grammatical units, respectively.

There was an interaction between the speaking condition and the grammatical unit for the dependent measures of duration ratio, $F(2, 18) = 4.49$, $p = .026$, $\eta^2 = .20$, F0 peak ratio, $F(1, 18) = 8.06$, $p = .004$, $\eta^2 = .31$, and intensity peak, $F(2, 18) = 4.79$, $p = .016$, $\eta^2 = .21$. As shown in

Table 1

Descriptive Statistics of Female Participants by Condition and Acoustic Parameter

		Subject		Verb		Object	
Acoustic Parameter	Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
Duration[a]							
	Baseline	0.14	0.02	0.21	0.02	0.21	0.02
	Target	0.13	0.01	0.21	0.03	0.23	0.02
F0 Mean[a]							
	Baseline	1.03	0.05	0.98	0.03	0.93	0.08
	Target	1.09	0.14	1.01	0.05	1.00	0.13
Intensity Mean[a]							
	Baseline	1.07	0.02	0.99	0.02	0.93	0.02
	Target	1.10	0.02	1.00	0.03	0.95	0.04
F0 Peak[a]							
	Baseline	1.25	0.16	1.16	0.10	1.14	0.10
	Target	1.16	0.11	1.13	0.12	1.28	0.19
Intensity Peak[a]							
	Baseline	1.13	0.01	1.07	0.02	1.04	0.02
	Target	1.15	0.02	1.05	0.05	1.04	0.05
F0 slope[b]							
	Baseline	770.07	210.52	66.57	47.95	272.97	70.64
	Target	775.40	215.31	246.77	98.43	255.28	126.14

Note. [a]Values are ratios. [b]Values are relative differences in Hz/second.

Table 2

Descriptive Statistics of Male Participants by Condition and Acoustic Parameter

		Subject		Verb		Object	
Acoustic Parameter	Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
Duration[a]							
	Baseline	0.14	0.02	0.19	0.02	0.22	0.04
	Target	0.13	0.02	0.20	0.02	0.22	0.03
F0 Mean[a]							
	Baseline	1.05	0.03	0.99	0.03	0.92	0.07
	Target	1.05	0.08	0.99	0.02	0.87	0.20
Intensity Mean[a]							
	Baseline	1.08	0.01	0.99	0.01	0.93	0.03
	Target	1.11	0.01	1.01	0.02	0.94	0.04
F0 Peak[a]							
	Baseline	1.34	0.23	1.14	0.10	1.15	0.16
	Target	1.16	0.05	1.06	0.05	1.35	0.35
Intensity Peak[a]							
	Baseline	1.13	0.02	1.07	0.01	1.03	0.03
	Target	1.16	0.02	1.06	0.03	1.04	0.03
F0 slope[b]							
	Baseline	690.07	107.95	123.36	47.95	159.65	86.02
	Target	648.08	84.05	161.77	98.43	203.79	97.82

Note. [a]Values are ratios. [b]Values are relative differences in Hz/second.

Figure 1, the duration ratio was lower for the subject grammatical unit (-.007) in the target condition, whereas the duration ratio was higher for both the verb (+.002) and object (+.012) grammatical units. As illustrated in Figure 2, the F0 peak ratio was lower for the subject (-.13) and verb grammatical unit (-.06) in the target condition, while the F0 peak ratio was higher in the object grammatical unit (+.13). As shown in Figure 3, the intensity peak ratio was higher for the subject grammatical unit (+.02) and the object grammatical unit of the target condition, while the intensity peak ratio of the verb grammatical unit was lower (-.01).

In addition, the analysis indicated a significant difference in the F0 slope between the male and female speakers, $F(1, 18) = 5.92$, $p = .026$, $\eta^2 = .25$. The female participants had a mean F0 slope within the target words of 413 Hz per second, while the F0 slope within the target words for male participants was 331 Hz per second.

The statistical analysis also indicated significant differences across various grammatical units for all of the dependent measures examined in this study (refer to Table 3 for the specific ANOVA statistics). As previously mentioned, a listing of the mean values of each dependent measure for both the female and male speakers can be found in tables 1 and 2, respectively. All other statistical comparisons were not found to be significant.

Discussion

The aim of the present study was to investigate the ways in which children use F0, duration, and intensity to mark contrastive stress in sentence production. The data used in this study, collected by Dromey (2010), were re-analyzed to examine the emphasis or stress of a target word relative to both a separate baseline production and the neighboring words within the same utterance. In addition, this study analyzed the intensity and F0 of a contrasted word or phrase in terms of not only mean, but also peak and slope values. A statistical analysis was

