Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes maybe granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

Classification of Infant Vocalizations using Untrained Listener Judgments

by

Candice J. Perry

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Communication Sciences and Disorders

Idaho State University

Summer 2016

Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Candice Perry find it satisfactory and recommend that it be accepted.

_____

Heather Ramsdell-Hudock

Major Advisor

_____

Kristina M. Blaiser

Committee Member

_____

Nicki Aubuchon-Endsley

Graduate Faculty Representative

Idaho State
UNIVERSITY

Office for Research Integrity
921 South 8th Avenue, Stop 8046 • Pocatello, Idaho 83209-8046

January 4, 2016

Candice Perry
Comm Sci Disorders/Deaf Educ
MS 8116

RE: regarding study number IRB-FY2016-124: Classification of Infant Vocalizations using Untrained Listener Judgments

Dear Ms. Perry:

I agree that this study qualifies as exempt from review under the following guideline: Category 4: Collection or study of existing data. This letter is your approval, please, keep this document in a safe place.

Notify the HSC of any adverse events. Serious, unexpected adverse events must be reported in writing within 10 business days.

You are granted permission to conduct your study effective immediately. The study is not subject to renewal.

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. Contact Tom Bailey (208-282-2179; fax 208-282-4723; email: humsubj@isu.edu) if you have any questions or require further information.

Sincerely,

Ralph Baergen, PhD, MPH, CIF
Human Subjects Chair

Acknowledgements

I would like to thank Dr. Heather L. Ramsdell-Hudock my thesis advisor for all of her encouragement, knowledge, and support in writing this thesis. I would also like to thank my committee members M. Kristina Blaiser, and Nicki Aubuchon-Endsley for their contribution. The biggest acknowledgement and thank you goes to my husband Trae Perry, my children Kyler and Jaden, and family and friends who believed in me and helped to support and encourage me through this process.

TABLE OF CONTENTS

List of Figures viii

List of Tables ix

Abstract x

Chapter I: Literature Review 1

Introduction 1

Chapter II: Classification of Vocal Development 3

Chapter III: Caregiver Report 9

Chapter IV: Early Identification for Infants 12

Chapter V: Purpose 15

Infants 16

Untrained Listeners 18

Procedure and Analysis 18

Chapter VI: Results 21

Chapter VII: Discussion 31

Study limitations 34

Clinical Implications 35

Future Directions 36

References 37

Appendix A 44

List of Figures

Figure 1: Tag cloud for listener response labels to vowel utterances 23
Figure 2: Tag cloud for listener response labels to squeal utterances.....24
Figure 3: Tag cloud for listener response labels to growl utterances 25
Figure 4: Tag cloud for listener response labels to raspberry utterances.....26
Figure 5: Tag cloud for listener response labels to whisper utterances 27
Figure 6: Tag cloud for listener response labels to cry utterances.....29
Figure 7: Tag cloud for listener response labels to laugh utterances.....30

List of Tables

Table 1: Number of Infant Utterance Types Presented to Untrained Listeners.....21
Table 2: Proportion of Occurrence for each Response Type for Vowel Utterances 23
Table 3: Proportion of Occurrence for each Response Type for Squeal Utterances 24
Table 4: Proportion of Occurrence for each Response Type for Growl Utterances.....25
Table 5: *Propo*rtion of Occurrence for each Response Type for Raspberry Utterances ...27
Table 6: Proportion of Occurrence for each Response Type for Whisper Utterances.....28
Table 7: Proportion of Occurrence for each Response Type for Cry Utterances 29
Table 8: Proportion of Occurrence for each Response Type for Laugh Utterances.....30

Classification of Infant Vocalizations using Untrained Listener Judgments

Thesis Abstract—Idaho State University (2016)

Purpose: Vocal communication follows a developmental sequence, however there is no acknowledged "gold standard" terminology to refer to infant vocalizations. The *purpose of this* research was to begin developing a gold standard for describing infant vocalizations by incorporating caregiver report. Methods: Untrained listeners were selected to report on vocalizations obtained from archived video/audio recordings infants from 4 through 12 months of age. Results: Listener responses fell into several distinct categories, some of which overlapped with published categorical descriptors, and some of which yielded alternate and additional terminology that may provide more complete coverage of the range of sounds infants produce. Conclusions: Outcomes from this and future studies could help to simplify the process of identifying infants who may be at risk by reinforcing parents and researchers (trained and untrained listeners) to use and understand the same terminology. Clinical implications, study limitations, and future directions will be discussed.

Classification of Infant Vocalizations using Untrained Listener Judgments

Chapter I- Literature Review

Introduction

There are multiple ways researchers describe infant vocalizations, however the terminology commonly used in the research lab and educational setting may not be the same as that used by caregivers. According to Stark, Bernstein, and Dermorest (1993), vocal communication follows a predicted and orderly developmental sequence in typically developing infants in the first 18 months of life. Although typically developing infants produce vocalizations frequently, especially as their age increases, there is no identified "gold standard" terminology that all researchers refer to and use consistently. Traditionally, researchers have used 3 terms to broadly define the development of speech: perlocutionary, illocutionary, and locutionary. Another common means of classifying infant vocalizations in more detail is through the prelinguistic stage model. The stages include the phonation (0-2 months), cooing (1-4 months), expansion (3-8 months), canonical babbling (5-10 months), and variegated babbling (11-14 months) stages. The sounds produced in these stages are not mutually exclusive, but show a great deal of variation in the frequency and quality within and between stages (Oller, 1986, 2000; Oller et al., 2013; Oller & Griebel, 2008).

For the past 20 years, several researchers and studies have turned to parent reports and found that these statements serve as an efficient technique for assessing and characterizing early language in children under age 3 (Dale, 1991; Rescorla & Alley 2001). Perhaps a good reason for this is that the prelinguistic infant's babbling of syllables is noticed by caregivers (Oller, Eiler, & Basinger, 2001; Ramsdell et al., 2012).

Some researchers describe that "parent reports of language and communication are an appealing option because parents have extensive experience with their children under a wide variety of naturalistic situations" (Feldman et al., 2005, p. 856).

A *gold standard* for describing infant vocalizations, given that different researchers use different terminology, would be beneficial in bridging the gap between research and clinical practice. If a gold standard were developed in accord with what caregivers report, clinicians could easily ask caregivers about prelinguistic vocal development and receive responses that could inform practice. As it stands, it is difficult to categorize infant vocalizations for many reasons, among which we could include classification confusion given multiple ways to categorize stages of development.

Chapter II- Classification of Vocal Development

One way to categorize stages of development that has been used in the past is through a broad classification scheme of speech, including the perlocutionary, illocutionary, and locutionary stages of development (Bates, Camaioni, & Volterra, 1975).

Perlocutionary communication refers to infants reacting innately to certain internal physiological states by producing certain behaviors (e.g., smiling or crying), which inevitably result in the satisfaction of material needs or maintenance of social interaction. The first step in developing communicative intent is getting an adult response to an infant signal, and therefore establishing a circular means-end relationship (Bates, Camaioni, & Volterra, 1975). Furthermore, Bates and colleagues (1993) reported that children start using intentional communication after the first 4 months of life. Perlocutionary acts are communicative acts that have some effect on the listener without having control over said effect (e.g., the hunger cry of a newborn infant resulting in a parent feeding him) (Bates, Camaioni, & Volterra, 1975). Until intentional communication is used by the infant, infants engage only in perlocutionary communication with adults. The perlocutionary stage occurs between approximately 0 and 9 months of age (Stark, Bernstein, & Dermorest, 1993).

