7-28-2010

Infant Vocal Development in a Dynamic Communication System: Vocal Exploration Activities in Various Social Settings

Kyounghwa Kwon

To the University Council:

The Dissertation Committee for Kyounghwa Kwon certifies that this is the final approved version of the following electronic dissertation: “Infant Vocal Development in a Dynamic Communication System: Vocal Exploration Activities in Various Social Settings”

D. Kimbrough Oller, Ph.D.
Major Professor

We have read this dissertation and recommend its acceptance:

Eugene. H. Buder, Ph.D.

Roger Bakeman, Ph.D.

Rick Dale, Ph.D.

Accepted for the Graduate Council:

Karen D. Weddle-West, Ph.D.
Vice Provost for Graduate Programs

INFANT VOCAL DEVELOPMENT IN A DYNAMIC COMMUNICATION SYSTEM: VOCAL EXPLORATION ACTIVITIES IN VARIOUS SOCIAL SETTINGS

by

Kyounghwa Kwon

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Audiology and Speech Pathology

The University of Memphis

August 2010

ACKNOWLDEGEMENTS

This work has been supported by a grant awarded to Dr. D. Kimbrough Oller, as Principal Investigator, and Dr. Eugene H. Buder, as co-Principal Investigator from the National Institute on Deafness and Other Communication Disorder (R01 DC006099-04) titled, “Infant Vocalization and the Foundations for Speech”.

I would like to express my sincere gratitude to my dissertation advisor, Dr. Kimbrough Oller for his thoughtful guidance, encouragement and mentorship during my entire Ph.D. journey. His encyclopedic knowledge, excitement, and integrity in conducting research have always inspired me to be an independent research scientist in this field. Along with the long journey, he has also showed great endurance and generosity with my decisions and allowed continuing the program. Without his generosity, this work would not have been possible.

I am also indebted to Dr. Eugene Buder for his enormous guidance and encouragement in every step of the whole program. His insightful advice and assistance in developing and analyzing research including this dissertation have been always critical to study’s completion. He also has been a good mediator and representative for me whenever I needed help.

I thank Dr. Roger Bakeman, whose expertise in the area of sequential analysis for social science has been indispensable in helping me completing this study. His concrete knowledge on lag sequential analysis helped me conceptualize and analyze this project with various statistical tools. I appreciate his willingness to serve my committee member and give me advice when I needed.

I thank Dr. Rick Dale, whose expertise in the area of statistical analysis, neural network and dynamical systems models of cognition and learning has helped me to analyze and look at my results in novel ways with new statistical tools. I also thank to him for his willingness to serve my committee member.

I am also grateful for the support and encouragement provided by the faculty, staff and students of the School of Audiology and Speech-Language Pathology. I specially thank Dr. Lesya Chorna, whose technical expertise helped me continue this research in many ways.

I am also grateful to my cohorts, Dr. Heather Ramsdell, Dr. Todd Gibson, Jamie Edrington, Maki Doiuchi, Anne Warlaumont, Jani Johnson, JingJing Xu, Kathy Shaughnessy, Julia Fitzer, and Doriann Berverly who have helped with proofreading and shared their knowledge and friendship over years.

I also want to express my gratitude to Sarah Lurthy, for her help in completing reliability coding for social engagement dimension.

I am also grateful to Mellissa Slade, Amy Myers, Sophia Jung, and John Swearengen for their careful proofreading.

Specific thanks also go to all members of Memphis Audio, especially Tony, Pam and Craig Yarbrough, for their support and prayers for our family and this study.

I cannot adequately express my thanks to the family and friends who have supported me and my immediate family during this process. I especially thank my parents and my parents-in-laws for their prayers.

I am grateful for my children, Anna and Elena for their understanding and sacrifice throughout this entire process. They were my biggest motivators and refreshment for this whole process. I thank for them for being there for me.

My greatest thanks and appreciation go to my husband, Kisang. Without his encouragement, endurance and sacrifice, none of this would have happened. He has always been with me during the peaks and valleys of this long journey and the best friend throughout my whole life. Without his love and support, none of this would have happened.

Finally I thank God for all his plans for me and giving me strength to complete this work.

Kyounghwa Kwon

July, 2010

ABSTRACT

Kyounghwa Kwon, Ph.D. The University of Memphis. August 2010.
Infant Vocal Development in a Dynamic Communication System: Vocal Exploration Activities in Various Social Setting. Major Professor: D. Kimbrough Oller, Ph.D.

Pre-linguistic vocal productions are the manifestations of a unique capacity of human infants. The uniqueness is characterized by voluntary vocal play and systematic repetition which serve no fixed functions. Such pre-linguistic infant vocalizations are the product of complex processes of internal vocal exploration activities and external social interactions. However, very few studies have incorporated both aspects to understand the nature of early vocal development. Therefore, the current study aims to investigate the relationship between vocal category development and exploration activities in varying social communication circumstances, using a developmentally appropriate coding scheme for infant vocalization.

The vocalizations of 7 infants at three different ages (approximately 4, 7, and 11 months) were used. The three vocal types (i.e., vocant, growl, and squeal), the three different engagement types (i.e., symmetrical, asymmetrical, and unengaged), and two proximity types (i.e., immediate and distant) were included. Based on observation and reports from prior research, the current dissertation incorporates fundamental frequency (f_0) and duration of utterances as parameters to measure vocal exploration activity in various social settings.

The findings in this study provides that infants build their vocal categories, producing the three vocal categories distinctively in the acoustic domains of f_0 and duration and showing systematic vocal repetition patterns among the categories. These results provide empirical evidence against Jakobson's view that pre-linguistic vocal

sounds are mere byproducts of biological functions, and presumably not affected by socialization.

The current results seem to suggest that internally driven vocal activities might tend to be more active when social demands are low, indicating that vocal play is a cognitively intensive process requiring attentional resources.

In terms of age effect on vocal exploration activities, at the earliest stage we see high exploration (and thus repetition), and at the latest stage we see consolidated categories (and thus high discriminability). This result provides a significant new perspective on the sequential logic of the two domains in vocal category development.

This research also showed infants' systematic tendency to utilize the vocal categories differentially in various social settings early on, suggesting infants' voluntary control efforts on their pre-linguistic vocal productions to respond to constantly changing social circumstances.

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

Introduction

Within the first year of life, rapid change in the domain of vocalization occurs. Infants begin by producing vegetative vocalizations and primitive phonations, and progress to producing meaningful words with mature syllabic forms. These changes in vocalizations may in part be due to the infant's physiological development. However, more significant non-physiological factors such as infants' voluntary efforts to control their vocalizations in the first year of life have been documented in longitudinal research (Oller 1980; Stark 1981). Playful vocal repetition which seems to function as controlled active practice for vocal repertoire and as a means to express their physical or emotional states has been reported uniquely in human infants (Oller, 2000).

As infants vocalize in repetitive manners, it has been noted by caregivers as well as researchers that infants establish vocal categories as early as the first months of life (Oller, 1980). As expected, such vocal category distinctions are fuzzier in early stages. As infants practice such categories through repetition, the category distinctions seem to become more contrastive. This is evident in various acoustic dimensions. Thus caregivers become more aware of such categories and they start to treat such categories as having specific functions or usages. The prominent vocal categories that are commonly mentioned by parents are vowel-like sounds (the term *vocant* is used to indicate such vocal category in the current dissertation), squeals, growls and raspberries. These early vocal categories are called *protophones* in the infraphonological view developed by Oller (2000).

Unlike fixed calls in other primates, such vocal categories freely associate with multiple functions, showing flexible usage. Flexible usage is a significant early milestone that distinguishes human speech from other species' fixed calls because no other primates show such capabilities in infancy or at any age (Oller, 2000). Parents commonly state that their infants express refusal or discomfort by producing squeal or squeak-like sounds but, infants also express exuberance with squeal sounds. These sounds serve as good examples of flexible functions of squeal sounds in early infancy. In contrast, sounds from primates are always associated to a class of functions such as alarm, mating or food announcement (Hauser, 1996; Jürgens, 1995). Young primates produce vocal calls to approximate mature functions of such fixed call types. Therefore, the presence of such early vocal categories with flexible uses in the first year of life indicates the uniqueness of human infants' vocalizations and suggests such categories are landmarks for speech-language development in infancy.

Although these early vocal categories have been recognized by caregivers and may have communicative significance, little systematic research has been conducted on them. Several researchers reported on vocal play in the literature, however, most of these reports contain merely subjective descriptions without any quantitative measures on concepts of active exploration as manifested in systematic repetition. Therefore, systematic research on the effects of active vocal exploration on early vocal category development is necessary to understand the emergence of speech and language capacity in human beings.

Unfortunately, the hypothesis that active and systematic vocal exploration affects the development of early vocal categories has not been testable or has been hard to prove for several reasons. First, due to the nature of the concept, relevant research requires methodological and technical complexities to properly operationalize both the notion of vocal exploration and systematic repetition. The concepts require developmentally appropriate vocal categories and advanced statistical tools to measure repetitive vocal activities.

Second, long-standing false beliefs derived from a Jakobsonian view on pre-linguistic vocalizations have discouraged research on relevant early vocal sounds (Jakobson, 1941/1968). Jakobson viewed infant vocalization as the mere byproduct of biological functions. Therefore, no communicative functions were attributed to pre-linguistic vocalizations.

The third reason that has misdirected or deemphasized the voluntary aspect of active exploration in infant vocalization is the popular Jakobsonian and Chomskyan idea of innate ability for leaning language in human infants. Such an idea has overlooked the possible roles of internal motivation as well as external learning. It is evident that humans have been endowed with more communicative capability than other species. However development of speech and language is a product of complex dynamical systems beyond endowment. Gibson (1988) suggested that such innate ability could not be realized without internally motivated exploratory behaviors or external learning.

The last reason stated above introduces another significant domain to consider in conducting research on infant vocalization. If vocal development is the product of the dynamic process of internal motivation and external learning in addition to endowed ability, proper research should examine both internal motivation and external learning to understand the very nature of vocal category development.

Research on external learning effects (i.e., social interaction effects) in infant vocalizations has been conducted in psychology and infant development for a long time. Quite a few studies reported that parental interaction enhances infant mental, linguistic and emotional growth (Butterworth, Hine, & Brady, 1977; Goldstein, King, & West, 2003; Goldstein & Schwade, 2008; Guitar & Marchinkoski, 2001; Lindblom, 1990; Papoušek & Papoušek, 1987,1989). Some research regarding social interaction effects on infant vocalizations has reported specific changes in perceptually judged vocal domains. Among the relevant studies, vocal quality and pitch change in dyadic communication has been well documented (Bloom,1975,1988; Bloom, Russell, & Wassenberg, 1987; Papoušek & Papoušek, 1979, 1989; Papoušek & Hwang, 1991).

Unfortunately, the validity of most relevant studies on vocal development considering social interaction effects is seriously weakened by methodological flaws. As will be detailed below, the majority of studies in this line of research have involved inappropriate vocal coding schemes or have shoe-horned immature pre-canonical sounds into mature phonological sound units. Additionally, this line of research has deemphasized the endogenous motivations of infants' own active exploration of their vocal capability that have been noted by longitudinal researchers. By simply ignoring such internal aspects, prior research unintentionally depicted infant vocalization as an exclusive result of social learning, ignoring the infant's active role in vocal development.

Statement of Problem

Pre-linguistic vocal productions are the manifestations of a unique capacity of human infants. The uniqueness is characterized by voluntary vocal play and systematic repetition which serve no fixed functions. Such pre-linguistic infant vocalizations are the product of complex processes of internal vocal exploration activities and external social interactions. However, very few studies have incorporated both aspects to understand the nature of early vocal development (Goldfield, 2000; Hsu & Fogel, 2001). Therefore, the current study aims to investigate the relationship between vocal category development and exploration activities in varying social communication circumstances, using a developmentally appropriate coding scheme for infant vocalization and state-of-art acoustic measurement technologies and analysis tools.

More specifically, if communication is a dynamic process characterized by timely coordination and mutual regulation by both infants and caregivers, we expect that vocal category development would be maximized in the context of close mutual interaction (Hsu & Fogel, 2001). This means that their exploration activities should be most active under close mutual interaction circumstances, showing voluntary involvement of both partners. This idea would contradict the claim of Jakobson that infant vocalization is a merely passive or unorganized byproduct of biological function.

Based on observation and reports from prior research, the current dissertation incorporates fundamental frequency (f_0) and duration of utterances as parameters to measure vocal exploration activity in various engagement settings (Bauer & Kent 1987; Bauer, 1988; Bloom, 1989; Henderson, 1998; Hsu, Fogel, & Cooper, 2000; Kwon, Buder, Oller, & Chorna, 2006; Kwon et al., 2008; Oller & Smith, 1977). The results of the study will provide new perspectives on the dynamics of the unique nature of human speech in early infancy. In the long run, the results of the current study can be compared to the results from at risk infants, thus contributing to identification of early red flags in speech and language development.

CHAPTER 2

Literature Review

Vocal Milestones in the First Year

Based on longitudinal observation, researchers in child phonology have suggested that pre-linguistic infant vocalizations proceed through several stages in the first year of life (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1981; Roug, Landberg, & Lundberg, 1989; Stark, 1980,1981; Zlatin, 1975). Although the descriptive terms used in the studies are diverse, the researchers have concluded similarly that there are four or five pre-linguistic vocal stages in the first year and that the core features of vocal development in each stage are undisputable. According to Oller (2000), infants' vocalizations exhibit substantial change while progressing through the following four stages: the phonation stage, the primitive articulation stage, the expansion stage and the canonical stage.

The Phonation Stage (0-2 months)

It is known that infants produce reflexive, vegetative sounds including cry, burp and hiccup in most of the first two months of life. Such sounds are not considered as similar to speech in terms of their sound quality and functions. In the first two months of life, infants also produce brief voluntary vocalizations with smooth voicing, or phonation. Such types of phonation are called quasivowels. The name quasivowel is given due to the fact that the vocal tract is in a relaxed breathing posture or at rest (Oller, 2000). At this stage, infants exhibit the capability of normal phonation without clear involvement of jaw, tongue or lip posturing.

The Primitive Articulation Stage (1-4 months)

Infants produce more clear and smooth voicing with roughly articulated margins produced by the back of the tongue being in contact with the back of the throat. The resulting sounds are called "cooing" or "gooing". Abundant studies have reported the vocalization characteristics in caregiver interaction at this age and have described vocalizations as being made through contact of tongue with the posterior vocal tract area (Anderson, Bard, Sotillo, Newlands, & Doherty-Sneddon, 1997; Anderson, Vietze, & Dokecki, 1977; Crnic, Ragozin, Greenberg, Robinson, & Basham, 1983; Masataka, 1993). However, the involvement of the anterior part of the vocal tract such as the lips and jaws seems very minimal.

The Expansion Stage (3-8 months)

As vegetative or distress vocalizations decrease after the first two months of life, voluntary non-distress vocalizations become more frequent in infant's vocal productions. Such voluntary vocal productions become prominent as infants go through the expansion stage. Infants display a wide range of new sounds and phonation types, including most prominently vocants, raspberries, squeals and growls. Infants produce vocants with the vocal tract open or lips or tongue positioned in various ways. The articulated vocants are more advanced than quasivowels in earlier stages because they approximate a variety of vowel qualities from mature languages. In addition, sequences of nuclei and margins are more slowly articulated than in mature syllables and this vocalization type is called marginal babbling. At this stage, the transition from an initial labial closure to a vocant is often longer than 200 ms (Papoušek & Papoušek, 1989)

The Canonical Stage (5 -10 months)

Infants produce closure and opening sequences with normal phonation in well-timed and repetitive patterns. The result is speech-like canonical babbling such as [baba] or [dada]. Since infants produce well-timed and repetitive sequences of syllables at this stage, these productions are hardly missed or unrecognized by caregivers. Around this stage, parents begin actively trying to negotiate the communicative functions or meaning of the canonical babbles (Papoušek, 1989).

A significant amount of research on canonical babbles has been done with various populations (for a review, see Oller, 2000). Since the time of the onset for canonical babbling is very robust, delayed onset after 10 months is considered as one of the strong red flags for profound deafness (Eilers & Oller, 1974,1994; Eilers, Neal, & Oller, 1996; Oller & Bull, 1984; Wallace, Menn, & Yoshinaga-Itano, 1998; Yoshinaga-Itano, Sedey, Coulter, & Mehl, 1998).

Although the presence of each vocal stage is often noticed by both parents and researchers and depicts clear progress in the vocal domain in young infants, some caveats about the sequential vocal stage models are as follows; first, although the four different stages are sequentially organized, infants show great individual differences in reaching each stage, as also is seen in crawling or walking in motor development. Secondly, many vocal types seem to be present across stages. For example, squeal sounds are observed most frequently at the expansion stage but are also observed in the primitive articulation and canonical babbling stages. These two factors imply a fuzzy nature of vocal category development stages.

Nevertheless, it is notable that vocal achievements progress logically on the basis of prior capabilities. In the earliest stage, infants show a simple production capability for nuclei that are identifiable and distinguishable from other vocal types in the earliest stage. Such capability will allow producing normal phonation in brief chunks, called nuclei (the phonation stage). Then, an articulation capability allowing production of chains or sequences of nuclei with margins is developed. The infant will produce phonation with modulated amplitudes and with very slowly articulated margins (the primitive articulation stage). The next step is to produce such phonation and articulation showing contrasts among nuclei and chaining contrastive nuclei through articulated margins (i.e., creating separate syllables out of articulated chains). However, sequences of nuclei and margins still might be slowly articulated at this point of development (the expansion stage). Therefore the next logical step requires faster transitions of sequences of nuclei and margins so that information can be conveyed within the optimized rhythmic capability of the human vocal tract (Kent, Mitchell, & Sancier, 1991). Such ability will allow the production of canonical syllables (the canonical stage).

While canonical babble sounds have been extensively studied with various populations, pre-linguistic vocalizations prior to canonical babbles have not been studied or considered to be equally significant vocal productions. One of the reasons is due to Jakobson's (1941) interpretation of early pre-linguistic vocalizations. In Jakobson's pre-linguistic vocalizations had been considered not connected or related to later speech. Jakobson deemphasized the pre-linguistic vocalizations as biological or physiological byproducts which do not serve any communicative functions. This theory of 'discontinuity' as well as other related claims about pre-linguistic vocalizations is discussed in the following section because the claims provided the motivation to investigate the notion of systematic and active exploration.

Jakobsonian View

Jakobson's (1941) monograph may be one of the most influential works on phonological development rooted in the framework of structural linguistics. Jakobson has been credited for offering universal principles in the field of child phonology. One of the most notable universal principles is that of "irreversible solidarity," meaning that acquisition follows a regular and invariant succession of stages of development. Irreversible solidarity works on speech sounds in the following ways: the universally occurring speech sounds of natural languages are favored and linguistically rare sounds are not favored by children. The related notion of "phonological universals," emphasizing phonological laws, is still empirically viable and has been applied to diagnostic considerations in the field of speech-language pathology (Dunn & Davis, 2008; Mohanan, 1992).

Although Jakobson (1941) has been credited for offering universal principles in the field of child phonology, some of Jakobson's claims, especially regarding discontinuity and randomness of pre-linguistic sounds, have lead to much subsequent empirical research in attempts to gain additional understanding as to children's phonological development. Jakobson divided phonological development into two separate periods that occur at different developmental stages. The first period is the pre-linguistic babbling period. According to Jakobson, during this period, vocal sounds do not exhibit any particular order of development. The second period is the acquisition of language proper, during which the child follows a relatively universal and invariant order in gaining intentional control over the sounds of the surrounding adult language.

Discontinuity Theory

Jakobson (1941) claimed that there is a discontinuity between the pre-linguistic and linguistic vocal development, meaning that in contrast with linguistic development, pre-linguistic infants produce all conceivable sounds, including rare and advanced sounds such as sibilants. He claimed that the sounds of pre-linguistic babbles do not have phonetic relation with later speech given that babbling is varied and consists of a random sample of sounds. Infants were believed to produce all sorts of sounds which were irrelevant to phonological universals, making it impossible to find any relation to later speech. It was also thought that the child loses nearly all of his ability to produce sounds from the pre-language stage to the first acquisition of words, showing a silent period that marks the border between babbling and language (p. 21). Grégoire (1948) also described pre-linguistic sounds as unformed and unorganized sounds, supporting the idea of discontinuity.

However, many researchers in the 1970s started to report phonetic similarity of babbling speech and later speech (Cruttenden, 1970; Oller, Wieman, Doyle, & Ross, 1975; Stoel-Gammon & Cooper, 1984). Stoel-Gammon and Cooper (1984) examined children’s inventories and frequency of use of consonants and phonotactic shapes in babbling and later meaningful speech. The results indicated that more frontal consonants than back consonants as well as more CV than CVC shapes were produced in all the children. It was also reported that the sounds used in pre-linguistic vocalizations served different functions (words, protowords, etc), both in different social settings and in different ambient settings (for a review, see Vihman, 1996). Studies by Oller, Eilers, Urbano, and Cobo-Lewis (1997) and Stoel-Gammon (1989) demonstrated that infants who start producing well-formed syllabic babbling relatively late tend to start talking late, in the second year. Later empirical research seems to agree that babbling is a precursor to later meaningful speech.

