Published in final edited form as:

Infant Behav Dev. 2018 February ; 50: 140–153. doi:10.1016/j.infbeh.2017.12.002.

Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings

Chia-Cheng Lee[1], **Yuna Jhang**[2], **George Relyea**[1], **Li-mei Chen**[3], and **D. Kimbrough Oller**[1,4]

[1]The University of Memphis, USA

[2]Chung Shan Medical University, Taiwan

[3]National Cheng Kung University, Taiwan

[4]The Konrad Lorenz Institute for Evolution and Cognition Research, Austria

Abstract

Canonical babbling (CB) is critical in forming foundations for speech. Research has shown that the emergence of CB precedes first words, predicts language outcomes, and is delayed in infants with several communicative disorders. We seek a naturalistic portrayal of CB development, using all-day home recordings to evaluate the influences of age, language, and social circumstances on infant CB production. Thus we address the nature of very early language foundations and how they can be modulated. This is the first study to evaluate possible interactions of language and social circumstance in the development of babbling. We examined the effects of age (6 and 11 months), language/culture (English and Chinese), and social circumstances (during infant-directed speech [IDS], during infant overhearing of adult-directed speech [ADS], or when infants were alone) on canonical babbling ratios (*CBR=canonical syllables/total syllables)*. The results showed a three-way interaction of infant age by infant language/culture by social circumstance. The complexity of the results forces us to recognize that a variety of factors can interact in the development of foundations for language, and that both the infant vocal response to the language/culture environment and the language/culture environment of the infant may change across age.

Keywords

infant; naturalistic home recordings; canonical babbling; social circumstance; age; cross-language

1. Introduction

1.1. Canonical babbling: A foundation for language

The minimal rhythmic units in spoken languages are syllables, the building blocks of words, and overwhelmingly these syllables are “canonical”, consisting of at least one consonant and one vowel. The ability to produce canonical syllables is critical in language because the vast majority of words in languages are composed of them. Thus the ability to produce canonical syllables must develop before the production of any substantial vocabulary of spoken words can be learned.

Infants in the first six months of life produce a wide variety of speech-like vocalizations, collectively called “protophones” (Oller, 2000), because they are precursors to mature phonology, clearly differentiable from cry, laughter, and vegetative sounds (e.g., sneeze or cough) (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980, 2000; Roug, Landberg, & Lundberg, 1989; Stark, 1980). These early protophones include quasi-vowels, full vowels, squeals, growls, yells, whispers, and raspberries. By the second half year, however, infants make the leap to frequent production of fully well-formed canonical syllables, having thus entered the stage of canonical babbling (CB, exemplified by well-formed syllables such as “mama” or “da”).

Longitudinal observations of infant development have confirmed that CB precedes the development of even a minimal vocabulary of spoken words. Oller, Levine, Cobo-Lewis, Eilers, and Pearson (1998) followed 42 infants and found a gap of > 8 months between the mean onset of CB (at 6 months) and the attainment of 5 words (at 14.5 months). Not a single infant in the study attained 5 expressive vocabulary items before the onset of CB—the minimal lag for an individual in the group was > 4 months.

The Oller et al. (1998) study deemed an infant to have reached the CB stage at the infant age reported in a telephone call by parents, if and only if the report of CB was confirmed by laboratory staff observations of regular occurrence of some CB in 5 laboratory sessions conducted during the following 2 weeks. This criterion aimed to designate the *earliest point* at which an infant manifested the capability for CB. In a more recent effort, McGillion et al. (2017) tracked 46 infants starting at 9 months, assessing CB a different way and designating a point of “babble onset” only after consistent production of two different supraglottal consonants in presumable canonical syllables (at least 10 occurrences in 3 of 4 consecutive recordings or at least 50 occurrences in one recording) that had been transcribed based on the samples. Babble onset was found using these measures to occur at a mean age of 9.8 months, preceding the attainment of 4 words (at 15.4 months) by > 5 months. Again, no infant in the study reached 4 words before the designated babble onset.

1.2. Canonical babbling: measure predicting developmental disorders

Determining age of CB onset has become standard practice scientifically and clinically because the emergence of CB clearly reflects a major advance in development of the speech capacity, and parents in interaction with infants are known often to treat canonical utterances as soon as they begin to occur as potential words (Papoušek, 1994). CB is salient to mature listeners, and parents have been found to be reliable informants in detecting its onset (Lewedag, 1995; Oller, Eilers, & Basinger, 2001; Oller, Eilers, Neal, & Cobo-Lewis, 1998).

Furthermore, there are sharp distinctions in onset of CB among groups of children with and without communication disorders. Profoundly hearing-impaired infants have been found to show very late onset of CB (Eilers & Oller, 1994; Koopmans-van Beinum, Clement, & van den Dikkenberg-Pot, 1998; Oller, Eilers, Bull, & Carney, 1985). Infants with autism (Patten et al., 2014), Down Syndrome (Lynch et al., 1995), Williams Syndrome (Masataka, 2001), cleft palate (Chapman, 2004; Chapman, Hardin-Jones, & Halter, 2003; Chapman, Hardin-Jones, Schulte, & Halter, 2001), and Fragile X syndrome (Belardi et al., 2017) have also been found to be delayed in CB onset.

Although infants with disorders have shown late occurrence of CB, the onset seems to be robust with regard to a variety of potential risk factors not accompanied by a diagnosed disorder. For example, CB has not been found to be delayed by premature birth (birth weights 1400–2100 g) in the absence of significant perinatal problems (Oller, Eilers, Steffens, Lynch, & Urbano, 1994). Törölä, Lehtihalmes, Heikkinen, Olsén, and Yliherva (2012) also found that preterm infants with extremely low birth weight (< 1000 g) showed similar CB onset to full-term infants, although after the onset of CB, they showed limited CB types and delayed first words. In addition, infants of low socio-economic status (SES) (Oller et al., 1994; Oller, Eilers, Urbano, & Cobo-Lewis, 1997) and very-low SES (Oller, Eilers, Basinger, Steffens, & Urbano, 1995) have not been found to be delayed in CB onset. Finally, both Spanish-learning monolinguals and English-Spanish bilinguals (Oller & Eilers, 1982; Oller et al., 1997) have been reported to show similar ages of CB onset to English-learning monolinguals.

1.3. Canonical babbling ratio: A widely-used convenient measure of infant vocal development

Evaluating whether an infant is in the CB stage can be determined in longitudinal tracking through parent report or through brief regular laboratory evaluation in most cases, and consequently CB can be used as a convenient scientific or clinical measure (Lanza & Flahive, 2008; Oller et al., 2001; Paul, Bruni, & Balweel, 2007). Other speech-like vocalizations of infancy (e.g., vowel-like sounds, isolated fricatives or affricates, raspberries) may be similarly tracked longitudinally and may also be important in development, but are not so commonly used in either research or clinical evaluation. A relatively fine-grained measure, the canonical babbling ratio (CBR=*canonical syllables/total syllables* in a recording session), takes account of the relative amount of CB and other speech-like syllables in samples of infant vocalization. Even though many studies have reported age of onset of CB (as cited above), the cost of such longitudinal research is high—note that both Oller et al. (1998) and McGillion et al. (2017) followed infants with many sampling dates and reviews of many recordings. CBR, in contrast, is a quantitative measure that can be obtained with a single recording followed by human coding, which can be reliably conducted in real time (Belardi et al., 2017; Patten et al., 2014). CBR can of course be used in longitudinal research as well, examining infants' vocal variability or developmental changes in the amount of CB and non-CB across recording sessions and ages.

Most recent research has designated an infant as having entered the CB stage (often by 5–6 months and very rarely later than 10 months of age) if the CBR is equal to or greater than 0.15 (Lynch et al., 1995) based on review of laboratory recordings. CBR is often taken into account in studies of various infant populations, such as infants with hearing impairment (Bass-Ringdahl, 2010), Down Syndrome (Lynch et al., 1995), autism (Patten et al., 2014), and Fragile X syndrome (Belardi et al., 2017). Some other studies using the CBR have examined Dutch-learning infants (Molemans, van den Berg, Van Severen, & Gillis, 2011), infants adopted from China (Price, Pollock, & Oller, 2006), Mandarin-learning infants (Chen & Kent, 2010), and infants with cleft palate (Jones, Chapman, & Hardin-Jones, 2003).

Researchers have also examined the relationship between CBR and later speech and language development. For example, Chapman, Hardin-Jones, and Halter (2003) evaluated infants with cleft palate. Given their anatomical limitations, it can be assumed that before corrective surgery, infants with cleft palate have particular difficulty in developing a variety of consonant-like onsets for canonical syllables. The researchers reported that typically-developing infants at 13 months showed a positive correlation between true canonical babbling ratio (TCBR=all canonical syllables except for those with glide or pharyngeal consonants/total syllables in a recording session) and speech and language development at 21 months. However, TCBR in infants with cleft palate before and after surgery did not significantly correlate with later speech and language measures other than mean length of utterance (MLU). This outcome was counterbalanced by the fact that infants' diversity of consonant-like onsets did indeed correlate positively with the speech and language measures. Chapman (2004) further found that TCBR in infants with cleft palate 9 months before the palatal surgery showed a *negative* correlation with MLU at 39 months. In contrast, a *positive* correlation between post-surgery TCBR at 13 months and MLU at 39 months was observed. The results suggest that velopharyngeal inadequacy before corrective surgery may thwart the typical positive prediction of speech outcomes by early CBR.

1.4. Possible influences on canonical babbling ratio: gap in the literature

CBR clearly reflects infant development of a phonological capacity because it differentiates among groups of typically developing and disordered infants. CBR is also a relatively convenient measure to obtain, and thus it may be one of the best candidates available as a potential predictor of language outcomes within groups. In order to place the possibility of prediction in perspective, it is important to evaluate CBR under a variety of conditions to assess possible influences on the measure that may need to be taken into account. We know that there is considerable within-infant variability in protophone measures, including CBR across recording sessions occurring in time frames within a single day or across as much as two weeks (Lewedag, 1995; Oller et al., 2007). It has been suggested that a critical factor in obtaining a stable measure is a relatively large a sample from each infant participant (Molemans et al., 2011). But sample size is clearly not the only issue that should be considered.

