An Exploratory Study to Examine a Drumming-to-Speech Intervention for Prosody Perception in Preschoolers with Cochlear Implants

Jessica MacLean

University of Miami, maclean.jess@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_theses

Recommended Citation
https://scholarlyrepository.miami.edu/oa_theses/790

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Theses by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
UNIVERSITY OF MIAMI

AN EXPLORATORY STUDY TO EXAMINE A DRUMMING-TO-SPEECH INTERVENTION FOR PROSODY PERCEPTION IN PRESCHOOLERS WITH COCHLEAR IMPLANTS

By

Jessica MacLean

A THESIS

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Master of Music

Coral Gables, Florida

December 2019
UNIVERSITY OF MIAMI

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Music

AN EXPLORATORY STUDY TO EXAMINE A DRUMMING-TO-SPEECH INTERVENTION FOR PROSODY PERCEPTION IN PRESCHOOLERS WITH COCHLEAR IMPLANTS

Jessica MacLean

Approved:

________________         _________________
Kimberly Sena Moore, Ph.D.  Shannon de l’Etoile, Ph.D.
Assistant Professor of Professional Practice, Professor, Music Therapy
Music Therapy

________________        _________________
Ivette Cejas, Ph.D.  Guillermo Prado, Ph.D.
Professor, Otolaryngology  Dean of the Graduate School

________________
Adam Tierney, Ph.D.
Birkbeck College
Children who utilize cochlear implants (CIs) often have trouble detecting and utilizing prosody, an element of spoken language which uses variance in the timing, pitch, and dynamics of speech to communicate meaning beyond semantics. Without a true grasp of prosody, children with CIs can miss conversational elements such as sarcasm and may not communicate effectively with others. Children with CIs match typically-hearing peers in measures of rhythm perception, but fall below their peers in measures of pitch, or melodic, perception. Evidence for behavioral and neural overlaps between prosody and music perception provide rationale for utilizing a drumming intervention to practice and strengthen detection of speech rhythm, which confers benefits for prosody perception. This exploratory study investigated the use of a Drumming-to-Speech (DTS) intervention to practice identifying stressed syllables in speech and nursery rhymes based on previous evidence demonstrating that improvements in speech rhythm perception can lead to improvements in prosody perception. Twelve children between the ages of three and five completed a five-week protocol. The intervention included pre- and post-tests as well as four in-person, student researcher-led sessions which involved drumming on a tubano drum to stressed syllables in repetitive
phrases, nursery rhymes, and songs as well as additional at-home practice in synchronization of speech and drumming. Participants completed Audie (Gordon, 1989), Profiling Elements of Prosody – Children 2015 (PEPS-C, 2015; McCann & Peppé, 2003), and synchronization assessments before and after the intervention in order to determine the intervention’s effects on music and prosody perception, specifically the perception of affective and contrastive stress prosody. Results demonstrated that participants significantly improved in perception of music, both melody and rhythm, and also in emotional prosody, but not in linguistic prosody. These findings support the presence of a strong behavioral and neural overlap between music and emotional prosody perception in synchronizing drum hits with speech rhythm. While participants improved in accuracy and variability of their phase and period synchronization abilities, only the improvements in average phase accuracy during the slow condition reached significance. More research and longer interventions which focus on speech rhythm synchronization may elucidate which aspects of rhythmic information are most important for perceiving speech rhythm and prosody. Additionally, age positively correlated with performance in the emotional prosody subtest at post-test, which may indicate a developmental window that is most appropriate for this intervention. Future recommendations for interventions and research are also addressed, such as increasing intervention dosage and incorporating measures which are sensitive to changes in dependent variables.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LIST OF FIGURES</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>LIST OF TABLES</strong></td>
<td>vi</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Need for the Study</td>
<td>4</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>5</td>
</tr>
<tr>
<td>Purpose Statement</td>
<td>6</td>
</tr>
<tr>
<td>2 RELATED LITERATURE</td>
<td>7</td>
</tr>
<tr>
<td>Prosody</td>
<td>7</td>
</tr>
<tr>
<td>Development of Prosody Perception</td>
<td>12</td>
</tr>
<tr>
<td>Music</td>
<td>16</td>
</tr>
<tr>
<td>Development of Music Perception</td>
<td>18</td>
</tr>
<tr>
<td>Overlaps Between Prosody and Rhythm Perception</td>
<td>20</td>
</tr>
<tr>
<td>Effects of Rhythm on Speech and Prosody Perception</td>
<td>25</td>
</tr>
<tr>
<td>Effects of Music Interventions on Speech and Prosody Perception</td>
<td>30</td>
</tr>
<tr>
<td>Drumming as a Tool for Improving Emotional and Focus Prosody</td>
<td>32</td>
</tr>
<tr>
<td>Importance of Preschool Language Development</td>
<td>33</td>
</tr>
<tr>
<td>Research Questions</td>
<td>33</td>
</tr>
<tr>
<td>3 METHODOLOGY</td>
<td>35</td>
</tr>
<tr>
<td>Design and Variables</td>
<td>35</td>
</tr>
<tr>
<td>Participant Recruitment</td>
<td>35</td>
</tr>
<tr>
<td>Participants</td>
<td>37</td>
</tr>
<tr>
<td>Measures</td>
<td>37</td>
</tr>
<tr>
<td>Drumming-to-Speech Intervention</td>
<td>43</td>
</tr>
<tr>
<td>Materials</td>
<td>45</td>
</tr>
<tr>
<td>Participant Recruitment</td>
<td>46</td>
</tr>
<tr>
<td>Experimental Procedure</td>
<td>49</td>
</tr>
<tr>
<td>Data Analyses</td>
<td>55</td>
</tr>
<tr>
<td>4 RESULTS</td>
<td>57</td>
</tr>
<tr>
<td>Descriptive Results</td>
<td>57</td>
</tr>
<tr>
<td>Research Questions</td>
<td>59</td>
</tr>
<tr>
<td>5 DISCUSSION</td>
<td>65</td>
</tr>
</tbody>
</table>
Effects of Drumming-to-Speech on Prosody Perception............................................ 66
Effects of Drumming-to-Speech on Music Perception.............................................. 71
Relationships between Changes in Music and Prosody Perception....................... 72
Effects of Drumming-to-Speech on Synchronization Ability .................................. 73
Links Between Demographic Variables and Music and Prosody Perception........... 74
Additional Findings and Implications...................................................................... 75
Limitations .................................................................................................................. 79
Conclusions and Future Directions ......................................................................... 80

REFERENCES ........................................................................................................... 82

APPENDICES ............................................................................................................. 92

  A: Recruitment Flyer .......................................................................................... 92
  B: Informational Letter to Parents .................................................................... 93
  C: Phone Recruitment Script ............................................................................ 94
  D: In-Person Recruitment Script ..................................................................... 96
  E: Email Recruitment Script ............................................................................. 98
  F: Demographic Questionnaire ......................................................................... 99
  G: List of Nursery Rhymes ................................................................................ 101
  H: Sample Pictures to Accompany Nursery Rhymes ...................................... 102
  I: Session Plan Script ....................................................................................... 103
  J: Sample At-Home Practice Sheet .................................................................... 107
  K: Sample Script for At-Home Practice Tracks ............................................... 108
  L: Data Collection Sheet ................................................................................... 109
  M: Adult Consent Form ..................................................................................... 111
  N: Child Assent Description and Script ............................................................ 115
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Example of Levels within the Prosody Pyramid</td>
<td>10</td>
</tr>
<tr>
<td>2. Methodology Flow Chart</td>
<td>36</td>
</tr>
<tr>
<td>3. Pre- and Post-Test Score Percentages for Music and Prosody</td>
<td>61</td>
</tr>
<tr>
<td>4. Performance on Emotional Prosody Post-Test by Participant Age</td>
<td>64</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Therapeutic Function of Music</td>
<td>46</td>
</tr>
<tr>
<td>2. Music-Based Intervention Reporting</td>
<td>47</td>
</tr>
<tr>
<td>3. Participant Demographics of Total Sample</td>
<td>58</td>
</tr>
<tr>
<td>4. Intervention Completion of Total Sample</td>
<td>59</td>
</tr>
<tr>
<td>5. Wilcoxon Signed-Rank Non-Parametric Tests for Significance</td>
<td>60</td>
</tr>
<tr>
<td>6. Spearman Correlation Coefficients Between Dependent Variable Changes</td>
<td>61</td>
</tr>
<tr>
<td>7. Wilcoxon Signed-Rank Non-Parametric Tests for Significance</td>
<td>62</td>
</tr>
<tr>
<td>8. Spearman Correlation Coefficients between Demographic Variables and Post-Test Scores for Dependent Variable Changes</td>
<td>63</td>
</tr>
</tbody>
</table>
Chapter One

Introduction

Statement of the Problem

Children with severe to profound hearing loss who receive cochlear implants (CIs) early in life (i.e., before two years of age) are capable of acquiring functional levels of speech and language (Nicholas & Geers, 2007; Niparko et al., 2010). Despite achieving typical proficiency in measures of vocabulary, grammar, and syntax use, children with CIs often perform well below typically-hearing (TH) peers on measures of prosody perception and production (Hopyan-Misaken, Gordon, Dennis, & Papsin, 2009; Kalathottukaren, Purdy, & Ballard, 2015; Meister, Landwehr, Psychny, Walger, & von Wedel, 2009; Peng, Tombling, & Turner, 2008).

Prosody perception and expression play crucial roles in communication through transmittance of nonverbal cues pertaining to the speaker’s emotional and semantic intent. Prosodic information provides the first cue to syntactic structure for infants (Pinker, 1987; Wennerstrom, 2001), and researchers have linked prosodic abilities to emotional development in typically-hearing and hearing-impaired populations (Hopyan-Misakyan et al., 2009). Perception and comprehension of prosodic information have cascading implications for perception and comprehension of other information in the speech stream, such as syntactic organization, as well as implications for social and communicative competence and speech development (Paul et al., 2005; Pinker, 1987; Wennerstrom, 2001). Given the demonstrated developmental importance of prosody, impaired prosody (as is the case for children with CIs) should be a target area for intervention and research.
Prosody is a suprasegmental aspect of speech; in other words, it conveys meaning that cannot be gleaned from semantic understanding alone (Ross, 2000). Also referred to as the “melody of speech” (Wennerstrom, 2001), prosody is characterized by variance in pitch, timing, loudness, and timbre (Aziz-Zadeh, Sheng, & Gheytanchi, 2010). Prosody provides information about the speaker’s affective intent (emotional prosody) and also provides cues to linguistic and semantic structure (linguistic prosody; Plante, Holland, Schmithorst, 2006). The present study will investigate both aspects of prosody, but will narrow the focus of the investigation within linguistic prosody, as it contains many subskills.

Linguistic prosody combines elements of three subcategories: turn-end prosody, chunking prosody, and focus prosody. Chunking prosody includes chunking a thought into segments, such as through segmenting complex noun groups (i.e., the difference between “blue bell” and “bluebell”). Turn-end prosody involves utilizing changes in pitch, timing, and timbre to regulate conversational behavior, such as when pitch goes up at the end of a question. Finally, focus prosody involves understanding and using word emphasis to make specific words or syllables stand out (i.e., I wanted that toy to distinguish the toy in question from another; Stojanovik, 2010). As children with CIs struggle with perception of focus prosody specifically (Most & Peled, 2007), the present study will selectively investigate this aspect of linguistic prosody.

Focus prosody, or stress perception, may rely largely upon fundamental frequency (F0) variance within an utterance (Raphael, Borden, & Harris, 2007; Cruttenden, 1997). However, some argue that for CI recipients, duration and loudness cues may be more reliable (Kochanski, Grabe, Coleman, Rosner, 2005; O’Halpin, 2010) as they are
transmitted through the CI more consistently than frequency cues due to the degradation of frequency information through the CI (Green, Faulkner, & Rosen, 2004; Limb & Roy, 2014). This degradation impairs not only the perception of prosody for CI recipients, but also music (Gfeller, Driscoll, Kenworthy, & Voorst Van, 2011).

Despite pitch processing deficits, young CI users are particularly motivated by music (Trehub, Vongpaisal, & Nakata, 2009). In fact, within the domain of music, children with CIs display a relative strength in rhythm perception compared to their weakened pitch perception. Children with CIs match TH peers in tasks of rhythm perception and production (Hsiao & Gfeller, 2012). Behaviorally, rhythm perception is utilized for both prosody perception and music perception (Hausen, Torppa, Salmela, Vaino, & Särkämö, 2013), and can mediate pitch perception in music (Gfeller et al., 2011; Hsiao, 2008). More specifically, rhythm provides temporal organization that can cue pitch change, one of the most salient aspects of prosody. For example, the familiar rhythmic pattern of “Happy Birthday” helps CI users recognize the song and creates expectations in the listener for the pitch melody to change when a new note occurs (Hsiao & Gfeller, 2012). Additionally, CI recipients use duration (or rhythm) cues to discriminate elements of prosody, such as the difference between a question and a statement (Marx et al., 2015). Therefore, rhythmic training may facilitate pitch perception for prosody.

Rhythmic training is most easily translated through musical training on percussion instruments. Playing percussion (such as a tubano drum) demands accurate perception and production of rhythm. In fact, researchers and clinicians recommend the use of rhythmically-based instrument playing activities for children with CIs because of their
strengths in rhythm perception (Gfeller et al., 2011). Thus, an intervention targeted at drumming to (and therefore perceiving the rhythm in) speech could theoretically improve prosody perception in children with CIs by boosting perception of durational prosodic cues.

**Need for the Study**

**Theoretical relevance.** This study expands understanding of the use of durational cues to improve prosody perception in children with cochlear implants. Evidence that rhythmic training can improve focus and emotional prosody perception suggests a hierarchical role for timing above frequency changes in the speech signal for perceiving prosody (Hausen et al., 2013). Additionally, a better understanding of the links between musical rhythm and prosody perception could improve understanding of music’s therapeutic mechanisms and of the mechanisms underlying prosody perception in the general population as well as in individuals with CIs. These findings support future research directed at constructing more effective interventions.

**Clinical implications.** This study primarily benefits children with CIs who experience poor prosody perception. Prelingually deafened children could better understand others’ prosody, which could in turn improve interactions with others for enhanced social, communicative, and emotional development. The present study may inspire future long-term, clinical studies utilizing rhythmic interventions to improve prosody perception, and may also benefit clinicians through stronger evidence-based practice and successful therapeutic outcomes.
Definition of Terms

**Rhythm.** Rhythm refers to the durational patterns of a stimulus (Grahn, 2012). It is often characterized by stimulus onsets, and the time intervals between those onsets (also referred to as Inter-Stimulus Intervals (ISIs)).

**Prosody.** Prosody refers to the variances in frequency, duration, and amplitude (perceptually, pitch, timing, and loudness) that convey suprasegmental meaning within speech utterances (Aziz-Zadeh, Sheng, & Gheytanchi, 2010; Wennerstrom, 2001). In other words, prosody conveys meaning that is beyond the literal meaning of the utterance’s semantics. Prosody can convey emotional or linguistic information.

**Emotional prosody.** Emotional (or affective) prosody communicates a speaker’s emotional intent (Plante et al., 2006) through variances primarily in pitch, timing, and loudness cues in spoken utterances. For the purposes of the present study, emotional prosody serves to cue happiness, sadness, and anger (Peppé & McCann, 2003).

**Focus prosody.** Focus prosody, a subskill of linguistic prosody, communicates relative importance of syllables or words through the addition of stress (Raphael et al., 2007). Stressed syllables are typically accented by higher pitch, longer duration, and greater intensity than unstressed syllables (Most & Peled, 2007). The use of stress to demarcate the importance of certain syllables provides information about sentence structure and directs the listener’s attention (Plante et al., 2006).

**Drumming-to-speech (DTS).** For the purposes of this study, drumming-to-speech consists of striking a tubano drum once with one hand to coincide with each stressed syllable in an utterance. As the stimuli used in this study are nursery rhymes, stressed syllables alternate with unstressed syllables in order to facilitate aural-motor
rhythmic entrainment (Leong & Goswami, 2014). This skill was practiced in the drumming-to-speech (DTS) intervention.

**Children with CIs.** For the purpose of this study, children with CIs included preschool-aged children three to five years of age with at least one cochlear implant who were prelingually deaf. These children were exposed to English either in the home or school setting, as determined through prescreening questions.

**Purpose Statement**

The purpose of this exploratory study was to examine the effect of a drumming-to-speech (DTS) intervention on the perception of linguistic (focus) and emotional prosody in preschool-aged children with CIs.
Chapter Two

Related Literature

This chapter will provide an overview of research literature regarding the definitions, development, and mechanisms of prosody and music perception. Additionally, this chapter will review the neural underpinnings of emotional and linguistic prosody, speech rhythm, musical rhythm perception, and will demonstrate an overlap in neural resources used for these tasks. Finally, this chapter will argue for using a rhythm-based music intervention to benefit emotional and focus prosody perception for children with CIs.

Prosody

Prosody, or the “melody of speech” (Wennerstrom, 2001), is an aspect of spoken language which provides structure, informs the listener of the speaker’s emotional state, and conveys the relative importance of information within the speech stream. More specifically, prosody conveys suprasegmental information beyond the semantics and syntax of an utterance though variances in acoustical characteristics such as frequency, duration, and amplitude. Perceptually, these characteristics correlate with pitch, timing, and intensity, respectively (Heffner & Sleve, 2015). Prosody also defines hierarchical levels of relative prominence or importance of syllables and words within an utterance (Shattuck-Hufnagel & Turk, 1996).

Characteristics of prosody. Prosody perception and expression play crucial roles in typical speech development by transmitting nonverbal cues that communicate a speaker’s intended emotion and semantic emphasis (Pinker, 1987; Wennerstrom, 2001). Researchers have linked prosodic abilities to social and communicative competence (Paul
et al., 2005), and emotional development (Hopyan-Misakyan et al., 2009). Acoustic features of prosody include pitch (melody, intonation), timing (speed, pauses, cadence), timbre (quality of sound), and loudness (Aziz-Zadeh et al., 2010; Ross, 2000).

**Pitch.** Pitch refers to the perceptual experience of high and low sounds. Tone and intonation in speech are the patterns of pitch variance within an utterance that contrast and organize words (Beckman & Venditti, 2010). Emphasis on certain syllables, or stress, is thought largely to rely upon variance in fundamental frequency (F0), though pitch variance is typically also accompanied by an increase in amplitude (Raphael et al., 2007).

**Timing.** Timing refers to the relative durations of speech sounds and pauses within an utterance. When perceiving stress, timing seems to predominate over loudness as the second most important cue that humans utilize, with pitch being first (Raphael et al., 2007). Timing within speech is determined by measuring acoustic intervals that correspond to phoneme- or syllable-level units (Fletcher, 2010). Often, stressed syllables are longer than unstressed syllables, differing by as much as 30 to more than 70 ms (Crystal & House, 1988). Speech rhythm refers to the alternation of strong and weak elements that are hierarchically structured based upon relative durations (Allen, 1975; Fletcher, 2010). Overall, awareness of timing is based upon knowledge of a language’s syllables, stress patterns, and speech rhythm structure (Gilbert, 2008).

**Loudness.** Loudness is the perceptual correlate of speech amplitude. Fluctuations in loudness determine relative importance of syllables within the speech stream; for example, stressed syllables are often spoken with greater intensity than unstressed syllables (Most & Peled, 2007). Loudness is important for prosodic perception, as
improving loudness can improve prosodic production, even in some clinical populations (Watson & Hughes, 2006).