An illocutionary speech act is a conventional speech act, or speech that is understood and recognized by both the speaker and the listener (Bates, Camaioni, & Volterra, 1975). An illocutionary act requires the intentional use of a conventional signal (e.g., pointing) to carry out some socially recognized function (such as commanding or indicating the presence of objects or events). The illocutionary stage occurs when

children intentionally use non-verbal signals to convey requests and to direct adult attention to objects and events. “An illocution is the conventional social act of ordering, advising, urging, etc.” (Bates, Camaioni, & Volterra, 1975, pg. 206). The illocutionary stage is from 10 to 12 months of age, during which the infant expresses intentional meanings through primarily prelinguistic vocalizations and gestures (Stark, Bernstein, & Dermorest, 1993).

The locutionary stage occurs when children use speech sounds to communicate their intentions, intentions that were previously expressed through prelinguistic vocalizations or nonverbally (e.g., using the word “no”) (Bates, Camaioni, & Volterra, 1975). This stage emerges around 13 months of age and older (Stark, Bernstein, & Dermorest, 1993). Many aspects of language skill overlap in development, such as body orientation, gesture, vocalization, level of intent, and age of onset (Rossetti, 1990). Therefore, it is difficult to apply stringent developmental milestones at specific ages, which is likely one reason developmental milestones are presented in age ranges, to account for individual variation (Bates, Camaioni, & Volterra, 1975).

The discrepancies brought to light in the proposed project are not primarily concerned with age of development. Rather, the project is directed toward the lack of consistency regarding infant vocalizations. According to Oller (2000), another way that many researchers have described early infant vocalizations is through the prelinguistic stage model. The first stage in this model is the *phonation stage*. During the first 2 months of life, typically developing infants produce nonreflexive and nondistress sounds. Most of these vocalizations fit into a category referred to as *quasi-resonant nuclei* (QRN), inherently not well-formed vowel-like sounds, often produced with the vocal

tract at rest, sounding fuzzy and imprecise. QRN do not include vocal fry, breathy voice, or similar quality vocalizations. Instead of using QRN, other researchers, such as Landberg (1989), use terms such as *vowels*, *syllabic consonants*, and *small throaty sounds*. "Auditory phonetic analysis suggests that the QRN ranges from a syllabic nasal to a high, mid, unrounded, nasalized vowel. A small proportion of fully resonant nuclei (FRN, often referred to as vowels or coos) occur along with QRN in nonreflexive vocalization during the first 2 months of life" (Oller, 2000, p. 95).

The *goo* stage is the second stage Oller (2000) discusses. This stage occurs around 2 to 3 months of age, and is shown to consist of primitive syllabification. Goos add a level of complexity to QRN through incorporation of a consonant-like element; they are considered primitive in the development of speech because of irregular transition timing between consonant and vowel parts when compared to canonical babbling. Goos can also be referred to as marginal syllables. At this stage, infants are not considered to have reached full phonetic competence (Oller, 2000).

Expansion is the third prelinguistic stage that occurs around the age of 3 to 8 months. Stark, Bernstein, and Dermorest (1993) refer to this stage as vocal play and Zlatin (1975) calls it exploratory phonetic behavior. Expansion utterances can consist of multiple sound types or combinations. The sound types include, but are not limited to raspberry vocalizations, squealing, growling, yelling, whispering, and marginal babbling. Further, it is during the expansion stage that FRN begin to be produced, or more articulate vowel productions. Since the infant can use multiple combinations of sound types, this stage is represented by an abundance of variation in sounds produced (Oller, 2000).

Canonical babbling or reduplicated babbling is the fourth prelinguistic stage and occurs around months 5 to 10. Oller explains that "the relatively rigid timing characteristics of syllabification in natural languages are introduced into the child's vocalization system at the onset of canonical babbling" (2000, p. 98). Some say that the onset of canonical babbling starts with reduplicated syllables, for example "mama," "dada," and "baba." Specifically, canonical syllable shapes contain clearly articulated vowel and consonant segments, with timely transitions between the two. Given these characteristics, canonical syllables are the first well-formed productions, matching syllables in the ambient language. Accordingly, it is during this stage that parents begin to report that their infants are speaking, even though there may not be intent behind the canonical syllables produced.

Variegated babbling, also referred to as conversational babbling or jargon, is when infants produce differing multisyllabic utterances with well-formed consonantal and vocalic characteristics such as "madaga budi," with adult like intonation (Paul, 2007). Vocalizations in this prelinguistic stage are sometimes referred to as gibberish and they emerge around 11 to 14 months of age (Oller, 2000).

Beyond the prelinguistic stage model, another alternative classification scheme for describing vocal development was used by Oller, Wetherby, and Prizant (2003). They described sounds using terms like spontaneous vocalizations, grunts, Bronx cheers, different cries (e.g., hungry, wet, tired, hurt, wants a toy, wants attention, and cries when lonely), sounds at the back of the throat (e.g., "k" and "g"), intonation patterns, screams, whines, laughs, babbles, sound play, imitates, sings, and words.

According to Nathani, Ertmer, and Stark (2006), the Stark Assessment of Early Vocal Development-Revised (SAEVD-R) describes infant vocalizations in a similar manner, with some alternate terminology. Level 1 (reflexive, from 0 to 2 months of age), primarily contains infant fussing, crying, and vegetative sounds (such as burps, coughs, or sneezes). In addition to non-speech sounds, protophones such as QRN are included. Level 2 (control of phonation, from 1 to 4 months of age) contains infant sounds that are no longer reflexive, and not produced in response to distress or need. Instead, the infant is voluntarily controlling some of the vocal tract. Consonant sounds, FRN, primitive closant-vocant combinations, chuckle or laughter, and occasionally glottal stops are produced. Level 3 (expansion, from 3 to 8 months of age) is primarily made up of protophones; infants produce mature vowels with full resonance that may be able to be transcribed as adult vowels. This is also the phase where infants start to experiment with their voices (e.g., squeals and vowel glides), and produce marginal babbling, containing consonant and vowel like elements that occur more slowly than mature syllables. Level 4 (basic canonical syllables, from 5 to 10 months of age) contains protophones recognized as canonical syllable forms by even naïve listeners. Level 5 (advanced forms, from 9 to 18 months of age) either precedes or coincides with first words (Nathani et al., 2006). Vocal production changes from vowel-like sounds, to closant-vocant combinations, to mature adult-like syllables, to mature adult language (Oller, 2000). Beyond the SAEVD-R model of development, Stark (1980) also provided a vocal model which included 5 levels: reflexive sounds (0 to 2 months), cooing and laughter (2 to 4 months), vocal play (4 to 8 months), reduplicated babbling (8 to 10 months), and nonreduplicated babbling and first words (10 to 14 months).

Prelinguistic vocalizations can be broadly classified into non-speech like vocalizations, and speech-like vocalizations, or protophones (Oller, 2000). Non-speech like vocalizations have fixed links to biological functions and are broken down into two categories: fixed signals (crying and laughter) and vegetative sounds (burps and hiccups). Protophones however, are precursors to speech and vary along a continuum, becoming more speech-like toward the end of the first year (Oller, 2000).

While all of the presented models are generally accepted as standard benchmarks for evaluating progress, and are similar in ages of emergence, numbers of levels, and vocalization types, there are differences in terminology and definitions across all of them. "These differences across laboratories, often unspecified, hamper the portability of vocalization types and make comparisons across studies difficult" (Nathani, Ertmer, & Stark, 2006, p. 2-3). Discrepancies in defining infant vocalizations create confusion. Given that different researchers use different terminology, creating a *gold standard* would be beneficial in bridging the gap between research and clinical practice. Again, caregiver report of prelinguistic vocalizations may aid in translation between research and clinical practice for this purpose.

Chapter III- Caregiver Report

"Munson, Edwards, Schellinger, Beckman, and Meyer (2010) showed that naïve listeners (untrained listeners) discern fine differences in children's speech" (as cited in Julien & Munson, 2012, p.1838). Johnson, Munson, and Edwards (2010) argued that both untrained listeners and speech-language pathologists are able to recognize details in a child's speech. Furthermore, the Rossetti (1990), which is a criterion referenced instrument intended to assess language in children ages birth to 36 months, states that a caregiver report or a clinician observation may be used for assessment purposes for they are both reliable.