We propose to break from prior evaluations of CB, which have been conducted overwhelmingly in short samples, usually in laboratory settings, and address the issue as naturalistically as possible by randomly sampling segments from all-day home recordings. Further, we propose to study the natural complexity of babbling development by considering effects of factors that may play significant roles in CBR and that may also interact with each other across age in determining CBR. The factors we have in mind are major differences in the ambient languages/cultures and circumstances of social interaction or lack of social interaction. No prior research has considered the development of CB in terms of possible interactions among infant age, infant language, and social circumstances. In the absence of any prior evidence, it seems plausible that interactions affecting CBR might occur since different cultures may influence infant vocal development in different ways at different ages, deploying IDS in ways that are adjusted to infant behavior.

1.5. Effects of language and culture on infant CBR

While there have been several studies of CBR in infants with disorders, as indicated above, research on CBR in infants learning non-European languages is severely limited. CBR has been used as a measure in Chen and Kent (2010) and Price et al. (2006) to observe Mandarin-learning infants in Taiwan and infants adopted from China, respectively. Still, as far as we know, there has been practically no research comparing CBR *across* any two languages—Oller et al. (1997) evaluated CBR in monolingual English-learning infants and bilingual Spanish- and English-learning infants. But CBR in infants exposed to phonological systems as different as Chinese and English has never been evaluated to our knowledge.

Research has shown that acquiring knowledge of phonotactic rules of languages may help children segment the stream of speech (Friederici & Wessels, 1993; Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006). Compared to English, the syllable structures of Chinese are streamlined. While syllable structures in English *can* be as simple as V or CV (a consonant [C] and a vowel [V]), they can also be as complex as CCCVCCC as in ‘strengths’, with all the other possibilities in between (CCV, VCC, CVC etc.). In contrast, Mandarin syllables only consist of V, CV, CVC, and VC types (Li & Liu, 1988), and the final C’s are limited to the nasal sounds, [n] and [ŋ] only. The syllable structures of other Chinese languages such as Southern Min or Cantonese are more complex than that of Mandarin, but all the Chinese languages are still far less syllabically complicated than English. Also, Chinese languages have phonemic tones, while English does not. Therefore, the tremendous differences between Chinese and English phonology provide us an opportunity to investigate whether the relative complexity of syllable structure in the two languages plays a role in early vocal development in English- and Chinese-learning infants. It is reasonable to hypothesize the possibility that the simpler syllabic structure in Chinese might facilitate and accelerate CB development or result in higher CBR during development.

In addition to the effect of language on infant CBR, infants exposed to different cultures might be influenced to vocalize differently and to progress at different rates with regard to CBR. Maternal responsiveness has been found to facilitate language development (Tamis-LeMonda, Bornstein, & Baumwell, 2001), but maternal responsiveness appears to be different across cultures (Keller et al., 2007). When talking to infants, both English and Chinese mothers use infant-directed speech (IDS), which in both cultures presumably shares similar functions, such as eliciting attention or teaching language (Farran, Lee, Yoo, & Oller, 2016; Liu, Tsao, & Kuhl, 2007). However, these functions may occur in different amounts and thus may have different impacts on babies. As far as we know, comparative research on CBR has to date only included Spanish- and English-learning infants.

1.6. Effects of circumstance on infant CBR

Researchers have investigated effects of social circumstance (with parent vocal interaction, while overhearing adult-to-adult talk, while alone) on infant *volubility* (Bloom, Russell, &

Wassenberg, 1987; Delack, 1976; Franklin, 2014; Iyer, Denson, Lazar, & Oller, 2016; McGillion et al., 2013; Roe & Drivas, 1997; Shimada, 2012). For example, Iyer et al. (2016) reported that infants produced fewer vocalizations when adults were talking to each other than when adults were talking to the infants. Also, infants produced more vocalizations when adults were in the room but not talking to the infants than when adults were talking to each other.

However, *CBR* as a function of circumstance has been only indirectly addressed. Gros-Louis, West, Goldstein, and King (2006) indicated that mothers responded to infants with more consonant–vowel syllables when infants produced consonant–vowel syllables than when they produced vowel-like sounds. Mother's contingent vocal feedback to infant babbling elicited more infant vocalizations that resembled mother's speech (Goldstein & Schwade, 2008). Hsu and Fogel (2001) also reported that infant syllabic vocalizations (presumably canonical) occurred more often when mother and infant were mutually engaged than when they were not engaged.

Thus we have no direct evidence about whether infants who have entered the CB stage produce similar CBRs in different interactive circumstances. It is possible that infants produce more speech-like vocalizations when other people are interacting with them (Goldstein & Schwade, 2008), but it is also possible that infants produce higher CBR when they are alone, a circumstance that to our knowledge has never been evaluated for CBR. The extent to which infants produce speech-like vocalizations when alone would perhaps provide a measure of the endogenous inclination to play with sounds that lay a foundation for learning language. The fact that some researchers have found (controversially) highest *volubility* when infants are alone (see citations above), suggests the possibility that CBR may also be elevated in infants who are alone.

1.7. Study goals

In the present study, we observed CBR across social circumstances in typically-developing English- and Chinese-learning infants at both 6 and 11 months of age. The data were obtained through all-day recordings in the infant homes. Our intention was to assess the three factors (age, language, circumstance) in a single design, making it possible not only to evaluate possible main effects but also to assess the possibility of interactions among the factors. A main effect of age would confirm that the CBR measure is consistent with a wide variety of prior research showing growth in CB across time—lack of a significant age effect would suggest a fundamental problem with the way the measure was implemented in the study. Main effects for language and circumstance would suggest infants adapt their vocal patterns to both linguistic and interactive circumstances. Interactions among the variables at statistically significant levels would suggest that growth in CBR is not merely a maturational phenomenon, but may also by a complex manifestation of infant adaptation to environmental factors that may themselves adapt to the infant and to infant age. On the practical side, interactions among the factors would suggest that CBR should be expected to vary in complex ways, and thus its utility as a clinical measure may need to be adjusted to account for these factors and perhaps others.

1.8. Hypotheses

1. **Age:** Infants at 11 months will have higher CBR than at 6 months.
2. **Language:** Chinese-learning infants will have higher CBR than English-learning infants at both 6 and 11 months because of the relative simplicity of syllable structure in Chinese.
3. **Social Circumstance:**
   a. Infants at both 6 and 11 months will have higher CBR when interacting with caregivers (infant-directed speech, IDS) than when overhearing adult-to-adult talk (adult-directed speech, ADS).
   b. Infants at both 6 and 11 months will have higher CBR when overhearing adult-to adult talk (ADS) than when they are alone (Alone).
4. **Interaction:** There will be an interaction of infant age x infant language x social circumstance.

2. Methods

2.1. Infants and Recordings

LENA (Language ENvironment Analysis) all-day home recordings were drawn from archives of longitudinal studies on infants from 13 English and 1 Chinese (Mandarin and Cantonese) middle-class households in Memphis and 7 from Chinese (Mandarin and Southern Min) middle-class households in Tainan, Taiwan. The parents were asked to record on a typical day in the home, and to avoid circumstances where the recorder might be picking up voices in public. Regarding mother's educational level, 4 of the 13 English-speaking mothers had some college-level instruction, 6 had a bachelor's degree, 2 had a master's degree, and 1 had a Ph.D. degree. Seven of the eight Chinese-speaking mothers had a bachelor's degree and one had a master's degree. All the recordings were made with permission from the University of Memphis Institutional Review Board for the protection of human subjects. The National Cheng Kung University agreed in writing to abide by the constraints of the University of Memphis IRB. Parents were given the option of having any recording erased before analysis of any kind, but the option was rarely invoked.

The LENA recorder is 3-3/8 × 2-3/16 × 1/2 inches, weighs less than 2 ounces, and can be worn in an infant vest all day long (Ford, Baer, Xu, Yapanel, & Gray, 2008). Using this system, recordings were acquired in maximally naturalistic interactions between the infants, family members, and other people. LENA offers the best method available for the representative sampling of infant vocalizations in their natural environments. The system has been used widely for research purposes (Caskey, Stephens, Tucker, & Vohr, 2014; Caskey & Vohr, 2013; Ford et al., 2007; Oller et al., 2010; Warren, Gilkerson, Richards, & Oller, 2010; Yoder, Oller, Richards, Gray, & Gilkerson, 2013; Zimmerman et al., 2009). The present study was based primarily on human coding rather than automated labeling of utterances, although automated analysis has been used in most prior studies with the LENA system. It is widely understood (and acknowledged by the LENA Research Foundation) that current automated labeling is not nearly as accurate as that of human listeners. Our approach used the automated analysis only as a screening for 5-minute segments to be coded by humans. Our human listeners were able to code the quality of each recorded syllable in real time by keystrokes in the Action Analysis Coding and Training (AACT) software (Delgado, Buder, & Oller, 2010). This software has been used in much previous research on infant vocal development (e.g., Franklin et al., 2014; Iyer et al., 2016; Iyer, Oller, & Neal, 2007; Oller, Buder, Ramsdell, Warlaumont, Chorna, & Bakeman, 2013)

Recordings from 6 and 11 months were selected (21 infants × 2 ages = 42 recordings) by the first author. Ten 5-minute segments with the highest child vocalization rate (HV) determined by LENA automated analysis (Xu, Richards, & Gilkerson, 2014) and twenty 5-minute randomly-selected (RS) segments (a total of 30 segments/150 minutes) were selected from each all-day home recording for the study. The LENA system automatically supplies access to all the 5-minute segments of any recording. Thus, a total of 1,260 (21 infants × 2 ages × 30 segments) 5-minute segments were human coded. The recordings at 6 and 11 months were selected in the hope of observing babbling both at the beginning of the canonical stage and after it should have been well underway in both groups of infants.