**Prosodic structure.** Prosody determines suprasegmental phonological structure and relative prominence of items within an utterance (Fletcher, 2010). In other words, prosody creates a hierarchy of syllables within a sentence that help the listener determine the most important syllable of the speech stream. In this way, the speaker utilizes prosodic variance through fluctuations in pitch, timing, and loudness in order to parse out salient parts of the speech stream for the listener.

The concept of a “Prosody Pyramid” (Gilbert, 2008) illustrates the structural organization of prosody. The base of the pyramid is “Thought Group,” which consists of a clause, phrase, or short sentence. The next level of the pyramid is “Focus Word,” which is the most important word of the thought group; focusing on this word is largely accomplished by contrasting it with the rest of the sentence. The next level of the pyramid is “Stress,” which occurs in one syllable of the focus word. Stress is achieved through lengthened vowel duration, increased vowel clarity, pitch change, or increased volume. The final level of the pyramid is “Peak,” as the stressed syllable functions as the peak of information for the thought group. This syllable must be easily recognized and clearly spoken. Figure 1 illustrates these levels in the sentence “How do you spell ‘easy’?”
Prosodic functions and neural mechanisms. Broadly speaking, prosody functions to convey meaning about language structure (linguistic prosody), including word segmentation, syntactic structure, and emphasis (Cutler, Dahan, & van Donselaar, 1997), as well as emotional state or affective intent (emotional prosody; Witteman et al., 2012). While the extant literature demonstrates a neural overlap between mechanisms utilized for processing both linguistic and emotional prosody (Wildgruber, Hertrich, Riecker, Erb, Andres, Grodd, & Ackermann, 2004), lateralization effects between both types exist (Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). While both hemispheres are utilized when perceiving linguistic and emotional prosody, the right hemisphere seems more specialized for emotional prosody perception (Witteman, Ijzendoorn, van de Velde, van Heuven, & Schiller, 2011), particularly in the lateral temporal lobes (Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003). Additionally, emotional prosody cues may receive temporal precedence in processing relative to

Figure 1. Example of levels within the Prosody Pyramid. Adapted from Teaching Pronunciation: Using the Prosody Pyramid p. 20), by J. Gilbert, 2008, New York: Cambridge University Press. Copyright 2008 by Cambridge University Press.
linguistic prosody; in other words, emotional prosodic information is processed before linguistic prosodic information (Paulmann, Jessen, & Kotz, 2012).

**Emotional prosody.** Emotional prosody uses variances in pitch, timing, and loudness to communicate the emotional state of the speaker (Scherer, 2003). This type of prosody recruits a bilateral, temporofrontal network that includes the auditory cortices and inferior frontal gyrus pars orbitalis, or Brodmann area 47 (Belyk & Brown, 2013; Hausen et al., 2013; Witteman et al., 2012). While short, non-verbal vocalizations may be processed initially by the amygdala via a fast route (Liebenthal, Silbersweig, & Stern, 2016), emotional prosody over longer timescales, such as in language processing, likely passes first through a slower cortical route, including the bilateral temporofrontal cortices (Witteman et al., 2012), before reaching the amygdala.

**Linguistic prosody.** Linguistic prosody uses variances in pitch, timing, and loudness to clarify semantic meaning. In effect, linguistic prosody is the aural equivalent of commas, colons, periods, and question marks in speech (Ross, 2000). Like emotional prosody, linguistic prosody also recruits right auditory areas of the brain, as well as the inferior frontal gyrus pars opercularis (Belyk & Brown, 2013). However, explicit processing of linguistic prosody distinctly recruits left hemisphere language areas, such as Broca’s area (Wildgruber et al., 2006). Linguistic prosody includes the following specific skills:

- **Turn-end.** Turn-end is the ability to understand and use prosody to regulate conversational behavior. For example, a downward inflection at the end of an utterance indicates a speaker is finished (Stojanovik, 2010).
• **Chunking.** Chunking describes the ability to use and understand prosody to segment complex noun groups. This type of prosody allows individuals to distinguish, for example, between the complex noun groups “bluebell” and “blue bell” (Stojanovik, 2010).

• **Focus.** Focus is the ability to understand or use prosody to emphasize specific words or syllables. Emphasis is often accomplished through the addition of “stress” to a certain sound within an utterance to give that sound importance. Often, the stressed syllable contains a higher pitch (fundamental frequency), longer duration, and greater intensity (amplitude) than unstressed syllables (Raphael et al., 2007; Most & Peled, 2007). As an example, think of the difference between saying “I’ll take the blue sock,” implying the choice of a blue sock over another colored sock and “I’ll take the blue sock,” implying the choice of a blue sock over another blue item. The current project will investigate focus prosody specifically.

**Development of Prosody Perception**

The development of prosody perception differs between typically-hearing (TH) children and children with cochlear implants (Most & Frank, 1994). Cochlear implants (CIs) are prosthetic devices that bypass damaged acoustic hearing for individuals who are severely to profoundly deaf. They work by electrically stimulating the cochlear nerve through implanted electrodes (Limb & Roy, 2014). Analyzing differences in developmental trajectories between TH children and children with CIs can provide suggestions for intervention.
Typical development of prosody. Infants learn about pitch, timing, and loudness before they are able to segment words. Through “prosodic bootstrapping” (Jusczyk, 1997), segmental boundaries in speech are marked off, which allows infants to extract underlying syntactic organization of an utterance at a later stage in development. Furthermore, infant-directed speech contains exaggerated prosodic characteristics that assist in the development of language acquisition by incorporating elements such as wide pitch ranges, higher pitches, more frequent rising intonation, longer pauses, and slower articulation rates (Cruttenden, 1994).

Typical development of emotional prosody. By five years of age, typically-developing children are able to understand, but not yet produce, emotional prosody. Production of emotional prosody is largely mastered by age eight (Wells et al., 2004); however, it can also be observed on a basic level in infants. That is, adults can reliably distinguish emotional prosody from non-linguistic vocalizations of infants, even those from a non-native culture. These early, non-culture-specific meanings in vocalizations suggest universality for typical development of certain aspects of vocal emotional prosody perception and production that is independent of gestures or other visual cues (Kerseken, Zuberbühler, & Gomez, 2017).

Typical development of focus prosody. Generally, the production of focus prosody is thought to precede understanding of focus prosody. For example, a child might say, “I wanted that toy,” to indicate desire of a specific toy, but may then have difficulty identifying the word that is emphasized. Use of focus prosody is largely established by the age of five and develops in tangent with grammatical comprehension and production (Wells et al., 2004). Processing of focus prosody is therefore not yet fully
developed in preschoolers (Wells, Peppé, & Goulandris, 2004). Focus prosody identifies words of importance within the speech stream, and while it largely relies upon perception of pitch information (Raphael, Borden, & Harris, 2007; Cruttenden, 1997), some information that communicates focus prosody extends to nearby words (i.e., words surrounding the word which receives focus also may have slightly raised pitches).

In contrast, lexical stress is a term used to identify specific syllables of importance within the speech stream (Segal, Houston, & Kishon-Rabin, 2016). While children from ages four to six are sensitive to stress, they cannot yet use stress location within a sentence to aid in processing meaning. Some research suggests that the ability to manipulate the position of a main accent in an utterance (focus prosody) may develop when multiword utterances begin to occur between two and three years of age (Crystal, 1987; Oller, Wieman, Doyle, & Ross, 1976), while others suggest that ten-year-olds are significantly worse than adults at perceiving focus prosody (Cruttenden, 1985).

**Development of prosody in children with CIs.** Children who are prelingually deaf and who receive a cochlear implant are younger in hearing age than their TH peers. Though length of deafness varies, children with CIs are automatically delayed in speech development compared to TH peers; for example, babbling is delayed for children who are prelingually deafened (Koopmans-van Beinum, Clement, & Den Dikkenberg-Pot, 2001). Though profoundly deaf infants do not differ significantly from TH infants in frequency (F0) production, they fall significantly behind in articulation, duration, and syllable production due to a lack of auditory feedback (Koopmans-van Beinum et al., 2001).
Prosodic development is also delayed, though children with hearing loss may be able to produce appropriate intonations between the ages of six and nine (Most & Frank, 1994). In addition, school-aged children with cochlear implants show a preference for trochaic syllables, in which the first syllable of a syllable pair is stressed, which mirrors preferences that children exhibit earlier in development. Therefore, promoting the development of appropriate rhythmic speech patterns for children with CIs is important (Titterington, Henry, Kramer, Toner, & Stevenson, 2006).

However, because the CI compresses spectral information, F0 and pitch are not readily available to the listener. These degraded pitch cues may impair CI users’ perception of emotional or focus prosody (Green et al., 2004; Limb & Roy, 2014). Despite this limitation, CI recipients can distinguish large F0 differences of over half of an octave, though everyday speech typically contains variances smaller than half an octave, with infant-directed speech being an exception (Golinkoff, Can, Soderstrom, Hirsh-Pasek, 2015). Interestingly, some studies suggest that children with CIs can produce stress and intonation contours even if they cannot yet perceive them. In these instances, when the speaker reaches a point in an utterance of semantic interest, increased muscular tension results in a physiological response that increases F0 (O’Halpin, 2010).

Instead of pitch, CI recipients likely rely on timing and loudness cues to infer prosodic meaning. Timing may provide a more important cue than loudness for the perception of stress and intonation (O’Halpin, 2010). The majority of children with CIs can only perceive amplitude changes greater than those found in everyday speech, implying that duration cues may be more valuable (Hegarty & Faulkner, 2013), though
other studies demonstrate the importance of amplitude for identifying sentence stress (Meister, Landwehr, Pyschny, Wagner, & Walger, 2011).

**Development of emotional prosody for children with CIs.** Though school-aged children with CIs demonstrate the ability to identify emotions through facial affect, they do not identify emotion in speech as well as TH peers (Hopyan-Misaken et al., 2009). Because children who are prelingually deaf rely upon restricted input from the CI to develop auditory areas of the brain, restricted input potentially impedes development of brain areas typically utilized for emotional prosody perception. In this way, the lack of expected auditory input may alter development of auditory processing areas, resulting in cortical reorganization (Gordon, Papsin, & Harrison, 2006).

**Development of focus prosody for children with CIs.** CI users’ processing of focus prosody is less efficient than that of TH peers. In one study, adolescents with CIs reacted significantly more slowly to target phonemes than did TH peers. Even with contextual semantic cues, CI recipients were slower to respond to stressed phonemes than their TH peers (Holt, Demuth, & Yuen, 2016). Additionally, though infants with CIs can detect lexical stress (Segal, Houston, & Kishon-Rabin, 2016), they fall below TH peers (Most & Peled, 2007). Deficits in the development of both emotional and focus prosody could be addressed through practice identifying meaningful auditory cues in a modality that resembles the complexity of speech sounds—music.

**Music**

Music shares similar characteristics and structural properties with prosody (Heffner & Slevc, 2015; Lerdahl & Jackendoff, 1983; Schön, Magne, & Besson, 2004). Music consists of meaningful fluctuations in pitch, timing, and loudness and is
hierarchically organized (much like the “Prosody Pyramid”). Furthermore, musical training may confer benefits related to processing prosody.

**Pitch.** Similar to prosodic pitch, musical pitch is the perceptual correlate of fundamental frequency, though pitch perception is also influenced by other acoustical elements, including amplitude (i.e., a pitch might sound slightly different depending on whether it is played loudly or softly). A sequence of pitches is referred to as a melody, and the overall shape and directions of changes in pitch is referred to as contour.

**Timing.** In music, the relative duration of musical utterances and rests is referred to as rhythm. Rhythm provides temporal organization that can cue pitch change. For example, the familiar rhythmic sequence of “Happy Birthday” can facilitate song recognition and expectation for a new melodic pitch when the next note in the melodic rhythm occurs (Hsiao & Gfeller, 2012). Within music, the organization of rhythm into a hierarchical structure is referred to as meter. Listeners expect acoustic events to occur more often on strong beats than on weak beats, indicating perceptual implications for metrical organization (Honing, Ladinig, Háden, & Winkler, 2009).

**Synchronization.** An aspect of timing, synchronization, also referred to as “rhythmic entrainment,” refers to the ability to temporally match movements or internal oscillators to external oscillators that occur at regular intervals. The temporal frequency between stimulus occurrences (such as 600 milliseconds between metronome stimuli at 100 beats per minute) is referred to as the “periodicity” of the stimulus. Stimulus responses (such as taps on a drum) that occur at the same time as the external oscillator are referred to as being “in phase,” whereas stimulus responses which occur before or after the external oscillator onset are referred to as being “out of phase” (Grahn, 2012).
The isochronous time intervals inherent in musical stimuli lend themselves to synchronization because of the regularity and predictability of information. The isochronous beats act as anchors for synchronization, facilitating entrainment with the music stimulus in a way that is difficult for speech, which does not contain isochronous intervals.

**Loudness.** Similar to loudness in speech, loudness in music is the perceptual correlate of the amplitude of the sound signal. Loudness is also often referred to as intensity. In music, amplitude fluctuations are referred to as dynamics. Often, strong beats in music are perceived to be louder than weak beats (Repp, 2010).

**Musical training.** Research indicates musical training may benefit skills involved in prosodic perception and production. For example, children who have received musical training are better than non-musically trained peers in perceiving word boundaries, an aspect of linguistic prosody (François, Chobert, Besson, & Schön, 2013). In addition, adults with musical training have enhanced emotional prosody processing (Thompson, Schellenberg, & Husain, 2004), while adults with amusia exhibit deficits in this area (Patel, Peretz, Tramo, & Labreque, 1998). Thus, because of characteristic similarities and potential transfer of benefit from music training to speech perception, music training could be utilized as a way to strengthen prosody perception.

**Development of Music Perception**

Both prosody and music perception develop via similar trajectories early in life (Brandt, Slevc, & Gebrarian, 2012). Infant-directed speech and infant-directed singing both contain exaggerated prosody (increased pitch, timing and loudness variance), and may offer mechanisms through which children learn prosody (Golinkoff et al., 2015).
**Typical development of music perception.** Early childhood musical development includes the growth of musical skills, such as pitch perception, rhythmic movements, and emotional responses to music (Gooding & Standley, 2011). By age four, the typically-developing child can synchronize movement to a beat (McAuley, Jones, Holub, Johnston, & Miller, 2006), though children as young as 2.5 years of age can entrain to a regular auditory pulse with a social partner (Kirschner & Tomasello, 2009). By age five, children are able to identify pitch changes as proceeding either “up” or “down,” even when the pitch change is less than a semitone (Stalinski, Schellenberg, & Trehub, 2008). Children five to six years of age distinguish between sections of music through melody-independent elements, particularly loudness (Schwarzer, 1997).

**Development of music perception for children with CIs.** Young CI users are particularly motivated by music, and often prefer musical activities over activities without music (Trehub et al., 2009). CI users are significantly less accurate than TH individuals in tasks of musical pitch perception, possibly due to decreased availability of spectral information through the CI (Hsiao & Gfeller, 2012). However, children with CIs display a relative strength in rhythm perception, even matching TH peers in rhythm perception. Rhythm may also help to mediate pitch perception for children with CIs, such as with the song “Happy Birthday” or “Jingle Bells” (Hsiao & Gfeller, 2012). Activities at which children with CIs may be successful include recognition and discrimination of rhythmic patterns or adapting to tempo changes.

While rhythm may help pitch perception for CI recipients, changes in loudness level can be disruptive. In one study, participants compared a reference pitch for which the intensity remained constant to a second tone that varied in intensity, but had the same
pitch. Most of the participants with CIs incorrectly perceived a change in pitch when stimulus intensity changed (Arnoldner, Kaider, & Hamzavi, 2006).

**Overlaps Between Prosody and Rhythm Perception**

Both prosody and music rely on the perception of pitch, timing, and loudness cues (Heffner & Slevc, 2015). Because of degraded pitch processing through the CI (Limb & Roy, 2014), and because research suggests that interventions for children with CIs should focus on strength areas such as rhythm perception (O’Halpin, 2010), the current review will focus on the behavioral and neural overlaps between prosody and rhythm perception.

**Behavioral overlaps.** CI recipients use timing cues to discriminate prosodic elements, suggesting that rhythm could play a role in improving prosody (Marx et al., 2015). Speech rhythm perception is critical to the perception of prosody, as prosodic changes rely upon durational variance in speech (Hausen et al., 2013). In a study investigating the relationship between music and prosody perception, healthy adults completed tasks of pitch, beat, and harmony perception as well as a task of word stress perception in order to discover potential behavioral overlaps between music and prosody perception. Results revealed that rhythm predicted most of the relationship between prosody and music perception. These findings support a strong perceptual overlap between music and prosody perception that is mediated by rhythm, particularly timing and stress. These findings support shared cognitive resources underlying music and prosody perception that largely depend on rhythm perception.

**Neural overlaps.** Overlaps between the neural resources utilized for the perception of rhythm and prosody support development of a rationale for improving rhythm perception to enhance prosody perception. In other words, training that
strengthens areas utilized for rhythm perception should confer benefits for prosody perception due to the use of shared neural resources.

Central nervous system structures utilized for rhythm perception. Grahn (2012) reviewed then-current research on timing mechanisms, entrainment models, and the neuroscience of rhythm. Results indicated similar neural recruitment for rhythm perception and production in the cerebellum, basal ganglia, parietal cortex, prefrontal cortex, premotor cortex, and supplemental motor area; however, neural recruitment may differ based on the rhythmic skill in question. For example, whereas the cerebellum is involved in “automatic” or subconscious timing at the subsecond level, such as what occurs during synchronization to a metronome, the basal ganglia are involved in “cognitively-controlled” timing at the suprasecond level, such as what occurs when drumming along with a repeated rhythmic pattern. Differences in location of automatic and cognitively-controlled timing mechanisms suggests differences between subsecond, or small-scale, and suprasecond, or large-scale, rhythm perception (Grahn, 2012).

While Grahn (2012) provided an overview of the neural substrates of rhythm, Thaut, Trimarchi, and Parsons (2014) investigated specific neural underpinnings of three subcategories of rhythm: rhythmic patterns, meter, and tempo. Healthy, adult participants (five musicians, five non-musicians) each completed discrimination tasks of rhythmic pattern, meter, tempo, and melody. Researchers adjusted the complexity of each task to match the participant’s music experience so that musicians and non-musicians attained comparable levels of accuracy on the tasks. Participants completed each discrimination task while undergoing positron emission tomography (PET) scans.
Results demonstrated that all three rhythmic subcategories involved the motor cortex, right posterior cerebellum, and bilateral areas of the frontal, cingulate, parietal, prefrontal, and temporal cortices, including the primary auditory cortex. In addition, each rhythmic element distinctly recruited other areas of the brain. Rhythmic pattern perception recruited right cortical auditory processing areas, meter perception the right prefrontal cortex and inferior frontal gyrus, and tempo perception the posterior insula and postcentral gyrus.

The authors suggested that meter perception involves more cognitive and abstract (i.e., global) processing, whereas pattern perception involves different auditory processing areas and is processed locally. Tempo, on the other hand, seems to recruit areas utilized for somatosensory and premotor processing. Despite the small sample size and varying levels of stimuli among participants, the authors interpret their findings as support for a neural network for rhythmic processing that uses differential processing patterns depending on the stimulus requirements (Thaut, et al., 2014). Justifying a music intervention to impact prosody perception requires identification of an overlap between neural areas utilized for rhythm perception and those involved in prosody perception.

