University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

12-3-2013

# Infant Volubility Across Circumstance Estimated from All-day Recordings

Barbara S. Franklin

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation

Franklin, Barbara S., "Infant Volubility Across Circumstance Estimated from All-day Recordings" (2013). *Electronic Theses and Dissertations*. 848.
https://digitalcommons.memphis.edu/etd/848

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

# INFANT VOLUBILITY ACROSS CIRCUMSTANCES ESTIMATED FROM ALL-DAY RECORDINGS

by

Barbara S. Franklin

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Communication Sciences and Disorders

The University of Memphis

December, 2013

## Dedication

I dedicate my dissertation work to my family and colleagues. I am especially grateful to Bruce Alexander for being supportive throughout this entire process. My daughter Emily Franklin has also been an emotional support and I appreciate her help.

I also dedicate this dissertation to my colleagues at CSD, Lauren Burrows for her insights and support, Edina Bene for being there to listen and provide statistics advice, and to Kathy Fulmer for editing and encouraging.

**Acknowledgments**

I wish to thank my committee members who were more than generous with their expertise and precious time. Special appreciation goes to D. Kimbrough Oller, my committee chair and primary mentor for his countless hours of reflecting, reading, encouraging, and most of all patience throughout the entire process. An additional thank you goes to Eugene H. Buder, Corrina Ethington, and Stan Franklin for agreeing to serve on my committee.

Grant support for data collection at the University of Memphis funding comes from NIDCD DC011027. The LENA Foundation was also a great support of information and advice. I wish to thank them as well.

## Abstract

Franklin, Barbara Sarah. Ph.D. The University of Memphis. December, 2013. Infant Volubility Across Circumstances Estimated From All-Day Recordings. D. Kimbrough Oller.

By analyzing infant volubility we may illuminate early communication development and aid early identification of disorders. LENA recordings provide a naturalistic method for gathering data all-day and yield automated labeling of speakers including the infant wearing the device. In Study 1, 10 six-month-old infants were studied for volubility in a naturalistic setting analyzed by LENA and a human coder using a traditional repeat-listening method based on a single all-day recording for each infant. Twelve randomly selected 5-minute segments (excluding presumed sleep) were coded for vocal type (vocants, squeals, and growls). This allowed for estimation of six-month-old infant vocal rate across an entire day, as well as rate of each vocal type in naturalistic settings and comparison between human coders and the LENA automated software analyses. These are the first human infant volubility data based on truly naturalistic sampling. This data will lay groundwork for comparing vocal rate in humanity and other species. Additionally, we may shed light on the role of volubility and the diversity of vocal types in human infancy in predicting development and disorders of language.

In Study 2, twenty-four randomly sampled 5-minute segments from the same recordings as Study 1 were coded in real-time to determine effects of circumstance on volubility. Additionally, 10 five-minute segments with highest child vocalization rate (determined by LENA) and random sampling of ten presumed sleep segments (very low volubility determined by LENA) were also coded. First, coders listened to segments, coding vocalizations in real time as vocants, growls, squeals, laugh, or cry. Each coder

then coded circumstance (including but not limited to vocalizations directed to the infant, vocalizations directed to others, infant alone, and infant asleep) by answering a set of questions with a scaled response.

This is the first reported naturalistic assessment of infant volubility across circumstances. The work will help determine the extent to which infants use both protophones and fixed signals (cry and laugh) spontaneously, instrumentally, and in social circumstances, offering perspective on both the endogenous tendencies for vocalization in human infants and tendencies for vocalization used for communicative purposes.

# Table of Contents

Section Page

## List of Tables

## List of Figures

Figure Page

## Statement of Problem

Infant volubility has typically been studied in short-term observations either at home or within the lab setting. Study of infant volubility with all-day recordings at home may provide much more important information regarding patterns of infant vocalization. The study of infant volubility levels in specific social or non-social circumstances may also provide information regarding the inherent endogenous motivation in the human infant to vocalize.

Prior research results regarding volubility for typically developing infants differ substantially. Mean infant volubility levels reported in the literature vary amongst typically developing 6-month-old infants from a low of 1.3 vocalizations per minute (Hsu, Fogel, & Messinger, 2001) to a high of 11.3 vocalizations per minute (Molemans, 2011). Differences in circumstances of recording or in measurement methods may be responsible for the wide disparity of results. The research suggests at least as large a range of differences across infants within studies (Hsu et al., 2001; Jaffe et al., 2001; Molemans, 2011; Oller, Eilers, Urbano, & Cobo-Lewis, 1997; Stark, Bernstein, & Demorest, 1993).

The studies presented in this dissertation were done in a completely naturalistic environment within the family setting across an entire day. This approach provides a look into a naturalistic 6-month-old infant volubility level, both in general, and in differing social or environmental circumstances, which may provide clues to infant learning, including the development of strategic vocalization, and the role of auditory (especially parental) input. Further research may help provide more sensitive assessments for both developmental evaluation and detection of possible anomalies. Also the ability to determine and assess infant volubility when infants are alone may provide information regarding the endogenous nature of infant vocal development.

Another goal of the dissertation is to compare the human coding of six-month-old infants to the LENA automated analysis. This will help determine the efficacy of the LENA analysis software and may make it possible to develop a preliminary means of adjusting LENA's automated estimate of volubility for any database of all-day recordings for 6-month-olds to match the volubility that would likely be found with human coding.

## Introduction

Prior to speaking in words, infants appear to vocalize as a way to practice motor patterns that are precursors to speech production and as an early form of social interaction (Locke, 1989). Early infant vocalizations produced in the home setting across an ordinary day may provide insight into the kinds of circumstances that might produce contingent social responses from infants, as well as those vocalizations that are produced spontaneously by infants.

Typical levels of infant volubility are not easy to estimate. Prior research has reported large ranges of volubility in six-month old infants (Hsu et al., 2001; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001; Molemans, 2011; Oller et al., 1997; Stark et al., 1993). These studies have all been based on short-term (often 20 min or less) sampling that may not be representative of infant vocalization in general. By measuring infant volubility using all-day recordings, now possible because of the development of the LENA system (Richards, Gilkerson, Paul, & Xu, 2008), we may acquire much more generalizable information about vocal development. By evaluating the all-day recordings for a variety of circumstances, we may help clarify differences across prior studies, and more importantly, help determine the roles of social, instrumental, or spontaneous circumstances in vocal development.

Two studies, using naturalistic LENA home recordings are presented here based on all-day recordings from 10 six-month-old infants. This will be the first time that an all-day infant volubility level will be estimated using at home LENA automated analysis and comparison human coding. The comparison between the LENA analysis and the human coding analysis should produce a way to convert results from the automated analysis system to be more consistent with human coding–this way the very large LENA Research Foundation database can conceivably be used as a basis for estimating infant volubility at much larger sample size. Furthermore, review of infant volubility in different circumstances (social or alone) will help in assessing the functions of infant vocalizations and the influences that drive their production.

## Literature Review

Human infants have richly differentiated vocalizations at a very early age. The precursors to speech, the protophones, include according to Oller (2000): vocants, quasi- or fully resonant vowel-like sounds, squeals, (higher pitched than normal sounds) and growls, (lower pitched or harsh vocal quality normal pitched sounds) as well as gooing, raspberries, and canonical babbling (Koopmans-van Beinum, & Van der Stelt, 1986; Oller, 1980; Stark, 1978). They appear to be produced flexibly in a variety of settings or situations.

Volubility is the amount of this speech-like vocalization produced by infants over a period of time, measured by either the number of vocalizations per minute or the percent of time spent vocalizing. High infant volubility may be perceived by caregivers as an indicator of fitness, providing information regarding their infants' future potential (Locke, 2006) as well as influencing the parental investment in the infant. Research has demonstrated that infants who are highly voluble are more likely to receive care, engaging adults with greater success (see review in Locke, 2006). Additionally, infants who monitor adult reactions (see review in Chisolm, 2003) and adjust their vocal output accordingly receive greater levels of care and caregiver commitment. The ability to flexibly adjust vocal output and increase volubility appears to be present only in human infants.

In contrast, non-human primates appear to vocalize relatively inflexibly, and in the case of apes, at low levels of volubility compared to humans. However, little is actually known about the details of volubility in non-human primates. There appears to be better information about the usage of various vocal types in some primates, however. Each vocalization type is presumed to occur in a specific social or emotional circumstance (Hauser, 1996). As an example, the vocal type or category for threat is not used when there is no threat intended. Thus, there appears to be a relatively fixed one-to-one relationship between each sound and its function. How the vocalization sounds and the way it is used are both constant. The non-human primate does not appear to develop vocalizations capable of conveying multiple functions (Marler, Evans, & Hauser, 1992) although more recent research may provide evidence of context-specific call sequences (Ouattara, Lemasson, & Zuberbühler, 2009).

While volubility levels in non-human primates are reported to range from minimal in tree shrews (Benson, Binz, & Zimmerman, 1992) and apes, to apparently much higher levels in the pygmy marmoset (Elowson, Snowdon, & Lazaro-Perea, 1998), vocalizations appear to be relatively constrained in terms of the functions they can serve. For example, infant rhesus monkeys in a separated circumstance have been noted to display two different coos, however: 1) one when the infant was separated but in view of the mother and 2) the other when the infant was in isolation (Bayart, Hayashi, Faull, Barchus, & Levine, 1990). The infant rhesus monkey coos differed in intensity, and it was reasoned that the isolation coos, which were of higher intensity, were indicative of a greater urgency than the lower intensity coos indicating a need for reassurance or contact when the mother was in view. The difference in vocalization appeared to be dependent upon the circumstance, although both were instrumental, potentially requesting contact from their mother (Bayart et al., 1990). Bonobos, both infants and adults, also appear to use different intensities of hooting when in isolation versus within sight of others. It was found that infant bonobos did use different vocalization types from adults (de Waal, 1988).

It appears that even though volubility may differ greatly across species, both human and nonhuman primate infant vocalizations are salient to their respective parents (Bayart et al., 1990; Chisolm, 2003, Locke, 2006; Rheingold, Gerwitz, & Ross, 1959). It has been observed that individual human infant vocalizations tend to elicit immediate adult responses (Keller, Lohaus, Volker, Cappenberg, & Chasiotis, 1999; Goldstein, King, & West, 2003; Goldstein & West, 1999; Gros-Louis,West, Goldstein, & King 2006; Warlaumont, et al., 2010). Also, greater numbers of human infant vocalizations are associated with greater numbers of adult replies (Gilkerson & Richards, 2009; Goldstein et al., 2003; Gros-Louis et al., 2006). Prelinguistic vocalizations appear to provide a framework within which an interactive relationship can develop (Iyer & Oller, 2008;

Papoušek & Papoušek, 1989). Goldstein, Schwade, and Bornstein (2009) suggest that parental responses to infant vocalization likely assist in language acquisition. There is a substantial literature supporting this claim, invoking the notion that human infant emotions and intellect develop in large measure within the context of face-to-face vocal and affective interaction (Anderson, Vietze, & Dokecki, 1977; Bakeman & Adamson, 1984; Beebe, Jaffe, Feldstein, Mays, & Alson, 1985; Bornstein & Tamis-LeMonda 1989; Cohn & Tronick, 1987; Fogel & Garvey, 2007; Hsu & Fogel, 2003; Jaffe et al., 2001; Papoušek & Papoušek, 1979; Stern, 1974; Tronick, 1982 ). In the first months of life this interactive pattern is said to result in a capability for primary intersubjectivity (dyadic interaction) and then later in the first year, secondary intersubjectivity (triadic interaction incorporating joint attention) (Trevarthen, 1977, 1979).

It is also possible that spontaneous infant vocalizations in alone situations play a significant role in the development of speech and language (Berger & Cunningham, 1983; Delack, 1978; Delack & Fowlow, 1978; Jones & Moss, 1971). Infant vocalizations produced when alone have been interpreted as vocal play, self-stimulation activities (Dodd, 1972; Locke, 1989; Masur & Rodemaker, 1999) or as motor practice (Locke, 1989). Early infant vocalizations occur while their phonetic capabilities are in the process of developing and these early productions can only be produced with the motor ability that exists at the time (Locke, 1989). This sort of vocal play is reminiscent of apparent practice in limb movement that appears to contribute to acquisition of the ability to reach and grasp (Thelen, 1991). Early vocalizations in a non-social setting may be the rhythmic practice that permits vocal control to develop (Bickley, Lindblom, & Roug, 1986; Iverson, 2010; Iverson & Fagan, 2004). It has also been proposed that the frequency of vocalization may play a role in later speech sound production (Scherer, Williams, & Proctor-Williams, 2008), implying that increased volubility in the form of motor practice may be related to speech development.

Research in infant vocalizations has produced two conflicting points of view regarding the roles of the social or alone circumstances in infant volubility. On the one hand there are results advocating a strong role for social interaction, reporting that infants increase rates of vocalization in response to contingent vocalization from adults (Bloom & Esposito, 1975; Todd & Palmer, 1968). The assumption that social interaction stimulates infants to vocalize more also appears to be supported by the fact that both parents and infants of higher SES tend to vocalize more than parents and infants of lower SES (Hart & Risley, 1995; Oller, Eilers, Basinger, Steffens, & Urbano, 1995). Infants are thought to "bond" to parents and vice versa during face-to-face interactions occurring as early as two months of age, wherein vocalizations and smiling of both parties seem to stimulate interaction (Stern, Jaffe, Beebe, & Bennett, 1975; Trevarthen, 2001). Also infants who experience more vocalizations directed to them are reported to become more efficient in processing familiar words and to develop larger expressive vocabularies (Hart & Risley, 1995; Hoff, 2014; Pan, Rowe, Singer, & Snow, 2005; Rowe & Goldin-Meadow, 2009; Weisleder & Fernald, 2013).

On the other hand, there are advocates of the view that infant volubility may be more endogenously motivated and that infants vocalize most when alone (Delack & Fowlow, 1978; Jones & Moss, 1971; Yang, 2005; see review in Locke, 1993). In accord with this viewpoint, spontaneous, solitary, self-stimulating vocalization plays a critical role in vocal development. Since all this prior research has been done largely in lab or somewhat contrived home settings, all-day fully naturalistic recordings may illuminate the importance of both vocalization in social interaction and spontaneous vocalization produced when alone for the development of language.

It does appear that by five to six months of age infants appear to have learned that their vocalizations produce contingent responses from the adults in their environment (Goldstein et al., 2009). Furthermore, 6-month-old infants utilize sounds flexibly both alone (i.e., while mother talks with another adult) and in interaction with others (Oller, 2000; Oller et al., 2013).

Interestingly, research in infant vocal interaction has tended to categorize vocalizations in a very rough way, usually distinguishing only cry and non-cry (Bornstein et al., 1992; Camp, Burgess, Morgan, & Zerbe, 1987; Hsu et al., 2001). Consequently, there has not been the opportunity to determine if particular speech-like vocalization types are utilized in specific ways in interaction. Does volubility across vocal type vary with circumstance, revealing that the different speech-like sounds have different functions? Thus far, research provides only the most preliminary answer to this question (Franklin et al., in press; Papaeliou, Minadakis, & Cavouras 2002; Scheiner & Fischer, 2011; Scheiner, Hammerschmidt, Jürgens, & Zwirner, 2006; Stark et al., 1993).

Further research on infant volubility in varying naturalistic circumstances may be of additional benefit clinically and scientifically. As reviewed above, vocalization has been thought to be influential in establishing the parent-infant bond during the first year. Volubility levels also appear to be strongly related to rate of language development, and exhibit some important group differences, particularly with respect to SES (Bornstein & Bradley, 2003; Craig & Washington, 2005; Farah et al., 2008; Hart & Risley, 1995, 1999;

Hoff, 2006, 2003; Hoff, Laursen, & Tardif, 2002; Hoff & Tian, 2005; Kelly, 2011; Pungello, Iruka, Dotterer, Mills-Koonce, & Reznick, 2009; Oller et al., 1995). By estimating infant volubility across an ordinary day in the home environment, along with indications about the interactive circumstances that influence volubility, we may lay groundwork for later studies that may be able to identify patterns of parent-infant interaction that are most effective in supporting the acquisition of language. Additionally, by assessing the amount of spontaneous non-social infant vocalizations that occur, we may shed light on the role of endogenous vocal exploration in language acquisition.

### Volubility Measures in LENA (Language ENvironment Analysis) Studies

The LENA software system was created in response to research, which indicated that variations in children's language abilities are partially predicted by the amount parents speak to their children (Hart & Risley, 1995). The study of infant vocalization has changed dramatically with the advent of all-day recordings and automated analyses. The LENA system provides a way to move out of the lab and into the child's natural environment. The ability to view the infant in their "real-life" provides a contrast with the results from lab studies, making it possible to analyze children's vocalization patterns as they occur in a natural setting. This ease of recording may also increase participant pools (Marchman & Weisleder, 2011). Additionally, the LENA software and analyses system decreases the time spent collecting and analyzing conversational data in large quantities (Aragon & Yoshinaga-Itano, 2012; Montgomery, Gilkerson, Richards, & Xu, 2009; Oller et al., 2010). The LENA data have augmented the Hart and Risley findings that children who are exposed to more words in infancy have better language skills (Montgomery et al., 2009).

Infants are recorded in their home environment up to sixteen hours per day using the LENA recorder. After recording, the contents of the recorder are uploaded and processed through the language environment analysis software. The LENA software analyzes every millisecond (ms) of the audio file and produces core reports as well as segmented data for additional analysis. The core reports provide counts for Adult Words, Conversational Turns (when a child vocalizes and an adult vocalizes shortly thereafter, or an adult speaks and a child vocalizes shortly thereafter), Child Vocalizations, as well as normative comparison information, automatic vocalization assessments, developmental age, and a breakdown of the components in the audio environment.

