Infant vocal imitation of music

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the
Graduate School of The Ohio State University

By

Lucia Pereira da Cruz Benetti

Graduate Program in Music

The Ohio State University

2017

Master's Examination Committee:

Dr. Eugenia Costa-Giomi, Advisor

Dr. Julia Shaw

Dr. Laura Wagner

Copyrighted by
Lucia Pereira da Cruz Benetti
2017

Abstract

Infant vocalizations have been described mostly from the perspective of language acquisition. There is little systematic research on infant production of music sounds. This study investigated infants' music production by identifying and analyzing vocal imitation of music produced by one infant throughout one day. One 15-month-old infant wore a portable recording device that captured 16 continuous hours of sounds produced within hearing distance of the infant, as well as sounds from the infant's own vocal production. Instances of music imitation were identified through extensive and intensive listening of the audio file. Physical parameters of the imitative vocalizations were collected through acoustic analysis. A selection of imitations was presented to adults in a validation study involving perceptual similarity judgments between the infant's imitation and the model that was imitated. Results from acoustic analyses and the validation study supported the perceptual identification of the imitations.

The findings show that the infant imitated music that he had heard that day. The infant imitated music features such as pitches, intervals, and rhythms of songs that were sung to him and of a melody produced by a toy. Some imitations occurred many hours after the infant had heard the melodies that served as models. These results reveal the ability of infants to (a) develop accurate representations of music melodies, (b) recall such representations at will hours later, and more importantly, (c) express these music

representations vocally. In summary, the findings show that infants are capable of producing music imitations.

Para Lalá – eles passarão, tu passarinho

Acknowledgments

Eugenia, for being inspiring and for taking care of me in so many ways;
Zoltán Koppányi, for being a pleasant surprise; and
My parents, Antonio and Silvia, and my sister, Fernanda, because most of this work is yours.

I would also like to thank the following individuals for their unconditional support: Deborah Goldenberg, LaRae Copley, Luis Cruz-Ortega, Luís Felipe Eick, Mayra Pacheco Pachado, and the Music Education faculty and graduate students at the Ohio State University.

Finally, thank you to James and his family. He will probably never know that a stranger spent so many months listening to his babbling, analyzing it, and admiring it.

Vita

2012.....B.M. Piano Performance, Universidade
Federal do Rio Grande do Sul, Brazil

August 2015 to May 2016.....Graduate Research Associate, School of
Music, The Ohio State University

May 2016 to August 2016Graduate Research Associate, Center for
Cognitive and Brain Sciences, Department
of Psychology, The Ohio State University

August 2016 to presentGraduate Teaching Associate, School of
Music, The Ohio State University

Fields of Study

Major Field: Music

Table of Contents

Abstract.....	ii
Dedication.....	iv
Acknowledgments.....	v
Vita.....	vi
Chapter 1: Introduction.....	1
Purpose of the Study	6
Chapter 2: Review of Literature	7
Infant Vocal Development	7
Vocal Imitation in Infancy	10
Infants and Music	12
Infant Music Production.....	14
Chapter 3: Method	22
Participant.....	22
Data Collection.....	23
Perceptual Identification of Imitations.....	24

Description of the Models and the Imitations	26
Other Data Collected	29
Acoustic Analysis.....	29
Validation: Perception Study with Adults.....	31
Chapter 4: Results	32
Music and Language Environment and Language Production	32
Music Environment	32
Language Environment and James' Language Production	36
Overview of Imitations.....	37
“Rain Rain Go Away”	38
Mother singing “Rain Rain Go Away” (12:25 p.m.).....	38
Infant Imitation and Father Singing “Rain Rain Go Away” (6:40 p.m.).....	42
“Happy Birthday”	54
“Happy Birthday” Melody from Electronic Toy (8:40 a.m.)	54
Infant Imitation of “Happy Birthday” (6:45 p.m.).....	55
Music Intervals.....	60
Imitation of Sister's Improvised Melody (8:07 a.m.).....	61
Imitation of Brother's Music Interval (8:12 a.m.)	63
Chapter 5: Validation Study.....	67

Method	69
Participants	69
Procedures	70
Results and Discussion.....	73
“Happy Birthday” Imitation	73
“Rain Rain Go Away” Imitation	76
Imitation of Brother’s Interval.....	79
Chapter 6: Discussion and Conclusion	82
Results	83
“Rain Rain Go Away” Imitation	83
“Happy Birthday” Imitation	83
Imitation of Music Intervals	84
General Discussion.....	86
Production of Music Sounds	86
Perception of Melodies.....	88
Mechanisms Involved in Vocal Imitation	93
Motivations for Imitating and Making Music	95
Limitations of the Study	97
Implications.....	99

Conclusion.....	100
References.....	102
Appendix A: Sequence of Events During the Day	112
Appendix B: Validation Study with Adults	114

List of Tables

Table 1. Characteristics and description of models imitated by the infant	27
Table 2. Characteristics and description of infant's imitations	28
Table 3. Infant's music environment and duration of exposure to music	35
Table 4. Frequency (Hz), duration (s), and intensity (dB) of infant's "Rain Rain Go Away" sounds	50
Table 5. Test items and response options	71
Table 6. Characteristics of audio excerpts presented to participants in the "Happy Birthday" question	74
Table 7. Participants' response choices for the two presentations of the "Happy Birthday" item	76
Table 8. Characteristics of audio excerpts presented to participants in "Rain Rain Go Away" imitation questions.....	77
Table 9. Participants' response choices for the "Rain Rain Go Away" comparisons	78

List of Figures

Figure 1. Music notation of mother singing "Rain Rain Go Away" (perceptual analysis)	40
Figure 2. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of mother's pitches from acoustic analysis of mother singing "Rain Rain Go Away"	41
Figure 3. Spectrogram of mother singing "Rain Rain Go Away"	42
Figure 4. Representation of mother's pitches using Melodyne	42
Figure 5. Music notation of infant's "Rain Rain Go Away" imitation (perceptual analysis)	44
Figure 6. Music notation of father singing "Rain Rain Go Away" (perceptual analysis)	45
Figure 7. Music notation of infant's imitation of father and father singing second part of "Rain Rain Go Away" (perceptual analysis)	46
Figure 8. Spectrogram of infant's vocalization that preceded the "Rain Rain Go Away" imitation and the beginning of the "Rain Rain Go Away" imitation	48
Figure 9. Spectrogram of infant's "Rain Rain Go Away" imitation	49
Figure 10. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of infant's pitches from acoustic analysis of the infant's "Rain Rain Go Away" sounds	52

Figure 11. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of father's pitches from acoustic analysis of father singing "Rain Rain Go Away"	52
Figure 12. Spectrogram of last measure of father singing "Rain Rain Go Away" and beginning of infant's imitation of father	53
Figure 13. Music notation of "Happy Birthday" melody from electronic toy	54
Figure 14. Music notation of infant's "Happy Birthday" imitation (perceptual analysis)	55
Figure 15. Spectrograms of infant's "Happy Birthday" imitation (original audio and edited audio)	57
Figure 16. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of pitches from acoustic analysis of infant's "Happy Birthday" imitation	58
Figure 17. Formant structure of infant's "Happy Birthday" imitation	60
Figure 18: Music notation of sister's improvised melody (perceptual analysis).....	62
Figure 19. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of pitches from acoustic analysis of sister's improvised melody and infant's imitation	63
Figure 20. Spectrogram of brother's interval and infant's imitations; and duration (s), frequencies (Hz), and equivalent pitch based on equal temperament tuning (rounded to the nearest quarter tone).....	65

Chapter 1: Introduction

I listened to the babbling of a baby nearly every day for four months. The babbling came from a 16-hour audio file of James, a 15-month-old infant who had worn a portable recording device in his pocket for an entire day. The recording had captured his world of sounds: the father trying to get him to say “bacon”, his older brother and sister screaming in the background, the mother singing to James in the sweetest tone of voice. And, of course, the recording device also captured James in all his vocal eloquence, making sounds that were mostly unintelligible.

To be fair, his sounds were unintelligible only at first. After some experience listening, I realized that James was extremely expressive. The variations in timing, pitch, duration, and especially timbre contained a wealth of information; I just had to learn to listen. Experiencing the qualities of the sounds, as opposed to identifying them as *something* – a certain word, a song, rain falling or a car backfiring – required a very special type of attention. It was overwhelming to make sense of all the sounds that surrounded him. After half an hour of listening to the audio file I would be exhausted.

When I began listening to the audio file, the initial plan was to focus on the music that was produced in the household by the family members, not the baby’s vocalizations. My focus soon shifted because I became fascinated with James’s own sounds. It was an entirely new world. I would listen over and over again to seconds of recording to try to

understand them – I had no visual clues at all, and no linguistic information either. Was he playing with a toy? Was anyone near him? Was he calling out to someone, or was he vocalizing for himself? The sounds that he made when by himself were remarkable. Here was this unknown baby vocalizing alone, and I had no idea what he was doing, but he seemed to be having fun most of the time. As a matter of fact, he seemed rather happy making sounds when no one was listening.

With time, a few things became evident about James's vocal production. First, many of the sounds James produced when he was alone were different from those he produced around others. In fact, many times his private sounds were more interesting and varied than his sounds in social interactions: When he was around other people, his sounds seemed to become shorter and speech-like whereas when he was alone, he experimented with different pitches, rhythms and timbres. Secondly, not all sounds sounded communicative or expressive. In other words, not all sounds sounded like he wanted to engage with those around him. This was striking for me, because the studies I read on infant vocalizations mostly emphasized the role of vocal production for the development of communication skills.

Lastly, it became evident that James imitated things that he heard a lot. Many times, the imitations were less salient – perhaps just an imitation of a falling contour or an isolated vowel. Other times, though, the imitations were elaborate and beautiful. My biggest surprise came near the end of the recording, when he produced a sequence of clearly pitched sounds that to my ears sounded like music. In the initial notes that I took while familiarizing myself with the recording, I described the vocalization with the

following words: “J sounds – this really sounds like happy birthday.” Only after a while did I remember that 10 hours earlier he had played with a toy that produced that melody.

The process of listening to James’s sounds, thinking about what the different sounds expressed, trying to interpret them, and identifying the imitations led me down rabbit holes. They were charming black holes, however, and Dr. Costa-Giomi was kind enough to listen to me ramble on about them as I read something new from a wide range of literature from unfamiliar disciplines.

One rabbit hole was related to intention. Recognizing an imitation required a judgment of perceptual similarity between the model that was supposedly imitated and the infant’s vocalization in response. James’s imitations were not identical to the models, though, which is to be expected considering that infants have a restricted repertoire of sounds and limited cognitive resources. But if the imitation is not perfect, how can I tell if James really responded to the model, as opposed to vocalizing random sounds? Does it matter if James intended to imitate? If I knew for sure that James intended to imitate a sound, however, then I could try to find the similarities between the model and the imitation. Clearly, this was not possible: I could not ask James first and analyze afterwards. As a listener, then, I could only assume that a vocalization was an imitation if the imitation was successful to my ears.

Another rabbit hole was related to music and language. When I began listening to James, I tried to classify his sounds in some way – sounds that were more musical, more prosodic, more syllable-like, more reflexive. It turned out to be impossible to label the vocalizations, though, because the more I classified the more I needed to define my

categories precisely. I couldn't define my categories without establishing what parameters I would use to define them. And therein lies the problem. Perceptually, it was not so complicated to say that a certain vocalization sounded more like music, or more like prosody, or like both (or like nothing, and then going back to listening again). But my perceptual experience corresponded to what features of the sound, exactly? And could these features be reliably measured? Thinking about the functions of the sounds was certainly very important, but the labels were unconvincing. Furthermore, my judgments tended to change the more I listened.

The literature did not help. There seems to be no objection to the fact that infant vocal production is musical: Infants play around with pitch contours, they use regular rhythms in their babbling, and they can be trained to match specific pitches. However, what is meant by musicality in the context of infant vocalizations depends on the framework provided by the different authors. Accounts of language development tend to emphasize the role that these elements play in learning to speak and to communicate: for instance, pitch contours can communicate specific internal states to the caregiver (Esteve-Gibert & Prieto, 2013), rhythmic babbling helps organize articulatory movements (Oller, 2000), and variations in pitch help with phonation (Oller, 2000). Accounts related to music production tend to refer to musical features without explaining why these musical features are music and not something else (e.g., Papoušek & Papoušek, 1981).

Furthermore, the same terminology is used differently depending on the discipline. Many music-oriented studies, for instance, refer to musical elements of vocalizations (e.g., Papoušek & Papoušek, 1981), but so do many language-oriented ones

too. Fernald (1989) concluded that in infant-directed speech, “the melody carries the message” (p. 1505) from the caregiver to the infant. The ambiguity of terms like *melody* and *musical* created a problem when analyzing James’s sounds. For example, when James imitated a melody of infant-directed speech, does his vocalization contain musical features? Would Papoušek and Papoušek (1981) describe the vocalization as containing musical features? And when James imitated a music melody, how is the musicality of this music imitation different from the musicality of a prosodic imitation? What does it mean to sound musical, as opposed to sounding like music?

These rabbit holes were very frustrating, but they were inevitable. We know very little about infant production of music, and there are almost no systematic investigations on the topic. The exploratory nature of this study required an open mind, but it also required a meaningful foundation from which to observe the phenomenon. As it turned out, the rabbit holes were helpful for creating a framework for understanding music production in infancy.

First, to begin to describe infant production of music, it was necessary to distinguish between vocal sounds that are *musical* and vocal sounds that are *music*. I will use the term *music* to describe musical vocal production that not only expresses musical features, but also organizes these features according to patterns and structures traditionally associated with music systems. Secondly, to begin to describe infant music production with more precise, music-related parameters, analyzing vocal imitations of music seemed like a good starting point. By identifying the model sound and the infant’s

imitative response, it was possible to compare the music features of the model and the music features of the imitation.

Purpose of the Study

The present case study investigated an infant's production of music, specifically his vocal imitations. I analyzed instances of vocal imitation of music by one infant recorded throughout an entire day. The infant wore a portable recording device for 16 continuous hours at home with his family. The device generated an audio file that documented all sounds produced within hearing distance of the infant, as well as all sounds produced by the infant himself. The advantages of this methodology included: (a) The infant was recorded in a naturalistic setting and in unstructured contexts. His vocalizations were spontaneous, and the responses by family members were not pre-established. (b) The recording device captured the infant's vocal production in private, when he was alone, as well as his production in the presence of others. (c) The extended recording time allowed for listening to the infant's vocal behavior in a variety contexts and arousal states. (d) By recording the infant at home with his family, I gained access to the music production of the family, including all music produced for the infant, as well as any other music to which the infant was exposed.

The study addressed the following questions: (1) Does the infant imitate music? If so, (2) in what interaction contexts does he imitate music (i.e., parent-infant dyad, communicative interaction with family members, in the presence of others but without communicative interaction, or alone), and (3) what features of the music he hears does he imitate?

Chapter 2: Review of Literature

This chapter presents an overview of (a) infant vocal development, (b) vocal imitation in infancy, (c) music in infancy, and (d) music production in infancy.

Infant Vocal Development

A developmental milestone in infancy is the onset of speech. Before they utter their first words at around 12 months, prelinguistic infants are already producing a variety of sounds that incorporate phonetic elements of speech. The continuity between infant babbling and meaningful speech has been the focus of much work (e.g., Boysson-Bardies, Halle, Sagart, & Durand, 1989; Nathani, Ertmer, & Stark, 2006; Oller, 2000; Oller et al., 2013; Vihman, Macken, R. Miller, Simmons, & J. Miller, 1985). Although there is considerable variation as to when and how children produce their first words (Vihman, 1993), there is a general agreement that infant vocal production follows a rough developmental timeline (for a review, see Kuhl et al., 2008, and Nathani et al., 2006; for a comprehensive discussion of vocal development, see Oller, 2000). During the first couple of months, infants' vocalizations consist mostly of sounds that are not considered to be direct precursors to speech. These non-speech sounds include vegetative sounds, such as burps and hiccups, and fixed signals, such as crying and laughter. As infants gain more vocal tract control and experience with the sounds of language, they gradually begin to produce vocalizations that sound more and more like speech, such as closants

(consonant-like sounds) and vocants (vowel-like sounds). Around the third month of life, infants produce vocalizations that resemble adult vowels, and at 8 months, they produce consonant-vowel (CV) combinations that sound like syllables of human speech. Because these syllables approximate mature forms, they are referred to as *canonical*. Infants frequently produce canonical syllables in sequences, a type of vocal behavior called canonical babbling. Starting at around 10 months, infants' consonant and vowel sounds become increasingly similar to the specific phonetic categories of their native language (Kuhl et al., 2008; Vihman & Boysson-Bardies, 1994). This process is thought to follow the perceptual narrowing of phonetic discrimination (Kuhl et al., 2008), in which infants lose their capacity to discriminate speech sounds in all languages because of their increasing focus on the speech contrasts of their native language (Werker & Tees, 1984).

Although infants' vocal production becomes progressively more complex throughout the first year of life, it is noteworthy that infants do not stop producing earlier types of vocalizations (Oller, 2000). For example, Nathani, Ertmer, and Stark (2006) found that more than half of the vocalizations produced by 13- to 15-month olds were vowels, even though by this age infants were already well capable of producing CV syllables and had been doing so for around six months.

In addition to learning to produce consonant and vowel sounds, infant vocal development also involves learning prosodic¹ features of speech, such as intonation and

¹ According to Fletcher (2010), "prosody or *prosodic* features are . . . synonymous with variations in *suprasegmental* parameters such as duration, intensity, and f_0 that contribute in various combinations to the production and perception of stress, rhythm and tempo, lexical tone, and intonation of an utterance. In other words it refers to those phonetic and phonological aspects of spoken language that cannot necessarily be reduced to individual

stress patterns. Even before they are producing meaningful words, infants' utterances already reflect intonation and phrasing organization typical of spoken language (Lynch, Oller, Steffens, & Buder, 1995; Snow, 2007). In addition to patterns of prosody common to many languages, infants also incorporate the characteristic pitch variations, voice quality, duration, and stress patterns of their native language into their babbling sounds (Boysson-Bardies, Sagart, & Durand, 1984).