Central Nervous System structures utilized for prosodic perception. Because speech rhythm is involved in the perception of prosody, understanding the neural networks involved in speech rhythm perception contributes to the understanding of prosodic neural circuitry. In a study investigating speech rhythm processing (Geiser et al., 2008), healthy adults listened for either rhythm or prosody within sentences consisting of fake words. Participants who listened to the speech rhythm discriminated between “isochronous” and “non-isochronous” sentences, based on whether or not the speaker had
spoken along with a metronome; this task constituted explicit rhythm processing. Other participants listened for prosody and discriminated between questions and statements; this task constituted implicit rhythmic processing, as the participant focused on prosody but also perceived rhythm. All participants listened to the sentences while undergoing functional magnetic resonance imaging (fMRI) scans (Geiser et al., 2008).

Comparisons between tasks demonstrated greater activation in the supplemental motor area (SMA), especially in the right hemisphere, for the explicit rhythm task. This activation extended into the cingulate gyrus and was also present in the insulae, right basal ganglia, and right inferior frontal gyrus. A direct contrast analysis between isochronous and nonisochronous sentences revealed an interaction with explicit and implicit rhythm processing. Isochronous speech, when explicitly processed, recruited the right post-superior temporal gyrus, supramarginal gyrus, and parietal operculum. Implicit processing of isochronous speech recruited the left post superior temporal gyrus, supramarginal gyrus, and parietal operculum. These results imply a role for the SMA, insula, and right inferior frontal gyrus for explicit processing of acoustic patterns, and also imply a role for the right auditory cortex in explicit processing of isochronous speech. As these roles arose during explicit tasks in which the participants were actively listening to rhythm, these areas are also involved in top-down processing of suprasegmental cues in speech (Geiser et al., 2008). These findings should be considered when developing speech rhythm tasks, as different brain areas are recruited depending upon whether or not rhythm is being explicitly or implicitly processed.

Other researchers have explored the neural correlates underlying the perception and production of certain prosody types. Witteman, Van Heuven, and Schiller (2012)
performed a quantitative meta-analysis of the neuroimaging literature on emotional prosody perception to determine which neural areas were most robustly activated across protocols. The researchers performed two meta-analyses to compare brain activation in similar areas across studies that utilize similar tasks. One meta-analysis looked at tasks in which participants actively extracted acoustic cues of emotional prosody (stimulus-driven processing), while the other looked at studies in which the tasks did not require explicit attention toward acoustic characteristics (task-driven processing).

For both meta-analyses, the authors found convergence in activation for a bilateral, temporofrontal network for the perception of emotional prosody. This network included the primary auditory cortex, mid-superior temporal gyrus (STG), and inferior frontal gyrus (IFG). The authors did not find the hemispheric lateralization previously found in lesion studies, suggesting that lateralization in the auditory cortex for processing prosody might be modulated by non-prosodic, acoustic aspects of the speech signal (Witteman et al., 2012). This study provides a useful glimpse into the differences in processing between stimuli with and without prosody and tasks that require focus on prosody or on other stimulus features.

Similarly, Belyk and Brown (2013) performed meta-analyses of neuroimaging studies involving emotional and linguistic prosody perception in order to determine the differences between neural networks for both types of prosody. Meta-analyses revealed that emotional prosody recruited the right post-STG, SMA, IFG, cerebellum, and middle frontal gyrus, as well as limbic areas such as the left amygdala, ventral anterior insula and ventral putamen. The authors also found that linguistic prosody tended to recruit neural areas typically associated with speech and language functions. Across studies, linguistic
prosody processing recruited bilateral IFG, post-STG, supramarginal gyrus, middle frontal gyrus, right SMA, IFG pars orbitalis, primary auditory cortex, and left caudate nucleus, as well as the bilateral insula and cerebellum. In summary, while linguistic and emotional prosody processing overlap in neural substrates such as right hemispheric auditory areas, emotional prosody tended to recruit more limbic areas while linguistic prosody recruited areas more typically associated with speech and language functioning.

**Neural overlaps between rhythmic and prosodic perception.** The neuroimaging literature of rhythmic and prosodic perception demonstrates neural overlaps between rhythmic and prosodic perception networks that involve the cerebellum, basal ganglia, SMA, bilateral auditory processing areas (such as the primary auditory cortex), and right inferior frontal gyrus (Belyk & Brown, 2013; Geiser et al., 2008; Grahn, 2012; Thaut et al., 2014; Wittemen et al., 2012). Rhythmic pattern and metrical perception in particular recruit right cortical auditory processing areas and right IFG (Thaut et al., 2014), areas also recruited for both emotional and linguistic prosody perception (Belyk & Brown, 2013). This overlap provides a rationale for utilizing rhythm training that focuses on rhythmic patterns and meter (i.e., stressed syllables) within speech to elicit a functional change in prosody perception for children with CIs.

**Effects of Rhythm on Speech and Prosody Perception**

To date, published research which explores the effect of drumming to speech in TH children or children with CIs does not exist; thus, this review focuses on research with adults. Before investigating the effect of drumming to speech on prosody perception, this review will investigate the effect of music in general on prosody. Patel’s (2011, 2012, 2014) OPERA hypothesis states that because music perception and speech
perception share overlapping neural networks, and because music places higher sensory and cognitive demands on these shared neural underpinnings than does speech, musical training can drive neuroplasticity to strengthen speech perception. In other words, music training expands auditory capabilities beyond that required for speech, resulting in improved speech processing.

Each letter of the OPERA acronym represents one of five mechanisms by which music training can benefit speech.

1. “Overlap” describes the shared neural networks used to perceive and/or produce the acoustical features of music and language. Because music and speech perception rely upon similar areas of the brain, music skills can strengthen speech skills.

2. “Precision” states that musical features such as pitch and rhythm require more precise processing than speech features, and therefore place higher demands upon neural networks.

3. “Emotion” holds that music which activates emotion networks in the brain will elicit emotions (e.g., happiness), and therefore strengthen overall neural processing by recruiting the additional neural networks implicated in emotion processing. Furthermore, as an inherently rewarding stimulus, music is associated with dopamine release, which results in pleasure and promotes neural plasticity.

4. “Repetition” posits that musical components need to be repeated in order to maximize linguistic benefit. In other words, the neural plasticity that begins
through emotional connections and pleasure is most robust when a stimulus occurs repeatedly, as occurs in music training.

5. Finally, “Attention” requires that the music engage frontal networks associated with focused attention. Attention facilitates learning in other areas such as language development; thus, engagement of attention networks supports learning of speech perception and production.

According to the OPERA hypothesis, if all five conditions are met, cross-domain neuroplasticity will drive the shared networks of speech and music processing to function with greater precision than is needed for speech perception alone, which in turn strengthens speech processing (Patel, 2011, 2012).

In an expanded version of the OPERA hypothesis, Patel (2014) offers preliminary confirmation for using music training to enhance speech perception in CI users. Patel shares results from two adult participants who completed a protocol based on prior research with CI users (Galvin et al., 2007). The participants, both non-musician CI recipients, received 10 hours of piano training over the course of one month in which they learned to play simple, five-note patterns with differing contours (i.e., rising, falling, rising-falling). Participants underwent several tests before and after training, including a prosodic, statement-question discrimination task (Chatterjee & Peng, 2008) and a melodic contour identification task (MCI; Galvin et al., 2007, 2008). After training, both participants improved on the MCI task, but differed in their speech perception results. One participant did not improve in prosodic perception while the other participant did improve. While these data are preliminary and may have been affected by the
participants’ already-typical speech perception during pre-testing, they also suggest that simple music training for non-musician CI users can improve prosody perception.

In another study, Falk and Dalla Bella (2016) examined the effect of aligning finger taps with speech rhythm. Native-German speaking, typically-hearing young adults listened for a target verb in pairs of German sentences that occurred on either a metrically strong or weak position. A metronome cued participants to begin tapping before the sentence began, and lined up temporally with either the strong syllables or the weak ones. Participants then indicated whether they detected a change in the target word after listening and tapping. A second group of participants identified changes in target words with a metronome in the background but did not tap along. In terms of significant findings, results indicated an interaction between alignment and task, whereby the tapping task resulted in better performance and faster detection rates for congruent alignment. In other words, changes that occurred in metrically-strong conditions were detected more reliably, more accurately, and more quickly than those in weak positions in both tapping and cueing tasks.

These results are the first to indicate that tapping along with syllables confers benefits for speech perception. The authors posited that the aligned motor rhythm may build temporal expectancies which strengthen attention at expected times in speech. The authors also argued that building metrical expectancies is important for the early interpretation of prosodic and semantic features of speech which may guide further processing (Falk & Dalla Bella, 2016). Further research needs to be done with typically-developing and hearing-impaired children in order to determine if this same effect could be seen in a younger population, or a clinical one.
While Falk and Dalla Bella (2016) found an enhancement for verbal processing when aligned with motor rhythms, Cason, Astésano, and Schön (2015) investigated effects following rhythmic priming on verbal processing in healthy, non-musician adults being trained to identify target syllables in speech. Participants in an auditory-only group heard rhythmic primes consisting of a short rhythm played by electric drum sounds before hearing sentences. An auditory-motor group received an additional period of auditory-motor training with the rhythmic primes in which they copied the prime rhythms and imitated the pattern of strong and weak beats through tapping.

During the experimental task, participants indicated whether they perceived a target vowel in the final syllable of the sentence. Sentences either matched or mismatched the prosodic (metrical) structure of the priming cue in terms of the number of beats, number of syllables, and stress patterns. From these two variables, the researchers devised four conditions: (a) number and stress match, (b) number match, stress mismatch, (c) number mismatch, partial (half of the sentence) stress mismatch, and (d) number mismatch, full stress mismatch.

Results indicated that participants performed significantly better on condition one, when the number and stress of the priming cue matched the prosodic structure of the sentence. Thus, the researchers argued that the rhythmic primes which matched the prosody of the sentence facilitated metrical expectation, and therefore facilitated speech processing. Additionally, training increased this effect, perhaps due to an enhanced sensitivity to metrical patterns (Cason et al., 2015). These findings demonstrate the ability of prosodic-meter priming to benefit speech processing, and thus support a paradigm that uses rhythm to facilitate prosody perception.
Summary. Research supports the use of music, specifically rhythm, as mechanisms for improving prosody perception. The OPERA hypothesis posits that music training can improve speech and language skills because of the higher cognitive and sensory demands it places on the brain in areas that overlap with speech perception and production (Patel, 2011). Tapping to metrically strong syllables can enhance speech processing and therefore may support the metrical expectancies needed to interpret prosodic and semantic features of speech (Falk & Dalla Bella, 2016). Additionally, rhythmic priming can improve speech perception when the primes match the prosody of the sentence, and this effect can be increased after auditory-motor training (Cason et al., 2015). Taken together, these studies support the use of a rhythm-based paradigm to isolate and improve speech rhythm perception in order to improve prosody perception.

Effects of Music Interventions on Speech and Prosody Perception

To date, published research does not exist that explores the impact of a music therapy intervention aimed at detecting speech rhythm and prosody (e.g. Rhythmic Speech Cueing or Developmental Speech and Language Training through Music) with children who have CIs. Furthermore, more research needs to be done to understand the perception of both prosody and rhythm in prelingually deaf children and the immediate effect that motor alignment with speech has on prosody perception.

At a more general level, research demonstrates effects of music training on prosody perception in TH populations and CI recipients. For example, TH adults with music training perform better than non-musically trained adults in emotional prosody identification (Thompson et al., 2004). Clinically, preliminary evidence shows that a
simple piano training paradigm focused on learning melodic contours can improve prosody perception in adult CI users (Patel, 2014).

Another study found that after a melodic contour training program, adult CI recipients improved in prosodic perception, specifically in their ability to discriminate questions from sentences (Lo, McMahon, Looi, & Thompson, 2015). The training consisted of five-note sequences that were formed into nine contour patterns; difficulty was manipulated either by changing duration or interval size. No significant differences existed between methods of difficulty modulation, suggesting that both methods of melodic contour training conferred benefits for speech perception (Lo et al., 2015).

In a different study, researchers created and implemented a musical ear training program with recent CI recipients in which participants received education in identifying and understanding melody and harmony. In addition to improving in musical perception measures, adult participants experienced quicker improvements in emotional prosody recognition than did participants in a non-musical control group (Petersen, Mortensen, Hansen, & Vuust, 2012). Following these results, a program aimed at refining perception of musical elements could benefit emotional prosody perception.

Music training that includes rhythm-focused lessons as well as learning broader musical concepts may also benefit prosodic perception for younger populations. For example, TH children who received one year of keyboard training outperformed those who did not receive music lessons in identifying emotional prosody (Thompson et al., 2004). In another study, after six months of individualized piano lessons, children with CIs improved perception of emotional speech prosody as well as discrimination of contour and rhythm and memory for melodies (Good, Gordon, Papsin, Nespoli, Hopyan,
Peretz, & Russo, 2017). Taken together, these studies indicate benefits for utilizing a music training program focused on rhythm to benefit prosody perception.

**Drumming as a Tool for Improving Emotional and Focus Prosody**

Prosody, the “melody of speech” (Wennerstrom, 2001), largely relies upon variances in pitch, timing, and timbre to organize speech, emphasize important words, and convey speaker intent (Aziz-Zadeh et al., 2010; Gilbert, 2008; Ross, 2000). Music also consists of variances in pitch, timing, and timbre, among other elements. The affordances and limitations of CIs affect both domains; CI users struggle with pitch processing in music and speech (Green et al., 2004; Limb & Roy, 2014), though they are able to utilize timing cues to identify aspects of prosody and music (Hsiao & Gfeller, 2012; O’Halpin, 2010). In fact, duration may be the most valuable cue to prosodic stress for prelingually deafened children with CIs (Hegarty & Faulkner, 2013). Training paradigms centered on rhythmic cues in speech (such as through drumming) may be able to strengthen prosody perception in children with CIs (O’Halpin, 2010), particularly since music training seemed to improve the perception of pitch (Chen, Chuang, McMahon, Hsieh, Tung, & Li, 2010) and prosody (Good et al., 2017; Thomson et al., 2014) for children with CIs.

Playing a percussion instrument requires consistent and accurate rhythm perception and production and is one method through which individuals can focus on and practice rhythmic abilities. Researchers and clinicians recommend the use of rhythmically-based instrument playing activities for children with CIs because of the children’s relative strength in rhythm perception (Gfeller et al., 2011). Thus, an intervention targeted at drumming to speech would demand attention to speech rhythm.
Drumming to repeated sentences would facilitate practice of prosodic timing at the syllable level, thereby fostering attention to durational cues that assist in prosody perception. Practice of timing cues would also assist in forming temporal hierarchies of speech that are important in understanding focus and emotional prosody (Gilbert, 2008). As timing cues are important for prosody perception for children with CIs (O’Halpin, 2010), and because rhythm may mediate the perception of pitch (Hausen et al., 2013), a drumming to speech paradigm could theoretically improve prosody perception for children with CIs through strengthening perception of durational and intensity contrasts that are characteristic features of prosody.

**Importance of Preschool Language Development**

The greatest gains in auditory habilitation occur within the first three years of cochlear implant use. As implantation typically occurs before 18 to 24 months of age, the preschool age range is a crucial period in language development for children with CIs (Fryauf-Bertschy, Tyler, Kelseay, & Gantz, 1992). Furthermore, prosody perception provides a foundation for phonological, pre-literary skills, which also develop during the preschool years (Lonigan, Burgess, & Anthony, 2000). Therefore, interventions targeting prosody perception in preschoolers may benefit suprasegmental speech perception and measures of reading ability, and so support holistic language development (Calet, Gutiérrez-Palma, Simpson, González-Trujillo, & Defior, 2015).

**Research Questions**

The purpose of this study was to examine the effect of a drumming-to-speech (DTS) intervention on the emotional and focus prosody perception of preschool-aged children with CIs. This is an exploratory study aimed at examining effects of the DTS
intervention and determine potential for further investigation. The following research questions were addressed:

1. What is the effect of drumming-to-speech on emotional and focus prosody perception in preschool-aged children with cochlear implants?
2. What is the effect of drumming-to-speech on music perception in preschool-aged children with cochlear implants?
3. How are changes in music perception related to changes in prosody perception in preschool-aged children with cochlear implants?
4. What is the effect of drumming-to-speech on synchronization abilities?
5. What is the relationship between demographic variables (e.g. age, gender) and benefit from drumming-to-speech in preschool-aged children with cochlear implants?
Chapter Three

Methodology

This chapter delineates the method utilized for this study. Participant inclusion and exclusion criteria, recruitment parameters, and measurements are described below. See Figure 1 for a depiction of the methodology.

Design and Variables

This exploratory study utilized a single group, pre-test/post-test design. The independent variable was the drumming-to-speech (DTS) intervention, and the dependent variables were music perception (rhythm and melody), synchronization ability, and prosody perception, both emotional and focus prosody. As this was an exploratory study, there was no control condition utilized for comparison; instead, participants’ pre-test scores served as a baseline for determining effects of the intervention.

Participant Recruitment

Upon obtaining approval from the University of Miami Human Subject Research Office and Institutional Review Board (#20180559), the student researcher recruited participants through email and flyer distributions and in-person visits to the Debbie School and the University of Miami Ear Institute, both sites which serve children with cochlear implants in Miami, Florida. The student researcher provided a recruitment flier (Appendix A) and informational letter (Appendix B) to send home to parents at both sites. The student researcher also recruited at family night events at the Barton G. Kids Hear Now Foundation, where she utilized a verbal script when discussing the study with parents (Appendix C). The student researcher used an email script to follow-up with parents met at both sites and to reach out to potential families found through contact lists.
provided by affiliates at the Barton G. Kids Hear Now Foundation (Appendix D). Participants’ legal guardians gave written, informed consent prior to study enrollment, and participants gave assent prior to each intervention session following the process described in Appendix N. The student researcher provided all recruitment materials in English and Spanish to accommodate parents’ language preferences.

**Figure 2.** Methodology flow chart. This figure illustrates the sequence of steps in this investigation following Consolidated Standards of Reporting Trials guidelines for social and psychological interventions (Grant et al., 2018).
Participants

Twelve children between the ages of three and five years (six female, \( M = 4.6 \) years old, \( SD = .899 \)) completed the study. Two other children participated in the first three sessions, but then did not provide verbal assent to further sessions and were removed from the study. Figure 2 provides an overview of study recruitment and enrollment.

Inclusion criteria. To meet inclusion criteria, participants’ parents had to respond yes to three prescreening questions. First, participants had to be prelingually, severely to profoundly deaf children aged three to five years who used at least one CI. Participants may have utilized bimodal hearing (one CI and one hearing aid) or two CIs. Diagnosis was confirmed by documentation or by demographic questionnaire. Participants’ parents also had to respond yes to the participant having exposure to English and to understanding simple directions in English, though this did not have to be the participant’s primary language.

Measures

The student researcher measured focus and emotional prosody perception through the contrastive stress and emotion subtests, respectively, of the Profiling Elements of Prosody in Speech-Children 2015 (PEPS-C 2015; Peppé & McCann, 2003) and measured rhythmic and melodic perception through the Audie test (Gordon, 1979). The student researcher administered both the PEPS-C and Audie through computer speakers positioned in front of the participant. The student researcher also assessed synchronization abilities through a social drumming paradigm (Kirschner & Tomasello, 2009).
Profiling Elements of Prosody in Speech-Children (PEPS-C) 2015. The PEPS-C 2015 (Peppé & McCann, 2003) is a computerized assessment of expressive and receptive prosody skills. The test is suitable for all ages above four years of age, but has a pre-test vocabulary check to ensure that participant responses (including those of younger participants) are not confounded by vocabulary knowledge. The PEPS-C 2015 consists of 14 subtests which assess distinct aspects of receptive and expressive prosody skills. Though this test is not standardized, normative data are available from the publisher. The PEPS-C demonstrates test-retest reliability ($r = .85$), as consecutive scores from administrations differing by six months do not differ significantly (Kalathouttukaren, Purdy, & Ballard, 2015), but does not have enough psychometric data available with typical and clinical populations to determine validity.

