Social and endogenous motivations in the emergence of canonical babbling: An autism risk study

Helen L. Long[1,2], Gordon Ramsay[3-5], Dale D. Bowman[1,2,6], Megan M. Burkhardt-Reed[1,2], & D. Kimbrough Oller[1,2,7]

Corresponding author: Helen Long, hlong@memphis.edu

Abstract

There is a growing body of research emphasizing the role of intrinsic motivation and endogenous activity to support the development of cognitive systems alongside the well-established role of social interaction. The present study longitudinally evaluated canonical babbling across the second-half year of life, when canonical babbling becomes well-established. We compared segments rated as having high and low levels of turn taking and independent vocal play in 98 children at low and high risk for autism spectrum disorder. Segments were extracted from all-day home audio recordings to observe infants in naturalistic settings. Canonical babbling ratios (CBR) were determined based on human coding along with Likert-scale ratings on the level of turn taking and vocal play in each segment. We observed highly significant differences in CBRs between risk groups during high and low vocal play, but high and low levels of turn taking yielded a weaker effect. There were also interactions of CBR with age, risk, and vocal function variables. We conclude that social and endogenous/exploratory motivations may drive both high- and low-risk infant tendencies to produce their most speech-like vocalizations.

[1] School of Communication Sciences and Disorders, University of Memphis, Memphis, Tennessee, USA

[2] Institute for Intelligent Systems, University of Memphis, Memphis, Tennessee, USA

[3] Marcus Autism Center, Children's Healthcare of Atlanta, Atlanta, Georgia, USA

[4] Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia, USA

[5] Center for Translational Social Neuroscience, Emory University, Atlanta, Georgia, USA

[6] Department of Mathematics, University of Memphis, Memphis, Tennessee, USA

[7] Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria

Lay abstract

Infants at low and high risk for autism produced similar babbling rates across various levels of social interaction. Infants at high risk for autism produced higher rates of mature, well-formed (canonical) syllables during bouts of independent vocal play. In this study, we compared rates of babbling in infants at low and high risk for autism across different levels of infant vocal turn taking and vocal play. Eight 5-minute segments were selected from each all-day home recording (98 infants, 483 recordings in the second half-year of life). Listeners coded infant syllables as canonical, i.e., "adult-like," or noncanonical, i.e., "immature," forms. After coding each segment, listeners rated each segment for its frequency of both infant turn taking and vocal play using a Likert scale. We predicted that low-risk infants would do more canonical babbling during social interaction and high-risk infants would babble more during independent vocal play. But infants at high risk for autism did not differ significantly in babbling rates across levels of social interaction from low-risk infants. Also contrary to our prediction both groups produced higher rates of canonical babbling during periods of high vocal play. Both groups produced similar canonical babbling rates over time. These findings may inform our understanding of internal motivations for canonical babbling and potentially early indicators of differences in the vocal production of infants with autism before they produce their first words.

Keywords

vocal development, social motivation, vocal play, endogenous activity, infant behavior, autism spectrum disorder

Introduction

Overview

Canonical babbling has been long established as a robust stage of prelinguistic vocal development occurring prior to the emergence of the first word, having been argued to constitute a necessary foundation for vocabulary development (Koopmans-van Beinum & van der Stelt, 1986; Oller, 2000; Stark, 1980). To our knowledge, there is no published research evaluating the role of exploratory motivation in infants' production of canonical babbling and no direct evaluation of the extent to which social engagement in vocal turn taking affects it. In the present research, we observed babbling in infants at low- and high-risk for autism in naturalistic settings. Segments extracted from all-day home audio recordings were rated for levels of infant turn taking and independent vocal play to measure the degree of social and non-social vocal activity (and thus, social and exploratory motivations, respectively). We examined these findings within an evolutionary developmental biology (evo-devo) framework (Bertossa, 2011; Carroll, 2005; Newman, 2000, 2012), in part to inform our understanding of how babbling may be used to signal developmental progress to caregivers (Locke, 2017; Oller & Griebel, 2005, 2008). Comparing differences between autism risk groups may help to elucidate exploratory tendencies and potential breakdowns in social motivation in autism, as well as providing clinically useful perspectives on the development of language foundations.