Passive and Random Production of Pre-linguistic Sounds

Jakobson described infants' babbling sounds as "purposeless egocentric soliloquy" and a "biologically-oriented period of tongue-delirium" (Jakobson, 1968, p. 24). He also used the phrase "ephemeral sound productions" for babbling sounds (p. 28). Such descriptions for the pre-linguistic sounds imply passive, random and non-communicative use of pre-linguistic sounds. Given the logic of such descriptions, pre-linguistic sounds appeared to have no significant communicative functions.

In contrast, researchers such as Stark (1980) and Oller (1980, 2000) have reported that infants seemed to actively control their vocalizations using systematic repetition. However, compared with the discontinuity claim, empirical studies regarding active exploration of pre-linguistic vocalizations have been extremely sparse. Testing the phonetic continuity or similarity hypothesis requires only simple methodological designs to count relevant phonological units from the vocal productions. In contrast, testing the hypothesis of active vocalization requires more complex methods to investigate the internal motivation for vocalization. In the next section, prior research on exploratory or voluntary vocal play is reviewed.

Active Exploration

Exploratory behaviors of infant development have been documented for a long time in psychology and infant development literature. Based on her prior studies, Gibson (1988) concluded that nature has not immediately endowed the infant with the ability to perceive the world. Instead, she firmly concluded that infants spend the majority of their entire first year finding out about the world around them.

In motor development literature, exploratory behaviors and self-organization processes derived out of such exploratory activities are well documented by Thelen (1995). An example of such behaviors would be arm movements to reach a toy by young infants. The first attempts to reach a toy might look very unorganized and diverse in their arm movement trajectories. As infants repeat their arm movement to reach a toy, the movement trajectories become more consistent and systematic, showing a narrower arm movement trajectory space.

Exploration of Phonatory/Acoustic Properties

The same phenomenon of "exploration" appears also to occur in the vocal domain. By controlling or exploring articulatory/acoustic properties, infants appear to construct the foundations on which later phonetic segments and other mature phonological units can be built. At first, infants are believed to produce sounds in disorganized patterns, showing an emergence of primitive vocal categories. During vocal exploration, infants contrast the phonetic qualities of each category and self-organize their vocal categories, so that caregivers come to recognize them. This process implies fuzziness in vocal category distinctions in early stages. Within the first year of life, through active and controlled vocalizations in systematic repetitive manners, infants appear to develop their own infraphonological categories.

De Boysson-Bardies (1999) described infant vocalizations as becoming progressively volitional, with infants seeking to extend their sound repertoire. De Boysson-Bardies stated that “infants seem to be infinitely delighted in playing with their voice” (p. 38). The author also reported that infants play at varying the intonation, succession and duration of these sounds. The author conjectured that the reason for repeating certain familiar types of vocalizations was that infants attempt to familiarize themselves with routines and become more capable of producing varied sound effects.

Likewise, Oller (2000) described infants as exploring various possibilities of vocal tract posturing to create differing vowel-like qualities as well as variations in amplitude, duration, and vocal quality. Koopmans van Beinum and van der Stelt (1986) illustrated vocalizations at this stage as productions of all kinds of variations in the phonatory domain especially concerning intonation, duration, and intensity. The ability to control such vocal productions was suggested as critical to later speech development (Roug et al., 1989). The authors presumed that sufficient exploration of vocal capacities would allow the child to modify and refine vocal capabilities to match to the ambient language. Thus the term exploration in the current study specifically refers to exploration of phonatory/acoustic properties in the vocal domain.

Systematic Vocal Repetition

Goldfield (2000) is one of the few studies that has examined acoustic characteristic of exploratory vocalizations at age six months. The underlying idea of the study is quite similar to the current study in the sense that the study examined the exploratory nature of pre-linguistic vocalizations with specific consideration of the repetitive nature of such exploratory vocal activity. Due to the relevance to the current dissertation, detailed discussion on Goldfield is allocated in the subsequent paragraphs.

Goldfield (2000) reported being intrigued by the serial production of vowels with "song-like" quality in pre-linguistic speech sounds (p. 422). The author attributed the emergence of vocal play to changes in the respiratory, laryngeal and supralaryngeal subsystems. By around six months of age, each of these subsystems was considered to change in ways that allowed infants to explore their own productions. The hypothesis of the study was that infants discover the relation between the dynamics of the vocal tract and variations in sound production by repetition of vowels. The specific research questions of the study were as follows; first, what vocal tract properties are being explored? Second, how does the infant modify vocal tract shape to explore the possibilities of sound production? The study specifically examined the resonances of the vocal tract created by tongue gestures that vary in height and anterior-posterior movement measured by the first formant (F1) and the second formant (F2) vowel space. In the study, vocalizations were considered as parts of a "sequence" in cases where utterances were separated by less than 1.0 second. Goldfield compared formant values for vowel-like sounds occurring in sequences and in isolation.

The results of the study revealed that the F1-F2 space for individually produced vowels was broader than that of vowels in sequences, meaning that vowels produced in sequence occupy a more limited region of the vocal tract than those produced singly. The results indicated that higher correlation coefficients for F1 and F2 during exploration of vowels in sequences are evidence of systematic vocal exploration within sequences. This can be interpreted as to minimize the number of degrees of freedom to be controlled. The author interpreted such a higher F1 and F2 correlation values for vowels in sequences as a result of a more restricted search of the vowel space for more selective exploration since increasing perceptual selectivity is considered a hallmark of developmental growth in infancy by Gibson (1988). The author suggested that the repetitive sequence might be an organizational constraint during vocal play as considered by the gestural dynamic model of speech production (Browman & Goldstein, 1992).

The study by Goldfield (2000) is significant in the sense that the exploratory activity in repetitive sequences was measured with objective acoustical variables. Since the primary focus was to examine the role of the tongue, the supralaryngeal structure, exploratory, and performatory actions during vocal play, the use of the F1 and F2 formants was appropriate. Similarly, the current study also aims to examine exploratory activity in repetitive sequences during vocal play. However, there are several differences in details. First, unlike Goldfield (2000), fundamental frequency (f_0), and duration of the utterances are used here as variables. Based on prior research and laboratory observation, these two variables appeared to be the most relevant and prominent acoustic variables that infants play with. Second, while Goldfield used a 1.0 second-rule to define the boundary for groups of vowels in a sequence, the current study incorporated a 7.0 second-rule for such “bouts”. The concept of “bout” is slightly different from “sequence” in Goldfield in the sense that bouts can encompass sequences of both infant and associated utterances of another speaker. Goldfield did not consider interactive contexts with other communication partners in infant vocal production. However current data were obtained when infants interacted with others in various social contexts, which may result in longer gaps between infant utterances due to vocal turn taking activities.

The 7-second rule was based on an intuitive empirical inspection of real interactions. Evaluation of many bouts revealed that gaps more than 7 seconds between vocalizations seemed to correspond to more than one interaction, whereas any gap less than 7 seconds seemed to occur within a single interaction. Thus the term "bout" in the current study corresponds to an intuitive notion of "interaction".

Another difference is that the current work incorporates various engagement settings to investigate exploratory activities. To my knowledge, no systematic research has been conducted regarding exploration activities in various engagement settings in young infancy. In the following section, relevant research on infants' vocal productions in social contexts is discussed.

Vocal Production in Social Context

Changes of vocal quality in a dyadic communication setting have been well documented (Bloom, 1975,1988; Bloom et al., 1987; Papoušek & Papoušek, 1979, 1989; Papoušek & Hwang, 1991). Papoušek and Papoušek (1989) reported vocalization characteristics in mother-infant vocal matching in young infants. The variables of the study were absolute pitch, pitch contour, duration, rhythm, and vowel-like resonance or consonant-like closure. The study reported reciprocal vocal matching of an average of 41% to 57% of infant non-cry vocalization in the six features, illustrating the dynamic nature of the caregiver-infant interaction in early vocal development.

Bloom and colleagues have done a series of studies on the social interaction effects on young infants in the vocal domain (Bloom, 1975, 1988; Bloom et al., 1987).

According to the studies, infants' non-distress vocalization in general increased significantly when the caregiver provided a social response to an infant's vocalization, regardless of whether the response was contingent or not. Contingent responsiveness meant timely appropriate looking, smiling, touching and verbalizing responses by adults. An interesting finding was that the occurrence rates of speech-like ("syllabic") infant vocalizations were higher only when contingent responses were provided by the caregiver.

Unfortunately, the validity of the results appeared to be vulnerable due to methodological flaws, especially the vocal coding scheme. In the series of studies, infant vocalizations were divided into two categories: syllabic and vocalic sounds. The syllabic sounds were considered as speech-like vocalizations. The speech-like sounds were reportedly judged by the criterion of parental descriptions such as "Baby is really talking" (Bloom et al., 1987, p. 215). Specifically, syllabic sounds were defined as "those that had grater oral resonance, pitch variation, and possible consonant-vowel contour" (p. 215). In contrast "the vocalic sounds had greater nasal resonance and were more often produced toward the back of the mouth. Compared with syllabic sounds, vocalic sounds were perceived as more uniform in pitch and more forced" (p. 215)

However, such articulator-based descriptions do not properly illustrate the commonly used notion of syllabic or vocalic sounds. The most prominent characteristics of syllabic sounds may be the timely transition of consonant-like closure to vowel-like nuclei as in a mature adult's speech (Crystal & House, 1990; Gillis, Schauwers, & Govaerts, 2003; Kent, 1991). The description for syllabic sounds in the studies by Bloom and colleagues fails to correctly characterize a critical aspect of the syllable definition in a general usage. In addition, according to their definition, syllabic sounds were associated with greater pitch variation. However, excessive pitch variations are frequently observed with vocalic sounds in young infancy. Moreover, this linguistically confusing vocal communication coding scheme has been adopted in other recent studies such as Hsu and Fogel (2001).

A similar study by Goldstein and Schwade (2008) reported that social feedback to infants' babbling facilitated rapid phonological learning, using developmentally more meaningful coding schemes and well-controlled experimental design. The contingent social feedback meant timely response to each vocalization "by speaking while moving closer to, smiling at, and touching their infants" (p. 516). The vocalizations were coded with four types: a quasi-resonant nucleus, a fully resonant nucleus, a marginal syllable and a canonical syllable. Fully resonant forms and canonical syllables were considered as mature adult-like forms whereas quasi-resonant forms and marginal syllables as immature forms. The authors argued that infants learn new vocal forms by discovering phonological patterns in their mothers' contingent speech and then generalize from these patterns.

The majority of studies attempting to examine social effects on infant vocal development have not considered infants' own exploratory activities. Earlier studies in the stimulus–response paradigm (Ferguson, 1986; Skinner, 1957; Staats, 1972), which associates an activity with specific consequence by external reward and punishment, have highlighted the learning process. Thus, such studies might give the wrong impression that infant vocalizations are the products of social learning and conditioning, not infants' active exploration of vocal range and ability. Such an impression is in contradiction with observations by longitudinal researchers who have emphasized endogenous factors in infants' vocal play.

The assumption that infants build their own categories by exploring their vocal range and ability is similar to a cognitive-constructive theory advocated by Ferguson (1986). Ferguson (1986) suggested that infants exercise their articulatory skills and respiratory control through active play and repeated self-exploration. However, this view deemphasizes social learning and conditioning processes in infant vocalizations through social interaction.

Therefore, it appears that the two models based on the stimulus-response paradigm and the cognitive-constructive theory fail to address the interrelated aspects of endogenous exploration and external social learning in infant vocal development. In my opinion we need a better model which comprehensively depicts the relation between internal and external factors in infant vocal development.

Dynamic Process Model

Hsu and Fogel (2001) approached this issue by suggesting a dynamic process point of view. The authors assumed that caregivers are potential partners in the co-creation of vocalizations. In their model, communication is a dynamic process characterized by contingent coordination and mutual regulation by both infants and caregivers, providing quality of mutual creativity and a co-creation of novelty (p. 88). This model is similar to the 'push-and-pull' effects hypothesis for animal calls by Scherer (Scherer,1985,1992). Scherer argued that qualitative differences in animal calls were affected by both internal physiological push elicited by cognitive processes and an external pull effect by environmental factors.

Similarly, Hsu and Fogel (2001) proposed a model encompassing both internal motivations of the infant and caregiver effects in human vocal development. Hsu and Fogel's hypothesis was that if infants' vocalizations are due to exploration of their vocal repertoire and capability and if caregivers play a role in this process, more advanced vocalizations during symmetrical communication (direct mutual interaction) settings could be expected. They measured the quantity and quality of vocalizations in varying social communication settings: Symmetrical, Asymmetrical, Unilateral, Disruptive, and Unengaged. Table 1 represents the definitions and examples in Hsu and Fogel.

This detailed and intuitive social context coding scheme allowed examination of the vocal productions of infants in terms of their quality and quantities in various engagement settings. According to their results, the rate of vocalizations was positively associated with the symmetrical communication setting but negatively associated with the unilateral communication setting. In addition, syllabic vocalizations were more likely to occur in the symmetrical mother-infant communication setting.

Hsu and Fogel(2001)'s study demonstrated that the qualitative and quantitative characteristics of infant vocalizations are closely coupled with the moment-to-moment dynamics of mother-infant communication. Unfortunately, the results of the study are hard to interpret because the vocal coding scheme of Bloom and colleagues' studies cannot be easily reconciled with standard conceptions of speech science and linguistic phonetics. The terms syllabic and vocalic in very unique ways that do not correspond clearly with concepts of standard description in other fields. Still their results show some consistency and consequently it appears that the fundamental distinction operationalized in these studies was related to more well-formed syllables as opposed to pre-canonical vocalizations. The results appear to suggest that parents responded more to the well-formed syllables. Although the coding definitions were hard to interpret, the use of the intuitive social context coding scheme and the new perspectives from the dynamic process model gave the study significance.

Interpersonal Proximity in Infant Vocal Development

Extremely few studies have been reported to examine the effect of physical distance on the quality and quantity of speech in human infants. The designs of most previous studies, including Hsu and Fogel (2001) and Goldstein and colleagues incorporated only immediate physical context settings (Cohn & Tronick, 1987; Goldstein et al., 2003; Goldstein & Schwade, 2008). For example, the infants were held by the mothers on their laps in the Hsu and Fogel (2001). In the study, only face-to-face interaction circumstances from mother's lap, with no toys, were included. Therefore, conditions of physical proximity did not vary significantly. Other studies also examined the early vocal development in a close face to face interaction. Further the sessions were only five minutes long and parents were instructed to interact during that period.

Anderson, Vietze, and Dokecki (1978) reported that infants vocalize significantly more when they were within arm's reach of their mothers (and thus face to face) than when they were held. However, both of the conditions can be viewed as immediate physical proximity circumstances. Morris (2005) reported vocal quality differentiation in one infant in an interactive social context (face to face) as opposed to a separated social context (infant and parent in different parts of the same room, with parent engaged in conversation with another adult).

Table 1

Definition and Examples of Different Mother-Infant Communication Patterns

Communication Pattern	Definitions and Examples
Symmetrical	This is characterized by the mutual coordination and elaboration of novelty both the partners in an ongoing interaction. Novel actions are created by the continuous coordinated exchange between the two. An example of this pattern of communication is a mother and infant actively engaged in a peek-a-boo game. The mother covers the infant's face and changes the timing of the uncovering during each round of the game. The infant smiles, laughs, and shows excited bodily movements. The actions of mother and infants are continuously coordinated with each other during the game.
Asymmetrical	The two participants are interested in the same activity; however, while one partner is actively elaborating an activity, the other is merely attending without active participation. An example is a mother wiggling her fingers and then beginning to walk her fingers on the infant's tummy. The infant observes the mother's actions but shows no other behavioral indications of participation in the game.
Unilateral	Unilateral communication occurs when one of the two partners is actively trying to engage the second in an activity while the second is engaging in other activities. There is no mutual coordination between the two participants. An example is a baby intensely sucking on her own thumb, disregarding the mother's attempt to get a smile from her by tickling and kissing her feet.
Disruptive	The first partner attempts to become involved in the activity of the second, who is not responsive. This typically occurs when the first partner acts in a way that is inconsistent with the flow of the second partner's activity and the second shows active avoidance or resistance. An example is an infant sucking on his thumb vigorously, and his mother suddenly grabbing his hand and pulling it away from his mouth. The infant then fusses to protest the mother's abrupt movement.
Unengaged	There is an absence of any communicative engagement between mother and infant: the two are not engaged in any mutual activity. For example, while mother is looking at the camera in the room, infant is looking at and chewing on his shirt.

Note. From "Infant vocal development in a dynamic mother-infant communication system" by Hsu, H.C., & Fogel, A. 2001, *Infancy, 2*, 94 Copyright 2001 by the Lawrence Erlbaum Associates, Inc. Reprinted with permission of the authors.

The results of the study showed that the mean duration of vocalizations for the infant was shorter in the interactive social context than in the separated social context. In addition, the infant vocalized more in the interactive social context than in the separated social context. Morris suggested that inclusion of more data at various ages will help to generalize the results.

Summary of Literature Review

Long standing false viewpoints that pre-linguistic infant vocalizations are completely random and unorganized byproducts of biological function have been present since Jakobson (1941/1968). However, recent infant vocalization research has shown that human infants' vocal activities are unique, demonstrating a capability for voluntary vocalization. In fact, voluntary vocalizations are significant early landmarks, since other primates do not show such a capability at any age (Hauser, 1996; Jürgens, 1995). Such active exploration of the vocal capacity by human infants results in their own infraphonological categories within their first year of life.

These infraphonological sounds of young infants are already different in their utilization from other species' signals as they can be produced with multiple functions. Therefore, establishment of their own vocal categories illustrates that infant vocalizations are not mere byproducts of biological function but a manifestation of active exploratory endeavor to produce controlled vocal patterns (Oller, 1978; Stark, 1981; Zlatin, 1975).

Although several researchers observed empirically such vocal play and addressed such phenomena in the literature, those were merely subjective descriptions without any quantitative measures on such concepts as active exploration in systematic repetition, with the exception of the recent work of Goldfield (2000) and Kwon, Oller, and Buder

(2009). In general, research based on cognitive-constructivist assumptions has deemphasized the external role of social learning in caregiver-infant interaction settings.

Another line of research has focused on caregiver-infant interaction effects in infant vocalization development. A series of studies by Bloom and colleagues reported that caregiver's contingent responses affect the quality of infants' vocalizations. Similarly Goldstein and colleagues (2003) have reported that contingent social feedback to infants' pre-linguistic sounds facilitates phonological learning. They argued that infants learn their vocalizations through social learning with statistical learning. The majority of the studies in this line of research adopted inappropriate coding schemes or shoe-horned immature pre-canonical sounds into mature phonological sound units. Additionally, this line of research deemphasized the internal motivations of infants' own active explorations of the vocal range and ability that were noted by longitudinal researchers.

Therefore, research is needed to examine infant vocal development, especially early vocal category development, considering both the infants' own exploration tendency and the external aspect of social interaction, using developmentally appropriate vocal schemes. Study of the interrelated processes of these two factors would enhance the description of the dynamic nature of vocal category development in infancy. In other words, the adaptation of age appropriate vocal coding schemes from longitudinal research on vocal exploration with relevant social coding schemes from social interaction/response research would improve on prior research. In the following section, relevant theoretical rationale and methodological design features for the current dissertation are discussed.

Theoretical Framework of the Current Dissertation

Coding of Vocalizations

In order to characterize young infants' early exploratory vocalizations, the current dissertation investigates pre-canonical and canonical infant syllables within the infraphonological framework developed by Oller (2000). Whereas most vocal development studies usually describe the vocalizations of infants phonetically using IPA transcription, a non-phonetic method to code infant vocalizations is adopted. This non-phonetic method emphasizes the infant's capabilities given the developmental, physiological mechanism at any given vocal/phonological stage. Therefore the framework emphasizes step by step vocal development at each stage forming foundations for the next developmental stages.

As discussed in the vocal milestone section, infants in the expansion stage produce speech-like sounds called protophones exemplified by squeals, growls, marginal babbles, and vocants. Squeals and growls as well as vocants are among the most prominent vocal categories in the expansion stage. These three vocal categories are non-reflexive vocal types and appear to be freely used in varying social interactions to express different emotional states. The term "category" implies that the infant produces these sounds with coherence and repetitiveness so that adult listeners come to recognize them categorically. Vocant sounds refer to quasi-vowels or fully resonated vowels in modal voice. Squeals refer to sounds with a salient high pitch component, clearly above the f0 of the baby's normal modal voice. Growls refer to sounds with a salient low pitch or a high harshness component, also clearly outside the range of the baby's normal voice. The presence of vocant, growl and squeal categories was observed by both laboratory staff and was also commonly noted by parents that we surveyed in the Infant Vocalization (IVOC) Laboratory. Parents reported that they recognized those sounds as part of their child's vocal system. In addition, most European languages have terms that refer specifically to these categories (Oller, 2000). The existence of such terms can be taken as evidence that these categories are recognizable and an infant builds categories by exploring their vocal repertoires.

Therefore, in the current study, infant vocalizations are coded with infraphonologically driven descriptive terms. By adopting the infraphonological framework, developmentally meaningful non-cry and non-vegetative pre-canonical babbling and canonical babbling sounds are used as analysis units. Most of the previous studies solely used utterances that yielded well-formed transcriptions even though they acknowledged that infants' vocalization at the earliest stages do not contain vowels or consonants. In addition, previous studies often selected vocalization data only to include well-formed or phonetically transcribable sounds, which occur only at later stages in the first year. This resulted in exclusion of a significant amount data which might represent developmentally meaningful vocalizations. The three vocal categories of Vocant, Growl, and Squeal are phonatory, therefore they make no reference to supraglottal articulation. Because such phonatory description is applicable to any developmental stage, pre-canonical babbling sounds from infants as early as 3 months old are included in the current dissertation.