Each subtest contains two training items, two practice items, and 14 test items formatted as two-alternative, forced-choice tasks. Responses on practice items and test items are scored out of 16. In order to “pass,” participants must answer at least 12 items correctly. Those who score below 12 received a “fail” rating. The present study utilized the Contrastive Stress Understanding subtest to assess focus prosody perception, and the Affect Understanding subtest to assess emotional prosody perception, as well as a preliminary vocabulary check to ensure that vocabulary exposure would not impact scores. Each of these tests take approximately five minutes to administer.

Vocabulary check. This subtest ensures that the participants have been exposed to all test items prior to testing, and that the participant hears the name of each item paired with the picture. During this subtest, the participant sees each potential test item
displayed on the screen, and hears the administrator say the name of this picture that was used in the subtests.

**Affect understanding.** The affect understanding subtest assesses the participants’ ability to detect the emotional intent of a speaker. During this subtest, the participant sees a food picture displayed on the screen and hears a female speaker saying the name of the food with either positive or negative affect. The participant receives instructions that say the person is telling them whether they like or dislike certain foods. During training items, the child is instructed to point to the smiling face when the person sounds like she likes the food, and to point to the frowning face when the person sounds like she does not like the food. The administrator utilizes a mouse to click the face selected by the participant, and the software records responses.

**Contrastive stress understanding.** The contrastive stress subtest assesses the participants’ ability to detect the word that is emphasized in an utterance, or focus prosody. During this subtest, participants must select the color which a female speaker emphasized in her sentence. Instructions state that the speaker is telling them which color sock she forgot to buy at the store, so the participant must point to the color of sock the speaker forgot. For example, the female speaker says, “I wanted BLUE and black socks,” and out of the two pictures, blue and black, the participant must point to the color blue on the screen. Conversely, when the speaker says, “I wanted blue and BLACK socks,” the participant must point to the color black. The administrator utilizes a mouse to click on the color that the participant selects, and the software records responses.

**Audie.** Audie (Gordon, 1989) is a test of developmental musical aptitude intended for children three to four years of age. The test is structured as a game in which children
are asked to verbally indicate whether they heard Audie’s “special song” (a descending D major triad), or whether they heard something different. Participants verbally respond “yes” for Audie’s song and “no” for a different song, and the administrator (i.e., the student researcher) transcribes the response on a provided slip of paper. The melodic and rhythmic subtests each contain ten test stimuli which vary by one pitch or one rhythm, respectively. Participants hear the answer to each question on the recording after it is presented. The Audie test in total takes about 15 minutes to administer.

Gordon (1989), in the Audie test manual, reports Kuder-Richardson reliability of .69 for the rhythm subtest and .68 for the melodic subtest. Taggart (1994) reports predictive validity to be .59 for the rhythm subtest and .41 for the melody subtest, and also suggests there is evidence of inverse and diagnostic validity. Audie has previously been used with preschool-aged children with cochlear implants to assess musical aptitude (Petersen et al., 2012).

**Synchronization assessment.** The synchronization assessment determined the participants’ ability to entrain movements with auditory cues provided by the student researcher in a social context, following an earlier paradigm introduced by Kirschner and Tomasello (2009). During the synchronization activities, the participant and student researcher each played a tubano drum. The participant’s drum contained a drum trigger sensor which recorded each drum hit and each presentation of the metronome stimuli simultaneously for later comparison. The fast condition contains an interstimulus interval (ISI) of 400 ms (or 150 beats per minute), and the slow condition contains an ISI of 600 ms (or 100 beats per minute). The student researcher listened to the metronomic stimulus for two trials of 20 seconds each of the fast and slow ISIs through one ear and drummed
at the same time as the metronome stimulus. The student researcher instructed the
participant to hit their drum at the same time she hit hers, so that the participant
synchronized to the student researcher’s drumming. The social presentation of the
metronomic stimulus provided the participant with an audiovisual version of the
metronome stimulus, which Kirshner and Tomasello (2009) demonstrated is more
appropriate for preschoolers than an auditory-alone presentation.

The administrator recorded all participant drumming for this portion of the
session through Audacity, and later analyzed the data through a computerized analysis
conducted in customized scripts in MATLAB (Mathworks, Inc., Natick, MA). Analyses
determined the amount of synchrony between the participant’s drumming and the
administrator’s drumming under the assumption that the latter’s drumming did not vary
significantly from trial to trial. This information helped to determine whether or not the
participant was able to synchronize to the rhythmic stimulus as measured by the
following, recorded during both fast and slow trials: offset accuracy, offset variability,
tempo accuracy, tempo variability, average phase, and phase consistency.

**Offset accuracy.** Offset accuracy is a measure of the participant’s ability to lock
into the phase of the administrator’s drumming hits. The program pairs and determines
the absolute value of the distance between participant drum hits and the stimulus
(metronome) events in milliseconds. A smaller onset value means the participant hit the
drum closer in time to the metronome and was therefore more accurate.

**Offset variability.** Offset variability is a measure of the overall consistency of the
temporal distance between participant drum hits and the metronome events across each
condition in milliseconds. A smaller value in variability means that the participant was
more consistent in the placement of drum hits in relation to the stimulus, suggesting
better phase locking.

**Tempo accuracy.** Tempo accuracy is a measure of the deviation between inter-
response intervals between participant drum hits from inter-stimulus intervals (ISI) of the
stimulus in milliseconds. During the fast condition, the metronome ISI was 400 ms, while
the slow condition utilized ISIs of 600 ms.

**Tempo variability.** Tempo variability is a measure of the overall consistency of
participant inter-response intervals in relation to metronome ISIs in milliseconds. A lower
value of variability indicates that the participant was more consistent in the durational
intervals between drum hits.

**Average phase.** Similar to offset accuracy, average phase is a measure of
temporal distance between participant drumming and stimulus (metronome) onsets.
However, whereas offset accuracy is given as a measure of milliseconds in temporal
distance, average phase is scaled by ISI and measured in radians to provide a picture of
where in time the participant is typically responding in relation to the periodic stimulus.
Like offset accuracy, absolute values of average phase are reported in the present study.
Drum hits which occur at the exact same time as the metronome are given a value of 0
radians, and drum hits which occur further away in time from stimulus events receive
larger values.

**Phase consistency.** Phase consistency is calculated via circular statistics and
represents the overall magnitude, from 0 to 1, of consistency in the timing of drum hits in
relation to paired metronome events for a given participant. The calculation involves
determining the vector of strength for all phase deviations performed by the participant,
and calculates an overall vector length of consistency. Higher values indicate better consistency.

**Drumming-to-Speech Intervention**

The Drumming-to-Speech (DTS) intervention is a novel rhythmic experience that involves aligning drum hits with stressed syllables in live speech samples, nursery rhymes, and song, with the intention of improving rhythm perception. Here, the intervention focuses on improving rhythm perception to confer benefits for prosody perception, a demonstrated clinical need area for children with CIs, (e.g., Hopyan-Misaken et al., 2009). The speech consisted of spoken phrases and nursery rhymes which alternated between stressed and unstressed syllables to facilitate rhythmic entrainment (Leong & Goswami, 2014). Participants who participated in the DTS intervention drummed to entire poems during the sessions and during at-home practice while receiving feedback and assistance. This intervention incorporated four nursery rhymes of four lines each (see Appendix F); two poems featured trochaic stress patterns (strong-weak-strong-weak, etc.) while two others featured iambic stress patterns (weak-strong-weak-strong, etc.). The student researcher selected poems based upon age-appropriateness of the nursery rhymes and upon alternation between stressed and unstressed syllables in the rhyme structure.

Within the domain of Neurologic Music Therapy, this intervention applies elements of Rhythmic Speech Cueing (RSC) in that the participants align drum hits with speech syllables (Mainka & Mallien, 2014). However, this intervention in other ways resembles Developmental Speech and Language Training through Music (DSLM), as participants did not speak along with the auditory cue, and the intervention is utilizing a
developmentally-appropriate rhythmic task to facilitate development of a linguistic ability related to understanding (prosody perception) for children with CIs who have not yet acquired this skill (LaGasse, 2014).

**Therapeutic Function of Music.** The Therapeutic Function of Music (TFM) is a conceptual framework for articulating the purpose and applications of specific musical elements to a treatment goal based upon prior evidence (Hanson-Abromeit, 2015). This framework ensures the consistent application of musical elements for desired outcomes and provides specific information about how and why a certain musical element should be used (or not used). TFM analysis results in Table 1 informed the creation of the rhythmic-speech stimuli to improve focus and emotional prosody perception in preschool-aged children with CIs for the DTS intervention.

**Synthesis of the music.** Following the element purposes and descriptions identified in Table 1, neither melody nor harmony were incorporated in this intervention, as they can detract from speech perception. The primary focus of this intervention was rhythm, as this element can facilitate perception of prosody, pitch (or in speech, intonation), and speech rhythm. Playing a simple rhythm on an instrument can provide additional sensory cues (tactile, visual, and auditory) to facilitate rhythmic perception, and the timbral difference between a drum allow for distinction between the two sound sources. As slight intensity differences occur between stressed and unstressed syllables, dynamics helped to facilitate identification of stressed syllables, but were kept within a comfortable range for each child. If the child indicated that a volume was too loud or that the child could not hear the sound, the student researcher modified the volume appropriately. The tempo of the exercises remained at a steady, moderate-slow speed.
through the use of a metronome set at 90 beats per minute (bpm), which helped to facilitate temporal processing and entrainment. The lyrics and timbre were produced in a natural style to provide ecological validity and facilitate transfer, so the student researcher recited the nursery rhymes in a speaking voice with exaggerated prosody, such as one might use when reading a story to a preschooler.

**Music-based intervention reporting.** Following guidelines set by Robb, Burns, and Carpenter (2011), reporting for music-based interventions should follow certain criteria. Providing information about the theory, content, delivery schedule, interventionist, treatment fidelity, setting, and unit of delivery helps to improve replication and implementation of music-based interventions. These criteria can be found in Table 2.

**Materials**

**Tests.** The student researcher administered PEPS-C (affect and contrastive stress understanding subtests) and Audie as pre-test and post-test with scripts, and with the order of subtests randomized across participants. The student researcher administered PEPS-C on a laptop which recorded responses through the program, and administered Audie through a CD player while the recording participant responses on separate sheets for each participant and test.

**Sentences.** The student researcher, a female, native American English speaker, presented each nursery rhyme to the participants live, reciting the poems to a metronome at a syllable rate of 1.5 Hz (roughly 90 bpm) based upon recommendations from the TFM analysis (see Table 1). Live speech samples were used because recorded voice is more difficult for CI users to process, and live speech also follows ecological validity and
<table>
<thead>
<tr>
<th>Musical Element</th>
<th>Theoretical Framework</th>
<th>Element Purpose</th>
<th>Element Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Melody</strong></td>
<td>Children with Cls are significantly less accurate than their NE peers in tasks involving pitch and melodic perception (Griesser et al., 2011).</td>
<td>N/A</td>
<td>Melodic perception is difficult for children with Cls, and therefore, this study did not utilize melody.</td>
</tr>
<tr>
<td><strong>Rhythm</strong></td>
<td>Children with Cls perform rhythm tasks comparably to NE peers when not observed by other means (Griesser et al., 2011; Enia &amp; Griesser, 2012).</td>
<td>To facilitate perception of perceived pitch, To facilitate perception of melody, To facilitate practice of speech rhythm identification.</td>
<td>Perceived rhythm with Cls can play simple rhythmic patterns. Use of an instrument (such as a drum) can further aid rhythmic processing by providing visual, auditory, and tactile cues. Rhythm should be utilized to facilitate processing of melody.</td>
</tr>
<tr>
<td><strong>Harmony</strong></td>
<td>Background noise can impede perception of lyrics (Griesser et al., 2011).</td>
<td>N/A</td>
<td>Harmony should not detract from lyric perception, and therefore was not used in this study.</td>
</tr>
<tr>
<td><strong>Form</strong></td>
<td>An background noise can impede perception, the therapist should start songs a cappella, such as with lyrics only, and then add instruments (Griesser et al., 2011). Form should be simple (Griesser et al., 2011).</td>
<td>To facilitate practice of prospective perception.</td>
<td>Activities should begin with words first, and then add harmonic elements such as rhythm and instrument play. The form should be simple.</td>
</tr>
<tr>
<td><strong>Dynamics</strong></td>
<td>Loud volume can distort vocal quality or create discomfort for children with Cls (Griesser, 2000, 2011). Intensity differences between verses and mainstem melodies provide available vocalic contrast throughout the CI (O'Hanlon, 2010).</td>
<td>To facilitate perception of tonal stress.</td>
<td>Volume should be kept within comfortable range for the child, however, slight volume differences can facilitate perception of melody.</td>
</tr>
<tr>
<td><strong>Tempo</strong></td>
<td>Rhythmic patterns should be played at a moderate tempo (Enia &amp; Griesser, 2012). Slower tempos, such as around 100 beats per minute (bpm), can facilitate sound processing for children with Cls (Griesser et al., 2011).</td>
<td>To facilitate proactive perception. Tempo should be moderate to slow (around 100 bpm) to facilitate accurate perception.</td>
<td></td>
</tr>
<tr>
<td><strong>Timbre</strong></td>
<td>Children with Cls may not be able to accurately discriminate between instrumental sounds (Enia &amp; Griesser, 2012). Children with Cls can better discriminate between volume if they are very similar (voice vs. drum; Griesser et al., 2011).</td>
<td>To distinguish between speech and instrument play. Timbre utilized should be very different in order to facilitate discrimination.</td>
<td></td>
</tr>
<tr>
<td><strong>Style</strong></td>
<td>Instrument playing can be structured to facilitate speech and language development (Griesser et al., 2011). Nursery rhymes form a strong rhythmic beat, and therefore facilitate enunciation (Lewis &amp; Georger, 2004).</td>
<td>To provide ecological validity. Nursery rhymes should be utilized to facilitate speech rhythm (prospect) perception; additionally, instrument playing can facilitate speech and language development.</td>
<td></td>
</tr>
<tr>
<td><strong>Lyrics</strong></td>
<td>Songs lyrics are faithfully transmitted through the CI, and can be recognized by children with Cls (Griesser et al., 2011). Lyrics should be appropriate to current development (Vasquez, Tredin, &amp; Schellenberg, 2006). Timing can alter prosody, timing, and production of lyrics; therefore, the therapist should sing with a &quot;natural&quot; vocal quality and pace song with visual aids (Griesser et al., 2011).</td>
<td>To facilitate proactive perception. Lyrics should be developmentally appropriate and produced with natural vocal quality and visual aids.</td>
<td></td>
</tr>
</tbody>
</table>
Table 2

**Music-Based Intervention Reporting**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Specific Information for DTS Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention theory</td>
<td>Due to neural overlaps between rhythmic and prosodic perception, a drumming to speech paradigm could strengthen speech rhythm perception to benefit linguistic and emotional prosody perception (see Chapter Two of this manuscript for a more in-depth review)</td>
</tr>
<tr>
<td>Intervention content</td>
<td></td>
</tr>
<tr>
<td>Person selecting music</td>
<td>The student researcher selected the music elements (i.e., rhythm) based upon the TFM analysis (see Table 1).</td>
</tr>
<tr>
<td>Music</td>
<td>The music consisted of simple rhythmic patterns played on a tubano drum to align with the stressed syllables of speech.</td>
</tr>
<tr>
<td>Music delivery method</td>
<td>The music was practiced live with the student researcher in intervention sessions, and was practiced with a parent/guardian at home through a recording.</td>
</tr>
<tr>
<td>Intervention Materials</td>
<td>DTS sessions and at-home practice required pre-recorded speech stimuli and a computer or phone to play stimuli.</td>
</tr>
<tr>
<td>Intervention strategies</td>
<td>The DTS intervention utilized music skills involved in beat extraction, rhythm perception, and auditory-motor integration.</td>
</tr>
<tr>
<td>Specific Information for DTS Intervention</td>
<td>The DTS intervention were administered in four, 20-minute weekly sessions. The participant was also asked to complete at least five, five-minute practice sessions at home each week, for a total of 20 practice sessions across four weeks.</td>
</tr>
<tr>
<td>Interventionist</td>
<td>The student researcher, who is receiving masters-level training in music therapy, implemented the DTS intervention.</td>
</tr>
<tr>
<td>Treatment fidelity</td>
<td>The student researcher followed an intervention script (see Appendix H) to ensure treatment fidelity. Additionally, the student researcher pilot tested before implementing protocol.</td>
</tr>
<tr>
<td>Setting</td>
<td>DTS intervention sessions took place at the University of Miami Frost School of Music graduate music therapy teaching assistant office or in a quiet office at the Debbie School at the University of Miami Mailman Center for Child Development, dependent upon the family’s preferred location.</td>
</tr>
<tr>
<td>Unit of delivery</td>
<td>Each participant received the DTS intervention through individual sessions with the student researcher.</td>
</tr>
</tbody>
</table>
promotes rapport between the student researcher and participant. The metronome
beat was only perceptible to the student researcher via an earbud in one ear. Based upon
Leong and Goswami’s work (2014), participants heard nursery rhyme sentences which
alternated between stressed and unstressed syllables; however, the DTS intervention
contained adapted sentences. The sentences were presented at a slower rate based on
recommendations for children with CIs. Additionally, the participants practiced poems in
their entirety at home to increase the amount of practice and therefore potential
therapeutic benefit. Each poem was presented to the participant four times, with short,
10-second breaks between each poem repetition. In addition, the participant and student
researcher practiced drumming to nonsense syllables in either iambic (e.g., DEE da, DEE
da) or trochaic format (e.g., da DEE, da DEE) to reflect the stress structure of the poem.

Pictures. The student researcher provided pictures to represent each line of poetry
(see Appendix G). The student researcher utilized these pictures to increase engagement
with the DTS intervention and therefore facilitate rhythmic practice. The pictures
featured animated images created in Adobe Photoshop and did not contain words. An
example can be found in Appendix G.

Instruments. The student researcher provided one Remo 10-inch tubano (Model
#TU-1112-PM) for the participant, and one 10-inch tubano drum for the student
researcher for each of the four weekly DTS sessions. The student researcher and
participant each utilized a separate drum to facilitate separate recording of participant
drumming responses and to distinguish between the drums through different timbres. To
record the participant’s drumming responses during synchronization activities, the
participant’s drum was equipped with a drum trigger sensor which was fed through a
single-channel mixer to the student researcher’s laptop, where responses were recorded through Audacity. Drum trigger recordings were processed in MATLAB to identify participant drum hits and stimulus (metronome) onsets, which were later analyzed to determine accuracy and variability of the participant’s responses in relation to the metronome.

**Experimental Procedure**

The participants each met with the student researcher for five individual meetings and received four total DTS intervention sessions. The first meeting lasted around 60 minutes and consisted of pre-test measures and the first DTS intervention session. The second, third, and fourth meetings lasted around 30 minutes and consisted of the second, third, and fourth DTS sessions. The fifth meeting, which also lasted around 30 minutes, consisted of post-test measures only. All five in-person meetings took place in a quiet office at the Debbie School (part of the University of Miami Miller School of Medicine Mailman Center for Child Development) or in the music therapy graduate teaching assistant office in the Frost School of Music at the University of Miami’s Coral Gables campus. The student researcher and families determined meeting locations based on family preferences.