All the infants were recruited by advertising or word-of-mouth. None of the infants had significant perinatal or hearing problems. There was no family history of speech or language issues. Eight of the thirteen English-learning infants and three of the eight Chinese-learning infants were males. Among the 13 English-learning infants, six were first-born, six were second-born, and one was third-born. As for the 8 Chinese-learning infants, only one infant was second-born, while the other seven infants were first-born. All the Chinese families in our study spoke Mandarin because Mandarin is the official language used in schools in Taiwan and China, but the ones living in Taiwan also spoke Southern Min, although much less frequently than Mandarin. A single Mandarin- and Cantonese-learning infant was born and lived in Memphis, but the parents took care of the infant at home during his first year of life and spoke Mandarin to the infant and to each other at home. Grandparents, who were also sometimes present, spoke Cantonese. The statistical analysis showed similar patterns with and without the Mandarin-Cantonese learning infant’s data.

Mandarin, Southern Min, and Cantonese are tone languages and share many linguistic characteristics, but they are not mutually understandable. The phonological systems of Southern Min and Cantonese are somewhat more complex than that of Mandarin. For example, Mandarin allows only/n/and/ŋ/in syllable-final position, while/n/,/ŋ/,/m/,/p/,/t/, and/k/are allowed at the end of syllables in Southern Min and Cantonese. Mandarin has four lexical tones: high-level, rising, falling-rising, and falling (Yip, 2002), while Southern Min has seven (Chung, 1996) and Cantonese has six tones (Ciocca & Lui, 2003). Southern Min is widely spoken in Taiwan and the Fujian province in China, while Cantonese is widely spoken in the Guangxi and Guangdong provinces and in Hong Kong.

2.2. Coders

Eight primary coders and three reliability coders participated in the coding tasks (10 females and 1 male). Seven of the eight primary coders and one of the reliability coders were monolingual English speakers, while the remaining coders (one primary coder and two reliability coders) spoke English and Mandarin. The primary Mandarin-English coder grew up in the U.S. and spoke Mandarin only at home. The two Chinese reliability coders grew up in Taiwan and were native speakers of Mandarin as well as speakers of English as a second language. All the coders were graduate students in the School of Communication Sciences and Disorders at the University of Memphis except for one of the Chinese reliability coders, who was a graduate student in the Foreign Languages and Literature Department at the National Cheng Kung University in Taiwan. None of the coders had a history of hearing loss. An ideal situation for coding would have included equal numbers of primary English- and primary Mandarin-speaking observers, with an additional group of balanced bilinguals. We did not have access to individuals meeting these criteria, but hope to develop a more balanced coding team for future research.

2.3. Coder training

The LENA automated analysis system does not determine the rate of CB, and thus human coding was necessary for the study. The emergence of CB is so salient that without training, parents have been found to be reliable informants in detecting its onset (Lewedag, 1995; Oller et al., 2001; Oller et al., 1998). Thus, judging CB vs. infant utterances without canonical syllables (non-CB) does not require much training. Our laboratory coders quickly grasped the intuitive idea of CB. Three one-hour training sessions were provided to the coders. In the first training session, the trainers (Lee and Oller) provided definitions and audio examples of utterances, syllables (i.e., CB and non-CB), cries, laughs, and vegetative sounds (e.g., sneezes, hiccups, burps, coughs, grunts due to body movements, or yawns). An utterance in our usage refers to a vocalization occurring within one breathing cycle/breath group, in accord with procedures defined in prior research (Lynch et al., 1995). Thus, an utterance may consist of a syllable, several syllables, a cry, a laugh, or a vegetative sound produced within one expiration. A syllable is the minimal rhythmic unit of an utterance. The coders were taught to count the number of syllables by listening to “beats” in the rhythm of vocalization, and during coding, they could supplement their judgments by observing changes in amplitudes of the waveform in the acoustic display, TF32 (Milenkovic, 2004).

A CB is “a rhythmic stereotypy involving jaw and/or tongue movement coordinated with vocal cord vibration” (Oller, 2000, p. 114). A CB consists of a consonant-like margin and vowel-like nucleus, forming a syllable meeting the following criteria: (1) a fully-resonant nucleus, (2) at least one supraglottally articulated (e.g., tongue, lip, or jaw) consonant-like element (i.e., a margin), and (3) a timely formant transition between the nucleus and the margin. Non-CB includes:

- **a.** Syllables lacking any margin (i.e., vowel-like sounds only);
- **b.** Syllables with vowel-like nuclei but no supraglottal articulation, such as syllables with glottal stops or fricatives only as margins
- **c.** Marginal babbles (MB) [if the syllable meets criteria (1) and (2) but not (3), OR (2) and (3) but not (1).] For example, a syllable is judged to be MB if the formant transition between the nucleus and the margin is slow (nominally < 120 ms), or if the vowel-like sound is not fully resonant (i.e., the nucleus is produced with the vocal tract at rest, yielding a “quasi-vowel” or “quasi-resonant nucleus’).

d. Syllables consisting throughout of supraglottally-generated sound sources such as in raspberries, isolated fricatives or affricates.

e. Cries, not coded, are distress vocalizations that are defined as having either 1)strong and long duration nuclei with dysphonation or very loud and tense vocal quality (we call this "wail" cry, or 2) short nuclei with a glottal burst at the onset or offset (a vocal type we call "whimper"). Cries also sometimes include a rapid ingress or "catch breath" at the end. Wails also sometimes include glottal bursts at the onset or offset. A glottal burst without a voiced nucleus sounds like a cough, also not coded. A key feature of cry is that it indicates infant distress or discomfort categorically. When the infant's stress level appears to be "out of control" in negative vocalization, coders typically agree easily that the utterances produced are cries. In such cases, the sounds are judged to be cries even though bursts may not be present (Yoo, Buder, Lee, & Oller, 2015).

f. Laughs, not coded, are treated as expressions of joy in infants and acoustically share characteristics with whimpers, although laughs tend to have nuclei that are distinctly non-nasal. Sometimes short laughs or whimpers are indistinguishable when coders cannot see the speakers' facial expression.

g. Vegetative sounds (sneeze, cough…), which are not coded either.

During the initial phase of training, the coders were assigned to listen to five audio samples with an acoustic display of waveform and spectrogram available. The samples were extracted from laboratory recordings of infants learning English. Each sample was 10- to 15-seconds long and composed of utterances with a total of 10–20 syllables. Cries, laughs, and vegetative sounds were not included in the samples. The coders counted the number of syllables (both CB and non-CB) while listening to each audio sample in real-time and wrote the count down after hearing each sample. They then listened to the same samples a second time, but this time, they listened in small chunks so that they could count the number of syllables more carefully in repeat listening, after which they wrote down a second count for each sample. Then they played the same sample a third time, counted the number of syllables in real-time while listening to the sample, and wrote down the count again. Thus, there were three counts for each sample of the total number of syllables. They then compared their syllable and CB syllable counts with a gold standard coding performed for the stimuli by Oller (the last author). The coders then compared one by one their coded syllables with those of the gold standard. The coders could play each audio sample as many times as they preferred, while they viewed the gold standard coding, and could read a corresponding commentary by Oller and Jhang (the second author).

Later, still in training, the coders listened to a 20-minute audio and video laboratory recording of an infant learning English. While listening to the recording, coders coded every infant syllable as either CB or non-CB in real time by keystrokes in the Action Analysis Coding and Training (AACT) software (Delgado, Buder, & Oller, 2010). AACT is the system that was also used in data collection once training was completed. Coding in real time within AACT allows coders to hear each utterance or syllable only once and make each judgment by a keystroke without stopping. CB coding in real time has been found in this task to be as reliable as the more usual but more time-consuming method of repeat-listening coding (Jhang, Yoo, & Oller, 2014) and the real-time coding method has been verified to yield group differences in CB (Patten et al., 2014). While coding during this training, the coders could see the waveform, spectrogram, and the videos of the speakers to help them differentiate the voices of the speakers. The coders were told not to code infant cries, laughs, or vegetative sounds. After coding, the coders determined the number of CB, non-CB, and CBR, and compared their numbers with the values the trainers provided. If the numbers did not fall within a criterion range (i.e., within the CBR values of the two gold-standard trainers, Oller and Jhang), the coders were required to repeat the process.

In a subsequent training session, the trainers reviewed the real-time coding results with the coders. Then the coders went on to the repeat-listening coding of the same 20-minute audio and video recording. In this pass, they indicated by cursor settings, the onset and offset of each infant syllable. In this case, the coders could listen to each infant utterance up to three times. Again, if the total numbers of CB, non-CB, and CBR of this 20-minute recording did not fall within the criterion range, the coders went back to review their codes and compare them with the gold standard coding. In cases of persistent disagreement with the gold standard, the coders set up an individual meeting with one of the trainers (Lee) to explain the discrepancies and improve the agreement level.

As a final training experience in preparation for data collection (which was based on LENA recordings), the coders were assigned to code in real-time five 5-minute segments extracted from LENA recordings of infants learning English and Chinese. After reviewing the coding results, the trainer (Lee) gave feedback to the coders based on individual discrepancies from gold standard coding by the trainers. After the feedback, the coders coded another ten 5-minute LENA segments of infants learning English and Chinese. At this point, the coders were deemed to have completed the CB coding training because the segment-level correlations between the CBR for the ten segments obtained by the primary coders and the three reliability coders were all above 0.8.

2.4. Coding procedure

Five primary coders were assigned two recordings (one at 6 and one at 11 months) from each of three infants, while the other three primary coders were assigned two recordings from two infants each (42 recordings in total). The 30 five-minute segments that were selected from each all-day recording were presented for coding in chronological order to help coders recognize contexts and thus more accurately identify speakers and circumstances. However, the order of presentation of recordings was randomized for each coder. A total of 1,260 segments were thus coded, blocked in groups of 30 according to the recording from which they were drawn.