Infants not only learn to produce speech-related sounds, but they also learn to express communicative intent and specific affective states through variations in pitch contours. Such communication seems to be partly independent of both phonetic content and prosodic features of human speech. For example, Papaeliou and Trevarthen (2006) found that prelinguistic infants' utterances when playing with their mothers had different pitch patterns compared to the utterances produced when infants were holding and exploring an object without maternal engagement. Further supporting the distinction between communicative and non-communicative utterances, Esteve-Gibert and Prieto (2013) found that the pitch patterns of infant vocalizations differed based on the specific communicative intention of the infant: Utterances tended to have wider pitch range and

consonants and vowels but generally extend across several segments or syllables" (p. 523). Prosody is described by many as musical; for example, Snow and Balog (2002) write that "speech prosody involves the 'musical' aspects of language – variations in the pitch, length and loudness of spoken utterances" (p. 1026). In particular, the term *intonation* refers to patterns of pitch changes that serve "to contrast and to organize words and larger utterances" (Beckman & Venditti, 2010, p. 603). Intonation is an important element of expression. Intonation patterns have grammatical functions (e.g., a pitch contour marking a phrase boundary), pragmatic meanings (e.g., a pitch contour expressing the communicative intent of the speaker, such a falling pitch contour indicating a command), and attitudinal meanings about the speaker's feelings regarding what is being said (Snow & Balog, 2002).

longer duration when the infant was requesting an object or expressing discontent, compared to the pitch contours of utterances in which the infant responded to the mother.

Infant vocal development has been studied mostly from the perspective of language acquisition. It is not clear how infants' vocal development relates to music production. Many of the speech production skills that infants acquire during the first year, such as regulation of fundamental frequency (Kent & Murray, 1982) and control of duration (Lynch et al., 1995), are also related to the skills necessary to produce sequences of sounds with the temporal and pitch organization of music. Although not much is known about infants' vocal behavior in terms of music, it is not unreasonable to assume that infants develop the necessary production skills and become capable of vocalizing meaningful sounds for music just as they do for language.

Vocal Imitation in Infancy

Imitation of actions produced by others is thought to be an important mechanism by which infants learn new behaviors (Jones & Herbert, 2006; Meltzoff, 1988; Meltzoff & Moore, 1977, 1983) and acquire social and cultural knowledge (Csibra & Gergely, 2009, 2011; Gergely & Csibra, 2006; Király, Csibra, & Gergely, 2013; Ray & Heyes, 2011). Infants have been shown to imitate body gestures, such as mouth and finger movements (Meltzoff & Moore, 1977, 1983), as well as novel acts, such as illuminating a light box by touching a panel with one's head (Gergely, Bekkering, & Király, 2002; Király et al., 2013; Meltzoff, 1988) and leading a toy animal into a toy house (Schwier, Maanen, Carpenter, & Tomasello, 2006; Southgate, Chevallier, & Csibra, 2009). In some cases, infants imitated actions that were modeled for them weeks before (Klein &

Meltzoff, 1999; Meltzoff, 1988). Such acts of deferred imitation are thought to reflect infants' capacity for retaining and recalling accurate memory representations (Jones & Herbert, 2006; Klein & Meltzoff, 1999).

Infants also imitate vocal sounds that they hear around them. Vocal imitation is thought to be an important process for learning to communicate through sounds (Gratier & Devouche, 2011; Kuhl & Meltzoff, 1982, 1996; Papoušek & Papoušek, 1989). Already in the first months of life, infants imitate the intonation patterns of their parents in turn-taking interactions; during these exchanges, parents also imitate their infant's intonation contours (Gratier & Devouche, 2011; Papoušek & Papoušek, 1989). Through such imitation and repetition of prosodic contours during social interactions, prelinguistic infants learn to engage in communicative vocal exchanges and to produce and master new sounds. In addition to intonation patterns, infants also imitate the phonetic forms of their language. The process of hearing ambient speech and attempting to imitate the sounds seems to help infants refine the articulatory movements necessary for speech production (Kuhl & Meltzoff, 1996).

One of the processes by which learning occurs during vocal imitation involves forging sensorimotor links between the sounds that are produced and the movements that caused them (Kuhl et al., 2008; Kuhl & Meltzoff, 1996; Vihman & Boysson-Bardies, 1994). As infants perceive and discriminate speech sounds, they store models with which to compare their own production. Each attempt at imitating a model generates auditory feedback as well as proprioceptive feedback, which the infant then uses to evaluate the degree to which the vocalization matched the target sound. The sensorimotor interactions

are shaped through constant practice and vocal play. As a result, the infant generates stable representations of motor sequences and the sounds that these movements produce. This process of auditory-articulatory mapping allows infants to become more proficient producers of speech sounds. A similar sensorimotor model has also been used to describe singing, but not in the context of infant vocal development and learning. Such a model has mostly been used to explain individual variability in pitch matching abilities in adults (Berkowska & Dalla Bella, 2009; Tsang, Friendly, & Trainor, 2011) so its relevance to infant singing development is not clear.

Not much is known about vocal imitation of music in infancy. Findings from one of the few studies that addressed imitation of musical sounds suggest that infants can be conditioned to match isolated pitches that are sung or played to them (Kessen, Levine, & Wendrich, 1979). Some anecdotal accounts mention instances in which infants imitated music, but in many cases the information is not presented systematically or is only mentioned casually as part of a hypothetical model of music development (e.g., Elmer, 2011; Trehub & Gudmundsdottir, 2015; Welch, 2002). Considering that imitation of vocal sounds is a common behavior in infancy, it is reasonable to suppose that infants also imitate vocal sounds that are music, not only those that are speech.

Infants and Music

Babies live in a world filled with music. Infants are exposed to music of many different styles and timbres, ranging from live singing and playing by family members to recorded music from audio recordings and music toys (Costa-Giomi & Sun, 2016; Custodero & Johnson-Green, 2003; Ilari, 2005; Trehub et al., 1997; Young, 2008). Vocal

music is especially important for infants, and singing to infants is a common practice in childrearing across many cultures (Trehub, Unyk, & Trainor, 1993). Caregivers sing special songs to their babies, and they seem to do this in a distinctive style from that of singing directed to adults (Trainor, 1996; Trehub et al., 1997). Importantly, babies are responsive to such infant-directed (ID) singing. Premature infants, for instance, increase pacified sucking movements when given musical reinforcement consisting of ID singing (Standley et al., 2010). Infants perceive ID singing as different from ID speech, and they attend to singing more (Costa-Giomi, 2014; Nakata & Trehub, 2004) or at least as much as they attend to ID speech (Costa-Giomi & Ilari, 2014; Tsang, Falk, & Hessel, 2016). Singing to infants is thought to be important for regulating the infant's emotional state (Corbeil, Trehub, & Peretz, 2016), modulating level of arousal (Shenfield, Trehub, & Nakata, 2003), communicating affect to the infant (Trainor, 1996; Trainor, Clark, Huntley, & Adams, 1997), and supporting musical development (Costa-Giomi, 2013).

Infants are well-equipped for perceiving and processing the diverse musical input that surrounds them (for a review, see Trainor & Hannon, 2013). Abilities that emerge during the first half year of life include sensitivity to consonant and dissonant sounds (Schellenberg & Trainor, 1996; Trainor, 1997), perception of pitch (He & Trainor, 2009) and relative pitch (Chang & Trehub, 1977a), sensitivity to strong and weak metrical positions (Winkler, Háden, Ladinig, Sziller, & Honing, 2009), perception of musical temporal grouping (Juszyk & Krumhansl, 1993), and discrimination of rhythmic patterns (Chang & Trehub, 1977b). Later abilities reflect a process of enculturation, in which infants learn the characteristics of their native musical system. This process is

reflected in certain culture-specific perceptual biases, such as a processing advantage for familiar scales (Lynch, Eilers, Oller, & Urbano, 1990), and preferences, such as preference for the musical meters of their culture (Soley & Hannon, 2010). By the end of the first year, infants have considerable resources for perceiving and processing music in a sophisticated manner. Can they express their perception and representation of music through music making? Can they produce music?

Infant Music Production

Although we know much about the development of music perception in infancy, we know very little about infant music making. Most developmental accounts of music production focus on the singing of school-age children (e.g., Rutkowski, 1990, 1997; Welch, 2016). For example, Welch (2002, 2016), describes a 4-phase model of vocal pitch-matching that applies to school children and that refers to their ability to reproduce songs. Welch's model of singing development is described in detail to illustrate how generalizing a bounded model to other situations can lead to misconceptions. Initially, younger children attend mostly to the words and not the notes of the song. As a result, phase 1 of the model is characterized by singing that is chant-like and has little pitch variation. As they gain more singing experience and more exposure to the song repertoire, children start attending to melodic contours in addition to the lyrics. Accordingly, children's singing in phase 2 incorporates the general pitch contours of the melody, but many pitches are still inaccurate. In phase 3, children produce accurate melodies and intervals, but the tonality between phrases is less stable. Finally, in phase 4, children can sing simple songs with accurate pitches and a stable tonal center. In

summary, vocal pitch-matching by children in the context of singing a song begins with the words of the song, followed by the melodic contours, the specific pitches within the contours, and finally, a stable tonal center.

The four phases of vocal pitch-matching – from chanting the lyrics to accurate pitches and stable tonal center – are part of a framework for understanding the developmental timeline of singing behavior in school-age children. Before this age, it is assumed that the distinction between singing and speaking is blurred (Welch, 2016; 2002). Welch (1994) describes two reasons why this is so. First, the lack of distinction between singing and speaking is explained in part by physiological maturation processes of the vocal tract (Welch, 1994). Because it is difficult for young children to produce sounds with different voicing qualities, speaking voice and singing voice sound similar. The other reason concerns psycho-social development and enculturation processes: Very young children do not yet understand that a *song* is a distinct object from speech, and that singing and speaking require attending to different features of vocalizations and have different social and cultural meanings (Welch, 1994). Because speech and singing are not differentiated in infants and toddlers, according to Welch (1994) it is inappropriate to evaluate singing development at this age because “the first stage in any assessment of singing is a recognition that the activity [singing] *is* singing rather than something else” (p. 3, emphasis in original). Infants’ vocal behavior is described as belonging to a *babbling song* phase, which is described vaguely as consisting of experimentation with glissandi and repetitive pitch variations (Welch, 1994).

Many other accounts of infant music development echo the view that infants do not sing (e.g., Trainor & Hannon, 2013; Trehub, 2016; Tsang et al., 2011). For example, Trainor and Hannon (2013) reason that “because most singing involves words, it [singing] necessarily involves interactions between music and language” (p. 447). If it is the case that singing necessarily involves speech, the logical conclusion is that prelinguistic infants do not sing. But if prelinguistic infants do not sing, what is it that they do in terms of music? Infants are certainly very receptive to music, and they clearly develop appropriate perception skills for processing features of music. In fact, music educators have argued that infants actively participate in musical exchanges with those around them, and that such early engagement with music does indeed reflect vocal responses that are specific to music stimulation (Costa-Giomi & Benetti, in press; Reigado, Rocha, & Rodrigues, 2011; Tafuri & Villa, 2002).

If vocal production follows perception, there is no reason to suppose that infants do not incorporate music features into their own vocalizations. From a developmental perspective, it seems premature to assume a discontinuity between early vocal behavior in infancy and later music production abilities in early childhood. Research on the music components of infant vocalizations is scarce, however, so there is no evidence supporting an account of music development that includes music production.

Some lines of research discuss infant musical production as vocal expressions of affect in the context of communication and social interactions. For example, in an extensive case study of their daughter’s vocalizations from birth to 16 months, Papoušek and Papoušek (1981) describe the process by which Tanya, their daughter, learned to sing

a song. At 2 months, Tanya imitated short intonation patterns from ID speech; at 3.5 months, she matched pitches that her parents sung or played to her on the piano; at 4 months, she imitated longer contours of ID speech; and at 11.5 months, she hummed pieces of the melody and rhythm of certain songs. Papoušek and Papoušek conclude that the musical features of her vocalizations served as vehicles for communicating emotional states and solidifying parent-infant affective ties. Similarly, Malloch (1999) describes in great detail how musical elements actually structure turn-taking and imitation processes in infant-parent dyadic communication. He refers to such powerful and universal exchanges brimming with music as *communicative musicality*.

Both Malloch (1999) and Papoušek and Papoušek (1981) identify features of infants' vocalizations that are explicitly musical, such as melodies with discrete pitches and regular rhythmic patterns. For these authors, though, the music features of infant vocalizations are valuable because they are means for more general developmental outcomes, such as communication through turn-taking and expression of creativity. In other words, the music features are not fully regarded as a means for music-making that is independent of other behavioral outcomes. Although it is likely that music features are good for musical and non-musical development alike, the primary focus on the social aspects of musical vocalizations tends to ignore the question of what makes the musical features unique for whatever it is that they do. For example, Papoušek and Papoušek (1981) support that infant imitation of music melodies is important for the development of language and of the self, but they do not address why these melodies must be *music* melodies for such developmental processes to occur. What makes a music melody

unique? How is imitating a music melody any different from imitating a prosodic pitch contour? And if music melodies and prosody melodies are indeed different in terms of the effect they have on infant development, then what characteristic of these stimuli are responsible for such differences? Are the differences to be found in the physical characteristics of the signal? In the context of use? In affective cues? In the infant's biological response?

Instead of focusing on the significance of infant music production for social interactions, some studies have begun to describe the role of such vocalizations for music learning (Addessi, 2016; Reigado et al., 2011; Tafuri & Villa, 2002). The few studies that take this approach suggest that infants do in fact make music during their interactions with adults. They also propose that the vocalizations that infants make in response to a music stimulus are different from the vocalizations they produce in response to non-musical sounds. Reigado, Rocha, and Rodrigues (2011) compared infants' vocalizations after they had heard a song and after they had heard a poem. They found that the vocalizations in response to music were shorter and had a more restricted range compared to the vocalizations in response to language. Tafuri and Villa (2002) analyzed infants' vocalizations in response to maternal singing, and found that many of the infant sounds contained musical features such as glissandi, music intervals, and repeated rhythmic patterns. They called such vocalizations *musical babbling*. The study also intended to compare the vocalizations of infants who attended music classes with their mothers with those who did not attend music classes. Unfortunately, the comparison could not be made because the control group was too small. Overall, the exploratory studies that attempted

to investigate infants' vocal behavior in response to music suggest that infants are vocally responsive to music, and that they may vocalize selectively depending on the nature of the stimuli.

In summary, the findings from the few studies on infant music production suggest that in many cases infants' vocalizations can be described and represented using musical terms and notation. They also show that such vocalizations are organized in a musically meaningful manner. Importantly, the findings also indicate that infants are indeed very attuned to music produced around them, and may even incorporate these music elements into their own vocalizations. Although past research has greatly contributed to our understanding of the largely neglected topic of infant music production, there are important limitations that must be addressed. First, it is not entirely clear if the musical vocal behavior described in these studies is specific to music. For example, in Reigado, Rocha, and Rodrigues' study (2011), infant vocalizations in response to music displayed a more limited range compared to the vocalizations in response to language. However, this was true for the stimuli as well: The pitch ranges of the songs were also smaller than the pitch ranges of the poems. It could be that the infants' adjustment of range was simply the result of the reduced range of the music stimuli. To distinguish between infants' adjustment of pitch range under the two conditions, the music and speech stimuli would need to be matched for range.

Conclusion

Two important issues emerge from the review of literature, both of which must be addressed to better understand infant music production. First, little is known about

infants' musical behavior in private contexts (i.e., when alone or without direct vocal engagement with another person). Analyzing infants' spontaneous and unprompted vocalizations, including those that occur in the absence of others, may be especially important for understanding infant music production because vocalizations in social situations may not be wholly representative of infant musical behavior. Indeed, infants' vocalizations seem to become more speech-like in the presence of verbal feedback from adults (e.g., Goldstein & Schwade, 2008; Gros-Louis, West, & King, 2014; Warlaumont, Richards, Gilkerson, & Oller, 2014), and it is possible that in social situations infants naturally gravitate towards vocalizing sounds that resemble speech even in the presence of music. Secondly, to begin to describe music production first it is necessary to operationalize relevant dimensions of music. One such dimension is melody: What makes a music melody different from a melody of speech prosody in the context of infant vocal development? The lack of consistent parameters (e.g., the distinction between a music melody and a prosodic melody) across studies makes it difficult to integrate previous findings. As a result, we do not yet have a systematic framework from which to raise meaningful hypotheses about vocal production of music by infants.

The present case study addressed these issues by focusing on imitations of music in a naturalistic setting. Imitations provide an opportunity for comparing the features of the infant's imitative response with the features of the music model produced around him by people or music devices. Because the model is known to be a music melody (for example, the mother singing a lullaby), we can attempt to determine what features of this model the infant integrated into his own imitative response, and how those features relate

to the perceived similarity between the model and the imitation. Furthermore, by capturing the infant's imitations that occur in a naturalistic setting, we can begin to understand the patterns of music behavior as it occurs spontaneously. A case study provides an opportunity for in-depth investigation of this complex and multi-faceted phenomenon as it manifests itself in this family.

Chapter 3: Method

As part of my involvement with a larger project on the soundscapes of American families (Costa-Giomi, 2016), I was assigned one 16-hour audio file of one family to listen and analyze the infant's home soundscapes. The present case study describes and analyzes instances of music imitation by the youngest child of this family, "James," during the one day in which he was recorded.

Participant

The participant in this study was James², a 15-month-old male infant. James lived with his mother and father, a 5-year-old brother, and a 3-year-old sister in a middle-class urban neighborhood in the Southwest region of the United States. The mother was a homemaker, and the father worked as a software engineer. Both parents held graduate degrees. Only English was spoken at home.

All three children were formally engaged with music from an early age. James, the infant, and "Diane," the sister, attended a Suzuki music class for infants, toddlers, and their parents. "Dale," the older brother, participated in a string program and was learning to play the cello. In addition, both parents had studied music for 10 years.

² The infant and all family members were assigned pseudonyms.

Data Collection

The data for this study were collected using a portable recording device worn by the infant throughout an entire day. The recording device, also known as Digital Language Processor (DLP), is part of a system developed by the LENA Foundation (*Language Environment Analysis*, www.lena.org) for studying the home language environment of infants and young children. The DLP captures sounds from the infant's perspective, including all sounds that are produced near the infant as well as all sounds from the infant's own vocal production. The DLP records up to 16 continuous hours. The device is turned on when the infant wakes up in the morning and placed inside the chest pocket of an infant vest or t-shirt. The clothing is especially designed with fabric that minimizes swishing noises when the infant moves. Because the DLP is placed in the infant's pocket, it also records the infant's vocalizations when he or she is alone. At night, the device is turned off. The result is a high-quality audio recording that can be analyzed for further study of quantitative and qualitative aspects of the infant's vocal production and home sound environment.

In the LENA system, the sound data collected from the DLP are analyzed by software developed specifically for assessing language development. The software measures the total number of adult words, the total number of conversational turns, child vocalization frequency, and the total time of ambient sounds (e.g., music from electronic devices, television, sounds from toys). The numbers are compared with normative data to provide an assessment of the infant's linguistic development and a general overview of the infant's home language input.