**Consent and demographic questionnaire.** Prior to the first session, parents gave written, informed consent, and the participant gave verbal assent. After receiving both parental consent and child assent, the student researcher continued with study implementation without the parents present in the sessions. The parent received a questionnaire (Appendix E) to complete during the session or to take home and bring back to the next session. The questionnaire asked for information regarding the child’s
demographics, speech/language experience, amount of auditory-verbal therapy received each week, and educational background.

**Pre-tests.** The student researcher administered pre-tests at the beginning of the first in-person meeting. Participants completed the Profiling Elements of Prosodic Speech – Children subtests for contrastive stress and affective understanding (PEPS-C 2; Peppé & McCann, 2003) and Audie test of music perception (Gordon, 1989) prior to beginning the intervention. Pre-test order (music or prosody) was randomized across subjects to control for potential order effects. After completing pre-test measures, the student researcher transitioned to the first DTS session.

**DTS intervention sessions.** Sessions lasted for about 30 minutes and occurred once a week for four weeks. Sessions followed a pre-determined structure of an introductory activity, synchronization activity, drumming-to-speech activity, and a closing activity (see Appendix H for a detailed session plan). The weekly poem order (Trochaic 1, Trochaic 2, Iambic 1, Iambic 2) presented to participants during the drumming-to-speech portions of the session was randomized across participants to control for potential order effects (see Appendix F).

**Introductory activity.** The introductory activity provided a transition for the participant into the study setting, provided basic instruction and practice on drum technique, and built rapport between the participant and student researcher. The participant practiced drumming along to the stressed syllables in one- and two-word utterances with prompting and guidance from the student researcher. The phrases consisted of the participant’s and student researcher’s names, one-word utterances selected by the child, and two-word utterances featuring contrastive stress.
Each week, the participant and student researcher practiced drumming to each other’s names four times each. Next, the student researcher facilitated practice of another one-word utterance category, the order of which was randomized across participants: color, movie, food, or animal. Both the student researcher and participant selected a word from this category and practice drumming to its stressed syllable four times in a row. Finally, the student researcher led the participant in a contrastive stress activity in which the participant and student researcher hit the drum during one word in a pattern, but not on the second word (for example, BLUE dog, BLUE dog, BLUE dog, BLUE dog), in total saying the pattern four times. The first word received the stress and drum hit during the first pattern of four times through the words, and then during the second repetition, the second word received the stress and drum hit (for example, blue DOG, blue DOG, blue DOG, blue DOG). Then, the student researcher and participant created another color and animal and practiced contrastive stress for those patterns as well. In total, the introductory activity took around five minutes for each participant.

**Synchronization assessment.** The synchronization assessment served to determine the child’s ability to entrain drum hits to a rhythmic stimulus. Following Kirschner and Tomasello (2009), the student researcher led the participant in a synchronization assessment. The student researcher listened to a metronome through an earbud in one ear that was inaudible to the participant throughout the synchronization activity and instructed the participant to hit the drum at the same time she did. The student researcher then began to drum on her tubano drum at the same time as the metronome stimulus in her ear, providing an audio-visual stimulus for synchronization. This process occurred for two, 20-second trials each of two different tempi: a fast tempo
of 400 ms ISI (2.5 Hz or 150 bpm) and a slow tempo of 600 ms ISI (1.67 Hz or 100 bpm). If the child did not synchronize during the first trial of a given tempo, the student researcher provided hand-over-hand assistance for the second trial of that tempo. The student researcher made records of assistance required during each trial in the data collection sheet. The synchronization activity in total lasted approximately 5 minutes. While the student researcher performed the assessment at each session, only the assessments from each participant’s first and last sessions were analyzed to assess overall change.

**Drumming-to-speech activity.** This task facilitated the practice of speech rhythm perception, an aspect of prosodic (specifically, linguistic prosodic) perception (Hausen et al., 2013) within the context of nursery rhymes. This portion of the experiment is adapted from work done by Leong and Goswami (2014) in which adults with dyslexia drummed to nursery rhyme sentences at the stress level; however, this intervention utilized full poems at a slower spoken rate based upon intervention recommendations for children with CIs (Gfeller et al., 2011). Each session centered around one nursery rhyme, and the order of rhymes was randomized across participants. During this activity, the student researcher presented the stimuli while listening to a metronome playing a rate of 90 beats per minute (bpm), or 1.5 Hz, through an earbud in one ear that was inaudible to the participant.

First, the student researcher demonstrated the ability to “drum along with speech” to the participant. The student researcher first presented a trochaic or iambic pattern of neutral syllables spoken with exaggerated prosody (e.g., da DEE da DEE da DEE da DEE for iambic) in which the syllable stress followed the type of patterning of the
nursery rhyme for that session. The student researcher synchronized her speech with the metronome to speak at a slightly slower rate than normal speech (90 bpm), as recommended for facilitation of rhythmic processing (Gfeller, 2000). Next, the student researcher guided the participant in drumming with one hand on the tubano drum to the stressed syllables in the line of neutral syllables, repeating four times with breaks between repetitions. When prompting the participant to tap or drum, the student researcher also drummed in order to provide a visual cue for the client, as visual cues are helpful for rhythm perception (Hsiao & Gfeller, 2012). If the participant did not engage in the activity or did not successfully align drum hits with stressed syllables, the student researcher provided hand-over-hand assistance.

The student researcher then read the nursery rhyme aloud while pointing to pictures representing each line of the rhyme and synchronizing the rhyme to a metronome played through an earbud in one ear at a rate of 90 bpm. The pictures (see Appendix G) provided a visual cue for the child for each line of the poem to further enhance processing of the semantic meaning of the nursery rhyme and were only used during this first presentation of the nursery rhyme when the student researcher and participant recited the poem together. Next, the student researcher prompted the participant to drum along to the stressed syllables on the tubano drum, providing hand-over-hand assistance when appropriate. The student researcher facilitated practice of each nursery rhyme four times.

Finally, the student researcher asked the participant to repeat each line of the nursery rhyme with her. The amount of each line that the participant repeated depended on their success. The student researcher started by presenting the first line of the nursery rhyme and asked the participant to repeat the line. If the participant did not remember the
whole line, or could not approximate the words in the line, the student researcher would present the participant with four syllables (half of a line). If this was also unsuccessful, the student researcher presented the nursery rhyme word-by-word. This process happened for each line of the nursery rhyme through two repetitions of the rhyme. The student researcher made note of the highest level at which the participant completed the task. This activity lasted approximately ten minutes.

**Closing activity.** The closing activity included “copy-cat,” a rhythm game in which the participant repeated rhythms presented by the student researcher on the tubano drums, and a goodbye song. The closing activity provided a method of practicing rhythm perception within a musical modality, instead of combining rhythm with speech as before. During “copycat,” the student researcher presented between three and five simple, 2- to 4-note patterns and asked the participant to play back the rhythm (i.e., quarter note, eighth note-eighth note, quarter note). The student researcher presented the same pre-composed rhythms to each participant. First, the student researcher presented the same rhythm to all participants first (three quarter notes), and based upon the participant’s success, selected the next rhythms from the pre-composed list. For example, if the participant did not succeed in repeating the first rhythm, the student researcher chose another simple rhythm (four quarter notes). However, if the participant did succeed, the student researcher chose a more complex rhythm (quarter note, two eighth notes, then one quarter note).

The student researcher then led the participant in a short goodbye song to serve as a transition out of the session. In this activity, the participant drummed on the second,
stressed syllable of “good-bye” when prompted during the context of a song. The closing activities lasted about five minutes in total.

**At-home practice.** After each DTS session, parents received an information sheet with details about that week’s at-home practice (see Appendix I). This intervention incorporated at-home practice to increase the amount of time participants had to practice identifying stressed syllables in speech. Parents also received a link to a Google Drive folder after the first session that contained four separate tracks. Each track guided the participant through a five-minute practice session of a particular poem, during which time they drummed to the stressed syllables of the poem practiced during that week’s DTS intervention session (see Appendix J). The student researcher asked parents to practice with their child at home five different days before the next session. In addition, the student researcher provided parents with a practice log and asked them to mark on the sheet the days their child completed the practice track.

**Post-test.** One week following the final DTS intervention session, participants completed post-test measures with the student researcher. The participant completed the same measures utilized during the pre-tests, contrastive stress and emotional prosody understanding subtests of the PEPS-C 2015 (Peppé & McCann, 2003) and the Audie test of music perception (Gordon, 1989).

**Data Analyses**

The student researcher performed descriptive analyses on the following demographic variables: gender, type of hearing device or amplification (bilateral v. unilateral v. bimodal), chronological and hearing ages, primary language, education type, number of music activities, amount of auditory verbal therapy, and previous music
therapy experience. The student researcher analyzed all synchronization data to assess the offset accuracy and variability, tempo accuracy and variability, average phase, and phase consistency of the participant’s drumming during synchronization activities. The student researcher conducted Wilcoxon signed-rank tests between pre-test and post-test results of prosody, music, and synchronization scores to determine the presence of significant changes following the DTS intervention. The student researcher also calculated Spearman correlation coefficients between age, age at implantation, type of hearing device or amplification, and synchronization measures with prosody and music scores.
Chapter Four

Results

This chapter contains the results of descriptive analyses, as well as statistical analyses that tested for differences between pre-test and post-test scores for the dependent variables of prosody perception (emotional and linguistic), music perception (melody and rhythm), and synchronization ability (offset accuracy, offset variability, tempo accuracy, tempo variability, and phase consistency). The student researcher used non-parametric, Wilcoxon signed-rank tests to determine group differences because of the small sample size and non-normal distribution. Pre-test and post-test scores for music and prosody perception were taken directly from the Audie and PEPS-C 2015 results, respectively. Pre-test and post-test synchronization data were processed through customized scripts in MATLAB which identified each drum hit and compared participant responses with stimulus onsets. The program then computed the participants’ offset accuracy, offset variability, tempo accuracy, tempo variability, average phase, and phase consistency for both the fast and slow stimulus conditions.

This chapter provides results of descriptive analyses performed on participants’ demographic questionnaire results, followed by inferential results from non-parametric statistical analyses and correlations to address research questions.

Descriptive Results

Participants. Participants were preschool-aged children between the ages of 3 and 5 years ($M = 4.622$ years, $SD = .899$). Detailed demographic information for participants is provided in Table 3.
### Table 3

*Participant Demographics of Total Sample (n=12)*

<table>
<thead>
<tr>
<th>Demographic Variables</th>
<th>Frequency in Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
</tr>
<tr>
<td><strong>Hearing Mode</strong></td>
<td></td>
</tr>
<tr>
<td>Hearing Aid + Cochlear Implant (Bimodal)</td>
<td>2</td>
</tr>
<tr>
<td>Bilateral Cochlear Implants</td>
<td>10</td>
</tr>
<tr>
<td><strong>Chronological Age (in Years)</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td><strong>Hearing Age (in Years)</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><strong>First Language</strong></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>6</td>
</tr>
<tr>
<td>Spanish</td>
<td>6</td>
</tr>
<tr>
<td><strong>Education Type</strong></td>
<td></td>
</tr>
<tr>
<td>Auditory/oral classroom</td>
<td>11</td>
</tr>
<tr>
<td>Mainstreamed classroom</td>
<td>1</td>
</tr>
<tr>
<td><strong>Number of Weekly Music Activities at Home</strong></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><strong>Hours of Auditory-Verbal Therapy Each Week</strong></td>
<td></td>
</tr>
<tr>
<td>0 – .9</td>
<td>1</td>
</tr>
<tr>
<td>1 – 1.9</td>
<td>6</td>
</tr>
<tr>
<td>2 – 2.9</td>
<td>4</td>
</tr>
<tr>
<td>3 – 3.9</td>
<td>1</td>
</tr>
<tr>
<td><strong>Previous Experience with Music Therapy</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>10</td>
</tr>
</tbody>
</table>

**Intervention completion.** Table 4 describes the amount of at-home practice completed by participants and the amount of overall assistance required during the interventions. The student researcher marked levels of assistance required during each session for each activity as “completed,” “completed with assistance,” and “did not do,” then recorded an overall amount of assistance for the entire session. Eight participants
completed the tasks without assistance during the first week and ten did so during the
final week.

Table 4

<table>
<thead>
<tr>
<th>Intervention Activity</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>At-Home Practice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of forms returned</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Average days of practice</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td><strong>Assistance Required During Session</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Completed with assistance</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Did not do</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Research Questions**

Table 5 summarizes results from the Wilcoxon signed rank tests to assess the
significance in differences between pre- and post-test scores for music perception and
prosody perception. Table 6 summarizes performance and inferential results for
synchronization ability. Next are results of Spearman correlations between demographic
and dependent variables paired with follow-up statistical analyses when relevant.

**Effect of DTS on prosody perception (RQ 1).** Wilcoxon signed-rank tests
revealed a significant difference between pre-test and post-test results for the emotional
prosody perception subtest of PEPS-C 2015, but not for the contrastive stress
understanding subtest. Participants performed significantly better on the emotional
prosody subtest post-test \( (M = 9.42, SD = 2.23) \) than pre-test \( (M = 7.75, SD = 2.05; Z(12) = 2.31, p = .021, r = .668) \). Though scores declined between contrastive stress pre-test
scores \( (M = 10.33, SD = 2.06) \) and post-test scores, \( (M = 9.92, SD = 2.64; Z(12) = .409, p = .682, r = .118) \) this result was not significant. Participants who scored 12 or more on
both subtests received a “pass” qualification, while participants who scored under this amount received a “fail” rating.

**Effect of DTS on music perception (RQ 2).** Wilcoxon signed-rank tests revealed a significant increase between pre-test and post-test scores for both melody and rhythm subtests of Audie (Table 5, Figure 3). Participants performed significantly better on the melody subtest post-test ($M = 6.67, SD = 1.50$) than pre-test ($M = 5.50, SD = 1.24; Z(12) = -2.23, p = .026, r = -.643$). Participants also performed significantly better on the rhythm post-test ($M = 7.08, SD = 1.73$) than they did on the pre-test ($M = 5.67, SD = 1.78; Z(12) = -2.70, p = .007, r = -.780$).

**Table 5**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>$Z$</th>
<th>$p$</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melody</td>
<td>$M = 5.50$</td>
<td>$M = 6.67$</td>
<td>-2.23</td>
<td>.026*</td>
<td>-.643</td>
</tr>
<tr>
<td>Rhythm</td>
<td>$M = 5.67$</td>
<td>$M = 7.08$</td>
<td>-2.70</td>
<td>.007**</td>
<td>-.780</td>
</tr>
<tr>
<td>Emotional</td>
<td>$M = 7.75$</td>
<td>$M = 9.42$</td>
<td>-2.31</td>
<td>.021*</td>
<td>-.668</td>
</tr>
<tr>
<td>Linguistic</td>
<td>$M = 10.33$</td>
<td>$M = 9.92$</td>
<td>-0.409</td>
<td>.682</td>
<td>-.118</td>
</tr>
</tbody>
</table>

*Note. *$p < .05$. **$p < .01$*

**Relationship between changes in prosody and music perception (RQ 3).** The student researcher performed non-parametric Spearman correlations to examine relationships between performance on music and prosody perception measures. Change scores reflect the difference between post-test and pre-test performance for each measure. Spearman correlations did not reveal any significant relationships between change scores for rhythm and melody perception within music and emotional and linguistic prosody perception in speech. Table 6 provides more detailed information.
Figure 3. Average score percentages for dependent variable measures of prosody and music perception during pre- and post-test. *$p < .05$, **$p < .01$.

Table 6

Spearman Correlation Coefficients between Dependent Variable Changes

<table>
<thead>
<tr>
<th></th>
<th>Emotional Prosody</th>
<th></th>
<th>Linguistic Prosody</th>
<th></th>
<th>Melody</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$p$</td>
<td>$r$</td>
<td>$p$</td>
<td>$r$</td>
<td>$p$</td>
</tr>
<tr>
<td>Emotional Prosody</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Linguistic Prosody</td>
<td>.007</td>
<td>.983</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Melody</td>
<td>.454</td>
<td>.138</td>
<td>-.152</td>
<td>.636</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Rhythm</td>
<td>-.161</td>
<td>.617</td>
<td>-.169</td>
<td>.599</td>
<td>.157</td>
<td>.626</td>
</tr>
</tbody>
</table>

Effect of DTS on synchronization ability (RQ 4). The student researcher performed Wilcoxon signed-rank non-parametric tests for significance between pre-test and post-test scores for offset accuracy, offset variability, tempo accuracy, tempo variability, and phase consistency for fast and slow conditions, and an “overall” condition. The overall condition is a composite score of fast and slow condition performances adjusted for differences in ISI and provides a score across both conditions for pre-test and post-test variables. Wilcoxon signed-rank tests revealed a significant
decrease between pre-test and post-test scores for phase accuracy in the slow condition ($Z = -2.118, p = .043, r = .611$), whereby participants hit the drum significantly more accurately (smaller time interval between drum hits and stimulus onsets) during the post-test ($M = .485, SD = .586$) than they did in the pre-test condition ($M = .836, SD = .843$). There were no other significant results. Table 7 contains results for all Wilcoxon signed-rank tests for synchronization ability.

Table 7

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Offset Accuracy (ms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast</td>
<td>.023</td>
<td>.023</td>
<td>.018</td>
<td>.012</td>
</tr>
<tr>
<td>Slow</td>
<td>.023</td>
<td>.017</td>
<td>.019</td>
<td>.021</td>
</tr>
<tr>
<td>Offset Variability (ms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast</td>
<td>.088</td>
<td>.001</td>
<td>.085</td>
<td>.001</td>
</tr>
<tr>
<td>Slow</td>
<td>.134</td>
<td>.002</td>
<td>.130</td>
<td>.001</td>
</tr>
<tr>
<td>Tempo Accuracy (ms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast</td>
<td>.509</td>
<td>.031</td>
<td>.456</td>
<td>.001</td>
</tr>
<tr>
<td>Slow</td>
<td>.665</td>
<td>.036</td>
<td>.603</td>
<td>.004</td>
</tr>
<tr>
<td>Tempo Variability (ms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast</td>
<td>.273</td>
<td>.142</td>
<td>.187</td>
<td>.009</td>
</tr>
<tr>
<td>Slow</td>
<td>.327</td>
<td>.076</td>
<td>.193</td>
<td>.014</td>
</tr>
<tr>
<td>Average Phase (rad)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast</td>
<td>.917</td>
<td>.501</td>
<td>.583</td>
<td>.471</td>
</tr>
<tr>
<td>Slow</td>
<td>.836</td>
<td>.843</td>
<td>.485</td>
<td>.586</td>
</tr>
<tr>
<td>Phase Consistency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast</td>
<td>.383</td>
<td>.010</td>
<td>.412</td>
<td>.059</td>
</tr>
<tr>
<td>Slow</td>
<td>.430</td>
<td>.049</td>
<td>.402</td>
<td>.068</td>
</tr>
</tbody>
</table>

Note. *$p < .05$.

**Relationship between demographic variables and DTS performance.**

Spearman correlations were performed between post-test results for prosody and music perception scores and demographic variables: gender, hearing mode, chronological age, hearing age, first language, education type, number of music activities, hours of auditory verbal therapy per week, and previous music therapy experience. Table 8 contains results
of all Spearman correlations; one significant positive correlation was found, and this occurred between chronological age and performance on the emotional prosody post-test ($r = .803, p = .002$).