The reliability coders, who were not among the 8 English coders, coded 10% of all the recordings (distributed across all infants at both 6 and 11 months and both languages) in real time independently. The segment-level correlation between the CB ratios for segments with at least 6 coded syllables obtained by the primary coders and the three reliability coders was 0.90 and 0.89, and 0.84, respectively.

All the coders filled out a questionnaire every time they listened to a 5-minute segment. The questions were: 1) Does any other person talk to the baby? This can be the parent or another adult or child. 2) Does any other person talk to someone else? For example, 2 adults might be talking to each other or a person might be talking on the phone. 3) Do you think the baby is asleep?

These questions provided information about the amount of time people talked to infants (infant-directed speech, IDS) or among each other (adult-directed speech, ADS), and whether infants were asleep (Sleep) within the 5 minutes. Coders used a 5-point scale indicating the amount of time during each segment that pertained to each circumstance. 1 meant no such activity, 2 some of the time, 3 half of the time, 4 most of the time, and 5 close to the whole time in each 5-min segment. Even though seven of the primary coders did not know Chinese, they still filled out the questionnaire after listening to each Chinese segment. The reliability coders also filled out a questionnaire every time they listened to a segment, that is, for 10% of all the segments coded for the research. The agreement as indicated by segment-level correlations between the primary coders and the reliability coders for the three questions was 0.75, 0.82, and 0.75.

2.5. Data analysis

The CBR (i.e., CB syllables/the number of all syllables) for each 5-minute segment was calculated. We eliminated from the circumstance categorization all segments with very low vocal activity, i.e., < 6 coded syllables, in order to prevent skewing of the data by segments where the infant was scarcely vocally active. For segments with 6 or more syllables, we assigned each segment to a circumstance: IDS, ADS, IDSADS (a mixture of IDS and ADS), and Alone. The rules to assign segments to circumstances were based on the 1 to 5 scaled answers on the questionnaire (See Fig. 1):

Rule 2 was used to identify 5-minute segments in which only or predominantly IDS was present, while rule 3 was used to identify 5-minute segments in which only or predominantly ADS was present. After identifying segments with predominantly IDS or ADS, rule 4 was used to identify the remaining 5-minute segments with some adult talk as IDSADS, that is, cases where similar amounts of IDS and ADS were present. Finally, rule 5 was used to identify the rest of the segments as Alone, but only if the infant was awake some of the time during the segment. There were other procedures that could have been used to categorize the data for circumstance based on the questionnaire answers, and alternative analyses based on two other procedures are considered below.

We began with a total of 1,260 five-minute segments to potentially be used in the circumstance analysis. Forty RS 5-minute segments had also been independently selected as HV based on the LENA automated analysis, so we included them only once in the analysis. We also excluded 396 segments that had fewer than six coded syllables (including segments where the infant slept the whole time). Also, an outlier 5-minute segment from one of the Chinese-learning infants' recordings at 11 months was excluded because it had an inordinately high CBR of .6. See Table 1 for average occurrence of the circumstances (in proportion of the 150 minutes, i.e., 30 coded 5-min segments) per infant at each age per all-day home recording. After exclusions, 823 segments remained for circumstance analysis.

2.6. Statistical analysis

Our mixed-model analysis (SAS Proc Mixed) included one between-subjects variable (infant language background [Chinese or English]), two within-subjects variables (infant age [6 and 11 months] and social circumstance [IDS, IDSADS, ADS, and Alone]), one random variable (individual infants), and one dependent variable (CBR) for 823 segments. Infant age was also considered a repeated measure in the model. The approach is based on West, Welch, and Galecki (2014).

3. Results

As presented in Table 2, the mixed model analysis showed a significant main effect of infant age [$F(1, 733) = 44.25$, $p < .001$], a significant two-way interaction of infant age by social circumstance [$F(3, 733) = 4.72$, $p < .01$], and a significant three-way interaction of infant age by infant language by social circumstance [$F(3, 733) = 5$, $p < .01$]. All other main effects and interactions were not significant.

In order to evaluate the robustness of the pattern of results, and in particular the three-way interaction, we tried two additional breakdowns of social circumstance, and the mixed model results were similar for all these analyses, both showing the significant three-way interaction. In the first additional analysis, we categorized a segment as pertaining to the IDS circumstance if IDS was present (according to the questionnaire answers for the segment) and ADS was absent in the 5-minute segment, regardless of the amount of IDS. Similarly, we categorized a segment as pertaining to the ADS circumstance if ADS was present and IDS was absent, regardless of the amount of ADS. If both IDS and ADS were present in a segment, no matter how much of each, it was categorized as IDSADS. The remaining segments were categorized as Alone segments. In the second additional analysis, we categorized a segment as pertaining to the IDS circumstance when (1) when IDS was present and ADS was absent, or (2) in cases where both IDS and ADS were present, IDS occurred more often than ADS. Similarly, we categorized a segment as pertaining to the ADS circumstance (1) when ADS was present and IDS was absent, or (2) in cases where both IDS and ADS were present, ADS occurred more often than IDS. If the amount of IDS and ADS occurred was equal in accord with the questionnaire answers, the segment was categorized as IDSADS. We prefer the categorization scheme reported in the main text because it balances the amounts of IDS and ADS in obtaining an IDSADS group, but as noted, all three analyses yielded the three-way interaction.

In an additional test to evaluate the robustness of the pattern of results, we bootstrapped/resampled the 21 individual infant's CBRs with replacement from the original 13 English- and 8 Chinese-learning infant's CBRs 200 times and ran the same mixed model analysis for the 200 bootstrapped samples in SAS. This sampling with replacement resulted in cases where any particular infant's CBR could appear multiple times or not at all, and also a different number of English- and/or Chinese-learning infants could be represented in any bootstrapped sample. For 148 (74%) of the 200 bootstrapped samples the three-way interaction was statistically significant.

We now consider the outcomes for the mixed model and each of the hypotheses along with graphic illustrations. Of course the highest order interaction (the three-way interaction) takes precedence in our interpretations.

Fig. 2 shows the three-way interaction, which corresponds to several complexities. The figure illustrates the CBR for the two groups of infants at 6 and 11 months in the four circumstances. The left-hand (red) bars represent the CBR and standard errors of the two groups of infants at 6 and 11 months when hearing IDS, the second from the left (purple) bars represent CBR of infants hearing IDSADS (a mixture of IDS and ADS), the third from the left (blue) bars represent CBR of infants hearing ADS, and the right hand (yellow) bars represent CBR of infants when Alone.

The only simple result to report is that CBR of both groups of infants at 11 months in each circumstance was higher than their CBR at 6 months, a significant age effect, as expected in Hypothesis 1, and as noted previously, confirmation of this hypothesis is required as an indication that our CB measurements were reliable. But age interacted with the other variables, a fact that could not have been predicted based on any prior research. One way to depict the three-way interaction is to say that every social circumstance corresponded to one language group having higher CBR at one age, while the other circumstance did not show higher CBR for the same language group at the other age. For example, the difference between CBR for English and Chinese infants during IDS was very minimal at 6 months, while that difference was very large at 11 months. For ADS, the Chinese infants had higher CBR than English infants at 6 months but lower CBR at 11 months. These examples make clear that there was no simple language/culture effect as suggested in Hypothesis 2: Instead language/culture interacted with age and circumstance.

Among the biggest surprises of the study, the circumstance of IDS did not correspond simply to higher CBR (Hypothesis 3a). In fact, the Chinese-learning infants at 11 months, produced the only comparison consistent with the hypothesis. Inconsistent with the hypothesis, the Chinese-learning infants at 6 months and the English-learning infants at both 6 and 11 months had higher CBR during ADS than IDS. All the comparisons showed higher CBR in ADS than in Alone (consistent with Hypothesis 3b), but the trend was not strong enough to be reflected in a main effect of circumstance.

Fig. 3 presents a collapsing of the data in Fig. 2 to illustrate the clear age effect and the pattern of the CBR outcome for language/culture. The slightly higher CBRs for Chinese-learning infants were of course not statistically significant.

Fig. 4 similarly provides a collapsing of the data to illustrate the interaction between circumstance and age. There is more than one way to consider the interaction. On the one hand, the ranking by circumstance of CBR at 6 months from highest to lowest was ADS > IDSADS > IDS > Alone, but the ranking of CBR at 11 months was IDS > ADS > IDSADS > Alone. On the other hand, the interaction also reflects the fact that the difference between 6 and 11 month CBRs was more prominent for IDS, IDSADS, and ADS than for the Alone circumstance.

4. Discussion

4.1 Multiple effects on infant CBR

Canonical babbling (CB) is a precursor to speech, but no prior studies have evaluated factors that might influence the amount of CB infants produce. The occurrence of CB in natural settings has thus not been evaluated to assess effects of social and linguistic circumstances across time. The present study investigated the effects of these factors in order to determine if they might interact with each other across age in determining CBR. While infant *volubility* as a function of circumstance has been extensively studied (see citations above), ours is the first to empirically investigate whether cultural or linguistic differences correspond to different amounts of highly speech-like vocalizations, namely, CB across circumstances. Both American English and Chinese groups showed higher CBR at 11 months, suggesting, as we should expect, that the effect of age on the development of CB was strong. But age interacted with the other factors, and social circumstances showed different relative amounts of CB for both languages/cultures at both ages.

The complexity of the interaction may be partly attributable to differences in home environments and parenting attitudes in the US and Taiwan, differences that may interact with age and tendencies to engage infants in vocal exchange. In other words, a possible interpretation for the complexity of the interaction invokes the possibility that parenting practices (including the tendency to vocally engage the infant) change across time in different ways across the cultures. This idea must at this point remain speculative, but we do have at least minimal evidence to begin to flesh out such speculation based on possible cultural/language differences as reflected in Table 1 compared with the CBR data in Fig. 2. Consider, for example, the relatively high CBR for Chinese infants at 11 months relative to the American infants at the same age. One might imagine that IDS would be more common for Chinese households at that age than for American households, but Table 1 suggests the opposite tendency, with less IDS occurring in the Chinese than the American households at 11 months (a tendency that was not present at 6 months). Continuing to speculate, one might suggest that Chinese-speaking parents with 11 month-old infants were less assertive than American parents in using IDS, and that the Chinese-learning infants responded by producing higher CBRs, perhaps trying to elicit more speech interaction from the parents. Along a similar vein, one might speculate that American parents, who tended to produce more IDS than Chinese parents at 11 months, had an inhibitory effect on CBR in their infants. Additionally, at 6 months IDS was the most common circumstance for both language/culture groups, but CBR was higher in the other circumstances at 6 months, again hinting that perhaps higher CBR was encouraged by circumstances where parents were less inclined to produce IDS.