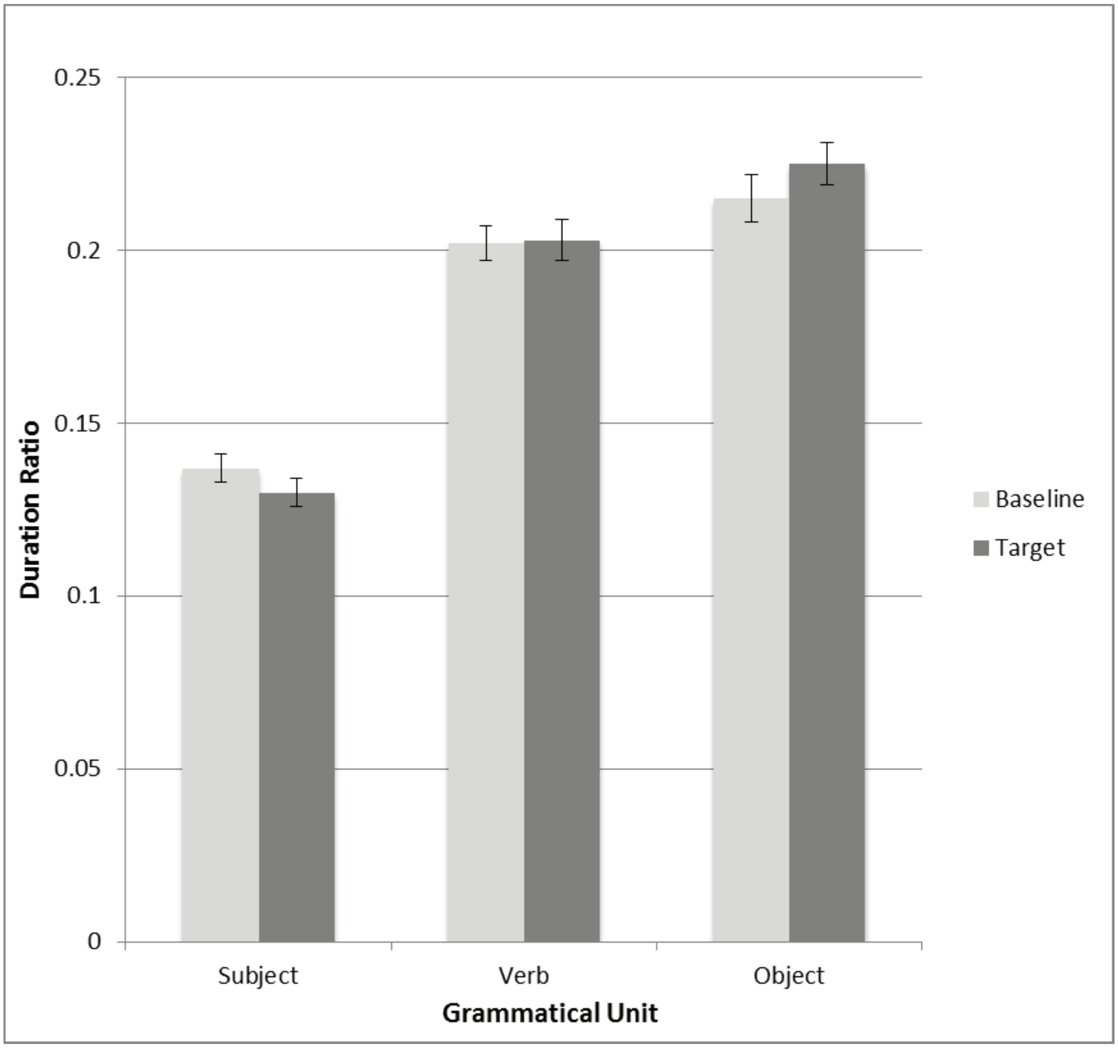

Figure 1. Duration ratio of the speaking condition by grammatical unit for all participants