Table 8

<table>
<thead>
<tr>
<th></th>
<th>Emotional Prosody</th>
<th>Linguistic Prosody</th>
<th>Melody</th>
<th>Rhythm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$p$</td>
<td>$r$</td>
<td>$p$</td>
</tr>
<tr>
<td>Gender</td>
<td>-.123</td>
<td>.703</td>
<td>-.146</td>
<td>.650</td>
</tr>
<tr>
<td>Hearing Mode</td>
<td>-.402</td>
<td>.196</td>
<td>-.132</td>
<td>.682</td>
</tr>
<tr>
<td>Age (y)</td>
<td>.803</td>
<td>.002*</td>
<td>.035</td>
<td>.913</td>
</tr>
<tr>
<td>Hearing Age (y)</td>
<td>.545</td>
<td>.067</td>
<td>.112</td>
<td>.728</td>
</tr>
<tr>
<td>First Language</td>
<td>-.123</td>
<td>.703</td>
<td>-.342</td>
<td>.277</td>
</tr>
<tr>
<td>Education Type</td>
<td>-.357</td>
<td>.255</td>
<td>-.221</td>
<td>.491</td>
</tr>
<tr>
<td>Music Activities</td>
<td>-.216</td>
<td>.524</td>
<td>-.035</td>
<td>.919</td>
</tr>
<tr>
<td>AVT</td>
<td>.345</td>
<td>.272</td>
<td>.051</td>
<td>.875</td>
</tr>
<tr>
<td>History of MT</td>
<td>-.298</td>
<td>.347</td>
<td>.131</td>
<td>.685</td>
</tr>
</tbody>
</table>

*Note. AVT = Weekly Auditory Verbal Therapy. MT = Music Therapy. **$p < .01.$

Because of the significant Spearman correlation, the student researcher performed a post-hoc, non-parametric Kruskal-Wallis test between age categories on performance on the emotional prosody post-test. Results revealed a significant difference between emotional prosody post-test performance between age categories, whereby five-year-old children scored highest ($M = 11.25$), followed by four-year-olds ($M = 9.50$), and then three-year-olds ($M = 7.50$; $H(3) = 6.594, p = .037$). Figure 4 provides a visual representation of these results.
Figure 4. Performance on emotional prosody post-test by participant age. Kruskall-Wallis tests revealed a significant difference between age groups on performance for the emotional prosody post-test ($p = .037$).
Chapter Five

Discussion

The purpose of this exploratory study was to examine the effect of a novel drumming-to-speech (DTS) intervention on prosody perception in preschool-aged children with CIs. Previous research suggests benefits of musical training for prosody perception in both typically-hearing children (TH; Thompson et al., 2004) and children who utilize CIs (Good et al., 2017), though these studies did not explore a rhythm-based intervention to benefit prosody perception. Therefore, this study sought to determine if a rhythm-based intervention focused on identifying and drumming to stressed syllables in repeated phrases, nursery rhymes, and song could significantly benefit emotional and linguistic prosody perception. This study also explored additional questions not yet answered in the literature, such as the effect of such an intervention on music perception and synchronization ability, as well as the relationship between changes in music perception and prosody perception, as well as between demographic variables and intervention outcomes.

Overall, results indicated significant improvements in the perception of emotional prosody and of melody and rhythm within music but did not reveal any significant change in linguistic prosody perception. Change scores on music and prosody measures between pre-test and post-test were not significantly correlated. Participants significantly improved in their accuracy of drumming close in time to the stimulus (phase information) during synchronization at a slower speed but did not significantly change in any other synchronization measures. Finally, participant age was significantly correlated on post
test performance for the emotional prosody perception subtest, while other demographic variables were not correlated with any other dependent variable scores.

**Effect of Drumming-to-Speech on Prosody Perception**

Children with CIs fall behind their TH peers on measures of prosody perception (Hopyan-Misaken et al., 2009), and this delay has cascading implications for later language outcomes and communicative development (Paul et al., 2005; Pinker, 1987; Wennerstrom, 2001). Previous research demonstrated improvements for emotional prosody perception following synchronization with a metrically strong syllables for TH children (Falk & Dalla Bella, 2016). The present study contributes to the literature by providing evidence this effect can also be seen in a population who needs improvements in prosody during a developmental period in which prosody perception is important.

Participants significantly improved in emotional prosody perception following the DTS intervention. This finding is consistent with other studies that have demonstrated effects of musical training for emotional prosody perception (Petersen et al., 2012; Thompson et al., 2004). Of particular note, most musical training paradigms to benefit prosody incorporate melodic training, whether that is melodic contour training (Lo et al., 2015) or training on a melodic instrument (Good et al., 2017). The DTS intervention only involves one “song” during the intervention, which is the final goodbye song at the end of every session. Aside from this song, all of the other activities involve drumming to the prosody of repeated words and phrases or nursery rhymes. Therefore, if truly due to the effects of the intervention, the improved emotional prosody perception is a result of the practice in rhythm training consisting of drumming to speech and synchronization practice, rather than practice in music training per se (including melody, harmony, etc.).
The ability for drumming to speech to improve emotional prosody perception in preschool-aged children with CIs supports a role for rhythm perception in the identification of emotional prosody. The DTS intervention required explicit processing of speech rhythm through tasks in which the participant was required to align drum hits with stressed syllables. Explicit processing of speech rhythm recruits different neural structures than does implicit processing of speech rhythm, and those structures recruited during explicit processing (SMA, insula, and right IFG) more closely align with neural structures utilized for prosody and music perception (Geiser et al., 2008). Therefore, since the present intervention consisted of explicitly synchronizing movements with stressed syllables in speech, overt practice of identifying stressed syllables in phrases and nursery rhymes served as the overlap between music and prosody perception, strengthening participants’ perception and detection of durational cues which in turn benefited their prosody perception.

Additionally, rhythm plays a role in the perception of prosody by mediating the perception of pitch (Hausen et al., 2013). Through practice in identifying where within a sentence stressed syllables occur, and through strengthening this identification through synchronization of motor movements and increased auditory feedback from playing a drum, participants in the DTS intervention strengthened their ability to tell when salient prosodic changes occur, and this mediated and improved their perception of what that information conveyed. Alternatively, the intervention may work through a more direct mechanism, whereby practice in synchronization with speech presented at isochronous intervals heightened children’s sensitivity to durational variances within speech, cues that are important for prosody (Campbell, 2000). The speech samples in this study facilitated
synchronization through presentation of isochronous intervals, which acted as anchors to synchronization. However, these are speculations and further research is needed to determine the exact mechanisms underlying the DTS intervention.

Other studies have demonstrated effects of musical training for certain skills within linguistic prosody perception, such as turn-end prosody which cues conversational turns (Lo et al., 2015), but have not demonstrated an effect for focus prosody perception. This study investigated the effect of drumming to speech on focus prosody perception through the contrastive stress perception subtest of the PEPS-C 2015. No significant difference occurred between pre-test and post-test performance on the contrastive stress perception subtest, which may have been due to any combination of several possible reasons.

First, while the participants drummed to contrasting levels of stress (i.e., hitting the drum on stressed syllables and taking the hand off for unstressed syllables), the alternating stressed and unstressed syllables does not truly approximate focus prosody, which uses word-level emphasis to communicate suprasegmental meaning. Instead, in the DTS intervention, participants practiced identification of lexical stress by synchronizing with alternating stressed and unstressed syllables. As the intervention practiced perception of lexical stress, but then assessed perception of focus prosody, this may explain a lack of improvement in linguistic prosody following the DTS intervention. Additionally, the stress was regular and did not communicate meaning beyond semantics; in fact, because it was periodic the lack of ecologically-valid, non-isochronous variances present in everyday speech may have desensitized the participants to detecting differences in stress in a meaningful way. Future studies should investigate potential
differences in benefit between stress patterns utilized during interventions (i.e., trochaic or iambic), as the order of the unstressed and stressed syllables may affect changes seen in focus prosody. As the stress differences did not have a direct application to focus prosody, this may account for why the intervention did not benefit focus prosody perception; therefore, future intervention should distinguish and assess the intervention’s efficacy for improving perception of both lexical stress and focus prosody.

Another reason for not seeing an effect of drumming to speech for linguistic prosody perception could be the test utilized. The PEPS-C 2015 contrastive stress understanding subtest assesses contrastive stress, which is technically different from focus prosody. While both skills utilize differences in stress to identify meaning, contrastive stress may rely more upon fundamental frequency changes (Bolinger, 1961) than focus prosody, which relies upon changes in fundamental frequency as well as duration and amplitude (Raphael et al., 2007; Most & Peled, 2007). Contrastive stress and focus prosody might be completely separate skills, and so by addressing contrastive stress rather than focus prosody, the dependent variable measure did not assess the skill practiced during the intervention. The PEPS-C also incorporates stimuli which are presented through speakers. As children with CIs have more difficulty perceiving recorded speech than they do perceiving natural speech (Gfeller, 2000), utilizing a test which incorporates natural stimuli or which provides direct transmission to the CI may be a more ecologically-valid and accessible method of testing prosody perception for children with CIs.

Additionally, the test utilized seemed to confuse some of the younger children; for example, those children responded by picking their favorite color of sock, departing from
the instructions that directed them to pick the color the speaker stressed. The test also
does not provide an adaptation for children who are colorblind, and the present study did
not include one, so it is possible this could have confounded the results. Another test, the
Perception of Prosody Assessment Tool (PPAT; Klieve, 1998), is intended for use with
children who use CIs. Though the PPAT has not been widely used, it is possible this test
may have been more appropriate for this study as it was developed to assess prosody in
the population specific to this study (i.e., children with CIs). Future research could
consider utilizing this measure to assess prosody perception.

Another reason the intervention did not impact linguistic prosody perception may
have been its length. Four, weekly 20- to 30-minute music therapy sessions is a relatively
short-term intervention, and an element of linguistic prosody such as focus prosody
perception may require a longer length of practice in order to see changes. Alternatively,
the children could participate in more intensive intervention scheduling, such as attending
music therapy sessions two or three times per week for 45 minutes each. More practice
even within the same time frame as this intervention may be enough to see benefits in
linguistic prosody perception. Previous research demonstrates temporal precedence for
the benefits of music on emotional prosody perception relative to linguistic prosody.
Emotional prosody is processed temporally before linguistic prosody (Paulmann et al.,
2012), and in adults, music training leads to quicker improvements in emotional prosody
recognition (Petersen et al., 2012). This intervention did involve an at-home practice
component, but only a few participants completed this requirement. A longer or more
intensive intervention with more compliance or more in-person practice with the therapist
may develop perceptual abilities even more to where they can also benefit linguistic prosody perception.

Despite improvements seen in emotional prosody and any individual changes in linguistic prosody perception, participants remained within the “fail” qualification for both prosodic subtests. The DTS intervention may have assisted in the development of prosody perception skills but following the intervention participants were still falling behind typically-hearing peers. It is possible that a longer or more intensive dosage of the DTS intervention may provide preschoolers with cochlear implants the ability to catch up to typically-hearing peers and “pass” in measures of prosody perception.

**Effect of Drumming-to-Speech on Music Perception**

Participants significantly improved in both melody and rhythm perception following the DTS intervention. Other research demonstrates a benefit for music perception in tandem with prosodic improvements (Galvin et al., 2007; Petersen et al., 2012). The DTS intervention involves practice in identifying the stressed syllable rhythm of speech and so, not surprisingly, the participants’ rhythm perception improved. With practice synchronizing to and perceiving speech rhythm in a variety of contexts, participants may have attended to rhythm more closely than before, and therefore performed better. Practice in synchronizing to speech at regularly stressed intervals may also have helped the participants identify salient places within the sentence for prosodic change, which improved their perception of relative durations and helped them achieve better performance in rhythm perception.

As mentioned previously, this intervention only utilized one song during treatment, and so it is somewhat surprising that participants also improved in melodic
perception. Improvements in melodic perception following the DTS intervention underscores the overlap in mechanisms underlying music and prosody perception. Participants did not practice identifying or performing melodies as in other musical trainings; they identified and synchronized to rhythmic speech. Behaviorally, practice in detecting important durational, prosodic information might strengthen participants’ temporal expectancies, thereby facilitating their perception of spectral information (Hausen et al., 2013; Marx et al., 2015; O’Halpin, 2010). Neurologically, the presence of an overlap in neural networks between prosody and rhythm perception may suggest a mechanism at work in the DTS intervention in a way like the OPERA hypothesis, in which the more demanding neural and behavioral precision and processing of music confers benefits for speech perception (Patel, 2011). In other words, strengthening areas utilized for both prosody and rhythm perception through the DTS intervention may not only confer benefits for prosody perception (Belyk & Brown, 2013; Geiser et al., 2008; Grahn, 2012; Thaut et al., 2014; Patel, 2011; Witeman et al., 2012), but also melody perception.

**Relationships between Changes in Music and Prosody Perception**

The difference in post-test and pre-test scores for prosodic and musical measures did not significantly relate to one another, despite the links between the shared behavioral and neural underpinnings of music and prosody perception reviewed above. If these skills truly are dissociable, different mechanisms may exist which support perception within each category (music or speech). Disparity between skills could also explain why linguistic prosody perception did not receive significant benefit from the DTS intervention. However, it is unlikely that music and prosody perception are this
dissociable, given the amount of evidence for behavioral and neural overlaps and similarity of processing even in amusic individuals (Patel et al., 1998). Music and prosody perception may change at different rates and in different ways, and a four-week intervention is not adequate to detect the relationship between such changes.

The limitations of the study may be responsible for not finding a relationship between intervention impact for music and prosody. As this was an exploratory study, results may not elucidate a link between change scores between music and prosody. However, findings do indicate significant benefits of the intervention for melody, rhythm, as well as emotional prosody perception. Benefits in music and speech domains suggest the intervention may improve these skills through practice of drumming to speech, a task which is not strictly linguistic nor musical but rather combines the two domains. Therefore, the intervention and/or study design may require more refinement in order to truly determine the link between changes in music and prosody perception.

Effects of Drumming-to-Speech on Synchronization Ability

Participants significantly improved in average phase in the slow condition following the DTS intervention. As offset accuracy reflects the participants’ ability to drum close in time to stimulus onsets, significant improvements in this area may suggest the importance of phase information as the rhythmic link between music and prosody perception. This improvement may indicate a benefit of practice in identifying the onset of stressed syllables in speech and the importance hand-over-hand assistance provided when the participant could not synchronize, as assistance may help the participant to “re-set” the phase of their entrainment in a way more akin to flexible entrainment models (Grahn, 2012). However, none of the other phase measures reached significance, nor did
any of the other synchronization measures. Lack of significant findings may be due to the more long-term nature of synchronization abilities, in which change is typically seen after many years of musical training (Repp, 2010), and which a short-term intervention may not impact. A longer or more intensive intervention could change synchronization abilities significantly throughout the course of treatment. Future studies should investigate this aspect, as the synchronization practice utilized in the DTS intervention may serve as a mechanism for improvements in both music and prosody perception.

**Links Between Demographic Variables and Music and Prosody Perception**

Age correlated significantly with emotional prosody post-test performance, and post-hoc results demonstrated significant differences between ages, whereby performance followed development; five year-olds performed the best, followed by four year-olds, and then by three year-olds. Anecdotally, three year-olds required the most re-direction and encouragement to perform intervention activities and did not seem to understand the tests, whereas five year-olds followed directions and seemed to understand the tests. Differences in performance across ages may point to a later window of development for which this intervention is most beneficial; this intervention may have been too complex for the three- and four year-olds, as their group average at post-test still fell well below the “pass” qualification for test scores (set at 12 correct out of 16, their group average post-test scores were 7.50 and 9.50, respectively).

A later developmental window for this intervention may be preferable because of chronological and language development. Children with CIs often experience language delays (Koopmans-van Beinum et al., 2001), and because of their younger hearing ages, have less experience with language than do typically-hearing children of the same
chronological age. Therefore, while typically-hearing preschool-aged children may be ready to acquire an understanding of prosody, the preschool age may be too early in language development for preschool-aged children with CIs.

Though the large number of correlations performed should be taken into consideration, other demographic variables did not correlate with performance in music or prosody measures. Though type of amplification did not correlate with performance, the two children who utilized bimodal hearing (one CI and one hearing aid) anecdotally improved more over the course of the treatment than did participants who utilized two CIs. As residual acoustic hearing benefits the processing of spectral information (Gfeller, 2000), future research should determine whether type of amplification has implications for candidates or progress in this intervention. Parents of 10 out of 12 participants noted no previous history of music therapy; however, the Debbie School classrooms receive music therapy services through hosting local university practicum students, so these numbers may not be accurate. Additionally, hours of weekly auditory-verbal therapy did not correlate with performance on either measure, though this amount of therapy may not indicate anything about the speech and language achievements of the participants. Future research should administer a standardized speech and language measure to determine relationships between language achievement and benefit from the DTS intervention; this may also identify relationships between linguistic development and windows for prosody interventions.

Additional Findings and Implications

Intervention feasibility. As this study was the first iteration of the DTS intervention, the student researcher recorded data regarding intervention feasibility. In
general, the number of participants who were able to complete all activities without assistance increased between the first and last weeks of the intervention (from eight to ten, respectively). Improved performance may reflect a practice effect, as the session structure each week was similar so participants may have learned what to expect from each session.

Despite the independence of some children during the last session, two of the younger participants still required assistance at the end of the intervention, and there were two additional children did not assent to any sessions beyond the third. These children who experienced compliance obstacles were on the younger end of the preschool age range and taken with the finding that the five year-olds performed the best at post-test, future interventions addressing prosody perception should involve upper preschool-aged to early school-aged children, or a simplified protocol for younger children.

**At-home practice.** The at-home practice component was not consistently completed across participants. The first week, parents returned six forms (out of 12), while in subsequent weeks, only three parents returned forms. Lack of at-home practice may have been due to the setting, as I did not have in-person contact with the parents each week and largely communicated with them by sending forms home in their children’s backpacks and through e-mail. Lack of at-home practice may also be difficult because of family schedules outside of school, and therefore should not be relied upon for therapeutic improvements. Future interventions could incorporate more sessions each week or better communication about the importance and expectations for practice at-home, if outside practice is deemed necessary.
Theoretical applications. Though an exploratory study, the findings from this study have broader theoretical implications. First, as drumming to the stressed syllables in speech benefited both emotional prosodic and musical understanding, identification of and synchronization with hierarchical rhythms may act as a shared behavioral and/or neurological mechanism between perception of prosody and music. This mechanism may not support linguistic prosody components, such as focus prosody, as practice in identifying stressed rhythms did not improve this area. However, the lack of improvement in focus prosody may also be due to typically-developing production of focus prosody by children with CIs as demonstrated in previous literature (Weiss, Carney, & Leonard, 1985), or possibly due to a stronger reliance of emotional prosody for variances in timing. Future research on speech rhythm tasks for focus prosody perception may better elucidate the similarities or differences between speech rhythm mechanisms underlying different types of prosody.

Second, participants experienced improvement in melody perception after synchronizing movements to speech rhythm, not to music itself. Participants did not practice any pitch-based tasks, and yet they improved in their ability to distinguish differences in melodies varying in pitch. The presence of the link between speech rhythm and melody perception supports previous research which suggests that rhythm is a facilitator for pitch perception, cueing the listener in to the arrival of salient information (Hausen et al., 2013). Further, this finding supports the use of rhythmic paradigms to improve prosody perception in children with CIs (O’Halpin, 2010). In addition, this study supports the overlap, precision, repetition, emotion, and attention mechanisms needed for
and strengthened during music perception that also benefit speech perception (Patel, 2011).