It seems likely that there were other influences on the three-way interaction that might have occurred due to factors we did not monitor, such as size of the household (Coale, Fallers, & King, 2017). It is well known that American household space tends to be large, around twice as large as Taiwanese (U.S. Department of Housing and Urban Development and U.S. Census Bureau, 2005; National Statistical Report of Taiwan, 2010a), and the inclination to interact vocally might thus be affected by likelihood of family member proximity. Further, the number of family members in households may differ across the language/cultures—Lofquist, Lugaila, O'Connell, and Feliz (2010) and National Statistical Report of Taiwan (2010b) suggest that Taiwanese households tend to have 25% more occupants (i.e., Memphis: 2.48 vs Tainan: 3.4). The evidence in Table 1 suggests that the Chinese-learning infants at least at 11 months spent less time alone than English-learning infants. Mindell, Sadeh, Kohyama, and How (2010) indicated that 57% of (> 9000) families from predominantly-Caucasian countries or regions (e.g., Australia, United States, United Kingdom, Canada, New Zealand) reported that they and their infants went to bed in separate rooms, while only 4% of (> 20,000) families from predominantly-Asian countries (e.g., India, China, Taiwan, Hong Kong, Singapore, Thailand) reported that infants slept independently in a room. Mindell et al.'s (2010) finding on parent-infant sleeping arrangement in predominantly-Asian and predominantly-Caucasian countries partially conforms to our dataset because the Chinese-learning infants at both 6 and 11 months also showed less time Alone than the English-learning infants.

Still, sorting out the influences on the three-way interaction will clearly require much more research. We did not directly monitor factors such as amount of household space, number of family members, infant gender, infant birth order, infant temperament (Gartstein et al., 2006; Hsu, Soong, Stigler, Hong, & Liang, 1981), or individual behavioral differences (Colombo & Fagen, 2014) which we intend also to balance between groups in future studies. The most important conclusion to be drawn here is that the development of the capability for CB, which is clearly of critical importance in laying foundations for language, seems to reflect a highly adaptive process involving infant responsivity to the social and linguistic environment, which itself appears to adapt to the infant across time.

4.1 Effects of language/cultural environment on infant CBR

Infants growing up with different language backgrounds produce similar canonical syllable types (CB, e.g., ba, ma, da). But do infants learning a language with relatively simple syllable structure produce higher CBR (*CBR=CB syllables/total syllables*)? The onset seems to be robust and is not delayed in Spanish-speaking families or with English-Spanish bilingualism in the home (Oller & Eilers, 1982; Oller et al., 1997). Here we investigated whether the relative simplicity of syllables in Chinese might cause CB development to be accelerated in Chinese-learning infants. The results provided no strong evidence of higher CBR in Chinese-learning infants although they had higher CBR at 6 and 11 months than the English-learning infants did. The Chinese-learning infants showed a quantitative advantage in CBR but it was not statistically significant.

It is worth noting that research on the effect of ambient language differences on babbling has focused not on CBR but on the phonotactic, phonetic or acoustic content of babbling (Kern, Davis, & Zink, 2009; Levitt & Wang, 1991; Vihman & de Boysson-Bardies, 1994; Whalen, Levitt, & Goldstein, 2007). We did not monitor the content of CB types (e.g., CV, CVC…) produced by the two groups of infants in the present study, but in another study (Lee, Jhang, Chen, Relyea, & Oller, 2017), we investigated possible ambient-language effects in babbling in English- and Chinese-learning infants at 8, 10, and 12 months. We found that neither English nor Chinese adult listeners could tell if a non-canonical utterance was produced by an English- or a Chinese-learning infant. Only utterances with CB or identified as words showed a small but discernible tendency to be identifiable by language. Subsequent work will evaluate the possibility that the phonotactic, phonetic or acoustic content of CB may vary in accord with language/culture and social/interactive circumstances.

4.2 Effects of social circumstance on infant CBR

While CBR has not been previously studied directly in research on social interactive circumstances, there is indirect evidence of relevance. For example, in Gros-Louis, West, Goldstein, and King (2006), mothers responded to infants with more consonant–vowel syllables when infants produced consonant–vowel syllables than when they produced vowel-like sounds. Goldstein and Schwade (2008) also found that mother's contingent vocal feedback to infant babbling elicited more infant vocalizations that resembled mother's speech. Following Gros-Louis et al.'s (2006) and Goldstein and Schwade (2008) findings, the present study hypothesized that as infants start to produce CB, parents may respond to those infants with more advanced vocalizations, such as words (e.g., 'mama' or 'mommy') or nonsense CB, or tend to imitate infants' CB. With more advanced vocalizations from parents, infants, we reasoned, might produce more CB in return. Therefore, the highest CBR was hypothesized to occur in the IDS (infant-directed speech) circumstance (i.e., Hypothesis 3a), an anticipated main effect of IDS. The results were not consistent with Hypothesis 3a. What occurred instead was a three-way interaction of infant age by infant language by social circumstance. At 6 months the two language groups showed similar CBR in IDS, but at 11 months a substantially higher CBR was seen for the Chinese-learning infants over the English-learning infants. The pattern for ADS (adult-directed speech) was, however, very different, with Chinese-learning infants showing higher CBR at 6 months, and English-learning infants showing higher CBR at 11 months. The Chinese-learning infants at 11 months showed higher CBR when hearing IDS than ADS, which was consistent with Hypothesis 3a, but the advantage was insufficient to support the hypothesis because they were the *only* group (by age and language) that showed that advantage. The outcome was far more complex than Hypothesis 3a anticipated.

We also hypothesized that infants would produce more CB when overhearing speech than when alone (i.e., Hypothesis 3b). The results showed only a trend consistent with Hypothesis 3b (there was no significant main effect of circumstance in spite of higher CBR in ADS than Alone for all four age/language comparisons). The Alone circumstance may often correspond to a low infant arousal state and thus may not encourage infants to produce advanced vocalizations. However, the infants were not silent when alone. Even though they were not engaged in interaction, infants produced both canonical and non-canonical syllables, presumably endogenously or in imitation of syllables the infant may have heard. The tendency to explore vocalization when alone suggests vocal play, an activity that has been posited to lay a foundation for language (Oller et al., 2013; Stark, 1980).

One possible explanation for why the results were predicted by Hypothesis 4 but not predicted by Hypotheses 3 concerns differences in the recordings made in laboratory or home settings. Our predictions were based in part on the findings of Gros-Louis et al. (2006). But in that study, parents were *instructed* to interact with their infants in the laboratory setting of the study for ten minutes. Within the ten minutes, only the parents and infants were in the room, and they were fully engaged in face-to-face interaction. In our all-day home recordings, however, adults were *not* instructed to interact with their infants. The parents may of course have interacted with their infants intermittently, and the length of the interactions varied. The parents' faces may not even have been visible to the infants when the parents interacted with them on many occasions. In addition, the infants may have been distracted by sounds from the TV or other people walking into the room. Thus, the circumstances in Gros-Louis et al. (2006) were not comparable with those of our study, and the results may have been correspondingly affected. Our approach emphasizes the need for both laboratory research and naturalistic observation based on all-day home recordings.

4.3 Limitations and future directions

In the present study, we used real-time coding. In the future, we plan also to locate both infant and adult vocalizations using the repeat-listening method (where coders listen to an infant utterance multiple times and set boundaries at the onset and offset of each utterance). If infant and adult vocalizations are coded this way, we will have the opportunity to evaluate response lags of both infant and caregiver and to assess the possible effect of canonical and non-canonical vocalizations on parental responsivity. In addition, our primary coders in the present study were mostly English speakers. Only one of the primary coders knew Chinese. Although the coder agreement between Chinese and English coders was high, we will seek to balance language background of coders in future studies.

There are serious challenges to interpretation of audio-only materials. In the future, we hope for all-day video and audio in the home, but technological (head-worn cameras, storage of video…) and social problems (invasion of privacy) will have to be surmounted. With audio only, we cannot determine the number of people present with the infant and in many cases cannot determine who they were (parent, sibling, friend…). For example, in the IDSADS segments, we often did not know how many people were talking to the infants nor how many people were talking to each other. These patterns are especially difficult to interpret when there are guests in the room. Also, it is difficult to judge the extent of an infant's aloneness during a segment without video. The fact that English and Chinese differ dramatically not only in syllable structure but also in the presence of phonemic tones in Chinese but not in English was directly addressed in this study. Future research might consider this difference by including more detailed coding to characterize internal syllable structure and tone.

5. Conclusion

We set about to evaluate the possibility that CBR might vary with age, language/cultural background, and social circumstance. The results give us pause in reflecting on the simple hypotheses that have previously been investigated regarding such phenomena as ambient language effects (Engstrand, Williams, & Lacerda, 2003; Lee et al., 2017) and the growth of CBR across age in various groups of infants (Oller & Eilers, 1982; Oller et al., 1994; Törölä et al., 2012). In fact the three factors (age, language/culture, social circumstance) interacted significantly, greatly complicating possible interpretations, and suggesting that the development of babbling is sensitive to social circumstances and that social circumstances may also adapt to infant development across time. The simple syllables of Chinese might, we thought, facilitate early CB and high CBR compared to the complex syllables of English. We also sought to determine simple effects of interactive circumstances on CBR in both groups of infants. But the results suggested a complex evolving system of vocalization and social niche adaptation with no simple effects of language/culture or social circumstance.