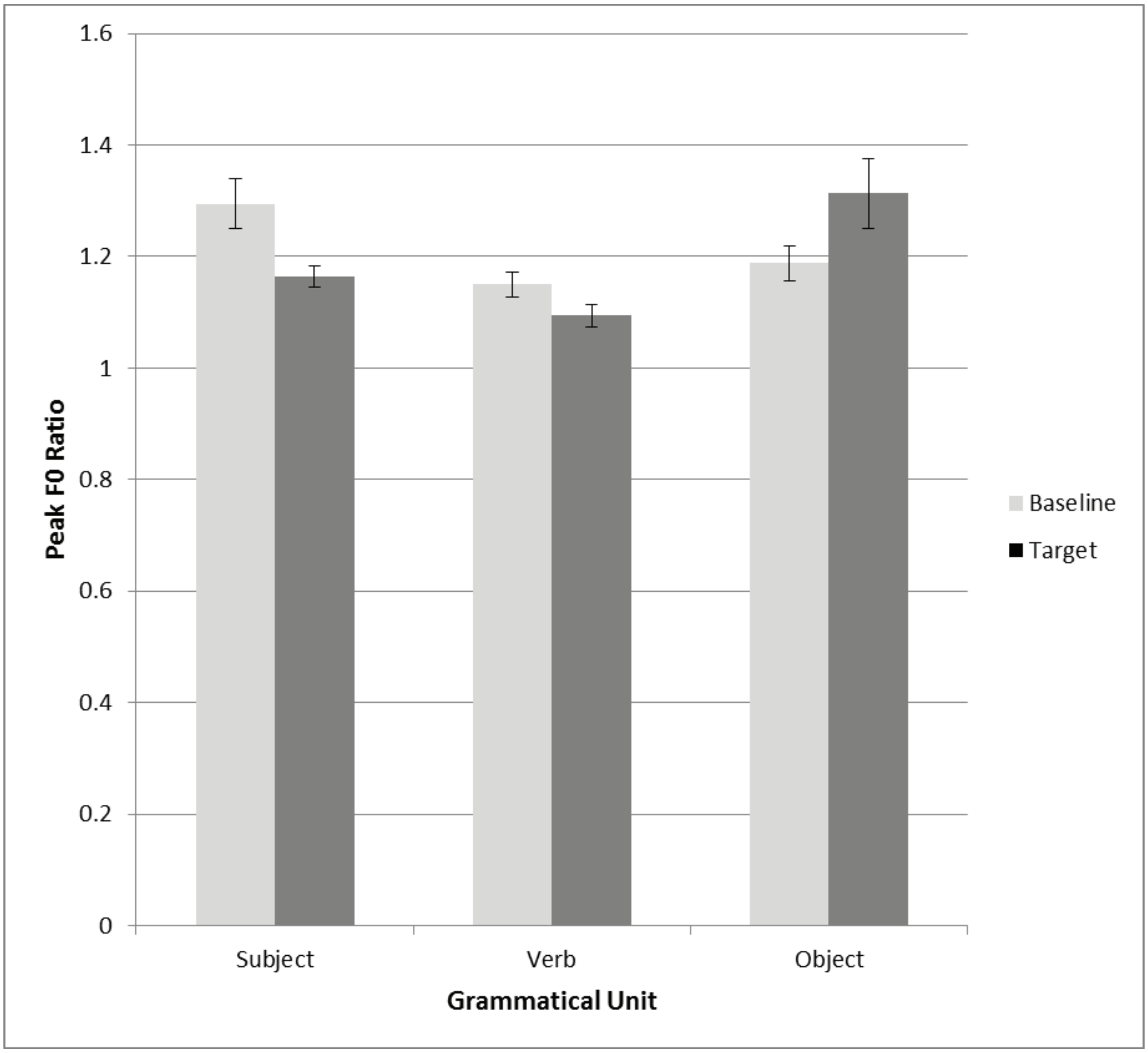

Figure 2. F0 peak ratio of the speaking condition by grammatical unit for all participants

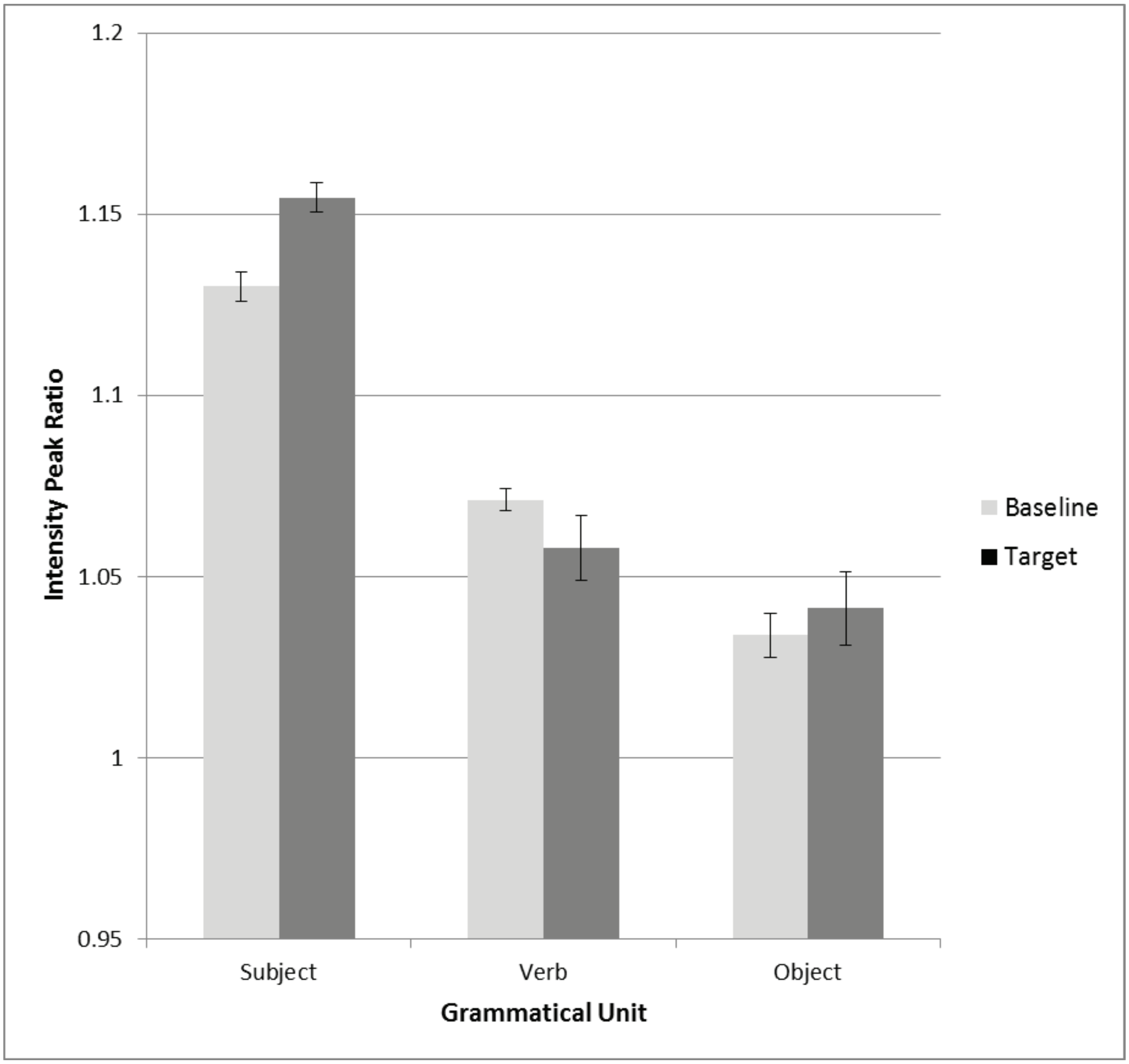

Figure 3. Intensity peak ratio of the speaking condition by grammatical unit for all participants