James wore the DLP during a weekday in January. The device was turned on at approximately 7:20 a.m., when James woke up in the morning, and turned off 16 hours later. Data for this study were collected from the audio file produced during this day. The audio file consisted of all sounds produced within hearing distance of James, as well as James's own vocalizations. It is important to note that the sounds captured by the device were produced spontaneously during natural situations in the home and not prompted by the researcher or selected by the parent. The parents kept a log describing the main activities of the day, such as *naptime* or *visit from friends*, and the people with whom James interacted. The log provided a general overview of the day's events and helped to identify the voices in the recording.

Perceptual Identification of Imitations

The focus of this study was on the infant's imitation of sounds produced around him. To identify and describe instances of imitation, I completed extensive listening of the recording. To facilitate listening and analysis, the 16-hour audio file was divided into shorter segments based on the events and interactions of the day. The duration of each segment varied from a few minutes to a few hours. For example, segments involving interactions between the mother and the infant were very short, whereas segments in which the infant was sleeping were considerably longer.

Each segment of the sound file was subjected to repeated listening. The strategy I employed involved shifting between listening for context and listening for details in each segment. First, it was necessary to familiarize myself with the general context of the household, the time sequence of events during the day, and the voices of the family

members and the infant. Initially, I registered who was present, where they were in relation to the infant (i.e., next to the infant, near the infant, in the background), and the nature of their interaction with the infant (speaking or singing directly to the infant, speaking or singing to someone else, speaking or singing to themselves, not speaking or singing at all). I described what the infant was doing and his vocal production during that time in more detail. I also documented the general time frame in which events occurred.

The next step involved repeated and in-depth listening of smaller fragments of the segment, particularly those parts in which the infant vocalized. I described the infant's vocalizations and other relevant sounds in greater detail, and I documented the exact start and end times of these events. I described all infant-directed speech, and all singing or music produced specifically for the infant as well as produced in the background. I focused not only on describing the content of the vocalizations, but also on what the vocalizations sounded like. When appropriate, I transcribed the vocalizations using speech symbols and music symbols. As a rough estimate, each hour of recording required approximately five hours of listening time. I completed this part of the study in four months.

Finally, it was necessary to isolate and analyze specific instances in which I suspected that the infant imitated a sound. Imitative responses were identified by perceptual judgement of musical relationships and similarities (e.g., relative pitch and rhythmic motifs) between the sounds of the model and the sounds of the imitative response. The pairing of model and imitative response was a result of a perceptual

similarity judgment by the author, who is a trained musician and who spent more than four months listening to this infant and his family.

The goal of this case study was to determine whether the infant imitated music and to provide an in-depth analysis of such imitations. Thus, it was not necessary to collect all instances of imitations. Instead, I selected a small number of imitations that occurred in a variety of contexts and that had a variety of sounds as models to describe in-depth and exhaustively.

Description of the Models and the Imitations

For each selected imitation event, I described the model (i.e., the sound which James imitated) and the imitative response by the infant. For the model, I described (a) the source of the sound, (b) the type of sound produced (e.g., timbre), (c) the nature and quality of the sound (e.g., speaking, singing with words, singing without words), (d) the location (e.g., next to the infant, near him, in the background) and context (infant-directed or non-infant directed) in which the sound was produced, and (e) the musical and linguistic features of the sound (Table 1). For the imitative response, I described (a) how much time elapsed between the model and the infant's imitation (ranging from seconds to hours after the model), (b) in what interaction context the infant produced the imitation, (c) who was present when the infant vocalized and how they responded, and (d) the music and linguistic features of the infant's imitations when applicable (Table 2).

Table 1. Characteristics and description of models imitated by the infant

Characteristic	Description	Examples
Source	What object or person produced the sound?	Human (mother, father, sister, brother) Acoustic instrument Electronic device
Type	What kind of sound was produced?	Human voice Musical instrument Synthesized sound
Quality	What was the nature and texture of the sound produced?	Speech Singing (with and without words) Other sounds (e.g., hissing or percussive sounds) Non-vocal melody (with and without accompaniment)
Location and Context	Where and in what context was the sound produced, in relation to the infant?	Context: infant-directed and non-infant-directed Location: next to the infant, near the infant, in the background
Features	What were the musical and linguistic features of the sounds?	Music: notes, melodic contour, key, starting pitch, rhythm, timbre, duration Speech: words, intonation, stress patterns

Table 2. Characteristics and description of infant's imitative responses

Characteristic	Description	Examples
Time	How much time elapsed between the model and the infant's imitation?	3 seconds 10 hours
Interaction context	In what social context was the imitation produced?	Parent-infant dyadic communication Social interaction with family members Playing with siblings In private (e.g., playing alone with toy)
Family members	Who was present when the infant imitated (even if not interacting with the infant)? How did they respond to the infant's imitation?	Mother present, no verbal response Mother in the background, no response Father near infant, imitates infant's sounds Siblings next to infant, no response
Features	What were the musical and linguistic features of the imitation?	Music: notes, melodic contour, key, starting pitch, rhythm, timbre, duration Speech: phonetic content, intonation, stress patterns, formant patterns

Other Data Collected

I identified and described all music to which the infant was exposed during the 16 hours of observation. The types of music included infant-directed singing, non-infant directed singing, music from electronic toys, and music from acoustic instruments. The specific repertoire was identified when possible. I counted the total time of exposure to each type of music, as well as how much music each family member produced throughout the day.

Acoustic Analysis

I completed an acoustical analysis of each of the selected imitations. The purpose of these analyses was to corroborate my perceptual findings by determining what parameters of the sound were manipulated by the infant and how such manipulations related to the parameters of the model and to the perceptual description of the sounds. Accordingly, only imitations that were identified by listening were analyzed for acoustical data.

Acoustic analyses were used to complement, not replace, my perceptual analysis of James's vocalizations. Although it is in reference to infant phonetic development, not music, the argument that Kuhl and Meltzoff (1996) make for using both perceptual and acoustic analyses is convincing:

Phonetic transcription is needed to establish whether an infant's response matches that of the adult phonetic category. Instrumental analyses establish which acoustic dimensions infants are capable of varying and establish how infants manipulate those dimensions when producing sounds that are perceived as different. (p. 2427)

The advantages of using both perceptual and acoustical analyses are clear in the discussion of pitch in music. Frequency is a physical attribute of the wave, whereas pitch is a psychological attribute arising from the listener's perception of a periodic sound. Sounds that give rise to the sensation of pitch usually have energy distributed among multiple harmonics (integer multiples of the lowest, or fundamental, frequency). Pitch extraction algorithms used in software such as Praat (Boersma & Weenink, 2016) detect periodicity in the acoustic signal and make an educated guess as to the fundamental frequency (for a comprehensive explanation of the algorithm used in Praat, see Boersma, 1993). However, the physical attributes and the perceptual experience sometimes do not seem to correspond to each other. For example, listeners perceive pitch categorically and not as a frequency continuum. They accommodate a range of frequencies when making similarity judgments between two sounds, even if the two sounds are considered out of tune based on physical measures of frequency (Siegel & Siegel, 1977). In fact, expert musicians judge melodic intervals that are up to 25 cents out of tune as being in tune (Vurma & Ross, 2006). Listeners also make use of patterns and relations between the information from a signal. In the well-known phenomenon of the missing fundamental, for instance, listeners identify the pitch of a sound by integrating information from the energy distribution among the harmonic spectrum, even when they cannot hear the fundamental frequency (Moore, 2008). These examples illustrate the need to examine the characteristics of sounds using both perceptual and acoustical analyses.

Acoustic data were collected using Praat, version 6.0.17 (Boersma & Weenink, 2016). The acoustic parameters that were measured included pitch, spectral information

(harmonic and formant patterns), duration, and intensity. For some excerpts the pitch information was collected with Melodyne, a professional software for audio pitch correction, in addition to Praat. Melodyne was developed specifically for analyzing and editing music. According to the developers of the software (www.celmony.com), Melodyne identifies music notes by considering parameters such as pitch center, pitch modulation, and pitch drift.

Some of the excerpts contained overlapping voices and noises that interfered with acoustic analysis. With the help of an audio engineer, unwanted sounds were made less salient by using iZotope, a professional software for sound editing (www.izotope.com).

Validation: Perception Study with Adults

The imitations were identified based on my judgment of similarity between models and the infant's sounds. Would the imitations be perceived as such by other adults? Would other adults perceive the similarities between the model and the imitative response? The answers to these questions were critical to lend validity to my perceptual analysis. After completing the perceptual and acoustical analyses, I conducted a validation study with adult participants to address these questions. The design and results of the validation study are described in detail in Chapter 5.

Chapter 4: Results

The results are divided into two sections. The first section is a general description of James's music and language environment and his language production throughout the day. The second section presents an overview of James's vocal imitations and the detailed analyses of four selected imitations: (1) "Rain Rain Go Away"; (2) "Happy Birthday"; (3) and (4) two music intervals, one sung by the sister and the other sung by the brother.

The infant wore the recording device on a weekday. The recording device was turned on at approximately 7:20 a.m. and stopped recording at approximately 11:20 p.m., four hours after James went to sleep, for a total of 16 hours of continuous recording time. School was cancelled on that day due to inclement weather, and the three children and the mother stayed at home. The father was home for most of the day. A friend of the mother and her toddler daughter visited the family before lunch. Appendix A shows an outline of the sequence of events of that day including naps, meals, outings, visits, and activities.

Music and Language Environment and Language Production

Music Environment

James was exposed to music from several sources, both human as well as electronic, throughout the day. He heard vocal music from human singing, vocal music from a synthesized voice from an electronic toy, non-vocal music from an acoustic

instrument, and non-vocal music in the form of synthesized sounds from electronic toys. Table 3 presents the various types of music to which he was exposed. They included (a) infant-directed singing (i.e., singing specifically produced for James) by the mother, the father, and possibly by his sister Diane; (b) non-infant directed singing (i.e., singing not produced specifically for James) by the parents and siblings; (c) songs from various electronic music toys; (d) plucking from a cello played by his brother Dale; and (e) vocal sounds used in musical ways (e.g., rhythmic speech).

I calculated the amount of time that James was exposed to the various types of music. Table 3 shows the total amount of time for each type of music. During the day, the mother sang to James for approximately eight minutes, and the father for 2 minutes. He heard some sporadic singing by his siblings that amounted to almost two minutes, and non-infant directed singing produced mostly by the father that totaled 40 seconds. James heard Dale plucking a cello for 31 seconds and melodies played by music toys for 24 minutes. It should be noted that when calculating the times for the various types of music, I did not include any pauses embedded in the musical interactions.

Specific songs and the descriptions of musical sounds are shown in Table 3. Overall, the music repertoire from singing and from the toys consisted predominantly of well-known children's songs, such as "London Bridge" and "Pop Goes the Weasel." The parents sang mostly children's songs to James. They sometimes sang the songs with traditional lyrics and at other times with invented lyrics. Singing was also spontaneously incorporated in contexts that did not involve a fixed repertoire, such as calling out James's name using a music interval or using a series of rising pitches to accompany the

phrase “up up up up up.” Most non-infant-directed singing captured by the recording device consisted of the father singing to the siblings, although the mother also sang with Diane briefly. Dale sang very little throughout the day. Diane sang mostly improvised melodies, but she also sang short excerpts of “London Bridge.”

The melodies from the music toys that James heard during the day of data collection also consisted of children’s songs, only one of which I did not recognize (see Table 3). This unfamiliar song about birthdays was the only one that included words to accompany the synthesized sounds. For a more detailed description of the music and the musical interactions observed in this household, see Costa-Giomi and Benetti (in press).

Table 3. Infant's music environment and duration of exposure to music

Music	Duration (min:s)	Content
Infant-Directed (ID) Sounds		
Mother singing	08:21	Traditional children's songs (e.g., "There Was an Old Lady Who Swallowed a Fly," "London Bridge," "Hush-a-bye Baby," "Hush, Little Baby," "Rain Rain Go Away"); Two unfamiliar children's songs
Father singing	02:02	Traditional children's songs ("Rain Rain Go Away"); Traditional songs with invented lyrics ("The Wheels on the Bus," "Rain Rain Go Away"); "Abiyoyo" (a children's song by Pete Seeger)
Other sounds	00:20	Rhythmic speech, rhymes, clapping, pitched speech (e.g., "up-up-up" with rising pitches)
Total	10:43	
Non-ID Sounds		
Mother to siblings	00:51	"London Bridge" with sister (00:08); Chanting and clapping rhythms with brother for cello practice (00:43)
Father to siblings	00:32	Traditional children's song ("There Was an Old Lady Who Swallowed a Fly"); "Abiyoyo" (Pete Seeger)
Brother	00:17	Traditional children's song ("Bingo"); "Abiyoyo" (Pete Seeger); Imitation of fire truck siren with pitches
Sister	01:37	Traditional children's song ("London Bridge"); Improvised melodies (00:13 possibly ID)
Total	03:17	
Acoustic Instruments		
Cello	00:31	Plucking with rhythms
Total	00:31	
Electronic Music		
Music toys	24:21	Traditional songs (e.g., "Happy Birthday," "Bingo," "Old McDonald," "Pop Goes the Weasel," "Oh Susanna," "Father John," "Clementine") without lyrics; Unknown song ("Birthday Song 2") with lyrics
Total	24:21	
Total	38:52	

Language Environment and James's Language Production

The 16-hour audio file was analyzed using the LENA software to gather information about James's home language environment. The language environment data yielded by the LENA system is useful because it provides an overview of the infant's linguistic input and output (i.e., the language he heard and that he produced) during that day. It is important to emphasize that the software's audio processing system provides quantitative measures of sounds produced within hearing distance of the infant, and does not differentiate between speech sounds and singing. During the day of the recording, 20,814 adult words were spoken to or near James. James produced 2,165 vocalizations (i.e., infant sounds that are not vegetative sounds or fixed signals such as crying).

Although I identified some instances of intelligible speech sounds by James during the 16 hours of recording, most of his sounds were prespeech vocalizations. This is in line with previous research suggesting that at 15 months infants produce mostly vowel sounds and syllables containing consonant-like and vowel-like elements but that do not yet resemble adult forms (Nathani, Ertmer, & Stark, 2006). James produced some identifiable words that were contextually relevant, such as "doggy," "daddy" and "dada," "mommy" and "mama," "bye," "nana" (i.e., banana), "go," "no," and "mine." The most frequent words were "doggy" and "daddy." In some instances, these two words seemed to be morphed into "daggy." The only two-word utterance was "bye daddy."

Overview of Imitations

James produced vocalizations that were perceptually similar to sounds he had heard previously, suggesting that throughout the day he engaged in acts of vocal imitation. The results presented here include a brief overview of the variety of sounds that James imitated, followed by a detailed description of a selection of music imitations. The purpose of the overview is not to describe the imitations in detail or to give an exact count of each type of imitation, but rather to give a general idea of James's imitative vocal life: What sounds were imitated? Who produced those sounds? When did the imitations occur and in what contexts? What features of the model sounds were imitated?

James imitated live vocal sounds produced by all four family members (mother, father, sister, and brother) and non-live music in the form of synthesized sounds from an electronic music toy. James imitated speech, melodies with words, melodies without words, and other vocal sounds (such as "ts" sounds). The sounds he imitated were both infant-directed and non-infant-directed. He imitated sounds that were produced next to him, near him, and in the background.

James imitated sounds immediately after hearing them as well as hours later. James's imitations occurred in private contexts (e.g., playing alone with a toy), in the presence of others but without communicative interaction (e.g., playing next to the other children, but not interacting with them), and in the presence of others with communicative interaction (e.g., turn-taking with a parent). In many cases, no one responded to James's vocalizations. Such lack of vocal feedback seemed to occur mostly in situations in which the other person did not hear James's vocalizations or the preceding

sound that he imitated. In other cases, the parents did acknowledge the imitation, sometimes by imitating the infant's vocalization themselves.

James imitated a variety of musical and linguistic sounds. The linguistic elements he imitated included phonetic features, formant structure, and intonation patterns. The music elements included absolute pitches of the melody, the interval structure (relative pitches), the general melodic contour, rhythms, and tempo. Sometimes only one feature of the model was salient in James's imitation. For example, in some cases the intonation pattern was especially noticeable but the phonetic content was not, or the pitches matched the melody but the rhythm was not as clear. Most imitations, however, incorporated more than one feature of the model.

“Rain Rain Go Away”

Two episodes are described below. The first one occurs soon after lunch, when the mother sings “Rain Rain Go Away” to James when putting him to sleep. The second one occurs approximately six hours later, when the father and James are alone. The infant calls out to the father and produces a series of vocalizations that resemble the melody of “Rain Rain Go Away.” The father recognizes the melody and proceeds to sing the song. While the father is singing, James imitates him three times.

Mother Singing “Rain Rain Go Away” (12:25 p.m.)

Description. It is naptime; James has just had lunch. As James and his mother get ready to leave the kitchen, she explains to Diane and Dale that she is taking the infant to the bedroom for his nap. Now in the bedroom, the mother sits down on a rocking chair and holds James in her arms. He is very still for a few minutes, and except for breathing

sounds from James and creaking sounds from the rocking chair moving back and forth, it is very quiet. James is calm and likely asleep. As the mother puts him in his crib, however, he starts complaining. To ease the transition, the mother begins to sing “Bah Bah Black Sheep” (key of G, starting on G3). Initially she maintains the traditional lyrics of the song, but about halfway through she changes the words:

James, James, you’re my little boy.

You are sweet and full of joy.

You’re really fat and you’re really sweet.

You look nice in your birthday suit.

James Cooper you’re my little boy.

James Cooper you bring us lots of joy.

The singing lasts approximately 85 seconds. James is quiet throughout – perhaps he is already sleeping.

After 20 seconds of silence, the mother begins singing again, this time the song “Rain Rain Go Away.” The melody of this song, as sung by the mother, consists of only two different notes which form an interval of a minor third. The first two notes of the melody, corresponding to the lyrics “rain rain,” form the descending minor third (Figure 1). When the mother begins singing, the minor third relationship is in tune: The pitches of the first and second notes are, respectively, a flat E4 and a flat C#4. These initial two notes created a stable reference tonal point. She is mostly consistent in maintaining the descending minor third interval for the entire 21 seconds of singing – even when humming during the last few seconds. Her intonation sometimes drifts slightly when

holding the longer notes or transitioning between notes when articulating consonants, yet the tonal center remains clear and stable.

Figure 1. Music notation of mother singing "Rain Rain Go Away" (perceptual analysis)

The mother sings the song two times. The first performance of the song lasts 8 seconds, and the second performance lasts 13 seconds. The tempo she chooses for the first time is rather fast – approximately 116 beats per minute. The expression of her voice is tender and sweet. She yawns while she sings the last part of the song. After a pause of 3 seconds, she sings the song again, but this time at a much slower tempo of approximately 72 beats per minute. She repeats the last two measures of the song, replacing the lyrics “little children want to play” with humming sounds.