Acknowledgments

Funding

This research was supported by NIH R01 DC011027 and R01 DC006099 to D. Kimbrough Oller, PI; the Chiang Ching-Kuo Foundation for International Scholarly Exchange in Taiwan RG005-D-10 to Li-mei Chen, PI; and Government Scholarship to Study Abroad from the Ministry of Education in Taiwan.

Thanks to Xuzi Zhou at Dell EMC for technical assistance

References

Bass-Ringdahl, SM. The relationship of audibility and the development of canonical babbling in young children with hearing impairment; Journal of deaf studies and deaf education. 2010. p. 287-310.https://doi.org/10.1093/deafed/enq013

Belardi K, Watson LR, Faldowski RA, Hazlett H, Crais E, Baranek GT, … Oller DK. A retrospective video analysis of canonical babbling and volubility in infants with Fragile X Syndrome at 9–12 months of age. Journal of Autism and Developmental Disorders. 2017; :1–14. DOI: 10.1007/s10803-017-3033-4 [PubMed: 27696184]

Bloom K, Russell A, Wassenberg K. Turn taking affects the quality of infant vocalizations. Journal of Child Language. 1987; 14(02):211–227. https://doi.org/10.1017/S0305000900012897. [PubMed: 3611239]

Caskey M, Stephens B, Tucker R, Vohr BR. Adult talk in the NICU with preterm infants and developmental outcomes. Pediatrics. 2014; 133:e578, e584.doi: 10.1542/peds.2013-0104 [PubMed: 24515512]

Caskey M, Vohr BR. Assessing Language and Language Environment of High Risk Infants and Children: a New Approach. Acta Pediatrica. 2013; 102:451–461. DOI: 10.1111/apa.12195

Chapman KL. Is presurgery and early postsurgery performance related to speech and language outcomes at 3 years of age for children with cleft palate? Clinical Linguistics & Phonetics. 2004; 18(4–5):235–257. http://dx.doi.org/10.1080/02699200410001693486. [PubMed: 15259571]

Chapman KL, Hardin-Jones M, Halter KA. The relationship between early speech and later speech and language performance for children with cleft lip and palate. Clinical Linguistics & Phonetics. 2003; 17(3):173–197. http://dx.doi.org/10.1080/0269920021000047864. [PubMed: 12858838]

Chapman KL, Hardin-Jones M, Schulte J, Halter KA. Vocal development of 9-month-old babies with cleft palate. Journal of Speech, Language, and Hearing Research. 2001; 44(6):1268–1283. DOI: 10.1044/1092-4388(2001/099)

Chen LM, Kent RD. Segmental production in Mandarin-learning infants. Journal of Child Language. 2010; 37(02):341–371. DOI: 10.1017/S0305000909009581 [PubMed: 19490748]

Chung, R-f. The segmental phonology of Southern Min in Taiwan. Taipei, Taiwan: Crane Publishing; 1996.

Ciocca V, Lui J. The development of the perception of Cantonese lexical tones. Journal of Multilingual Communication Disorders. 2003; 1(2):141–147. http://dx.doi.org/10.1080/1476967031000090971.

Coale, AJ., Fallers, LA., King, PB. Aspects of the analysis of family structure. Princeton, NJ: Princeton University Press; 2017.

Colombo, J., Fagen, J. Individual differences in infancy: Reliability, stability, and prediction. New York, NY: Psychology Press; 2014.

Delack JB. Aspects of infant speech development in the first year of life. Canadian Journal of Linguistics/Revue canadienne de linguistique. 1976; 21(01):17–37. https://doi.org/10.1017/S0008413100008033.

Delgado, RE., Buder, EH., Oller, DK. Action analysis coding and training (AACT). Miami, FL: Intelligent Hearing Systems; 2010.

Eilers RE, Oller DK. Infant vocalizations and the early diagnosis of severe hearing impairment. The Journal of Pediatrics. 1994; 124(2):199–203. http://dx.doi.org/10.1016/S0022-3476(94)70303-5. [PubMed: 8301422]

Engstrand O, Williams K, Lacerda F. Does babbling sound native? Listener responses to vocalizations produced by Swedish and American 12-and 18-month-olds. Phonetica. 2003; 60(1):17–44. DOI: 10.1159/000070452 [PubMed: 12802087]

Farran LK, Lee CC, Yoo H, Oller DK. Cross-cultural register differences in infant-directed speech: an initial study. PloS one. 2016; 11(3):e0151518. http://dx.doi.org/10.1371/journal.pone.0151518. [PubMed: 26981626]

Ford, M., Baer, CT., Xu, D., Yapanel, U., Gray, S. The LENA™ language environment analysis system: Audio specifications of the DLP-0121. Boulder, CO: Lena Foundation; 2008. Retrieved from https://pdfs.semanticscholar.org/44d1/08871b090c846d40fb1c096cdd279a627c2c.pdf

Franklin, BS. Unpublished doctoral dissertation. University of Memphis; TN: 2014. Infant volubility across circumstances estimated from all-day recordings.

Franklin BS, Warlaumont AS, Messinger DS, Bene ER, Iyer SN, Lee CC, … Oller DK. Effects of parental interaction on infant vocalization rate, variability and vocal type. Language Learning and Development. 2014; 10(3):279–296. doi: http://dx.doi.org/10.1080/15475441.2013.849176. [PubMed: 25383061]

Friederici AD, Wessels JM. Phonotactic knowledge of word boundaries and its use in infant speech perception. Attention, Perception, & Psychophysics. 1993; 54(3):287–295. DOI: 10.3758/BF03205263

Gartstein MA, Gonzalez C, Carranza JA, Ahadi SA, Ye R, Rothbart MK, Yang SW. Studying cross-cultural differences in the development of infant temperament: People's Republic of China, the United States of America, and Spain. Child Psychiatry and Human Development. 2006; 37(2): 145–161. DOI: 10.1007/s10578-006-0025-6 [PubMed: 16874564]

Goldstein MH, Schwade JA. Social feedback to infants' babbling facilitates rapid phonological learning. Psychological Science. 2008; 19(5):515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x. [PubMed: 18466414]

Gros-Louis J, West MJ, Goldstein MH, King AP. Mothers provide differential feedback to infants' prelinguistic sounds. International Journal of Behavioral Development. 2006; 30(6):509–516. https://doi.org/10.1177/0165025406071914.

Hsu, C-c, Soong, W-t, Stigler, JW., Hong, C-c, Liang, C-c. The temperamental characteristics of Chinese babies. Child Development. 1981; :1337–1340. DOI: 10.2307/1129528 [PubMed: 7318527]

Hsu HC, Fogel A. Infant vocal development in a dynamic mother-infant communication system. Infancy. 2001; 2(1):87–109. DOI: 10.1207/S15327078IN0201_6

Iyer SN, Denson H, Lazar N, Oller DK. Volubility of the human infant: Effects of parental interaction (or lack of it). Clinical Linguistics & Phonetics. 2016; 30(6):470–488. http://dx.doi.org/10.3109/02699206.2016.1147082. [PubMed: 27002533]

Iyer SN, Oller DK, Neal R. On the robustness of vocal development: an examination of infants with moderate-to-severe hearing loss and additional risk factors. Journal of Speech, Language, and Hearing Research. 2007; 50:1425–1444. DOI: 10.1044/1092-4388(2007/099)

Jhang, Y., Yoo, H., Oller, DK. Canonical babbling judged in real-time as a screening tool. Paper presented at the American Speech-Language-Hearing Association (ASHA) Convention; Orlando, FL. 2014.

Jones CE, Chapman KL, Hardin-Jones MA. Speech development of children with cleft palate before and after palatal surgery. The Cleft palate-craniofacial journal. 2003; 40(1):19–31. http://dx.doi.org/10.1597/15451569(2003)040<0019:SDOCWC>2.0.CO;2. [PubMed: 12498602]

Keller H, Abels M, Borke J, Lamm B, Su Y, Wang Y, Lo W. Socialization environments of Chinese and Euro-American middle-class babies: Parenting behaviors, verbal discourses and ethnotheories. International Journal of Behavioral Development. 2007; 31(3):210–217. https://doi.org/10.1177/0165025407074633.

Kern, S., Davis, B., Zink, I. From babbling to first words in four languages. In: D'Errico, F., Hombert, JM., editors. Becoming Eloquent: Advances in the Emergence of Language, Human Cognition, and Modern Cultures. 2009.

Koopmans-van Beinum, FJ., Clement, CJ., van den Dikkenberg-Pot, I. Influence of lack of auditory speech perception on sound productions of deaf infants. Paper presented at the International Society for the Study of Behavioral Development; Berne, Switzerland. 1998.

Koopmans-van Beinum, FJ., van der Stelt, JM. Early stages in the development of speech movements. In: Lindblom, B., Zetterstrom, R., editors. Precursors of Early Speech. Hampshire, England: The Macmillan Press; 1986. p. 37-50.

Lanza, J., Flahive, L. Guide to communication milestones. East Moline, IL: LinguiSystems; 2008.

Lee CC, Jhang Y, Chen L-m, Relyea G, Oller DK. Subtlety of ambient-language effects in babbling: A study of English- and Chinese-learning infants at 8, 10, and 12 months. Language Learning and Development. 2017; 13(1):100–126. DOI: 10.1080/15475441.2016.1180983 [PubMed: 28496393]

Levitt AG, Wang Q. Evidence for language-specific rhythmic influences in the reduplicative babbling of French-and English-learning infants. Language and Speech. 1991; 34(3):235–249. DOI: 10.1177/002383099103400302 [PubMed: 1843525]

Lewedag, VL. Unpublished doctoral dissertation. University of Miami; FL: 1995. Patterns of onset of canonical babbling among typically developing infants.

Li, G., Liu, R. A dictionary of Chinese character information. Beijing, China: Science Press; 1988.