Canonical Babbling Development in Typical Development and Autism

Throughout the first half year of life, infants evidence an emerging capacity to control and coordinate the respiratory, phonatory, and articulatory mechanisms. Within the second half year, and rarely later than 10 months, infants begin canonical babbling (Oller, 1980; Stark, 1980), defined as the production of mature consonant-vowel syllables with well-formed transitions between the consonant- and vowel-like elements (e.g., [baba], [dada]). These syllables provide a basis for interaction and play with repeated and varied syllables, foundational for the production of first words (Oller, 2000). The onset of canonical babbling is known to be a robust predictor of typical speech development (Oller et al., 1998; Nathani et al., 2006), with delays observed in several disorders including deafness (Eilers & Oller, 1994; Oller & Eilers, 1988), Down syndrome (Lohmander et al., 2017; Lynch et al., 1995), Fragile X syndrome (Belardi et al., 2017), cerebral palsy (Levin, 1999; Nyman & Lohmander, 2018), and William syndrome (Masataka, 2001). Lang et al. (2019) reviewed the mixed evidence on canonical babbling onset in autism, summarized below.

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by deficits in social communication and restricted interests and repetitive behaviors (American Psychiatric Association, 2013). Diagnosis is common nowadays by 18-24 months of age

(Zwaigenbaum et al., 2015). Symptoms in early infancy include reduced or absent dyadic interaction, social responsiveness, and joint attention (Kellerman et al., 2019; Mundy, 2017; Ozonoff et al., 2010), and there is some evidence suggesting prelinguistic vocal developmental anomalies (e.g., Sheinkopf et al. (2012)). Two studies previously analyzed *canonical babbling ratios* (CBRs) in infants with ASD. Patten et al. (2014) showed significantly lower ratios in children with autism at 9-12 months and 15-18 months compared to controls, and Paul et al. (2011) found lower ratios at 9 months in infants at high risk for autism compared to low-risk infants, but not in a 12-month group. Two retrospective video analysis studies also found mixed results when analyzing *canonical syllables per minute*. Werner et al. (2000) showed no differences between infants later diagnosed with autism relative to typically developing controls between 8-10 months but significant differences at 12 months in complex babbling rates, and Chericoni et al. (2016) found no differences between the two groups at ages 6-12 months. Two other studies observed *ages of onset* for the canonical babbling milestone in infants at low and high risk for autism. Iverson & Wozniak (2007) reported that high-risk infants had a wider range for age of onset for canonical babbling (5-18 months) compared to the low-risk group (5-9 months), but LeBarton & Iverson (2016) found 33/37 infants at high risk for autism reached the canonical babbling stage by 14 months, with a typical average mean age of onset (7.67 months). In a feasibility study analyzing *syllable complexity*, Pokorny et al. (2017) found that an equal number of neurotypical and autistic infants in each group (4/10) produced more complex types of utterances than single canonical syllables by 10 months.

Overall, there is a lack of conclusive evidence on canonical babbling developmental differences in children at risk for autism or later diagnosed with ASD. Inconsistent findings across studies may be attributed to the well-established variability in autism characteristics, the varying methodologies used to analyze babbling development, or the differing group types included (as pertaining to retrospective analysis of children diagnosed versus prospective risk studies). Additional research including larger sample sizes is also necessary to provide a smaller margin of error when comparing typically developing groups and groups with autism. In this study, we compare the emergence of canonical babbling for infants at low and high risk for autism using the largest sample size to date (98 infants) and with evaluation based on sampling from all-day recordings across the second half-year of life (483 total recordings).

The Social and Endogenous Nature of Infant Vocalizations

When evaluating the emergence of canonical babbling, there is reason to consider potential differences in social and endogenous motivations behind the production of these advanced vocal forms. Considerable research has evaluated the role of social interaction in infant vocal development and the emergence of language (Franklin et al., 2013; Gros-Louis et al., 2014; Hsu & Fogel, 2001; Iyer et al., 2016; Lee et al., 2018). Caregivers are known to elicit and maintain "protoconversations" (Bateson, 1975), supporting the emergence of mature vocal stages such as canonical syllables and words (Bråten, 1988; Golinkoff et al., 1992; Rochat et al., 1999). Experimental studies using the still-face paradigm have also shown effects of social interaction on infant volubility and vocalization types (Delgado et al., 2002; Franklin et al., 2013; Goldstein et al., 2009). It is important to note that this body of research has primarily examined the effects of parental interaction on infant behavior. To our knowledge, there is no published evidence directly examining the relative roles of interaction and endogenous vocalization on infant vocal development, including canonical babbling.