Parent report has been said to be a reliable and valid measure of communication development, as well as a sensitive indicator of language delays in young children (Bates, Bretherton, & Snyder, 1988). This finding proves helpful because a child's level of communication is the best early indicator of a developmental delay (Wetherby & Prizant, 2003). Therefore, because of the naturalistic environment and communication partner, as well as the argued validity and reliability of the parent report, there should be a shift toward child and family-centered assessments with infants and young children. This could ensure higher quality evaluation and assessment practices that yield more meaningful measures (Rossetti, 1990).

Additionally, there is strong evidence that mothers specifically, are verbally responsive to infants' early vocalizations, and that they play an important role in eliciting vocal imitations from their infants around the ages of 1 to 2 years (Gros-Louis, West, Goldstein, & King, 2006). Brent and Siskind (2001) claimed that early word learning is dependent on exposure, which typically includes input from caregivers. According to

Clark and Bernicot (2008), and Olson and Masur (2012), mothers play a vital role in furthering their infants' language development, specifically for early word learning, by responding to their vocalizations. Still, the way mothers respond may depend on the kind of vocalizations their infants produce and on their infants' developmental level. Certain behavioral responses from social partners can generate specific vocalizations (Bloom 1993). Masur and Eichorst (2002) found that infants with more self-generated speech had more extensive vocabularies. They claimed that when the infant imitates, it is not only a sign of advanced language skills, but may also be a means by which infants learn words. Furthermore, frequent exposure to these vocalizations and imitations can solidify and improve the infants' original semantic representation (Gershkoff-Stowe, 2002).

Mothers' verbal responses to infants' early vocal behaviors have been linked with language outcomes for children (Brady, Marquis, Fleming & McLean, 2004). Maternal responses to imitations showed structural details that could facilitate early word learning (Olson & Masur, 2012). Vocal imitation in early infancy has been repeatedly linked with vocabulary acquisition. In the second year of life, infants who imitate more frequently have larger vocabularies (Bates, Bretherton & Snyder, 1988). Masur and Eichorst (2002) argued that infant imitation is not simply a sign of advanced language skills, but may also be a mechanism by which infants learn words.

Gros-Louis, West, Goldstein, and King (2006) reported that limited focus has been paid to the details of change in the infants' prelinguistic development. The lack of focus on early vocal development is surprising given that maternal responsiveness to infants during the first 2 years has been found to influence later language development. Also, social feedback is essential for vocal development. "Previous research

demonstrated that maternal feedback to prelinguistic vocalizations influenced the production of more developmentally advanced vocalizations suggesting that effects of maternal responsiveness on vocal development may start during the prelinguistic phase" (Gros-Louis, West, Goldstein, and King, 2006, p. 509).

Given that untrained listener/parent report is easy to obtain and has been proven reliable in tracking aspects of speech and language development, we propose that caregivers may also enable development of a more consolidated means for classifying infant vocalizations. It is expected that development of a gold standard for describing infant vocalizations through incorporation of caregiver (or untrained listener) report will enable identification of atypical patterns earlier in age, and therefore facilitate early intervention.

Chapter IV- Early Identification for Infants

According to Wetherby and Prizant, (2003), there is mounting evidence that intervention beginning during infancy or preschool has a greater impact on outcomes for children and families than providing services at school age. Furthermore, they reported that for every dollar spent on early intervention, up to 7 times that amount can be saved in later costs. "Recent advances in brain research show that the environment sculpts the young child's brain, as neurons form connections and mature in response to stimulation. The environment has the greatest potential to influence a child's developing brain during the first few years of life. Early experiences affect brain structure because the brain operates on a *use it or lose it* principle". If a child does not have adequate emotional, physical, cognitive, and language stimulation, neurons (and therefore functional ability) can be lost permanently" (Wetherby & Prizant 2003, pg. 8). Parents can be involved in early identification as early as age 2, however, language delay is typically identified by early school age (preschool or kindergarten) if not noticed, or reported by parents earlier (Zambrana, Pons, Eadie, & Ystrom, 2014).

We know that infant prelinguistic vocal behaviors predict later language abilities (Heimann, Strid, Smith, Tjus, Ulvund, & Meltzoff, 2006; Watt, Wetherby, & Shumway, 2006; Wetherby, Allen, Cleary, Kublin, & Goldstein, 2002; Wetherby, Goldstein, Cleary, Allen, & Kublin, 2003), but identifying those infants and toddlers at risk for future speech and language difficulties is challenging (Määttä, Laakso, Tovanen, Ahonen, & Aro, 2012) because *normal* vocal development is variable (Bates, Dale, & Thal, 1995; Darrah, Hodge, Magill-Evans, & Kembhavi, 2003; Fenson *et al.*, 2000; Thal, Bates, Goodman, & Jahn-Samilo, 1997). There are vital changes that occur in language during

the first 2 years. These changes are demonstrated in the transition from prelinguistic communication to linguistic communication. There is research indicating that prelinguistic abilities predict later language abilities, and by their first birthday, children can share attention and emotion, as well as communicate intentionally using a variety of gestures and speech-like sounds that have shared meanings with caregivers (Bates, 1976; Stern, 1985).

Language expression encompasses the child's use of verbal and nonverbal behaviors to communicate (Rossetti, 1990). An effective assessment of a child's expressive language begins with an accurate description of his expressive language history and current language expression skills by the parent or caregiver (Rossetti, 1990). When administering the Rosetti, for example, a child and his caregivers are observed as they interact and play, with particular attention given to the child's language productions. The examiner observes the child, and also directly interacts with him throughout the test to obtain more information about expressive skills. Beyond the Rosetti, expressive and receptive language have been shown to be the best indicator of developmental performance in children under 3 years of age (Capute, Palmer, & Shapiro, 1987). There is a strong correlation between a child's early language skills and their school performance later in life (Rossetti, 1990). However, the behaviors a child may display during an assessment, on any given day, represent only a sample of the child's full range of skills and behaviors (Rossetti, 1990).

Wetherby and Prizant (2003) found 3 main limitations of formal communication assessments which are based on current theories of language development beginning with questionnaires for young children. First, most of the formal instruments focus on

language milestones (e.g., number of different words) instead of social-communicative and symbolic foundations of language (e.g., gaze/point following). Second, examiners are primarily using clinician rather than child directed assessment tools which are actually limiting the chances of observing a spontaneous, child-initiated communication act. Third, most instruments limit the client's family from participating during the assessment process.

Early identification is key to a successful intervention. For young children who have, or are at risk for delay or disorder, early intervention has been shown to positively impact developmental outcomes (Hebbeler *et al.*, 2007; McLean & Cripe, 1997). Accordingly, it is imperative to identify infants and toddlers in need of services as early as possible so that intervention can be provided during pivotal developmental periods, when the brain is most capable of change due to plasticity (Hebbeler, 2009).

Chapter V- Purpose

The *long term goal* of this line of research is to develop a gold standard for describing infant vocalizations through incorporating the caregiver's perspective of sounds infants produce. The *objective* of the proposed study is to have untrained listeners, with no educational background in speech-language pathology or child development, name or describe infant vocalizations. The *central hypothesis* is that we will be able to generate a list of sound types in development that are commonly identified by untrained listeners, some of which will overlap with published categorical descriptors. This hypothesis is based on knowledge of infant vocal development and research showing that caregivers (untrained listeners) are good at reporting developmental milestones. The *rationale* for the proposed research is that, once a list of sound types has been generated based on untrained listener report, further exploration of the functionality of these sounds can be initiated. We can refine the list by having an additional group of untrained listeners and researchers classify presented vocalizations according to the sound types previously generated, and therefore, develop a more consolidated means to classify infant vocalizations. It is expected that development of a gold standard for describing infant vocalizations through incorporation of caregiver (or untrained listener) report will enable identification of atypical patterns earlier in age, and therefore facilitate early intervention.