The LENA "Child Vocalization Count" (Gilkerson & Richards, 2009) includes infant or child sounds that are identified by the automated algorithm as being "speech-related vocalizations" (see Oller et al., 2010). This designation excludes cries and vegetative sounds, which are also identified by the algorithm. The Child Vocalization Count provides LENA's estimate of infant volubility. However, that estimate is predictably lower than in the case of human coding. The reason is that the automated algorithm cannot identify speakers in overlapping voices, nor in cases of voices overlapping with noise. In these cases, LENA labels the sounds as "Overlap", and many infant and child sounds are thus left out in the speaker identifications of LENA (Xu, Yapanel, & Gray, 2009). This fact implies that the LENA algorithm is necessarily conservative with regard to volubility estimation.

In our human coding, child vocalizations that can be discerned despite overlap of other voices or noise are counted and included in volubility estimates. We presume that vocalizations in overlap may contribute to social interaction, and thus should be considered in our counts. Cry and laugh are also coded but are not included in counts contributing to volubility level. Our study will make it possible to make comparisons between the conservative volubility estimates of LENA, and the typical laboratory-based methods of estimation.

The LENA organization has reported volubility estimates by month of age from a huge sample involving hundreds of infants and many thousands of hours of recording (Gilkerson & Richards, 2008). There are 51 recordings for 6-month old infants, and the average volubility level according to the automated analysis was 1.29 vocalizations per minute.

The LENA algorithms represent an attempt to simulate human listener identification of speakers. The algorithms are trained to mimic the performance of human listeners, and the judgment of their performance is always made in terms of the extent to which they match human listener judgments. It is clear that, while the algorithms are extremely useful for rapid analyses of large databases of recordings that could not be coded practically by human listeners, they fall far short of human listener performance. The present work will provide a variety of measures of the extent of human to human and human to LENA agreement on coding of infant vocalization.

## Rationale

Our research focuses on infant volubility in the infant's natural home environment. The research seeks to estimate volubility with human coding and to compare that estimate with results obtained through the LENA method, the first fully automated, fully naturalistic approach to determining infant volubility. The work will make it possible to compare human coding with LENA analysis in terms of agreement of humans to humans and humans to LENA.

We also evaluate three different types of circumstances that may be influential on infant volubility. The first requires vocalization directed to the infant (VDI), including in this case an adult or other child vocalizing to or attempting to interact with the infant. The second is one where other people are speaking to each other but not to the infant (vocalization directed to other, VDO). The third circumstance is one where the infant is alone (infant alone, IA). We also identify segments of infant sleep in order to make it possible to exclude them and focus attention for each infant on wakeful periods.

A widely cited study (Delack, 1978; Delack & Fowlow, 1978) on infant volubility in different circumstances gave the impression that infants vocalized most when alone and without toys, with much lower rates of vocalization occurring in vocal interaction with either a parent or a stranger. The study was, however, quite unclear with regard to methods of determining circumstance and categorization of infant sounds. For example, were cries included in the volubility rate, and what did 'infant alone' actually mean? It is also unclear whether the infant spent a greater amount of time alone than with the mother or a stranger, in which case the larger number of vocalizations reported in the alone circumstance could merely be a reflection of the amount of time that was spent alone by the infants in this study. Many additional studies on vocal interaction yield findings that vary greatly, some suggesting high rates of vocalization when parents interact vocally with the baby (Todd & Palmer, 1968), and some that question that conclusion (Jones & Moss, 1968). Additional research is clearly in order regarding infant volubility in non-social (or alone) circumstances, as compared to circumstances of interaction. The key issue for us is that such research on infant volubility has never yet been conducted in a truly naturalistic setting without contrived circumstances or interference from experimenters. What is needed is random selection of samples for volubility measurement from naturalistic, all-day recordings where volubility in varying circumstances can be assessed as these circumstances occur freely in the home.

Volubility as a measure may provide a perspective regarding the nature of infants' motivation to vocalize and their expectations regarding other's vocalization. The volubility of infants in response to changes in circumstance is a key matter related to the degree to which infant vocalizations are endogenously produced as opposed to influenced by social interaction or other environmental factors (Locke, 1993). An infant's ability to respond systematically to vocalizations from caregivers may also be an important developmental achievement (Farah et al., 2008; Goldstein et al., 2009; Tamis-LeMonda, Bornstein, & Baumwell, 2001).

Prelinguistic vocal development appears to be regulated through infant-caregiver interaction and the caregiver's contingent responses to infant vocalization (Pelaez, Virues-Ortega, & Gewritz, 2011; Tamis-LeMonda et al., 2001). It is possible, however, that infants are driven to vocalize as a way to display health, thus encouraging parental care and engagement. Additionally, infants may vocalize spontaneously either as play or practice (Dodd, 1972; Masur & Rodemaker, 1999; Papoušek & Papoušek, 1989).

Since infant volubility measures have been gathered in a variety of circumstances, a comparison of naturalistic infant volubility across a day might provide insight into both the inherent motivation to develop language skills and the importance of parental input. An additional focus is to compare the human coding of six-month-old infants to the

LENA automated analysis. Automated measures of volubility would decrease the amount of time spent in coding and analyzing and may assist in gathering greater amounts of data. The ability to analyze long naturalistic recordings in a short amount of time could facilitate differentiating between typical and atypical volubility rates.

This dissertation consists of two studies. The primary goal of Study 1 was to estimate typical all-day levels of 6-month-old infant volubility in a naturalistic setting using human coding.

The research questions in Study 1 included:

1) What is a 6-month-old infant's volubility across a day in a naturalistic setting?

2) How do the LENA analysis results compare to human coding results for infant vocalization rates?

Study 2 was a comparison of 6-month-old infant volubility across variations in degrees to which vocalizations were directed to the infant (VDI), vocalizations were directed to others (VDO), and the infant was alone (IA). The research questions in the second study included:

1) Does infant volubility change across differing circumstances for segments selected at random across the day or for segments selected specifically to represent periods of high vocal activity?
2) Are there differences in vocal type that are related to different circumstances?
3) Do natural environmental variations influence infant volubility (TV/radio, traveling in the car or stroller, and being out of doors)?

# Methodology

## Participants

Ten (5 female, 5 male) 6-month-old infants participated in all-day LENA recordings in their home. One such all-day recording was selected for each infant to be analyzed in this work. The families of all ten infants were categorized as mid-SES based on mother's educational level (Hollingshead, 1978, 1975). Eight of the subjects were White Non-Hispanic and 2 were White Hispanic. Five of the families reported using both English and Spanish in the home setting, three of these in an effort to have their infants learn a second language.

Six of the infants were part of a larger longitudinal study while 4 had 2 all-day recordings during the infant's sixth month. For the 6 infants, many longitudinal LENA recordings were available, and one recording for each of them at 6 months with near 12 hours duration was selected. For the 4 infants not in the longitudinal study, the second of the two available recordings was chosen in each case to avoid any Hawthorne effect, which has been reported for first recordings using the LENA system. All 10 infants were considered typically developing and had not been diagnosed with any disorder. All 10 infants could be deemed mid to high SES based on mother's educational level.

# Procedure

## Study 1 Overview

The LENA system facilitates extraction of samples according to the user's needs for random or semi-random sampling, because every 5-minute period is directly accessible within the software for each recording. Random sampling provided a view of the entire day, and reduced coding cost. There were 110 – 169 5-minute segments available for the 10 infants (Table 1).

Table 1
*Participant Information*

| Participant | Sex | Total # of 5-minute Segments | Study 1-#sleep segments excluded | Race | Language spoken at home | SLP Mother |
|---|---|---|---|---|---|---|
| **1** | F | 146 | 57 | Caucasian | Eng | N |
| **2** | F | 137 | 41 | Caucasian | Eng/Span | Y |
| **3** | M | 169 | 44 | Caucasian | Eng | N |
| **4** | M | 110 | 17 | Caucasian | Eng/Span | Y |
| **5** | M | 135 | 30 | Caucasian | Eng | N |
| **6** | F | 164 | 82 | Caucasian | Eng | N |
| **7** | M | 156 | 51 | Hispanic | Eng/Span | Y |
| **8** | M | 160 | 24 | Hispanic | Eng/Span | N |
| **9** | F | 138 | 74 | Caucasian | Eng | N |
| **10** | F | 134 | 28 | Caucasian | Eng/Span/Ger | N |

For Study 1, twelve 5-minute periods of presumed awake time were extracted in a semi-random fashion from the all-day infant recordings. We excluded any segment occurring within a sequence of at least three consecutive 5-minute periods where the LENA system had found no infant vocalizations (Child Vocalization Count = 0), on the assumption that these were likely periods that the infant was asleep. The number of remaining segments was divided by 12 (number of segments to be extracted). This dividend was used to locate 12 equidistant 5-minute samples across the entire recording day, with Table 1 indicating the number of sleep segments excluded. This procedure allows for both semi-random selection and dispersal of sampling across the day. Segments with zero infant vocalizations according to LENA were included as long as they were not part of three consecutive 5-minute segments considered sleep time—there were 10 of these in the final selection. The coders (principal and reliability, PC and RC respectively) coded in repeat- listening mode (see explanation under coding).

**Study 2 Overview**

In the second study, infant vocalization rate was studied using the same recordings as in Study 1. A key focus here was differing circumstances of vocalization: infant alone (IA), vocalizations directed to the infant VDI, vocalizations directed toward others (VDO), or infant sleeping (IS). Twenty-four 5-minute segments were randomly selected from each infants' recording for a total of 240 segments. This set we refer to as the Randomly Selected sample (RS). The total number of segments from each infant's recording was divided by 24 (number of segments to be extracted) in order to locate 24 equidistant 5-minute samples across the entire recording day. Five primary coders (C1-5) were utilized with each coder assigned two infants, and coding was conducted in real-time (see below, coding). A reliability coder (RC) coded a subset of the data from all the coders.

An additional set of 100 segments was extracted consisting of the ten highest volubility segments from each infant's recording as determined by LENA automated coding (Child Vocalization Count), in order to make possible evaluation of periods of high vocal activity. This set we refer to as the High Volubility sample (HV). The same five coders utilized for the RS segments were again utilized with the same two infants assigned to each. These segments were then coded in real time identifying infant vocalizations.

In order to determine if the presumed sleep segments from Study 1 did in fact consist of sleep, ten 5-minute segments meeting the same requirements as in Study 1 (occurring within three consecutive segments where LENA labeled no infant vocalizations) were selected from each infant's recording (Low Volubility or LV, with a total of 100 segments). The total number of presumed sleep segments was divided by 10 (number of segments to be extracted) in order to locate 10 equidistant 5-minute samples across the designated sleep segments across the day. The same five coders utilized in the prior task were utilized with the same two infants assigned to each.

The total number of segments coded in Study 2 was thus 440. For each infant, the 24 RS segments, the 10 HV Segments, and the 10 LV segments were randomized. Each coder was presented with 44 randomly ordered segments from each of two infants, for a total of 88 5-minute segments per coder. Each set of 44 was then coded in real time identifying infant vocalizations. The coders were blinded as to the reason for the study.

Upon completion of coding of each 5-minute segment in Study 2, each coder answered a set of questions concerning the circumstance present within the segment. The questions were:

1. Does any other person talk to the baby?

2. Does any other person talk to someone else?

3. Do you think any other speaker besides the baby is in the same room with the baby?

4. Do you think the baby is asleep?

5. Are any baby vocalizations clearly audible?

6. Do you think another person whose voice you hear is next to or holding

the baby?

7. Do you think the baby nursing or being fed?

8. Do you think the TV or radio is on?

9. Do you think the TV or radio is in the same room?

10. Is there considerable noise throughout the segment besides TV or radio as for example from toys, slamming doors, construction equipment…?

11. Do you think the recording is occurring outside the home (for example at the grocery store, or in a daycare center)?

12. Do you think the recording is occurring while driving in a car?

The questions were answered using a scaled response from 1 to 5. Five represented the entire time (e.g., adult to adult talk all throughout the 5 minutes), and 1 represented total lack of occurrence of the circumstance during the five minutes. For questions 1 and 2 the scale was as follows; 1 = No adult vocalizations present, 2 = Adult interaction occurs less than half the time, 3 = Adult interactions occur about half the time, 4. Adult interactions occur more than half the time, and 5 = Adult interaction occurs throughout. The coders were asked to attend to the amount of time that the adult was engaged in vocal interaction or attempted vocal interaction, not the amount time actually spent in vocalizing (the focus thus was on the amount of conversation time, not the amount of vocalization per se). The first three questions did not require focus on infant vocalizations, but rather on the vocalization environment of the child.

For question 5 (infant vocalization audibility) the scale was as follows; 1 = No vocalizations, 2 = Short and low (sounds produced at very low intensity or with very short duration, < 100 ms) vocalizations only, none coded, 3 = some vocalizations are audible just above the short and low threshold, 4 = about half well above the threshold, and 5 = lots of clearly audible (well above the threshold) vocalizations. For questions 4 and 6 through 12 (environmental) the following scale was utilized; 1= Never, 2 = Less than half the time, 3 = About half the time, 4 = More than half the time, and 5 = Close to the whole time. Volubility levels were then calculated by circumstance, VDI, VDO, and IA. Circumstances were generally collapsed into three categories for analysis: 1. Less than 25% of the time, 2. 25-50% of the time, and 3. More than 50% of the time.

**Differences between Study 1 and Study 2**

Differences in the methodologies of Study 1 and Study 2 include the following. 1) In Study 1 human coding was based on repeat listening (starting and stopping the recording and listening repeatedly to code each utterance) which allows for specification of onset and offset of vocalizations. In Study 2, real-time coding was utilized with each utterance judged by a single keystroke as it occurred during playback. Thus, no value for onset or offset or duration of vocalizations was obtained. 2) In Study 1 vocalizations that were deemed very low in intensity or very short in duration were included in the human coding. Such vocalizations were often < 50 ms in duration and only noticeable when the listener was very attentive. In Study 2, these "short and low" vocalizations were *not* coded, partly on the assumption that real-time listening should simulate natural parent listening, where it would seem such vocalizations would be ignored. In addition, real-time listening is more compatible with a strategy where very short or very low intensity events are ignored. 3) In Study 1 there was single primary human coder for all the segments, while in Study 2 there were five human coders (C1-5), each assigned to code the data from two of the infants. 4) The most comparable data across studies were based on the randomly selected segments from both—the *N* for Study 1 was 120 segments, 12 per infant, while for Study 2 it was 240 segments, 24 per infant. 5) In Study 1 we attempted to limit our focus to segments where the infant was asleep by eliminating from the possibility of random selection any segment that the LENA analysis showed to be part of three consecutive 5-minute segments with zero Child Vocalization Count. In Study 2 on the other hand we eliminated sleep segments based on coder judgments about infant sleep (questionnaire item four). And 6) Study 2 included two sampling types not present in Study 1, the 100 High Volubility segments (10 per infant) and the 100 Low Volubility segments.

These differences were partly the result of the way the work was developed. Study 1 was conducted first and provided an initial view of naturalistic volubility levels based on random sampling and human coding to compare with LENA derived values. Study 2 was designed afterwards and took advantage of the experience of coding and results in Study 1. But perhaps more important, a second study offered the opportunity to address the issue of possible circumstance variations. In addition we increased the sample size seeking greater stability of data, and added the HV and LV samples to provide additional perspectives. We shifted to real-time coding because of the substantial increase in the amount of coding that would be required in Study 2. Prior informal efforts had suggested we would lose little if any accuracy of coding using the real-time method.

**Coding for both Study 1 and Study 2**

The coding for both studies was completed off-line, using the audio recordings. Speech-like vocalizations (such as speech or protophones, i.e., the presumed precursors to speech), were coded, including: 1) vocants, either fully resonant  or quasi resonant vowel- like sounds, 2) squeals, sounds of higher than typical pitch, and 3) growls, sounds of lower than typical pitch or having a raucous vocal quality. Cries and laughs were also coded and only considered for certain aspects of the analysis.

Coding was performed using the Action Analysis Coding and Training software (AACT, 1996). For Study 1 the AACT system was used to code the vocalizations auditorily by protophone type, to review spectrographic displays of the audio recordings, and to mark onset and offset of each utterance produced by the infants during the 5-minute segments (providing a duration measure).

For Study 2 the AACT system was used for real-time coding of the same categories used in Study 1 with a larger sample and expanded segment selection procedures. Real-time coding was completed by listening to the digital recording and selecting the vocal type as the recording was played, pressing a particular key each time one of the five vocalization types occurred. There was no onset or offset recorded. The time recorded was determined by when the coder selected the vocalization type.

**Inter-rater Agreement**

**Study 1**, To assess coding reliability in Study 1, the RC categorized 17% of the Randomly Selected segments (two segments per infant). This included the re-coding of two 5-minute periods from each of the 10 participants (100 minutes overall). The 5-minute segments were chosen using a random number assignment for each of the participant's segments. Agreement was measured by comparing the number and type of infant vocalizations categorized as well as an event-based point-to-point analysis allowing for the use of the kappa statistic. The kappa statistic takes into account any agreement occurring by chance and is generally a more appropriate measure than simple percent agreement because it adjusts for imbalances in the base rate of occurrence of categories.

The correlations between volubility for the two coders were as follows (Figure 1): 1) for data collapsed across the two segments for each infant: Principal Coder or (PC) to Reliability Coder (RC), $r = .98$, $p < .001$, $n = 10$, $r = .96$, $p < .001$, $n = 20$. On the point-to-point comparison, if both coders determined there was a vocalization present within plus or minus one second, it was considered agreement; in contrast, if either the PC or RC coded a vocalization when the other did not, it was considered a disagreement. Additionally, if both coders determined there was no vocalization present within plus or minus one second of any coded vocalization, this was considered agreement (but only one agreement up to the next coded vocalization). The two coders' agreement on vocalizations present or not present was at 80% overall (kappa = .60, $p <.001$). See Appendix A for scatterplots of inter-rater reliability for Study 1.