Contradicting the perhaps implicit assumption that infant vocalizations are simply interactive, several researchers have recently emphasized the role of intrinsic motivation in the development of emotional and cognitive systems, including those related to vocal development (Davis & Panksepp, 2018; Moulin-Frier et al., 2014; Moulin-Frier & Oudeyer, 2013). Infants produce more speech-like vocalizations, or "protophones," (including both canonical and precanonical babbling) without person-directed gaze (both when alone and in the presence of caregivers) than they produce socially-directed sounds (Harold & Barlow, 2013; Oller et al., 2013). More recently, several authors of the present study found that approximately 75% of all infant protophones in laboratory recordings were endogenously produced (Long et al., 2020). We know that social interaction influences infant babbling, phonological learning, and complex language skills (Albert et al., 2018; Elmlinger et al., 2019; Goldstein et al., 2003; Goldstein & Schwade, 2008; Kuhl, 2007), but social and endogenous motivations for infant vocal activity require additional research to elucidate their relative roles. Thus, our research evaluates canonical babbling across segments with high and low levels of both infant turn taking and exploratory vocal play.

An Evolutionary-Developmental Perspective on the Role of Social Motivation in Canonical Babbling

We and others have hypothesized selection pressures on the production of endogenously produced protophones. Baby sounds can be seen as fitness signals selected to elicit long-term investment from caregivers, required across the lengthy period of relative helplessness, or altriciality, of infant humans (Locke, 2017; Long et al., 2020; Oller et al., 2016, 2019). In accord with the fitness signaling hypothesis, the *quality* of infant vocalizations can be considered a salient and reliable signal of fitness. Following this reasoning, it might be seen as advantageous for infants to produce their most advanced vocal forms during periods of caregiver attention. Empirical evidence has been presented to show that caregivers are keenly aware of their infants' developmental capabilities, including in the vocal domain (Bodnarchuk & Eaton, 2004; Lyytinen et al., 1996; Oller et al., 2001). Higher rates of canonical syllables (as opposed to less advanced protophones) during social interaction than during periods of aloneness could suggest a social motivation for producing the more advanced protophones. If the idea is on target, we might conclude that canonical babbling was selected as a salient signal of developmental progress, especially during social interaction. Furthermore, a breakdown in the social motivation of infants as a result of a neurodevelopmental condition such as that seen in autism could potentially result in lower rates of canonical babbling during social interaction than in those of typically developing infants.

The *social motivation theory* (Chevallier et al., 2012) posits that reduced social attention in infancy leads to the social-cognition developmental differences observed in autism spectrum disorder. Additional research supports this notion, showing social information is less salient in individuals with autism (Chevallier et al., 2013; Schultz et al., 2000; Weeks & Hobson, 1987) and less intrinsically rewarding in individuals with autism compared to typical controls (Bottini, 2018; Gray et al., 2018; Scott-Van Zeeland et al., 2010; Sepeta et al., 2012). Reductions in social orienting can also affect language development (Baranek et al., 2013; Dawson et al., 2004; Su et al., 2020), a supposition supported by speculations predicting positive associations between social motivation and language emergence; these speculations have yielded, for example, the continuity hypothesis (Bruner, 1974), the speech attunement framework (Shriberg et al., 2011), and the elicited bootstrapping hypothesis (Camarata & Yoder, 2002), which has been recently elaborated by Su et al. (2020). This body of research and theory highlights the importance of identifying early differences in social interaction in infants at risk for autism in order to provide support and intervention as early as possible.