Table 3

Analysis of Variance Statistics as a Function of the Grammatical Unit

Main Effects	*F*-statistic	*p*-value	Partial Eta Squared
Duration[a]	141.49	< .001	0.89
F0 Mean[a]	30.68	< .001	0.63
Intensity Mean [a]	287.35	< .001	0.94
F0 Peak[a]	5.76	.007	0.24
Intensity Peak[a]	107.83	< .001	0.86
F0 Slope [b]	229.32	< .001	0.93

Note. [a]Values are in ratios. [b]Values are relative differences in Hz/second.

conducted to examine possible differences in the acoustic measures as a function of the speaking condition, speaker gender, and grammatical unit.

Summary of Acoustic Findings

Speaking condition. When comparing the baseline speaking condition to the speaking condition eliciting contrastive stress, significant differences were only found for the acoustic measure of mean intensity. These findings are unlike results reported in previous studies with adult speakers, which have found that differences in contrastive stress are primarily expressed through changes in relative F0, with some secondary differences in intensity and duration (Bolinger, 1978). The results of the present study may indicate that preadolescent children rely on changes in relative intensity to mark contrastive stress more than changes in F0 or duration. While the inferential statistics did not show significant differences in speaking condition for any of the F0 or durational measures, the descriptive data did show trends in the data that may suggest the children are using these acoustic features to mark contrastive stress in a perceptually salient manner.

Gender-related differences. Gender-related differences in how the child speakers marked contrastive stress were only found for the dependent measure of F0 slope. Female speakers produced the target words with a greater degree of F0 slope than the male participants. This finding is supported by previous research conducted with adults indicating that male speakers rely less on changes in F0 to mark contrastive stress when compared to females (Cooper, Eady, & Mueller, 1985). This finding also supports the suggestion by Patel & Grigos (2006) that children may use F0 contours to mark contrastive stress in place of other prosodic markers.

Grammatical unit. In addition, results of the data analysis showed significant acoustic differences between the three grammatical units. This finding was expected, considering that each grammatical unit was at a different, but invariable, location of the sentences. For example, the subject targets were always toward the beginning of the sentence, the verbs in the middle of the sentence, and the objects toward the end of the sentence. Thus, considering inherent acoustic differences in sentential F0 (e.g., pitch declination for declarative sentences), aerodynamic pressures, and sentence final lengthening, it stands to reason that the measures examined in this study would also vary.

Limitations of the Current Work and Directions for Future Research

There are a number of methodological and developmental concerns with the current study that need to be considered.

Lack of adult-like development. The age at which children develop the ability to express contrastive stress in an adult-like manner remains unclear. Although a number of researchers have found that children as young as 6 months of age use prosodic markers to express stress across syllables or single words (Bauman-Waengler, 2009; Cutler & Swinney, 1987; Davis et al., 2000; Galligan, 1987; Snow, 1995; Whiteside & Hodgson, 2000), the results of this study support the findings of a number of other studies that indicate that children probably don't develop adult-like contrastive stress within sentences until after ten years of age (Cruttenden, 1974, 1985; Patel & Brayton, 2009; Wells et al., 2004). According to Cruttenden (1986), the core features of contrastive stress are used both receptively and expressively by young children in early stages of development. However, other more subtle features of intonation, including true contrastive stress in sentence production, cannot be understood or produced in adult-like ways until a child is at least 10 years of age. Considering the age of participants in the current study and the paucity of literature in the area of prosodic development, additional research is needed that examines the acquisition of contrastive stress in children across a wider range of ages and linguistic backgrounds.

Individual speaker differences. It is possible that a significant effect in speaking condition was not found due to individual differences in how each child chose to mark contrastive stress. For example, one child may have used duration to mark contrastive stress, while another may have preferred to use changes in F0 or intensity. Since the data was grouped according to speaker gender during the statistical analysis, it is possible that such differences were undetectable. In addition, it is possible that individual children might have used the same acoustic parameter to mark contrastive stress in an opposing manner. For example, one child may have contrastively marked a word by producing it with a relatively higher mean F0, whereas another child may contrastively mark the same word with lowering of mean F0. Although both children may have marked contrastive stress in a perceptually effective manner, the ability to use contrastive stress would likely not be revealed by a statistical analysis which averages across individual data.

Complex nature of contrastive stress. It is possible that the design of the current study was unable to capture the complex nature of contrastive stress. Prosodic functions like contrastive stress are suprasegmental in nature, with variations being expressed at the syllable, word, or even sentence level. In addition, a wide variety of acoustic cues or combinations of cues can be used to signal emphasis or stress. Not only are there differences in how stress is marked between individual speakers, but also within a speaker depending on the linguistic context and communication environment. As has been found in adults, children may use a type of *cue trading* strategy to express prosodic stress. Because of this complexity, it is difficult to design and execute a study which accounts for all of the possible factors involved. This complexity is accentuated by the fact that the participants in the current study were children with developing speech and language abilities.