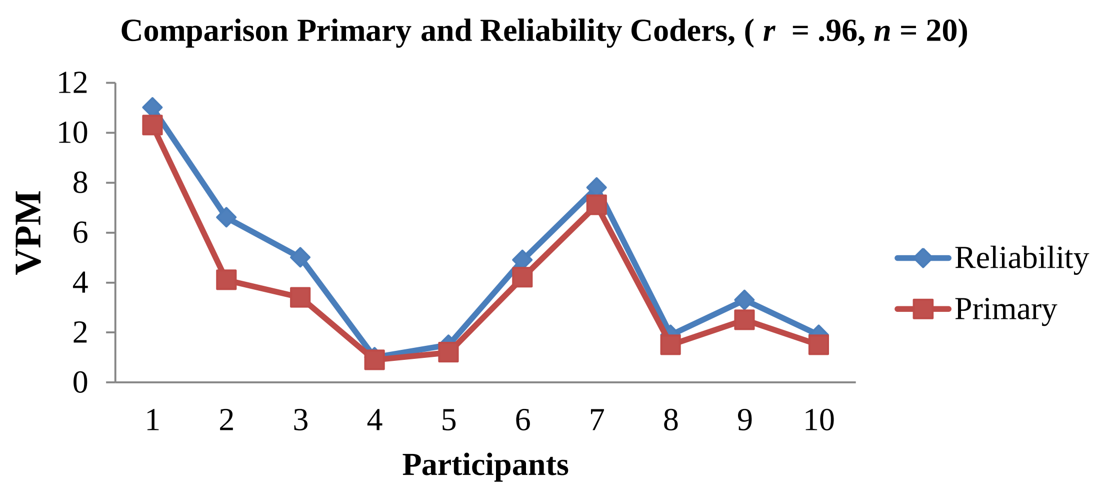



*Figure* 1**.** Comparison of PC and RC infant vocalizations per minute (VPM) by participant for two 5-min segments per participant.

The same reliability coding was used to determine agreement on infant vocalization type (vocant, squeal, and growl) in Study 1. Each coder was given a forced choice of vocant, squeal or growl, so a three by three comparison was made. Vocalizations where there was no corresponding code from the opposite coder were not included in this analysis. The two coders' overall agreement was 92% (kappa = 0.64, $p < .05$). Consistent with prior data (Oller; 2000, Stark et al.,1993), vocants were the most common codes assigned, and the higher base rate for vocants helps explain the discrepancy between the percent correct and kappa values.

**Study 2**: To assess reliability of the coding within Study 2, the RC coded 30% of the Randomly Selected segments. See Appendix B for scatterplots of inter-rater reliability for Study 2. This included approximately 7 5-minute segments from each of the 10infants. The 5-minute segments for the RC were chosen at random from the 24 segments for each infant participant. Agreement was measured by comparing volubility at the session level as well as in point-to-point analysis allowing for the use of the kappa statistic. Session-level agreement on volubility was indicated by very high correlations (mean = .96) between the values obtained by the 5 coders and the RC (see Table 2).

The point-to-point agreement on volubility (presence or absence of an infant vocalization at each point in time) between the five coders' and the RC's coding was completed by using a window of ± one second between potential vocalization code points. The presence of a vocalization for one coder along with the absence of a vocalization marked by the other coder within the one-second window was deemed a lack of agreement for that point. The five coders' overall agreement (point-to-point agreement on presence or absence of a vocalization) with the RC was 76% (Cohen's kappa= 0.52).

Table 2
*Inter-rater Agreement on the presence or absence of infant vocalization*

| | Vocalizations present or not present | | |
|---|---|---|---|
| | Correlation *n*=14 | Kappa | Percent Agreement |
| C1 to RC | .97 | 0.57 ($p < .001$) | 79% |
| C2 to RC | .95 | 0.72 ($p < .001$) | 86% |
| C3 to RC | .97 | 0.61 ($p < .001$) | 81% |
| C4 to RC | .98 | 0.42 ($p < .001$) | 71% |
| C5 to RC | .95 | 0.27 ($p < .001$) | 62% |
| ***Mean*** | ***.96*** | ***.52*** | ***76%*** |

The coders and the RC also coded for infant vocalization type (vocant, squeal, growl, cry and laugh). The coders' overall agreement with the RC was 79% (kappa = 0.45). While vocant agreement was high, there were small numbers of squeals, growls, cries, and laughs.

Table 3
*Inter-rater kappa and agreement on infant vocal type*

| | Vocal Type Inter-rater Agreement | |
|---|---|---|
| | Kappa | Percent Agreement |
| C1 to RC | 0.47 | 88% |
| C2 to RC | 0.53 | 81% |
| C3 to RC | 0.65 | 82% |
| C4 to RC | 0.26 | 67% |
| C5 to RC | 0.36 | 77% |
| **Mean** | **0.45** | **79%** |

Additionally, a comparison of circumstance agreement was made both at the session level with correlation and in percent agreement. We assessed exact circumstance agreement (both coders recorded exactly the same number from 1 to 5 for a question) as well as ± 1 level of agreement (the numbers assigned by the two coders differed by 1) on circumstance. For each of the twelve survey questions between the five coders and the RC the results were as follows:

Table 4

*Circumstance agreement (for judgments rated from 1 to 5), exact and within ±1 between the coders and the RC.*

| Question | Exact agreement for 140 segments (proportion) | Agreement for 140 segments to within ±1 (proportion) | Correlation between RC and C1-C5 for 140 segments |
|---|---|---|---|
| 1 | 0.70 | 0.92 | .86 ($p$ < .001) |
| 2 | 0.76 | 0.92 | .86 ($p$ < .001) |
| 3 | 0.82 | 0.92 | .84 ($p$ < .001) |
| 4 | 0.83 | 0.89 | .79 ($p$ < .001) |
| 5 | 0.46 | 0.94 | .83 ($p$ < .001) |
| 6 | 0.61 | 0.70 | .53 ($p$ < .001) |
| 7 | 0.92 | 0.94 | .68 ($p$ < .001) |
| 8 | 0.79 | 0.82 | .72 ($p$ < .001) |
| 9 | 0.76 | 0.79 | .56 ($p$ < .001) |
| 10 | 0.48 | 0.66 | .26 ($p$ < .001) |
| 11 | 0.94 | 0.97 | .92 ($p$ < .001) |
| 12 | 0.96 | 0.97 | .88 ($p$ < .001) |

*Note.* The coders' mean agreement across the 12 questions within ±1 with the RC was 87%.

## Results

### Study 1

The 6-month-old infant volubility level estimated using the LENA system was 2.4 vocalizations per minute (VPM) for the 120 five-min segments selected semi-randomly for Study 1. The infant volubility levels estimated from the 120 extracted human coded samples was 4.7 vocalizations per minute, nearly twice as high as the LENA estimate. In all ten cases the human coding produced higher volubility levels (see Figure 2).

Human coding of the 12 5-minute segments correlated with the LENA results for the same segments at $r = .85$, $n = 10$ ($r = .68$, $n = 120$). Comparisons of each participant's vocalizations per minute are displayed in Figure 2 for the PC and LENA analyses for the extracted segments and for the all-day LENA analysis. The human coding for each of the infants (red) yielded a greater number of vocalizations per minute when compared to either the LENA automated analysis of the extracted segments for the sample (blue) or the all-day automated analysis (black).

In a LENA to LENA comparison, volubility estimated per infant for Study 1's extracted 5-minute segments (which excluded segments designated by the LENA analysis as occurring as part of any consecutive three segments with zero Child Vocalization Count) and the estimate from the all-day recording results per infant (with the same zero segments excluded, but hundreds of additional non-zero segments included) were highly correlated ($r = .93$, $p < .001$, $n = 10$) and were not significantly different ($t(9) = 1.126$, $p = .269$). These indicate that the extracted segments which make up the sample were representative of the all-day recording according to LENA.

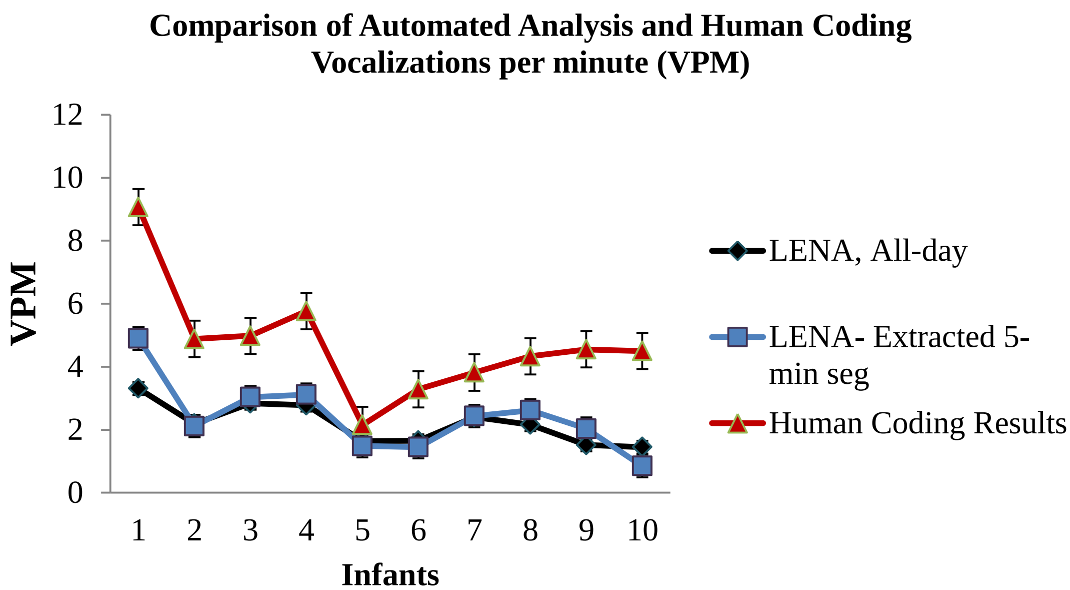



*Figure* 2. Comparison of Automated and Human Coding by Infant

The human coder values and the LENA values for the extracted segments were significantly correlated ($r$ = .85, $p$ < .001, $n$ = 10; $r$ = .68, $p$ < .001, $n$ = 120) and significantly different ($t$(9) = - 7.04 , $p$ < .001). The human coded volubility and the all-day automated results were also highly correlated ($r$ = .76, $p$ < .05) and also significantly different ($t$(9) = - 5.75, $p$ < .001).

The durations of vocalizations per minute averaged across segments are presented by participant in Table 5. As in the volubility analysis, the durations of vocalizations based on human coding were greater than those produced by the automated LENA analysis. The durations for the human coded 5-minute segments and the LENA durations for the 5-minute segments were highly correlated ($r$ = .53, $p$ < .001, $n$ = 120, $r$ = .86, $p$ = .001, $n$ = 10), and the LENA values were significantly lower than those of the human coder ($t$(119) = -7.18, $p$ < .001). Thus the LENA-estimated volubility was a little more than half that of the human coder value (see above), while the total *duration* of infant vocalizations per minute estimated by LENA was *less* than half the estimate based on human coding. LENA's 5-min sample results also significantly and very highly correlated with the all-day LENA automated results for duration ($r = .93, p < .001, n = 10$; $t(9) = 1.42$, $p = .19$).

Table 5

*Comparison of LENA analysis and human coding for duration of voc/min.*

| | **Mean PC human, (*SD*)** | **Mean LENA, extracted segments (*SD*)** | **Mean LENA All-day (*SD*)** |
|---|---|---|---|
| 1 | 9.44 (3.96) | 3.81 (2.71) | 2.69 (2.84) |
| 2 | 5.81 (5.49) | 1.82 (2.03) | 1.92 (1.95) |
| 3 | 4.50 (5.10) | 2.38 (2.37) | 2.24 (2.53) |
| 4 | 5.21 (6.13) | 2.42 (2.42) | 2.09 (2.31) |
| 5 | 1.47 (2.02) | 1.10 (0.93) | 1.34 (1.58) |
| 6 | 2.30 (1.91) | 1.31 (1.42) | 1.30 (1.54) |
| 7 | 7.59 (6.46) | 3.05 (4.81) | 2.55 (3/09) |
| 8 | 3.11 (3.32) | 1.79 (1.62) | 1.65 (1.89) |
| 9 | 4.53 (5.45) | 1.61 (2.84) | 1.05 (1.53) |
| 10 | 3.98 (4.29) | 0.69 (0.81) | 1.13 (1.68) |
| **Overall** | 4.80 (4.41) | 2.00 (2.20) | 1.80 (2.09) |

*Note*. For columns 1 and 2, the values represent the mean of the sums of durations per minute for segments coded (12 per infant). For column 3, the values represent the mean of the sums of durations for all segments from the day aside from those excluded as potential sleep time (see above).

**Study 2 Analysis**

**Agreement on volubility coding across Study 1 and Study 2.** The most useful measure we know of for agreement between human coders across Study 1 and Study 2 is based on the 23 segments that were the very same ones in the two cases, having been fortuitously selected at random both times. The correlation between the volubility values obtained in the two cases was $r = .78$, $p < .001$, $n = 23$, with one very salient outlier point—after removal of that point the correlation between values obtained across the studies was $r = .94$, $p < .001$, $n = 22$. These correlations provide a rough estimate of the agreement between two different methods of coding for determination of volubility: repeat listening and real-time coding. Two additional factors could have contributed to any lack of agreement, viz., the fact that different coders were involved in the two studies and that short and low vocalizations were included in one study but not the other.

**Volubility level compared with Study 1**. The infant volubility estimated by the human coders in Study 2 was 3.3 vocalizations per minute (Figure 3). This value differs notably from that estimated in Study 1 (4.7), but the reasons are likely due to a sampling factor. As a result of the differing random selection methods, there were more than four times as many segments with zero infant vocalizations as determined by the human coders in Study 2 as in Study 1. This difference was presumably caused by the criterion in Study 1 eliminating any segment from possible random selection if it was one of three consecutive segments with a zero value on Child Vocalization Count according to the LENA analysis. This procedure was intended in Study 1 to eliminate sleep segments and appears to have eliminated many zero segments where sleep was not involved. The result was a small number of segments where the human coder found zero vocalizations (10 of 120 = .08). In contrast, Study 2's selection procedure was truly random, and sleep segments were empirically determined and eliminated subsequently from the calculation of volubility. Still infants appeared to be silent often, even when not asleep, and this fact was reflected strongly in Study 2, with a relatively large number of zero's (79 out of 240 = .33).

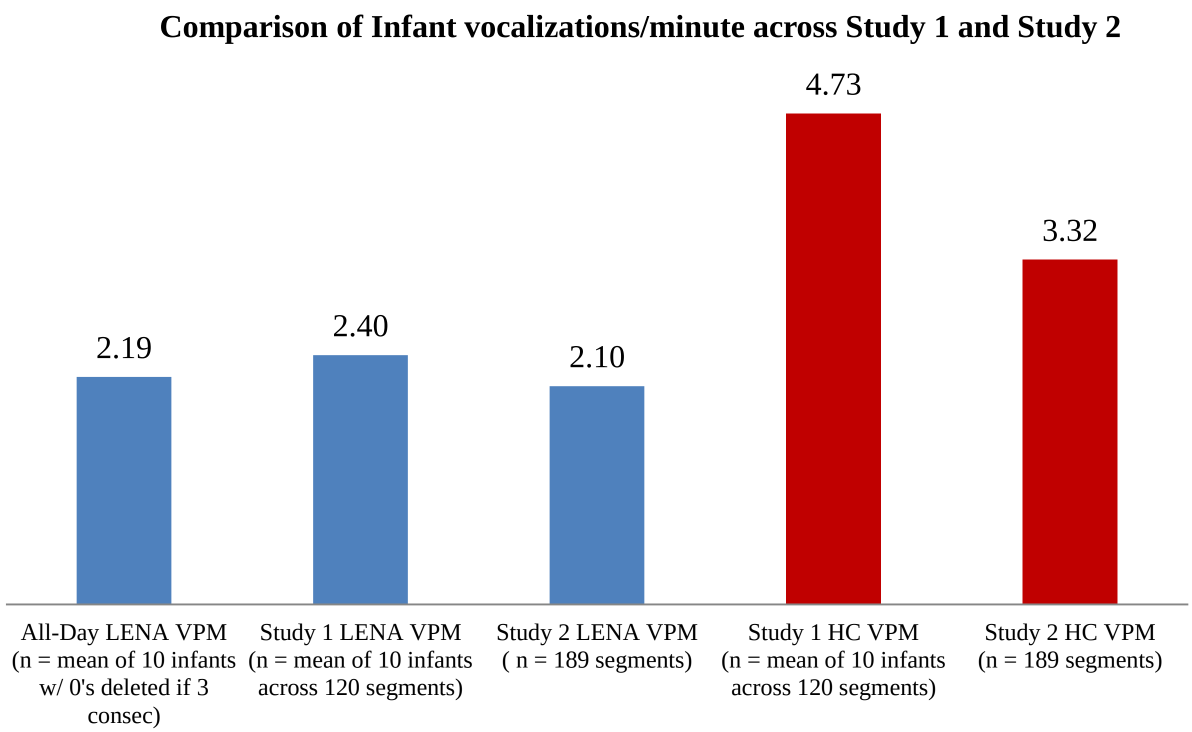
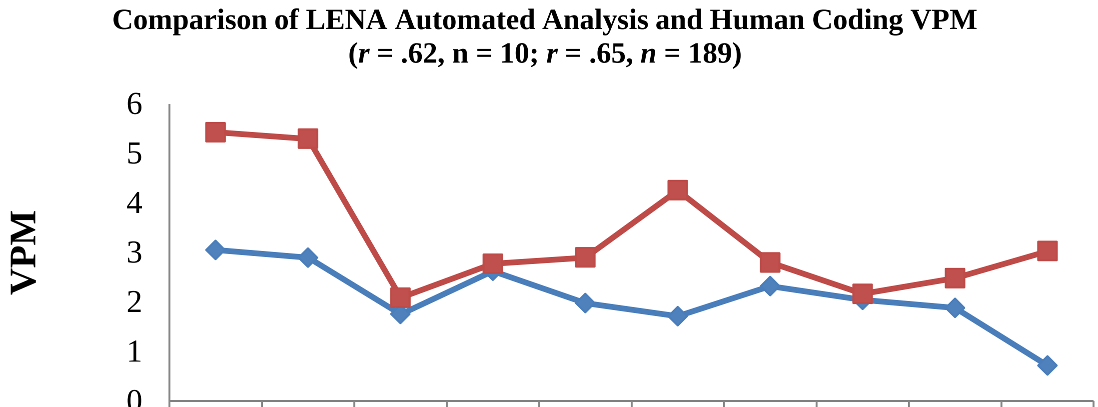



*Figure* 3. Comparison of volubility levels (VPM) across the studies and methods of estimation.