Acoustic analysis. The first time the mother sings, the pitches of the first two notes are 324.34 Hz (E4 -28 cents) and 272.04 Hz (C#4 -32 cents), respectively. When she sings the song again, the pitches of the first two notes are 325.85 Hz (E4 -20 cents) and 273.67 (C#4 -22 cents), respectively. In both cases, the pitches of the subsequent notes remain mostly within a quarter tone of these two pitches. In Figure 2, the pitches of the second rendition of the song are represented as semitones (rounded to the nearest quarter tone). The exact frequencies of each note are listed below the lyrics.

Figure 2. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of mother's pitches from acoustic analysis of mother singing "Rain Rain Go Away"

Figure 3 shows the spectrogram corresponding to the beginning of the same segment. The blue line indicates the pitch contour and the yellow line indicates intensity. Figure 4 shows the pitches of the same segment of singing, but as detected by the software Melodyne. The visual representation of the minor third interval so fundamental in the song "Rain Rain Go Away" is very clear. It is interesting to observe the degree to which the pitch as sung by the mother drifts within each note. This is illustrated by the wavering line within each red-yellow blob. This variation is minimal and, as shown in Figures 1 and 2, does not detract from the identification of the pitches of the song.

Figure 3. Spectrogram of mother singing "Rain Rain Go Away" (blue line indicates the pitch contour; yellow line indicates intensity)

Figure 4. Representation of mother's pitches using Melodyne

The acoustic analysis supports my perception of the mother's singing. The mother's rendition of "Rain Rain Go Away" is indeed accurate and consistently in tune.

Infant Imitation and Father Singing "Rain Rain Go Away" (6:40 p.m.)

Description. The father and James are alone. The father is interacting with him, making playful rhythmic hissing sounds and pausing every once in a while to let James answer. James responds with similar sounds when it is his turn. After 10 seconds of this

game, James vocalizes two syllables with a long, drawn-out melodic contour. The sound resembles the word “daddy”. Except for the imprecise articulation of the second consonant – a fusion between /g/ and /d/ –, the syllables are well-produced and the word comes very close to canonical form.

The two syllables of “daddy” are followed by another eight sounds consisting of consonant-like sounds and vowels but that do not resemble any identifiable word. The eight sounds are also qualitatively different from the “daddy” vocalization in terms of sonority: The timbre is dull and muffled, and the intensity within each sound is unstable. These characteristics contribute to the perception that these sounds appear more introverted compared to the syllables of “daddy.” With respect to the timing, however, the eight sounds are clearly defined and consistent. James’s breathing pattern is regular, and his breaths divide the eight sounds into three groups of roughly equal duration. The first group consists of two short sounds followed by a long one, the second group consists of two sounds of medium duration, and the third group consists of three short sounds. The predictable timing of the breaths and the temporal organization of the sounds give rise to the sensation of a predictable rhythm but with imprecise execution. It is almost as if one could infer what rhythm James intended to produce but was not able to, much like one can infer that what he intended to say with the two preceding syllables was “daddy”, even though his immature articulatory movements do not allow him to do so as well as an adult.

This sequence of vocalizations, consisting of the two syllables of “daddy” and the eight subsequent sounds, last approximately eight seconds.

James is fully engaged in vocalizing. Compared to the eight previous sounds, the two sounds that follow are louder and more sonorous, giving the impression that something new is beginning. These two sounds are part of a sequence of 11 consisting of both isolated vowels as well as combinations of consonant-like sounds and vowels. The sequence of 11 sounds lasts approximately seven seconds. It is naturally divided into four groups of approximately equal duration, each group ending with a short pause for breathing. The first two sounds, corresponding to E₄ and C₄ (a descending minor third), form Group 1. Group 1 is slightly longer than the others. Groups 2, Group 3, and Group 4 consist of three short sounds each. The sounds of Groups 2, 3, and 4 do not sound as clearly pitched as those of Group 1.

The temporal organization and the pitches of the 11 sounds closely resemble the rhythm and melody of the song “Rain Rain Go Away.” Figure 5 illustrates how the sounds can be represented using music symbols, in the key of A flat major. The prominent descending minor third interval of these first two notes leaves a strong melodic imprint that affects the interpretation of the next nine sounds, which are not as clearly defined in terms of pitch.

Figure 5. Music notation of infant's "Rain Rain Go Away" imitation (perceptual analysis)

The father, who has been listening to James the entire time, is apparently also reminded of the song “Rain Rain Go Away”: He interrupts James and says, “It sounds like you’re singing...,” and then proceeds to sing the song with the traditional lyrics: “Rain rain go away / Little children want to play.” He sings in the key of G Major, starting on D3 and descending to B2 (Figure 6). He only sings the first half of the song. This lasts approximately 6.5 seconds.

Figure 6. Music notation of father singing "Rain Rain Go Away" (perceptual analysis)

After finishing the first half of the song, the father pauses, almost as if waiting for James to respond. James seems to understand the cue, and perhaps even anticipates it: Immediately after the father stops singing he imitates the rhythm and the melodic contour of the last measure of the song (i.e., “want to play”). James produces three vowels with a descending interval pattern whose durations are equivalent to two quarter notes followed by a half note. This rhythm is identical to the one notated in measure two of Figure 5. Interestingly, James does not begin singing on E_b, as he had done initially, but rather on D, his father’s starting pitch (Figure 7). Although James is successful in matching his father’s first pitch, he is not able to match the lower note of the minor third interval (i.e.,

B3). Instead, James produces a sound with a pitch that is very hard to define. This note is depicted with the note head “x” in Figure 7. Nevertheless, the pitch is clearly lower than D4, and the descending melodic contour of the tune is preserved.

Figure 7. Music notation of infant's imitation of father and father singing second part of “Rain Rain Go Away” (perceptual analysis)

James produces two more sounds with durations equivalent to a quarter note, but because the father begins to sing the last part of the song (i.e., “come again another day”) at the same time, the infant’s sounds are obscured. When the father finishes singing, James produces three sounds with the same rhythm as the last part of the melody sung by his father (i.e., “-nother day”) but with unclear pitch. In addition to imitating the rhythm of the song, James also imitates phonetic features of the lyrics. He produces consonant-like sounds that resemble /n/ and /d/, as well as vowels with formant patterns that approximate the general shape of the father’s vowel sounds.

James then produces another two sounds, both with well-articulated consonants (/d/) that sound almost like “dada.” When James stops vocalizing, the father resumes singing. This time he changes the words slightly: “Rain rain go away / Come again

another day / Little James wants to play.” The father pauses for a few seconds and then says, “Or...,” with a long undulating contour that suggests he is about to introduce another possibility. As expected from the prosodic cues, he presents a variation of the song by adding improvised lyrics to the melody: “Octopus, octopus / Riding on the city bus.” The father then pauses for a moment and mumbles something to James, who starts crying. Yet again, the father says, “Or... we can do...,” and introduces a new variation of words for the same melody: “Red balloon / Red balloon / Floating flying to the moon.” James vocalizes short sounds of distress, and the father begins again: “Scrambled eggs / Scrambled eggs / Makes big muscles in my legs.” After this version of the lyrics, James vocalizes three sounds that resemble the rhythm of “in my legs,” followed by short moans of frustration. The rhythm of these three sounds is not quite as precise as the rhythm of his previous imitations. Despite the less accurate musical reproduction, the three sounds in response to the music are easily distinguished from the sounds of distress that preceded them and the moans of frustration that followed them. Compared to the sounds of discomfort, the timing of these three sounds is more organized, and the same quality of timbre is maintained throughout each sound.

James now begins to cry, and the father again sings with improvised lyrics: “Butterfly / Butterfly / See them flutter in the sky.” After a few seconds, James produces a series of “uh-oh” sounds that seem unrelated to the song. The father continues interacting with James, but does not continue singing.

The father sang to James for a total of 47.5 seconds during the entire 2-minute interaction.

Acoustic analysis. The vocalizations represented in Figure 5 (the infant’s first imitation of “Rain Rain Go Away”), Figure 6 (the father singing “Rain Rain Go Away”), and Figure 7 (the infant’s imitation of the father singing the song) were analyzed in terms of specific acoustic parameters. Additionally, the infant’s vocalization that preceded the first “Rain Rain Go Away” imitation, including the word “daddy,” were also analyzed. The spectrogram of James’s vocalization that preceded “Rain Rain Go Away” is shown in Figure 8. The spectrogram illustrates the sequence of vocal events leading up to the imitation, and depicts the rhythmic grouping of James’s vocalization described in the previous section.

Figure 8. Spectrogram of infant's vocalization that preceded the “Rain Rain Go Away” imitation and the beginning of the "Rain Rain Go Away" imitation (blue line indicates pitch contour; yellow line indicates intensity)

The infant’s first “Rain Rain Go Away” vocalization (Figure 9) consists of 11 sounds organized into four groups of roughly equal duration. The groups are separated by pauses during which the infant breathes. Group 1 consists of two sounds. Groups 2, 3,

and 4 each consist of three sounds. Across individual sounds, the intensity is mostly uniform. Within each sound, the peak of intensity usually occurs close to the onset of the sound. In each group, the maximum intensity occurs during the first sound of the group. The increase in intensity marks the beginning of the group. Table 4 presents the intensity measure of each sound.

Figure 9. Spectrogram of infant's "Rain Rain Go Away" imitation (blue line indicates pitch contour; yellow line indicates intensity)

Table 4. Frequency (Hz), duration (s), and intensity (dB) of infant's "Rain Rain Go Away" sounds

Group	Intensity (dB)	Frequency (Hz)	Pitch ^a	Duration ^b (s)
Group 1				
Sound 1	80.30	308.08	E _♭ 4 -17¢	0.83
Sound 2	72.72	260.54	C4 -7¢	0.80
Total				1.63
Group 2				
Sound 1	76.41	290.34	D4 -20¢	0.47
Sound 2	75.87	296.56	D4 +17¢	0.51
Sound 3	75.07	258.94	C4 -18¢	0.77
Total				1.75
Group 3				
Sound 1	79.47	294.81	D4 +7¢	0.35
Sound 2	71.66	297.98	D4 +25¢	0.40
Sound 3	73.24	272.77	C _♯ 4 -15¢	0.84
Total				1.59
Group 4				
Sound 1	80.04	307.71	E _♭ 4 -19¢	0.45
Sound 2	79.44	287.35	D4 -38¢	0.48
Sound 3	79.30	297.62	D4 +23¢	0.87
Total				1.80

^aPitch equivalent of frequency based on equal tempered scale. ^bBased on onset time of the sound in relation to total duration of group.

The time that elapses between the onset of successive sounds suggests that the durations of the sounds are related to an underlying temporal structure (Table 4). In Group 1, the onset of the second sound divides the group into two parts of equal duration.

In Group 2, the onset of the third sound divides the group into two parts, and the onset of the second sound divides the first half of the group in two. Groups 3 and 4 share the same temporal organization as Group 2. The temporal organization suggested by these subdivisions, although not perfect, aligns with the two rhythmic patterns that are a prominent feature of the “Rain Rain Go Away” melody: the two half notes, and the two quarter notes followed by a half note (Figure 5).

In Group 1, the pitches of the first and second sounds are 308.08 Hz (E₄ -17 cents) and 260.54 Hz (C₄ -7 cents), respectively (Table 4). These frequencies corroborate the perceptual judgment of pitch represented earlier in music notation (Figure 5). Importantly, the intervallic relationship between these two frequencies is a descending minor third. Table 4 lists the frequencies of the sounds of all four groups. In Groups 2, 3, and 4, the descending minor third interval becomes progressively smaller, decreasing from a major second in Group 2 to a minor second in Groups 3 and 4. Note, however, that in all groups the descending melodic contour is preserved.

Figure 10 shows the musical representation of the pitches of the sounds. Pitches are represented as equal-tempered semitones, rounded to the nearest quarter step. The rhythmic value of each sound is interpreted based on the duration of the sound (as estimated from the onset time) in relation to the duration of the group. For example, the first sound of Group 1 is represented as a half note (two beats) because its duration (0.83 s) is approximately half the duration of the group (1.63 s).

Figure 10. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of infant's pitches from acoustic analysis of the infant's "Rain Rain Go Away" sounds

When the father sings for the first time, the pitches of the first two notes ("rain rain") are 145.86 Hz (D3 -12 cents) and 121.79 (B2 -24 cents), respectively. As he continues to sing, the pitches fluctuate slightly. When he sings the last measure, immediately before James vocalizes in response, the father is back to D3 and B2. Figure 11 shows the exact frequencies of the pitches, as well as the equivalent music notation.

Figure 11. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of father's pitches from acoustic analysis of father singing "Rain Rain Go Away"

The pitches of James's three sounds in response to the father are 292.93 Hz (D4 -4 cents), 287.38 Hz (D4 -37 cents), and 231.03 Hz (B3 -15 cents), respectively. The

spectrum, pitch contour, and intensity of the sounds of the infant, as well as the father's singing that immediately preceded these sounds, is shown in Figure 12.

Figure 12. Spectrogram of last measure of father singing “Rain Rain Go Away” and beginning of infant's imitation of father (blue line indicates pitch contour; yellow line indicates intensity)

The father's next notes partially overlap with James's vocalizations, so it was not possible to collect accurate pitch information for this segment. The pitches of the last three notes, corresponding to the lyrics “-nother day,” are 142.16 Hz ($D_{\flat}3$ +44 cents), 144.25 Hz ($D3$ -31 cents), and 117.34 Hz ($B_{\flat}2$ +12 cents). The infant produces three sounds in response. The pitches of the first and second sounds are 300.22 Hz ($D4$ +38 cents) and 322.84 Hz ($E4$ -36 cents), respectively. It was not possible to determine the pitch of the last sound because of noise from the infant's clothes.

The acoustic analysis of the infant's sounds and the father's singing supports my perception: The infant's initial “Rain Rain Go Away” vocalization resembles the

intervals and rhythm of the song, and the infant's imitation in response to the father matches the father's notes.

"Happy Birthday"

"Happy Birthday" Melody from Electronic Toy (8:40 a.m.)

Description. James is by himself, playing with an electronic music toy. The toy plays one of two songs: the traditional song "Happy Birthday," and an unfamiliar song about birthdays ("Birthday Song 2"). The toy's version of "Happy Birthday" consists of synthesized keyboard sounds for the melody and has no lyrics. The melody is in the key of G major, and the starting pitch is D4 (Figure 13). The melody lasts 11 seconds from beginning to end. "Birthday Song 2" consists of synthesized sounds for the melody and accompaniment, and an artificial-sounding child's voice singing the lyrics. James did not imitate this song.

Figure 13. Music notation of "Happy Birthday" melody from electronic toy

James plays with the toy for 10 minutes. He does not vocalize while playing with the toy. During the 10 minutes, the infant listens to 11 executions of the entire "Happy Birthday" song, and one partial execution that lasts 7 seconds, for a total of 128 seconds. During the remaining 8 minutes of play, the toy plays "Birthday Song 2" and a short recording of a child's voice speaking about birthdays and blowing out birthday candles.

James also hears his mother sing the first six notes of “Happy Birthday” for 4 seconds while the toy plays the other song (“Birthday Song 2”). The mother sings with “doo” syllables instead of words.

Infant Imitation of “Happy Birthday” (6:45 p.m.)

Description. It is almost time for James to go to sleep. The father has been singing the song “Rain Rain Go Away” to James and is still next to him as the mother approaches them. When the mother is near, the infant vocalizes, making sounds that resemble variations of “ma.” During a few seconds, both parents play with James by imitating his syllables. Still near the infant, the parents begin to discuss the father’s upcoming job-related engagement. While they are talking amongst themselves, James vocalizes nine distinct sounds. The sequence of vocalizations lasts approximately five seconds.

The first six sounds have equal duration (Figure 14). A pause of approximately the same sound duration separates the third and fourth sounds. The six sounds are pitched (Figure 14). The last three sounds of James’s vocalization have unclear pitch and are depicted with an “x” as the note head in Figure 14. James produces the first three sounds with different vowel sounds as indicated in the figure.

Figure 14. Music notation of infant's "Happy Birthday" imitation (perceptual analysis)

His vocalization seems to go unnoticed by the parents, who continue their private talk. Soon after, the father leaves and the mother begins to talk and read a book to the infant as she prepares him for bedtime.

Acoustic analysis. The recording captured the infant's vocalization as well as the parent's voices. Before measuring frequency and duration with Praat, iZotope's noise removal and spectral repair tools were used to reduce the intensity of the parent's voices. Figure 15 shows the spectrograms and pitch contours of the original audio and of the edited audio. The brackets below the figures indicate the infant's vocalizations. In the spectrogram of the original audio, the pitch contour in blue does not always correspond to the actual pitches of the infant's sounds as experienced by the listener. The pitches shown for the second and third vocalizations, for example, are lower than James's notes. This error in pitch extraction sometimes occurs when the frequency range of the recorded voices is unusually large, such as the range between James's high fundamental frequency and the father's low fundamental frequency. To correct this, the pitch extraction settings were changed. The corrected pitch contour is shown in the spectrogram image of the edited version of the audio. The contour has sharper spikes compared to the contour of the original audio, which is possibly an artifact of the altered pitch settings.

Figure 15. Spectrograms of infant's "Happy Birthday" imitation (original audio and edited audio)

The pitches of the vocalization are shown as musical notation in Figure 16. All sounds except for the last match the pitches of the melody from the toy. James's first vocalization is divided in two parts based on the vowel quality. It is represented as two separate notes, each eighth note indicating a different vowel sound. The pitch of the first part of the vocalization, in which James produces an "ah" sound, is 300.42 Hz (D4 +39 cents). The pitch of the second part, in which he produces a vowel sound closer to "eh," is 291.37 Hz (D4 -14 cents). The fourth sound begins at 424.44 Hz (G#4 +38 cents) and drops to 405.59 Hz (G#4 -41 cents). Although the sound begins on a high G#, it drops 79 cents by the end – a difference close to one semitone. It is represented as a G natural to facilitate the visual comparison between James's vocalization and the music from the toy.

The seventh sound begins at 318.61 Hz (D $\sharp$ 4 +41 cents) and rises to 320.28 Hz (E4 -50 cents). It is represented as an E to facilitate comparison.

Figure 16. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of pitches from acoustic analysis of infant's "Happy Birthday" imitation

Pitch analysis of James's first sound prompted the investigation into James's vowel sounds. The first sound maintains the same pitch, yet it seems to be divided into two parts that corresponds to the initial rhythm of the melody of "Happy Birthday." To investigate further, vowels were extracted from the audio using the Vocal Toolkit plugin for Praat (www.praatvocaltoolkit.com). It must be noted that the description of infants' vowels, and infant vocalizations in general, is difficult and requires training. Infants' vowel space is different than the vowel space of adults, and infants are not yet capable of rapid formant transitions. Nevertheless, a rough acoustic description of James's vowels is of interest because it provides an initial glimpse into the parameters of vocal production that he is capable of manipulating. Additionally, it may reveal how such manipulations relate to the original sounds that he was imitating.