The central hypothesis for this project will be tested by pursuing the following aim: from 4 to 12 months of infant age, across typically developing infants, we will identify the sound types described by untrained listeners. The *working hypothesis* for this aim is that sound types reported will fall into several distinct categories that can be used to reliably describe infant vocal development in the future. We expect the outcome for

the aim to show that untrained listener description will yield some alternate and additional terminology that may provide more complete coverage of the range of sounds infants produce as well as make training of coding infant vocalizations more natural. Our hypothesis is based on prior documentation that caregiver report is reliable for describing infant vocalizations (Bates, Bretherton, & Snyder, 1988).

Infants

Vocalizations for this study were obtained from 10 of 15 typically developing infants, video/audio recorded monthly in a study conducted by Dr. Heather Ramsdell-Hudock at East Carolina University (ECU)[1]. All infants were from 4 to 18 months of age at the beginning and termination of the study, respectively. Flyers advertising the study were sent to addresses (obtained from Register of Deeds records at the Pitt County Court House, Greenville, NC) of families with infants born between November, 2010 and March, 2011. Parents interested in participating in the study with their infants were interviewed, and details of the study, along with informed consent, were discussed. Inclusion criteria for the study consisted of caregivers who experienced normal pregnancies and no significant history of prenatal or perinatal problems; families where English was the primary language spoken in the home; families who were able to travel to the laboratory monthly; families who did not expect to move away from the surrounding area within 2 years of beginning participation in the study; and infants not at risk for developmental disorders. For further clarification, infants considered at risk would have been those who had experienced one or more of the following conditions prior to 7 months of age: pre- and/or perinatal problems; ear, nose, and throat problems;

[1] Only 10 of the 15 original participants had data at 4 and 5 months of age. These infants were selected to participate in the present study.

swallowing/sucking problems; and/or a family history of speech and/or language problems (Brady et al., 2004; Girolametto et al., 1999; Goldstein & Schwade, 2008; McDuffie & Yoder, 2010). For the purposes of this project, we explored data from 4 through 12 months of infant age. Following previous approval from the University Medical Center Institution Review Board at ECU, caregivers voluntarily gave informed consent for participation in the study. Further, exemption was obtained from the Human Subjects Committee at Idaho State University (ISU), as the purpose of the proposed study is covered in the original consent.

All families were of middle socioeconomic status (as determine through parent self-report on participant history interview). There were no infant participants born to single parent homes, and both mothers and fathers participated in the original study. Five of the infants were first born, three had one older sibling, and two had three older siblings. Siblings ranged in age from 2 years to 5 years at the time of infant participants' births.

Three of the ten infant participants were male, and seven were female. One female infant was African American, one male infant was Asian American (father of East Indian descent and mother of Vietnamese and Hawaiian descent), and one male infant was Palestinian and the rest were Caucasian. One male infant was from a home where English and Arabic were spoken, and a second male infant was from a home where English, Indian, and Vietnamese were spoken. All infants had normal hearing; they all passed an automated auditory brainstem response newborn screening (ALGO 3 or ALGO 5 Newborn Hearing Screener System) to click stimuli presented at 35 dB nHL. In addition, full hearing evaluations including tympanometry, transient evoked otoacoustic

emissions, and visual reinforcement audiometry were conducted at 6 and 18 months of age, with follow-up testing as needed for instances where results were abnormal (i.e., middle ear dysfunction) or testing was incomplete. One of the infants received bilateral myringotomy and pressure equalization tubes during their enrollment in the study.

Untrained Listeners

We sought to include 40 untrained listener participants in the study. Listeners qualified as untrained if they had not been previously educated in linguistics, protophone terminology, speech-language pathology, child development, or music (to avoid bias based on previously trained listening skills). Additionally, the listeners were native speakers of American-English with normal hearing. We inquired about gender, age, and parenting/caregiver experience. The listeners volunteered to participate in the study in response to fliers posted on the ISU campus in Pocatello, ID and word of mouth. Informed consent (see Appendix A) was obtained prior to participation.

Procedure and Analysis

A randomized set of infant vocalizations from the archived data base was located based upon a breathe group criterion (each change in the direction of airflow corresponds with a new utterance, Oller & Lynch, 1992) and extracted from the original recording sessions to eliminate extraneous caregiver and lab staff productions, toy sounds, and vegetative infant sounds. Once infant utterances were located, they were coded for vocal type. Generally, laboratory staff were instructed to use as few listening opportunities as possible before assigning codes (no more than three). The reason for this instruction was for the researchers to assign codes intuitively based on salient characteristics; the most prominent impression of each utterance was used to determine judgment. For vocal type,

utterances were coded as vowel, growl, squeal, raspberry, whisper, laugh, or cry. The coding was conducted without viewing video or TF32 spectrographic display of the utterances, as no visual support was allowed, so that viewing social interaction and the acoustic display did not skew coder judgment. Modal pitch across infants was judged intuitively by laboratory staff upon listening to vocalizations produced by each infant, and vocal type was coded accordingly. *Vowel* was coded if the utterance was perceived as predominantly produced in modal phonation, in the mid pitch range of the infant. *Growl* was coded if the most salient pitch of the utterance was notably lower than the infant's modal phonation, or if the pitch of the utterance was in the normal range but the utterance was produced with very high tension. *Squeal* was coded if the utterance was notably higher in pitch than the infant's modal phonation. *Raspberry* was coded if the infant produced any sort of lip or tongue trill. *Whisper* was coded if the infant produced a voiceless utterance with audibly perceptible articulatory movement. *Laugh* was coded if laughing was the most salient characteristic of the utterance, and *cry* was coded if uncontrollable crying (more than fussing) was perceived as the most salient characteristic of the utterance. The vocalizations were vetted to make sure there was diversity of sound type, the full range of sound types in the first year, and the same sounds (e.g., squeals, growls, vowels, etc.) from every infant (to the extent possible, as some sounds were not produced by some infants at certain ages). To the extent possible, one of each vocal type was selected from each infant at each age for presentation to untrained listeners.

The untrained listeners were all read the same script (which is shown below in Appendix B) to ensure all participants received the same information. All participants were given a hearing screening by a Speech-Language Pathology graduate Master's

student. If they passed the hearing screening, they were given further directions and administered the infant audio recordings through Microsoft Windows Media Player. The untrained listeners were played 193 vocalizations, some from each infant at each age, presented randomly. They were prompted to respond to the probe, "What would you call this type of sound? Please answer in one to three words." To avoid biasing, they were not provided any examples. Before beginning with the audio recordings, any questions they had, gender, age, and number of children were documented. Their responses were recorded for analysis. The listeners were blinded to infant and infant age. Listener responses were explored for themes and patterns within and across reports were identified.

Chapter VI- Results

We aimed to identify how untrained listeners would describe infant vocalizations, from ten infants between 4 and 12 months of age. Archived data files were coded based on infant utterance and age, with vowels, squeals, growls, raspberries, whispers, cries, and laughter selected for presentation to listeners. The distribution of these sounds across infant ages is displayed below in Table 1.

A total of 40 listeners participated in the study, with 5 listener responses excluded from the final exploration as a result of failed hearing screenings (2 participants) or data entry error (3 participants). Data entry errors consisted of the graduate clinician skipping one or more lines when recording responses which attached recorded responses to the wrong utterance types. Retesting could not be administered for it would affect reliability of outcomes/results. Of the remaining 35 participants, 23 were female and 12 were male. The average participant age was 30 years, ranging from 23 to 49 years. In addition, participants had one to six children, with an average of two.