Fundamental Frequency (f_0) as an Acoustic Variable

Anatomical Evidence

Independent from studies on infant speech perception and vocal output, there are also physiological and anatomical studies indicating special focus on f_0 in young infants' vocalizations.Tucker and Tucker (1979) reported that laryngeal growth is extremely active during the first 18 months of postnatal development. Interestingly, a study conducted by Kahane and Kahn (1984) measured the weight of the four intrinsic laryngeal muscle groups of 9 young infants (6 to 24 weeks old) and 12 adults from autopsy specimens (the infants' lateral cricoarytenoid (LCA) and thryoarytenoid (TA) muscles were removed as a unit due to the difficulties of determining the boundary that divides the two muscles). The results indicated that infant intrinsic muscles are proportional in weight to values in the adult. The average values were about 12 % of the adult values. An interesting finding was that the weight of the Cricothyroid (CThy) muscles was the largest portion of the values. The CThy muscle is the primary lengthener and thinner of the vocal folds, thus controlling $f_{0.}$ According to the report, the weight of the CThy muscle is 97% of the infant laryngeal muscle group, as contrasted to 59% in adults.

The authors commented that "on the basis of weight alone, the cricothryroid muscle appears to hold a preeminent position in the hierarchy of infant laryngeal muscles" (Kahane & Kahn, 1984, p.132). They also commented that "this developmental status may be predetermined by the need for the tensor of the vocal folds to be of adequate mass to regulate length and tension adjustments during infant reflexive and vegetative vocalization, and also during vocal play" (p. 132). As the authors suggested, the disproportionately large absolute weight value of the designated structure that controls f_0 serves as empirical anatomical evidence supporting the idea that infants actively utilize the f_0 mechanism early on in their production of vocalization, and later in language.

Previous Research on f_0 in Infancy

As reported by longitudinal researchers, variations of intonation and pitch were among the most notable variables at the expansion stage. Descriptive terms such as "song-like" quality as in Goldfield (2000) indicate the appreciation of pitch in vocal play. A study by Keating and Buhr (1978) may be one of the earliest studies that dealt with f_0 in vocalizations of infants and children. The study reported that infants and toddlers' f_0 values ranged from 30 to 2500 Hz, which are well outside the values in previous reports. In addition, the study discussed the distinction among three registers (fry, modal, and falsetto) in infants and children, distinctions that also occur in adult vocal registers and are relevant to categorizations to be employed in the current study.

Several other studies on f_0 shift or intrinsic fundamental frequency have also been reported. Both Bauer (1988) and Whalen, Levitt, Hsiao, and Somoodinsky (1995) reported that infant babbling shows the same intrinsic f_0 phenomenon as adult speech. Studies by Bauer (1988), Whalen et al. (1995), and Goldfield (2000) utilized phonetic vowel categories as data even though they acknowledged the 'shoehorning' of immature infant sounds into the mature vocal coding system. In addition, none of the above studies considered social context effects to assess the possibility that f_0 may vary in different social circumstances.

In summary, significant roles of f_0 in pre-linguistic vocalizations have been observed and reported in the literature. However, objective measurements of f_0 have been infrequent (Bauer, 1988; Goldfield, 2000; Robb & Saxman, 1988; Whalen, Levitt, Hsiao, & Somoodinsky, 1995). The validity of a few relevant studies was vulnerable due to transcription of phonetically immature vowel categories. In addition, none of the studies distinguished social contexts in considering variations of f_0.

Previous Research on Duration of Vocal Sounds in Infancy

Studies have reported on the acoustically oriented objective measurement of the duration of pre-linguistic vocalizations (Bloom, 1988, Bloom, 1989; Hsu et al., 2000; Laufer & Horii, 1977; Masataka & Bloom, 1994; Oller, 1980). Mean durations reported in the prior research ranged from 0.5 to 1.5 seconds for 3 to 5 months of age. According to Laufer and Horii (1977), mean duration of vocalizations increases as infants get older. Similarly, according to Oller (1980), the mean duration of vocalizations in infancy increased from about .7 seconds at 3 months of age to about 1.4 seconds at 5 months. It was suggested that such an increase in the duration may be related to the increase of syllabicity of sounds as well as the increased vocal breath control of the infant. Hsu et al. (2000) reported that syllabic sounds were longer than vocalic sounds, independent of social context.

Social context effects on duration of infant vocalizations have been investigated in several studies (Bloom, 1988, 1989; Bloom et al., 1987; Hsu et al., 2000; Masataka & Bloom, 1994; Oller, 1980). According to Bloom (1989), the mean duration of syllabic sounds was reported to be 1.22 seconds whereas that of vocalic sounds was 0.89 seconds in baseline periods of vocalization before social contingencies were applied. During social input, the duration of syllabic sounds was reported to be 1.34 seconds whereas that of vocalic sounds was .93 seconds. According to the study, the main effect of social context (Turn-Taking versus Random Social Input) was statistically non-significant, meaning that the duration of utterances were not statistically different across the two social contexts. However, it was reported that infants produced significantly longer vocalizations when adults provided nonverbal inputs not verbal inputs.

However, above results should be taken cautiously due to the vocal coding scheme that was used. In Bloom and colleagues'(1987, 1988) works and Hsu et al. (2000), syllabic vocalizations were defined as sounds with greater oral resonance and pitch contour and longer duration whereas vocalic vocalizations were defined as sounds with more nasal resonance and sounds produced toward the back of the mouth. These definitions do not conform to standard phonological conceptions from the literature in infant vocal development or child phonology. Again, due to the coding definitions of pre-linguistic vocalizations, the significance of this research in vocal development is difficult to assess.

Rationale and Research Questions of the Current Study

As discussed so far, it is assumed that infants build their own vocal categories by exploring the vocal space. This is an assumption that is consistent with other motor development theory, and is part of a framework of dynamic systems thinking. It is also assumed that as infants vocalize in a repetitive manner, caregivers become aware of the categories. Prior research on parental reports and both auditory and acoustic coding of infant sounds suggest active vocal exploration and categorical vocal contrasts in infancy. These impressions are supported in two ways: in acoustic contrasts among vocal categories and in the apparent practice of vocal categories by infants as seen in temporal analysis of repetitive sequences. I propose to quantitatively investigate such vocal activities manifested in the two different domains: acoustic discriminability and temporal repetition of vocal categories.

However, infant vocalizations are assumed to be the product of a complex process of interplay between internally-motivated vocal exploration activities and social interactions. The primary goal of the current dissertation is to evaluate vocalization categories in terms of acoustic discriminability and temporal repetition in three conditions of social engagement: one when the infant interacts with the caregiver mutually, one when only one of the potential participants to the interaction is engaged, and one when the infant and caregiver are both disengaged.

If communication is a dynamic process characterized by timely coordination and mutual regulation by both infants and caregivers, infant vocal development may be maximized in the context of mutual interaction. This means that their activities (that we may interpret as exploratory) should be most apparent, showing most distinctive categorization in the acoustic domains, and showing the strongest vocal repetition patterns under mutual interaction circumstances. The extent to which infant vocal categories manifest themselves in acoustic distinctiveness and patterns of repetition is the empirical focus of this effort.

A secondary goal is to address discriminability and temporal repetition at three ages (3 to 4 months, 6 to7 months, and 10 to11 months) with the expectation that both variables will demonstrate increase at later ages. If infants develop their own vocal categories through active exploration and repetitive practice, we predict that infants should show most distinctive categorization in the acoustic domains and stronger repetition patterns at later ages.

In addition to the primary goals, subsidiary goals related to frequency of occurrence of the three vocal categories in various social and physical proximal settings are addressed to investigate possible usage patterns of the vocal categories. The current data collection afforded a less formal examination of vocal development in another domain, namely that of "physical proximity of the child to the caregiver" with two categories: Immediate (the infant within 2 feet of the caregiver) and Distant (the infant more than 2 feet away from the caregiver). The current study will provide an overview of potential effects of proximity. The specific research hypotheses and subsidiary questions of the current study are as follows:

Background Hypotheses Regarding Vocal Category Development in the Acoustic Dimensions of f_0 and Duration

1. Infants will produce the auditorily coded vocal categories distinctively in the acoustic domains of f_0 and duration.

2. Infants will show systematic repetition or alternation patterns of the vocal categories.

Main Hypotheses Regarding Infant Vocal Categories in Various Engagement Settings

3. The vocal category distinctions will be most apparent when infants and caregivers engage in mutual interaction.

4. Vocal repetition patterns will be more apparent as infants get older.

Subsidiary Questions Regarding the Usage (frequency) of Vocal Categories

5. Does the usage of vocal categories differ by engagement settings?
6. Is there any age related change in the usage of vocal categories?
7. Are there any patterns or trends in the usage of vocal categories regarding the two physical proximal settings?
8. Does the duration of utterances differ by vocal type, engagement setting and age?

CHAPTER 3

Methods

This chapter describes the research methods used in the current study. It is divided into nine sections: (a) research participants, (b) recording procedure, (c) utterance selection, (d) coding of vocalizations, (e) coding of social engagement, (f) coding of physical proximity, (g) acoustic analysis, (h) analysis of sequential dependency, and (i) statistical analysis related to vocal exploration activities

Research Participants

The research participants for the current study were recruited through word of mouth and advertisement in local child birth education classes for a longitudinal research project. The vocal recordings are a portion of the data from longitudinal research of the IVOC Laboratory at The University of Memphis. The participants had no history of pre- or peri-natal neurological or physiological problems based on parental reports. All the parents of the participants were reported as middle class or above in their socio-economic status. Seven infants (6 females and 1 male) were included (Table 2).

Recording Procedure

Vocalizations of the infants at three different ages, 3 to 4 (Age 1), 6 to 7 (Age 2), and 10 to 11 (Age 3) months were used for the current study. Each recording was conducted for approximately one and a half hours in a laboratory setting. Each session ran around 20 minutes unless other difficulties happened to shorten sessions, such as infants' distress or technical problems. When infants showed distress or discomfort, the session was terminated and resumed later. Accordingly, cry or fuss vocalizations were limited in the recordings.

Table 2

Information on the Participants

Participant	Sex	Ethnicity
AD	Female	Caucasian
EA	Female	Hispanic
KA	Male	Caucasian
KS	Female	African American
LO	Female	Caucasian
LP	Female	Caucasian
SM	Female	Caucasian

During each visit for vocal recordings at the IVOC laboratory, we attempted to record three different session types: *interactive, separated, and alone* sessions. In the interactive sessions, infants and caregivers played together in a natural setting and engaged in interaction. In the separated sessions, infants were with at least one caregiver and one staff member, who held a conversation with the caregiver. Although they were in the same room, the protocol for the separated sessions was for caregivers and laboratory staff to avoid direct interaction with infants unless infants were extremely fussy. In the alone session, infants played alone in the room. Although the alone session was attempted at each recording day, it was often very short because the infants did not like to be alone and often fussed or cried. Caregivers were allowed to soothe the infants, therefore changing the session type to interactive.

The IVOC laboratory used for these recordings was equipped with four remote controlled cameras, and two different camera views selected from among the four cameras were recorded simultaneously at all times. The laboratory had an observation window between the control room and the nursery. The staff controlled the cameras and audio in the observation room. During the recordings, infants were watched through one-way glass in the observation room and, at the same time, two camera angles were selected from the fours views and individually recorded as video files in MPEG format and in most cases in AVI formats on separate computers. One camera angle was intended to present a close-up of the infant’s face and one a wider view of the interaction.

In order to make reliable acoustic measurement of amplitude possible, a calibration tone was recorded at the beginning of each session at roughly the same distance from the child’s mouth as the child’s microphone, and the observed sound pressure level was logged from a meter (RadioShack model 33-2205). In order to avoid mouth-to-microphone distance variability, the infants were fitted with a custom-built vest that housed a wireless microphone system (Samson Airline UHF ALI). The custom-built vest maintained the mouth-to-microphone distance between approximately 5 to 7 cm: this equipment provided excellent signal-to-noise ratio and allowed us to compare amplitudes of infants’ vocalization against the fixed reference level. The stereo audio signals were generally recorded at 48 kHz, although some early sessions for one infant were recorded at 44.1 kHz. Audio signals from the infant and the mother were recorded separately into two channels.

Utterance Selection

Utterance selection was conducted with recorded files using a recently updated version of Action Analysis Coding and Training (AACT 3.45), a software system designed to facilitate audio-video analysis. The onset and offset boundaries of each utterance were inspected and labeled in TF32, which stands for "Time-frequency 32 bits", a very flexible program for acoustic analysis of speech and other audio-frequency waveforms within AACT 3.45. Identification of utterances followed *breath group* rules developed by Oller and Lynch (1992). According to Oller and Lynch (1992), an utterance boundary should be placed when a ingress within a vocalization is notable (p. 519). However, when an utterance had a sudden phonation stop accompanying a breath halt and resumed a phonation as in a glottal stop, it was considered one utterance, regardless of the length of the stop. If no breath was detected, a breath group was defined as a vocalization separation from other vocalizations by at least 200 ms unless a glottal stop intervened. If the duration of an utterance was less than 50 ms, which is significantly below the average duration of a syllable, the utterance was excluded from analysis. For the current study the term *utterance* and *vocalization* are interchangeable. All vocalizations including reflexive or vegetative vocalizations were identified as potentially analyzable utterances for other research purposes. However, reflexive and vegetative vocalizations were excluded from any analysis in the current study.

Coding of Vocalizations

Coding of vocalization types was also conducted off-line using AACT 3.45. The coders for vocalization types were Ph.D. and master's students and a Ph.D.-level scientist in speech language pathology program. All coders had been trained to use the protophone categorization system of Oller (2000). At the expansion stage, the protophone types include vocants, growls, squeals, raspberries, marginal babbles, yells, etc. For the current study, coders were asked to judge among the most prominent three categories of Vocant, Squeal, and Growl and were limited to listening to each utterance three times only during coding. This limitation was hoped to prevent coders from “overanalyzing” the utterances, and thus was hoped to make the coding resemble parental judgments. No spectral display of utterances was allowed. The coding definitions of each vocalization type that were roughly agreed upon in the IVOC laboratory are presented in Table 3. Coders were encouraged to use their intuitive judgments, but in general they were expected to follow the definitions in Table 3.

As shown in the definitions in Table 3, perceived f_0 (pitch) plays a significant role in classifying the categories. However, pitch was not the only criterion in judging the vocal categories. For example, the growl category could be classified by a salient low pitch or by harshness characteristics. Furthermore, when a very brief period of squeal-like or growl-like components occurred within an utterance, coders were asked to judge the utterance intuitively, and in some instances, they chose to ignore the very brief occurrences of special features, interpreting them as accidental.

Table 3

Vocal Coding Definitions

Vocal Category	Definition
Vocant	The vocant category includes quasi-vowels and full vowels. a. Quasi-vowel: vowel-like sound produced with the vocal tract at rest or closed usually with modal voice. b. Vowel: vowel sound produced with the vocal tract open or postured usually with modal voice.
Sqeual	The squeal must have a salient high pitch component, clearly above the normal pitch range of the infant's modal voice. The utterance does not have to require squeal quality throughout to be coded as a squeal.
Growl	The growl must have a salient low pitch or high harshness component, so low or harsh that it is clearly outside the range of the infant's normal voice. The utterance does not have to have growl quality throughout to be coded as a growl.

Inter-rater Reliability of Vocal Codes

In a prior study, reliability of these vocal codes was assessed with a smaller data set (Kwon et al., 2006). Each analyst was assigned to two sessions that had previously been coded independently by other analysts. The overall Cohen's Kappa between the coders was approximately .7. Fleiss (1981) reported that .40 - .75 represents as fair to good agreement among raters. Because Cohen's Kappa value takes account of imbalances in the numbers of utterances pertaining to each category, it is more conservative than a simple agreement method, which the majority of the current infant vocalization studies adopt (Fleiss, 1981). Given the fact that infant vocalization categories are not preformed but are built during the infant's exploratory activities, some fuzziness of the categories at earlier stages is inevitable. Therefore, it may be unrealistic to expect very high reliability in coding these utterances. Considering these two factors, a Cohen's Kappa value of .7 appears to constitute a satisfactory level of reliability.

Initially each session was assigned to one member of a team of coders. After all sessions had been coded, a senior member of the team reviewed every coded utterance and changed codes in cases that were deemed incorrect. Seven sessions (one from seven infants) were analyzed for changes that were made by the senior coder. A total of 108 changes (out of approximately1300 codes) were made. Code changes from vocants to growls were most frequent (50%) followed by changes from squeals to vocants (26%). The analyses were conducted on these final codes including the corrections.

Coding of Social Engagement

The social engagement coding system was adapted from Hsu and Fogel (2001). As discussed in the prior section, five patterns of communication dynamics were used by Hsu and Fogel: Symmetrical, Asymmetrical, Unilateral, Disruptive, and Unengaged. In the study, the Disruptive and Unilateral categories refer to cases where the first partner attempts to be involved in the activity of the second, who is not responsive. When the first partner uses disrupting behaviors to get the infant's attention and the infant becomes fussy, the context is called Disruptive. If the first partner tries to be involved in the infant's activity but the infant still avoids interacting, the context is called Unilateral. In the current study, these types were combined within the Asymmetrical category. In the current study, the five categories were collapsed into three engagement categories. Due to infrequency of disruptive cases, Disruptive was collapsed into Asymmetrical. Unilateral was also collapsed into Asymmetrical for conceptual simplicity of coding.

Consequentially, the three engagement types of Symmetrical, Asymmetrical, and Unengaged were used. The engagement types were determined by the following factors: 1) the presence of prior relevant actions including vocalizations, physical touches or facial expressions by the other participant (usually a caregiver) within seven seconds before the target (infant) vocalization and 2) the directivity of the target vocalization toward the partner or an activity held in common with the partner. When the antecedent activity and following vocalization happened within seven seconds, the following vocalization was considered a *contingent* or *timely* response. 'Seven seconds' was selected to be consistent with the 'seven second rule' used to define a *bout* in temporal vocal repetition domain analysis (as conducted in our laboratory), which will be discussed later. Additional video examination was conducted to investigate if there was any instance where contingent vocal exchange happened outside the seven second interval. Visual inspection of the data confirmed that such cases were very rare, less than 10 instances out of all the data. Therefore, the use of the seven seconds as a rule appeared to be appropriate.

Within 7 seconds, when two participants were directing actions to one other or the same activity, the vocalization of the infant was coded as Symmetrical. When one of the participants' actions was not directed toward the other participant or the same activity, the vocalization was coded as Asymmetrical. When neither of the participants' actions was directed toward each other or the same activity, the vocalization was coded as Unengaged.

The same criteria were applied when infants produced consecutive sequences of vocalization without interruption by caregivers, the so-called infant's bouts. If the caregiver showed continued facial, physical, or vocal directedness towards the infant during the second and following vocalizations, and the infant still directed her activity toward the caregiver, the utterances were coded as Symmetrical. If either one of the participants did not engage in the activity during the second or subsequent vocalizations, they were coded as Asymmetrical. When neither of the participants was engaged during the relevant vocalizations, the vocalizations were coded as Unengaged. Table 4 represents the detailed definitions and examples of the three engagement settings.

Reliability of Social Engagement Codes

Social engagement coding was done off-line through AACT 3.45, in an engagement field (or dimension) that included the same time designations as each utterance in the vocal type field. The primary author coded the social engagement dimension of each utterance. In order to assess reliability of the engagement codes, a second rater, who had undergone social engagement coding training with the primary rater, independently coded approximately 10% of the entire sample. A Cohen's Kappa coefficient was computed between the two raters. The value of .64 indicated fair to good agreement between the two raters (Fleiss, 1981).

Table 4

The Coding Definitions and Examples of the Three Social Engagement Settings

Engagement Type	Definition
Symmetrical	Two participants are engaged in the same activity. Antecedent communicative actions (facial, vocal, physical actions) must occur within seven seconds of target vocalizations. Both participants show an active contribution in the communication exchange. An example of this type is a situation where the mother vocalizes and the infant also vocalizes in a face to face interaction setting.
Asymmetrical	Two participants are not engaged in the same activity. One of the partners is actively trying to engage the second while the second is engaging in other activities. No mutual coordination is observed. An example of this type is a situation where the mother tries to engage the inattentive infant.
Unengaged	Neither of the participants is engaged in the same activity. The absence of any communicative engagement between the mother and the infant is noted.

Coding of Physical Proximity

The three recording settings (interactive, separated, and alone sessions) yielded varying physical proximity conditions. However, the recording protocols for the interactive and the separated sessions created significant bias with regard to engagement and proximity; in the interaction session, the caregiver was asked to interact with the infant whereas in the separate session, the caregiver engaged with a laboratory staff member, who was entering caregiver responses to questions in a log. In separated sessions, the infant was seated or playing away from the caregiver most of the time, which usually resulted in asymmetrical or unengaged engagement circumstances. Therefore, symmetrical engagement circumstances tended to occur with immediate proximity between participants whereas asymmetrical and unengaged circumstances occurred more often with distance between participants. However, there were many occasions where the mother engaged with the infant when the infant bid for attention in separated sessions. Therefore, independent coding of physical proximity throughout the recording session was completed for each utterance.