By tracking the formants³ and listening to the isolated vowels of James's vocalization, it became clear that James's sounds resemble the vowels of the lyrics to the song "Happy Birthday." The acoustic analysis shows patterns of formant changes that are remarkably similar to what occurs in those lyrics. In Figure 17, the formants of the first three sounds are represented with red dots. At the beginning of the first sound (1a in Figure 17), F1 is at its highest, which suggests a low vowel sound. In 1b, F1 is lower and F2 higher, and the distance between the two is considerably large. This suggests that James is now producing a higher and more fronted vowel. In the second sound, F1 remains stable. F2 lowers considerably and the distance between F1 and F2 decreases. From 1b to 2, then, the sound becomes more like a mid central vowel. In the third sound, F1 is slightly lower than in the second sound, and the distance between F1 and F2 increases as F2 becomes higher. As a result, the vowel becomes slightly higher and more fronted.

³ According to Joliveau, Smith, and Wolfe (2004), "the different vowel sounds in normal speech are produced by adjusting the position of the tongue, jaw and lips so that the vocal tract resonates at certain specific frequencies (R_i). The vocal folds vibrate at the pitch frequency, f_0 , which is usually lower than the R_i . Those harmonics of f_0 that fall near the R_i produce associated peaks, or formants (F_i), in the speech spectrum" (p. 116). Vowels are identified primarily based on the relationship between the locations of the three lower formants F1, F2, and F3 (Gelfand, 2009).

Figure 17. Formant structure of infant's "Happy Birthday" imitation

The pattern of formants described above are similar to the formants of the vowels of the word “happy” (the near-low front vowel /æ/ and the high front vowel /i/) and the word “birthday” (the low-mid central vowel /ɜ/ and the high-mid front to near-high near-front diphthong /eɪ/). Importantly, the vowel sounds used by the infant occur with the exact notes where one would expect to hear the syllables of the lyrics: “happy” with the first two D notes and “birthday” with the following E and D. This is remarkable, especially considering that James did not hear the melody of “Happy Birthday” sung with lyrics on that day. The melody from the toy did not include lyrics, and during the four seconds in which the mother sang the melody while the other song was playing, she used “doo” syllables.

Music Intervals

Two instances of imitation are described below. In the first one, James imitates the last two notes of his sister Diane’s improvisation. In the second one, James overhears his brother Dale sing two notes that form an interval of a descending major third. James vocalizes two utterances, each consisting of two tones. In the first utterance, James

imitates the major third interval but not the exact pitches. In the second utterance, James produces a gliding and continuous pitch contour with an initial pitch peak that corresponds to Dale's first note, and a lower pitch peak that corresponds to Dale's last note.

Imitation of Sister's Improvised Melody (8:07 a.m.)

Description. It is shortly after 8 a.m., and the family has just finished breakfast. James is in the kitchen with the father, who is cleaning up the dishes. The siblings are running around the house. Dale screams from the other room, Diane responds with another scream, and then James screams too from his seat in the kitchen. The father replies to James with playful sounds. In the background, Diane suddenly begins to sing an improvised melody using nonsense syllables. The melody is very short and consists of only six notes. It is difficult to identify some of the notes of the improvisation because Diane is farther away from the infant. Because of the distance from the recording device, Diane's voice is softer and more susceptible to interference from overlapping sounds such as the voices of the father and Dale and the sounds of dishes in the sink. Some notes are not as affected by the interfering sounds, such as the first note (E4) and the last two notes (E4 and B3, respectively). A tentative sketch of the improvised melody is shown in Figure 18.

One second after Diane finishes her improvisation, James vocalizes a descending interval pattern on /n/ sounds that resembles the perfect fourth interval formed by the last two notes of Diane's song (E4 and B3). James does not match the higher note of Diane's perfect fourth interval, producing a sound that is close to G#4. The second note James

produces is B3, identical to Diane's last pitch. James's two sounds have the same durations as the corresponding notes in Diane's song.

Figure 18: Music notation of sister's improvised melody (perceptual analysis)

No one acknowledges James's vocalizations. He continues producing /n/ syllables but the resulting sounds do not incorporate noticeable music features.

Acoustic analysis. The recording of Diane's improvisation and James's imitation contains many overlapping voices and sounds from the kitchen, making it difficult to obtain reliable acoustic measures. Dale's voice in particular is problematic because it is loud and high-pitched. Noise removal tools were used to improve the quality of the sounds, but in some cases the procedure was not effective in isolating the desired voices. It was not possible to measure the frequency and duration of the second and fourth notes of Diane's melody.

The pitches of Diane's song and the two sounds produced by James are shown in Figure 19. The pitches of the last two notes of the melody, which form a perfect fourth interval, are 331.21 Hz (E4 +8 cents) and 244.23 Hz (B3 -19 cents). The pitches of James's first and second sounds are 404.48 Hz (G#4 -46 cents) and 251.71 Hz (B3 -33

cents). The acoustic analysis for this excerpt is consistent with the perceptual description of Diane's melody and James's sounds presented earlier (see Figure 18): James's vocalization resembles the ending of Diane's improvisation.

Figure 19. Frequencies (Hz) and music notation (rounded to the nearest quarter tone, equal temperament tuning) of pitches from acoustic analysis of sister's improvised melody and infant's imitation

Imitation of Brother's Music Interval (8:12 a.m.)

Description. It is 8:12 a.m., approximately five minutes after James vocalized in response to Diane singing in the background. The siblings are playing, and James is by himself in the kitchen. While playing, Dale produces a very loud shriek followed by two syllables of equal duration that resemble “wah-wah” sounds. The notes of the two syllables are F4 and D $\flat$ 4, forming a descending major third interval. One second later, James vocalizes two vowel sounds similar to “ia-ia.” The interval between the pitches of James's vocalizations resembles Dale's descending interval, although the pitches are not the same. This is particularly evident for James's first sound, which is clearly higher than Dale's. However, James maintains the descending major third interval by raising the

pitch of the second note as well. The durations of the sounds produced by the infant are also very close to the durations of Dale's syllables.

James vocalizes another two sounds that also outline a descending pitch pattern. Unlike his preceding vocalization, however, the two target pitches seem to be connected by a continuous pitch glide from the first sound to the second one. Additionally, the two sounds in the second utterance differ from the sounds of the first utterance in terms of duration. Whereas in the first pair of sounds both sounds were of similar duration, in the second pair the first sound is much shorter and the second sound longer.

Acoustic analysis. Figure 20 shows the spectral and intensity measures of Dale's two syllables and James's two utterances in response. The duration, frequency, and equivalent pitch in semitones are indicated below each sound. The pitches of Dale's two syllables, 347.81 Hz (F4 -7 cents) and 277.00 Hz (D $\flat$ 4 -1 cent), form an interval of a descending major third. Both notes have approximately equal duration.

Figure 20. Spectrogram of brother's interval and infant's imitations (blue line indicates pitch contour; yellow line indicates intensity); and duration (s), frequencies (Hz), and equivalent pitch based on equal temperament tuning (rounded to the nearest quarter tone)

James's first imitation of the interval consists of two sounds. The pitch of the first sound begins at 353.00 Hz (F4 +19 cents), but quickly ascends to 378.85 (F#4 +41), where it plateaus. The intensity of the part of the sound that corresponds to F#4 is greater compared to the intensity associated with F4. The pitch of James's second sound, 292.34 Hz (D4 -8 cents), forms an interval of a descending major third with the previous F#4. James's two sounds are approximately equal in duration, but slightly shorter than Dale's two notes.

In the second utterance, the frequency peaks at 347.09 Hz (F4 -11 cents) and then glides down to 272.69 Hz (D₄ -28 cents). These pitches match those produced by Dale. As indicated by the pitch contours in Figure 20, the contour of the first utterance seems to outline a clear peak, whereas the contour of the second utterance is more continuous and with more gradual changes in frequency values.

The acoustic analysis for this excerpt is consistent with the perceptual description of the brother's interval and James's sounds presented earlier: James's two vocalizations resemble the music interval sung by the brother.

Chapter 5: Validation Study

I identified James's music imitations through a process of repeated listening and comparison of sounds, including the ones produced by James and those that had occurred many hours earlier. After so much time listening to short segments of audio and comparing them, I started doubting my own ears. Was James truly imitating melodies that he had heard hours earlier? Perhaps listening to the segments so intensely had made me perceive similarities that were just not there. Such phenomena, in which certain spoken phrases sound like music after a certain number of repetitions, have been shown to occur in adult listeners (Deutsch, 2010; Deutsch, Henthorn, & Lapidis, 2011; Falk, Rathke, & Dalla Bella, 2014). It should be noted that I did have some evidence that my perception of the similarities between James's vocalization and music he had heard earlier was not of my own fabrication; this perception was shared by someone else: the father. During the "Rain Rain Go Away" vocalizations that James produced while interacting with him, the father identified James's sounds as similar to the melody of the song.

During the many weeks that I observed this baby and his family, I frequently reflected on my findings with others, including extensive discussions with Dr. Eugenia Costa-Giomi. During our talks, I had the opportunity to explain the context of the vocalizations, to describe in detail the features of the sounds that I identified as imitation,

and to replay the audio segments as many times as needed. It could be the case that I also influenced Dr. Costa-Giomi with my perceptions of the infant's vocalizations. Would other adult listeners, unfamiliar with the voices of this family and with the task of comparing audio excerpts for similarity, also agree with my perception?

A validation study was conducted to assess how naïve adult listeners would respond to the audio excerpts of the imitations. The study was designed so that participants listened to James's imitations and to the models and compared them. For some questions, participants were presented with the imitation and asked to identify the model from a set of options; for other questions, they were presented with the model and asked to identify the imitation.

There were a number of problems that needed to be addressed when designing the study. One concern was that the vocalizations were very short and somehow strange; infant babbling was probably unfamiliar to those without child rearing experience. Another concern was that listeners would be confused as to the task they were asked to complete considering its ambiguity (i.e., perception of similarities). To address these issues, participants first completed a set of five questions for familiarization purposes. These introductory items were designed to acquaint the participants with tasks that involved listening to short audio clips with overlapping voices and allocating attention to individual voices. Additionally, some of the study questions included comparisons between James's sounds and words spoken by other family members. Speech items were considered to be a more familiar type of stimuli for most listeners to identify, and would thus be helpful for understanding the nature of the task.

I did not include all the imitations that I had identified in the validation study, and I only presented to listeners those instances of imitation that I considered especially significant and representative of James's vocalizations. Additionally, because the study was completed online, I sought to maintain the duration of the study between 10 and 15 minutes to maximize the number of respondents.

Method

Participants

Participants were recruited from the School of Music at the Ohio State University, as well as by personal invitation in cases in which the participant was not affiliated with the School of Music or the university. They were included in the study if they agreed to participate and if they were between the ages of 18 and 65. Participants completed the study anonymously, and 25 received partial course credit in a Music Theory/Aural Skills undergraduate course. The study was part of the lab protocol approved by the Institutional Review Board.

The study was created using Qualtrics and presented online from November 16, 2016 to December 15, 2016. The number of participants who completed the study was 66. Questions were timed to assure that participants were spending a reasonable amount of time listening to the audio excerpts, and participants were excluded if they spent less than 5 minutes completing the study ($n = 1$). The final number of participants included in analysis was 65. Of these participants, 39 were undergraduate music majors.

Procedures

The study was administered online. The completion of the consent form and questionnaire took approximately 15 minutes. The format of the online questionnaire consisted of three sections (see Appendix B for an outline of the study and questions): (1) introduction and familiarization, (2) test, and (3) feedback. In the first section, participants were given a brief introduction to the study. They were told that they would be listening to short audio excerpts involving children and babies, and that they would answer questions about what they heard. After the introduction, participants answered five multiple choice questions for familiarization purposes. The questions involved tasks such as listening to an audio excerpt of the family playing a board game and identifying the number of different voices they heard. Participants were also asked to identify words spoken by the infant (“bye daddy” and “mommy”) and by the brother (“he has no teeth”). To familiarize them with music stimuli, participants were presented with an excerpt of James’s sister singing “London Bridge” and asked to match it to the correct melody from excerpts of melodies played by the electronic toys.

The second section of the study – the test – consisted of three blocks of multiple choice questions. The order of block presentation was the same for all participants, but the order of presentation of the questions within each block was randomized. For each question, the order of choices was also randomized. Table 5 lists the test items and response options presented to participants in each test question.

Table 5. Test items and response options

Item presented	Options for comparison
1) Infant: HB vocalization	a) Toy: “Happy Birthday” b) Toy: “Birthday Song 2” c) Toy: “London Bridge” d) Toy: “Bingo” e) Toy: “Ten Little Indians”
2) Infant: RRGa vocalization (before father sings)	a) Father: RRGa b) Brother: “Bingo” c) Mother: “Bah Bah Black Sheep”
3) Infant: RRGa vocalization (before father sings)	a) Mother: RRGa b) Sister: “London Bridge” c) Father: “Abiyoyo”
4) Father: ID speech (“blue car”)	a) Infant: vocalization before father’s words b) Infant: vocalization after father’s words c) Infant: unrelated vocalization
5) Brother: music interval	a) Infant: first vocalization after brother b) Infant: unrelated vocalization c) Infant: unrelated vocalization
6) Brother: music interval	a) Infant: first vocalization after brother b) Infant: second vocalization after brother c) Infant: unrelated vocalization
7) Sister: sentence spoken in the background	a) Infant: vocalization after sister b) Infant: unrelated vocalization c) Infant: unrelated vocalization
8) Father: ID speech (“blue car”)	a) Infant: vocalization before father’s words b) Infant: vocalization after father’s words
9) Infant: RRGa vocalization	a) Mother: RRGa b) Father: RRGa
10) Brother: music interval	a) Infant: first vocalization after brother b) Infant: second vocalization after brother
<i>Note.</i> HB = “Happy Birthday”; RRGa = “Rain Rain Go Away”	

The first block (Items 1 to 7, Table 5) included seven questions in which participants listened to audio excerpts and were asked to make similarity judgments between them. Participants listened to the test item and then to three or five possible responses. They were to choose which response most resembled the test item. In some cases, the infant's imitation was presented as the test item (e.g., Item 1, Table 5) and in others, the model was the test item (e.g., Item 5, Table 5). The infant's imitations included in this block consisted of imitations of music ("Happy Birthday," "Rain Rain Go Away," and the brother's music interval) and of speech (stress pattern of ID speech from the father and the sister's intonation contour).

In the second block of the test, two items and response options from the first block were presented again: Item 1 ("Happy Birthday" imitation) and Item 7 (sentence spoken by sister). The questions were repeated to check if participants were consistent with their answers and attentive to the task.

Each of the three questions in the third block consisted of a test item presented previously in the first block. But whereas in the first block participants had three possible responses to choose from, in the third block they had only two options (Items 7 to 10, Table 5). The response options were chosen with the objective of isolating important comparisons for participants. For example, when presented with James's "Rain Rain Go Away" (Item 9, Table 5), participants were forced to choose between the model of the mother singing "Rain Rain Go Away," in the key of A, and the father singing the same

song in the key of G. The direct comparison between the two parents' singing had not been available to participants in the first block.

In the third and last section of the study, participants were asked to give feedback regarding the clarity of instructions, the quality of the audio excerpts, and their level of involvement in answering the questions. They were also given the opportunity for leaving suggestions and general comments if desired.

Results and Discussion

Reliability was assessed by comparing participants' responses to items that were presented twice (Items 1 and 7, Table 5), once in block one and again in block two. For each of the two items that appeared twice, I calculated the proportion of participants that gave the same response (i.e., number of same responses divided by the total number of responses). Reliability was 0.74 for Item 1 and 0.88 for Item 7.

In the next sections, the test items and responses of seven questions that were included in the study are described in detail. The seven questions were based on the following imitations: "Happy Birthday," "Rain Rain Go Away," and the brother's music interval.

"Happy Birthday" Imitation

Shortly before going to bed, James produced a sequence of sounds that resembled the pitches of the beginning of the melody of "Happy Birthday." In the morning, 10 hours before the vocalization in question, James had played with a toy that produced this song. The objective of this question in the study was to assess if participants agreed with my judgment of similarity between the infant's vocalization and the music from the toy.

Participants compared James's "Happy Birthday" vocalization with five excerpts of music. I opted for maintaining ecological validity by presenting listeners with the data that I had had as reference for making my own judgment. Thus, I decided against creating stimuli that would allow for controlling any confounders, such as tonality, starting pitch, or range. The five possible responses consisted of short segments of the following tunes, which were produced by two different toys with which the infant had played during the day (Table 6): "Happy Birthday," "Birthday Song 2" (an unknown song from the same toy that played "Happy Birthday"), "London Bridge," "Bingo," and "Ten Little Indians."

Table 6. Characteristics of audio excerpts presented to participants in the "Happy Birthday" question

Excerpt	Duration (s)	Key	Starting Pitch	Lyrics?
Happy Birthday	3.0	G	D4	No
Birthday Song 2	3.7	C	C4	Yes
London Bridge	3.0	C	G5	No
Bingo	3.8	C	G4	No
Ten Little Indians	3.0	C	C5	No
Infant vocalization	3.5	G	D4	—

These excerpts were chosen because the melodies could serve as ecologically valid controls with which to compare the infant's vocalization. All melodies had been part of the infant's active listening experience during the day from extensive playing with the two toys. Additionally, the melodies are all from typical children's songs and share

many melodic turns and rhythmic features. This reduced the possibility that participants' responses would be drawn to one melody over the others because of salient differences in style or music content. Lastly, the use of melodies in the form of synthesized sounds from James's toys was a way to ensure that participants would not favor sounds that were produced by a human voice, instead of attending to the actual melodies. This addressed the concern that given options that included both sounds from toys and sounds from human singing, participants might opt for an excerpt of singing because it was produced by a human and not because the melody resembled James's vocalization.

The question was presented twice, once in the first block and again in the second block. Table 7 shows the results from both presentations of the question. Chi-squared analyses showed that there was a significant difference in participants' choice of responses (first presentation: $\chi^2(4) = 83.385, p < .0005$; second presentation: $\chi^2(4) = 148.615, p < .0005$). Most selected the excerpt of "Happy Birthday," and less than 14 percent chose any of the remaining responses. In fact, for the second presentation of the test item, 80% of the listeners identified James's vocalization as "Happy Birthday."

In summary, the results for the "Happy Birthday" question show that participants agreed with the perceptual and acoustical analyses of the infant's imitations. When participants were given a choice between five melodies that James had heard while playing with electronic toys, most opted for the response with the "Happy Birthday" melody as most similar to the infant's vocalization.