Liu HM, Tsao FM, Kuhl PK. Acoustic analysis of lexical tone in Mandarin infant-directed speech. Developmental Psychology. 2007; 43(4):912. [PubMed: 17605524]

Lofquist, D., Lugaila, T., O'Connell, M., Feliz, S. Households and Families: 2010. United States Census Bureau; 2010. Retrieved from https://www.census.gov/prod/cen2010/briefs/c2010br-14.pdf

Lynch MP, Oller DK, Steffens ML, Levine SL, Basinger DL, Umbel V. Onset of speech-like vocalizations in infants with Down syndrome. American Journal on Mental Retardation. 1995; 100(1):68–86. [PubMed: 7546639]

Masataka N. Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate limiting constraint on the emergence of canonical babbling. Developmental Science. 2001; 4(2):158–164. DOI: 10.1111/1467-7687.00161

McGillion ML, Herbert JS, Pine JM, Keren-Portnoy T, Vihman MM, Matthews DE. Supporting early vocabulary development: What sort of responsiveness matters? IEEE Transactions on Autonomous Mental Development. 2013; 5(3):240–248. DOI: 10.1109/TAMD.2013.2275949

McGillion ML, Herbert JS, Pine JM, Vihman MM, DePaolis R, Keren-Portnoy T, Matthews DE. What paves the way to conventional language? The predictive value of babble, pointing, and socioeconomic status. Child Development. 2017; 88(1):156–166. DOI: 10.1111/cdev.12671 [PubMed: 27859008]

Milenkovic, P. TF32 [Computer Program]. Madison, WI: 2004.

Mindell JA, Sadeh A, Kohyama J, How TH. Parental behaviors and sleep outcomes in infants and toddlers: a cross-cultural comparison. Sleep medicine. 2010; 11(4):393–399. http://dx.doi.org/10.1016/j.sleep.2009.11.011. [PubMed: 20223706]

Molemans I, van den Berg R, Van Severen L, Gillis S. How to measure the onset of babbling reliably? Journal of Child Language. 2011; 39(03):523–552. https://doi.org/10.1017/S0305000911000171. [PubMed: 21892989]

National Statistical Report of Taiwan. Square meter per person per household in the Taiwan Min area. Taipei, Taiwan: Statistical Bureau; 2010a. Retrieved from http://www.stat.gov.tw/public/Attachment/53115143371.pdf

National Statistical Report of Taiwan. Number of households, population, and average occupancy per household in the Taiwan Min area. Taipei, Taiwan: Statistical Bureau; 2010b. Retrieved from http://www.stat.gov.tw/public/Attachment/5311419271.pdf

Nazzi T, Iakimova G, Bertoncini J, Frédonie S, Alcantara C. Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. Journal of Memory and Language. 2006; 54(3):283–299. http://dx.doi.org/10.1016/j.jml.2005.10.004.

Oller, DK. The emergence of the sounds of speech in infancy. In: Yeni-Komshian, GH.Kavanagh, JF., Ferguson, CA., editors. Child phonology. Vol. 1. New York: Academic Press; 1980. p. 93-112.

Oller, DK. The emergence of the speech capacity. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.

Oller DK, Buder EH, Ramsdell HL, Warlaumont AS, Chorna L, Bakeman R. Functional flexibility of infant vocalization and the emergence of language. Proceedings of the National Academy of Sciences. 2013; 110(16):6318–6632. DOI: 10.1073/pnas.1300337110

Oller DK, Eilers RE. Similarity of babbling in Spanish- and English-learning babies. Journal of Child Language. 1982; 9(03):565–577. DOI: 10.1017/S0305000900004918 [PubMed: 7174757]

Oller DK, Eilers RE, Basinger D. Intuitive identification of infant vocal sounds by parents. Developmental Science. 2001; 4(1):49–60. DOI: 10.1111/1467-7687.00148

Oller DK, Eilers RE, Basinger D, Steffens ML, Urbano R. Extreme poverty and the development of precursors to the speech capacity. First Language. 1995; 15(44):167–187. https://doi.org/10.1177/014272379501504403.

Oller DK, Eilers RE, Bull DH, Carney AE. Prespeech vocalizations of a deaf Infant: A comparison with normal metaphonological development. Journal of Speech, Language, and Hearing Research. 1985; 28(1):47–63. DOI: 10.1044/jshr.2801.47

Oller DK, Eilers RE, Neal AR, Cobo-Lewis AB. Late onset canonical babbling: A possible early marker of abnormal development. American Journal on Mental Retardation. 1998; 103(3):249–263. DOI: 10.1352/0895-8017(1998)103<0249 [PubMed: 9833656]

Oller DK, Eilers RE, Steffens ML, Lynch MP, Urbano R. Speech-like vocalizations in infancy: an evaluation of potential risk factors. Journal of Child Language. 1994; 21(01):33–58. https://doi.org/10.1017/S0305000900008667. [PubMed: 8006094]

Oller DK, Eilers RE, Urbano R, Cobo-Lewis AB. Development of precursors to speech in infants exposed to two languages. Journal of Child Language. 1997; 24(02):407–425. [PubMed: 9308425]

Oller, DK., Iyer, SN., Buder, EH., Kwon, K., Chorna, L., Conway, K. Diversity and contrastivity in prosodic and syllabic development. Paper published at the Proceedings of the International Congress of Phonetic Sciences (ICPhS); Saarbrücken, Germany. 2007 Aug.

Oller, DK., Levine, SL., Cobo-Lewis, A., Eilers, RE., Pearson, BZ. Vocal precursors to llinguistic communication: How babbling is connected to meaningful speech. In: Paul, R., editor. Exploring the speech-language connection. Vol. 8. Baltimore, MD: Paul H. Brookes Publishing; 1998. p. 1-25.

Oller DK, Niyogi P, Gray S, Richards JA, Gilkerson J, Xu D, … Warren SF. Automated vocal analysis of naturalistic recordings from children with autism, language delay and typical development. Proceedings of the National Academy of Sciences. 2010; 107(30):13354–13359.

Papoušek, M. Vom ersten Schrei zum ersten Wort: Anfänge der Sprachentwickelung in der vorsprachlichen Kommunikation. Bern: Verlag Hans Huber; 1994.

Patten E, Belardi K, Baranek GT, Watson LR, Labban JD, Oller DK. Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. Journal of Autism and Developmental Disorders. 2014; 44(10):2413–2428. DOI: 10.1007/s10803-014-2047-4 [PubMed: 24482292]

Paul, R., Bruni, P., Balweel, A. Language disorders from infancy through adolescence: Assessment & intervention. 3. St. Louis, MO: Elsevier Health Sciences; 2007.

Price JR, Pollock KE, Oller DK. Speech and language development in six infants adopted from China. Journal of Multilingual Communication Disorders. 2006; 4(2):108–127. http://dx.doi.org/10.1080/14769670601092622. [PubMed: 23204925]

Roe KV, Drivas A. Reciprocity in mother-infant vocal interactions: Relationship to the quantity of mothers' vocal stimulation. American Journal of Orthopsychiatry. 1997; 67(4):645. http://dx.doi.org/10.1037/h0080262. [PubMed: 9361871]

Roug L, Landberg I, Lundberg LJ. Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. Journal of Child Language. 1989; 16(01):19–40. https://doi.org/10.1017/S0305000900013416. [PubMed: 2925811]

Shimada YM. Infant vocalization when alone: Possibility of early sound playing. International Journal of Behavioral Development. 2012; 36(6):407–412. https://doi.org/10.1177/0165025411431408.

Stark, RE. Stages of speech development in the first year of life. In: Yeni-Komshian, GH.Kavanagh, JF., Ferguson, CA., editors. Child phonology. Vol. 1. New York: Academic Press; 1980. p. 73-90.

Tamis-LeMonda CS, Bornstein MH, Baumwell L. Maternal responsiveness and children's achievement of language milestones. Child Development. 2001; 72(3):748–767. DOI: 10.1111/1467-8624.00313 [PubMed: 11405580]

Törölä H, Lehtihalmes M, Heikkinen H, Olsén P, Yliherva A. Early vocalization of preterm infants with extremely low birth weight (ELBW), part II: from canonical babbling up to the appearance of the first word. Clinical Linguistics & Phonetics. 2012; 26(4):345–356. http://dx.doi.org/10.3109/02699206.2011.636500. [PubMed: 22404864]

U.S. Department of Housing and Urban Development and U.S. Census Bureau. American Housing Survey for the Memphis Metropolitan Area: 2004 Current Housing Reports, Series H170/04-8. United States Census Bureau; 2005. Retrieved from https://www.census.gov/prod/2005pubs/h170-04-8.pdf

Vihman MM, de Boysson-Bardies B. The nature and origins of ambient language influence on infant vocal production and early words. Phonetica. 1994; 51(1–3):159–169. DOI: 10.1159/000261967 [PubMed: 8052670]

Warren SF, Gilkerson J, Richards JA, Oller DK. What automated vocal analysis reveals about the language learning environment of young children with autism. Journal of Autism and Developmental Disorders. 2010; 40:555–569. DOI: 10.1007/s10803-009-0902-5 [PubMed: 19936907]

West, BT., Welch, KB., Galecki, AT. Linear mixed models: a practical guide using statistical software. Boca Raton, FL: CRC Press; 2014.

Whalen DH, Levitt AG, Goldstein LM. VOT in the babbling of French- and English-learning infants. Journal of Phonetics. 2007; 35(3):341–352. http://dx.doi.org/10.1016/j.wocn.2006.10.001. [PubMed: 19641636]

Xu D, Richards JA, Gilkerson J. Automated analysis of child phonetic production using naturalistic recordings. Journal of Speech, Language, and Hearing Research. 2014; 57(5):1638–1650. DOI: 10.1044/2014_JSLHR-S-13-0037

Yip, M. Tone. Cambridge, UK: Cambridge University Press; 2002.

Yoder PJ, Oller DK, Richards JR, Gray S, Gilkerson J. Stability and validity of an automated measure of vocal development from day-long samples in children with and without autism spectrum disorder. Autism Research. 2013; 6(2):103–107. DOI: 10.1002/aur.1271 [PubMed: 23436778]

Yoo, H., Buder, EH., Lee, C-C., Oller, DK. Differentiating acoustic properties of cry and speech-like vocalizations in the first year. Paper presented at the American Speech-Language-Hearing Association (ASHA) convention; Denver, CO. 2015.