Finally, this study may also point to a window in linguistic and cognitive development where children reap the most benefits from an intervention directed at prosody. As the oldest age category in this study performed significantly higher on the post-tests than did the younger ones, the older children may possibly receive greater benefit from the intervention. Performance on the post-test did not correlate with hearing age but instead correlated with chronological age, which suggests a level of cognitive development may be necessary to complete the intervention. Anecdotally, the older children were more compliant and independent within the sessions than the younger children were. Future research could identify cognitive or linguistic skills (i.e., following multi-step commands, understanding of syntax) which must be in place in order to learn prosody perception.

**Clinical applications.** This study has applications for music therapy as well as other therapies which target prosody perception with children with CIs. This study suggests that drumming to rhythmic speech may benefit emotional prosody perception for children with CIs. Synchronization with speech that is spoken to a metronome helps the client to practice identification of speech rhythm, which in turn is beneficial for prosody perception. The present study identifies one such method of speech rhythm practice within a modified Rhythmic Speech Cueing (RSC) or Developmental Speech and Language Training through Music (DSLM) paradigm but does not yet represent a finalized protocol. Development of a technique of a more specific method for prosody perception and communication of this method with other professionals should be a
priority for music therapists working with this population and should involve collaboration with other disciplines such as speech-language pathology.

Speech therapists may also benefit from the incorporation of motor synchronization to identify stressed syllables in rhythmic patterns into their sessions. Practice in drumming or tapping to hierarchical speech rhythm in nursery rhymes or repeated phrases may increase clients’ perception of speech rhythm and therefore improve prosody perception. However, music therapists and speech therapists should communicate and co-treat in order to provide the greatest overlapping benefit of rhythmic and linguistic expertise.

**Limitations**

The present study is not without limitations. As the design of the study did not involve the use of a control group, observed outcomes may not truly be the result of the intervention, and could instead be the result of a practice effect. The small sample size makes it difficult to extrapolate findings to the general population of children with CIs difficult. Additionally, the measure used to assess prosody, though previously utilized for children with CIs, had not been utilized with preschoolers with CIs, and presented some difficulties both in administration and in communication of the intent of the research questions. Anecdotally, younger children did not seem to understand the instructions of the PEPS-C subtests, so future studies may need to select a more appropriate measure for this age group. The synchronization analyses also presumed that the student researcher’s tempo did not deviate at all from the metronome stimulus, which may not have been true. Future research with control groups, larger sample sizes, and more sensitive tests are needed to confirm these preliminary results.
Improvements can also be made to the DTS intervention itself. Though participants demonstrated improvements in music and emotional prosody perception after only four, 20- to 30-minute sessions of music therapy, this is a relatively short-term treatment length. In order to determine the impact of DTS on long-term linguistic prosody or synchronization abilities, future applications of this intervention could utilize an intervention period longer than one month and more frequent sessions within that period (such as two sessions per week). If at-home practice is utilized in future interventions (or if future research demonstrates that the at-home practice component is necessary), the researcher or clinician should check with the family regularly to ensure that the participant or client is practicing at home, so that practice is effective.

**Conclusions and Future Directions**

The present study implemented a novel drumming-to-speech (DTS) intervention with the intent of improving prosody perception for preschool-aged children with cochlear implants (CIs). Participants received weekly individual music therapy sessions which focused on drumming to stressed syllables in repetitive phrases or nursery rhymes following research suggesting the importance of speech rhythm identification for prosody (Hausen et al., 2013), and received additional practice in synchronization with a student researcher (Kirschner & Tomasello, 2009). After four weeks of sessions, participants significantly improved in their perception of emotional prosody and melody and rhythm within music but did not improve in perception of linguistic prosody (specifically, focus prosody). The intervention used only one song to close each session, and so benefits in emotional prosody and music perception seem due to practice in synchronization to speech rhythm. This study suggests an important role for synchronization with speech
rhythm in the development of prosody perception for children with CIs and provides additional support for the degree of overlap between neural and behavioral mechanisms involved in perceiving prosody and music, especially the rhythmic components of both.

Older participants performed best at the end of the intervention in understanding emotional prosody, which may suggest that older preschool-aged children could benefit most from an intervention targeting emotional prosody perception. Further research should be conducted with older children to determine the window of development during which an intervention targeting suprasegmental features of speech such as prosody may be most beneficial. Additionally, future research with controls and more robust interventions may elucidate other links that were not seen in this short-term intervention, such as links between speech rhythm synchronization and linguistic prosody or most synchronization abilities.
References


Kersken, V., Zuberbühler, K., & Gomez, J. C. (2017). Listeners can extract meaning from non-linguistic infant vocalisations cross-culturally. *Scientific Reports* 7, 41016. doi:10.1038/srep41016


Appendix A: Recruitment Flyer

- Study #: 20180569   Effective Date: 9/7/2018

MUSIC STUDY

PRESCHOOLERS WITH COCHLEAR IMPLANTS ARE BEING ASKED TO TAKE PART IN A RESEARCH STUDY THROUGH UNIVERSITY OF MIAMI:
Drumming to Speech to Improve Prosody Perception in Children with Cochlear Implants: A Pilot Study

FREE drumming program, once a week for 5 weeks!

REQUIREMENTS & ELIGIBILITY
OPEN TO ANY PRESCHOOLER (AGES 3-8) USING A COCHLEAR IMPLANT AND EXPOSED TO ENGLISH CHILD PARTICIPATES IN 4 MUSICAL SESSIONS AND 1 SESSION OF MUSICAL AND SPEECH GAMES PARENT PARTICIPATES IN AT-HOME PRACTICE AND QUESTIONNAIRE

IF YOU HAVE QUESTIONS, PLEASE CONTACT
Jessica MacLean, UM Graduate Teaching Assistant XXX-XXXX-XXXX

Yes, I am interested in having my child participate in the music study!

Parent's Name: ____________________________
Child's Name: ____________________________
Parent Phone Number: ______________________
Parent Email Address: ______________________
Appendix B: Informational Letter to Parents

August 1, 2018

Dear Preschool Parents,

My name is Jessica MacLean. I am a graduate student in music therapy at the University of Miami.

This fall I will be doing a research study as part of my graduate education. The purpose of this study is to help me learn how music can be used to help children better understand the meaning behind others’ speech. This study is open to any preschool child (ages 3-5) who uses at least one cochlear implant and who is exposed to English. I hope to have thirty children in the study and participation is completely voluntary.

Children who participate will receive individualized music sessions. We will be doing activities that center around drumming and hearing the “rhythm” in speech. Each session will be about 30 minutes long and will take place during a time that is convenient for you and your child. I will see each child a total of 5 times (once a week for five weeks). The first session will be a little longer (60 minutes) as this session will also include some music and speech games that will allow me to see if our music sessions have helped your child understand speech better. In addition, I will be asking you to fill out a questionnaire and help your child with drumming homework for five minutes, once a day during the study.

Please find enclosed a flyer for the study with my contact information. You are welcome to contact me if you have questions or are interested in participating. I will also be at the preschool on Tuesday, August 21st from 7am to 9am and Thursday, August 30th from 2pm to 4pm. During these on-site visits I will be available to talk with you about the study in detail and can answer any questions.

I look forward to meeting you. Feel free to contact me (317-430-2399) if you have any questions or would like more information.

Thank you for your consideration.

Jessica MacLean
Appendix C: Phone Recruitment Script

Hello, this is Jessica MacLean calling from the University of Miami. May I please speak with [parent name]?

**At any point during this conversation, if parent is no longer interested:** Thank you so much for your time and consideration. Have a wonderful day.

*If parent is not available:* What would be a better time for me to call back and reach him/her?

*If parent is available:* Hello, [parent name]. I am a graduate student in music therapy. I received your contact information from [staff member] at [UM Ear Institute/Debbie School] (OR we spoke at [event] at [UM Ear Institute/Debbie School] about enrolling [child name] in a current research study). I am calling to see if you and your child [name] would be interested in a music research study that I am conducting, would you like to learn more information?

This fall I will be doing a research study as part of my graduate education. The purpose of this study is to help me learn how music can be used to help children better understand the meaning behind others’ speech. This study is open to any preschool child (ages 3-5) who uses at least one cochlear implant and who is exposed to English. I hope to have thirty children in the study and participation is completely voluntary.

Children who participate will receive individualized music sessions. We will be doing activities that center around drumming and hearing the “rhythm” in speech. Each session will be about 30 minutes long and will take place during a time that is convenient for you and your child. I will see each child 5 times, once a week for five weeks. The first session will be a little longer (60 minutes), as this session also includes some music and speech games that will allow me to see if our music sessions have helped your child understand speech better. In addition, I will be asking you to fill out a questionnaire and help your child with drumming homework for five minutes, once a day during the study.

Would you and your child be interested in participating?

If so, I have a few questions to ask you. One, does your child use a cochlear implant?

How old is your child?

Does your child hear English at home, at school, or in the community?

Does your child follow simple instructions in English?

(If the parent answers yes to the above questions) Thank you so much for your interest in participating! I would like to schedule a 60-minute appointment for the first meeting, since we have a couple of extra games to play the first time around. What day/time would
work best for you for the first appointment? Also, what location would you prefer? We are conducting the study at the University of Miami Frost School of Music on the Coral Gables campus, the UM Ear Institute downtown in the medical district, the UHealth clinic in Plantation, and at the Debbie School also in the medical district.

[Student researcher will give the parent the address of the preferred location and describe parking information for that site.] Would you like to schedule the following four sessions now, or when we meet on [meeting date]? Ideally, it would be great to meet on the same day each week.

Great! Would you prefer a confirmation call, text, or email? I will send you confirmation of this appointment the day before.

Thank you so much, [parent name]. I look forward to seeing you and [child name] on [date]. If you have any questions, please reach out to me. Have a good day!
Appendix D: In-Person Recruitment Script

Good evening. My name is Jessica MacLean. I am currently a graduate student in music therapy at the University of Miami. Earlier this month I sent a flyer and informational letter home with your child that introduced a research study I am doing as part of my education. I am here today to share some information about my study, answer questions, and invite those with children who are three- to five-years old, who understand English, and who use at least one cochlear implant to participate.

The purpose of this study is to help me learn how music, specifically drumming, can help your child perceive prosody, or the “melody of speech” that helps us understand others’ meaning. I hope to have thirty children in this study. Your participation is completely voluntary. If you agree, your child will receive five weekly music sessions. One set of games will occur immediately before the first music session, and the last set of games will occur one week after the fifth, or last, session. I will be leading all music sessions. We will be doing music activities that involve drumming and listening to speech. Each session will be about 30 minutes long, except for the first session, which will be 60 minutes long. Sessions will take place during a time that you and I will schedule together. I will see each child 5 times, once a week for five weeks. This study also involves some participation from you in the form of filling out questionnaires and guiding your child in some at-home drumming practice to reinforce concepts we go over in our sessions. This will all take no more than five hours of your child’s time in total over the next five weeks.

My purpose in talking with you today is to see if you are interested in learning more about the study and to answer any questions you may have. We can either discuss any questions you may have about the study and determine if your child might be the right fit now, or I can call or email you at a later time if that is more convenient.

If a parent wants to discuss at that moment: Would you and your child be interested in participating? If so, I have a few questions to ask you. One, does your child use a cochlear implant? How old is your child? Does your child hear English at home, at school, or in the community? Does your child follow simple instructions in English? (If the parent answers yes to the above questions) Thank you so much for your interest in participating! I would like to schedule a 60-minute appointment for the first meeting, since we have a couple of extra games to play the first time around. What day/time would work best for you for the first appointment? Also, what location would you prefer? We are conducting the study at the University of Miami Frost School of Music on the Coral Gables campus, the UM Ear Institute downtown in the medical district, the UHealth clinic location in Plantation, and at the Debbie School also in the medical district. I will send you a confirmation email today with your time, my contact information, and more details about the study location (including parking, accessibility, etc.).
If a parent wants to discuss later: If your child would be willing to participate, please write down your name, email, phone number, and general availability on the sheet I have provided, and I will contact you to discuss the study and determine if your child might be the right fit at a later date.

(Answer questions.)

If a parent says “yes”: Thank you (parent name) for your time today. I will be contacting you via (phone or email) to schedule a session and answer any other questions you may have.

If a parent says "yes" and a session is scheduled: Thank you (parent name) for your time today. I look forward to meeting you on (insert date) at (insert location). If you have any questions between now and then, please do not hesitate to contact me at 317-430-2399.

If a parent says "no": Thank you (parent name) for your time today and your consideration. Have a wonderful day.
Appendix E: Email Recruitment Script

Dear (parent name),

My name is Jessica MacLean, and I am a graduate student in music therapy at the University of Miami. [I am reaching out to you to follow-up on our conversation at (place name) about my research study.] OR [I received your contact information from (name) and am reaching out to you to invite your child to participate in a music research study. For my master’s thesis, I am looking at how music (specifically, drumming) can help children with cochlear implants to better understand the meaning behind what others say (prosody). The study will consist of five total sessions that occur once a week for five weeks. The first session will last 60 minutes, while the other sessions will last 30 minutes. Sessions will consist of drumming games that focus on helping children to better hear the emotions and meaning behind others’ speech. During the first and last sessions, we will play music and speech games that will allow me to see if this intervention is helping your child to better understand the meaning behind speech. I will be asking for your participation filling out questionnaires, and for your assistance in completing drumming homework with your child to reinforce concepts learned in the intervention.

In order to determine if your child may be eligible for this study, please answer the following questions:
   1. Does your child use a cochlear implant?
   2. Is your child between the ages of 3 and 5?
   3. Does your child hear English at home, at school, or in the community?
   4. Does your child follow simple instructions in English?

If you answered yes to the above questions, your child is eligible to participate in our study! If you are interested, please send me your preferred meeting times and location preference(s) from the options below. I look forward to hearing from you soon!

1. Barton G. Kids Hear Now Foundation
   Clinical Research Building, 5th Floor
   1120 NW 14th St, Miami, FL
2. Debbie School
   Mailman Center for Child Development
   1601 NW 12th Ave, Miami, FL
3. Frost School of Music
   University of Miami, Coral Gables Campus
   1314 Miller Drive, Coral Gables, FL
4. UHealth at Plantation
   8100 SW 10 St
   Plantation, FL 33324

Thank you,
Jessica MacLean
Appendix F: Demographic Questionnaire

Please answer the following questions regarding your child and return to the student investigator (Jessica) before the end of the first session.

Participant Number: __________________________________________

| 1) Child’s Date of Birth: ___ - ___ - ___ Date of Visit: ___ - ___ - ___ |
|-------------------|-----------------|
| month     day      year | month    day    year |

<table>
<thead>
<tr>
<th>2) Gender:</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Male</td>
<td>□ Female</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3) CI Configuration:</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Unilateral CI:</td>
</tr>
<tr>
<td>□ Bimodal (CI+HA):</td>
</tr>
<tr>
<td>□ Bilateral CIs:</td>
</tr>
</tbody>
</table>

| 4) Date of 1st CI surgery: _____________ Activation: _____________ |

| 5) Date of 2nd CI surgery (if applicable): ____________ Activation: _______ |

**Educational History (please circle)**

| 6) Does your child attend full-time or part-time pre-school? |
|-------------------|-----------------|
| □ Full-time       | □ Part-time     | □ None |

<table>
<thead>
<tr>
<th>7) What type of preschool are they enrolled in?</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Regular preschool classroom</td>
</tr>
<tr>
<td>□ Auditory/oral classroom</td>
</tr>
<tr>
<td>□ Total communication classroom</td>
</tr>
<tr>
<td>□ School for the deaf and hard of hearing</td>
</tr>
<tr>
<td>□ Other:</td>
</tr>
<tr>
<td>___________________________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8) Does your child receive auditory-verbal therapy?</th>
<th>□ Yes</th>
<th>□ No</th>
</tr>
</thead>
<tbody>
<tr>
<td>If so, how many times per week?</td>
<td>_______</td>
<td></td>
</tr>
<tr>
<td>How long is each session?</td>
<td>_______</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9) Does your child attend any other activities to assist with their speech or language development?</th>
<th>□ Yes</th>
<th>□ No</th>
</tr>
</thead>
<tbody>
<tr>
<td>If so, what services, groups, or activities does your child attend?</td>
<td>_______</td>
<td></td>
</tr>
</tbody>
</table>
Music History

10) Has your child ever received music therapy services?  □ Yes         □ No
   If so, how long was the period of services received? _________________________

11) What is your child’s music exposure?
   □ Listens to radio □ Caregiver sings to child □ Mommy-and-me groups
   □ Music Together  □ KinderMusik  □ Other: ____________________________

Child’s Language History

12) What is the first language your child learned: □ English □ Spanish □ ASL
   □ Other: __________________________

13) What is the second language your child learned: □ English □ Spanish □ ASL
   □ Other: __________________________

14) Percentage of language spoken at home

   First language: _____%
   Second language: _____%
   Comments: ____________________________________________________________

15) Percentage of language spoken at school

   First language: _____%
   Second language: _____%
   Comments: ____________________________________________________________

16) Music / TV exposure

   First language: _____%
   Second language: _____%
   Comments: ____________________________________________________________
Appendix G: List of Nursery Rhymes

[Nursery Rhyme 1 – Trochaic 1]
Jack and Jill went up the hill
To fetch a pail of water. (pause)
Jack fell down and broke his crown and
Jill came tumbling after.

[Nursery Rhyme 2 – Trochaic 2]*
Pe-ter, Pe-ter pump-kin ea-ter,
Had a wife but could not keep her,
Put her in a pump-kin shell, and
There he kept her, ve-ry well.

[Nursery Rhyme 3 – Iambic 1]*
The Queen of Hearts she made some tarts
All on a sum-mer’s day; (pause)
She took them to the ci-ty square
And gave them all a-way.

[Nursery Rhyme 4 – Iambic 2]*
What’s that? It’s cheese up on the clock,
The mice ran up to check the stock
The clock struck one, the mouse ran down,
And now there’s no more cheese, he frowned!

*These nursery rhymes have been adapted in order to make the syllable structure alternate directly between stressed and unstressed. Simple, alternating structure facilitates entrainment (Leong & Goswami, 2014), makes the syllable structure identical across nursery rhymes, and decreases complexity, making the words more developmentally appropriate (Gfeller, et al., 2000).
Appendix H: Sample Pictures to Accompany Nursery Rhymes

Trochaic 1 (Jack and Jill)
Appendix I: Session Plan Script

I. Introductory Activity (5 minutes)
   a. Personal introductions: Hi, my name is Jessica, what is your name?
   b. Orientation to the task: Today, we will be playing some word games with the drum! We will talk about how to hit the drum, then drum along with the rhythm of a nursery rhyme, and then play a music game at the very end.
   c. Instructions: To begin, we will talk about how to hit the tubano drum. We will only drum with one hand; which hand are you going to use? When we hit the drum, we take our hand off of the drum right after we hit, like so (demonstrates). Now you try. (Repeat four times)
   d. Practice drumming to stressed syllables
      i. Names: Let’s practice drumming to words by drumming to our names. I’ll start. My name is Jessica, so I would drum Jes-si-ca (hitting the drum on first and third syllables). Try it with me! (repeat four times)
      ii. Now let’s try drumming your name, so say your name first. I think the rhythm of that is (x), let’s try that together! (repeat four times)
      iii. One word utterances: Let’s practice some more! This time, let’s practice to favorite colors! What is your favorite color? I think the rhythm of that is (x), let’s try it together! (repeat four times)
      iv. My favorite color is purple, so let’s try that together! (drumming on first syllable, repeat four times)
         1. Vary utterances week to week: colors, foods, animals, movies
      v. Let’s practice drumming to only one word, even when we have two words in our pattern! We will practice hitting our drum when we say the word that is more important, for example if I say BLUE dog, BLUE dog, I will hit the drum when I say blue, like this (provides example). Try that with me!
         1. Practice BLUE dog BLUE dog BLUE dog BLUE dog.
         2. Practice blue DOG blue DOG blue DOG.
         3. What is your favorite color? What is your favorite animal? Let’s practice drumming to that! (Practice contrastive stress as above, with the participant-selected color and animal.)
   e. The student researcher will track which components of the activity the participant was able to complete successfully on a data collection sheet specific to each participant.