Table 7. Participants' response choices for the two presentations of the "Happy Birthday" item

Responses	No. of Participants	
	First Presentation	Second Presentation
Happy Birthday	42 (64.6%)	52 (80.0%)
Birthday Song 2	2 (3.1%)	1 (1.5%)
London Bridge	4 (6.2%)	0 (0.0%)
Bingo	9 (13.8%)	6 (9.2%)
Ten Little Indians	8 (12.3%)	6 (9.2%)
Total	65	65

“Rain Rain Go Away” Imitation

The mother sang the song to James before his nap, at around 12:15 p.m. James’s vocalization occurred more than 6 hours later. The father heard James and acknowledged that it sounded like the song “Rain Rain Go Away.” These excerpts – the mother singing, James vocalizing, and the father singing – were included in the validation study. Without any experience with the voices of the members of the family, and without any contextual information, would listeners identify all three excerpts as the same song? Without any temporal information as to how the events unfolded, would listeners perceive James’s vocalizations as more similar to the mother singing or to the father?

The study included three questions that asked listeners to compare James’s “Rain Rain Go Away” vocalization with a set of singing excerpts. Table 8 lists the characteristics of the excerpts included in the questions. In the first question, the three

excerpts for comparison consisted of the father singing “Rain Rain Go Away” and two instances of family members singing other songs: Dale singing “Bingo,” and the mother singing “Bah Bah Black Sheep.” The results are shown in Table 9. Chi-squared analysis showed that there was a significant difference in participants’ choice of responses ($\chi^2 (2) = 91.969, p < .0005$). Nearly 90 percent of participants chose the excerpt of the father singing “Rain Rain Go Away.”

Table 8. Characteristics of audio excerpts presented to participants in "Rain Rain Go Away" imitation questions

Excerpt	Origin	Duration (s)	Key	Starting Pitch
Rain Rain Go Away	Father	3.2	G	D3
Rain Rain Go Away	Mother	3.8	A	E4
Bingo	Brother	2.8	B	D#4
Bah Bah Black Sheep	Mother	3.7	G	G3
London Bridge	Sister	2.9	B	F#4
Abiyoyo	Father	2.9	A $\flat$	D $\flat$ 3
Infant vocalization	Infant	3.8	A $\flat$	E $\flat$ 4

The second question that required the participants to identify James’s “Rain Rain Go Away” imitation provided them with three excerpts from which to choose: the mother singing “Rain Rain Go Away,” the sister singing “London Bridge,” and the father singing “Abiyoyo” (Table 8). The results are shown in Table 9. Chi-squared analysis showed that there was a significant difference in participants’ choice of responses ($\chi^2 (2) = 101.754, p$

< .0005). More than 90 percent of participants chose the excerpt of the mother singing “Rain Rain Go Away.”

Table 9. Participants' response choices for the "Rain Rain Go Away" comparisons

Question 1		Question 2		Question 3	
Responses	No. of Participants	Responses	No. of Participants	Responses	No. of Participants
Father: “Rain Rain Go Away”	58 (89.2%)	Mother: “Rain Rain Go Away”	60 (92.3%)	Mother: “Rain Rain Go Away”	25 (38.5%)
Brother: “Bingo”	6 (9.2%)	Sister: “London Bridge”	3 (4.6%)	Father: “Rain Rain Go Away”	40 (61.5%)
Mother “Bah Bah Black Sheep	1 (1.5%)	Father: “Abiyoyo”	2 (3.1%)		
Total	65	Total	65	Total	65

The third question about James’s “Rain Rain Go Away” imitation had only two possible responses, the mother and the father singing “Rain Rain Go Away.” Chi-squared analysis failed to show a significant difference in the proportion of respondents who chose each excerpt: almost 40 percent of participants opted for the mother, and approximately 62 percent opted for the father, $\chi^2(1) = 3.462, p > .05 (p = .063)$.

In summary, the results for the three “Rain Rain Go Away” questions support those of the perceptual and acoustic analyses. Participants matched James’s vocalization

of “Rain Rain Go Away” to the models of the song produced by the mother and the father. Given a choice between the father singing “Rain Rain Go Away” and two other instances of singing by family members, almost 90 percent of participants opted for the excerpt of the father singing the song. When the choices for comparison included the mother singing “Rain Rain Go Away” and two other instances of singing by family members, more than 90 percent of participants opted for the excerpt of the mother. When participants were forced to choose between the mother and the father, they were evenly split between the two parents. These results suggest that participants identified all three excerpts – James, mother, and father – as being similar.

Imitation of Brother’s Interval

James heard his brother in the background singing two notes. James then produced two pairs of sounds that resembled the brother’s notes. The pitches of the first pair were not the same as the pitches of the second pair, however. Based on the pitch extracted from acoustic analysis, James’s second pair of sounds matched the brother’s exact pitches, while the first pair of sounds was one semitone higher than the brother’s. Would other adults identify both pairs of sounds as imitations of the model? Three questions were designed to address this issue.

The objective of the first question was to assess if participants matched James’s first pair of sounds with the brother’s interval. Participants listened to an excerpt of the brother’s interval and three other excerpts for comparison: James’s first pair of sounds, and two other pairs of sounds that James had produced during the day but in a different context. Chi-squared analysis showed that there was a significant difference in

participants' choice of responses ($\chi^2 (2) = 96.492, p < .0005$). Of a total of 65 participants, 59 (90.8%) selected the excerpt with James's first pair of sounds, 3 (4.6%) selected the excerpt with an unrelated pair of sounds, and 3 (4.6%) selected the excerpt with the other unrelated pair of sounds. In other words, more than 90 percent of participants judged James's first imitation as being most similar to the brother's model.

Participants were presented with the brother's vocalization again in the second question. The excerpts from which they could choose consisted of James's first pair of sounds, James's second pair of sounds, and one pair of sounds produced in a different context. Chi-squared analysis for this question also showed that there was a significant difference in participants' choice of responses ($\chi^2 (2) = 30.400, p < .0005$). Of a total of 65 participants, 35 (53.8%) selected the excerpt with James's first pair of sounds, 29 (44.6%) selected the excerpt with James's second pair of sounds, and 1 (1.5%) selected the excerpt with the unrelated pair of sounds. In other words, most participants judged James's first imitation as being most similar to the brother's model, although a considerable number of participants opted for the second imitation as most similar to the model. This pattern of results lends support to my perception of both vocalizations as imitations of the brother.

The third question was a direct comparison between James's first pair of sounds and his second pair. This question was presented in the last block of the survey so that it would not influence participants' decisions in the previous two questions. Chi-squared analysis failed to show a significant difference in the proportion of respondents who chose each excerpt: Of a total of 65 participants, 35 (53.8%) selected the excerpt with

James's first pair of sounds, and 30 (46.2%) selected the excerpt with James's second pair of sounds, $\chi^2(1) = .385, p > .05$ ($p = .535$). When forced to choose between James's first pair and second pair of sounds, participants did not favor one response over the other, suggesting that to a naïve listener, both sounded similar to the model.

In summary, the results of the validation study with adults support the conclusion that James did indeed imitate music that he had heard. Although participants heard very short excerpts and were given no contextual information, they matched the models for "Rain Rain Go Away," "Happy Birthday," and the music interval with the infant's imitations.

Chapter 6: Discussion and Conclusion

The present case study investigated infants' music production by identifying and analyzing instances of vocal imitation of music produced by one infant throughout one day. James, a 15-month-old infant, wore a portable recording device that captured 16 continuous hours of sounds produced near him as well as sounds from his own vocal production. The results show that James imitated sounds that he had heard previously. James imitated phonetic and prosodic features of speech, as well as music features of songs and melodies, such as pitches, intervals, melodic contour, and rhythms. The music that he imitated consisted of infant-directed singing, singing that was heard in the background, and a synthesized tune from a toy. Some imitations occurred while he was playing by himself, and others while he interacted with an adult. In some instances, the imitations occurred hours after hearing the model. In summary, it is clear that James was capable of producing music, as demonstrated by his vocal imitations of melodies.

Four instances of vocal imitation of music by James were described and analyzed in detail: "Rain Rain Go Away," "Happy Birthday," a music interval produced by the sister Diane, and a music interval produced by the brother Dale. I will first review and interpret the analyses of each of the imitations and then discuss the findings in the context of infant music development and infant vocal imitation. Lastly, I will address the limitations of the study and implications of the findings.

Results

“Rain Rain Go Away” Imitation

The “Rain Rain Go Away” imitation (hereafter, Rain Rain imitation) consisted of a sequence of sounds produced by the infant that matched the intervals and rhythm of the song, which his mother had sung to him 6 hours earlier that day. When James produced these sounds, the father, who was next to the infant at that moment, heard the infant’s vocalization, identified the melody, and proceeded to sing the song to him. James then imitated the father’s rendition of the song. Each family member – James, the mother, and the father – initially sang the song on a different starting pitch and in a different key. James’s imitation of the mother was a half-step lower than the mother’s model, and his imitation of the father matched the father’s starting pitch. In addition to imitating the starting pitch, James also imitated the minor third interval structure and the rhythmic patterns of the “Rain Rain Go Away” melody. My perception of James singing, the acoustical analysis of the stimuli, and the validation study support these findings. James’s deferred imitation of the melody suggests that he remembered these music features and purposefully recalled them hours later without hearing the model.

“Happy Birthday” Imitation

The “Happy Birthday” imitation (hereafter, Happy Birthday imitation) consisted of a series of six pitched sounds that matched the pitches of the “Happy Birthday” melody produced by an electronic toy with which James had played 10 hours earlier that day. The melody from the toy was in the form of synthesized sounds, and it did not contain lyrics. Although both parents were near the infant when he imitated the sounds,

neither of them seemed to have noticed such remarkable accomplishment. The lack of encouragement and feedback from the parents raise the possibility that James's motivation for vocalizing was non-social: James did not persist in vocalizing to call for their attention. Instead, he remained quiet for a few seconds after he finished imitating. Perhaps the infant vocalized as a form of vocal play that was not intended for communicative purposes. My perception of James singing, the acoustical analysis of the stimuli, and the validation study support these findings. James's deferred imitation of the melody suggests that he not only remembered the music features of the model, but also recalled and reproduced them at will hours later.

An interesting finding from the analysis of the formant structures was that the relationships between the first, second, and third formants of the infant's first three sounds were similar to those that would be expected for the vowels of the words "happy birthday." James did not hear the song with lyrics on that day, so it is likely that he remembered this information from exposure to the lyrics prior to the day of the recording. Despite the different timbre and the absence of words, James heard the electronic toy and recognized its melody as being the same as a song that was sung to him on a previous day.

Imitation of Music Intervals

James imitated two music intervals, one produced by the sister and the other by the brother. The first interval consisted of the last two notes of an improvised melody sung by the sister in the background. When the sister finished the song, James immediately vocalized two sounds that resembled the final interval. The second interval

was produced by the brother in the background, approximately five minutes after the sister's improvisation. James, who was playing by himself, vocalized two pairs of sounds that resembled the brother's major third interval. James's first pair of sounds in response to the brother formed an interval of a major third, but the interval was a half step higher than the brother's model. The pitches of James's second pair of sounds were identical to the brother's model. My perception of James singing, the acoustical analysis of the stimuli, and the validation study support these findings.

James did not interact with anyone when imitating his siblings and did not seem intent in initiating any interactions. James's spontaneous and private imitation of the music intervals produced far away from him suggests that perhaps he was simply practicing or enjoying the vocalizing of music sounds he had heard. Throughout that day, neither sibling engaged with James in a significant way; the only direct interactions between them involved yelling at the infant to stop doing something. Thus, the fact that he imitated his siblings suggests that James was not restricted to imitating only sounds produced by a significant caretaker or only sounds that were intended for affective turn-taking.

In conclusion, the results of the perceptual description, the acoustic analysis, and the validation study with adults indicate that James engaged in acts of vocal imitation of music. He imitated various features of the music models, such as absolute pitches, interval structure, and rhythms. That the imitations occurred hours later suggests that he represented these features and recalled them at will without the presence of the model. The acts of imitation were not restricted to social contexts, considering that they also

occurred in situations in which the infant was not interacting with anyone. Given the diverse contexts and interactional settings in which James imitated music, it is likely that his motivations to vocalize were also diverse.

General Discussion

James's music imitations are significant from several perspectives: (a) production of music sounds, (b) perception of melodies, (c) mechanisms involved in vocal imitation, and (d) motivations for imitating and making music. Each perspective is discussed below.

Production of Music Sounds

James produced sounds that were good exemplars of music. He vocalized clearly articulated sounds with stable pitches. In the Happy Birthday imitation, for example, the sounds were perceptually distinct, the timbre was rich and resonant, and the pitches were easy to identify. James also had good control of the timing of his breaths and the duration of the sounds. In the Rain Rain imitation, for instance, his breaths occurred at the end of each measure of the song, after the longer note. Though his executions were not always accurate, James's timing of breaths in this example suggest that he had prepared the movements necessary for producing a continuous sequence of sounds with coherent rhythms.

The fact that James was successful in producing vocalizations that sounded like music indicates that he had motor representations of the movements necessary to produce those sounds. It is possible that the motor representations are a product of auditory-motor mapping acquired during vocal play and imitation, similar to the articulatory-auditory mapping processes that occur in infant phonetic learning (e.g., Kuhl & Meltzoff, 1996;

Vihman & Boysson-Bardies, 1994). This model proposes that infants are constantly attempting to produce and match the speech sounds that they hear and store in memory. Infants become more accurate in their production as they discriminate speech sounds better and as they relate the sounds that they make when they practice with the pattern of movements that caused them. It is possible that this mechanism is also relevant to musical vocal sounds. For example, in one study 3-month-old infants were able to match rather reliably a pitch that was sung or played to them, after repeated exposure to the tones in daily sessions and vocalization practice (Kessen, Levine, & Wendrich, 1979). Such consistent imitation of specific pitches suggests that these infants had some form of control of their vocal tract movements, and that they associated specific patterns of movements with specific auditory outputs. The instance in which James was alone and attempted to match an interval that he heard his brother Dale produce in the distance could be considered an example of musical vocal play and imitation. By imitating Dale and by practicing this imitation, James might have been learning the consequences of the motor patterns and storing these associations, which would then be available for later use in different sequences of sounds.

Perhaps infants are better music-makers than previously acknowledged by some accounts (e.g., Davidson, 1985; Trainor & Hannon, 2013; Trehub, 2016; Trehub & Gudmundsdottir, 2016; Tsang, Friendly, & Trainor, 2011; Welch, 1994, 2014, 2016). James, at least, produced music sounds that were controlled and organized, and many of his vocalizations did not seem to be unstructured explorations of pitches and rhythms.

Perception of Melodies

Absolute and relative pitch. James's imitation of the exact pitches of "Happy Birthday" from the electronic toy indicates that he encoded the absolute pitches of the melody. The Rain Rain imitation indicates that he encoded relative pitch as well: He imitated the song in different keys. Transposing a melody requires extracting and representing the invariant properties of the music model. In music, the identity of a melody rests not so much on the exact pitches but rather on the interval structure represented by the distance between successive pitches. James's imitation of "Rain Rain Go Away" suggests that he perceived and recognized the structure of the two models sung by his parents in different keys.

Perception studies indicate that infants attend to both absolute pitch (AP) information (Saffran, 2003; Saffran & Griepentrog, 2001; Saffran, Reeck, Niebuhr, & Wilson, 2005) and relative pitch (RP) information (Lynch & Eilers, 1992; Plantinga & Trainor, 2005; Saffran et al., 2005; Trainor & Trehub, 1992). It seems that the preferential processing strategy – AP or RP – depends partly on the task and on the nature of the input. Infants demonstrated sensitivity to RP cues when they were given the opportunity to learn the melody during the week prior to the test (Plantinga & Trainor, 2005) and to familiarize themselves with transposed melodies (Trainor & Trehub, 1992). In statistical learning tasks, which involve discriminating stimuli that are shorter and sound less natural compared to melodies infants normally hear, infants attended to AP when all RP cues were removed from the tone stream (Saffran, 2003), and RP cues when all AP cues were removed (Saffran et al., 2005). When both RP and AP cues were

present, however, infants tended to rely on AP information (Saffran & Griepentrog, 2001). Adults have been also shown to process AP information to some extent, but it seems that adults' primary form of pitch encoding is RP-based (e.g., Saffran, 2003; Takeuchi & Hulse, 1993). This shift in pitch processing strategy suggests that there may be developmental changes, perhaps related to learning and enculturation.

Saffran, Reeck, Niebuhr, and Wilson (2005) suggested that what infants learn and encode depends on the information available to them: Although infants showed a preference for tracking AP probabilities, they also tracked RP probabilities when that was the only source of information for performing the task. If infants learn from what they are given, it may be the case that regular exposure to input that prioritizes relative pitch information provides important learning experiences. By hearing a melody performed in different keys, infants might consolidate RP processing skills that are so necessary for melodic categorization.

Singing to infants may be a naturally occurring source of such melodic variability. The songs that caregivers sing to infants are highly repetitive in melody and rhythm patterns, and are usually chosen from a limited repertoire of children's songs (Longhi, 2009; Trainor, 1996). By singing familiar melodies in different keys, caregivers provide exemplars from which infants learn the similarities of the melody despite differences in absolute pitches. Thus, exposure to transposed melodies – in James's case, "Rain Rain Go Away" sung by the mother and by the father – might promote melodic categorization based on relative pitch information. It is unlikely that a toy would provide the same source of variability. James heard many exemplars of familiar melodies from the toys,

but each melody was always reproduced in the same key (and all but one – “Happy Birthday” – were in the key of C Major).

Timbre. Singing also provides another type of variation that may be beneficial for promoting music learning: Timbre. Considering that identifying the invariant properties of words despite timbre variation of different talkers is fundamental for understanding spoken language (Houston & Jusczyk, 2000; Jusczyk, Pisoni, & Mullennix, 1992), timbre is a particularly important dimension of sound that infants must learn to discriminate and to ignore when necessary. Can infants also ignore timbre differences and recognize melodies played by different instruments or sung by different voices? Less is known about the effects of timbre variability on infants’ categorization of music, but it seems that infants behave differently depending on whether the melody is vocal or instrumental (Costa-Giomi, 2013). Infants categorize melodies that differ in vocal timbre at a younger age than they do with melodies that differ in instrumental timbre: Infants recognized the same melody sung by different women at 11 months, but only at 13 months did they categorize melodies played by different instruments (Costa-Giomi, 2013).