Zimmerman F, Gilkerson J, Richards J, Christakis D, Xu D, Gray S, Yapanel U. Teaching by listening: The importance of adult-child conversations to language development. Pediatrics. 2009; 124:342–349. DOI: 10.1542/peds.2008-2267 [PubMed: 19564318]

Highlights

- Canonical babbling was observed at 6 and 11 months in American English- and Chinese-learning infants with all-day home recordings, during circumstances where parents used infant-directed or adult-directed speech, and where infants were alone.
- Amount of canonical babbling was influenced by a significant interaction of infant age, language/culture, and social circumstance.
- The research suggests babbling observed in its natural setting develops in a complex way with mutual influences of the social environment on infant behavior and of infant behavior on the social environment.

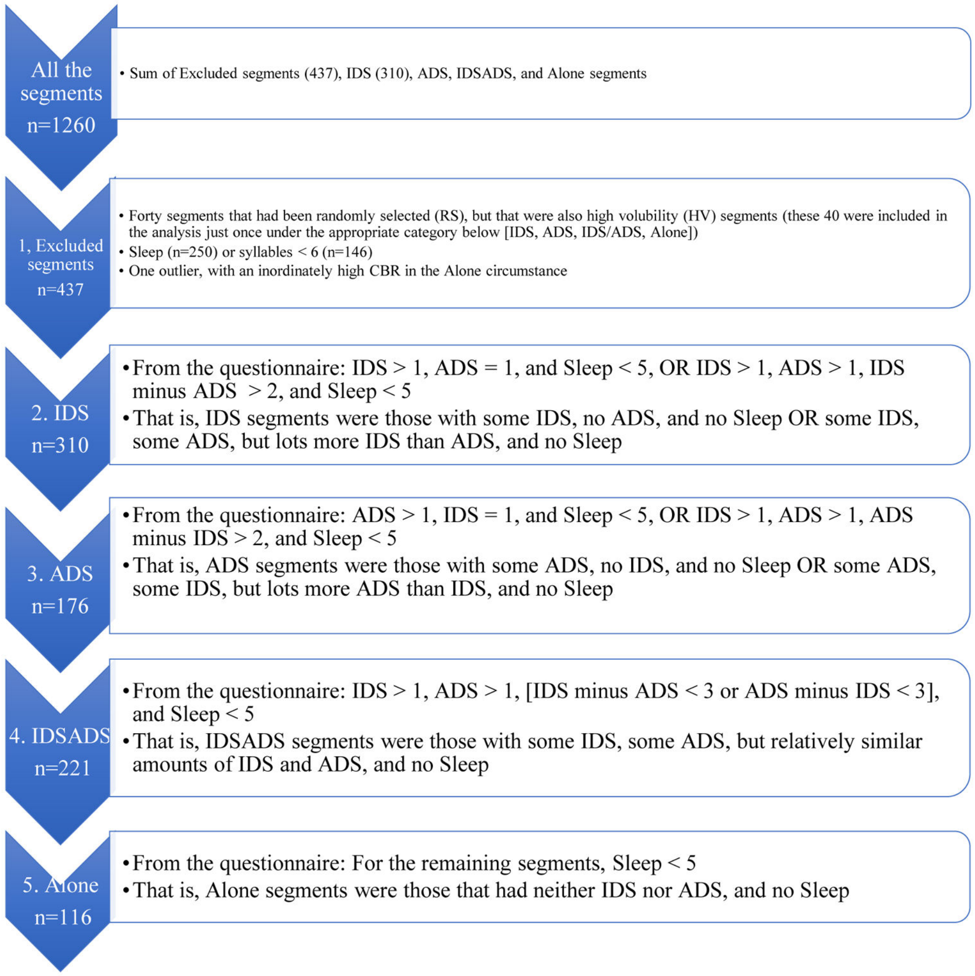

Fig. 1.
Assignment of each segment to a circumstance

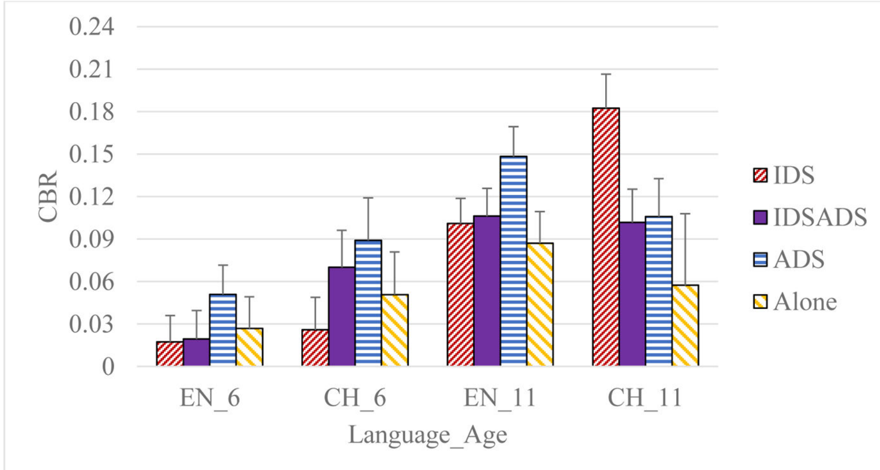

Fig. 2.
Canonical babbling ratio (CBR) and standard error (SE) by infant language background infant age, and social circumstance.

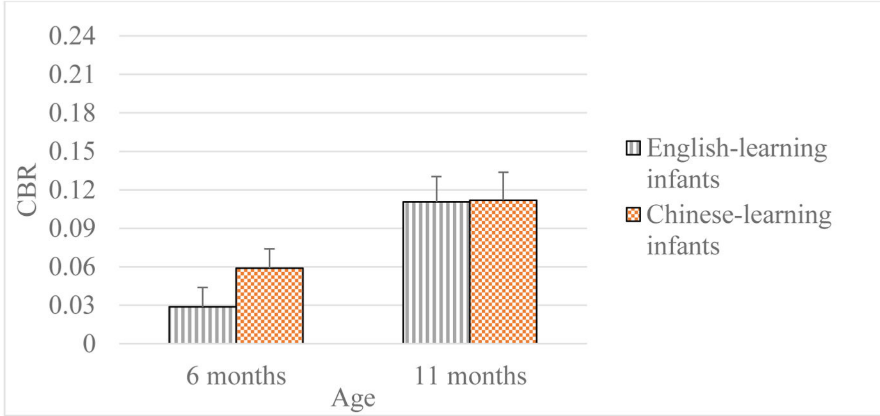

Fig. 3.
Canonical babbling ratio (CBR) and standard error (SE) by infant age and infant language background.

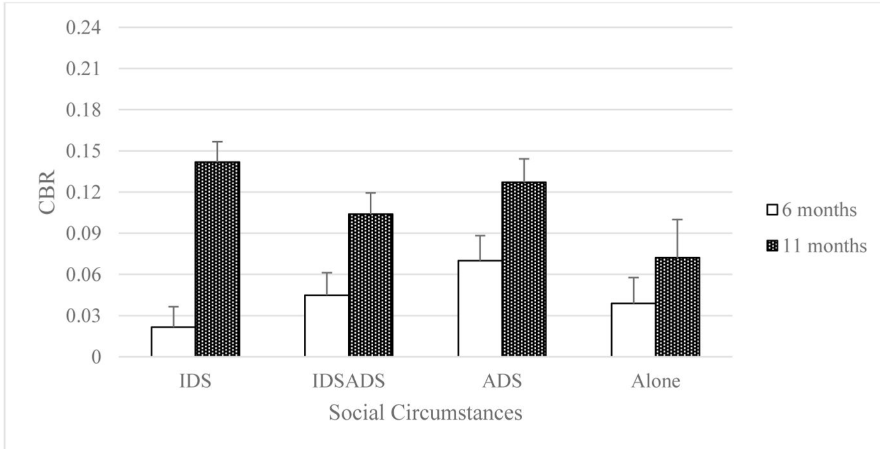

Fig. 4.
Canonical babbling ratio (CBR) and standard error (SE) by infant age and social circumstances.

Table 1

Mean proportion and number of 5-min segments (standard error) occurring in each of the social circumstances. Values computed across 13 English-learning and 8 Chinese-learning infants (30 segments and 150 min per recording)

	EN 6 mo		CH 6 mo		EN 11 mo		CH 11 mo	
	Prop.	***# of Seg(SE)***	***Prop.***	***# of Seg(SE)***	***Prop.***	***# of Seg(SE)***	***Prop.***	***# of Seg(SE)***
IDS	0.22	6.62 (1.12)	0.25	7.63 (1.29)	0.29	8.69 (1.59)	0.21	6.25 (1.16)
IDSADS	0.16	4.69 (0.61)	0.15	4.50 (1.03)	0.18	5.83 (0.75)	0.23	6.75 (0.96)
ADS	0.16	4.85 (1.39)	0.10	3.29 (0.60)	0.14	5.09 (0.92)	0.14	4.86 (1.00)
Alone	0.11	3.38 (0.64)	0.10	3.29 (0.83)	0.11	3.58 (0.53)	0.03*	1.00* (0.29)
Subtotal	0.65	19.54	0.60	18.70	0.72	23.20	0.60	18.86
Sleep or Syllables < 6	0.35	10.46	0.40	11.30	0.28	6.80	0.40	11.14

Note:

* Excluding the outlier with a CBR of .6 (an extremely high value) in a single Alone segment

Table 2

Type III test of fixed effects

Effect	**Num df**	**Den Df**	**F-value**	***p***
Age	1	733	44.25	<.0001
Language	1	733	0.44	0.5064
age*language	1	733	2.06	0.1512
Circumstance	3	74	2.1	0.1081
age*circumstance	3	733	4.72	0.0028
language*circumstance	3	733	1.34	0.2615
age*language*circumstance	3	733	5	0.0019