Interestingly, there is limited research examining endogenously motivated vocal learning in infancy (Syal, 2011). Instead, the great majority of research has focused on parental activity rather than internal motivations of the infant as influencing vocal development. In a salient recent example of such research, Su and colleagues found early social motivations around 23 months predicted language skills 2 years later—specifically, higher performance on social motivation tasks was significantly correlated with functional language abilities (Su et al., 2020). Such literature is consistent with the expectation that reduced social attention and inclinations in early infancy may affect the infant's motivation to produce advanced vocal forms during interaction, and thus may yield reductions in vocal fitness signaling in infants with low social motivation. It is thus consistent with the social motivation theory and also with our evo-devo approach, to predict that infants with typically developing levels of social motivation will produce higher rates of canonical syllables during periods of high vocal interaction than infants with low social motivation. The present body of data offers the opportunity to evaluate this possibility during periods where caregivers and infants engage in high or low amounts of vocal turn taking, and while comparing canonical babbling rates of infants who are at low risk for autism (presumably with typical levels of social motivation) and infants who are at high risk for autism (presumably with lower levels of social motivation). In accord with the social motivation theory, we anticipate that low-risk infants will show higher rates of canonical babbling (with respect to their own baselines) during periods of high turn taking, but that high risk infants will not show higher rates during high turn taking.

On the Role of Exploratory Vocal Play in Typical Development and Autism

We are influenced by the literature-based hypothesis that infants at low autism risk should be expected to produce more canonical syllables during social interaction, while high-risk infants should not be expected to do so, but recent evidence suggests a contrasting possibility. Research by Long et al. (2020) has shown that typically developing infants produce protophones (both canonical and precanonical) predominantly endogenously. Even in laboratory recordings, during periods when parents seek social interaction with infants, most protophones (~60%) appear not to be directed to parents, and this predominance of endogenous vocalization is even stronger (~80%) when parents are present with infants but not attempting to engage them. The results suggest that research on infant tendencies to vocalize at varying levels of advancement should compare circumstances showing high vocal turn taking with circumstances showing high endogenous vocal activity, which we shall refer to here as *vocal play* (Stark, 1980, 1981). Thus, we deem it important to examine not only social motivations for the production of canonical syllables but also intrinsic, exploratory motivations.

During vocal play infants explore sensorimotor aspects of the vocal apparatus and practice with various properties of sounds such as syllabic structure, amplitude, and pitch control. Play has been well established to be important throughout development. Piaget treated play as necessary for children to understand and learn about the world (Piaget, 1952). Vygotsky also viewed play as necessary for the development of cognitive systems and interpersonal relationships (Berk, 1994; Vygotsky, 1978). Panksepp and colleagues proposed play as a fundamental neurobehavioral process, motivated by a play “emotion”, distributed widely among social animals (Davis & Panksepp, 2018; Panksepp, 2005; Panksepp et al., 1984; Panksepp & Biven, 2012). Stark described vocal play as highly variable, with infants producing sounds in new and repeated combinations, modifying patterns and features during bouts of independent infant vocal activity (Stark, 1980). Stark’s description of vocal play evokes the notion that its occurrence can also be considered a sensorimotor exploration of the vocal mechanism necessary to learn and master speech production.

Although it appears that infants in general are endogenously motivated to produce protophones, the social motivation theory of autism hints at the intriguing possibility that infants with autism may be relatively more inclined to vocalize independently/endogenously than neurotypical infants. Further, the reasoning might be extended to suggest that the rate of canonical babbling would be relatively higher (with regard to their own baselines) for infants with autism than for typically developing infants. In the context of the present dataset, it might be predicted that infants at high risk for autism will produce relatively higher rates of canonical babbling during independent vocal play than infants at low risk for autism. In contrast, infants at low risk for autism would not be expected to show higher rates of canonical syllables during high vocal play.

These speculations are perhaps supported by the fact that children with autism have been shown to spend more time participating in isolated play with objects and to produce more repetition of physical actions in play compared to typically developing peers (Atlas, 1990; Naber et al., 2008; Sigman & Ungerer, 1984; Williams et al., 2001). These patterns suggest that as infants with autism begin to produce canonical syllables, they may be particularly interested in the physical, articulatory properties of these sounds—not unlike their often intense interest in the physical characteristics of objects—and they may produce these sounds with greater repetition, perhaps in enjoyment of the self-stimulatory nature of the repetition.