Not only are the underlying mechanisms of contrastive stress highly complex, but the perceptual salience of such acoustic cues or combination of cues is not well understood. Although the measurement ratios in the current study were not found to be statistically significantly across speaking condition, it remains unclear if the actual differences would be perceptually significant (e.g., Furrow, 1984). For example, a statistically insignificant difference in intensity ratio of 1.01/1.03, may result in a perceptual *just noticeable difference* to a listener. Thus additional perceptual studies should be conducted to address this question more directly. In a perceptual experiment, listeners may be able to perceive subtle differences in F0, duration, and intensity that are undetectable by a production analysis alone.

Participant training. An additional possibility is that participants were somewhat uncertain of how they were supposed to participate in the experiment. This uncertainty could have led to participants feeling uncomfortable or tentative, causing them to use prosody in unexpected ways. Dromey (2010) stated that the F0 values collected from the participants support the hypothesis that they had some uncertainty, as the children and adults lowered their F0 on the subject, increased their F0 when saying the verb targets, and showed the largest increases in F0 on object targets. If the participants received more training prior to data collection, they may have performed in ways that conformed more closely to what previous research outcomes predicted.

Measurement methodology. The F0 and intensity ratios were calculated based on a comparison between a baseline and target production of the same sentence. This type of measurement calculation was designed to accommodate for the F0 and intensity variations that naturally occur in sentence production. To prompt a minimal degree of word specific emphasis or stress, the baseline sentences were elicited by having the participants read a printed version of the sentences. The target sentences, however, were elicited through a picture description task. Mixing these two types of elicitation may have affected the production patterns of the children in an atypical manner. Thus it may be beneficial to conduct additional research that more closely matches the methods of elicitation.

In addition, the children may have been mimicking the pattern of the administrator rather than producing the sentences in an individual manner. Previous studies have indicated that the order of mastery in the comprehension and production of stress is imitation, comprehension, and then production (Atkinson-King, 1973). It is possible children who participated in the present study had mastered only the imitation component of contrastive stress, thereby affecting the results of the study. The children's speech productions may have also been affected by the redundant nature of the elicitation task. Because one of the functions of prosody is to convey affect, the child's mood would potentially alter the acoustic parameters of his or her speech. Studies have shown that sadness or negative experiences, including boredom, are characterized by a decrease in F0 variability (Snow & Balog, 2002). Thus it may be of value to examine the use of contrastive stress from a fewer number of sentences which are elicited in a more naturalistic manner. This type of design might prevent the children from losing interest and becoming bored with the elicitation task.

Despite the limitations previously mentioned, it is hoped that the findings of this study will promote greater understanding of the development and use of contrastive stress by preadolescent children. In addition, it is anticipated that the methodological insights discussed in this study will facilitate future studies in the area of prosodic development.

References

Atkinson-King, K. (1973). Children's acquisition of phonological stress contrasts. *Working Papers in Phonetics, 25*, 1-161.

Bauman-Waengler, J. (2009). *Introduction to phonetics and phonology: From concepts to transcription.* Boston: Pearson.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences, 17,* 97–110. University of Amsterdam.

Boersma, P., & Weenink, D. (2009). Praat (Version 5.1.20). [Computer Software]. University of Amsterdam.

Bolinger, D. (1978). Intonation across languages. In J. H. Greenberg (Ed.), *Universals of human language: Vol. 2. Phonology* (pp. 471-524). Hillsdale, NJ: Erlbaum.

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America*, 77, 2142-2156.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language, 12,* 643-661.

Cutler, A., & Swinney, D. (1987). Prosody and the development of comprehension. *Journal of Child Language, 14*, 145-167.

Davis, B. L., MacNeilage, P. F., Matyear, C. L., & Powell, J. K. (2000). Prosodic correlates of stress in babbling: An acoustic study. *Child Development, 71*(5), 1258-1270. doi: 10.1111/1467-8624.00227

Dromey, A. (2010). An acoustic and perceptual investigation of contrastive stress in children (Master's thesis). Brigham Young University, Provo, UT.

Furrow, D. (1984). Young children's use of prosody. *Journal of Child Language, 11*, 203-213.

Galligan, R. (1987). Intonation with single words: Purposive and grammatical use. *Journal of Child Language, 14*, 1-21.

Hornby, P. A., & Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research, 13*, 395-399.

Ladd, D. R. (1993). Constraints on the gradient variability of pitch range, or, Pitch level 4 lives! *Phonological Structure and Phonetic Form* (pp. 43-63). Cambridge: Cambridge University Press.

Lehiste, I., & Peterson, G. E. (1959). Vowel amplitude and phonemic stress in American English. *The Journal of the Acoustical Society of America, 31*(4), 428-435.

Patel, R., & Brayton, J. T. (2009). Identifying prosodic contrasts in utterances produced by 4-, 7-, and 11-year-old children. *Journal of Speech, Language and Hearing Research, 52*, 790-801. doi: 10.1044/1092-4388(2008/07-0137)

Patel, R., & Grigos, M. (2006). Acoustic characterization of the Question-Statement Contrast in 4, 7, and 11 year-old children. *Speech Communication, 48*, 1308-1318. doi: 10.1016/j.specom.2006.06.007

Panagos, J., & Prelock, P. A. (1997). Prosodic analysis of child speech. *Topics in Language Disorders, 17*(4), 1-10. doi: 10.1097/00011363-199708000-00003

Snow, D. (1995). Formal regularity of the falling tone in children's early meaningful speech. *Journal of Phonetics, 23,* 387-405. doi: 10.1006/jpho.1995.0030

Wells, B., Peppe, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31*, 749-778. doi: 10.1017/S030500090400652X

Whiteside, S. P., & Hodgson, C. (2000). Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication, 32*, 267-285. doi: 10.1016/S0167-6393(00)00013-3

Annotated Bibliography

Atkinson-King, K. (1973). Children's acquisition of phonological stress contrasts. *Working Papers in Phonetics, 25*, 1-161.