James’s Happy Birthday sounds matched both the exact pitches of the toy and the vowel sounds of the lyrics. This suggests that he perceived two instances of the melody as being the same song, despite considerable differences in timbre between the synthesized sounds of the toy and the vocal sounds of singing. In other words, James identified the melody and generalized not only across timbre variation, but also across other important differences such as presence/absence of speech sounds. James’s parents

probably sang the song many times to him; considering that infants' sensitivity to timbre variability develops earlier for vocal sounds than for instrumental sounds (Costa-Giomi, 2013), it is likely that James recognized the toy melody because he was already capable of identifying it across different voices. Furthermore, that James recognized the melody without the words suggests that he encoded music information as partially separate from the speech information of the lyrics. Perhaps singing to infants provides an important source of timbre variability. By first learning to recognize melodies sung by different people, infants can then do the same for non-vocal melodies.

Importance of singing to infants. Infant perception studies and the nature of James's imitative behavior provide support for the claim that singing models promote development of melodic categorization in infants. It is likely that singing *to* infants, not just playing music for them or in the background, is an important component for music learning. In one study, infants that were randomly assigned to music classes that involved singing activities with their parents demonstrated greater familiarity with culture-specific music structures after 6 months, compared to infants assigned to music classes that involved only passive listening (Gerry, Unrau, & Trainor, 2012). Similarly, recent studies with language acquisition suggest that verbal interactions with infants are more important for creating a rich language learning environment than mere exposure to a large quantity of speech in the household (Weisleder & Fernald, 2013). It is likely that passive exposure to music is not a substitute for infant-directed singing.

In fact, there is evidence that infants' responses to singing are enhanced if the auditory stimuli are coupled with videos of the person's facial movements as they sing.

For example, 11-month-old infants did not discriminate between the singing voices of unfamiliar women unless they were given the opportunity to watch a video of the singers' faces (Costa-Giomi, 2013). Similarly, infants in one study preferred ID singing over ID speech only in the audiovisual condition (Costa-Giomi, 2014). These findings suggest that perhaps it is important that infants associate the singing voice with the person who is producing it. Infants respond differently to singing that is produced especially for them, and it may be that such personal interactions promote richer experiences for music learning. Interestingly, James did not imitate "Birthday Song 2" – the only toy song that contained words – on that day. There are likely to be many explanations for why some melodies were imitated and others were not, but it seems reasonable to speculate that one reason may be that no one sang this song for him on a regular basis.

The advantages of cross-modal presentation of auditory stimuli have also been shown in the speech domain (e.g., Kuhl & Meltzoff, 1982). Findings from another speech-related study suggest that learning experiences are superior if the infant-directed information is presented *live*: Infants exposed to ID Mandarin Chinese speech from live speakers during lab sessions performed better on tests of Mandarin speech perception, compared to infants exposed to videos of the same speakers presenting the same information (Kuhl, Tsao, & Liu, 2003). Perhaps this pattern of results would also hold for music learning in infancy, which would lend further support for the importance of live singing to infants as a music learning experience.

Mechanisms Involved in Vocal Imitation

Imitation in infancy. The previous considerations assume that James truly attempted to imitate certain melodies that he had heard earlier in the day. But how do we know that James is responding to specific stimuli, as opposed to producing random sounds that are unrelated to the songs? And if his production is not random, how do we know that the processes underlying the vocal behavior are not solely the result of learning a stereotyped pattern of responses, one in which the imitations do not adapt to occasional changes in the model? The difficulty of determining the underlying processes of imitation has led to a considerable amount of research with newborns and very young infants. For some time, young infants were thought not to possess the perceptual and cognitive capacities for complex matching acts (Meltzoff, 1988; Meltzoff & Moore, 1977, 1983). The experimental designs of these studies attempted to distinguish imitation from two other types of responses: stereotyped acts that occur as a physiological response to a general arousal state, and acts that result from fixed stimulus-response linkages learned through contingent feedback or associative learning from repeated reinforcement.

Nature of the infant's music imitations. The contexts and contents of James's imitations suggest that his behavior was, in fact, a specific and intentional response to the musical characteristics of the models. James exhibited a pattern of differential responding, in which the form and features of his response depended upon the form and features of the model. For example, when his father sang "Rain Rain Go Away," James responded with sounds that incorporated music features of that song, as opposed to, for instance, responding with sounds that resembled "Happy Birthday." Thus, James did not

merely respond to a music model with vocalizations that resembled music; James responded to a model by incorporating the identifying features of that particular model. In addition, the imitative responses did not seem to be solely a product of conditioning. James's imitative acts occurred not only in communicative exchanges, but also in non-communicative contexts in which the behavior was not elicited nor rewarded – for instance, the parents did not encourage nor acknowledge his “Happy Birthday” vocalization. Furthermore, his acts of imitation occurred many hours later, in the absence of the auditory stimulus, suggesting that he recalled his representations of the model at will.

Perception and cognitive processes for imitating music. It is important to describe the nature of the music imitations because it is then possible to speculate about the underlying perceptual and cognitive processes that are likely to be involved. We know very little about infant music imitation and production. The lack of research raises the question of whether despite the well-developed music perceptual abilities during the first year of life, they are not yet capable of producing music. Production processes are much more complex than the ones necessary for perceptual discrimination of music features. A successful act of music imitation requires not only perceiving the features of the music signal, but also representing the features as a sound percept, organizing an appropriate motor response to match the percept, and comparing the vocal output with representations of the model sound (e.g., Berkowska & Dalla Bella, 2009; Pfordresher et al., 2015). Thus, James's imitations of music reflected sophisticated production processes that build on earlier skills of music perception. This suggests that the acquisition of music

production abilities occurs along a continuum during development. In other words, James was not simply matching sounds that he heard around him; James was also learning to make music.

Motivations for Imitating and Making Music

If James's vocalizations were intentional and organized responses to music models, then what was his motive for imitating? It could be that his primary motivation for this behavior was purely social in nature. Perhaps a vocal act that is music is only one means, out of many others, to engage in, maintain, or reinforce interactions with other people. This view is shared by Papoušek and Papoušek (1981), who propose that music functions primarily as a demonstration of affect between infants and caregivers, and that musical abilities develop so as to allow the infant to express internal affective states. And it may very well be, at least initially in early development, that infants vocalize musical sounds because they seem to be effective in creating and solidifying affective bonds with adults that are very important to them. In fact, James did produce music vocalizations in social interactions, and these seemed to have important non-musical functions. In the Rain Rain imitation with the father, for instance, the song drives and structures the dyadic turn-taking and imitation.

However, James also produced imitative music behavior privately and in non-communicative contexts. James imitated his brother while playing by himself. He also imitated synthesized sounds from an inanimate object, which directly speaks to a motivation for imitation that is not exclusively explained by social or linguistic factors.

This challenges the view that the primary end of infants' music production is the interaction and not the music itself.

James heard only 14 minutes of singing on that day, approximately 10 minutes of which was infant-directed. That James imitated music despite the low frequency of musical input is impressive. He imitated simple melodic turns from the siblings' models after hearing them only once. He reproduced the distinct melodic patterns of a song that he heard his mother sing hours earlier. Even though exposure to these sung music models was scarce on this day, James was so receptive that he remembered and reproduced them. In comparison, on that same day James was exposed to more than 20,000 adult words produced near him. Imitating music in the absence of recurrent and extensive exposure suggests that young infants are capable of perceiving, representing, recalling, and reproducing the music in which they are immersed.

Infants could be motivated to produce music sounds for a number of reasons, and it is very plausible that one such motivation is that the act is intrinsically appealing to infants. We know that (a) infants are exposed to many different types of music (e.g., Costa-Giomi & Sun, 2016; Custodero & Johnson-Green, 2003), and they are attentive to it (e.g., Costa-Giomi 2014; Merkow & Costa-Giomi, 2014); (b) infants become increasingly better listeners as they develop refined music perception skills such as relative pitch and perception of temporal patterns (e.g., Trehub, 2001); and (c) all humans eventually come to produce the music of their culture (e.g., Dalla Bella, Giguère, & Peretz, 2007). Combined, these considerations suggest that music-loving infants hear a lot of music, develop and learn skills for listening to music gradually, and eventually

grow up to be children and adults that are proficient in the music language of their culture.

Limitations of the Study

This study was designed as an in-depth exploration of infant vocal imitation of music, a phenomenon that has not been systematically analyzed in previous literature. As a case study, the intended outcome was to thoroughly describe this complex behavior in detail. Although the results do not support statistical generalization to a broader population, they do inform and complement previous theoretical frameworks of infant music production and development (e.g., Davidson, 1985; Elmer, 2011; Papoušek & Papoušek, 1981; Welch, 2002, 2016). Specifically, the results suggest that infants may be more active music-makers than previously acknowledged, that they represent and reproduce features of melodies that are explicitly musical (e.g., relative pitch and music rhythmic structures), and that perhaps they are motivated to produce such music vocalizations for purposes that cannot be solely attributed to communicative functions.

Nonetheless, the results and conclusions of the study should be viewed in light that this investigation is based on only one infant and of only one day in his life. The findings are not generalizable to all infants and all families, or even to James during other days. James's parents provided a nurturing environment for musical engagement, language development, and social interactions. All children in the family, including James, had music lessons, and both parents had previous music experience. His parents demonstrated affection constantly, and many times they did this with music (Costa-Giomi & Benetti, in press). Some of the most meaningful moments in all 16 hours of recording

occurred when the parents were singing to James. Compared to other interactions, these moments were usually very short – sometimes lasting only seconds –, which attests to the power of music and to the importance of quality, and not quantity, of musical interactions with infants. Families are very different for many reasons, and they engage with music also in a variety of ways (Costa-Giomi, 2016). It is important to note, however, that (a) the pattern of musical activities in this family is comparable to that found in other families that share similar demographic characteristics (Costa-Giomi, 2016), and (b) the 16-hour timeframe in which the infant was observed proved sufficient to identify instances of the desired behavior – namely, vocal imitation of music.

Another limitation was that the identification of the imitations ultimately rested on my perceptual judgment. Although acoustic analysis and the validation study provided support for my findings, the original identification of what to analyze was based on my repeated listening of James's many vocalizations and the many sounds that he heard throughout the day. It is important to highlight, however, that the methodology employed in this study proved to be crucial for identifying the music imitations. The identification of music imitations was only possible because the data collection and analysis methods allowed for in-depth and exhaustive listening of the continuous audio recording of an entire day of James's life. As an example, I would not have identified James's imitations of the intervals produced by his siblings without repeatedly listening to the file. These vocalizations lasted less than three seconds and the models produced by the siblings in the background were muddled by a multitude of other sounds. Through repeated listening I developed a sensitivity to the nuances of James's vocalizations that allowed me to

identify even short imitations. Furthermore, if I had not listened to all 16 hours, I would not have identified certain vocalizations as instances of imitation; for example, without listening to the entire file, I would not have known that 10 hours before vocalizing “Happy Birthday” James had played with a toy that produced the melody that served as his model.

Extensive recording time was also crucial, because it allowed for capturing all the infant’s vocalizations, including those that occurred in the absence of others. Without access to the sounds he made in private, when he was alone, I would not have identified his imitations of the siblings’ music intervals. It is also interesting to note that some imitations went unnoticed by family members even in cases in which an adult was near the infant, such as the Happy Birthday imitation. Thus, to identify music production by infants, it is not sufficient to rely on parental observation of their children’s vocal behavior.

Implications

James was responsive to the music produced for him and around him, and his imitations reflected an understanding of key features of music. Furthermore, James seemed to be engaged with the music itself independently of the many non-musical functions associated with singing to infants. It is likely that singing to the infant played an important role in fostering such engagement with music by creating a rich music learning environment. In fact, it may be that singing promoted the acquisition of complex music skills such as melodic categorization (Costa-Giomi, 2013).

Musical parenting seems to come naturally to caregivers (Custodero & Johnson-Green, 2003; Ilari, 2005; Malloch, 1999; Papoušek & Papoušek, 1981; Trehub et al., 1997; Young, 2008). Awareness of the importance of singing to infants in terms of musical development may be informative for parents. Rather than worrying about the quantity and type of music to which their child is exposed, parents may wish to simply share meaningful moments of singing with their child – a behavior that is likely to be spontaneous, natural, and familiar to them. It seems that there is no need for prescriptive music interventions but rather a recognition of the value of singing to and with infants. Ultimately, the fundamental goal of developing music skills is not to attain high levels of music expertise. Music is about experiencing meaning in its many forms, and infants' early experience with music may provide an opportunity for developing an affective and intuitive relationship with music that will last a lifetime.

Conclusion

It has long been argued that musical behavior is universal, and that just like language, music is an extremely complex system found only in the human species. Given that music is so pervasive, natural, and specialized, any minimally satisfactory account of human musicality needs to describe how this behavior develops and manifests itself during a person's life. It is important to emphasize that humans are not merely music listeners; humans are music makers. Yet the issue of infant music production has been largely ignored. Accounts that make no attempt to describe infants' music production beyond vague descriptors such as *musical babbling* and that do not take into consideration the possibility of developmental change, lead to a tacit assumption that

individuals magically become competent musicians. But how can this be? Even if musical behavior is stripped of all its cultural, social, and emotional significance, it is still necessary to explain how infants become producers of what is “arguably the most complex auditory stimulus in existence” (Limb & Roy, 2014, p.13) from a signal processing perspective.

Studies of infant music learning (e.g., Addessi, 2016; Reigado, Rocha, & Rodrigues, 2011; Tafuri & Villa, 2002) were some of the first to directly describe the development of music production skills. Such work is laudable and provides starting points for future research. However, many studies of music production in infancy and early childhood share a common problem: Without clear operational definitions and meaningful measurements that reflect such definitions, it becomes very difficult to make appropriate comparisons and to collect data that are generalizable. The findings of this study provide initial clues regarding infant music production and imitation. By beginning to understand how infants manipulate their vocal production to express music and in what contexts such music vocalizations occur, we can start to raise hypotheses that can be tested with empirical designs.

The infant in this study was responsive to the music he heard and was motivated to imitate the music around him. The fact that infants actively engage with music, and that they are inclined to imitate it, perhaps even when they are alone, further suggests that from a very early age humans desire and spontaneously pursue meaningful music experiences.

References

- Addessi, A. (2016, July). *Vocal development in infancy: The role of reflexivity during interactions with parents, grandmother, and educator during routine care*. Paper presented at the International Conference on Music Perception and Cognition, San Francisco, CA.
- Beckman, M. E., & Venditti, J. J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 603-652). Malden, MA: Blackwell Publishing.
- Berkowska, M., & Dalla Bella, S. (2009). Acquired and congenital disorders of sung performance: A review. *Advances in Cognitive Psychology*, 5.
- Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences*, 17(1193), 97–110.
- Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer (Version 6.0.17). Retrieved from <http://www.praat.org/>
- Boysson-Bardies, B. de, Halle, P., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16(1), 1–17.
- Boysson-Bardies, B. de, Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, 11(1), 1–15.
- Chang, H.-W., & Trehub, S. E. (1977a). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, 24(2), 324–331.
- Chang, H.-W., & Trehub, S. E. (1977b). Infants' perception of temporal grouping in auditory patterns. *Child Development*, 48(4), 1666–1670.
- Corbeil, M., Trehub, S. E., & Peretz, I. (2016). Singing delays the onset of infant distress. *Infancy*, 21(3), 373–391.

- Costa-Giomi, E. (2013). Infants' perception of timbre in music. In J.-L. Leroy (Ed.), *Actualités des universaux en musique / Topics in musical universals* (pp. 183–197). Paris: Editions des Archives Contemporaines.
- Costa-Giomi, E. (2014). Mode of presentation affects infants' preferential attention to singing and speech. *Music Perception: An Interdisciplinary Journal*, 32(2), 160–169.
- Costa-Giomi, E. (2016, July). *Infants' home soundscapes: A portrait of American Families*. Paper presented at the International Conference on Music Perception and Cognition, San Francisco, CA.
- Costa-Giomi, E., & Benetti, L. (in press). Through a baby's ears: Musical interactions in a family community. *International Journal of Community Music*.
- Costa-Giomi, E., & Ilari, B. (2014). Infants' preferential attention to sung and spoken stimuli. *Journal of Research in Music Education*, 62(2), 188–194.
- Costa-Giomi, E., & Sun, X. (2016). Infants' home soundscape: A day in the life of a family. In J. A. Bugos (Ed.), *Contemporary research in music learning across the lifespan: Music education and human development* (pp. 87–96). New York, NY: Routledge.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–153.
- Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 366(1567), 1149–1157.
- Custodero, L. A., & Johnson-Green, E. A. (2003). Passing the cultural torch: Musical experience and musical parenting of infants. *Journal of Research in Music Education*, 51(2), 102–114.
- Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, 121(2), 1182–1189.
- Davidson, L. (1985). Tonal structures of children's early songs. *Music Perception: An Interdisciplinary Journal*, 2(3), 361–373.
- Deutsch, D. (2010). Speaking in Tones. *Scientific American Mind*, 21(3), 36–43.
- Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *The Journal of the Acoustical Society of America*, 129(4), 2245–2252.

- Elmer, S. S. (2011). Human singing: Towards a developmental theory. *Psychomusicology: Music, Mind and Brain*, 21(1–2), 13–30.
- Esteve-Gibert, N., & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants. *Journal of Child Language*, 40(5), 919–944.
- Falk, S., Rathcke, T., & Dalla Bella, S. (2014). When speech sounds like music. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1491.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60(6), 1497–1510.
- Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 523–602). Malden, MA: Blackwell Publishing.
- Gelfand, S. A. (2009). *Hearing: An introduction to psychological and physiological acoustics* (5th ed.). Boca Raton, FL: CRC Press.
- Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, 415(6873), 755.
- Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N. J. Enfield & S. C. Levenson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 229–255). Oxford: Berg.
- Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Developmental Science*, 15(3), 398–407.
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19(5), 515–523.
- Gratier, M., & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Developmental Psychology*, 47(1), 67–76.
- Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, 19(4), 385–408.

- He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *Journal of Neuroscience*, 29(24), 7718–8822.
- Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1570–1582.
- Ilari, B. (2005). On musical parenting of young children: musical beliefs and behaviors of mothers and infants. *Early Child Development and Care*, 175(7–8), 647–660.
- Joliveau, E., Smith, J., & Wolfe, J. (2004). Acoustics: Tuning of vocal tract resonance by sopranos. *Nature*, 427, 116.
- Jones, E. J. H., & Herbert, J. S. (2006). Exploring imitation in infancy: Deferred imitation and the development of declarative memory. *Infant and Child Development*, 15(2), 195–205.
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, 19(3), 627–640.
- Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, 43(3), 253–291.
- Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America*, 72(2), 353–365.
- Kessen, W., Levine, J., & Wendrich, K. A. (1979). The imitation of pitch in infants. *Infant Behavior and Development*, 2, 93–99.
- Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology*, 116(2), 471–486.
- Klein, P. J., & Meltzoff, A. N. (1999). Long-term memory, forgetting, and deferred imitation in 12-month-old infants. *Developmental Science*, 2(1), 102–113.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1493), 979–1000.

- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218(4577), 1138–1141.
- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, 100(4), 2425–2438.
- Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100(15), 9096–9101.
- Limb, C. J., & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing Research*, 308, 13–26.
- Longhi, E. (2009). ‘Songese’: Maternal structuring of musical interaction with infants. *Psychology of Music*, 37(2), 195–213.
- Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*, 52(6), 599–608.
- Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science*, 1(4), 272–276.
- Lynch, M. P., Oller, D. K., Steffens, M. L., & Buder, E. H. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychobiology*, 28(1), 3–25.
- Malloch, S. N. (1999). Mothers and infants and communicative musicality [Special issue]. *Musicae Scientiae*, 3, 29–57.
- Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, 24(4), 470.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), 74–78.
- Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, 54(3), 702–709.
- Merkow, C. H., & Costa-Giomi, E. (2014). Infants’ attention to synthesized baby music and original acoustic music. *Early Child Development and Care*, 184(1), 73–83.
- Moore, B. C. J. (2008). Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society B* 363, 947–963.

- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, 27(4), 455–464.
- Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics & Phonetics*, 20(5), 351–369.
- Oller, D. K. (2000). *The emergence of the speech capacity* (1st ed.). Mahwah, N.J: Psychology Press.
- Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences of the United States of America*, 110(16), 6318–6323.
- Papaeliou, C. F., & Trevarthen, C. (2006). Prelinguistic pitch patterns expressing “communication” and “apprehension.” *Journal of Child Language*, 33(1), 163.
- Papoušek, M., & Papoušek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language*, 9(6), 137-157.
- Papoušek, M., & Papoušek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. *Advances in Infancy Research*, 1, 163–224.
- Pfordresher, P. Q., Demorest, S. M., Dalla Bella, S., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Theoretical perspectives on singing accuracy: An introduction to the special issue on singing accuracy (Part 1). *Music Perception*, 32(3), 227-231.
- Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, 98(1), 1–11.
- Ray, E., & Heyes, C. (2011). Imitation in infancy: The wealth of the stimulus. *Developmental Science*, 14(1), 92–105.
- Reigado, J., Rocha, A., & Rodrigues, H. (2011). Vocalizations of infants (9-11 months old) in response to musical and linguistic stimuli. *International Journal of Music Education*, 29(3), 241–255.
- Rutkowski, J. (1990). The measurement and evaluation of children's singing development. *The Quarterly*, 1(1-2), 81-95. Retrieved from <http://www-usr.rider.edu/~vrme/v16n1/visions/spring10>

- Rutkowski, J. (1997). The nature of children's singing voices: Characteristics and assessment. In B. A. Roberts (Ed.), *Proceedings of the Phenomenon of Singing International Symposium I* (pp. 201-209). Retrieved from <http://journals.library.mun.ca/ojs/index.php/singing/article/view/945>
- Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, 6(1), 35-43.
- Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, 37(1), 74-85.
- Saffran, J. R., Reeck, K., Niebuhr, A., & Wilson, D. (2005). Changing the tune: The structure of the input affects infants' use of absolute and relative pitch. *Developmental Science*, 8(1), 1-7.
- Schellenberg, E. G., & Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. *The Journal of the Acoustical Society of America*, 100(5), 3321-3328.
- Schwier, C., Maanen, C. van, Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy*, 10(3), 303-311.
- Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal singing modulates infant arousal. *Psychology of Music*, 31(4), 365-375.
- Siegel, J. A., & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception & Psychophysics*, 21(5), 399-407.
- Snow, D. (2007). Polysyllabic units in the vocalizations of children from 0;6 to 1;11: Intonation-groups, tones and rhythms. *Journal of Child Language*, 34(4), 765-797.
- Snow, D., & Balog, H. L. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua*, 112(12), 1025-1058.
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology*, 46(1), 286-292.
- Southgate, V., Chevallier, C., & Csibra, G. (2009). Sensitivity to communicative relevance tells young children what to imitate. *Developmental Science*, 12(6), 1013-1019.

- Standley, J. M., Cassidy, J., Grant, R., Cevasco, A., Szuch, C., Nguyen, J., ... Adams, K. (2010). The effect of music reinforcement for non-nutritive sucking on nipple feeding of premature infants. *Pediatric Nursing*, 36(3), 138–145.
- Tafuri, J., & Villa, D. (2002). Musical elements in the vocalisations of infants aged 2–8 months. *British Journal of Music Education*, 19(1), 73–88.
- Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, 113(2), 345–361.
- Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development*, 19(1), 83–92.
- Trainor, L. J. (1997). Effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals. *Journal of Experimental Psychology: Human Perception and Performance*, 23(5), 1427–1438.
- Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, 20(3), 383–396.
- Trainor, L. J., & Hannon, E. E. (2013). Musical development. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 423–498). San Diego, CA, US: Academic Press.
- Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, 18(2), 394–402.
- Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences*, 930(1), 1–16.
- Trehub, S. E. (2016). Infant musicality. In I. Cross, S. Hallam, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 387–398). New York, NY: Oxford University Press.
- Trehub, S. E., & Gudmundsdottir, H. R. (2015). Mothers as singing mentors for infants. In G. Welch, D. M. Howard, & J. Nix (Eds.), *The Oxford handbook of singing*. Oxford University Press. Retrieved 15 Mar. 2017, from <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199660773.001.0001/oxfordhb-9780199660773-e-25>.
- Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., & Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology*, 33(3), 500–507.

- Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, 16(3), 285–295.
- Tsang, C. D., Falk, S., & Hessel, A. (2016). Infants prefer infant-directed song over speech. *Child Development*. Advance online publication. doi:10.1111/cdev.12647
- Tsang, C. D., Friendly, R. H., & Trainor, L. J. (2011). Singing development as a sensorimotor interaction problem. *Psychomusicology: Music, Mind and Brain*, 21(1–2), 31.
- Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, 21(1–2), 61–82.
- Vihman, M. M., & Boysson-Bardies, B. de. (1994). The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica*, 51(1–3), 159–169.
- Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A re-assessment of the continuity issue. *Language*, 61(2), 397–445.
- Vurma, A., & Ross, J. (2006). Production and perception of musical intervals. *Music Perception: An Interdisciplinary Journal*, 23(4), 331–344.
- Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, 25(7), 1314–1324.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143–2152.
- Welch, G. F. (1994). The assessment of singing. *Psychology of Music*, 22(1), 3–19.
- Welch, G. F. (2002). Early childhood musical development. In L. Bresler & C. Thompson (Eds.), *The arts in children's lives: Context, culture and curriculum* (pp. 113–128). Dordrecht, NL: Kluwer.
- Welch, G. F. (2016). Singing and vocal development. In G. McPherson (Ed.), *The child as musician: A handbook of musical development* (2nd ed., pp. 441–461). Oxford: Oxford University Press.

- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63.
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7), 2468–2471.
- Young, S. (2008). Lullaby light shows: Everyday musical experience among under-two-year-olds. *International Journal of Music Education*, 26(1), 33–46.

Appendix A: Sequence of Events During the Day

7:20 – 7:50 a.m.: Wake-up.

James wakes up. The mother gets him ready for breakfast (diaper change, bath, clothes) and reads two story books to him.

7:50 – 8:10 a.m.: Breakfast.

James is in the kitchen with the entire family.

8:10 – 10:30 a.m.: Playtime.

James is playing, sometimes with siblings and at other times by himself. The mother joins the children frequently. The father interacts with the children as well, but leaves at 8:40 am.

10:30 – 11:50 a.m.: Playtime and visitors.

The mother's friend and her toddler daughter visit. The children are playing, James mostly by himself, while the mother and friend chat. Later, everyone plays a board game and hide-and-seek. The friend and her daughter leave.

11:50 a.m. – 12:20 p.m.: Lunch.

James is in the kitchen with the mother. The siblings are playing in the living room.

12:20 – 3:25 p.m.: Nap.

Mother takes James to the bedroom and sings lullabies. James vocalizes during sleep – rhythmic repetitions of vowels and repetitions of the word “doggy.”

3:25 – 3:50 p.m.: Wake-up and snack.

James wakes up. The father gives him a snack, and then the mother gives him milk.

3:50 – 4:25 p.m.: Playtime.

The father and children are playing. James plays alone sometimes. The father reads books to the children. The mother is also present sometimes.

4:25 – 4:40 p.m.: Walk outside.

James and the father go outside for a walk.

4:40 – 5:20 p.m.: Playtime.

The children are playing. The father plays a board game with them.

5:20 – 5:40 p.m.: Playtime.

The mother plays with the children.

5:40 – 5:50 p.m.: Getting ready for dinner.

The father and James are in the kitchen preparing dinner.

5:50 – 6:10 p.m.: Dinner.

The entire family is together.

6:10 – 6:50 p.m.: Playtime.

The children are playing. Dale practices cello at 6:20 p.m.; the parents help him practice rhythms. The mother plays with Diane and James.

6:40 – 6:50 p.m.: Father with James.

The father and James are alone. The father sings to him.

6:50 – 7:10 p.m.: Getting ready for sleeping.

The mother and James are alone. She changes his clothes and gets him ready for sleeping. She sings and reads him a story.

7:10 – 11:20 p.m.: Sleeping.

James is sleeping. Sometimes he vocalizes. Recording stops at 11:20 p.m.

Appendix B: Validation Study with Adults

Part 1: Consent, Instructions, and Familiarization Questions

Consent

1. Consent to participate

Hello, my name is Lucia Benetti and I'm a graduate student at the School of Music at the Ohio State University. I am a student of Dr. Eugenia Costa-Giomi, who is the principal investigator for this project. Thank you for your interest in our work. The purpose of this project is to understand how children learn music. You will be asked to listen to and compare short audio excerpts. The duration of this study is approximately 15 minutes. All the data collected from your participation in this experiment is confidential. Participation in this research is voluntary. You can leave the study at any time or refuse to participate without any penalty or loss of benefits to which you are otherwise entitled. There are no anticipated risks or benefits to your participation in this study. Your participation will contribute to our knowledge of how children learn music. By participating in this study, you give consent. Do you agree to participate?

(1) Yes

(2) No

2. Age

Are you between the ages of 18 and 65?

(1) Yes

(2) No

Instructions

3. Instructions

You will hear short audio excerpts, many involving children and babies. You will be asked questions about what you hear. Some of the excerpts contain overlapping voices or background noise. Some of the excerpts are very short. Listen to the excerpts as many times as you want. Adjust the volume if necessary. You will not be allowed to skip questions or go back. This study only lasts about 15 minutes, and you must complete it all in one sitting. Your answers will be very valuable to help us understand how children learn music. Thank you for listening to the audio excerpts as many times as you need to complete the study.

Familiarization: 5 questions

Question 1. Listen to the following audio excerpt of a family playing a board game [audio: James, siblings, and father playing a board game]. How many different voices do you identify?

- a) 3
- b) 4
- c) 5
- d) 6
- e) 7

Question 2. Listen to this short excerpt of an infant vocalizing [audio: “bye daddy,” infant]. What is he saying?

- a) "Bye, daddy"
- b) "Bye, ready"
- c) "I, ready"
- d) "Hi, daddy"
- e) "I, daddy"

Question 3. Listen to this short excerpt of an infant vocalizing [audio: “mommy,” infant]. What is he saying?

- a) "Mommy"
- b) "My knee"
- c) "Ma, me"
- d) "Money"
- e) "Mine, see"

Question 4. Listen to the following audio excerpt of a child speaking [audio: “he has no teeth,” brother]. What is he saying?

- a) "He has no teeth"
- b) "He has no team"
- c) "Yes, no team"
- d) "Yes, no tea"
- e) "He has no tea"

Question 5. Listen to the following songs carefully and as many times as you need to become familiar with them. After you have familiarized yourself with the songs, click next. Now listen to a child singing one of the songs [audio: sister singing “London Bridge”]. Which one is it?

- a) [audio: “Happy Birthday” from toy]
- b) [audio: “Birthday Song 2” from toy]
- c) [audio: “London Bridge” from toy]
- d) [audio: “Bingo” from toy]
- e) [audio: “Ten Little Indians” from toy]

Part 2: Test Questions

Block 1: 7 questions

Question 1. Listen to the following songs carefully and as many times as you need to become familiar with them. After you have familiarized yourself with the songs, click next. Now listen to a child singing one of the songs [audio: infant's Happy Birthday vocalizations]. Which one is it?

- a) [audio: "Happy Birthday" from toy]
- b) [audio: "Birthday Song 2" from toy]
- c) [audio: "London Bridge" from toy]
- d) [audio: "Bingo" from toy]
- e) [audio: "Ten Little Indians" from toy]

Question 2. Listen to a child singing a song [audio: infant's Rain Rain vocalizations]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child singing again [audio: infant's Rain Rain vocalizations]. Which of the following songs is the child singing?

- a) [audio: father singing "Rain Rain Go Away"]
- b) [audio: brother singing "Bingo"]
- c) [audio: mother singing "Bah Bah Black Sheep"]

Question 3. Listen to a child singing a song [audio: infant's Rain Rain vocalizations]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child singing again [audio: infant's Rain Rain vocalizations]. Which of the following songs is the child singing?

- a) [audio: mother singing "Rain Rain Go Away"]
- b) [audio: sister singing "London Bridge"]
- c) [audio: father singing "Abiyoyo"]

Question 4. Listen to this short excerpt of an adult speaking [audio: father saying "blue car"]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the adult speaking again [audio: father saying "blue car"]. Which one of the following excerpts sounds most similar to the adult speaking?

- a) [audio: infant's vocalization before father]
- b) [audio: infant's vocalization after father]
- c) [audio: unrelated infant vocalization 1]

Question 5. Listen to this short excerpt of a child vocalizing [audio: brother singing music interval]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child vocalizing again [audio: brother singing music interval]. Which of the following excerpts sounds most similar to the child's vocalization?

- a) [audio: infant's vocalization after brother]

- b) [audio: infant's repetition of vocalization after brother]
- c) [audio: unrelated infant vocalization 1]

Question 6. Listen to this short excerpt of a child vocalizing [audio: brother singing music interval]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child vocalizing again [audio: brother singing music interval]. Which of the following excerpts sounds most similar to the child's vocalization?

- a) [audio: infant's vocalization after brother]
- b) [audio: unrelated infant vocalization 1]
- c) [audio: unrelated infant vocalization 2]

Question 7. Listen to this short excerpt of a child vocalizing [audio: brother intonation pattern]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child vocalizing again [audio: brother intonation pattern]. Which of the following excerpts sounds most similar to the child's vocalization?

- a) [audio: infant's vocalization after brother]
- b) [audio: unrelated infant vocalization 3 with cello added in the background]
- c) [audio: unrelated infant vocalization 4 with cello added in the background]

Block 2: 2 questions from block 1

Question 1. Listen to the following songs carefully and as many times as you need to become familiar with them. After you have familiarized yourself with the songs, click next. Now listen to a child singing one of the songs [audio: infant's Happy Birthday vocalizations]. Which one is it?

- a) [audio: "Happy Birthday" from toy]
- b) [audio: "Birthday Song 2" from toy]
- c) [audio: "London Bridge" from toy]
- d) [audio: "Bingo" from toy]
- e) [audio: "Ten Little Indians" from toy]

Question 7. Listen to this short excerpt of a child vocalizing [audio: brother intonation pattern]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child vocalizing again [audio: brother intonation pattern]. Which of the following excerpts sounds most similar to the child's vocalization?

- a) [audio: infant's vocalization after brother]
- b) [audio: unrelated infant vocalization 3 with cello added in the background]
- c) [audio: unrelated infant vocalization 4 with cello added in the background]

Block 3: 3 questions

Question 1. Listen to this short excerpt of an adult speaking [audio: father saying "blue car"]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the adult speaking again [audio: father saying "blue car"]. Which one of the following excerpts sounds most similar to the adult speaking?

- a) [audio: infant's vocalization before father]

b) [audio: infant's vocalization after father]

Question 2. Listen to a child singing a song [audio: infant's Rain Rain vocalizations]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child singing again [audio: infant's Rain Rain vocalizations]. Which of the following excerpts sounds most similar to the song that the child sings?

a) [audio: mother singing "Rain Rain Go Away"]

b) [audio: father singing "Rain Rain Go Away"]

Question 3. Listen to this short excerpt of a child vocalizing [audio: brother singing music interval]. Listen as many times as you need to familiarize yourself. Click next when you are ready. Here is the child vocalizing again [audio: brother singing music interval]. Which of the following excerpts sounds most similar to the child's vocalization?

a) [audio: infant's vocalization after brother]

b) [audio: infant's repetition of vocalization after brother]

Part 3: Participant Feedback

Instructions

You will now be asked a few questions about your experience with the survey. Your answers are very important for our research. Please answer honestly. You will not be penalized for any answers you give, and remember that your participation in this study is confidential. Thank you for taking the time to give us feedback.

Question 1. The instructions were clear.

a) Yes

b) Sometimes

c) No

Question 2. [shown if "yes" is not selected] What would help make the instructions clearer? Please select all that apply.

a) Rewording the questions.

b) Giving more precise instructions regarding what to listen for.

c) Organizing the questions and alternatives differently.

d) I don't know.

e) Other (Please state reason below)

Question 3. The audio excerpts were confusing to listen to.

a) Yes

b) Sometimes

c) No

Question 4. [shown if "no" is not selected] Why were some or all of the audio excerpts confusing to listen to? Please select all that apply.

- a) There were too many overlapping voices.
- b) The quality of the recording was not good.
- c) The background noise made it hard to focus on the voices.
- d) The excerpts were too short.
- e) I don't know.
- f) Other (Please state reason below)

Question 5. It was difficult to familiarize myself with the audio excerpts.

Yes (1)

Sometimes (2)

No (3)

Question 6. [shown if “no” is not selected] Why was it difficult to familiarize yourself with the audio excerpts? Please select all that apply.

- a) The excerpts were too short.
- b) I did not listen to them as many times as I needed.
- c) I was not able to familiarize myself, even after listening to the excerpts a reasonable amount of time.
- d) The excerpts were too confusing.
- e) I don't know.
- f) Other (Please state reason below)

Question 7. I answered the questions as best as I could.

a) Yes

b) Sometimes

c) No

Question 8. [shown if “yes” is not selected] Why did you not answer all or some of questions as best as you could? Please select all that apply.

- a) I was very confused.
- b) I did not spend enough time listening.
- c) I did not pay attention to the audio excerpts or to the questions.
- d) The survey took too long.
- e) I don't know.
- f) Other (Please state reason below)

Question 9. Do you have any other comments or suggestions? Any feedback is greatly appreciated.