For the physical proximity setting, three categories of Immediate, Distant and Alone were used. When the caregiver and infant were in close physical distance (less than two feet), or within arm's reach (Anderson et al., 1978), the setting was defined as Immediate. Cases where infants were more than 2 feet away from the caregiver were coded as Distant. When infants were left alone in the room, the setting was coded as Alone.

Reliability of Physical Proximity

Reliability of physical proximity codes was assessed by comparing the primary rater's codes with an independent rater's codes. The independent coder completed a brief training on judging physical proximity. Approximately 10% of the entire sample was coded by the independent rater. A Cohen's kappa was computed and a value of .98 was obtained, indicating excellent reliability.

Acoustic Analysis

f_0 Analysis

The primary author and one Ph. D. level research associate, both with prior training in acoustic analyses of speech, conducted f_0 analysis. In order to trace f_0 in infant vocalizations, a special version of the TF32 program developed within the AACT environment was used. This version of TF32 was designed specifically to allow for flexible parameter settings for infant vocalization f_0 traces (Milenkovic, 2005).

Settings for the spectrographic display were based on the vibratory regime study by Buder et al. (2008), which analyzed acoustic patterns made by infants, interpreted in terms of vocal tissue movement. A very narrow band spectrographic display was necessary to see harmonic contours, which helped to identify the dominant f_0 values for the regime analysis by Buder et al. (2008). For the same reasons, f_0 contours were produced with a narrow band spectrogram in the current study. The bandwidth was set at 10 Hz, the floor at 95 dB, and the dynamic range at 64 dB in most cases. These spectrographic settings changed slightly depending on the recording quality of sessions.

The harmonic structures and regime codes helped to guide f_0 determination. When there were intervening or ambiguous harmonics, analysts were advised to refer to the regime codes, if available, to decide which harmonics to trace as the f_0 of the utterance. For example, when a portion of an utterance was coded as *subharmonics*, which was primarily defined by the abrupt appearance in the narrowband spectrogram of intervening harmonics (Buder et al., 2008), doubling, or higher integer multiples in relation to the surrounding set, analysts traced the higher harmonics as the f_0 for that segment. However, when a segment of an utterance was coded as pulse, showing a very similar spectral display to subharmonics, analysts traced the lower harmonics as f_0 for the segment. When a regime code was not available, analysts were required to judge and trace among ambiguous harmonics based on spectrographic display and auditory impressions.

Earlier work from our laboratory has reported the inclusion of four f_0 measures (the mean, the minimum, the maximum, and the Standard Deviation) to characterize the three vocal categories. Therefore, once f_0 contours were inspected individually, the four different f_0 measures were extracted.

Figure 1 shows an example of the TF 32 display within AACT 3.45 used to code the relevant dimensions. The two waveforms for each channel of the two microphones are shown at the top. Then, the spectrogram shows a f_0 contour based on the microphone inside a special vest worn by the infant. In this example, the infant's microphone is in channel 2. On the upper left corner of the f_0 contour display panel, a numeric readout of the four different f_0 values for this whole utterance is shown. The maximum and the minimum f_0 value of this utterance are highlighted by arrows. The standard deviation value and the mean value were computed using all values throughout the utterances. The next rows in the display present vocal categories, regimes and engagement codes. In the given example, Squeal is the vocal category type and Symmetrical is the engagement type. Between the two codes of Squeal and Symmetrical, a sequence of vocal regime codes is shown. The bottom row shows the special parameter setting options for infant vocalizations for TF 32 in AACT3.45.

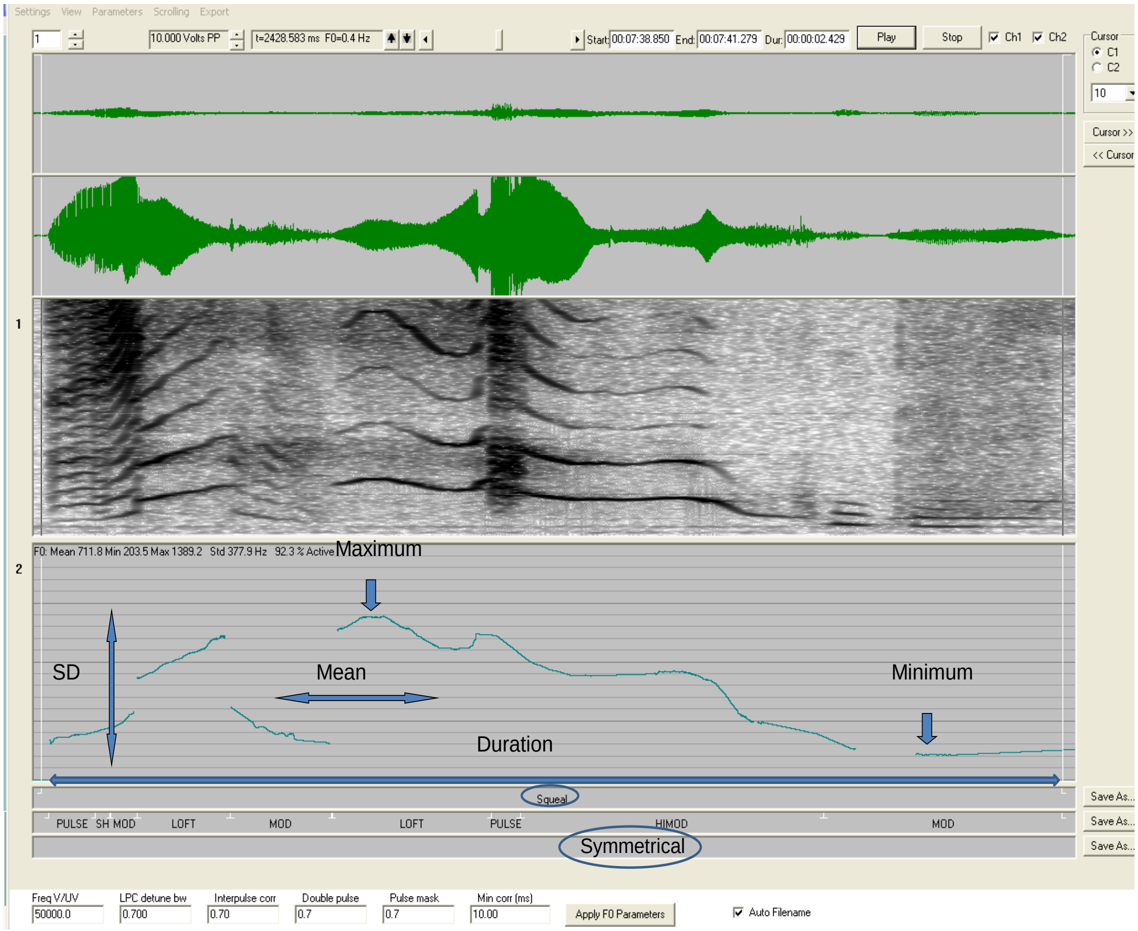

Figure 1. An example of the TF 32 interface within AACT 3.45.

The majority of f_0 values from the data were automatically extracted, especially for vocant sounds. However, when the phonation was not modal, it was often necessary to use hand inspection and marking. The experienced Ph.D. level analyst completed f_0 analysis on approximately 30% of all data and the primary author completed approximately 70%. Approximately 1,300 vocalizations were excluded from f_0 analysis due to overlapping vocal qualities or artifacts.

Inter-rater Reliability of f_0 Analysis

Reliability of f_0 traces was assessed by comparing the two analysts' f_0 measures. Two analysts were assigned to code 40 utterances from a training file which was not included in the current study. Each analyst individually completed tracing the f_0 contours and extracted four f_0 measures from the 40 utterances (a total of 160 f_0 measures from each analyst). Then the 160 f_0 values were compared across analysts. The Pearson correlation of the four f_0 measures between the two analysts was .79. Results indicated best agreement on the mean f_0 values (r = .872, n = 40, p < .001) and lowest agreement score on the minimum f_0 values (r = .637, n = 40, p < .001). The mean difference for the mean and minimum f0 values was 20 Hz (SE =10 Hz) and 51 Hz (SE = 25Hz), respectively. These results appeared to represent difficulty in tracing low f_0 contours where hand corrections were most needed.

Duration Analysis

Utterance boundaries were marked in TF32 within AACT 3.45, by locating cursors at the onset and offset of each vocalization. Then AACT 3.45 automatically yielded beginning and ending time information for each utterance identified by coders.

Inter-rater Reliability of Duration of Vocalizations

Reliability of duration of vocalization was assessed by comparing the two analysts' duration measures. An independent analyst, who had prior training and experience on identifying utterance boundaries, was assigned to identify approximately first 20 utterances (a total of 108 utterances from five sessions) from randomly selected five infants in the current study. The analyst independently completed identifying utterance boundaries and the duration values were compared to the ones that were originally created by other analysts. The Pearson correlation of the duration measures between the coders was .957, indicating outstanding agreement (r =. 957, n = 108, $p < .001$). The mean difference in the duration between the coders was 3.10 ms.

Analysis of Sequential Patterns

The current AACT 3.45 provides a useful export function of vocal coding files to Sequential Data Interchange Standard (SDIS), which is a standard format for sequential data. The Generalized Sequential Querier (GSEQ) program then reads SDIS data input and produces a variety of sequential statistics as output to investigate the tendency for repetition of protophones (Bakeman & Quera, 1995). Each vocalization was tagged with onset and offset time information, which allowed investigation of temporal sequence characteristics. In order to confine the sequential lag analysis to sequentially occurring vocalizations, not isolated ones, gaps longer than seven seconds between utterances were treated as breaking sequences based on prior work (Kwon et al., 2009). For the current study, sequential patterns were analyzed at lag one, which encompass a sequence of two adjacent utterances.

Overview of Statistical Analysis Related to Vocal Exploration Activities

Acoustic Domain: Distinctiveness of the Vocal Categories

A linear discriminant function analysis (DA) was conducted to assess the discriminability (distinctiveness) of the vocal categories. DA is used to determine optimal discrimination functions (linear combinations of variables) to characterize or separate two or more classes, based on a set of predictor variables. In the current study, the discriminating (grouping or dependent) variable was the three vocal categories and the independent variables were the five acoustic measures. The five acoustic measures are treated as one set of variables and the set of dummies generated from the grouping variable as another set of variables; the DA then performs a canonical correlation analysis on these two sets. Thus a canonical correlation is a measure of the degree of relationship between two factors and the square of the canonical correlation is equivalent to the percent of total variance explained by the variables on a particular dimension (function) (Stevens, 2007).

The background question of the extent to which vocal categories could be distinguished by the five acoustic measures (i.e., infants will build their vocal categories, showing the ability to produce the categories distinctively in the acoustic domain) was assessed by a DA analysis across infants, ages, and engagement settings. The primary outcome measures were Wilks' Lambda values to test significance of the discrimination and the square of the canonical correlation (i.e., the percent of total variance explained) to assess the degree of discriminability of vocal categories by the five acoustic measures.

The test of the main hypothesis in the acoustic domain (i.e., vocal exploration activity will be most apparent when infants and caregivers engage in mutual interaction) involved two steps in statistical analyses. First, a DA analysis of the vocal categories' discriminability in the three engagement settings and at the three ages with 7 infants (a total of 63 discriminant function analyses) was conducted. In this step, the squares of the canonical correlation values from the first (more significant) functions of the 63 discriminant analyses were obtained.

Then the degree of discriminability in the three engagement settings and at the three ages was assessed by a two way analysis of variance (ANOVA) with 7 infants. The three engagement settings and the three ages were the independent variables and the dependent variable was the square of the canonical correlation values obtained from the first step.

Temporal Domain: Systematic Sequential Patterns

The background hypothesis that pre-linguistic vocalizations exhibit systematic sequential patterns (i.e., infants will build their vocal categories, showing systematic repetitions or alternation patterns) was examined by lag sequential analysis (LSA) on the temporal sequential patterns across infants and ages. LSA assesses the significance of transitional probabilities in sequential data. The analysis units were the three vocal categories and the analysis results were produced with (3 vocal types x 3 vocal types) contingency tables of adjusted residuals at lag one. Unlike the standardized residual, the adjusted residual takes into account the N for each category and gives a fairer indication of how far off the observed count is from the expected count, thus indicating the degree of sequential associations of the vocal categories at lag 1. Given the significant imbalance of the sample sizes of the vocal categories (Vocant:4536, Growl:1394, and Squeal:755) and small sample sizes in the cells for many contingency tables, the adjusted residuals appear to be more appropriate than standardized residuals in finding significant contributors to explaining associations in the contingency tables (Bakeman & Quera, 1995). Therefore, the degree of sequential patterns was assessed by the adjusted residual scores in each cell of the contingency tables. In order to test the background hypothesis of sequential patterns in general, the adjusted residual values were averaged across infants and ages.

The main hypothesis in the temporal domain (i.e., infants exhibit stronger temporal sequential patterns in a mutually active circumstance than in an asymmetrical or an unengaged circumstance and at later ages) involved two steps in statistical analyses. In the first step, the contingency tables of the adjusted residuals at lag one in the three engagement settings and at the three ages of each infant were produced. Then the adjusted residuals of the three repetition patterns (i.e., vocant-vocant, growl-growl, and squeal-squeal repetitions), which were on the diagonal (cells go from the upper left corner cell to the bottom right corner cell) of the table were obtained. In the second step, a two way ANOVA of the adjusted residual values was conducted to examine if there was any significant difference in the adjusted residual values depending on the three engagement settings and the three ages. Subsequent posthoc analyses and the calculation of effect sizes were conducted.

Subsidiary Questions: Analyses on the Usage of the Vocal Categories and Duration by Age, Engagement, and Proximity

First, engagement setting, age and proximity effects on the usage of vocal categories were examined across infants by chi-square tests. Because standardized residual scores have been conventionally used to investigate significant associations for chi-square tests, standardized residual scores instead of adjusted residual scores are reported for the subsidiary questions. Secondly, chi-square tests and standardized residuals scores of each infant were reported to determine if the overall results were driven by outlier individual infants.

ANOVAs were run to determine if there was any difference in the duration of utterances by age, engagement and proximity. Since three-way interaction effects among age, engagement setting, and vocal types were not a primary concern and they are hard to interpret, two separate two-way ANOVAs, instead of a three-way ANOVA, were conducted to determine whether the duration of utterances differed by age, vocal type and engagement setting. The dependent variables were the log-transformed duration measures. The independent variables were age, engagement setting, and vocal type.

CHAPTER 4

Results

Data Preparation and Transformation

DA is computationally very similar to the multivariate analysis of variance (MANOVA) and subject to all the same assumptions. Box's test of equality of covariance matrices examines the assumption of multivariate homogeneity of covariance. According to the results, there was a significant difference in their covariance matrices, violating an assumption of the homogeneity of covariance (F (30, 16750566) = 48.8, p <. 001). However, DA as well as MANOVA is robust against the violation of the assumption of homogeneity and normality, especially with a large sample size like in the current study (Olson, 1974).

Since DA is sensitive to outliers, potential outliers were checked and excluded based on Stevens (2007). The skewness of the six variables' distribution was examined (Figures 2 - 6). Visual inspections of the histograms revealed significant positive skewness for all five variables, indicating that the data distributions have extended right tails. The skewness ratios (skewness values divided by the standard error) in Table 5 were more than the critical value of 2.58 on all of the six variables, indicating severe asymmetry in their distributions.

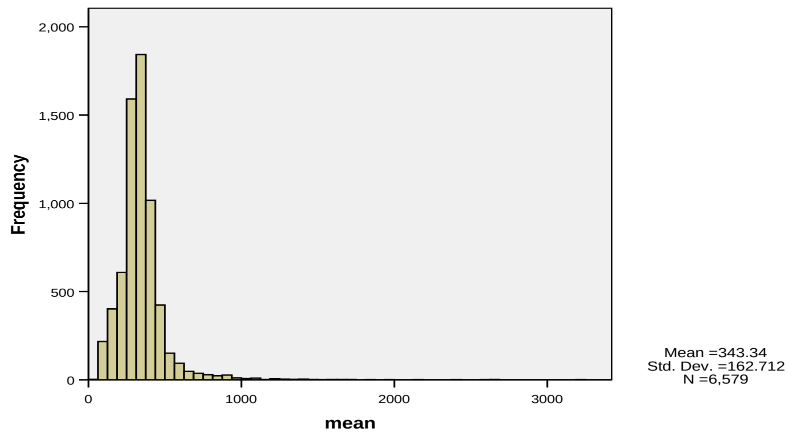

Figure 2. Histogram of the mean f_0 values of the raw data

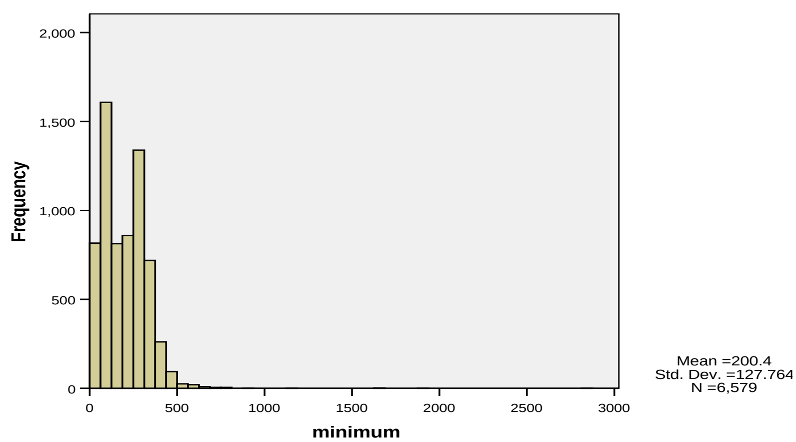

Figure 3. Histogram of the minimum f_0 values of the raw data

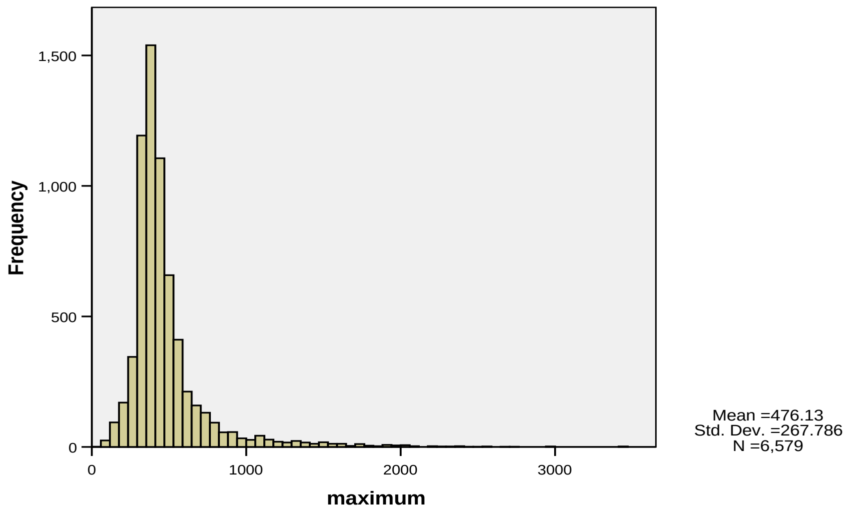

Figure 4. Histogram of the maximum f_0 values of the raw data

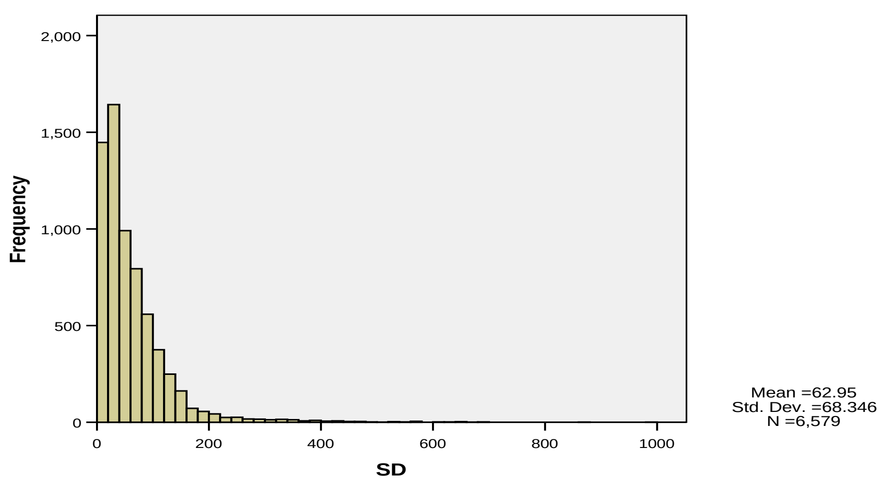

Figure 5. Histogram of the standard deviation of f_0 values of the raw data

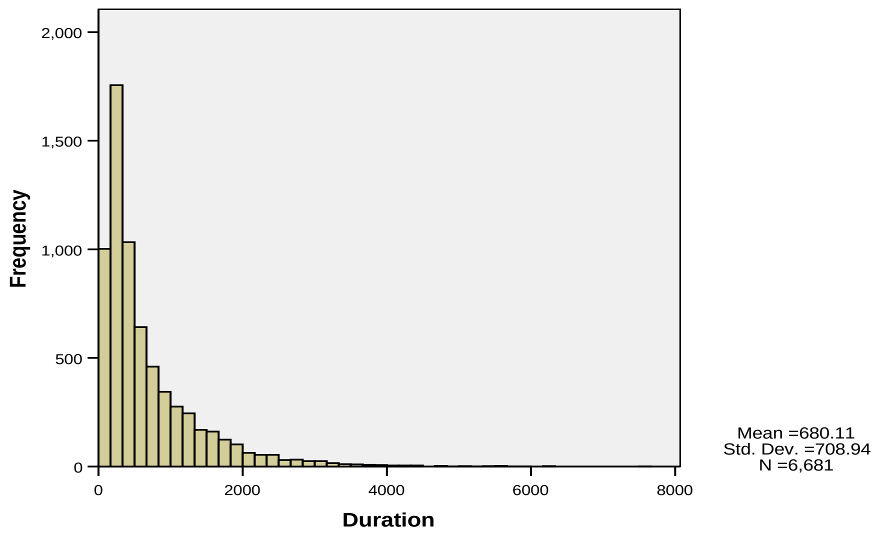

Figure 6. Histogram of duration values of the utterances of the raw data

Table 5

Skewness Values and Skewness Ratios of the Five Acoustic Variables

Acoustic Measure	Skewness (Skewness Ratio) with Raw Data	Skewness (Skewness Ratio) with Transformed Data
Mean f_0	4.759 (157)	-.285(9.5)
Minimum f_0	2.666 (86)	-.332 (11)
Maximum f_0	3.658 (121)	.744(24)
SD of f_0	3.847 (128)	-.263(8.7)
Duration	2.482 (82)	.155(5.2)

Although DA is fairly robust against the violation of the normality assumption, a data transformation was applied to the six variables since the skewness ratios were significantly higher than the critical value of 2.58. Among the transformations using the base 10 logarithm, natural logarithm, square root, and Arcsine, the natural logarithmic transformation yielded best results on the distributions as seen in the following histograms. However, the skewness ratios were still higher than the critical values of 2.58 (Figures 7 -11 and Table 5).