Patterns of stress in speech are used to signal differences of meaning and mark grammatical distinctions. In this study by Atkinson-King, the linguistic abilities of 300 children from the ages of 5 to 13 were studied in terms of the following topics: the acquisition of unemphatic stress patterns, differences between the ability to comprehend and produce stress contrasts, the relationship between syntax, semantics, and phonology in acquisition of stress patterns, the possibility of children's learning stress patterns by lexical item rather than rule, the relative order of acquisition of various stress rules, and the possibility of children creating idiosyncratic rules at any time before mastering the adult rule. Comparison data were also elicited from a group of adults. The author concluded that the correct production of prosodic stress is learned gradually and advances with a child's age throughout preadolescence. Atkinson-King also concluded that the order of mastery in the comprehension and production of stress is imitation, comprehension, and then production.

Bauman-Waengler, J. (2009). *Introduction to Phonetics and Phonology: From Concepts to Transcription.* Boston: Pearson.

This book by Bauman-Waengler provides a comprehensive review of the areas of phonetics and phonology. The publication discusses the relevance of clinical phonetics to the professional practice of individuals who are studying linguistics, communication disorders, education, or other related disciplines. The author provides examples in the text of functional situations in which segmental and suprasegmental speech processes are developed and typically expressed by speakers of American English. Although the majority of the book is devoted to discussing the typical production of sound segments and the associated phonetic transcription, one chapter does review prosodic development and expression. The work only briefly reviews communication disorders.

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America*, 77, 2142-2156.

Contrastive stress was examined in this study to determine its influence on intonational patterns of duration and fundamental frequency (F0). Researchers conducted a series of experiments designed to compare intonational patterns in oral reading with transcriptions of spontaneous speech. When the expected general declination of F0 peaks was not observed, the researchers designed two additional experiments to determine the influence of the location of the stressed word (sentence focus) in the sentence on duration and F0 peak declination. The researchers found that the elongation effects in sentence focus and utterance-final lengthening were found to be less than what would be found in an additive model. Cooper et al. concluded that the results of this study provide support for the notion that a constraint exists on the lengthening of segments in the production of speech.

Beach, C., Katz, W., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *Journal of the Acoustical Society of America, 99*(2), 1148-1160.

The study by Beach, Katz, and Skowronski investigated how children perceptually use prosodic cues to process phrasal interpretation. The authors presented different versions of the phrase *pink and green and white* to a group of child listeners. An analysis of the study's data indicated that all of the participants used duration and intonation cues for phrasal interpretation. The authors further suggest that children use acoustic-prosodic information to interpret syntactic phrases in a way that is very similar to adult listeners.

Bolinger, D. (1978). Intonation across languages. In J. H. Greenberg (Ed.), *Universals of human language: Vol. 2. Phonology* (pp. 471-524). Hillsdale, NJ: Erlbaum.

This book chapter by Bolinger is a comprehensive review of intonation across a number of different languages. The author discusses how intonation in language is made up of a variety of different linguistic elements, including pitch, intonational breaks, registers, intensity, rhythm, accent, tone, and stress. The chapter also reviews the various purposes of intonation, such as forming closures and accents. Bolinger defines pitch in terms of horizontal and vertical intonational breaks. The vertical breaks within an utterance are described as the prominences of accent and are used primarily to focus attention on a specific part of the utterance, whereas horizontal breaks are used for separation. The author also reviews the use of vocal registers (pulse, loft, and modal) during discourse to express emotion. Bolinger defines rhythm in speech as the system of repeating or alternating durations and the rate of succession. Although rhythm often plays a supporting role to other aspects of prosody, it is important in signaling pauses and marking intervals. Accent and tone are also discussed in depth. The author also concludes that stress within a word or phrase is important for the melodic line, which in turn determines the position of stress in an utterance.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language, 12,* 643-661.

In this study, the researcher hypothesized that children at age 10 would exhibit more competence on tones and nucleus placement than in groupings in a comprehension task involving different intonation groupings, nucleus placements and tones. To test this hypothesis, Cruttenden presented 20 ten-year-old children and 20 adults with a series of comprehension tasks. Results showed that intonation comprehension was significantly lower in children compared to the adult listeners. The author concluded that children probably do not interpret the intonation of utterances in a real life setting as consistently as adults. Results also suggested that there was no reason to support any priority of intonation production over comprehension, and that some use of intonation will be acquired early, while more complex uses that involve grammatical and situational context are not acquired until later in development. Cruttenden indicated that developmental models of time that address intonation comprehension were overly simplistic. The author concluded that the age at which prosodic elements are acquired or the order in which such elements are acquired cannot be reliably determined.

Cutler, A., & Swinney, D. (1987). Prosody and the development of comprehension. *Journal of Child Language, 14*, 145-167.