**LENA to human coding comparison.** As in Study 1, the data for Study 2 showed considerably higher vocalization rate as estimated by the human coders compared to LENA. Again for each of the infants, volubility was found to be higher with human coding than with the LENA analysis (Figure 4).




| | 1 (n = 20) | 2 (n =21) | 3 (n =18) | 4 (n = 21) | 5 (n = 18) | 6 (n = 14) | 7 (n =15) | 8 (n =23) | 9 (n = 17) | 10 (n =22) |
|---|---|---|---|---|---|---|---|---|---|---|
| LENA | 3.05 | 2.90 | 1.76 | 2.63 | 1.98 | 1.71 | 2.32 | 2.04 | 1.88 | 0.72 |
| C1-5 | 5.43 | 5.30 | 2.09 | 2.77 | 2.90 | 4.26 | 2.8 | 2.17 | 2.48 | 3.03 |

*Figure* 4. Comparison of Automated and Human Coding by Infant (sleep excluded).

A comparison of the LENA automated analysis results for volubility per Randomly Selected 5-minute segment and the human coding of volubility on the same 5-minute segments in Study 2 revealed that the five coders (each coding two infants) and the RC had similar patterns of correlation with the LENA results, and perhaps more importantly tended to agree with each other to a much greater extent than any of them agreed with the LENA results. To illustrate this point, the data from Table 5 have been presented here again in Figure 5. In addition the 5 human coders and the RC showed a remarkably similar pattern of relative correlations across the five infant pairs, with identical rank orders, suggesting again that the human coders had very high agreement with each other.

The correlation between the LENA and human coding values for the Randomly Selected segments Study 2 was lower ($r$ = .65 for $n$ =189; $r$ = .62 for $n$ = 10) than in Study 1 ($r$ = .69 for $n$ = 120; $r$ = .85 for $n$ = 10). This discrepancy could be due to procedural differences or could be the result of differences in the individual segments selected at random in both studies. Repeat listening coding (Study 1) vs. real-time coding (Study 2) is one possible source of the differences. The fact that a single coder produced the human data in Study 1 while 5 different coders were involved in Study 2 is another. The striking difference in number of segments yielding zero volubility across the studies could also have played a role, and of course a combination of these possible factors could have contributed.

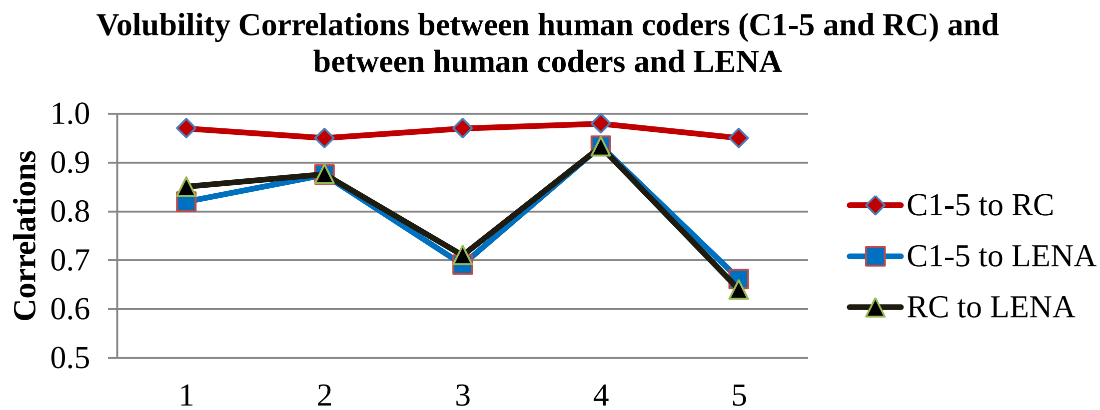



*Figure* 5. Correlations between all the human coders (C1-5 and RC) and LENA results for the 24 segments per infant (48 per coder) on infant volubility, $n = 48$.

*Note*. The red line reflects the correlations between the five human coders C1-5) and the reliability coder (RC). The blue reflects the correlations between the five human coders (C1-5) and the LENA estimates for the same segments. The black line is the correlation between the RC and LENA using the same breakdown of data (two infants per human coder) across the same segments.

**Circumstance Analyses for Study 2**

For the Randomly Selected samples, data on each of the circumstance parameters (VDI, VDO, and IA) was statistically analyzed by collapsing the five degrees of each circumstance parameter into three levels. This was done to increase the Ns for both infants and segments within circumstance categories allowing for more robust statistical comparisons. In the VDI and VDO circumstances, level 1 = 100% and > 50%, level 2 = ~50% and < 50%, and level 3 = 0%. In the IA circumstances level 1 = 0% and < 50% of the time infants were alone, level 2 = ~50% and > 50% of the time infants were alone, and level 3 = 100% of the time infants were alone within the segment. For descriptive clarity, both the five levels for each circumstance parameter as well as the three collapsed categories used for statistical analysis are presented graphically below.

**Randomly Selected Segments Where Vocalizations were Directed to the Infant to Varying Degrees (VDI).** A two-way ANOVA was conducted that examined the effects of three levels of VDI (fixed effect) and infants (random effect) on volubility. There was a significant main effect of the differing levels of VDI ($F(2, 9) = 7.68$, $p = .003$, $\eta_p^2 = .92$) as well as a significant interaction between the effects of VDI and infant ($F(2, 9) = 2.04$, $p = .01$, $\eta_p^2 = .96$). Levene's test indicated the assumption of homogeneity was violated, $F(27, 161) = 2.66$, $p < .001$. Because of the violation of homogeneity, a more stringent than usual alpha level was used (.01 versus .05). Tukey's post hoc tests for the VDI effect revealed that level 3 (0% VDI) showed significantly lower volubility than either level 1 or 2. It can be concluded that VDI was related to the amount of infant volubility at 6 months, since segments with vocalizations directed towards infants showed higher volubility in the 6-month olds when compared to the circumstance with no vocalizations directed to the infant. The significant interaction indicates that the levels of VDI affected infant volubility in different degrees for different infants. The data used for the statistical analysis are displayed in Figure 6.

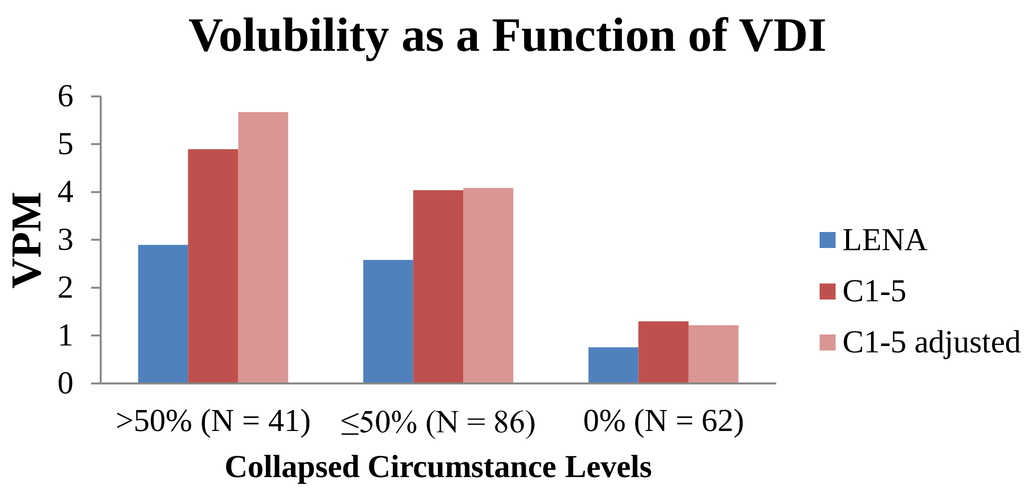
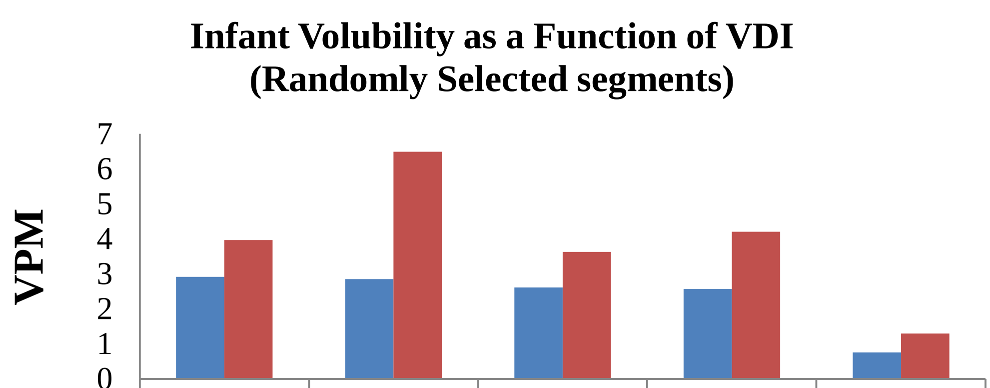



*Figure* 6. Infant volubility by degree of VDI collapsed into three levels as described in the text for Randomly Selected segments.

*Note*. LENA and C1-5 values were computed at the segment level with *N*'s as indicated in the figure. The C1-5 adjusted values are based on estimated marginal means computed by SPSS in the *F*-test.

The data for all five levels of VDI are displayed in in Figure 7 and include the LENA automated estimates as well.




| | 100% (N= 26) | >50% (N = 15) | ~50% (N = 25) | <50% (N = 61) | 0% (N = 62) |
|---|---|---|---|---|---|
| LENA | 2.92 | 2.85 | 2.62 | 2.57 | 0.75 |
| C1 - 5 | 3.97 | 6.49 | 3.62 | 4.20 | 1.30 |

*Figure* 7. Infant volubility by VDI (Randomly Selected segments) for all 5 levels of VDI

**Randomly Selected Segments Where Vocalizations were Directed to Others to Varying Degrees (VDO).** A two-way ANOVA was conducted that examined the effect of three levels of VDO (fixed effect) and infants (random effect) on volubility. There was a significant main effect of the differing levels of VDO ($F(2, 9) = 6.74$, $p = .003$, $\eta_p^2 = .89$) for the 189 Randomly Selected segments (excluding all sleep segments). Levene's test indicated the assumption of homogeneity was violated, $F(27, 161) = 3.10$, $p < .001$. Because of the violation of homogeneity, a more stringent alpha level was used (.01 versus .05). Tukey's post hoc testing revealed significantly higher volubility in level 2 compared to both level 1 and level 3. This indicates that VDO was strongly related to infant volubility. The data used for the statistical analysis are displayed in Figure 8.

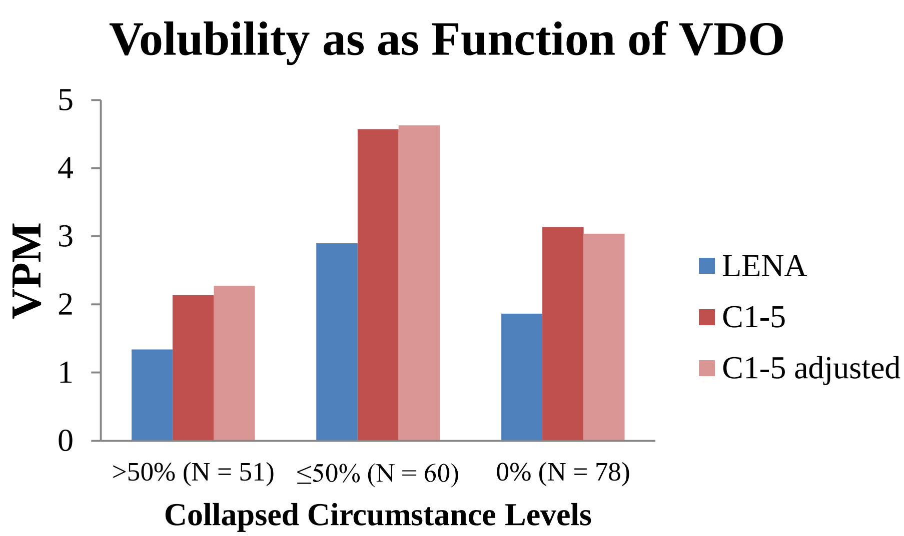



*Figure* 8**.** Infant volubility by VDO collapsed into three levels as described in the text for Randomly Selected samples.

*Note*. LENA and C1-5 values were computed at the segment level with *N*'s as indicated in the figure. The C1-5 adjusted values are based on estimated marginal means computed by SPSS in the *F*-test.

The data for all five levels of VDO are displayed in in Figure 9 and include the LENA automated estimates as well.

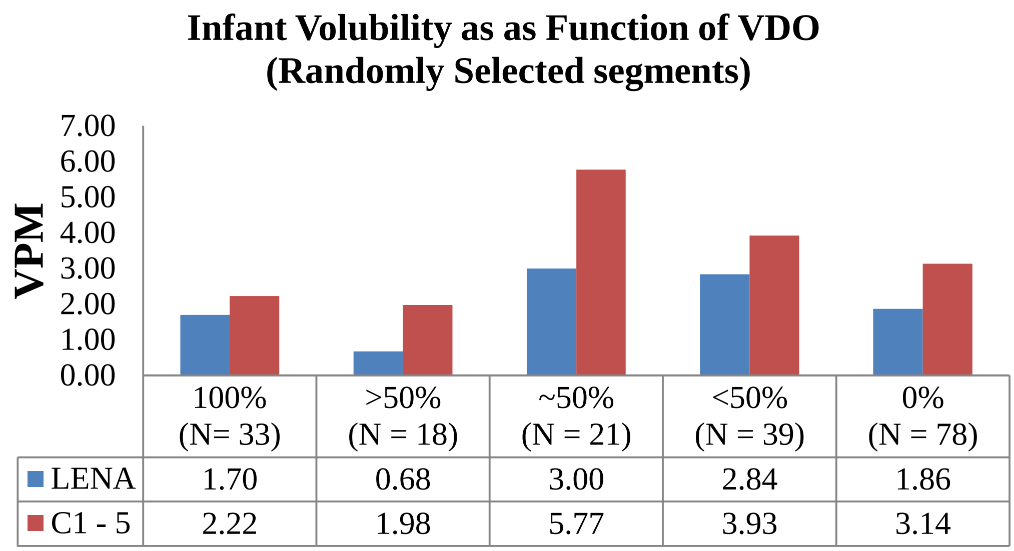



*Figure* 9. Infant volubility by VDO (Randomly Selected segments) for all 5 levels

**Randomly Selected Segments When the Infant was Alone to Varying Degrees (IA).** A two-way ANOVA was conducted that examined the effect of three levels (level 1 = 0% and < 50% of the time infants were alone, level 2 = ~50% and > 50% of the time infants were alone, and level 3 = 100% alone) of IA (fixed effect) and infants (random effect) on volubility. There was a significant main effect of differing levels of IA ($F(2, 9) = 4.67$, $p = .02$, $\eta_p^2 = .30$ ) as well as a significant interaction between the effects of IA and infant on volubility ($F(2, 9) = 1.80$, $p = .04$, $\eta_p^2 = .15$). The significant interaction indicated that the levels of IA did not affect infant volubility in the same way for all of infants in the study. Levene's test indicated the assumption of homogeneity was violated, $F(27, 161) = 2.19$, $p = .001$. Because of the violation of homogeneity, a more stringent alpha level was in order (.01 vs, 05); under the more stringent alpha level, both the main effect of IA and the interaction were only marginally significant. Tukey's post hoc testing was performed to test for significant differences between the levels. There was significantly lower volubility in the 100% alone (level 3) than in either other level. Infant volubility levels varied with differences in the percent time the infant spent alone as seen in Figure 10.

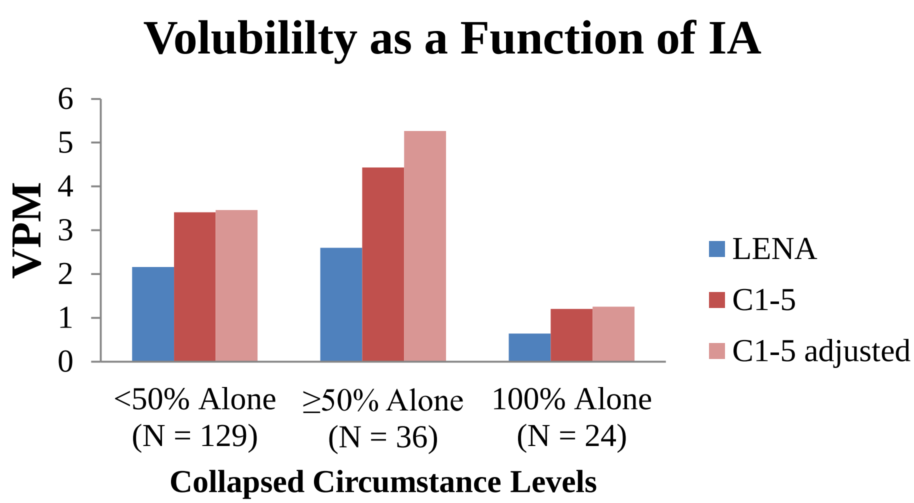



*Figure* 10. IA collapsed into three levels, as indicated in the text.

*Note.* LENA and C1-5 values were computed at the segment level with *N*'s as indicated in the figure. The C1-5 adjusted values are based on estimated marginal means computed by SPSS in the *F*-test.

The data for all five levels of IA are displayed in Figure 11 and include the LENA automated estimates as well.