Table 1
Number of Infant Utterance Types Presented to Untrained Listeners

		Infant Age in Months								
	Across Ages	4	5	6	7	8	9	10	11	12
Vowel	58	8	6	7	6	4	5	7	8	7
Squeal	36	4	4	2	5	7	5	0	3	6
Growl	34	5	3	3	4	4	3	5	3	4
Raspberry	9	1	1	1	0	2	0	1	3	0
Whisper	7	0	0	0	0	0	2	2	1	2
Cry	25	7	1	4	4	4	2	0	0	3
Laugh	23	2	0	3	3	3	1	3	4	4
Total	192	27	15	20	22	24	18	18	22	26

Tag clouds of the top 100 labels from listener responses were generated (using tagcrowd.com) to aid in visual presentation and reveal trends and patterns in responses.

Tag (or word) clouds are increasingly used for data analysis (Hearst, & Rosner, 2008). These tag clouds offer a visual sense of the key labels used by listeners in response to the infant vocalization stimuli. Each cloud visually represents the number of times a label occurred, with labels listed alphabetically and weighted to allow easy identification of the use frequency. The larger the font size, the more frequent the label. In addition to the tag clouds, the three most frequently used labels are presented in tabular form for each utterance type. Throughout the results and discussion, descriptive labels provided by listeners will be referred to as responses, labels, or tags. Themes from listener responses are as follows.

When untrained listeners were presented with VOWEL stimuli, "coo," "talk," and "happy" (or some derivative of these tags, e.g., "cooing") were the most frequently elicited labels, as displayed in Figure 1 and Table 2. Of all listener responses across infant ages, the VOWEL stimuli were tagged as "coo" 334 times (16.5% of the responses), as "talk" 269 times (13.3% of the responses), and as "happy" 226 times (11. 4% of the responses). In total, these three responses accounted for 41.2% of all of the labels provided for VOWEL stimuli.

If we look at these responses more closely, we can see that the pattern of labels is largely represented within all infant ages. "Coo" was the most frequently occurring response at every infant age except 7 months ("talk" was tagged 0.9% more often) and 11 months ("happy" was tagged 0.4% more often). Some other labels elicited for VOWEL stimuli included "tired," "content," "babble," "noise," and "sigh." The actual label "vowel" only appeared once across all listener responses to VOWEL stimuli.

Figure 1. Tag cloud for listener response labels to vowel utterances.

Table 2
Proportion of Occurrence for each Response Type for Vowel Utterances

	Across Ages	Infant Age in Months								
		4	5	6	7	8	9	10	11	12
Coo	16.5	15.0	19.5	15.1	18.1	20.0	16.0	20.0	12.5	14.7
Talk	13.3	12.9	10.5	13.9	19.0	15.0	11.4	15.1	11.1	12.2
Happy	11.4	8.6	14.3	10.2	8.1	12.9	9.7	13.1	12.9	13.1
Total	41.2	36.5	44.3	39.2	45.2	47.9	37.1	48.2	36.5	40.0

*Cell values are presented as a percentage of the total number of responses obtained.

As shown in Figure 2 and Table 3 the stimuli coded as SQUEAL were most often labeled as “squeal,” “happy,” and some form of “high.” “Squeal,” or a derivative was used 333 times (26.7% of the responses), “happy” 213 times (17.1% of the responses), and “high” had 72 responses (5.7% of the responses). In total, these three responses accounted for 49.5% of all of the labels provided for SQUEAL stimuli. Note that the label “high” does not appear exclusively in the tag cloud. However, when exploring the tag frequencies in excel, it was observed that “high” was the third most frequently occurring label. Our conclusion is that “high” did not appear in the tag cloud because it was always used as a qualifier for other listener responses, such as “high pitch”, “high

squeal", or "high noise". "Excited", "coo", "pitch", and "squeak" were also common responses to SQUEAL stimuli. "Squeal" was the most frequently occurring response at every infant age (except for 10 months where no coded utterances were used). Accordingly, the term SQUEAL appears to be mutually understood across trained and untrained listeners alike.

Figure 2. Tag cloud for listener response labels to squeal utterances.

Table 3
Proportion of Occurrence for each Response Type for Squeal Utterances

	Across Ages	Infant Age in Months							
		4	5	6	7	8	9	11	12
Squeal	26.7	23.6	22.1	31.4	22.3	27.3	25.7	35.2	30.0
Happy	17.1	17.1	14.3	15.7	10.9	16.3	20.0	24.8	19.0
High	5.7	7.1	3.6	5.7	6.9	5.3	6.3	6.7	4.8
Total	49.5	47.8	40.0	52.8	40.1	48.9	52.0	66.7	49.0

*Cell values are presented as a percentage of the total number of responses obtained.

When the untrained listeners were presented with GROWL stimuli, "grunt," "growl," and "play" were the three highest yielded responses, as displayed in Figure 3 and Table 4 below. "Grunt," or some derivative, was provided 199 times (17.1%),

"growl" 110 times (9.2%), and "play" 98 times (8.2%). This pattern of labels is largely represented across all infant ages. "Grunt" was the most frequently occurring response at every infant age except 6 months ("play" was tagged 8.5% more often and "growl" was tagged 4.7% more often). Some other labels elicited for GROWL stimuli included "happy," "talk," "frustrated," "coo," "tired," "noise," and "cry." The actual label "growl" appeared second most frequently to "grunt", suggesting it as a strong contender for describing low pitch/harsh quality infant vocalizations by untrained listeners.

Figure 3. Tag cloud for listener response labels to growl utterances.

Table 4
Proportion of Occurrence for each Response Type for Growl Utterances

		Infant Age in Months								
	Across Ages	4	5	6	7	8	9	10	11	12
Grunt	17.1	10.9	18.1	4.8	24.3	18.6	9.5	17.1	32.4	18.6
Growl	9.2	7.4	4.8	8.6	11.4	8.6	4.8	12.6	8.6	13.6
Play	8.2	5.1	4.8	13.3	8.6	9.3	6.7	10.9	5.7	9.3
Total	34.5	23.4	27.7	26.7	44.3	36.5	21.0	40.6	46.7	41.5

*Cell values are presented as a percentage of the total number of responses obtained.

When untrained listeners were presented with RASPBERRY stimuli, “raspberry,” “spit,” and “bubble” (or some derivative of these tags, e.g., “spitting”) were the most frequently elicited labels, as displayed in Figure 4 and Table 5. Of all listener responses across infant ages, the RASPBERRY stimuli were tagged as “raspberry” 42 times (13.3% of the responses), “spit” 40 times (12.7% of the responses), and “bubble” 23 times (7.3% of the responses). In total, these three responses accounted for 33.3% of all of the labels provided for RASPBERRY stimuli.

If we look at these responses more closely, we can see that the pattern of labels is largely represented within all infant ages except 7 months (“spit” and “bubble” were each tagged 2.9% more often) and 11 months (“spit” was tagged 1.9% more often). Some other labels elicited for RASPBERRY stimuli included “play,” “blow,” “grunt,” “fart,” and “coo”.

Figure 4. Tag cloud for listener response labels to raspberry utterances.

Table 5
Proportion of Occurrence for each Response Type for Raspberry Utterances

	Across Ages	Infant Age in Months						
		4	5	6	8	10	11	12
Raspberry	13.3	5.7	5.7	11.4	12.9	28.6	14.3	13.3
Spit	12.7	8.6	11.4	5.7	11.4	17.1	16.2	12.7
Bubble	7.3	8.6	11.4	5.7	7.1	8.6	5.7	0
Total	33.3	22.9	28.5	22.8	31.4	54.3	36.2	26.0

*Cell values are presented as a percentage of the total number of responses obtained.

WHISPER stimuli were presented to untrained listeners from only 9 to 12 months of infant age, as examples of whispers could not be found in the recordings from infants 4 through 8 months of age. When untrained listeners were presented with WHISPER stimuli, “whisper,” “talk,” and “breath” (or some derivative of these tags, e.g., “talking” or “breathing”) were the most frequently elicited labels, as displayed in Figure 5 and Table 6. Of all listener responses across infant ages, the WHISPER stimuli were tagged as “whisper” 43 times (17.6% of the responses), “talk” 41 times (16.3% of the responses), and “breath” 21 times (8.6% of the responses). In total, these three responses accounted for 42.5% of all of the labels provided for WHISPER stimuli. Other common responses included “babble,” “coo,” “quiet,” “play,” “noise,” “mama,” “soft,” “sigh,” “happy,” and “content.”