Adults and children use prosody in communication. In children, it has been suggested that production of prosody is more advanced than comprehension. At age 3, children have been proven to use stress to distinguish new from given information and at age 4, children have proven to be able to correctly assign contrastive stress during picture discrimination tasks. Regarding comprehension, infants show sensitivity to prosody in adult speech, which seems to be key in the acquisition of prosody in speech. However,, children under age 6 fell short of adult performance in tasks which measured response time in detecting accented target words. Although previous literature has suggested that children's productive skills in prosody are greater than those in comprehension, this appears not to be the case, as children's prosody is produced without their knowledge of a relationship between prosody and semantics.

Davis, B. L., MacNeilage, P. F., Matyear, C. L., & Powell, J. K. (2000). Prosodic correlates of stress in babbling: An acoustic study. *Child Development, 71*(5), 1258-1270. doi: 10.1111/1467-8624.00227

This study was an investigation into the premise that babbling in infants is *speech-like* as a result of its prosodic components. The researchers elicited speech samples, disyllabic sequences, from five infants and five adult speakers. The samples were acoustically analyzed to detect the use of prosodic correlates of stress, namely relative changes in pitch, intensity, and duration of the utterances. Results of the analysis indicated that the use of all three prosodic correlates was expressed in a similar manner by the adults and children. The researchers concluded that infants produce stress in the same manner as adults during production of single-syllable words and that adult-like stress patterns appear in children before lexical items.

Dromey, A. (2010). An acoustic and perceptual investigation of contrastive stress in children (Master's thesis). Brigham Young University, Provo, UT.

This study examined children's use of prosody to mark contrastive stress by eliciting simple sentences from 20 children who were 8 to 10 years of age. The author elicited speech recordings from the participants using a series of sentences linguistically constructed to allow children to contrastively stress the subject, verb, or object of the sentence (e.g., *The boy is carving the pumpkin*). The resulting speech recordings were then acoustically analyzed with Praat software to examine how or if the children marked the contrastive element of the sentence through relative changes in the F0, intensity, and duration of the target word. These measures were compared to measures extracted from a baseline production of the sentences with no intended emphasis. A comparison set was obtained from a set of data extracted from 10 adult speakers collected in a similar manner by Matthews (2010). The results of the experiment showed that speakers decreased their F0 when emphasizing the subject of the target sentences, yet increased their F0 when emphasizing verb and object parts of the sentence. No consistent changes were observed

with intensity during the tasks, but in general children were observed to shorten the duration of words intended to be contrastively stressed. In part, these results did not follow findings from previous research. Dromey (2010) concluded that this may have been due to the methods of analysis employed in the study.

Furrow, D. (1984). Young children's use of prosody. *Journal of Child Language, 11*, 203-213.

Some research has shown that children use prosody for communicative intent in the same manner as an adult. Other research argues that no apparent consistency exists in children's use of intonation in utterances. To address this issue, the present study used psychophysical judgments of pitch, loudness, and pitch range for each utterance to quantitatively describe prosody while relating these scores to social behaviors such as eye contact and interaction. Results showed that utterances made while making eye contact were louder, higher, and more variably pitched on average, indicating the impact of prosody use on communication.

Galligan, R. (1987). Intonation with single words: Purposive and grammatical use. *Journal of Child Language, 14*, 1-21.

The transition to purposive use of intonation with single words was examined in the study by Galligan. Two children were observed over the course of multiple six-week periods, during which time speech samples were recorded. An analysis of the children's speech samples indicated that one child asked questions using intonation by the age of 1:3 years, while the other child sought informative replies at age 1:6 years. Grammatical distinctions were used by both children at the single-word stage by 1:5 years of age. These results indicated that for the two children involved in the study a transition to the grammatical use of intonation was gradual, with a purposive use of rising tones to indicate the naming of objects.

Hornby, P. A., & Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research, 13*, 395-399.

In an experiment with twenty preschool children, Hornby and Hass investigated the development and use of contrastive stress in describing a referential contrast. The study used two pairs of pictures to elicit speech samples from the children. Results of the study indicated a mastery of contrastive stress patterns, as evidenced by children's tendency to stress the element of the description which corresponded to the contrasting element. This tendency was shown to be learned early through observation, as it was assumed that preschool-age children had not been instructed on the use of contrastive stress through metalanguage.

Ladd, D. R. (1993). Constraints on the gradient variability of pitch range, or, Pitch level 4 lives! *Phonological Structure and Phonetic Form* (pp. 43-63). Cambridge: Cambridge University Press.

In this chapter, the author describes various theories regarding intonational phonology. The author discusses pitch range, specifically describing the theory that a speaker's pitch range can vary in both a paralinguistic and gradient manner, assuming that pitch can be used to vary speech and to emphasize certain words and phrases. Ladd explains that pitch can vary with the degree of emphasis, and that greater emphasis leads to greater height of the accent. A speaker's use of loudness and high or low pitch contours to emphasize part of an utterance are also discussed by the author.

Lehiste, I., & Peterson, G. E. (1959). Vowel amplitude and phonemic stress in American English. *The Journal of the Acoustical Society of America, 31*(4), 428-435.