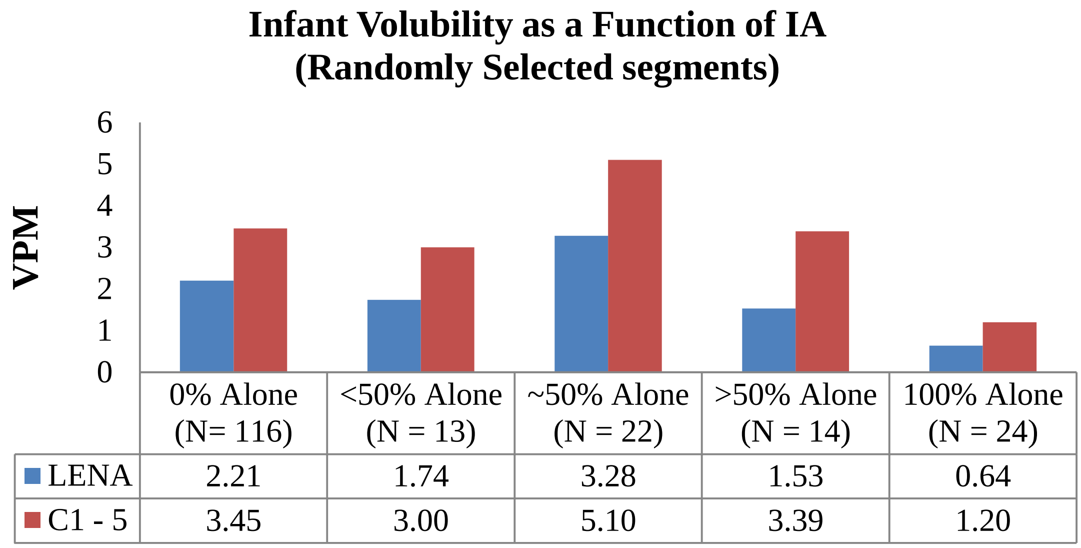
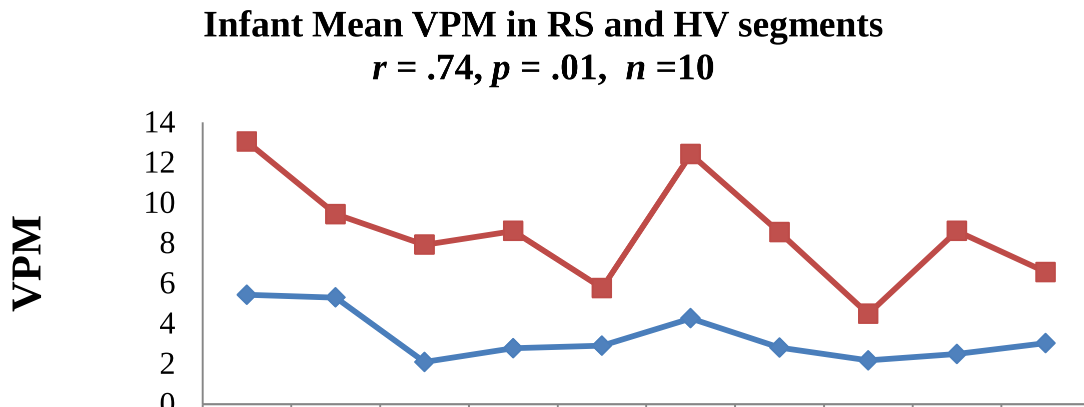



*Figure* 11. Infant volubility by IA (Randomly Selected segments) for all 5 levels of IA.

**High Volubility (HV) segments, Overall comparisons.** The volubility level in the High Volubility sample was higher, of course, than in the Randomly Selected (HV overall mean VPM = 8.6, RS overall mean VPM excluding sleep = 3.3), and the increase from LENA to human coded values was proportionally smaller for HV (46%) than for RS (58%). The correlation between LENA and human coding in the HV segments was also considerably lower than in the case of the RS segments, presumably in part because the range of values was lower in the HV segments ($r$ = .74, $p$ = .01, $n$ =10; $t(9)$ = 8.45, $p$ < .001).

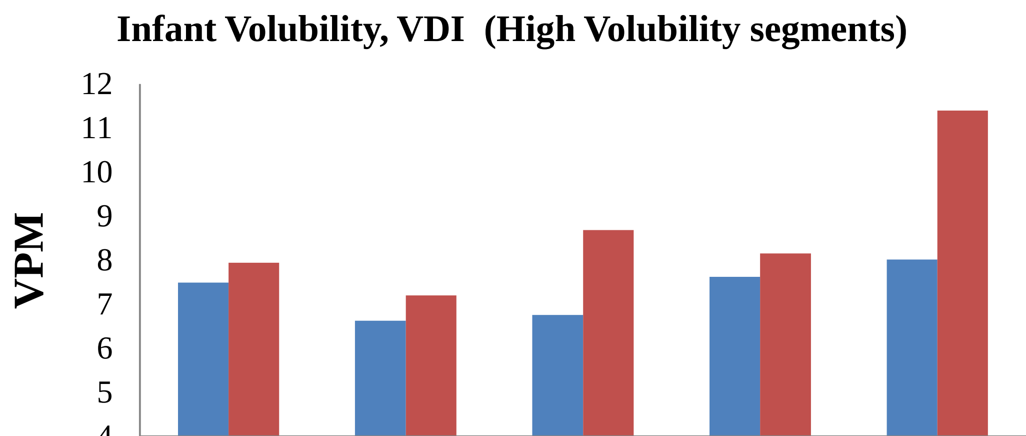



| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| C1-5 RS | 5.43 | 5.30 | 2.09 | 2.77 | 2.90 | 4.26 | 2.80 | 2.17 | 2.48 | 3.03 |
| C1-5 HV | 13.04 | 9.44 | 7.92 | 8.6 | 5.76 | 12.42 | 8.54 | 4.48 | 8.6 | 6.56 |

*Figure* 12. Comparisons of Infant mean VPM from both Randomly Selected and High Volubility segments

**High Volubility Segments for Different Levels of Vocalizations Directed to the Infant (VDI).** The High Volubility segments were distributed across infants and circumstances quite unevenly and in some levels of circumstance the total number of segments was small or included only a few of the infants. As a result it was not deemed appropriate to conduct ANOVA evaluations as had been done for the Randomly Selected segments. Instead data are displayed below for descriptive purposes only in figures for all five levels of each circumstance parameter along with the LENA automated estimates for the same segments.

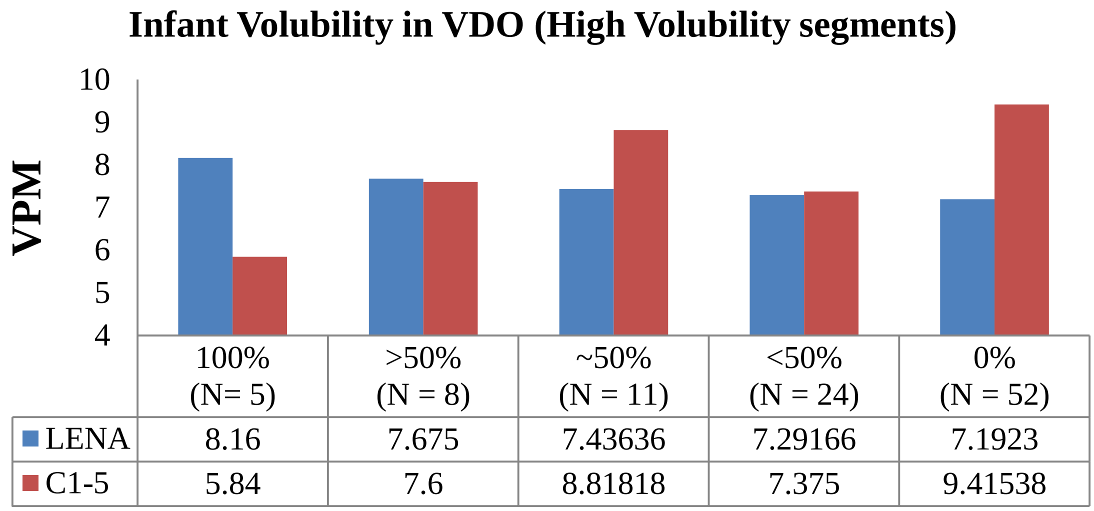
As indicated in Figure 13, the pattern of volubility for VDI with HV was quite different from the one seen for the Randomly Selected samples. Indeed, for human coded samples the highest level of volubility occurred in HV for the samples where there was 0% talk to the infants, precisely the opposite from the case of the RS samples. But the low *N* in the HV case for 0% makes us hesitate to draw a strong conclusion here until more data are available. The pattern of highest volubility in 0% VDI was even more weakly present in the case of the LENA values than for the human coding in the HV samples.




| | 100% (N= 29) | >50% (N = 13) | ~50% (N = 21) | <50% (N = 23) | 0% (N = 14) |
|---|---|---|---|---|---|
| LENA | 7.48966 | 6.63076 | 6.7619 | 7.62608 | 8.01428 |
| C1-5 | 7.94482 | 7.2 | 8.68572 | 8.15652 | 11.4 |

*Figure* 13. Infant volubility by VDI (High Volubility segments) for all 5 levels of VDI.

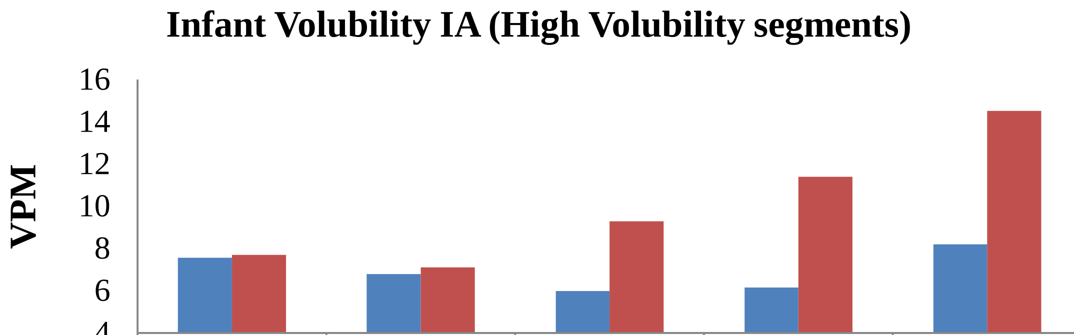
**High Volubility Segments for Different Levels of Vocalizations Directed to Others (VDO).** Again the pattern across categories of VDO was not the same in the HV samples as it was in the RS samples. Here, the human coding showed the highest rate at 0% VDO, whereas in the RS samples, the highest rate was for ~50% VDO. The LENA results for HV segments were not patterned across levels of VDO as in the case of human coding.




| | 100% (N= 5) | >50% (N = 8) | ~50% (N = 11) | <50% (N = 24) | 0% (N = 52) |
|---|---|---|---|---|---|
| LENA | 8.16 | 7.675 | 7.43636 | 7.29166 | 7.1923 |
| C1-5 | 5.84 | 7.6 | 8.81818 | 7.375 | 9.41538 |

*Figure* 14. Infant volubility by VDO (High Volubility segments) for all 5 levels of VDO.

**High Volubility Segments at Different Levels of the Infant Being Alone (IA).** Infant volubility levels varied with differences in the percent time the infant spent alone as seen in Figure 15. The results of the HV segment analysis showed that in contrast with the Randomly Selected segments the condition of the infant alone or not being spoken to did *not* correspond to low volubility, but rather to high volubility. Still the *N* for the low IA categories was very low, and the pattern will need to be evaluated with a larger sample in the future. The pattern of high volubility when the infant was alone was not strongly seen in the LENA values.




| | 0% Alone (N= 67) | <50% Alone (N = 11) | ~50% Alone (N = 9) | >50% Alone (N = 4) | 100% Alone (N = 9) |
|---|---|---|---|---|---|
| LENA | 7.56 | 6.78 | 5.98 | 6.15 | 8.20 |
| C1-5 | 7.70 | 7.11 | 9.29 | 11.40 | 14.51 |

*Figure* 15. Infant volubility by IA (High Volubility segments) for all 5 levels of IA.

**Low Volubility Segments.** Low Volubility segments were evaluated in Study 2 in order to determine the extent to which the three consecutive zero's criterion of Study 1 had actually effectively excluded infant sleep. To remind the reader, in Study 1, segments were not included in random sampling in any case in which they occurred in sequences where LENA assigned zero Child Vocalization Count to at least three consecutive segments.

In this section we address the question: Did the presumption of three consecutive zero-child-vocalization 5-minute segments according to LENA being sleep time turn out to be an appropriate method of separating out sleep time?

The data from human coding indicated that approximately 24% of the time, the infant was *not* asleep during any portion of these presumed zero-vocalization segments. Still, the infants *were* asleep for some portion of 76% of the segments according to the coders, and were considered asleep for the entire 5-minute segment in approximately

66% of the segments. It is also of interest that across the 97 Low Volubility segments that were actually fully coded (3 were not completed for unexplained reasons), there were 161 infant protophones coded along with 9 cries. 42 protophones occurred in a single 5-minute segment even though LENA designated the Child Vocalization Count as zero.

**Analysis on Vocal Types.** To conduct an analysis of vocal type and its possible variation across circumstances, we began by segregating the 189 Randomly Selected sessions (sleep excluded) into three classes. In the first there was high VDI ($n$ = 66), in the second there was high VDO ($n$ = 71) and in the third the infant was alone a significant portion of the time ($n$ = 31). Each 5-min segment in this analysis was uniquely assigned to one of the three circumstance classes.

A univariate two-way analysis of variance on infant volubility for the three protophone types and two reflexive vocalization types (cry or laugh) from the three classes of circumstance (high VDI, high VDO, or high IA) revealed a significant main effect of vocal type, $F(1, 330) = 75.86$, $p < .001$, $\eta_p^2 = .19$, and circumstance class, $F(2, 330) = 3.69$, $p = .03$, $\eta_p^2 = .02$. The circumstance-vocal type interaction was not significant, $F(2, 330) = 2.45$, $p = .09$, $\eta_p^2 = .02$. The significant vocal type effect corresponds to the more frequent production of vocants than other vocal types, and the circumstance effect supports the prior findings that high VDI corresponds to high volubility and high IA corresponds to low volubility. The lack of a significant interaction suggests the vocal types were not rigidly associated with the circumstance classes.

In order to compare the different vocal types, groupings which included 1) the protophones: vocants, squeals, and growls and 2) the reflexive sounds: cries and laughs, were evaluated in pairwise two-tailed t-tests with a Bonferroni adjustment ($p$ = .017). For the protophones, vocants were not significantly more frequent in VDI than in VDO ($t(9) = 1.55$, $p = 0.16$). However the higher frequency of vocants in VDI ($t(9) = 7.72$, $p < .001$) and VDO ($t(9) = 5.04$, $p = .001$) than in IA was significant. There were no other statistically significant pairwise comparisons within the protophones after the Bonferroni adjustment (Figure 16).

Considering the fixed signals, it is worthy of note that the coders found very little laughter in any of the sessions—a grand total of 55 cases, representing less than one half of one percent of the total vocalization in the sample. Cries on the other hand constituted about 11% of the infant vocalizations.

Analysis of the High Volubility segments showed similar outcomes. There was more crying in the HV segments (1.36 cries per minute) than in the RS segments (.4 cries per minute). In addition there was a .31 correlation between the number of protophones and the number of cries per segment for the 100 HV segments, compared to a .10 correlation for the 189 RS segments. Thus crying predicted high volubility in the segments chosen as having HV to a greater extent than in RS segments. This pattern suggests that infant distress may have contributed to high volubility, especially in the HV segments.

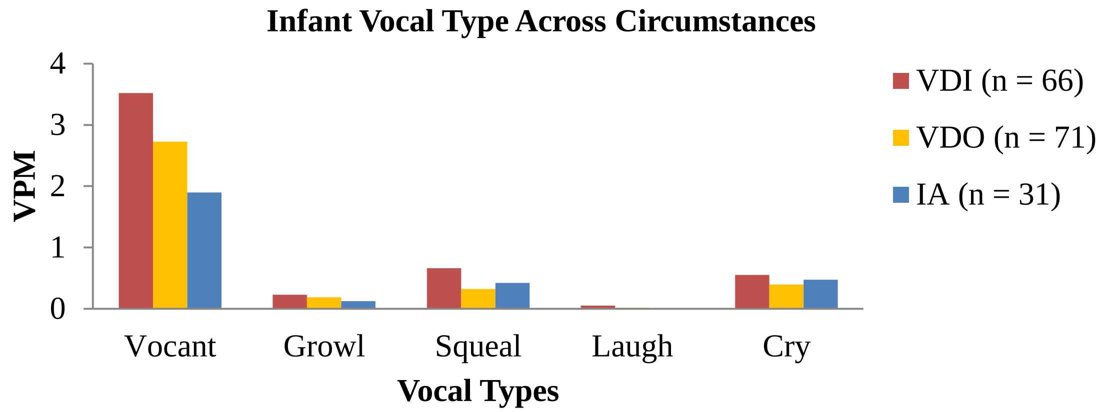
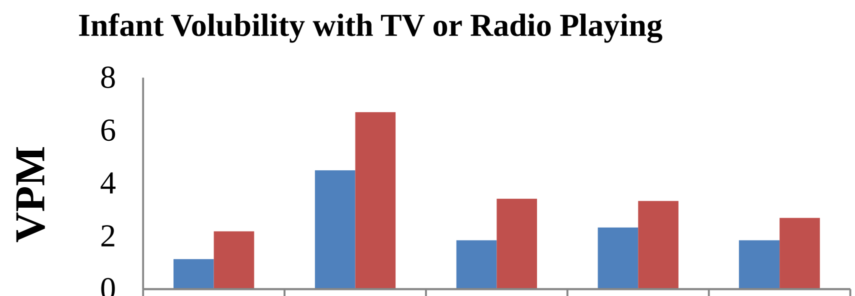



*Figure* 16. Infant vocal type across all three circumstance classes (RS segments)

## Environmental Effects

A comparison of infant volubility when there was no TV or radio playing to having the TV or radio playing 100% of the time indicated there was no significant difference in infant volubility levels related to TV or radio ($t(9) = -.27$, $p = .79$). Amount of TV did not correlate highly with either protophone volubility or the amount of crying in the Randomly Selected segments ($r = .04$, -.04, $n = 240$), but in the High Volubility segments, the correlations were higher ($r = .25$, $p = .002$ for protophones, $r = .37$, $p < .001$ for cry).