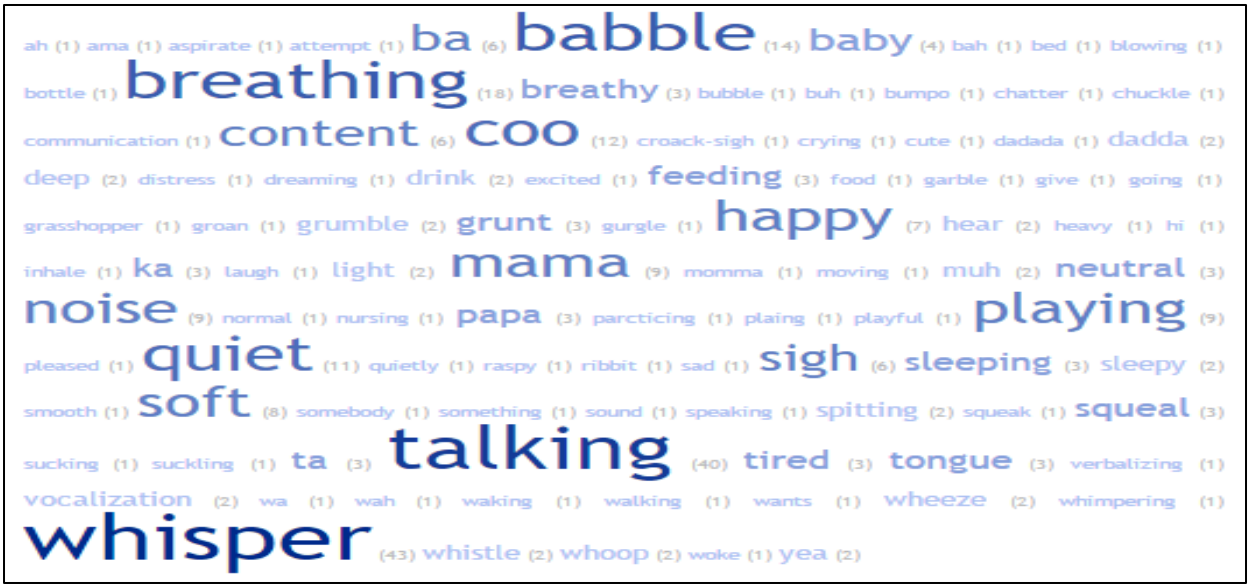

Figure 5. Tag cloud for listener response labels to whisper utterances.

Table 6
Proportion of Occurrence for each Response Type for Whisper Utterances

		Infant Age in Months			
	Across Ages	9	10	11	12
Whisper	17.6	27.1	20.0	8.6	10.0
Talk	16.3	12.9	14.3	17.1	21.4
Breath	8.6	18.6	4.3	2.9	5.7
Total	42.5	58.6	38.6	28.6	37.1

*Cell values are presented as a percentage of the total number of responses obtained.

When untrained listeners were presented with CRY stimuli, "cry," "sad," and "upset" (or some derivative of these tags, e.g., "crying") were the most frequently elicited labels, as displayed below in Figure 6 and Table 7. Of all listener responses across infant ages, the CRY stimuli were tagged as "cry" 313 times (35.7% of the responses), as "sad" 132 times (15% of the responses), and as "upset" 72 times (8.2% of the responses). In total, these three responses accounted for 58.9% of all of the labels provided for CRY stimuli. "Cry" was the most frequently occurring response at every infant age (except for 10 and 11 months where no coded utterances were used). Accordingly, the term CRY

appears to be mutually understood across trained and untrained listeners alike. Other tags elicited in response to CRY stimuli included “tired,” “mad,” “hurt,” “angry,” “hungry,” “frustrated,” “unhappy,” and “coo”.

Figure 6. Tag cloud for listener response labels to cry utterances.

Table 7
Proportion of Occurrence for each Response Type for Cry Utterances

	Across Ages	Infant Age in Months						
		4	5	6	7	8	9	12
Cry	35.7	36.7	34.3	33.6	36.4	39.3	30.0	34.3
Sad	15.0	13.1	11.4	12.1	22.1	16.4	18.6	10.5
Upset	8.2	6.9	11.4	9.3	8.6	6.4	8.6	10.5
Total	58.9	56.7	57.1	55	67.1	62.1	57.2	55.3

*Cell values are presented as a percentage of the total number of responses obtained.

When untrained listeners were presented with LAUGH stimuli, “laugh,” “happy,” and “giggle” (or some derivative of these tags, e.g., “laughing”) occurred most frequently among elicited labels, as displayed below in Figure 7 and Table 8. Of all listener responses across infant ages, the LAUGH stimuli were tagged as “laugh” 271 times (33.7% of the responses), as “happy” 129 times (16% of the responses), and as “giggle”

124 times (15.4% of the responses). In total, these three responses accounted for 65.1% of all of the labels provided for LAUGH stimuli. As with squeal and cry, the term "laugh" appears to be mutually understood across trained and untrained listeners alike. "Laugh" was the most frequently occurring response at every infant age (except for 5 months where no coded utterances were used). Some other labels elicited for LAUGH stimuli include "coo," "play," and "chuckle."

Figure 7. Tag cloud for listener response labels to laugh utterances.

Table 8
Proportion of Occurrence for each Response Type for Laugh Utterances

		Infant Age in Months							
	Across Ages	4	6	7	8	9	10	11	12
Laugh	33.7	35.7	30.5	23.1	41.0	51.4	21.0	27.9	47.9
Happy	16.0	18.6	12.4	15.7	16.2	25.7	10.5	14.3	21.4
Giggle	15.4	10.0	18.1	13.9	19.0	20.0	16.2	15.7	15.7
Total	65.1	64.3	61	52.7	76.2	97.1	47.7	57.9	85

*Cell values are presented as a percentage of the total number of responses obtained.

Chapter VII- Discussion

The intent of this study was to generate a list of terms in development that are commonly identified by untrained listeners as labels for infant vocalizations. Accordingly, from 4 to 12 months of infant age, across typically developing infants, sound types were identified and associated with infant vocalizations as described by untrained listeners. The following is our resulting list: coo, talk, happy, squeal, high, grunt, growl, play, raspberry, spit, bubble, whisper, breath, cry, sad, upset, laugh, and giggle. As expected, the responses fell into several distinct categories, some of which overlapped with published categorical descriptors, and some of which yielded alternate and additional terminology that may provide more complete coverage of the range of sounds infants produce. Overlapping labels include squeal, growl, raspberry, whisper, cry, and laugh. Vowel is the only published categorical descriptor not provided by untrained listeners. Now, further exploration of the functionality of these sounds can be initiated. It is expected that development of a gold standard for describing infant vocalizations through incorporation of caregiver (or untrained listener) report will enable identification of atypical patterns earlier in age, and therefore facilitate early intervention.

When untrained listeners were presented with VOWEL stimuli, it seems they did not quite know how to respond. Rather than naming the utterances with a descriptive label (e.g., vowel or consonant), the listeners seemed to describe a quality of the vocalization. For example, the vocalizations were often described as happy sounds or sad sounds, attributing a physical state to the infant. It was as if the vocalizations themselves, without the presence of a visual, elicited a similar response from listeners as early stage perlocutionary acts do. Recall that perlocutionary acts are communicative acts that have

some effect on the listener without having control over said effect (e.g., the hunger cry of a newborn infant resulting in a parent feeding him). The tags associated with VOWEL stimuli varied most across listeners. The listeners reported difficulty in labeling these vocalizations because it was hard to know if the utterance/sound was preceding a happy or unhappy emotion. Most participants assumed it was a happy sound and tagged the vocalization as a "coo," "talk," or "happy" sound. It is interesting to note this urge to label the infant vocalizations dependent upon physical state attributes. The only true vocalizations tied to physical states in human communication are cries and laughs (Oller, Buder, Ramsdell, Warlaumont, Chorna, & Bakeman, 2013), which is perhaps why listeners struggled with classification of VOWEL stimuli.