The article proposes the idea that the physical properties of speech are organized into a symbolic code, which is made up of various components, which can be analyzed at the physiological, acoustical, or perceptual level. The authors explain that in order to obtain correct information about listener judgments of speech, it is most effective to use the acoustical signal. Lehiste and Peterson also describe the interpretation of a speech signal by a listener, which is a complicated set of auditory parameters, including physiological effort, rate of vocal fold vibration, mode of laryngeal vibration, pharyngeal and oral articulation, palatopharyngeal closure, and duration. Regarding the use of stress in speech production, the authors propose a theory that it is the judgments of physiological effort involved in producing vowels that allows a listener to identify stress. These judgments of physiological effort are best represented via relative amplitude of the value.

Patel, R., & Brayton, J. T. (2009). Identifying prosodic contrasts in utterances produced by 4-, 7-, and 11-year-old children. *Journal of Speech, Language and Hearing Research, 52*, 790-801. doi: 10.1044/1092-4388(2008/07-0137)

The present study investigated the acquisition of prosody in children. In order to assess prosody acquisition, the researchers had 36 adult listeners rate linguistic contrasts produced by children 4, 7, and 11 years of age. Listeners rated the children in terms of both question-statement contrasts and sentential contrastive stress. The listeners answered questions regarding the ease of the task and impressions about which acoustic cues the listeners relied on to make their decisions about classification. Listeners had more difficulty classifying questions from statements in the 4-year-old age group, but were able to identify contrastive stress in all 3 age groups. Based on their findings, Patel and Brayton indicated that children's prosodic control may begin to stabilize between the ages of 4 and 7 years.

Patel, R., & Grigos, M. (2006). Acoustic characterization of the Question-Statement Contrast in 4, 7, and 11 year-old children. *Speech Communication, 48*, 1308-1318. doi: 10.1016/j.specom.2006.06.007

The authors examined the question-statement contrast used in children's utterances. The present study examined the use of prosodic cues used by 12 children from age groups of 4, 7, and 11 years of age through elicited productions of declarative statements and questions. 4-year olds were found to be unable to reliably signal questions using rising fundamental frequency, and instead used final syllable duration to mark this contrast. Children 7 years of age used all cues to contrast statements from questions and the oldest group primarily used changes in fundamental frequency. This suggests that an age-related pattern exists across development in the combination of acoustic cues to mark the question-statement contrast.

Panagos, J., & Prelock, P. A. (1997). Prosodic analysis of child speech. *Topics in Language Disorders, 17*(4), 1-10. doi: 10.1097/00011363-199708000-00003

Authors of this article explain that prosody has been described by linguists in terms of tone, stress, rhythm, and pause. This does not, however, account for sudden shifts that occur in the flow of speech, which involve changes in rate, pitch, stress, and emotion during speech. Panagos and Prelock state that in English, content words carry full stress placed on one syllable, and function words receive weak stress in grammatical phrases. In addition to this classifying role, prosody facilitates the organization of language into temporal sequences using rhythm. Rhythm exists in prosody due to the fact that within an utterance, there is an organization of tone groups. They also detail the organization of tone groups, which includes utterances organized by foot, and within the foot, utterances organized in syllables. These tone groups, feet, syllables, and phonemes are the units of language structure. The authors conclude that rhythm, stress, tone, and emotion vary during a speaker's discourse in order to exchange information about the social context, conversational topic, linguistic meaning, and grammar of what is being said, and that in this way, prosody is key to the organization of speech.

Snow, D. (1995). Formal regularity of the falling tone in children's early meaningful speech. *Journal of Phonetics, 23,* 387-405. doi: 10.1006/jpho.1995.0030

The purpose of this study by Snow was to examine if intonation could correspond with the length of an utterance in children, as has been previously found with adult speakers. The author investigated the speech patterns of nine children between 18 and 24 months of age. The utterances of the child participants were analyzed using a schematic continuum, showing variation in pitch as well as the complexity of pitch contours. Results indicated that pitch change did not vary with syllable length of utterances for the young child participants. The author concluded that pitch change in falling tones is acquired by children very early and may be one of the first expressive features of the grammar that they acquire.

Wells, B., Peppe, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31*, 749-778. doi: 10.1017/S030500090400652X

In this study, researchers tested the hypothesis that intonation is used differently in children as they increase in age. Four groups of thirty children with mean ages of 5:6, 8:7, 10:10, and 13:9 were administered a battery of prosodic tasks to investigate their use of comprehension and production of intonation in four communicative areas. The areas assessed included chunking, or prosodic phrasing, affect, interaction, and focus. The data indicated that 5-year old children have acquired many functional prosodic skills. However, prosodic comprehension and other aspects of intonation continue to develop between 5:0 and 8:7 and through age 10:10. These results led to the conclusion that the skill with which children of different ages used prosody varied across children, but tended to correlate with measures of expressive and receptive language development.

Whiteside, S. P., & Hodgson, C. (2000). Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication, 32*, 267-285. doi: 10.1016/S0167-6393(00)00013-3

The purpose of this study by Whiteside and Hodgson was to examine the speech patterns of both children and adults. The researchers used a picture-naming task to elicit a series of speech recordings from male and female children between 6 and 10 years of age. The speech samples from the child participants and a comparison group of adults was subsequently analyzed in terms of formant frequency values, coarticulation, and temporal speech patterns. In both perceptual and motor skills, the researchers found evidence of age and sex differences within the speech patterns. The authors concluded that the development of speech production co-occurs with maturation in the vocal mechanism. During this time, children learn perceptual and motor skills that will develop into coordinated and automated speech systems that are seen in adult speakers.