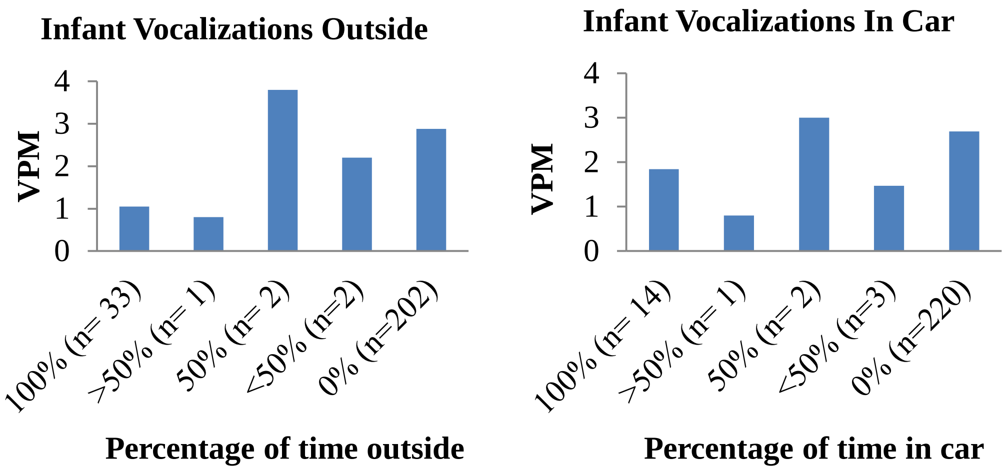
To assess the possible implications for TV watching on vocal interaction with infants, we determined that 53 0% VDI segments occurred among the 80 5-minute segments when the TV was on 100% of the time. In contrast only 9 5-minute segments with 50% or greater VDI occurred when the TV was on 100% of the time. Furthermore, 52 0% VDI segments occurred among  the 145 five-minute segments when the TV was not playing, and a similar number, 51 segments showed 50% or greater VDI when the TV was not playing. The obvious conclusion is that when the TV was on, people were much less likely (9/80 = 11% of segments) to talk to the infant 50% of the time or more, than when the TV was off (51/145 = 35% of the segments). This conclusion is supported by $\chi^2$ analysis ($\chi^2 = 15.1$, $p < .001$)




| | 100% (n=80) | >50% (n=4) | ~ 50% (n=8) | <50% (n=3) | 0% (n=145) |
|---|---|---|---|---|---|
| LENA | 1.13 | 4.50 | 1.85 | 2.33 | 1.85 |
| Human Coding | 2.18 | 6.70 | 3.43 | 3.33 | 2.70 |

*Figure* 17. Infant volubility with TV or radio in the environment.

Only 5 of the participants had segments which were coded as outside the home or in the car (Figures 18 and 19). Participant parents had been asked to try to confine recordings to the home, but some recording periods nonetheless occurred outside the home. *N*'s were small, but there was a trend for volubility to be lower for cases where 100% of the time was spent outside or in the car.




*Figures* 18 and 19. Infant vocalizations outside or in the car (RS segments)

## Summary of Results

### Study 1

Six-month old infant volubility from a home setting across a typical day was estimated at 4.7 vocalizations per minute based on 12 randomly selected samples per day per infant (excluding the time presumably spent in sleep, three consecutive five minute periods of zero child vocalization count as indicated by LENA). The rate of infant vocalization estimated by LENA was about half that estimated by the human coder.

Human coders (PC and RC) agreed with each other ($r$ = .98, $n$ = 10) on volubility to a much greater extent than the primary coder agreed with LENA on volubility ($r$ = .85, $p$ < .001, $n$ = 10). The correlation between mean volubility estimated by LENA for the exported segments on the 10 infants and volubility provided by LENA for the all-day results ($r$ = .93, $p$ < .001, $n$ = 10) indicated that the exported segments were highly predictive of the all-day recordings.

**Study 2**

Volubility in Study 2 based on human coding of 24 randomly selected 5-minute samples per infant (and then excluding segments where the infant was deemed by the coder to be sleeping) was lower than in Study 1 (2.6 compared to 4.7). Study 2's estimate (3.3) may be the more valuable one, because sleep segments were excluded based on direct monitoring rather than on the 3-consecutive-0-segment criterion of Study 1, which seems to have produced a skewing of the Study 1 data such that low volubility segments were too often excluded.

As in Study 1, the human coders in Study 2 agreed with each other on infant volubility to a much greater extent (mean for 5 coders, $r$ = .96, $n$ = 14) than they agreed with LENA (for data aggregated across 5 coders, $r$ = .65, $n$ =189).

For Randomly Selected segments, the six-month-old infants displayed low volubility when there was 0% VDI (1.3 VPM), significantly lower than in cases where vocalizations were being directed to the infant (3.6 to 6.5 VPM). Similarly, infants showed significantly lower volubility in the Randomly Selected segments when they were alone (1.2 VPM), than when others were present (3.0 to 5.1 VPM). Finally, it appeared that infant volubility was highest when talk between others occurred about 50% of the time (5.8 VPM), with both very high and very low levels of talk between others corresponding to lower infant volubility (for high levels of talk 2.0 to 2.2 VPM, for low 3.1 to 3.9 VPM).

In the High Volubility segments, infants exhibited the highest level of volubility when they were in the IA 100% circumstance (14.5 VPM), although this pattern was not significantly higher than when others were present (range 7.1 to 11.4 VPM), perhaps because the *n* for infants alone was low (9). This pattern may also have been partly due to fussiness. Several of the IA 100% segments included considerable crying, and indeed cry amounts correlated with protophone volubility positively ($r = .31$, $p = .002$, $n = 100$) in the High Volubility segments. Still, there were numerous examples in the data of segments where the infant was deemed alone, or not being spoken to at all, and where nonetheless, there were considerable numbers of infant vocalizations, perhaps instances of infant practice of vocalization.

In Study 1 presumed sleep segments were eliminated from possible sampling by excluding any segment that was one of at least three consecutive segments with zero Child Vocalization Count according to LENA. To test how well this procedure worked, in Study 2 we selected 100 segments pertaining to one of these sequences of at least three consecutive zero's, and we coded them. Seventy percent of these Low Volubility segments proved to be, as expected, cases where the infants were asleep at least part of time according to the coders, and where very low volubility occurred according to the human coders. In 15% of the segments, on the other hand, the infants were awake and produced a mean of 1.6 VPM, actually higher than the rate for infants alone and awake in the RS sample.

Vocal types as seen in the Randomly Selected segments revealed that protophones (vocants, growls and squeals) accounted for 89% of all infant vocalizations while cry accounted for 11% and laughs less than one-half of one percent. The very low volubility for laughter was not expected.

Also protophone types showed roughly similar patterns across three circumstance groupings of the data. There were significantly more infant vocalizations for vocants and growls when VDI was very high and significantly fewer when the infants were alone, with segments where others were talking to each other very frequently showing an intermediate level of infant volubility. Squeals also showed highest volubility when VDI was high, but did not show the pattern of vocants and growls for high IA and high VDO. Vocants were most prevalent among the protophones, as has been reported in prior research (Koopmans-van Beinum, 1986; Oller,1980, 2000; Stark, 1978).

Environmental factors did not seem to significantly affect volubility levels in Study 2. There were no significant differences between the volubility when the TV was on for the entire 5-minute segment ($n$ = 80) or when the TV or radio was not on at all ($n$ = 145). Similarly, no notable differences occurred in volubility for infants out of doors or riding in cars.

## Discussion

The use of the LENA automated recording and analysis system provides a way of gathering infant volubility across the day in the home environment with no intrusions or adjustments in daily routine. Because of the existence of this system, it is now possible to find all-day infant volubility levels and to estimate them both directly with LENA and separately by coding randomly selected segments that can be very conveniently obtained in the system.

The six-month volubility level according to our human coded random sampling method overall was 3.3 (Study 2 RS) or 4.7 (Study 1, excluding three consecutive zero segments). The results for Study 1 suggested that six-month-old infant volubility (4.7 VPM) fits in the middle for values previously reported for face-to-face interaction (2 to 6.5 VPM) (Delgado, Messinger, & Yale, 2002; Franklin et al., in press; Yale, Messinger, Cobo-Lewis, Oller, & Eilers, 1999). Still, face-to-face interaction might be expected to show considerably higher rates than would be obtained by sampling across an entire day. This expectation is more in line with the estimate of Study 2, and that estimate may be the more reliable one, because it was based on a sampling method that was truly random at the outset, and where 'sleep' segments were eliminated afterwards. As a result of this difference in sampling procedure, there were 79/240 segments with zero vocalizations according to the human coders included in Study 2, while there were only 10/120 zero segments included in Study 1.

The volubility estimate based on this all-day sampling (4.7 VPM) in Study 1 was nearly twice the value estimated by the LENA software (2.4 VPM) for the same 12 segments per infant. This finding is quantitatively informative, yet it was predictable that the volubility values here would be higher since the LENA software does not categorize sounds as infant vocalization unless they are free of overlay from other voices or sounds. The human listener usually discerns the infant voice despite overlay. In Study 2 the volubility estimates based on the all-day random sampling (3.3 VPM) were again notably higher that the values estimated by the LENA software (2.1 VPM).

When comparing infant volubility across different circumstances, there were differences in numbers of protophones produced when vocalizations were directed to infants or to others, or when infants were alone. Our findings suggest that contrary to the widely cited suggestion of Delack and Fowlow (1978), infants do *not* vocalize most when alone, but rather when they are engaged in vocal interaction. The presence of other individuals speaking with each other appears to have a complex effect, where greatest infant volubility corresponds to moderate amounts of adult-to-adult talk.

In contrast to the results for the Randomly Selected sample (Study 2), the High Volubility Segments showed highest infant volubility when IA was 100%. These patterns in the HV sampling may be thought to support the widely cited claim that infants vocalize more when alone than in social interaction. But interpreting this result is precarious, because the infants in the alone situation may have been calling for attention or fussing and thus not vocalizing in a truly solitary fashion. They may have been vocalizing instrumentally, and this possibility is supported by the positive correlation between amount of cry and protophone volubility in the HV segments—the correlation for the RS segments was considerably lower. Another possible influence on this pattern may be that infants sometimes produce high rates of vocalization in practice-like play while alone. Others have indeed suggested that infants vocalize as either motor practice (Locke, 1989) or as a form of vocal play (Dodd, 1972; Masur & Rodemaker, 1999). The infants in a few of the highest volubility segments (VPM > 7) were indeed alone and not crying. This is a pattern that suggests practice. Our continued research on this topic will attempt to tie this speculation to firmer ground with more detailed assessments of how infants vocalize (for example how much they fuss or seem to call for attention) when they are alone.

The LENA automated analyses consistently produce lower estimates of volubility than human coding. This was consistent across all circumstances and in all of the segment types (see Appendix C), although the differences were greater for Randomly Selected segments than for High Volubility segments. While the LENA automated analyses provides counts that correlate highly with human coding, the automated methods do not yet reflect the finer level of detail that human coders are able to recognize.

At this point we have a tentative conversion factor for LENA estimates of 6-month-old infant volubility designed to produce a value that should approximate human coding where all infant vocalizations including those in overlap with other voices or noise are included. The conversion factor is based on the totally random sampling of Study 2, with an *N* of 240 five-min segments. In this case, the human coding yielded a value (2.6 VPM), that was 58% higher than the LENA value for the very same segments (1.7 VPM). The inclusion of all 240 in this comparison is necessary because we have no means of determining when sleep occurred in the LENA all-day samples that we wish to use as a basis for larger scale estimation of volubility. If we use the reported value from 51 twelve-hour recordings from the LENA Research Foundation Natural Language Study Technical Report accessible on line, applying this conversion factor, we acquire an estimate of 2.0 VPM, lower than our human coding estimate of 2.6, perhaps a result of higher SES in the Memphis sample (the LRF sample was stratified to represent the US Census, while the Memphis sample was clearly mid to high SES). This value (2.0 VPM) is, we think, the best current estimate, consistent with standard laboratory counting procedures, of naturally occurring volubility for six-month-old Americans across the national range of SES.

## Limitations and Future Directions

Even though the combination of these two studies resulted in the coding of more than 9,000 infant utterances, the sample seems small in retrospect because so many of our questions ultimately would have profited from larger numbers. For example, the infant alone circumstance is of substantial interest, but we found few 5-minute segments through RS, HV or LV sampling where the infant was both alone and awake. A larger sample size is clearly in order to determine the role of spontaneous vocalizations when alone.

In these studies a single all-day recording from each infant was obtained, and it is unclear how much variation might have occurred with additional coding on another all-day recording at six-months of age for each infant. In addition our data pertain to one age, and there is considerable evidence (Caskey, Stephens, Tucker, & Vohr, 2011; Oller et al., 1995; Yale et al., 1999) to suggest that the values that will be obtained in our future efforts at other ages will be different. Having volubility estimates from different ages (across the whole first year and also in prematurely born infants still not at full term ages) to compare with LENA automated estimates will provide a range of conversion factors upon which we can base age-specific estimates of volubility based on LENA's large stratified sample.

Tracking the progression across ages for volubility and vocal type in a variety of circumstances such as those considered here should provide valuable information regarding the endogenous motivation of the infant as well as the role of parental input in the acquisition and development of language. Additional information regarding infant state and intentions might be obtained through additional questions about, for example, fussiness or body movement (Jones & Moss, 1971). Since cry and protophones in the HV segments did correlate significantly, it would be beneficial to determine how much of the vocalization (especially when the infant was alone) was spontaneously generated versus the amount that was inspired by discomfort or boredom (attention seeking or fussing).

While it may be some time before automated analyses can be exclusively relied upon for volubility data, the combination of automated analysis and human coding of as little as 20% of the 5-minute segments across the day has proven to provide us with enough information to begin to determine a typical range of volubility at six months. With larger samples of infants at each age, it seems within reach to develop clinical criteria for early identification based on very low (and perhaps very high) volubility.

The current LENA recorder is audio only. Clearly it would be valuable to verify circumstance judgments through coding of simultaneously obtained video signals. While all-day audio/video is still prohibitively expensive from the standpoint of power and storage requirements, it will presumably not be so for much longer. In the meantime short-term video recordings could be added to spot check audio-based circumstance coding. In addition, parents can supply hour by hour information about circumstances in the home, especially regarding when the infant is alone and/or asleep.

The ability to gather data in the natural home environment without disturbing the family routine provides a better depiction of the infant or child's true volubility as well as the distribution of different vocal types than traditional sampling methods. Naturalistic sampling may be particularly informative for infants and children with disorders who often find the laboratory setting constricting and do not perform to the best of their abilities except when they are at home. Our research thus can be viewed as laying foundations in both scientific and clinical realms.

## References

Action Analysis Coding and Training (Version 2.1) [Computer software, by Rafael Delgado]. (1996), Miami, FL: Intelligent Hearing Systems.

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

Aragon, M., & Yoshinaga-Itano, C. (2012). Using Language Environment Analysis to improve outcomes for children who are deaf or hard of hearing. *Seminar In Speech and Language, 33,* 340-353.

Austin, J. L. (1962). *How to do things with words*. London, UK: Oxford University Press.

Bakeman, R., & Adamson. L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development, 55*, 1278-1289.

Bayart, F., Hayashi, K., Faull, K., Barchus, J., & Levine, S. (1990). Influence of maternal proximity on behavioral and physiological responses to separation in infant rhesus monkeys (Macaca mulatta). *Behavioral Neuroscience, 104*, 98-107.

Beebe, B., Jaffe J., Feldstein S., Mays, K., & Alson, D. (1985). Interpersonal timing: The application of an adult dialogue model to mother-infant vocal and kinesic interactions. *Social perception in infants (*pp. 217-247). In T. Field & N.A. Fox (Eds.) Norwood, NJ: Ablex.

Benson, B., Binz, H., & Zimmerman, E. (1992). Vocalizations of infant and developing tree shrews (Tupaiabelangeri). *Journal of Mammology*, *73*, 106-119

Berger, J., & Cunningham, C. (1983). Development of early vocal behaviors and interactions in Down's syndrome and non-handicapped infant-mother pairs. *Developmental Psychology, 19,* 322-331

Bickley, C., Lindblom, B., & Roug, L. (1986). Acoustic measures of rhythm in infants' babbling, or 'all god's children got rhythm'. *Proc. Int. Congr. Acoust*. *12,* Canada, A6–4.

Bloom, K., & Esposito, A. (1975). Social conditioning and its proper control procedures. *Journal of Experimental Child Psychology, 19*, 209-222.

Bornstein M. H., & Bradley, R. H. (Eds.). (2003). *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *Maternal responsiveness: Characteristics and consequences* (Vol 43, pp 49-61), New Directions of Child Development, M.H. Bornstein (Ed.). San Francisco, CA: Jossey-Bass.

Bornstein, M., Tamis-Lemonda, C., Tal, J., Ludemann, P., Todaa, S., Rahn, C., Pecheux, M., Azuma, H., & Vardi, D. (1992). Maternal responsiveness to infants in three societies: The United States, France, and Japan. *Child Development, 63,* 808-821.

Buder, E. H., Warlaumont, A. S., & Oller, D. K. (2013). An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. In B. Peter & A. N. MacLeod (Eds.), *Comprehensive perspectives on child speech development and disorders: Pathways from linguistic theory to clinical practice*. NOVA.

Camp, B., Burgess, D., Morgan, M., & Zerbe, G. (1987). A longitudinal study of infant vocalization in the first year. *Journal of Pediatric Psychology, 12*, 321-331.

Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2011). Adult-infant conversations in the NICU and language and cognitive outcomes in preterm infants. Poster presentation at the LENA Conference.

Chisolm, J. (2003). Uncertainty, contingency, and attachment: A life history theory of theory of mind. In K. Sterelny & J. Fitness (Eds.) *From mating to mentality: Evaluating evolutionary psychology.*

Cohn, J., & Tronick, E. Z. (1987). Mother-infant face-to-face interaction: The sequence of dyadic states at 3,6, and 9 months. *Developmental Psychology*, *23*(1) , 66-71.