II. Synchronization Assessment (10 minutes)
   a. Student researcher and participant will now be seated across from each other with their own tubano drums. Student researcher will start the recording on Audacity (which records from the sensors placed on the underside of each drum) and will prepare the metronome track on the iPod.
b. *Now, we are going to play a game! I want you to try to hit your drum at the exact same time that I hit mine. I will start drumming first, and then I want you to hit your drum at the same time that I do! Are you ready?*

c. The student researcher will be listening to a metronome track set at 2.5 Hz (20 seconds, 50 drum hits). The student researcher will drum along at the rate of the metronome, heard through an earbud in her ear, while encouraging the participant to drum on his/her drum at the same time.

d. When finished, the student researcher will ask the participant, *Ready to try that one more time?* and then replay that track.

e. If the child is unable to synchronize during the first trial of each speed, the student researcher will provide hand-over-hand assistance for the second trial.

f. The student researcher will stop and save the recording for the two 2.5 Hz trials, and then start a new recording.

g. The student researcher will repeat the previous two steps for a second speed, 1.67 Hz (20 seconds, 33 drum hits), again with two trials. These parameters are based upon Kirschner & Tomasello (2009).

h. The student researcher will say, *Great job! Thanks so much for drumming along with me!* and save the recording of the 1.67 Hz trials.

i. The student researcher will track which components of the activity the participant was able to complete successfully on a data collection sheet specific to each participant.

III. Drumming to Speech (10 minutes)

a. The student researcher will present the syllable structure of one of the four nursery rhymes, but will use neutral syllables. The student researcher will listen to a metronome through one earbud at a rate of 90 bpm (1.5 Hz) to ensure maintenance of temporal structure in relation to the recorded stimuli.

   i. Instructions: *Now, we are going to practice drumming along to poetry! I am going to say some silly nonsense words, but I want you to listen for the rhythm in my words. Just listen, don’t tap the drum for now.* Recite syllable structure of the nursery rhyme, following template below:

   ii. Trochaic: DEE da DEE da DEE da DEE da

   iii. Iambic: da DEE da DEE da DEE da DEE da

b. Instructions: *Now, let’s tap on our knee to the rhythm of those nonsense words. Listen to my words and tap your hand on your knee, just like I am doing.* Repeat four times, and aid participant in tapping with hand-over-hand assistance, if necessary.

c. Instructions: *Now, we will listen to the actual nursery rhyme! Follow along with these pictures.* The student researcher will read aloud the nursery rhyme to the participant at a metronome rate of 90 bpm, pointing to the respective pictures for each line of the nursery rhyme.

d. Instructions: *This time, watch when I tap on my drum to the rhythm of the rhyme.*
e. The student researcher will recite nursery rhyme with the metronome, and will tap to each stressed syllable in a line, following the structure in Session Script.

f. The second time, the student researcher will say, Now, tap on your drum with me, just like I am doing! The student researcher will provide verbal or non-verbal assistance if child cannot perform this task alone. Repeat three times.

g. Instructions: Now, let’s play the rhythm of the whole rhyme on the drum! Listen closely to me and try to hit the drum when I do.
   i. The student researcher will then present the nursery rhyme four times with short (two second) breaks between repetitions.
   ii. The student researcher will provide verbal or non-verbal assistance if needed.

h. Instructions: Great job! Now, I want you to say the nursery rhyme with me. I will say one line, and then I will say it again, the second time with you! Try to say the words at the same time that I do.
   i. The nursery rhymes will be taught in three steps; the goal with each participant is to reach the third step, but the student researcher will remain at whichever level the child is able to achieve.
      1. Step one is repeating the nursery rhyme word by word (ex. student researcher says Jack, participant and student researcher say Jack together).
      2. Step two is repeating the nursery rhyme by small chunks, or half lines (ex. student researcher says Jack and Jill went, participant and student researcher say Jack and Jill went together).
      3. Step three is repeating the nursery rhyme line by line (ex. student researcher says Jack and Jill went up the hill to, participant and student researcher say Jack and Jill went up the hill to together).

   ii. This process will occur two times for each line of the nursery rhyme, so that the participant says 8 total lines of synchronized speech.

i. The student researcher will track which components of the activity the participant was able to complete successfully on a data collection sheet specific to each participant.

IV. Closing activity (5 minutes)

a. Rhythm game
   i. Instructions: I am going to play a rhythm, and I want you to play it back to me! Listen while I play. Try to make your rhythm sound exactly like mine. The student researcher will present ten, simple rhythmic patterns consisting of between two and four notes.
   ii. The student researcher will prompt participant to play with rhythmic speech: My turn first! (researcher plays rhythm) Now your turn! (participant mimics rhythm)

b. Goodbye Chant
i. Instructions: Great job drumming along to the rhymes today! To end our session, we are going to sing and play a goodbye song with our drum. When I point to you, I want you to hit the drum and say “Bye!”

ii. The student researcher will cue participant when to hit the drum and say “Bye!”

c. The student researcher will track which components of the activity the participant was able to complete successfully on a data collection sheet specific to each participant.
Appendix J: Sample At-Home Practice Sheet

Hello!

This week, your child practiced the rhythm of the nursery rhyme “Jack and Jill.” Please plan to practice this with your child at home once a day so that s/he can better learn the rhythms!

Your child can “drum” to the rhythm of the rhyme on any surface, but here are some suggestions:
- A big bucket turned upside down
- A pot turned upside down
- A table

This week, please assist your child in following the instructions for the track “Jack and Jill.” This should take about 5 minutes. The words are printed below for your convenience.

Jack and Jill went up the hill
To fetch a pail of wa-ter. (pause)
Jack fell down and broke his crown and
Jill came tumbling after.

Please mark below which days your child played along with the track, and on what “drum” your child practiced the rhythm (i.e., pot, bucket)! Thank you for your help in this study!

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix K: Sample Script for At-Home Practice Tracks

Example script for Track 1:

Hello! I’m excited to play the drum with you today! Today we will be practicing the nursery rhyme “Jack and Jill.” Listen carefully the first time around. Don’t touch the “drum” yet!

[Nursery Rhyme, said to metronome at 120 bpm as in session protocol]

Great listening! This time, tap on your knee to the sounds that are louder and longer. Mom or dad, please help with this part! Your child should just be tapping their knee on the stressed syllables.

[Nursery Rhyme]

Great tapping! This time, tap on your pretend drum in front of you to the sounds that are louder and longer. You will be drumming along to the same sounds as last time, but on the pretend drum this time. Parents, please help with this part! You will hear the nursery rhyme twice, please play along both times.

[Nursery Rhyme x2]

Great drumming! Thanks for playing with me today, see you again tomorrow!
Appendix L: Data Collection Sheet

Participant Number: ______________

Pre-Test Scores

*Audie (order: __*)*
- Melody: ______
- Rhythm: ______

*PEPS-C (order: __*)*
- Emotional: ______
- Contrastive Stress: ______

First Session:  ____ / ____ / ____

Second Session:  ____ / ____ / ____

Third Session:  ____ / ____ / ____

Fourth Session:  ____ / ____ / ____

Fifth Session:  ____ / ____ / ____

Post-Test Scores

*Audie (order: __*)*
- Melody: ______
- Rhythm: ______

*PEPS-C (order: __*)*
- Emotional: ______
- Contrastive Stress: ______

Session Tracking
On the next page, indicate the level of completion achieved by the participant.
For Synchronized Speech, circle the highest level the participant was able to complete successfully.

C = Completed without assistance
CWA = Completed with hand-over-hand assistance
NC = Not completed
<table>
<thead>
<tr>
<th>Activity 1: Introduction</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Names</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-word utterance</td>
<td>Word:_____</td>
<td>Word:_____</td>
<td>Word:_____</td>
<td>Word:_____</td>
</tr>
<tr>
<td>Contrastive stress</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity 2: Synchronization</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-paced</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-150 BPM</td>
<td>Trial 1: __</td>
<td>Trial 1: __</td>
<td>Trial 1: __</td>
<td>Trial 1: __</td>
</tr>
<tr>
<td></td>
<td>Trial 2: __</td>
<td>Trial 2: __</td>
<td>Trial 2: __</td>
<td>Trial 2: __</td>
</tr>
<tr>
<td>-100 BPM</td>
<td>Trial 1: __</td>
<td>Trial 1: __</td>
<td>Trial 1: __</td>
<td>Trial 1: __</td>
</tr>
<tr>
<td></td>
<td>Trial 2: __</td>
<td>Trial 2: __</td>
<td>Trial 2: __</td>
<td>Trial 2: __</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity 3: Nursery Rhymes</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral syllables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nursery Rhyme</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synchronized Speech</td>
<td>Word by word Chunks</td>
<td>Word by word Chunks</td>
<td>Word by word Chunks</td>
<td>Word by word Chunks</td>
</tr>
<tr>
<td></td>
<td>Line by line</td>
<td>Line by line</td>
<td>Line by line</td>
<td>Line by line</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity 4: Closing</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhythm game</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodbye Chant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
Appendix M: Adult Consent Form

University of Miami
CONSENT TO PARTICIPATE IN A RESEARCH STUDY

The effect of a drumming-to-speech intervention on prosody perception of preschool-aged children with cochlear implants

Jessica MacLean, Graduate Student
Masters of Music Therapy Student, University of Miami

Kimberly Sena Moore, Ph.D.
Assistant Professor of Professional Practice, Music Therapy
Frost School of Music, University of Miami

The following information describes the research study in which your child is being asked to take part. Please read the information carefully. At the end, you will be asked to sign if you agree to allow your child to take part in the study.

PURPOSE OF STUDY:

You are being asked to take part in a research study. The purpose of this study is to understand how drumming can help your child understand speech.

Your child is being asked to be in the study because he/she is between the ages of 3 and 5 and uses a cochlear implant.

PROCEDURES:

This study will have four steps: 1) Meeting with parent(s)/legal guardian(s) and child; 2) Music sessions with the child; 3) At-home practice; and 4) Final session.

Meeting with parent(s)/legal guardian(s) and child

1. This study will recruit 30 to 40 participants.
2. Parent(s) and/or legal guardian(s) will read the consent forms and decide whether they want to allow their child to be in the study.
3. The student researcher will explain to the child what they will be asked to do, and ask the child if they want to be in the study.

After getting your consent:

4. Parent(s) and/or legal guardian(s) and the child will meet with the student researcher at one of three locations chosen by the parent: University of Miami Ear Institute, Debbie School, or the Frost School of Music at the University of Miami Coral Gables campus.
5. Parent(s) and/or legal guardian(s) will be asked to fill out a form to provide information about their child like age, name, and gender.
6. Parent(s) and/or legal guardian(s) will be asked to sign a research agreement which allows the student researcher to access information about the child’s cochlear implant.
(such as age at amplification, implant type, and dates of CI surgeries) and the child’s speech and language scores.

7. It should take about 15 minutes to complete the forms.
8. It should take about 15 minutes to complete the forms. While the parent completes these forms, the student researcher will take the child into a separate room to complete a music session.

**Music session with the child**

1. During the first session, the child will complete two assessments. The first assesses how your child listens to music and the second assesses how your child listens to prosody, or the inflection in speech. This will later help the student researcher find out if the music sessions helped your child to understand speech better. The expected time of these assessments is one hour.

2. After completing the assessments, the student researcher and child will start the first music session. All music sessions will begin with the student researcher checking to make sure the CI is turned on and the sound loud enough for the child. Music sessions will last about 40 minutes and include four drumming activities.
   a. The first activity will include drumming to words and short rhythm patterns.
   b. The second activity will involve the child drumming on a drum at the same time as the student researcher at a regular pace.
   c. The third activity will involve the child drumming on the drum to nursery rhymes and saying the nursery rhymes at the same time as the student researcher.
   d. The final activity will include repeating simple musical patterns on a drum and drumming along to a “goodbye” song.

3. There will be a total of four 30-minute music sessions over four weeks. The first music session will happen on the same day as the first assessments. The other sessions will occur weekly on the same day and time. Each week the child will practice a new nursery rhythm. All the children will have the same music sessions, but the order they learn a new nursery rhythm is randomly assigned. Random assignment occurs by chance, like flipping a coin.

**At-home practice**

1. After the first three music sessions, the child will be asked to complete a short practice session every day using CD recording and drum provided by the student researcher.
2. The parent will start the CD track for the child. The track will guide the child through five minutes of drumming to the nursery rhyme practiced in that week’s music session.
3. The parent will be asked to mark the days when the child practiced on a sheet of paper provided by the student researcher.

**Final session**

1. One week after the final music session, the participant will be asked back to complete the final assessments. These will again assess how your child hears music and speech in order to see if the music sessions helped your child to understand speech better.
2. This final assessment session will take about one hour.
RISKS AND/OR DISCOMFORTS:

We do not expect that your child will have any personal risk or discomfort from taking part in this study. This research is considered to be minimal risk. This means that the risks of taking part in this research study are not expected to be more than the risks in your child’s daily life. There are no other known risks to your child if you choose to take part in the study.

There may be uncommon or previously unknown risks. You should report any problems to the researcher. You may skip any question you do not wish to answer on the forms.

BENEFITS:

It is possible that your child will benefit from this study by participating in a music therapy experience. Your child may benefit from this study by showing improved understanding of music and speech, which may in turn improve your child’s communication with others. Your involvement in the study may also help other children who use cochlear implants. The study will help us learn how music can be used as a way to help children with cochlear implants improve their understanding of certain areas of speech.

ALTERNATIVES:

The parent(s) or legal guardian(s) have a right to not allow their child to be in the research study. If you and your child decide not to join the study, their current treatment and future treatment options at the University of Miami Ear Institute or the Debbie School will NOT be impacted.

CONFIDENTIALITY:

The student researcher will do her best to keep the information you share confidential; however, it cannot be absolutely guaranteed. Individuals from the University of Miami Institutional Review Board (a committee that reviews and approves research studies), Research Protections Programs, and Federal regulatory agencies may look at records related to this study to make sure we are doing proper, safe research, and protecting human participants. The results of this research may be published or presented to others. You will not be named in any reports of the results.

To ensure confidentiality, the student researcher will lock all paper document data collected on you (e.g. responses to questions) in a locked filing cabinet. All data collected will be assigned a code number. The master list of participant names and code numbers will be held separate from the raw data by the student researcher in a locked filing cabinet. All electronic data will be stored on the University of Miami’s Box cloud storage system (like the Google drive and Dropbox systems). Access to the data will require a University of Miami userid and password and will only be able to be accessed by the researcher(s).

If you withdraw from the study at any time during the study procedures, your data will not be kept. It will be destroyed once you withdraw.

By signing this consent, you authorize the researcher(s) to access all information gathered throughout the study as may be necessary for purposes of this study.

COSTS:
There are no costs associated with your child’s participation in this study.

Although injury is unlikely, if injury should occur, treatment will in most cases be available. If your child has insurance, your child’s insurance company may or may not pay for these costs. If your child does not have insurance, or if your child’s insurance company refuses to pay, you will be expected to pay. Funds to compensate for pain, expenses, lost wages and other damages caused by injury are not routinely available.

**COMPENSATION:**

There will be no compensation for participation in this research study.

**RIGHT TO DECLINE OR WITHDRAW:**

Your child’s participation in this research study is voluntary. Your child is free to refuse to be in the study or withdraw his/her consent at any time during the study. Your child’s withdrawal or lack of participation will not affect the treatment they are receiving at the University of Miami Ear Institute or the Debbie School. The researcher reserves the right to remove your child without your child’s consent at such time that they feel it is in the best interest for your child.

**CONTACT INFORMATION:**

Jessica MacLean (317-430-2399; jam667@miami.edu) under the supervision of Dr. Kimberly Sena Moore (305-284-3943; kms132@miami.edu) will gladly answer any questions you may have concerning the purpose, procedures, and outcome of this project. If you have questions about your rights as a research subject you may contact Human Subjects Research Office at the University of Miami, at (305) 243-3195.

**PARTICIPANT AGREEMENT:**

I have read the information in this consent form and agree (to allow my child) to be in this study. I have had the chance to ask any questions I have about this study, and they have been answered for me. I am entitled to a copy of this form after it has been read and signed.

____________________________
Name of Child

____________________________
Name of Parent/ Legal Guardian

____________________________                               __________________
Signature of Parent/ Legal Guardian       Date

____________________________
Name of Person Obtaining Consent

____________________________                               __________________
Signature of person obtaining consent                                    Date
Appendix N: Child Assent Description and Script

Project title: The effect of a drumming-to-speech intervention on prosody perception of preschool-aged children with cochlear implants

Written child assent will not be sought due to the age of the child participants (aged 3-5). Instead, child assent to participate will be obtained at regular intervals verbally, through child’s participation in the drumming-to-speech sessions, and through the student researcher’s continual monitoring of child behavior’s that indicate assent. The following script will be utilized at the beginning of the assessment session to determine child assent to participate in the assessment session:

*My name is Jessica, and I am learning about music and speech. I would like your help to play a music game and a computer picture game with me. You will show me what you know about music and speech with these games. This will take about one hour. We will go to [insert location of where the assessment will take place at the site] to play these games. If you don't feel like playing, you don't have to. You can stop at any time and that will be okay. Do you want to join me?*

Child assent to continue to participate in the study will be determined verbally at the end of the assessment session using the following script:

*Thank you for joining me and helping me learn about what you know about music and speech. I would like you to join me again for some music. We will be playing music games, drumming on a drum, and drumming along with nursery rhymes. Each music session will take about 30 minutes and they will help me learn more about how music helps your speech. If you don’t feel like being in music sessions, you don’t have to. You can stop at any time and that will be okay. Do you want to be in the music group?*

If a child does not agree to participate in or does not complete an assessment session, the student researcher will ask the child if he or she would like to try again another day. If the child says or indicates “yes,” the assessment session will be rescheduled. If the child says or indicates “no,” then it will be determined that the child has indicated a desire not to participate and he or she will be removed from the study.

Child assent to continue study participation will be obtained on a session-by-session basis prior to each drumming-to-speech session and through the researcher’s continual monitoring of child behaviors that indicate assent. Participating children will indicate assent to participate verbally (e.g. stated “yes”) or nonverbally (e.g. smiling, nodding “yes”) and by following the student researcher’s directions to transition out of the therapy rooms (e.g. waiting by the door, standing next to the student researcher). If a participating child does not indicate a willingness to participate in music, the student researcher will initially attempt techniques such as redirection or positive reinforcement to encourage the child to join music. If the participating child continues to exhibit an unwillingness to join the music group, the child will be given the opportunity to return to their parent/legal guardian and will resume with the next scheduled music session.