When SQUEAL stimuli were presented, untrained listeners responded as researchers would, with a "squeal" designation. The three most common words used to describe this utterance set were "squeal," "happy," and "high." Untrained listeners often attached some type of descriptive label prior to the "squeal" tag, such as *high pitched* squeal. As with VOWEL stimuli, listeners seemed to prefer to describe the quality of the SQUEAL stimuli by defining a physical or emotional state (e.g., happy squeak). The tags associated with SQUEAL stimuli varied far less than with VOWEL stimuli. However, listeners continued to report difficulty in labeling these vocalizations, stating it was hard to determine if the utterances were preceding or following a happy or unhappy emotion. The untrained listeners reported that they would have been better able to decipher the single utterance if they were given either more known context or longer utterances. Regardless, "squeal" appears to be a consistent way of classifying high pitched infant vocalizations, because it was the top response given by untrained listeners.

The utterances coded as GROWL were commonly depicted by the untrained listeners as “grunt,” which appears similar in description to the term GROWL. Responses such as “clear throat,” “cough,” “weed whacker,” “roar,” and “dog bark” were also recorded. It appeared at times, some of the utterances evoked a response that was unexpected and occasionally not human generated. Considering all of these responses, the top three labels to describe GROWL stimuli were “grunt,” “growl,” and “play.” These three responses appeared to represent this vocal type well, which confirms that GROWL is probably an appropriate way of referring to these low pitch/harsh utterances, for researchers, clinicians, and/or untrained listeners.

It also appears that RASPBERRY stimuli would be labeled as “raspberries” by both researchers and untrained listeners alike. Here, the top three reported labels were “raspberry,” “spit,” and “bubble.” Out of all vocal types presented to listeners, this group engendered the most colorful responses. The responses varied from “poop,” “toot,” “sneeze,” and “fart,” to “slobber,” “blow,” “cough,” and “buzzing”. The listeners would occasionally verify with the researcher that these sounds were indeed generated by an infant through the oral cavity. For most of the RASPBERRY stimuli, “raspberry” seemed to be the common way to refer to the vocal type.

WHISPER stimuli appeared to confuse listeners, who often considered these vocalizations as just “breathing”. The most common terms used for this utterance category were “whisper,” “talk” and “breath.”

Understandably, CRY stimuli were another vocal type with which listeners attributed a physical state to the infant, including emotional detail in their responses (such as “upset cry,” “fussy cry,” “mad cry”). The three most common terms used to label CRY

stimuli were “cry,” “sad,” and “upset.” Listeners seemed to assume the label “cry,” and attach some type of unhappy feeling to the vocalizations (e.g., “not happy,” “annoyed,” “angry,” “wants attention,” “tired,” “hungry,” “bored”). Although “cry” would umbrella all additional words used to describe the overall negative affect in these utterances, the untrained listeners/caregivers were more descriptive in their labeling. The label “cry” would likely be equally understood by researcher or untrained listener and may be the best tag for this utterance type.

Finally, emotional tags were provided for LAUGH stimuli as well (such as “happy,” “excited,” and “pleasure”). The top three most common responses for LAUGH stimuli were “laugh,” “happy” and “giggle.” It is likely that “laugh” would be equally understood by researcher or untrained listener.

Study Limitations

This study has a number of limitations; primarily it was limited by sections of data that were incorrectly entered because of spelling errors such as “hassy” instead of “happy.” Before proceeding with the long term goal of this study, the data will need to be examined and the necessary corrections made. A secondary limitation would also be location, only using untrained listeners currently located in Eastern Idaho and vocalizations from infants in Eastern North Carolina. Language, dialect, and culture could have each skewed the responses. In the future, researchers could collaborate with colleagues across the country to recruit participants from diverse locations, backgrounds, and dialects to give the study a more rich variety and diverse population. In addition to this, the sample only included ten infants. A larger sample size of 16 infants could have been obtained from archived data, but the time needed to code infant utterances was

unrealistic for the scope of a thesis project. In addition, the archived data only contained recordings at 4 through 6 months for some of the 16 infants, so a new corpus of data would need to be obtained for further analysis at early ages. Further, given that infant vocalizations are highly variable (ranging in such things as volubility and sound types produced), sampling only 60 minutes of recorded vocalizations from each infant at each age may not provide a representative sampling of the sounds these infants could produce. A larger dataset from a wider selection of participants/untrained listeners and infants, including a non-English speaking, and/or an urban-centered community could provide a more in-depth insight into common responses. Finally, no quantitative statistical procedures were used in tabulating the results. It is possible that Chi Square or Odds Ratio Effect Sizes may provide statistical information to aid in interpretation of the results. As it stands, the results should be generalized with caution.

Clinical Implications

Results from this and future studies could help to streamline the process of identifying developmental status by bolstering parents and researchers (trained and untrained listeners) to use and understand the same terminology in identification of infants who may be at risk. It is crucial to identify infants and toddlers in need of services as early as possible when the brain is most capable of change due to plasticity (Hebbeler, 2009). Parents can be involved in early identification and they can inform practice. Early identification is key to a successful intervention. For young children who have, or are at risk for delay or disorder, early intervention has been shown to positively impact developmental outcomes. (Hebbeler *et al.*, 2007; McLean & Cripe, 1997).

Future Directions

The future plan is to have additional trained and untrained listeners (caregivers *and* individuals that have and have not been educated in child development, protophone terminology, linguistics, speech-language pathology, child development, or music) describe infant vocalizations. In the follow-up study, listeners will be provided with a list of terms and asked to select the most representative term to describe sounds presented. The terms will be generated from the most frequently occurring responses in the present study (e.g., coo, happy, high, etc.), and also the commonly used research terms (e.g., vowel, squeal, growl, etc.). From this follow-up study, the most consistently used terms to describe the infant sounds will highlight the best descriptors for infant sounds. The most frequently identified terms will likely be those that both parents and researchers (trained and untrained listeners) use and understand.

By documenting untrained listener responses, it helps to inform practice and have an important *positive impact* on clinicians, researchers, caregivers, and clients, by providing a necessary first step toward developing a gold standard for describing infant vocalizations through incorporating the caregiver's perspective of sounds infants produce. It is helpful and more applicable if naive listener and researcher alike could understand and use the same terminology. This will help to begin bridging the gap between research and clinical practice, moving us one step closer to streamlining the process of identifying developmental status for early identification of speech and/or language delay/disorder, and early intervention.

References

Bates, E. (1976). *Language and context: The acquisition of pragmatics.* New York, NY: Academic Press.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge, MA: Cambridge University Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development, 21*, 205-226.

Bates, E., Dale, P.S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (p. 96-151). Oxford, United Kingdom: Blackwell.

Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression.* Cambridge, England: Cambridge University Press.

Brady, N., Marguis, J., Fleming, K. & McLean, L. (2004). Prelinguistic predictors of language growth in children with developmental disabilities. *Journal of Speech, Language & Hearing Research, 47,* 663-677.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81,* B33-B44.

Capute, A., Palmer, F. B., & Shapiro, B. K. (1987). Using language to track development. *Patient Care, 21(19),* 60-70.

Clark, E. V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language, 35(2),* 349-371. doi: 10.1017/S0305000907008537.

Dale, P.S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research, 34(3),* 565-571.

Darrah, J., Hodge, M., Magill-Evans, J., & Kembhavi, G. (2003). Stability of serial assessments of motor and communication abilities in typically developing infants – Implications for screening. *Early Human Development, 72,* 97-110.