Cohn, J., & Tronick, E. Z. (1988). Mother-Infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology*, *24*, 3, 386-392.

Craig, H., & Washington, J. (2005). *Malik goes to school: Examining the language skills of African American students from preschool-5th grade*. Mahwah, NJ: Lawrence Erlbaum.

Delack J. (1978). Aspects of infant speech development in the first year of life. *Canadian Journal of Linguistics, 21*, 17–37.

Delack J., & Fowlow P. (1978). The ontogenesis of differential vocalization: development of prosodic contrastivity during the first year of life. In N. Waterson & S. Catherine (Eds) *The Development of Communication* (pp. 93–110). New York, NY: Wiley.

Delgado, C., Messinger, D., & Yale, M. (2002). Infant responses to direction of parental gaze: A comparison of two still-face conditions. *Infant Behavior and Development*, *25*, 311-318.

de Waal, F. (1988). The communicative repertoire of captive bonobos (Pan paniscus), compared to that of chimpanzees. *Behavior, 106,* 183-251.

Dodd, B. (1972). Effects of social and vocal stimulation on infant babbling. *Developmental Psychology,* 7(1), 80-83.

Dykstra, J., Sabatos-DeVito, M., Irvin, D., Boyd, B., Hume, K., & Odom, S. (2012). Using the language environment analysis (LENA) system in preschool classrooms with children with autism spectrum disorders. *Autism, 0,* 1-13.

Elowson, M., Snowdon, C., & Lazaro-Perea, C. (1998). Infant 'Babbling" in a non-human primate: Complex vocal sequences with repeated call types. *Behaviour, 135*(5), 643-664.

Farah, M., Betancourt, L., Shera, D., Savage, J., Giannetta, J., Brodsky, N., Malmud, E., & Hurt, H. (2008). Environmental stimulation, parental nurturance and cognitive development in humans. *Developmental Science*, *11*, 793-801.

Fleiss, J. S. (1981). Statistical methods for rates and proportions. New York, NY: Wiley.

Fogel, A. (1988). Cyclicity and stability in mother-infant face-to-face interaction: A comment on Cohn and Tronick (1988). *Developmental Psychology,* 24(3), 393-395.

Fogel, A., & Garvey, A. (2007). Alive communication. *Infant Behavior and Development*, *30*, 251-257.

Franklin, B., Warlaumont, A.S., Messinger, D.S., Bene, E., Iyer, S.N., Lee, C-C., Lambert, B., Oller, D.K. (in press). Effects of Parental Interaction on Infant Vocalization Rate, Variability and Vocal Type. *Language Learning and Development*.

Gerken, L. (2002). Early sensitivity to linguistic form. *Annual Review of Language*

*Acquisition, 2,* 1-36.

Gilkerson, J., & Richards, J. (2009). *The LENA natural language study*. LENA Foundation Technical Report, LTR-02-2.

Gilkerson, J., & Richards, J. (2009). *The power of talk (2nd Ed): Impact of adult talk, conversational turns, and TV during the critical 0-4 years of child development*. LENA Foundation Technical Report, ITR-01-2.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci*. U.S.A., *100*(13), 8030–8035.

Goldstein, M., Schwade, J., & Bornstein, M. (2009). The Value of Vocalizing: Five month old infants associate their own non-cry vocalizations with Responses from Caregivers. *Child Development, 80*, 636-644.

Goldstein, M., & West, M. (1999). Consistent responses of human mothers to Prelinguistic infants: The effect of prelinguistic repertoire size. *Journal of Comparative Psychology, 113*, 52-58.

Gros-Louis, J., West, M., Goldstein, M., & King, A. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30,* 5, 112-119.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Hart, B., & Risley, T. R. (1999).*The Social World of Children Learning to Talk.* Baltimore, MD: Paul H. Brookes.

Hauser, M. (1996). *The evolution of communication.* Cambridge, MA: MIT Press

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368-1378.

Hoff, E. (2006a). How social contexts support and shape language development. *Developmental Review, 26*, 55-88.

Hoff, E. (2006b). How social contexts support and shape language development. *Developmental Review, 26*, 55-88.

Hoff, E. (2014). *Language Development* ($4^{th}$ Ed). Wadsworth; Belmont, CA.

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (2nd ed., pp. 231–252). Mahwah, NJ: Lawrence Erlbaum Association.

Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders*, *38*, 271-278.

Hollingshead, A. (1975). Four factor index of social status. Unpublished manuscript, Department of Sociology, Yale University.

Hollingshead, A. (1978). Two factor index of social status. Unpublished manuscript, Yale University.

Hsu, H., & Fogel, A. (2003). Stability and transitions in mother-infant face-to-face communication during the first 6 months: A microhistorical approach. *Developmental Psychology*, *39*(6), 1061-1082.

Hsu, H., Fogel, A., & Messinger, D. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction: factors associated with quantitative and qualitative differences. *Infant Behavior and Development*, *24*, 107-128.

Iverson, J. (2010). Developing language in a developing body: the relationship between motor development and language development. *Journal of Child Language, 37,* 229-261.

Iverson, J., & Fagan, M. (2004). Infant vocal – motor coordination: Precursor to the gesture—speech system? *Child Development, 75,* 1053-1066.

Iyer, S. N., & Oller, D. K. (2008). Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss. *Volta Review, 108,* 115-138.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). *Rhythms of dialogue in infancy: Coordinated timing in development*. Chicago, IL :University of Chicago Press.

Jones, S. J., & Moss, H. A. (1971). Age, state and maternal behavior associated with infant vocalizations. C*hild Development, 42*, 1039-1051.

Kaye, K., & Fogel, A. (1980). The temporal structure of face-to-face communication between mothers and infants. *Developmental Psychology, 16*(5), 454-464.

Keller, H., Lohaus, A., Völker, S., Cappenberg, M., & Chasiotis, A. (1999). Temporal Contingency as an Independent Component of Parenting Behavior. *Child Development*, *70*(2), 474–485.

Kelly, D. (2011). Language acquisition challenges for preschoolers residing in low-SES households: implications for speech-language pathologists and developmental researchers. *Perspectives on Language Learning and Education,* 41-48.

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, NY: Stockton Press.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infant by 6 months of age. *Science, 255,* 606-608.

Locke, J. (1989). Babbling and early speech: continuity and individual differences. *First Language*, *9,* 191-206.

Locke, J. L. (1993). *The child's path to spoken language* (First ed.). Cambridge, MA: Harvard University Press.

Locke, J. L. (2006). Parental selection of vocal behavior: Crying, cooing, babbling, and the evolution of language. *Human Nature*, *17*(2), 155-168.

Marler, P., Evans, C. S., & Hauser, M. D. (1992). Animal signals? Reference, motivation, or both? In H. Papoucek U. J. Jurgens, & M. Papoucek, (Eds.), Nonverbal vocal communication: Comparative and developmental approaches (pp. 66–86). Cambridge, UK: Cambridge University Press.

Masur, E., & Rodemaker, J. (1999). Mothers' and infants' spontaneous vocal, verbal, and action imitation during the second year. *Merrill-Palmer Quarterly, 45,* 392-412.

McCauley, A., Esposito, M., & Cook, M. (2011). Language environment analysis of preschoolers with autism: validity & applications. Poster presentation at LENA Conference.

Molemans, I. (2011). *Sounds like babbling*. Antwerp: University Press.

Montgomery, J., Gilkerson, J., Richards, J., & Xu, D. (2009). Infants around more talk become toddlers with better language skills. Technical Presentation at the ASHA Convention 2009.

Oetting, J., Hartfield, L. , & Pruitt, S. (2009, May 05). Exploring LENA as a Tool for Researchers and Clinicians. *The ASHA Leader*.

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, NY: Academic Press.

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*. doi: doi:10.1073/pnas.1300337110/-/DCSupplemental.

Oller, D. K., Eilers, R., Basinger, D., Steffens, M., & Urbano, R. (1995). Extreme poverty and the development of the precursors to the speech capacity. *First Language, 15*(2), 167-187.

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

Oller, D. K., Niyogi, P., Gray, S., Richards, J., Gilkerson, J., Xu, D., Yapanel, U., & Warren, S. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings from the National Academy of Science*, 1-6.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences, 106*(51), 22026-22031.

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*, 763–782.

Papaeliou, C., Minadaki, 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. Thomas (Ed.) *Origins of the infant's social responsiveness* (pp. 175-208). Hillsdale, NJ: Erlbaum Associates.

Papoušek, M., & Papoušek, H., (1989). Forms and functions of vocal matching in interactions between their mothers and precanonical infants. *First Language, 9*(6) 137-157.

Pelaez, M., Virues-Ortega, J., & Gewritz, J. (2011). Reinforcement of vocalizations through contingent vocl imitation. *Journal of Applied Behavior Analysis, 44,* 33-40.

Pungello, E., Iruka, I., Dotterer, A., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental Psychology, 45,* 544-557.

Rheingold, H., Gerwitz, J., & Ross, H. (1959). Social conditioning of vocalizations in the infant. *National Institutes of Mental Health.*

Richards, J., Gilkerson, J., Paul, T., & Xu, D. (2008). *The LENA automatic vocalization assessment.* LENA Foundation Technical Report.

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, *323*, 951–953.

Scheiner, E., & Fischer, J. (2011). Emotion expression: The evolutionary heritage in the human voice. In *Interdisciplinary Anthropology* (pp. 105-129). Springer Berlin Heidelberg.

Scheiner, E., Hammerschmidt, K., Jürgens, U., & Zwirner, P. (2006) Vocal expression of emotions in normally hearing and hearing-impaired infants. *Journal of Voice, 20*(4),585-604.

Scherer, N., Williams, A., & Proctor-Williams, K. (2008). Early and later vocalization skills in children with and without cleft palate. *International Journal of Pediatric Otorhinolaryngology, 72,* 827-840.

Stark, R.E. (1978). Features of infant sounds: the emergence of cooing. *Journal of Child Language, 5,* 379-390.

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

Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In M. Lewis & L.A. Rosenblum, (Eds.) *The effect of the infant on its caregiver* (pp. 187-213), New York, NY: Wiley.

Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: two modes of communication within the mother-infant dyad. *Annals of the New York Academy of Sciences*, *263*, 89-100.

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

Thelen, E. (1991). Motor aspects of emergent speech: A dynamic approach. In N. Krasnegor, (Ed), *Biobehavioral foundations of language* (pp. 339-362). Hillsdale, NJ: Erlbaum.

Todd, G. A., & Palmer, B. (1968). Social reinforcement of infant babbling. *Child Development*, *39*, 591-596.

Trevarthen, C. (1977). Descriptive analyses of infant communicative behavior, (pp. 227-270). In H.R. Shaffer (Ed.), *Studies in mother-infant interaction*, New York, NY: Academic Press.

Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersuhjectivity. In M. M. Bullowa (Ed.), *Before speech: The beginnings of interpersonal communication* (pp. 321 -348). New York: Cambridge University Press.

Trevarthen, C. (2001). Infant Intersubjectivity: Research, Theory, and Clinical Applications. *Journal of Child Psychology and Psychiatry, 42*, 3-48.

Tronick, E. Z. (1982). *Social interchange in infancy.* Baltimore, MD: University Park Press.

Warlaumont, A. S., Oller, D. K., Dale, R., Richards, J. A., Gilkerson, J., & Xu, D. (2010). Vocal interaction dynamics of children with and without autism. In S. Ohlsson & R. Catrambone (Eds.), Proceedings of the 32nd Annual Conference of the Cognitive Science Society (pp. 121–126). Austin, TX: Cognitive Science Society.

Weinberg, K., & Tronick, E. Z. (1996). Infant Affective Reactions to the Resumption of Maternal Interaction after the Still-Face. *Child Development, 67*(3), 905-914.

Weisleder, A. & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science,* doi:10.1177/0956797613488145.

Wong, J., Stewart, P., & MacDonald, D. (1999). Vocal repertoire in the European badger (Meles meles): Structure, context, and function. *Journal of Mammalogy, 80,* 570-588.

Xu, D., Yapanel, & Gray, S. (2009). Reliability of the LENA Language Environment Analysis system in young children's natural home environment. LENA Technical Report, 05-2 retrieved electronically from http://www.lenafoundation.org/Research/TechnicalReports.aspx.

Yale, M. E., Messinger, D. S., Cobo-Lewis, A. B., Oller, D. K., & Eilers, R. E. (1999). An event-based analysis of the coordination of early infant vocalizations and facial actions. *Developmental Psychology, 35*(2), 505-513.

Yang, H.-C. (2005). The mutual effect of the vocalization quality and quantity in mother-infant dyads. Unpublished Master's Thesis.