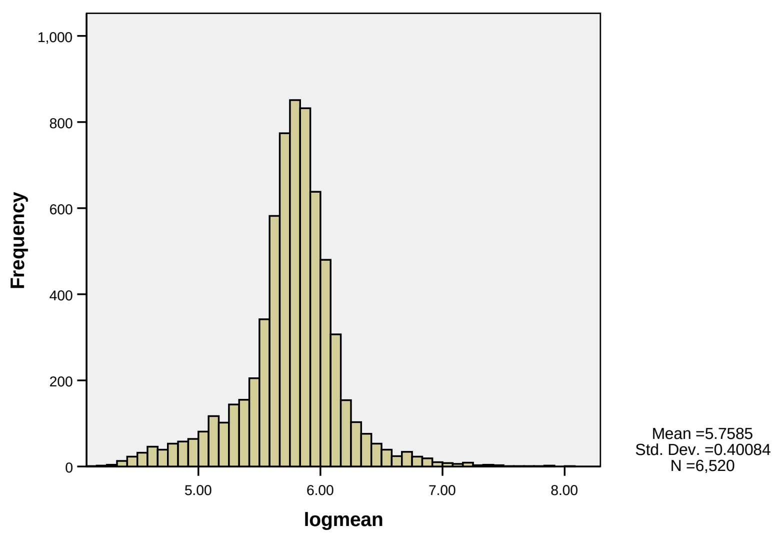

Figure 7. Histogram of the mean f_0 values of the data after logarithmic transformation and exclusion of potential outliers

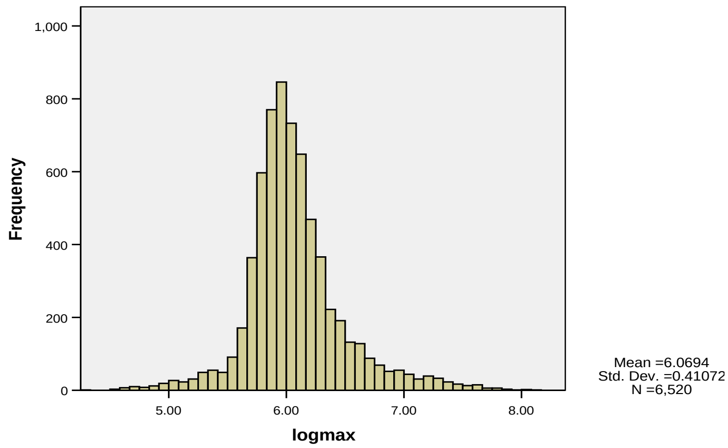

Figure 8. Histogram of the maximum f_0 values after logarithmic transformation and exclusion of potential outliers

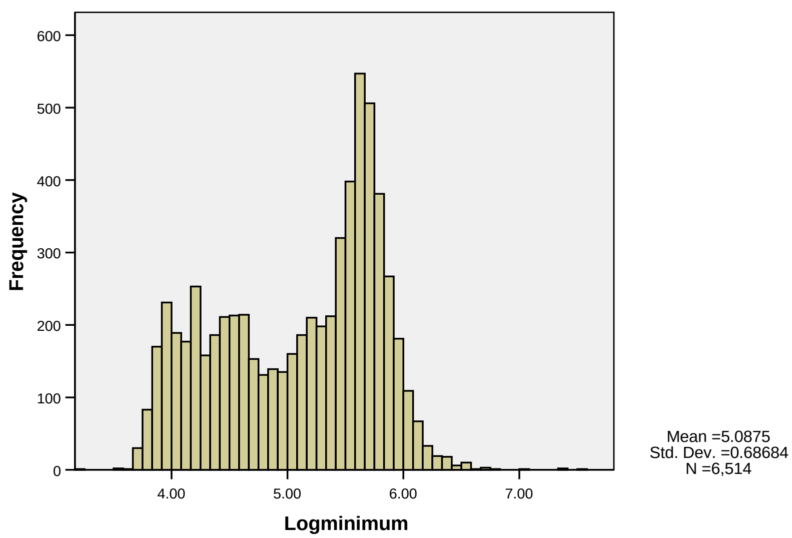

Figure 9. Histogram of the minimum f_0 values of the data after logarithmic transformation and exclusion of potential outliers

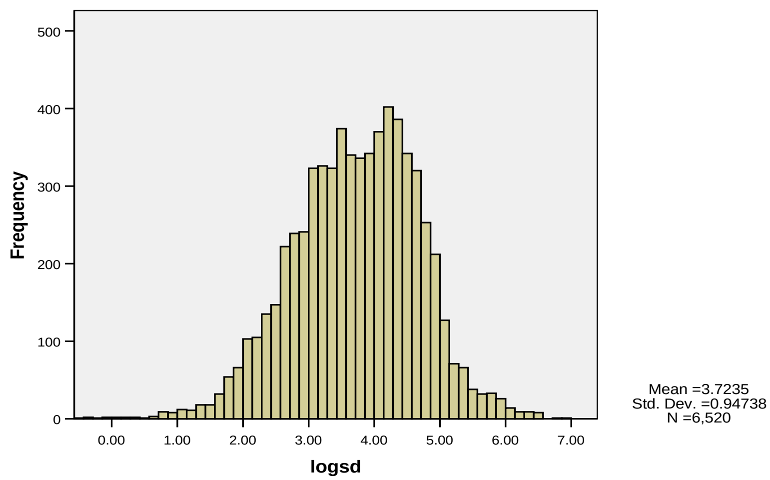

Figure 10. Histogram of the standard deviation of f_0 values of the data after logarithmic transformation and exclusion of potential outliers

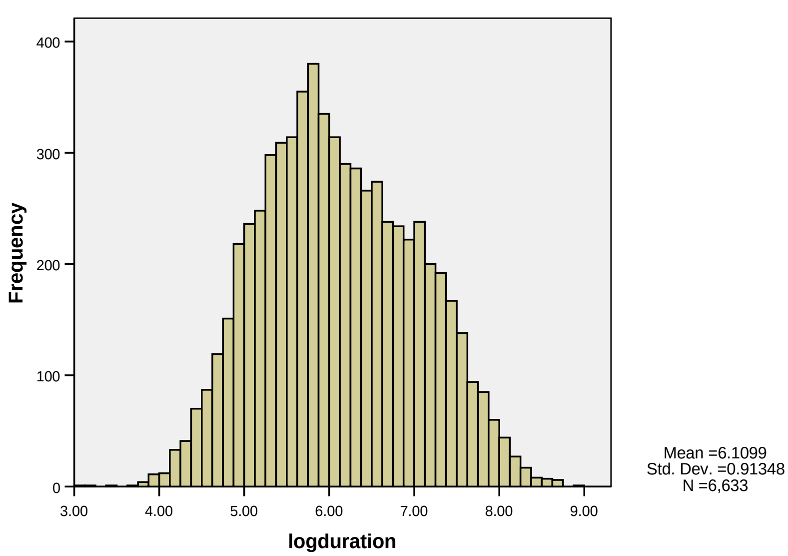

Figure 11. Histogram of duration of the vocalizations of the data after logarithmic transformation and exclusion of potential outliers

The possibility of serial dependency from utterance to utterance was examined. If the autocorrelation was strong, the tendency for repetition of vocalization types by the infants could have constituted a violation of the independence assumption for data to be submitted to ANOVA and DA. To evaluate autocorrelation, the data were submitted to a linear regression analysis for each of the five acoustic variables and the Durbin-Watson value assessing autocorrelation of the values was determined for each (Durbin & Watson, 1950). All of the five variables showed values lower than the critical value of 2, thus indicating that overall serial dependency of the data was negligible (mean f_0:1.30, maximum f_0:1.47, minimum f_0: 1.51, standard deviation of f_0:1.55, and duration of vocalization: 1.37).

Acoustic Domain: Distinctiveness of the Vocal Categories in the Acoustic Domains of f_0 and Duration

Two separate DA analyses, one with the log-transformed data with the exclusion of potential outliers and one with the original data with the inclusion of potential outliers, were conducted based on Stevens (2007). The frequency of each vocal type from the raw data and the log transformed data are presented in Table 6. The means, stand deviations and 90% confidence intervals (CIs) of the four f_0 measures and the duration values for each vocal type are reported in Table 7.

Table 6

Frequency of Vocal Types from the Raw Data and the Log-transformed Data

Vocal Category	Raw Data	Logarithmic -transformed Data without Outliers
Vocant	4536	4436
Growl	1394	1341
Squeal	755	737
Total	6685	6514

Table 7

Means, Standard Deviations, and 90% CIs of the Five Acoustic Measures

Acoustic Measure	Vocal Category	Mean (SD)	90% CI
Mean f_0	Vocant	323(87)	321-325
	Growl	271(102)	269 - 274
	Squeal	589 (317)	582 -595
Minimum f_0	Vocant	204 (109)	202 -206
	Growl	150 (98)	148 - 152
	Squeal	266 (214)	262 -271
Maximum f_0	Vocant	426 (161)	423 - 430
	Growl	416 (193)	412 - 420
	Squeal	874 (473)	864 - 884
SD of f_0	Vocant	47 (39)	41- 48
	Growl	62 (46)	61- 65
	Squeal	874 (133)	871- 877
Duration	Vocant	634 (647)	620 - 647
	Growl	699 (765)	684 - 715
	Squeal	941 (884)	923 - 959

Because there were three vocal categories, two discriminant functions (number of categories-1) were produced with both the raw data and the log-transformed data set. According to the results from the raw data set, the two functions were statistically significant (Wilks' Λ = .60, χ^2 (df = 10) = 3312, p < .001, Wilks' Λ = 0. 957, χ^2 (df = 4) = 288, p < .001) indicating that the five acoustic variables successfully distinguished the three vocal categories to a high degree. In other words, the results confirmed that the infants produced the vocal categories distinctively in the five acoustic domains; thus the categories were contrastive in terms of the five acoustic variables.

Likewise, according to the results from the log transformed data set without the potential outliers, the two functions were also statistically significant (Wilks' Λ = .64, χ^2 (df = 10) = 2917, p < .001, Wilks' Λ =. 928, χ^2 (df = 4) = 487, p < .001), also confirming that the infants produced the categories distinctively in the five acoustic domains.

The subsequent results based on the squares of the canonical correlations, the standardized coefficients, and the structure matrices did not differ across the two data sets. Since there were no substantial statistical differences, the results from the original raw data set are reported.

Given the results from the original raw data, Function 1 accounted for 92.9 % of the between group variance (this is the proportion of discriminating ability of the five acoustic variables found in Function 1) and 37% of the total variance whereas Function 2 accounted for 7.1 % of the between group variance and 4.2% of the total variance. For Function 1, the standardized coefficients indicated standard deviation of f_0 was the most influential variable in distinguishing the vocal categories, followed by mean f_0, minimum

f_0, duration, and maximum f_0 respectively. For Function 2, the standard deviation of f_0 was still the most influential variable in distinguishing the vocal categories, followed by mean f_0, maximum f_0, minimum $f_{0,}$ and duration. However, these standardized coefficient values do not tell us which groups were discriminated by the two functions. In order to identify the nature of the two functions across the groups, the locations of the group centroids were inspected using structure matrices (see Figure 12).

Function 1 had positive correlations with mean f_0, maximum f_0 and duration. The group centroids placed Squeal on the positive side of Function 1(x-axis) indicating that Squeal had a higher mean, maximum and standard deviation of f_0 while Vocant was in the middle, and Growl was on the negative side (see Figure 12). Thus, Function 1 appeared to differentiate Squeal from the other two groups. Function 2 (y-axis) had negative correlations with the minimum f_0 and positive correlations with the maximum f_0 which were the two major variables. The group centroids placed Growl on the positive end of the dimension while Squeal was in the middle and Vocant on the negative end, indicating Growl had the lowest minimum f_0 values (remember the negative correlation between the minimum f_0 and Function 2) among the three vocal types. Thus, Function 2 appeared to discriminate Growl from the other two vocal types.

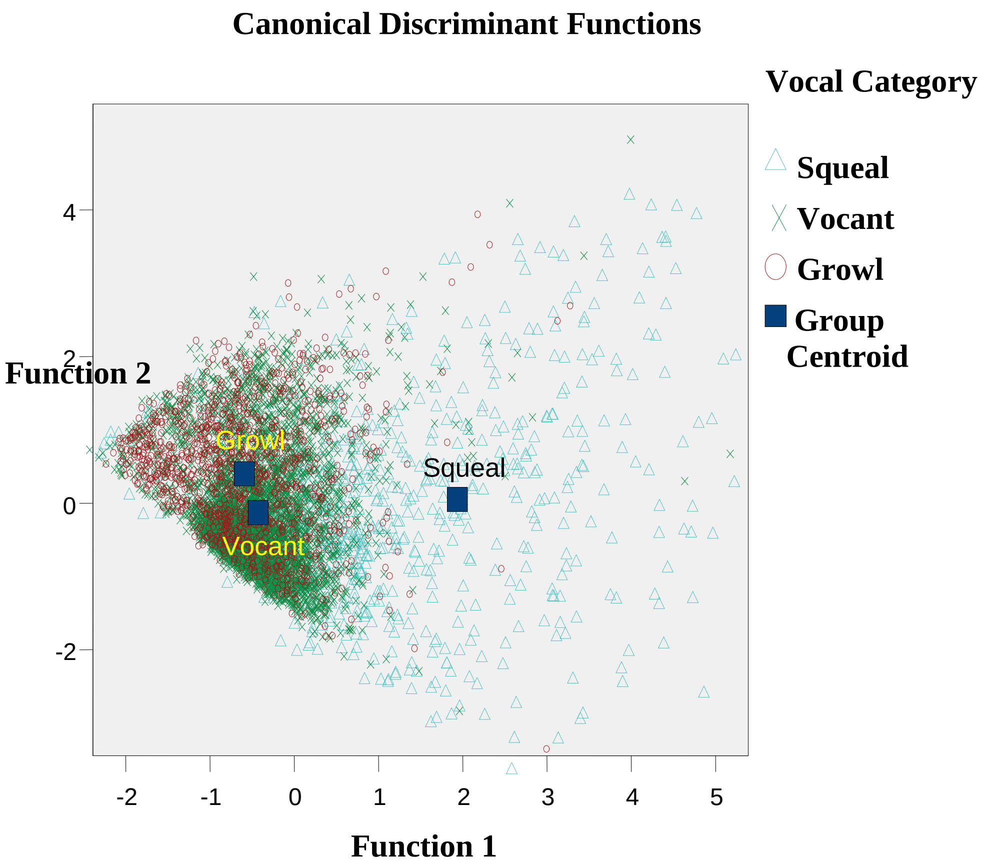

Figure 12. Plots of Three Vocal Categories' Centroids

Table 8

Standardized Coefficients, Correlations, and Groups' Centroids of Vocal Category

Acoustic Measure	Function 1	Function 2
Mean f_0	.572 (.882)	-.778 (-.448)
Maximum f_0	-.001(.837)	.084 (.174)
Minimum f_0	.220 (.269)	-.052 (-.739)
SD of f_0	.591(.721)	.877 (.686)
Duration	.061(.167)	-.017 (.228)
Centroids	Function 1	Function 2
Squeal	2.130	.051
Growl	-.398	.401
Vocant	-.234	-.130

Note. Numbers in parentheses represent the correlations

All pairwise *F*-tests were significant indicating that in the two dimensional space defined by the functions, all vocal category groups were different from each other (Squeal vs. Vocant: F (4, 6573) = 897, $p < .001$; Squeal vs. Growl: F (4, 6573) = 782, $p < .001$; Vocant vs. Growl: F (4, 6573) = 80, $p < .001$). Effect sizes were large for Squeal and Vocant and Squeal and Growl, while effect sizes were small for the Growl and Vocant comparison ($\eta^2_{Squeal/Vocant} = 0.35$; $\eta^2_{Squeal/Growl} = 0.33$; $\eta^2_{Growl/Vocant} = 0.046$) in accord with the rules of thumb recommended by Cohen (1988)

Temporal Domain: Vocal Category Exploration Manifested in Systematic Repetition or Alternation Patterns

A lag sequential analysis (LSA) was conducted to examine the tendency for systematic sequential patterns of vocal categories. For this analysis, 7895 utterances across infants and ages were included. According to the results presented in Table 9, 29 sessions out of a total of 42 sessions (about 69%) displayed significant repetition or alternation patterns of vocalizations in lag one statistics given the adjusted residual values in cells of the contingency tables ($p < .05$). Only four instances of statistically significant alternation patterns (2 instances of vocant-growl sequences and one instance of growl-vocant and squeal-growl sequences) by two infants were detected. All other significant patterns were repetitions of the same vocal types. The positive or negative adjusted residual values for diagonal (repetition patterns) and off – diagonal cells (alternation patterns) were used as dichotomous variables for a binomial test. There were 103 positive residual values (82%) and 23 negative values (18%) out of 126 diagonal cells ($p < 10^{-12}$). There were 61 positive residual values and 191 negative values out of 252 off-diagonal cells (24% and 76% respectively) ($p < 10^{-16}$). The results indicate that there was significant preference for positive values on the diagonals and negative values in the off-diagonals, meaning that repetition patterns were significantly preferred while alternation patterns were avoided. Given that there were 61 off-diagonal values that were positive, it seems plausible that the detected four cases of “significant” alternation patterns could have occurred by chance. Among the 13 sessions in which significant sequential patterns were not found, 6 sessions had fewer than 60 ‘utterances in sequences’ (total utterances used for LSA analyses), suggesting low power to detect any significance.

Table 9

Utterances in Sequences, Total Utterances and Sequential Patterns in Each Session of Each Infant

Infant	Session	Utterances in Sequences (Total Utterances)	Sequential Pattern
AD	1	227(260)	V, G, S
	2	163(201)	S
	3	79 (132)	
	4	78 (158)	V, G, S
	5	80 (110)	S
	6	126 (165)	V, G, S
EA	1	95(203)	V, G, S
	2	229 (269)	
	3	84 (95)	S
	4	273 (321)	
	5	55 (98)	
	6	140 (196)	V, S
KA	1	38 (83)	
	2	59 (76)	
	3	18 (47)	
	4	55 (89)	
	5	66 (74)	V, G
	6	50 (84)	V, G
KS	1	217(247)	V,G,S
	2	87 (126)	S
	3	60 (93)	V,G
	4	205(234)	
	5	114 (147)	V,G,S
	6	83 (128)	V,G
LO	1	215 (258)	V,S
	2	188 (228)	
	3	140 (194)	S,
	4	89 (154)	G-V
	5	160 (200)	G-V,V-G,S
	6	136 (182)	V,G,S
LP	1	164 (196)	V,G,S
	2	276 (308)	V,G
	3	33 (63)	
	4	137 (183)	V,G
	5	40 (80)	S-G
	6	32 (61)	
SM	1	190 (231)	V,G,S
	2	177 (214)	V,S
	3	178 (222)	V,S
	4	99 (133)	V,S
	5	45 (99)	G
	6	69 (115)	

Note. V= Vocant repetition, G= Growl repetition, S= Squeal repetition, V-G = vocant growl alternation, G-V= growl vocant alternation, S-G = squeal growl alternation

A (three vocal types x three vocal types) table collapsed from the 42 sessions (7 subjects x 6 sessions) is presented in Table 10. The raw entries are antecedent vocal events and the column entries are subsequent vocal events. The cells contain two forms of lag 1 statistics (adjusted residuals and standard deviation of the adjusted residuals in parentheses) and 90% confidence intervals of the adjusted residuals. The repetition tendencies of the same vocal types were indicated by larger positive residual values on the diagonal accompanied by smaller or negative residuals on the off-diagonal. The positive mean adjusted residuals on the diagonals ranged from 1.22 for growl repetition to 1.71 for vocant and squeal repetition, indicating a strong tendency for the infants to repeat the same vocal categories.