Feldman, H. M., Dale, P. S., Campbell, T. F., Colborn, D. K., Kurs-Lasky, M., Rockette, H. E., & Paradise, J. L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development, 76(4),* 856-868.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J.S., & Thal, D. (2000). Measuring variability in early child language: Don't shoot the messenger. *Child Development, 71,* 323-328.

Gershkoff-Stowe, L. (2002). Object naming, vocabulary growth, and the development of word retrieval abilities. *Journal of Memory and Language, 46(4),* 665-687.

Girolametto, L., Weitzman, E., Wiigs, M., & Pearce, P.S. (1999). The relationship between maternal language measures and language development in toddlers with expressive vocabulary delays. *American Journal of Speech-Language Pathology, 8,* 364-374.

Goldstein, M.H., & Schwade, J.A. (2008). Social interaction to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19,* 515-523.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30(6),* 509-516.

Hebbeler, K. (2009). *First five years find briefing. Presentation given at a Congressional briefing on June 11, 2009, to discuss Education that works: The impact of early childhood intervention on reducing the need for special education services.* Retrieved from http://www.sri.com/neils/pdfs/FFYF_Briefing_Hebbeler_June2009_test.pdf

Hebbeler, K. Spiker, D., Bailey, D., Scarborough, A., Mallik, S., Simeonsson, R., & Singer, M. (2007). *Early intervention for infants & toddlers with disabilities and their families: Participants, services, and outcomes. Final report of the National Early Intervention longitudinal Study (NEILS).* Retrieved from http://www.sri.com/neils/pdfs/NEILS_Report_02_07_Final2.pdf

Heimann, M., Strid, K., Smith, L., Tjus, T., Ulvund, S.E., & Meltzoff, A.N. (2006). Exploring the relation between memory, gestural communication, and the emergence of language in infancy: A longitudinal study. *Infant and Child Development, 15,* 233-249.

Hearst, M. & Rosner, D. (2008). Tag Clouds: Data Analysis Tool or Social Signaller? Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS) http://doi.ieeecomputersociety.org/10.1109/HICSS.2008.422

Johnson, J. M., Munson, B., & Edwards, J. (2010). *The role of clinical experience in listening for phonetic detail in children's speech.* Poster session presented at the Symposium for Research in Child Language Disorders, Madison, WI. doi: 10.1044/1058-0360(2011/11-0009).

Julien, H.M., & Munson, B. (2012). Modifying speech to children based on their perceived phonetic accuracy. *Journal of Speech, Language, and Hearing Research, 55,* 1836-1848. doi: 10.1044/1092-4388(2012/11-0131)

Kent, R. D. (1981). Sensorimotor aspects of speech development. *Development of Perception: Psychobiological Perspectives, 1,* 161-189.

Masur, E. F., & Eichorst, D. L. (2002). Infants' spontaneous imitation of novel versus familiar words: Relations to observation and maternal report measures of their lexicons. *Merrill-Palmer Quarterly, 48(4),* 405-426.

Masur, E. F., & Olson, J. (2008). Mother's and infants' responses to their partners' spontaneous action and vocal/verbal imitation. *Infant Behavior and Development, 31(4),* 704-715. doi: 10.1016/j.infbeh.2008.04.005.

Määttä, S., Laakso, M.-L., Tovanen, A., Ahonen, T., & Aro, T. (2012). Developmental trajectories of early communication skills. *Journal of Speech, Language, and Hearing Research, 55,* 1083-1096.

McDuffie, A., & Yoder, P. (2010). Types of parent verbal responsiveness that predicts language in young children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 53,* 1026-1039.

McLean, L.K., & Cripe, J.W. (1997). The effectiveness of early intervention for children with communication disorders. In M.J. Guralnick (Ed.), *The effectiveness of early intervention* (p. 349-428). Baltimore, MD: Brookes.

Munson, B., Edwards, J., Schelinger, S. K., Beckman, M. E., & Meyer, M. K. (2010). Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an

extraterrestrial view of Vox Humana. *Clinical Linguistics & Phonetics, 24(4-5),* 245-260. doi: 10.3109/02699200903532524.

Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics & Phonetics, 20(5),* 351-369.

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

Oller, D.K., Buder, E.H., Ramsdell, H.L., Warlaumont, A.S., Chorna. L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences, 110,* 6318-6323.

Oller, D. K., Eilers, R., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science, 4(1),* 49–60.

Olson, J., & Masur, E.F. (2012). Mothers respond differently to infants' familiar versus non-familiar verbal imitations. *Journal of Child Language, 39,* 731-752. doi: 10.1017/S0305000911000262

Prizant, B., & Wetherby, A. (1990). Assessing the communication of infants and toddlers: Integrating a socioemotional perspective. *Zero to Three Bulletin, 11,* 1-12.

Ramsdell, H. L., Oller, D. K., Buder, E. H., Ethington, C. A., & Chorna, L. (2012). Identification of prelinguistic phonological categories. J*ournal of Speech, Language, and Hearing Research, 55(6),* 1626-1639. doi: 10.1044/1092-4388(2012/11-0250)

Rescorla, L., & Alley, A. (2001). Validation of the Language Development Survey (LDS): A parent report tool for identifying language delay in toddlers. *Journal of Speech, Language, and Hearing Research, 44(2),* 434-445.

Rossetti, L. M. (1990). *The Rossetti Infant-Toddler Language Scale: A measure of communication and interaction.* East Moline, Il: LinguiSystems.

Stark, R. E., Bernstein, L. E., & Dermorest, M. E. (1993). Vocal communication in the first 18 months of life. *Journal of Speech, Language, and Hearing Research, 36(3),* 548-558.

Stern, D. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology.* New York, NY: Basic Books.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72(3),* 748-767.

Thal, D.J., Bates, E., Goodman, J., & Jahn-Samilo J. (1997). Continuity of language abilities: An exploratory study of late- and early-talking toddlers. *Developmental Neuropsychology, 13,* 239-273.

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic predictors of language outcome at 3 years of age. *Journal of Speech, Language, and Hearing Research, 49,* 1224-1237.

Wetherby, A.M., Allen, L., Cleary, J., Kublin, K., & Goldstein, H. (2002). Validity and reliability of the Communication and Symbolic Behavior Scales Developmental Profile with very young children. *Journal of Speech, Language, and Hearing Research, 45,* 1202-1218.

Wetherby, A.M., Goldstein, H., Cleary, J., Allen, L., & Kublin, K. (2003). Early identification of children with communication disorders: Concurrent and predictive validity of the CSBS Developmental Profile. *Infants and Young Children, 16,* 161-174.

Wetherby, A. M., & Prizant, B. M. (1992). Profiling young children's communicative competence. *Causes and Effects in Communication and Language Intervention, 1,* 217-253.

Wetherby, A. M., & Prizant, B. M. (2003). *Communication and Symbolic Behavior Scales: Normed Edition.* Baltimore, MD: Paul H Brookes.

Zambrana, I. M., Pons, F., Eadie, P., & Ystrom, E. (2014). Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. International Journal of Language & Communication Disorders, 49(3), 304-316. doi:10.1111/1460-6984.12073

Appendix A

Script for Participants

First of all, thank you for participating in this project. It may seem odd that I'm reading to you right now, but I have to follow a script to ensure that all participants receive the same information.

We'll begin today by conducting a hearing screening. If you pass the hearing screening, I'll play a number of infant vocalizations to you, and have you describe the sounds. Some vocalizations will be loud; some quiet. Specifically, you'll be responding to the question, "What would you call this type of sound? Please answer in one to three words."

For each vocalization, you will provide a response that I will record on my computer. We will spend 30 to 60 minutes doing this. Ideally, it would be best if I only play it once, but replaying it one more time would be permitted if needs be.

The task may be challenging, as you probably haven't always thought of how to describe infant vocalizations. However, I can't provide you with an example because I don't want to skew your response. Please just try to describe the sounds as best you can. We would like a response for everything.

Do you have any questions?

I need to begin by recording your gender, age, and number of children.