## Appendices

### Appendix A

### Study 1 Reliability Scatter Plots

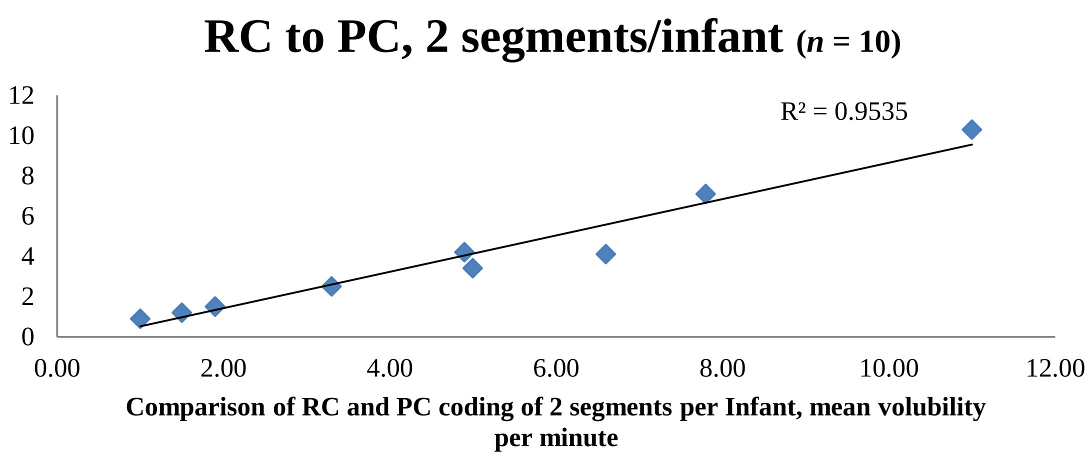



*Figure* 20. Scatterplot Study 1 RC to PC

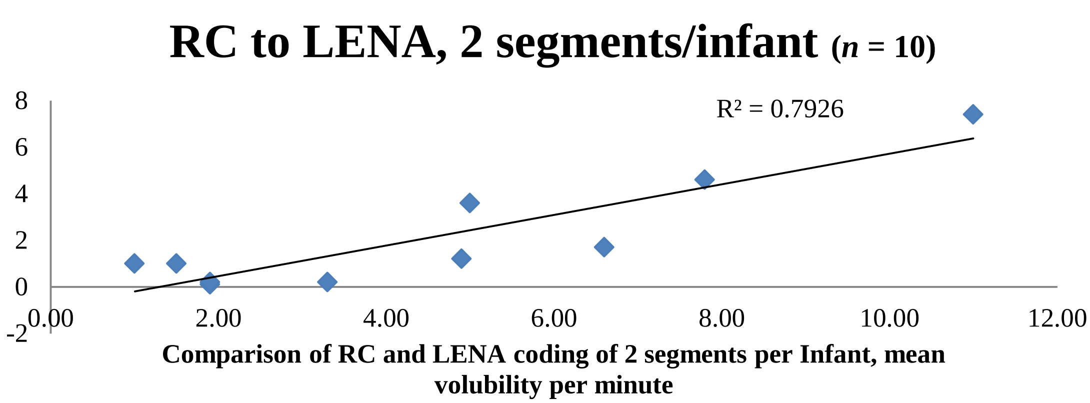



*Figure* 21. Scatterplot Study 1 RC to LENA

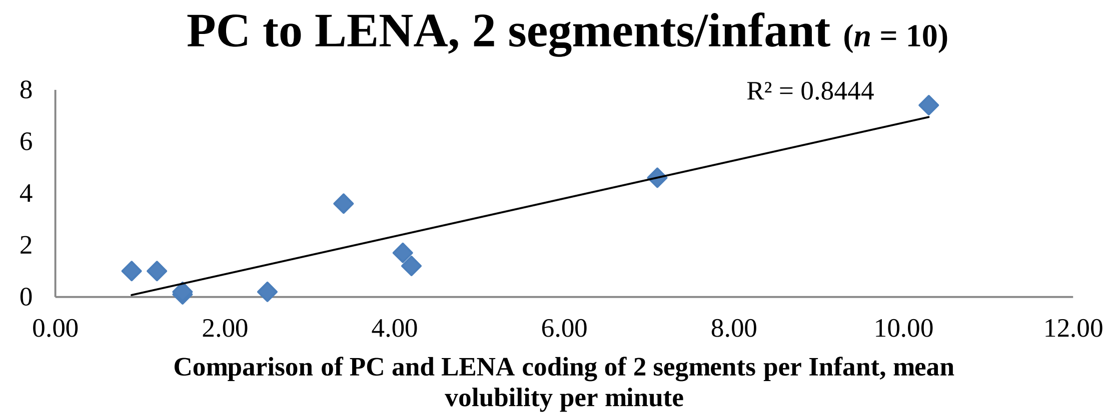



*Figure* 22. Scatterplot Study 1 PC to LENA

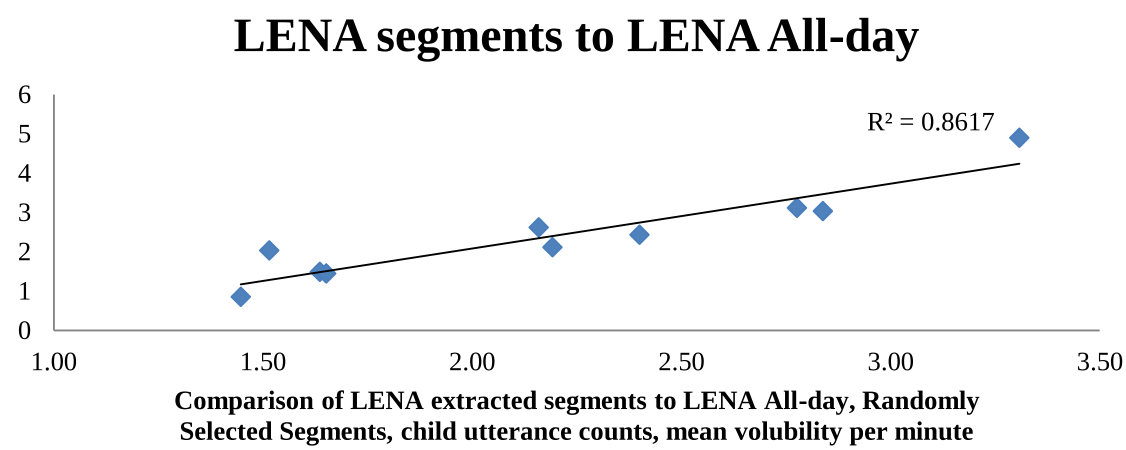



*Figure* 23. Scatterplot Study 1 LENA segments to LENA All-day

## Appendix B

### Study 2 Reliability Scatter Plots

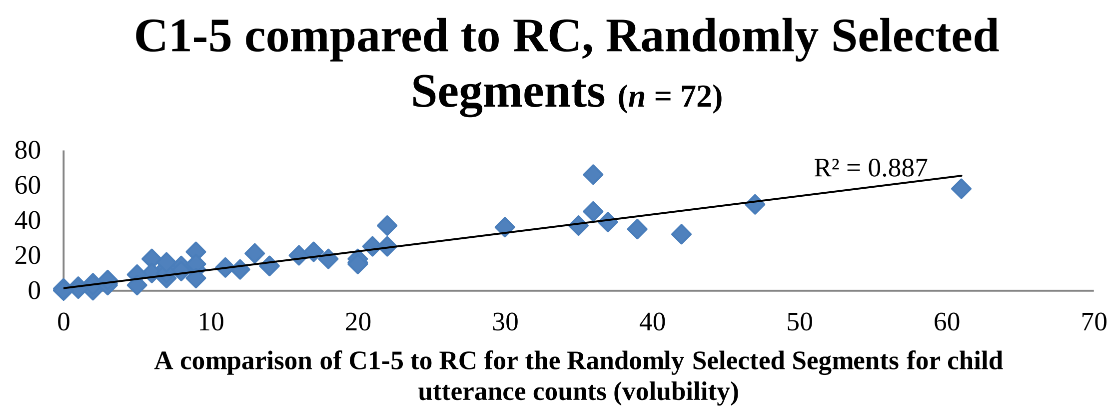



*Figure* 24. Scatterplot Study 2 C1-5 to RC, RS segments

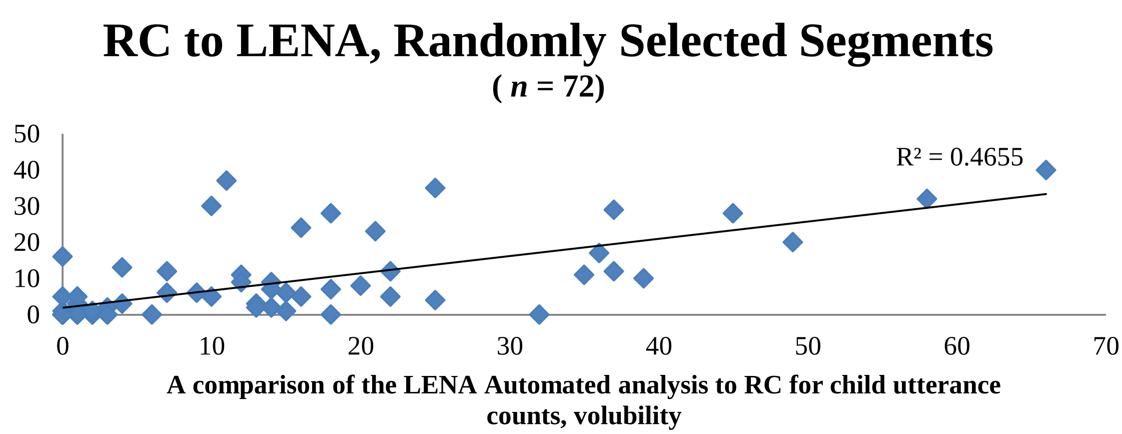



*Figure* 25. Scatterplot Study 2 RC to LENA, RS segments

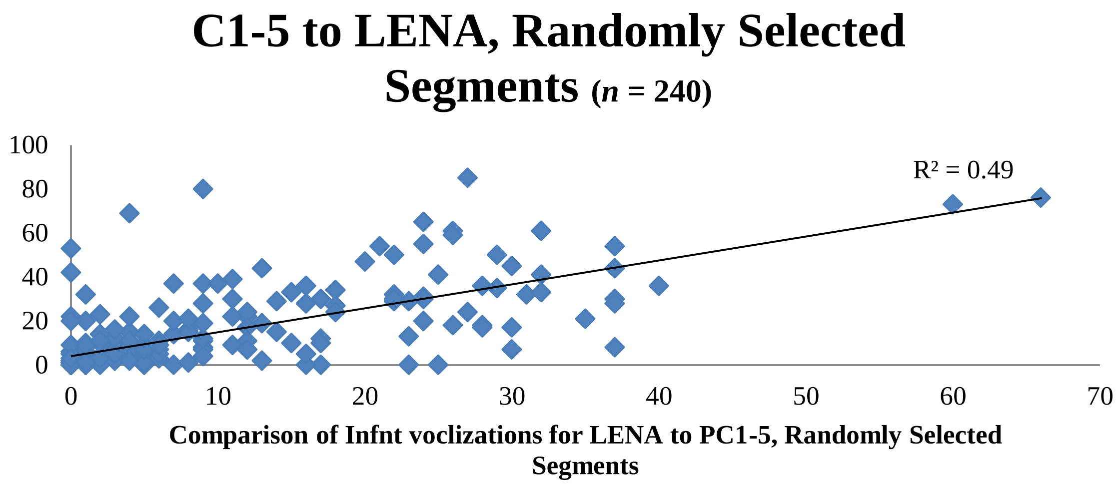



*Figure* 26. Scatterplot Study 2 C1-5 to LENA, RS segments

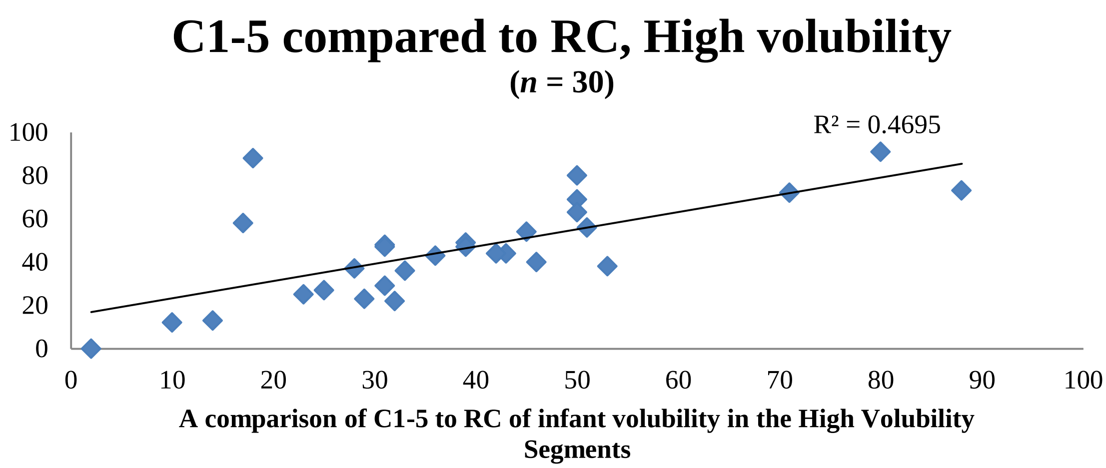



*Figure* 27. Scatterplot Study 2 C1-5 to RC, HV segments

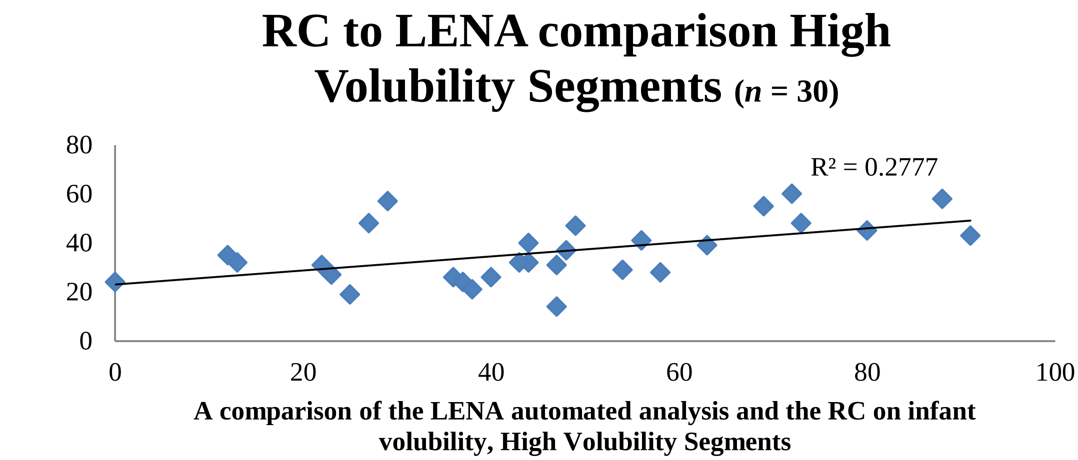



*Figure* 28. Scatterplot Study 2 RC to LENA, HV segments

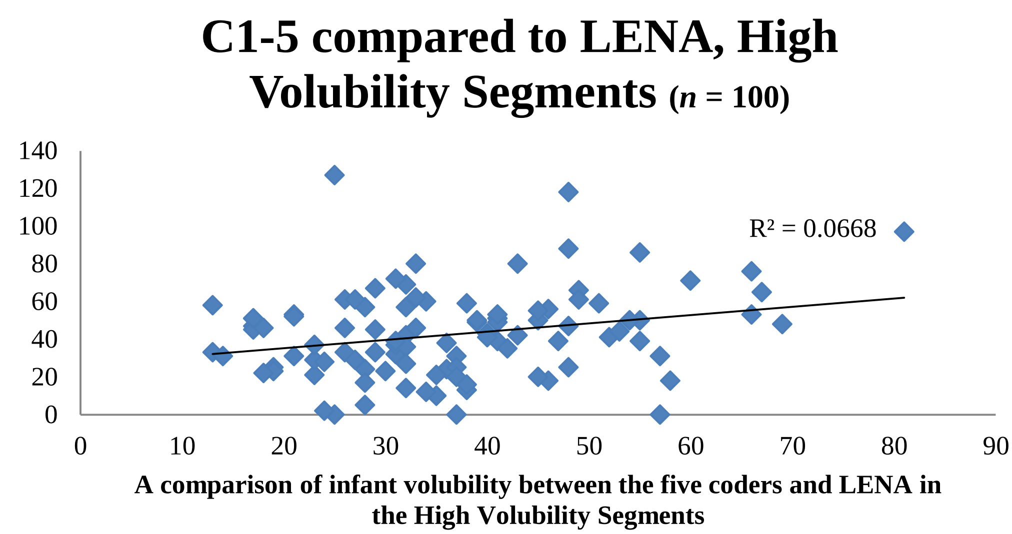



*Figure* 29. Scatterplot Study 2 C1-5 to LENA, HV segments

# Appendix C

## Study 2: Within-circumstance correlations

## and t-tests for human coding vs. LENA analysis

The LENA automated child vocalization count data is compared here to the human coding across the degrees of each circumstance parameter for both the Randomly Selected segments and the High Volubility segments.

**For levels of VDI in the RS sample.** When comparing LENA outcomes to those for the 5 coders for the Randomly Selected segments sorted by VDI, the correlations of volubility as estimated by the two methods (human and LENA) ranged from .55 to .66. The dependent t-tests indicated that even within levels of VDI, the human coding sometimes estimated significantly higher volubility rates than LENA. These values are deemed of interest because LENA research is now being conducted widely across differing circumstances of recording without empirical information about the robustness of LENA measures across circumstances.

Table 6

*Comparison of Infant Speech-like Vocalizations ~ LENA coding to Human Coding for the 5 degrees of VDI*

| VDI | Correlation | Dependent two tailed, t-test comparison |
|---|---|---|
| **100%** | **$r$ = .57, $p$ = .002** | $t$(25) = -.84, $p$ = .41 |
| **>50%** | **$r$ = .61, $p$ = .015** | **$t$(14) = -3.77, $p$ = .002** |
| **~50%** | **$r$ = .55, $p$ = .005** | $t$(24) = -1.89, $p$ = .07 |
| **<50%** | **$r$ = .62, $p$ < .001** | **$t$(60) = -3.87, $p$ < .001** |
| **0%** | **$r$ = .66, p < .001** | ***t(61) = - 3.32, p = .002*** |

(bold = statistically significant)

**For levels of VDO in the RS sample**. When comparing the LENA to human coding in the Randomly Selected segments at differing levels of VDO the correlations of infant protophones measures ranged from .43 to .72 (Table 7) and three levels of adult to other interaction revealed significantly greater numbers of vocalizations coded by humans (>50%, ~50%, and 0%).

**Table 7**

*Comparison of Infant Speech like Vocalizations ~ LENA to Human Coding, VDO circumstance*

| VDO | Correlation | Dependent two tailed t-test comparison |
|---|---|---|
| **100%** | **$r$ = .43, $p$ <.01** | $t(32) = -1.43$, $p = .16$ |
| **>50%** | **$r$ = .73, $p$ < .001** | **$t(17) = -3.92$, $p = .001$** |
| **~50%** | **$r$ = .61, $p$ = .003** | **$t(20) = -3.15$, $p = .005$** |
| **<50%** | **$r$ = .72, $p$ < .001** | $t(38) = -2.26$, $p = .029$ |
| **0%** | ***$r$ = .62, $p$ < .001*** | ***$t(77) = -3.61$ , $p = .001$*** |

**For levels of IA in the RS sample.** When comparing LENA to human coding for the Randomly Selected segments sorted by IA, the correlations of volubility as estimated by the two methods ranged from .56 to .88 (Table 8). The dependent t-tests indicate that even within levels of IA, the human coding sometimes estimated significantly higher volubility rates than LENA.

**Table 8**

*Comparison of Infant Speech like Vocalizations ~ LC to HC, IA circumstances*

| IA | Correlation | Dependent two tailed t-test comparison |
|---|---|---|
| **100%** | **$r$ = .56, $p$ = .005** | $t$(23) = 1.95, $p$ = .06 |
| **>50%** | **$r$ = .84, $p$ < .001** | **$t$(13) = 3.63, $p$ = .003** |
| **~50%** | **$r$ = .66, $p$ = .001** | **$t$(21) = 2.40, $p$ = .03** |
| **<50%** | **$r$ = .88, $p$ < .001** | $t$(12) = 2.01, $p$ = .07 |
| **0%** | ***r* = .58, *p* < .001** | ***t*(115) = -3.61 , *p* < .001** |

**For levels of VDI segments in the HV sample**. The correlations of infant protophone measures for LENA vs. human coding ranged from .10 to .99; only the ~50% VDI level was significant (see Table 9). There was a significantly greater number of infant vocalizations coded by human coders than LENA at the ~50% interaction level.

**Table 9**

*Comparison of infant speech-like vocalizations, VDI circumstance*

| VDI | Correlation | Dependent two-tailed t-test comparison |
|---|---|---|
| **100%** | $r$ = .31, $p$ = .10 | $t$(28) = -.60, $p$ = .55 |
| **>50%** | $r$ = .10, $p$ = .76 | $t$(12) = - .44, $p$ = .67 |
| **~50%** | **$r$ = .48, $p$ = .03** | **$t$(20) = -2.31, $p$ = .03** |
| **<50%** | $r$ = .27, $p$ = .20 | $t$(22) = -.59, $p$ = .56 |
| **0%** | $r$ = -.0002, $p$ = .99 | $t(13)$ = 1.69 , $p$ = .12 |

**For levels of VDO in the HV sample.** The correlations of LENA to human coding of volubility ranged widely from negative to positive (-.58 to .36), but Ns were very small in some cases, and only the 0% VDO (N=52) was significant. There was also a significant difference noted between LENA and human coding at 0% VDO ( $p$ = .001). See Table 10.

**Table 10**
*Comparison of LENA Automated analysis to human coding for VDO*

| VDO | Correlation | Dependent two-tailed t-test comparison |
|---|---|---|
| **100%** | *r* = -.58, *p* = .30 | *t*(4) = .63, *p* = .56 |
| **>50%** | *r* = .36, *p* = .38 | *t*(7) = .05, *p* = .96 |
| **~50%** | *r* = .20, *p* = .56 | *t*(10) = -1.11, *p* = .29 |
| **<50%** | *r* = .35, *p* =.11 | *t*(23) = - .10, *p* = .93 |
| **0%** | ***r = .33, p = .015*** | ***t(51) = - 3.42, p = .001*** |

**For levels of Infant Alone in the HV sample.** The correlations of LENA to human coding for volubility ranged from -.01 to .77; all but the ~50% IA were non-significant. There were sometimes significant differences noted between LENA and human coding. See Table 11.

**Table 11**
*Comparison of LC to HC infant volubility coding in the IA circumstance*

| IA | Correlation | Dependent two-tailed t-test comparison |
|---|---|---|
| **100%** | *r* = -.01, *p* = .99 | ***t*(8) = - 2.85, *p* = .02** |
| **>50%** | *r* = .37, *p* = .63 | ***t*(3) = - 3.39, *p* = .04** |
| **~50%** | ***r* = .77, *p* = .02** | ***t*(8) = - 3.17, *p* = .01** |
| **<50%** | *r* = .43, *p* =.19 | *t*(10) = - .31 *p* = .76 |
| **0%** | *r= .20, p = .11* | *t(66) =- .26, p = .79* |