Positive residuals off the diagonal would indicate systematic alternation between the categories. This pattern occurred for the growl-vowel, vowel-growl, and squeal-growl sequences in LO and LP (Table 9), but these were very rare and the mean adjusted residuals of the sequences across all the infants were negative or close to zero (mean adjusted residuals of growl-vowel:-1.0, vowel-growl:-1.13, squeal-growl: .03). Other possible alternations among the categories, especially the squeal-vocant or vocant-squeal sequences (mean adjusted residuals: -1.15, and -1.07 respectively), were unlikely to happen as indicated by the high negative mean residuals off the diagonal. Thus results of the LSA analyses confirmed the background hypothesis that the infants show vocal exploration activities through systematic repetition of the vocal categories.

Table 10

Mean Adjusted Residuals, Standard Deviations, and 90% CIs of the mean adjusted residuals of Vocalizations in Lag 1 Statistics across Infants and Sessions

	Vocant / 90% CI	Growl / 90% CI	Squeal/90% CI
Vocant	1.71(0.65)/1.54 ~1.89	-1.13(0.48)/-1.24 ~-1.0	-1.07(0.61)/-1.22 ~ .91
Growl	-1.00(0.68)/-1.17 ~ -1.0	1.22(0.73) /-1.0 ~ -1.4	-0.11(0.23)/ -.1 ~ -.05
Squeal	-1.15(0.33)/-1.23 ~ 1.06	0.03(0.38) /-.06 ~ 0.13	1.71(0.73)/ 1.5 ~1.9

Note. Numbers in parentheses represent SD of the adjusted residuals. "~" instead of "-" was used to represent the range of CIs due to possible confusion with negative values.

In summary, as hypothesized, the infants appeared to show significant voluntary vocal exploration activities to build their vocal categories by both utilizing the acoustic variables to contrast the vocal categories and by systematically repeating the categories.

Primary Questions Regarding Vocal Exploration Activity in Various Engagement Settings

As a first step to examine the disciminability of the vocal categories in various engagement settings, 63 discriminant analyses (7 infants x 3 engagement settings x 3 ages) were run. From the 63 discriminant functions the squares of the canonical correlation values (the percent variance accounted for), were obtained from the first (most significant) functions.

In the next step, an analysis on the degree of discriminability in the three engagement settings and at the three ages using a two-way ANOVA test with 7 infants was conducted. The two independent variables were engagement setting and age. The dependent variable was the squares of the canonical correlation value obtained from the first step. Means and standard deviations are presented in Table 11.

A univariate outlier procedure was run and one potential outlier was removed from the data. The results from the test for homogeneity of variance was not significant (Levene F (8, 53) =1.889, p = .081), indicating that this assumption underlying the application of ANOVA was met. The ANOVA produced a statistically non-significant interaction between the engagement setting and the age (F (4, 53) = .629, p = .644), indicating that the age effect on the square of the canonical correlation was not dependent on whether subjects were in different engagement settings. Main effects analyses were conducted to test differences in the degree of the discriminiability by age and engagement. The results indicated that there was no significant mean difference between engagement settings (F (2,53) =.856, p =.431) or ages (F (2,53) = 2.11, p =.130), suggesting that there was no credible evidence that infant vocal exploration activities, manifested by the degree of the discriminabilty among the vocal categories, are more active in mutually interactive settings or at later ages.

Primary Questions Regarding Sequential Patterns of Vocal Categories

In order to test if infants exhibit stronger temporal sequential patterns in a mutually active circumstance than in an asymmetrical or an unengaged circumstance and at later ages, tables of adjusted residuals at lag one in the three engagement settings and at the three ages of each infant were produced. However, the low N in cells for unengaged settings caused a power problem. The mean and standard deviation of the

Table 11

Means, Standard Deviations, and 90% CIs of the Squares of Canonical Correlation Values for the Three Engagement Settings and the Three Ages

Engagement	Age	Mean	Std. Deviation	N	90% CI
Symmetrical	Age 1	35.29	16.67	7	24 -45
	Age 2	49.62	14.75	7	40 – 58
	Age 3	53.1	18.95	7	41 -64
	Total	46	17.85	21	39 -52
Asymmetrical	Age 1	39.48	23.19	7	25 -53
	Age 2	48.17	13.43	7	39 -56
	Age 3	55.32	15	7	45 – 64
	Total	47.66	18.08	21	31 – 54
Unengaged	Age 1	50.98	22.34	7	37 -64
	Age 2	60.15	29.64	6	41 -78
	Age 3	49.7	17.81	7	38 – 60
	Total	53.28	22.6	20	44 -61
Total	Age 1	41.92	20.99	21	28 -54
	Age 2	52.27	19.63	20	40 -64
	Age 3	52.71	16.62	21	42 -63
	Total	48.91	19.51	62	44 – 52

adjusted residuals in the unengaged settings were reported but the data from unengaged settings could not be incorporated for further analysis. The adjusted residual values of the diagonals (same vocal type repetitions) on the contingency tables were obtained.

A two-way ANOVA test was conducted to determine whether there was a difference in the adjusted residual values by engagement setting and age. The dependent variable was the adjusted residual value. Means and standard deviations of the adjusted residuals are presented in Table 12. One potential outlier was excluded from the analysis.

The skewness value was .623 and skewness ratio was 2.87, which was slightly over the critical value of 2.58. The raw data was used since the skewness ratio was not significant. The test for homogeneity of variance was not significant (Levene F (5,119) = 1.97, p = .089), indicating that this assumption underlying the application of ANOVA was met.

The ANOVA produced a statistically non-significant interaction between engagement setting and age (F (2, 119) = .264, p = .769), indicating that the age effects on the adjusted residual values were not dependent on different engagement settings. The main effects analyses were conducted to test differences in the degree of repetitions in the two engagement settings and at the three ages. The results indicated that there was a significant mean difference among age groups (F (2,119) = 7.681, p = .001).

The posthoc analyses indicated that there were significant group mean differences between Age 1 and Age 2, and Age 1 and Age 3. The effect sizes were large between Age1 and Age 3 (d = .81) and medium between Age 1 and Age 2 (d = .52) in accord with the rules of thumb recommended by Cohen (1988). However, contradicting the hypothesis, the strongest repetition patterns were observed at the earliest age followed by the latest age and the middle age, respectively (Table 12). The main effects of the two engagement settings on the adjusted residuals were not significant (F (1,119) = 1.75, p = .188). Thus, no statistical evidence was found favoring the hypothesis that infant vocal exploration activities through sequential vocal repetitions are more active in mutually interactive settings.

Table 12

Means, Standard Deviations and 90% CIs of the Adjusted Residual Values of the Three Engagement Settings and the Three Ages

Engagement	Mean	SD	N of Sessions Used	90% CI
Symmetrical	.64	.33	42	.52 -.76
Asymmetrical	.52	.17	42	.45 -.58
Unengaged	.89	.27	17	.79 -.99

Age	Mean	SD	N of Sessions Used	90% CI
1	1.12	1.0	42	.87-1.37
2	.34	.73	42	.15 -.53
3	.61	.97	41	.36 -.89

Subsidiary Questions Regarding Vocal Category Development in Various Engagement Settings

Usage (frequency of Occurrence) of Vocal Categories in Various Engagement Settings

A potential difference in the usage of the three vocal categories in various engagement settings was assessed by a chi-square test. According to the result, there was a significant association between the vocal category frequency and the engagement settings (χ^2 (4, N = 6559) = 53.135, $p < .001$). Standardized residual values for each cell of the contingency (Table 13) showed that infants showed a systematic association pattern between the Unengaged setting and the Vocant category, and between the Symmetrical setting and the Squeal Category. These patterns suggest infants produced significantly more vocants when they were unengaged and significantly more squeals when they were mutually interactive. Likewise, the strong negative residual values indicated systematic fewer usages of vocants in mutual interaction and squeals in unengaged interaction settings.

Table 13

Frequency and Standardized Residuals of Vocal Categories in the Three Engagement Settings

	Engagement Settings			
Vocal Type	Symmetrical	Asymmetrical	Unengaged	Total N
Vocant	**-2.0** (2093)	0.6 (1609)	**2.6** (757)	(4459)
Growl	0.8 (682)	-0.2 (476)	-1.1 (193)	(1351)
Squeal	**3.8** (439)	-1.1 (248)	**-5.0** (62)	(749)
Total N	(3214)	(2333)	(1012)	(6559)

Note. Numbers in parentheses represent standardized residuals.

The adjusted residual scores of the individual infants (Table 14) revealed strong associations between the two variables in all infants except KA who had the smallest sample sizes (less than 66 utterances per session) (refer to Table 9). In addition, the results confirmed the general trends that were observed in the collapsed data presented in Table 13. Five infants showed negative adjusted residual values for the Symmetrical – Vocant and Unengaged-Squeal associations, indicating fewer usages of such patterns. In contrast, four infants showed significantly strong tendencies toward the Squeal-Symmetrical association.

Table 14

Chi-square Test Results and Contingency Table of Standardized Residuals of Each Infant

Chi-square Test	*AD:* χ^2 *(4, N=965)= 9.472, p =.05** *EA:* χ^2 *(4, N=1096)= 41.659, p* < .001 * *KA:* χ^2 *(4, N=438)= 3.697, p =.449* *KS :* χ^2 *(4, N=967)= 18.683, p =.001** *LO:* χ^2 *(4, N=1216)= 16.109, p =.003** *LP :* χ^2 *(4, N=889)= 12.772, p =.012** *SM :* χ^2 *(4, N=965)= 18.829, p =.001**		
	Symmetrical	Asymmetrical	Unengaged
Vocant	-0.4/-1.4/0.3/.0/-0.8/-0.7/-**2.7**	-0.4/ 0.9/0.4/0.6/0.1/0.6./0.8	1.4/**2**/**2**/-1.0/-1.0/1.6/-1.2/1.2
Growl	-0.9/0.2/-0.4/1/-0.2/-0.9/-0.7	0.2/0.7/-0.6/**-2.0**/0.8/0.6/1.0	-1.7/-0.9/1.3/**3.2**/-1.1/0.4/-0.3
Squeal	0.1/**3.2**/-0.3/-0.3/**2.1**/**2.0**/**2.4**	1.6/0.1/1.0/0/-1.3/**-2.4**/**-2.2**	-0.3/-**4.5**/0.3/-1.4/-**2.2**/1.3/-**1.8**

Note. Adjusted residuals of each infant are distinguished by slashes. The bolded numbers represent significant contributors for chi-square results in each individual.

Usage of Vocal Categories by Age

A significant association between the vocal categories and the three ages (χ^2 *(4, N = 6684) = 32.1*, *p* < .001) was also found. Infants produced significantly more growls at the earliest age while they exhibited more vocants at the latest age (Table 15). It was also shown that infants had a strong tendency to produce fewer vocants at the earliest age and fewer growls at the latest age. Individual chi-square test results revealed that all infants showed significant associations between vocal categories and ages (Table 16). Standardized residual scores of the contingency tables of the individual infants confirmed the same trends that were found in Table 15. However, individual differences in the usage of squeals were evident across the ages (Table 16).

Table 15

Frequency and Standardized Residuals of Vocal Categories at the Three Ages

Vocal Type	Age Age 1	Age 2	Age 3	Total N
Vocant	**-2.0** (1828)	.5 (1454)	**2.0** (1253)	(4535)
Growl	**3.2** (667)	-.5 (431)	**-3.6** (296)	(1394)
Squeal	0.6 (329)	-.5(231)	-.1(195)	(755)
Total N	(2824)	(2116)	(1744)	(6684)

Note. The numbers in parentheses represent the frequency. The bolded numbers represent the significant contributors of the chi-square result.

Table 16

Chi-square Test Results and the Contingency Table of Standardized Residuals of Each Infant at the Three Ages

Chi-square Test	AD: χ^2 (4, N=1028)= 20.935, $p < .001$ * EA: χ^2 (4, N=1104)= 22.349, $p < .001$ * KA: χ^2 (4, N=456)= 9.778, $p = .044$* KS: χ^2 (4, N=957)= 27.946, $p < .001$ * LO: χ^2 (4, N=1216)= 34.229, $p < .001$ * LP: χ^2 (4, N=891)= 27.599, $p < .001$ * SM : χ^2 (4, N=1014)= 9.429, $p = .051$*		
	Age 1	Age 2	Age 3
Vocant	-0.7/-1.3/1.0/-**1.9**/0.2/-1.1/-0.5	-1.2/1.4/-0.1/0.9/-**2.5**/0.6/-0.3	**2.1**/-0.2/-0.9/1.3/1.0/-**3.6**/**2.4**
Growl	0.8/1.6/-.07/**2.4**/0.5/**2.9**/0.9	1.9/**-3.0**/0/-1.2/**3.2**/**2.7**/0.9	**-3.0**/1.7/0.7/-1.5/-**3.6**/**-1.8**/**-2.5**
Squeal	0.2/1.6/**-2.0**/**2.8**/-1.1/-1.0/0.2	-1.0/-0.5/0.4/-1.2/-1.3/**2.0**/-0.3	0.8/-1.3/1.6/-**1.9**/**2.4**/-0.7/0

Note. The bolded numbers represent the significant contributors of the chi-square results.

Usage of the Vocal Categories by Physical Proximity

Among the three proximity categories of Immediate, Distant and Alone, the Alone types were very rare in the data. Thus, vocalizations coded as Alone were excluded from the analysis. The physical proximity effects on the usage of vocal categories were assessed by a chi-square test. There was a significant association between vocal category and proximity (χ^2 *(2,* N =5569*)* = 52.49, *p < .001*; Table 17). Infants produced significantly more growls when they were close to caregivers and more vocants when they were separated from caregivers. It was also shown that the infants had strong avoidance tendencies towards vocants when they were close to caregivers and growls, and squeals when they were separated from the caregivers.

Table 17

Frequency and Standardized Residuals of the Vocal Categories in the Two Proximity Conditions

	Proximity		
Vocal Type	Immediate	Distant	Total
Vocant	**-1.9** (3295)	**3.5** (1164)	(4459)
Growl	**2.8** (1121)	**-5.0** (230)	(1351)
Squeal	1.0 (597)	**-1.8** (152)	(749)
Total N	(5013)	(1546)	(6559)

Note. The numbers in parentheses represent the frequency. The bolded numbers represent the significant contributors of the chi-square results.

Four out of the seven infants showed a significant association between proximity settings and vocal types (Table 18), with results of LO and KS approaching the significance level of .05 (p = .065 and p = .06, respectively). The overall positive standardized residual scores for Vocant and Distant and Growl and Squeal with Immediate associations confirmed the general association tendencies observed in Table 18.

Table 18

Chi -square Test Results and the Contingency Table with Standardized Residuals Score of Each Infant in the Two Proximity Conditions

Chi-square Test	AD: χ^2 (2, N=965) =8.629, p = .013 * EA: χ^2 (2, N=1096) =17.535, $p < .001$* KA: χ^2 (2, N=438) = 2.647, p = .266 KS : χ^2 (2, N=967) = 5.465, p = .065 LO : χ^2 (2, N=1218) = 5.628, p = .06 LP: χ^2 (2, N=889) =10.998, p = .004* SM : χ^2 (2, N= 988) = 13.902, p = .001*	
	Immediate	Separated
Vocant	-0.7/-1.4/0.4/-0.4/-0.6/-0.1/-0.7	**1.8**/1.3/-0.6/1.0/1.2/0.4/1.5
Growl	0.7/1/-0.6/0.8/0.4/0.9/1.4	**-1.8**/-1.0/1.1/**-1.8**/-0.8/**-2.2/-3.0**
Squeal	0.4/**2.5**/0.4/0.3/0.8/-0.9/0.1	-0.9/**-2.3**/-0.6/-0.7/-1.6/**2.1**/-0.2

Note. The bolded numbers represent the significant contributors of the chi-square results.

Association between Engagement and Proximity

As acknowledged earlier, the recording protocol that was used in the IVOC laboratory created a significant association engagement and proximity. Symmetrical engagement circumstances tended to occur with immediate proximity between participants whereas asymmetrical and unengaged circumstances occurred more often with distance between participants. In order to test the degree of association between engagement and proximity, a chi-square test was conducted. As expected, the result indicated a very strong association (χ^2 *(2,* N =6559*)* = 1470*, p < .00*; Table 19).

Table 19

Frequency and Standardized Residuals of Engagement setting at the Two Proximity Settings

	Engagement Setting			
Vocal Type	Symmetrical	Asymmetrical	Unengaged	Total N
Immediate	8.9 (2898)	.3 (1796)	-16.3(319)	(5013)
Distant	-16 (316)	-.6 (537)	29.4 (693)	(1546)
Total N	(3124)	(2333)	(1012)	(6559)

Note. The numbers in parentheses represent the frequency.

Effect of Age, Engagement Setting and Vocal Categories on Utterance Duration

Age by Vocal Type

Means and standard deviations of the three ages and the three vocal types are presented in Table 20. The dependent variable was the logarithmically transformed duration measure. The ANOVA produced a statistically significant interaction between vocal type and age (F (4, 6624) = 3.145, p =.014), indicating that the duration of each vocal type was dependent on age (Figure 13). The effect size indicated by partial eta square was .002, representing a small effect by Cohen's rules of thumb. Simple main effect analyses were conducted to test the duration difference across the vocal types for each of the three ages. Results indicated that there was a significant duration difference between the vocal categories at all ages (age 1: F (2, 6624) = 28.16, age 2: F (2, 6624) = 40.35, age 3: F (2, 6624) = 12.37) given a critical value ($F_{0.05;6624}$ = 2.62). Squeal had the longest duration followed by Growl and Vocant respectively at all ages. The effect sizes for the group comparison at the three ages were calculated using Tukey (Table 21). All the comparisons showed significant group differences except the Vocant and Growl comparison at the middle age.

Table 20

Means, Standard Deviations and 90% CIs of Duration and Frequency of Vocal Types at the Three Ages

Vocal Type	Age	Mean	Std. Deviation	N	90% CI
Vocant	1	632	658	1827	607-657
	2	618	635	1454	590- 645
	3	645	636	1251	616 - 675
	Total	631	645	4532	615 - 647
Growl	1	687	751	667	639 -735
	2	669	783	431	607 -731
	3	763	784	295	687 - 838
	Total	697	769	1393	663- 731
Squeal	1	882	808	329	809 - 955
	2	1005	828	231	910 -1095
	3	954	1043	195	831 -1071
	Total	938	881	755	886 - 991
Total	1	674	704	2823	652 - 694
	2	671	701	2116	645 - 696
	3	700	725	1741	671 -728
	Total	680	708	6680	665 - 694

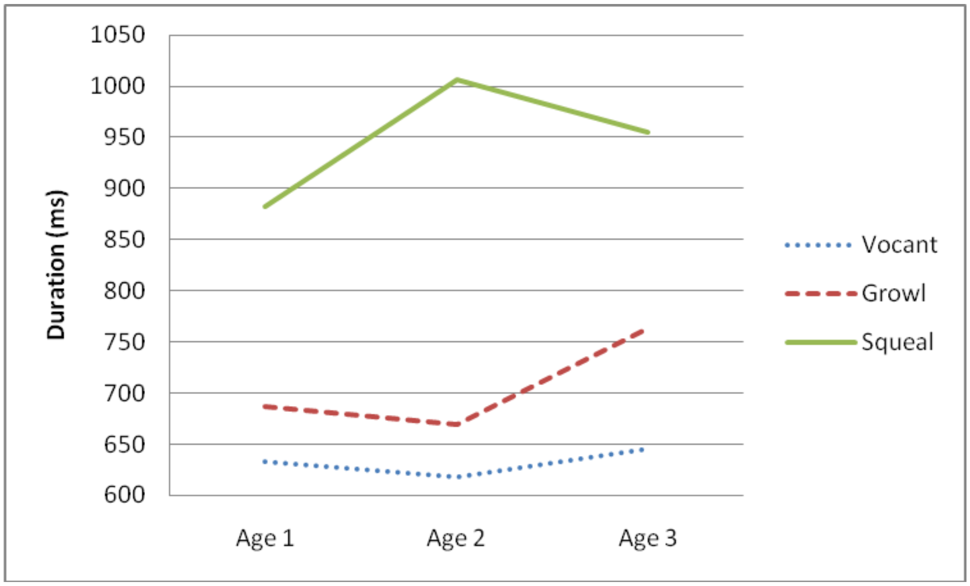

Figure 13. Interaction plot of duration of utterances by Age and Vocal Type

Table 21

Tukey Values for Vocal Type Contrasts in Duration at the Three Ages

	Age 1	Age 2	Age 3
Vocant-Growl	-3.772*	.251	-5.23*
Vocant-Squeal	-16.09*	-22.55*	-13.21*
Growl-Squeal	-12.31*	-22.80*	-7.98*

Note. Tukey critical value: $Q_{.05;3,6624} = 2.77$. * indicates significant Tukey values given the critical value.

Age by Engagement Setting

Means and standard deviations of the duration of utterances by the three ages and the three vocal types are presented in Table 22. The dependent variable was the logarithmically transformed duration measure. The independent variables were age and engagement setting. The ANOVA produced a statistically significant interaction between engagement setting and age (F (4, 6549) = 4.98, p = .001), indicating that the duration of vocalizations at each age was dependent on the settings of engagement (Refer to the interaction plot of Figure 14). The effect size reported by partial eta square was .003, representing a small effect by Cohen's rules of thumb. The simple main effect analyses indicated that there was a significant duration difference between the engagement settings at all ages (Age 1: F (2, 6549) = 8.10, Age 2: F (2, 6549) = 10.61, Age 3: F (2, 6549) = 8.72) given a critical value ($F_{0.05;\,2000}$ = 2.61). Effect sizes for group comparison at the three ages were calculated using Tukey (Table 23). All the comparisons showed significant group differences at all ages.

Table 22

Means, Standard Deviations and 90% CIs of Duration in the Three Engagement Settings

Age	Engagement Setting	Mean	Std. Deviation	N	90% CI
1	Symmetrical	704	722	1544	674-734
	Asymmetrical	614	620	956	581- 647
	Unengaged	795	915	234	696 -893
	Total	680	709	2734	658- 703
2	Symmetrical	743	752	975	703- 783
	Asymmetrical	627	653	742	588- 667
	Unengaged	571	637	390	518- 624
	Total	670	701	2107	645- 696
3	Symmetrical	759	760	695	711- 806
	Asymmetrical	601	632	635	560- 642
	Unengaged	734	775	387	669 -799
	Total	695	722	1717	666- 724
Total	Symmetrical	728	739	3214	706- 749
	Asymmetrical	615	634	2333	593- 636
	Unengaged	685	766	1011	645- 725
	Total	681	710	6558	666 -695

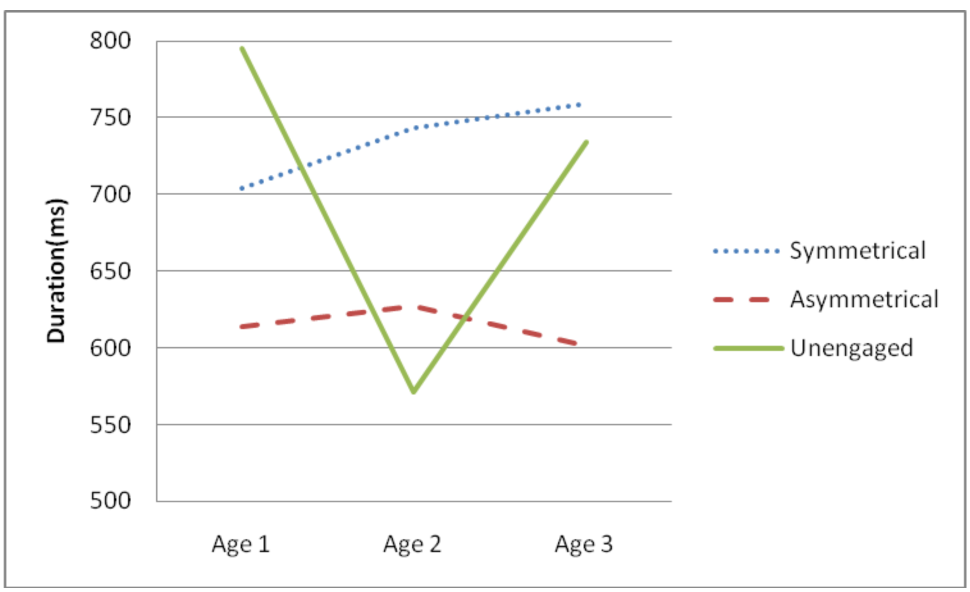

Figure 14. Interaction plot of duration of utterances by Age and Engagement

Table 23

Tukey Values for Engagement Group Contrasts of the Duration Measure by Age and Engagement

	Age 1	Age 2	Age 3
Symmetrical -Asymmetrical	10.34*	13.31*	18*
Symmetrical- Unengaged	-10.48*	19.80*	2.88*
Asymmetrical- Unengaged	-20.81*	6.49*	15.30*

Note. Tukey critical value: $Q_{.05;3,2000}$ = 2.77. * indicates significant Tukey values given the critical value.

Further Exploration

The main hypothesis regarding the discriminability of the three vocal categories in the three engagement settings and at the three ages were further investigated due to possible sample size effects in the discriminant analyses. Out of the total 63 discriminant analyses, 25 cases had less than 60 utterances. Based on Stevens (2007), a sample size of 60 (20 per categories) is recommended to provide sufficient power for DA. Therefore, additional analyses were run to determine if the small sample sizes had affected the obtained results.

First, an ANOVA without the 25 cases of low N was conducted. No statistical significance was detected for either interaction or main effects (interaction effect: F (4, 29) = .391, p = .813, age effect: F (2, 29) = .089, p = .915, engagement effect: F (2, 29) = 595, p = .558). Second, in order to achieve larger sample sizes, DAs were run collapsed across age for the analyses on three engagement settings effect and collapsed across engagement setting for the analyses on the three ages. This method allowed six data points (three squares of canonical correlation values from the three ages and three from the three engagement settings) for each infant. A total of 42 scores representing percent variance accounted for from the seven infants was collected. Then paired t- tests between the age groups and the engagement groups were conducted (Table 24). According to the results, there was a significant group difference in the square of the canonical correlation values between Age 1 and Age 3, indicating that infants showed stronger discrimination of the categories in terms the acoustic variables at the latest age. No other comparisons were statistically significant.

Table 24

Means, Standard Deviations, and t -test Results

	Mean	SD	*t*	df	Significance (2 tailed)
Symmetrical vs. Asymmetrical	.013259143	.151087160	.232	6	.824
Asymmetrical vs. Unengaged	-.037401429	.228004148	-.434	6	.679
Symmetrical vs. Unengaged	-.024142286	.166627157	-.383	6	.715
Age1 vs. Age 2	-.087364857	.123203239	-1.876	6	.110
Age1 vs. Age 3	-.118014000	.052482948	-5.949	6	.001*
Age2 vs. Age 3	-.030649143	.124904795	-.649	6	.540

* $p < .05$.

CHAPTER 5

Discussion

Summary of Results

In summary the results suggest 1) that infants build their vocal categories, showing the ability to produce three vocal categories distinctively in the acoustic domains of f_0 and duration, 2) that infants show systematic repetition patterns among these three categories, 3) no statistical evidence for the hypothesis that vocal exploration is most active when infants and caregivers engage in mutual interaction, 4) that vocal exploration activity is more apparent as infants grow older as indicated by the degree of the distinctiveness of the categories when the data were collapsed across engagement settings, 5) that the degree of systematic vocal repetition is strongest at the earliest age, and 6) that infants show systematic patterns in the usage of vocal categories varying by engagement settings, proximity conditions, and ages as well as in the duration of utterances by ages, engagement settings

Background Hypotheses: Development of Vocal Categories by Exploratory Activities

Acoustic Domain

The background questions regarding voluntary vocal exploratory activities were inspired by Jakobson's views on pre-linguistic infant vocal sounds. Jakobson considered pre-linguistic vocal sounds byproducts of biological functions, and claimed that infants showed random usage of all vocal types in early infancy. This view is not compatible with the frameworks of longitudinal researchers who have noted voluntary and systematic vocal exploratory activities in articulatory/acoustic and temporal domains (Oller, 1980; Stark, 1981). Oller (2000) argues that infants build their own pre-linguistic vocal categories through active vocal exploration activities. The goal of the background question was to assess whether infants build and contrast pre-linguistic vocal categories through exploration activity.

The results of the DA indicated that the two functions composed of five acoustic measures (mean f_0, maximum f_0, minimum f_0, SD of f_0, and duration) explained 41.2 % of the total variance among the three infant vocal categories. Significant segregations of each vocal group by the four subcomponents of f_0 and duration of utterances are visualized in Figure 12. Furthermore, pairwise *F*- tests showed all vocal category groups were different from each other. Large effect sizes for the comparisons between Squeal and the two other categories (Vocant and Growl) suggested that infants strongly utilize the five acoustic variables for distinguishing Squeal from the others. The effect size for the Growl and Vocant comparison was moderate given Cohen's (1988) rules of thumb. This relatively moderate effect size is not surprising considering the following two facts. First, both Vocant and Growl are on the lower end of the f_0 range while Squeal is on the high end, and the auditory perceptual difference between the two categories is not as strong (in listener agreement tests) as the difference between Squeal and the other vocal categories. Secondly, growls are commonly described in terms of other salient vocal characteristics such as harsh or tense vocal quality. These vocal qualities were not included as discriminating acoustic variables in the current study. Inclusion of these vocal qualities as additional variables such as regime could reveal better discrminability for the Growl group from other vocal groups.

Temporal Domain

In the temporal domain, the vocal exploratory activities were assessed in terms of systematic sequential patterns of the three vocal categories. 29 sessions out of 42 (69%) sessions from the seven infants displayed significant vocal repetitions or alternation patterns. Among the 13 sessions that failed to meet a statistical significance level of .05, seven sessions had small Ns, suggesting low power for the statistical analyses.

The adjusted residual values at lag one from LSAs also indicated strong repetition patterns for the same vocal types, demonstrating infants' tendency to repeat pre-linguistic vocal sounds. Goldfield (2000) suggested that the use of acoustically similar sequences in vocalization exploration activities may be due to perception-action coupling during development. In the study, vowels in sequence occupied a more limited space in the F1-F2 space compared to singleton vowels. According to Goldfield (2000), infants listen to their own sound productions within sequences, and then they discover how the particular posture of the articulators modulates sound in particular ways. Likewise, the repetition of the same vocal types found in the current study may also be understood as reflecting the infant' exploratory constraint principle to select a narrower range of the phonatory space within consecutive utterances than within nonconsecutive utterances. A possible interpretation of this tendency is that infants master selected vocal types by purposefully limiting the phonatory space they explore, thus increasing the consistency of targeted vocal productions.

Given the fact that there was no significant serial dependency in individual acoustic variables from the Durbin-Watson autocorrelation tests, the presence of significant serial dependency of the vocal types in the LSAs appears to constitute a contradiction. The infraphonological view by Oller (2000) distinguishes operational units (higher level phonological categories) and prime parameters (dimensions of description for higher units such as amplitude, duration, etc). The infraphonological approach assumes that the operational units are composed of the hidden dimensions and that they may be fundamentally different from the hidden dimensions (p. 8). In the current study, the five acoustic variables are the prime parameters and the three vocal categories are the higher level units. The seemingly contradictory results (no serial dependency on acoustic parameters, but strong repetitive tendencies for vocal categories) appear to correspond with the distinction of higher level categories and the parameters of which they are composed. Only the categories show the repetitive tendency. Individual acoustic parameters do not show the tendency, perhaps because categories are not uniquely defined by any one parameter, but rather constituted particular groupings of multiple parameter values.

The composite of the five acoustic variables explained approximately 44% of the total variance in the DA results. A portion of the unaccounted for 66 % of variance among the vocal types may be explained by other variables such as amplitude and vocal regime type. Consequently there is much further research to be done to help define the vocal categories of early infancy fully in terms of acoustic and physiological parameters.

Main Questions: Vocal Exploratory Activity in Various Engagement Settings and by Age

Engagement Effect

Much of what the infant does vocally is endogenously driven or responsive to constantly changing social circumstance based on his or her unique capabilities, intents and environment. Many recent studies have emphasized the positive effect of social feedback to infants' babbling (Goldstein & Schwade, 2008; Hsu & Fogel, 2001). The main question in this study was to examine the interrelated dynamics of the two factors (internal motivation and external social learning) in pre-linguistic vocal development. One of the primary goals was to evaluate vocal categories in terms of acoustic discriminability and temporal repetition in the three conditions of social engagement. The most contrastive vocal category distinction and vocal repetition patterns were expected in the mutual interaction setting given the assumption that communication is a dynamic process characterized by contingent coordination and mutual regulation by both infants and caregivers (Hsu & Fogel, 2001).

However, the results from the current study did not support the prediction. There was no statistically significant difference in the degree of category distinction and sequential vocal repetitions across the three engagement settings. In fact, the mean degree of the discriminability was largest in the unengaged setting followed by the asymmetrical and symmetrical respectively but the difference between the two engagement settings was very small (symmetrical:46 and asymmetrical:48). Similarly, the mean degree of sequential vocal repetition was largest in the unengaged setting followed by the symmetrical and asymmetrical settings.

A possible explanation for the outcome, contrasting with the explanation suggested by the hypothesis is offered as follows: The unengaged setting is free from social demands, and may provide an optimal environment where infants can focus on playing vocally by manipulating their vocal tract configuration. It has been empirically documented that deaf infants who have limited auditory feedback show significant delay in producing canonical babbling (for review, see Oller, 2000). Clear auditory feedback from vocalizations without intervening sounds from others in the unengaged setting may provide the infant with better opportunities to modulate vocal quality than socially engaged settings.

In contrast, the symmetrical setting is influenced by social demands associated with interactions (Bloom, 1988; Bloom et al., 1987; Papoušek & Papoušek, 1987). Infants regulate speech rhythm and show contingent reciprocal responses in the dynamic mutual interaction settings. This moment to moment social obligation may lead infants to increased attention to the dynamics of the turn taking activities and may lead to less opportunity for vocal exploration.

The degree of the vocal repetitions was smallest in the asymmetrical setting. In this setting, infants' vocalizations were produced while one of the partners did not engage. In most of these cases, caregivers were being interviewed by laboratory staff and infants bid for their attention. Other cases were where infants lost their interest in interaction distracted by other objects. In both cases, the infants' vocalizations may have required a heavy allocation of attention. In the first case, the infant tried to get the caregiver's attention, but she was unresponsive. In the second, the infant wished to interact with an object but was distracted by the caregiver. This attentional confusion may have caused less repetitive and less categorically distinctive vocalizations compared to other contexts. It is worthwhile to recall how the original hypothesis was formulated with respect to an existing literature. The majority of related prior studies dealing with social effects on vocalizations have reported that social feedback facilitates phonological development. The notion of phonological development was operationalized with production of diachronically diverse vocal types (i.e., some vocal types are observed earlier in chronological order and therefore are assumed to be more primitive than others). For example, studies focused on the distinction between either syllabic versus vocalic sounds (Bloom and colleagues, and Hsu and colleagues), or canonical vs. marginal syllables (Goldstein & Schwade, 2008). Vocalic sounds and marginal babbling are more primitive vocal types, occurring earlier in development than syllabic or canonical sounds.

The current work is about the whole first year, and the phonological units that were examined here were the three protophones of Vocant, Growl, and Squeal, which are primarily differentiated based on phonation, a factor that is prominently variable in precursors to speech. The three vocal types are observed across the whole first year, and especially it is clear all three categories occur throughout the Expansion stage (around 4 to 8 months of age). The research here does not then ask the same question that was posed by Goldstein and Schwade or by Hsu and Fogel. The present research is about discriminability and repetitiveness of vocalizations that occur at the same point in development, whereas the prior studies focused on vocalizations emerging at different stages. The results of the prior studies seem to suggest that when infants have reached the canonical stage, they are particularly able to be stimulated in social interaction circumstances to produce canonical sounds.

The results of the present study are not about the degree of advancement of vocal categories that might be elicited in social circumstances, but instead they are about the extent of vocal exploration that is seen in various engagement circumstances. The present work provides the first real empirical evidence of how endogenously driven exploration activities are affected by social settings. In other words the question is about the relationship between the internal and external aspects of vocal development, not about facilitation of diachronically advanced vocal categories in interaction.

The current results seem to suggest that internally driven vocal activities might tend to be more active when social demands are low. The results also may suggest that exploration activity is not a trivial process. Rather, it seems to indicate that vocal play is a cognitively intensive process requiring attentional resources.

Age Effect

Based on the assumption that infants develop their own vocal categories through active exploration and repetitive practice, stronger discriminability and temporal vocal repetition at later ages was expected. The results showed a different pattern. On the one hand, the supplemental analyses by t-tests in the acoustic domain supported the prediction, indicating stronger discriminability of categories at the last age. On the other hand, the strongest vocal category repetition occurred at the earliest age. In contrast, Buder, Oller, and Bakeman (2010, April) reported a general *increase* in vocal category repetitions as infants get older. However, the samples were different and the methods in analyzing data were different from the current study. Especially the criteria for "bout" were not identical. Therefore, additional detailed considerations of methods will be needed to reconcile these contrasting results.

The current results may suggest different stages of exploration activity. Consider an analogy with hand and arm movement. In order to consistently reach a target object, infants practice arm movements, which begin as disorganized. Then as a consequence of such practice, infants show more consistent and narrower arm movements within the trajectory space. Likewise, in order to build consolidated contrastive vocal categories, preparatory vocal practice of the vocal categories may be necessary. The current findings of the strongest degree of vocal repetition at the earliest stage and the largest discriminabilty at the last age seem to reflect a logical sequence. Like the practice of the arm movement to reach a target object, the systematic practice of vocal category to build consolidated vocal categories is deemed to represent the cause, with the effect being vocal category development. At the earliest stage we see high exploration (and thus repetition), and at the latest stage we see consolidated categories (and thus high discriminability). This result provides a significant new perspective on the sequential logic of the two domains in vocal category development.

Subsidiary Questions

The goals of the subsidiary questions were to examine possible systematic association patterns in *the frequency of occurrence of* the three vocal categories in various contexts in early infancy. It should be noted that the detected strong associations between the vocal categories and engagement or proximity do not mean that the categories carry fixed illocutionary intents. Functional flexibility of these protophone categories has been claimed to be a property of the three vocal categories, thus serving as a significant landmark for the development of underpinnings for the human speech and language system (Oller, 2000). The subsidiary questions ask a further question: is there tendency for the three vocal categories to occur differentially in the different engagement or proximity circumstances or at different ages?

Usage of Vocal Categories in Various Engagement Setting

The results indicated that infants produced significantly more vocants when they were unengaged and significantly more squeals when they were mutually interactive. In addition, significantly less use of vocants in mutual interaction and squeals in unengaged interaction settings were also found. The general trend in the usage of growls was in the same direction as in the usage of squeals, showing significantly more use in the mutual interaction setting.

These patterns appear to suggest that infants prefer squeals or growls when their arousal level is high or emotionally active in the mutual interaction setting. When infants are alone, infants are less emotionally provoked. Therefore the need to express emotion is less than in the interaction setting. Thus, infants seem to prefer vocants when they are emotionally neutral. However, in the current study, there was no direct measure of emotionality, expressivity or arousal. Another line of research regarding infant vocalizations and facial affect in the IVOC laboratory is on the verge of examining related issues and results of the research may help clarify the sense in which squeals and growls may be more associated with emotional expression than vocants.

Usage of Vocal Categories by Age

It was found that infants had a strong tendency to produce fewer vocants at the earliest age (more at later ages) and fewer growls at the latest age. Although no statistical significance was detected at a p value of .05, squeals showed the same patterns as growls.

The current results appear to suggest a phase shift in the use of vocal categories by age probably affected by both social learning and internal intents. As Hsu et al. (2000) mentioned, the phase shift may lead to regressions or attenuation in occurrence of certain sounds. In the earlier stages, infants appear to explore all the possible range of the vocal repertoire, resulting in more productions of peripheral elements such as squeals and growls. Once infants consolidate their vocal achievements and learn the social convention of using vowel-like vocal quality in mature speech through interaction with caregivers, regression of squeals and growls may occur. The increase in vocants may be a logical step for infants to advance to a higher level of vocal skills, the canonical syllable.

Usage of the Vocal Categories by Physical Proximity

The recording protocols for the interactive and the separated sessions created significant bias with regard to engagement and proximity; in the interaction session, the caregiver was asked to interact with the infant, whereas in the separate session, the caregiver engaged with a laboratory staff member, who was entering caregiver responses to questions in a log. In separated sessions, the infant was seated or playing away from the caregiver most of the time, which usually resulted in asymmetrical or unengaged engagement circumstances. Therefore, similar results to those found in the engaged circumstances on the use of vocal categories were expected in the immediate proximity circumstance.

Consistent with the expectation, infants produced significantly more growls and squeals when they were close to caregivers and more vocants when they were distant from caregivers. It was also shown that the infants used significantly fewer vocants when they were close to caregivers and growls and squeals when they were distant from the caregivers. However, the degree of association was different. In terms of the proximity condition, the association between Growl and Immediate was statistically significant, but Squeal and Immediate was not. In contrast, in the engagement domain, the association between Squeal and Symmetrical was statistically significant, but not Growl and Symmetrical. Therefore, the results regarding proximity add a new perspective specifically on the usage of vocal types in varying proximity conditions. In general, growls can be characterized by lower amplitude as well as creaky voice quality and tenseness. The lower amplitude is inevitable in part due to the physical dynamics to exert tensed vocal quality. Therefore, it is relatively hard to recognize growls if the communicators are located apart. At least the strong tendency of using growls in the immediate context seems to indicate infants' capability to modulate amplitude properties of the vocal categories in varying proximal conditions.

Effects of Age, Engagement Setting and Vocal Categories on Utterance Duration

In a sense, the duration effect can be considered as a subcomponent of the main question in the acoustic domain. The discriminant function analyses indicated that the vocal categories were well distinguished in terms of the composite of the five acoustic variables. The standardized coefficients indicated that duration was the least or second to the least significant contributor for the discriminant functions. Therefore, the results from the DAs can be used as a close approximation of the major role of f_0, although the exact discriminant power with the four subcomponents of f_0 was lower than the reported results from the five composite variables. Thus the results of utterance duration effects in various settings allowed determining the unique role of duration (accounting for significant unique variance) in the development of the three vocal categories.

Age by Vocal Type

Results indicated that there was a significant duration difference between the vocal categories at all ages, showing increase in duration from Age 1 to Age 3. The result indicated that the range of utterance duration (.6 to 1.05 seconds) and the increase by age were comparable to the range of .5 to 1.34 reported in prior work (Bloom, 1988, 1989; Oller 1980).

The current study provides further details in utterance durations by each vocal type. The Squeal was significantly longer followed by Growl and Vocant, respectively. This result is compatible with results of Magoon (2003) who reported longest duration for squeals. Again, these results seem to support perspectives on infants' voluntary efforts on mastering more exotic sounds such as squeals and growls by elongating their durations, thus providing intensified auditory feedback and articualtory motor experience. Although Papaeliou, Minadakis, and Cavouras (2002) did not include squeals and growls in their coding scheme, they coded emotional status of each vocalization and reported fairly similar results. Their findings indicated that vocalizations that expressed emotions showed higher f_0 values (peak, final, SD) and longer durations than vocalizations that did not convey emotions.

Age by Engagement Setting

The duration of vocalizations at each age was dependent on the settings of engagement (Refer to the interaction plot of Figure 13). All the age comparisons showed the same pattern of significant duration differences in the three engagement settings. The current study showed consistently and significantly longer durations in the symmetrical than in asymmetrical settings.

The duration of utterances in the unengaged setting varied from age to age, starting with longest duration values at the first age, shortest duration values at the middle age, and medium duration values at the latest age. It is hard to reason about this variation in duration without further analysis. Perhaps infants' vocalizations in unengaged settings are subject to more moment to moment variability based on endogenous variability.

In summary, although the role in discriminating the three vocal categories was smaller compared to the four f_0 measures, duration alone was still a significant variable in characterizing the categories, demonstrating dynamic interactions between age and engagement.

Conclusions

The results of the current study indicated that infants build their vocal categories, producing the three vocal categories distinctively in the acoustic domains of f_0 and duration and showing systematic vocal repetition or alternation patterns among the categories. These results provide empirical evidence against Jakobson's view that pre-linguistic vocal sounds are mere byproducts of biological functions, and presumably not affected by socialization.

The trend toward stronger vocal exploratory activities (both in the acoustic and temporal domains) in the unengaged setting suggests that vocal exploration activity may be a cognitively intensive process which requires dedicated self-monitoring of vocal capabilities through clear auditory feedback. Social interaction with others seems to lead infants to focus on the social interaction such as the dynamics of the turn-taking activities and leads to less opportunity for vocal exploration activities.

The results regarding age effects on vocal exploration provided a new perspective on how the two concepts of discriminability and systematic repetition of the vocal categories may be sequentially related. The results imply that in order to build consolidated contrastive vocal categories, antecedent vocal practice is necessary. These ideas were reflected in the strongest repetition patterns at the earliest age and the strongest degree of distinctiveness of the vocal categories at the latest age.

Infants also showed a preference for squeals and growls over vocants when they were mutually interactive and physically close to caregivers. This pattern suggests that squeals and growls are used significantly more when infants are emotionally aroused. In addition, the squeals and growls were used more when infants were younger whereas vocants were used more when infants were older. This result suggests a phase shift in the use of the vocal categories by age probably affected by both social learning and internal intents. The decrease in use of squeals and growls seems to be the natural logical step for infants to advance to master higher level vocal skills.

The role of duration in discriminating the three vocal categories was still significant in the three engagement settings although the effect size was smaller compared to the four f_0 measures. Consistently longer durations in the symmetrical than asymmetrical settings were found whereas fluctuating durations of utterances were observed in unengaged setting, suggesting moment to moment variability in infants' endogenous intents in the unengaged settings.

In summary, this research supports the idea that infants actively seek to build their own vocal categories based on their vocal capabilities. Evidence that they build consolidated vocal categories through vocal practice was manifested by strongest vocal repetition at the earliest age and strongest discriminability at the last age. The degree of exploration activity was not statistically different across the three engagement settings but showed a trend of strongest degree when social demands were low, suggesting that exploration activity is an endogenously driven intensive cognitive process. This research also showed infants' systematic tendency to utilize the vocal categories differentially in various social settings early on. In conclusion, all results suggest infants' voluntary control efforts on their pre-linguistic vocal productions to respond to constantly changing social circumstances, indicating that pre-linguistic vocal productions are the products of systematic and intentional vocal exploration, not merely the byproducts of biological maturation.

Further Directions

The goal of the current study was to understand the relationship between vocal category development and exploration activities in varying social communication circumstances, using developmentally appropriate methodology. Exploration activities were evaluated in terms of distinctiveness of vocal categories by the five acoustic variables in the acoustic domain and by vocal practice manifested in repetition in the temporal domain. Future research could expand upon the current study by considering the following suggestions.

First, inclusion of more acoustic variables, such as vocal regime, phonation quality and amplitude could help us in understanding the nature of the three vocal categories, especially the growl. While inspecting the data, strong patterns in utilizing amplitude were noted: high amplitude for squeals and low amplitude for growls in general. In addition, acoustic correlates of vocal regimes, which are the basic vibratory patterns of vocal tissue, has been suggested as a potential variable that might have strong associations with vocal categories (Buder et al., 2008). Phonation quality includes vocal quality such as harshness, tenseness, breathiness etc. Inclusion of this acoustic correlates of these variable should specifically enhance the discriminability of the growl group from others since the growl is often characterized by tense vocal quality.

Secondly, it may be useful to reanalyze data, excluding utterances below some minimum of duration such as 100 ms or with very low amplitude. This could enhance the reliability of f_0 measures and vocal coding, revealing stronger trends. In the current study, utterances longer than 50ms and any audible utterances were included if they were not reflexive or vegetative sounds. When utterances were short and soft, automatically produced f_0 traces were more difficult to determine due to confusion with environmental sounds. In addition, it was hard to judge vocal categories when utterances had short durations. Given the fact that a normal syllable length in speech is typically not less than 120 ms, comparing the results after exclusion of utterances shorter than 200 ms seems to be a valid step to enhance the ecological validity of coding.

Third, inclusion of diachronically/developmentally different vocal categories such as quasi-vowels, marginal babbles and canonical babbles will allow tests of whether social interaction facilitates phonological development. Prior work by Bloom and colleagues, Hsu and colleagues, and Goldstein and colleagues reported positive relationships between mutual interaction and phonological development.

Fourth, more data collected in natural settings could show more precise dynamics of vocal development. The current data were collected from a laboratory setting which tried to simulate a home environment as closely as possible. However, the representativeness of the samples is questionable. Sometimes, infants were shy and quiet or mothers were overactive due to the rolling cameras. All day vocal recording and automated analysis technology is now available, and could help supplement longitudinal laboratory recordings. The patterns of vocalization in infants have been observed to be changeable from moment to moment. Use of an all day recording device may reveal more clear and exact dynamics of the vocal development by capturing changes that might have been missed in laboratory sampling. Based on data from all day recordings, cross recurrence quantification analysis, an extension of lag sequential analysis, will allow assessing possible nonlinear patterns of vocal development in infancy, and may help clarify differences between overall repetition patterns and patterns of local clustering of particular vocal types in time (Dale, Warlaumont, and Richardson, in press).

Fifth, further analysis on individual differences in vocal category development may add important details. The current study did not emphasize individual differences in utilization of the acoustic variables, vocal categories repetition, or usage of the vocal categories. Prior work (Kwon et al., 2006, 2009) reported significant individual differences in usage of vocal types from session to session within age. Significant shifts were expected in individual exploratory inclinations, expressive interests or in patterns of social stimulation occurring within each session in each infant. Again, cross recurrence quantification analysis will allow assessing nonlinear dynamics of such shifts in vocal behaviors in each infant.

Lastly, future research on the vocalizations of the children with disorders will help to determine early vocal characteristics of disordered populations. Among populations with disorders, children with autism are well known to have anomalies in communication skills. Recently, significant efforts to find early red flags for autism have been made. Compared to research on social skills, research on early vocal characteristics has been very limited (Ornitz & Ritvo, 1976; Sheinkopf, Mundy, Oller, & Steffens, 2000; Simmons & Baltaxe, 1975). Application of the current methodology to at-risk infants or children with autism could reveal differences in the development of the categories as well as the usage of the categories in various engagement settings.

REFERENCES

Anderson, A. H., Bard, E. G., Sotillo, C., Newlands, A., & Doherty-Sneddon, G. (1997). Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception & Psychophysics, 59*, 580-592.

Anderson, B. J., Vietze, P., & Dokecki, P. R. (1977). Reciprocity in vocal interactions of mothers and infants. *Child Development, 48*, 1676-1681.

Anderson, B. J., Vietze, P., & Dokecki, P. R. (1978). Interpersonal distance and vocal behavior in the mother-infant dyad. *Infant behavior & Development, 1*, 381-391.

Bakeman, R., & Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS and GSEQ*. Cambridge, UK: Cambridge University Press.

Bauer, H. R., & Kent, R.D. (1987). Acoustic analysis of infant fricative and trill vocalizations. *Journal of the Acoustical Society of America, 81*, 505-511.

Bauer, R. B. (1988). Vowel intrinsic pitch in infants. *Folia phoniatrica, 40*, 138-146.

Bloom, K. (1975). Social elicitation of infant vocal behavior. *Journal of Experimental Child Psychology*, *20*, 51-58.

Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language*, *15*, 469-480.

Bloom, K. (1989). Duration of Early Vocal Sounds. *Infant behavior & Development*, *12,* 245-250.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, *15*, 211-227.

Boysson-Bardies, D. (1999). *How Language Comes to Children: From Birth to Two Years* Cambridge, MA: MIT Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory Phonology: An overview. *Phonetica*, *49*, 155-180.

Buder, E. H., Chorna, L., Oller, D. K., & Robinson, R. (2008). Vibratory Regime Classification of Infant Phonation. *Journal of Voice*, *22*, 553-564.

Buder, E.H., Oller, D.K., & Bakeman, R. (2010, April). Vocal repetition increases during the first year of development. Paper presented at the International Child Phonology Conference, Memphis, TN.

Butterworth, B., Hine, R. R., & Brady, K. D. (1977). Speech and interaction in sound-only communication channels. *Semiotica, 20*, 81-99.

Cohn, J. F., & Tronick, E. Z. (1987). Mother-infant face-to-face interaction: The sequence of dyadic states at 3,6, and 9 months. *Developmental Psychology, 23*, 68-77.

Crnic, K. A., Ragozin, A. S., Greenberg, M. T., Robinson, N. M., & Basham, R. B. (1983). Social interaction and developmental competence of preterm and full-term infants during the first year of life. *Child Development*, *54*, 1199-1210.

Cruttenden, A. (1970). A phonetic study of babbling. *British Journal of Disorders of Communication*, *5*, 110-118.

Crystal, T. H., & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America, 88*(1), 101-112.

Cucchiarini, C. (1996). Assessing transcription agreement: Methodological aspects. *Clinical Linguistics & Phonetics, 10* (2), 131-155.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (accepted). Nominalcross recurrence as a generalized lag sequential analysis for behavioral streams. *International Journal of Bifurcation and Chaos.* (special issue on recurrence)

Dunn, C., & Davis, B. L. (2008). Phonological process occurrence in phonologically disordered children. In *Applied Psycholinguistics* (Vol. 4, pp. 187-207). New York, U.S.A. Cambridge Univ Press.

Durbin, J & Watson, G.S. (1950). Testing for Serial Correlation in Least Squares Regression I. *Biometrika* 37, 409–428

Eilers, R. E., Neal, A. R., & Oller, D. K. (1996). *Late onset babbling as an early marker of abnormal development* (Poster presentation). Providence, RI: International Conference on Infant Studies.

Eilers, R. E., & Oller, D. K. (1974). *Babbling in a hard-of-hearing, a Down's syndrome, and a normal child: A comparative study*: Paper presented at the Western Conference on Linguistics.

Eilers, R. E., & Oller, D. K. (1994). Infant vocalizations and the early diagnosis of severe hearing impairment. *Journal of Pediatrics, 124*, 199-203.

Ferguson, C. A. (1986). Discovering sound units and constructing sound systems: It's child's play. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and Variability in Speech Process* (pp. 36–51). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fleiss, J. L. (1981) *Statistical methods for rates and proportions*. 2nd ed. New York: John Wiley .

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Anuual Review of Psychology, 39*(1), 1-42.

Gillis, S., Schauwers, K., & Govaerts, P. (2003, August). *Ontogeny and phylogeny of syllable organization*. Barcelona, Spain: International Conference to Provide a Festschrift for Peter MacNeilage.

Goldfield, E. C. (2000). Exploration of vocal tract properties during serial production of vowels by full term and preterm infants. *Infant behavior & Development, 23*(3-4), 421-439.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Science*, *100*(13), 8030.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515.

Grégoire, A. (1948). L'apprentissage du langage. Originally published in *Lingua*, 1.162-164, 168-169, 170-172 1948) (Leopold?, Trans.). In A. Bar-Adon & W. F. Leopold (Eds.), *Child language: A book of readings, published in 1971*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Guitar, B., & Marchinkoski, L. (2001). Influence of mothers' slower speech on their children's speech rate. *Journal of Speech, Language, and Hearing Research, 44*(4), 853-861.

Henderson, A. B. (1998). The development of pitch in early vocalizations: A longitudinal study of infants from two language communities. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 58*(12-B), 6532.

Hsu, H. C., & Fogel, A. (2001). Infant vocal development in a dynamic mother-infant communication system. *Infancy, 2*(1), 87-109.

Hsu, H. C., Fogel, A., & Cooper, R. B. (2000). Infant vocal development during the first 6 months: Speech quality and melodic complexity. *Infant & Child Development, 9,* 1-16.

Kahane, J. C., & Kahn, A. R. (1984). Weight measurements of infant and adult intrinsic laryngeal muscles. *Folia phoniatrica, 36*(3), 129-133.

Kent, R. D., Mitchell, P. R., & Sancier, M. (1991). Evidence and role of rhythmic organization in early vocal development in human infants. In J. Fagard & P. H. Wolff (Eds.), *The development of timing control and temporal organization in coordinated action*: Elsevier Science Publishers.

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 37-50). New York: Stockton Press.

Kwon, K., Buder, E. H., Oller, D. K., & Chorna, L. B. (2006). *The role of fundamental frequency characteristics in classifying prebabbling vocalizations*. Miami, FL: Annual meeting of the American Speech-Language-Hearing Association.

Kwon, K., Oller, D. K., & Buder, E. H. (2006). *Contextual flexibility in precanonical infant vocalizations:Its role as a foundation for speech*. Paper presented at the International Child Phonology Conference.

Kwon, K., Oller, D. K., & Buder E, H. (2009). *Systematic Vocal Repetition in Infant Vocalization.* Paper presented at the American Speech-Language-Hearing Association.

Kwon, K., Oller, D. K., Buder, E. H., Cleary, J., Sheinkopf, S. J., & Mundy, P. (2008). *Pitch (Fundamental Frequency, i.e. f0) Development of Prelinguistic Vocalizations in Children with Autism*. Paper presented at the ASHA

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Laufer, M. Z., & Horii, Y. (1977). Fundamental frequency characteristics of infant non-distress vocalizations during the first twenty-four weeks. *Journal of Child Language, 4*, 171-184.

Lindblom, B. (1990). On the communication process: Speaker-listener interaction and the development of speech. *Augmentative and Alternative Communication*, 220-230.

Magoon, J. C. (2003). *Phonatory Control in the Developing Infant Protophone.* Unpublished master's thesis, University of Memphis, Memphis, USA.

Masataka, N. (1993). Relation between pitch contour of prelinguistic vocalizations and communicative functions in Japanese infants. *Infant Behavior and Development, 16*, 397-401.

Masataka, N., & Bloom, K. (1994). Acoustic properties that determine adults' preferences for 3-month-old infant vocalizations. *Infant Behavior & Development, 17*(4), 461-464.

Mohanan, K. P. (1992). Emergence of complexity in phonological development. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 635-662). Timonium, MD: York Press.

Morris, M. E.(2005). *A Preliminary Methodology for the Study of Infant Vocalizations in Contrasting Social Contexts.* Unpublished master's thesis, University of Memphis, Memphis, USA.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G.Yeni-Komshian, J. Kavanagh & C. Ferguson (Eds.), *Child phonology, Volume 1, Production* (pp. 93-112). New York: Academic Press.

Oller, D. K. (1981). Infant vocalizations: Exploration and reflexivity. In R.E.Stark (Ed.), *Language behavior in infancy and early childhood* (pp. 85-104). New York: Elsevier North Holland.

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum and Associates.

Oller, D. K., & Bull, D. (1984). *Vocalizations of deaf infants* (Poster Session). New York, NY: International Conference on Infant Studies.

Oller, D. K., Eilers, R. E., Urbano, R., & Cobo-Lewis, A. B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of Child Language, 27*, 407-425.

Oller, D. K., & Smith, B. (1977). The effect of final syllable position on vowel duration in infant babbling. *Journal of the Acoustical Society of America, 62*, 994 997.

Oller, D. K., Wieman, L., Doyle, W., & Ross, C. (1975). Infant babbling and speech. *Journal of Child Language, 3*, 1-11.

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association, 69*(348), 894-908.

Ornitz, E., & Ritvo, E. (1976). Medical assessment. In E. Ritvo (Ed.), *Autism, Diagnosis, current research, and Management* (pp. 7–26). New York: Spectrum Publications.

Papaeliou, C., Minadakis, G., & Cavouras, D. (2002). Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *Journal of Speech, Language and Hearing Research, 45*(2), 311-317.

Papoušek, H., & Papoušek, M. (1979). The infant's fundamental adaptive response system in social interaction. In E. B. Thoman (Ed.), *Origins of the infant's social responsiveness* (pp. 175-208). Hillsdale, NJ: Erlbaum Associates.

Papoušek, H., & Papoušek, M. (1987). Intuitive parenting: a dialectic counterpart to the infant's integrative competence. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 669-720). New York: Wiley.

Papoušek, M., & Hwang, S. (1991). Tone and intonation in Mandarin baby talk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics, 12*, 481-504.

Papoušek, M., & Papoušek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language, 9*, 1989.

Robb, M. P., & Saxman, J. H. (1988). Acoustic observations in young children's non-cry vocalizations. *Journal of the Acoustical Society of America, 83*, 1876-1882.

Roug, L., Landberg, I., & Lundberg, L.-J. (1989). Phonetic development in early infancy: a study of four Swedish children during the first eighteenth months of life. *Journal of Child Language, 16*, 19-40.

Scherer, K. (1992). Vocal affect expression as symptom, symbol and appeal. In H.Papoušek, U.Jürgens & M.Papoušek (Eds.), *Nonverbal vocal communication* (pp. 43-60). New York: Cambridge University Press.

Scherer, K. R. (1985). Vocal affect signaling: a comparative approach. *Advances in the Study of Behavior, 15*, 189-244.

Sheinkopf, S. J., Mundy, P., Oller, D. K., & Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders, 30*(4), 345-353.

Simmons, J. Q., & Baltaxe, C. (1975). Language patterns of adolescent autistics. *Journal of Autism & Childhood Schizophrenia, 5*(4), 333-351.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts, Inc.

Staats, A. W. (1972). Language behavior therapy: A derivative of social behaviorism. *Behavior Therapy, 3*(2), 165-192.

Stark, R. (1981). Infant vocalization: A comprehensive view. *Infant Medical Health Journal, 2*(2), 118-128.

Stark, R. E. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology, vol. 1* (pp. 73-90). New York: Academic Press.

Stevens, J. P. (2007). *Intermediate statistics: A modern approach*: Lawrence Erlbaum.

Stoel-Gammon, C. (1989). Prespeech and early speech development of two late talkers. *First Language, 9*, 207-224.

Stoel-Gammon, C., & Cooper, J. (1984). Patterns of early lexical and phonological development. *Journal of Child Language, 11*, 247-271.

Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist, 50*(2), 79-95.

Tucker, J. A., & Tucker, G. F. (1979). A clinical perspective on the development and anatomical aspects of the infant larynx and trachea In G.B.Healy & T.J.I.McGill (Eds.), *Laryngo-tracheal problems in the pediatric patient.* Springfield, IL: Charles C Thomas.

Wallace, V., Menn, L., & Yoshinaga-Itano, C. (1998). Is babble the gateway to speech for all children? A longitudinal study of children who are deaf or hard of hearing. *Volta Review, 100*(5), 121-148.

Whalen, D. H., Levitt, A. G., Hsiao, P., & Somoodinsky, I. (1995). Intrinsic F0 of vowels in the babbling of 6-, 9-, and 12-month-old French- and English-learning infants *Journal of Acoustic Society of America, 97* (4), 2533-2539.

Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics, 102*, 1161-1171.

Zlatin, M. (1975). *Preliminary descriptive model of infant vocalization during the first 24 weeks: Primitive syllabification and phonetic exploratory behavior* (Final Report, Project No 3-4014, NE-G-00-3-0077): National Institutes of Health Research Grant.