Specific Aims and Hypotheses

The present research compares canonical babbling ratios (CBRs) of infants at low and high risk for autism during recorded segments with high and low levels of both turn taking and vocal play across three age ranges during the second half-year of life. We preliminarily analyzed for possible differences in CBR between infants of high and low socioeconomic status (SES) and found no significant differences; therefore, we do not report SES effects in the data below. Sex differences were evaluated in a recent study from our laboratory using the present dataset and no significant sex differences for CBR were found (Oller et al., 2020); therefore, we do not include sex as a variable in the present work. Findings from this study may inform our understanding of social and exploratory motivations in the emergence of advanced prelinguistic vocalizations in typical and atypical development. Furthermore, risk group differences could suggest early signs of social language impairments in the first year of life. The following are hypotheses to be evaluated:

Predicted Interactions

Our initial analyses conducted using Generalized Estimating Equations (GEE) addressed Turn Taking (TT) and Vocal Play (VP) separately. Consequently, one analysis included three variables: Age, Risk, and TT, and another: Age, Risk, and VP. Based on the social motivation theory of autism, we predicted interactions of:

1. Risk and TT: CBRs in low-risk (LR) infants will be higher during segments with high TT than low TT while high-risk (HR) infants will not show higher CBRs during high TT.
2. Risk and VP: CBRs in HR infants will be higher during segments with high than low VP while LR infants will not show higher CBRs during high than low VP.

We also examined the possible interaction between Risk and Age in a GEE analysis including only those two independent variables. We predicted:

3. Risk and Age: CBRs will increase to a greater extent in LR infants across the three ages than in HR infants.

Predicted Main Effects

For interpretive perspective, we also analyzed main effects for CBR in a final GEE including Age, Risk, TT, and VP. We predicted:

1. Age: Higher CBRs will occur at the later ages than earlier ages, highlighting infants' increasing ability to control the speech mechanism (Lee et al., 2018; Nathani et al., 2006; Oller, 2000).

2. Risk: Higher CBRs will occur in the LR group compared to the HR group, a prediction based on the predominant, albeit inconsistent findings of the existing literature (Lang et al., 2019).

3. TT: Higher CBRs will occur during segments with high TT compared to low TT.

4. VP: Higher CBRs will occur during segments with low VP compared to high VP.

Methods

The institutional review boards of the University of Memphis and Emory University approved the procedures used in this study. Families provided written consent prior to participation in this study.

Participants

As part of an NIH-funded Autism Center of Excellence conducted at the Marcus Autism Center in Atlanta, GA, 100 families of newborn infants were recruited via flyers, advertisements, social media and community referrals to participate in a longitudinal sibling study of development across the first three years of life. We analyzed data from 98 infants (two infants did not complete recordings at the ages studied in this paper). Infants were recruited as being either at high risk (HR, n=49) or low risk (LR, n=49) for autism. Infants were deemed HR if they had at least one older biological sibling with a confirmed autism diagnosis, and LR if they had no familial history of autism in 1^{st}, 2^{nd}, or 3^{rd} degree relatives. Sex and socio-economic status (SES) measures[1] were balanced to the greatest extent possible in accord with known autism male-to-female ratios (Loomes et al., 2017) and SES make-up of participants living in the greater Atlanta, GA area who were willing and able to participate in a 3-year longitudinal study. Table 1 presents demographic information for the infants included in this study.

[1] SES was measured using maternal education. Low/High-SES groups were based on a median split of maternal education in the entire cohort.

Table 1. Numbers of infants by Risk, Sex, and SES

*Number of participating infants by risk status, sex, and socio-economic status (SES). *One infant's family did not report an SES level.*

		High-Risk	Low-Risk	Total
Total		**49**	**49**	**98**
Sex	**Male**	34	30	64
	Female	15	19	34
SES*	**Low SES**	26	18	44
	High SES	22	31	53

Families were asked to complete audio recordings once a month between 1-36 months of age. This study used data collected between 6.5 and 13 months of age to represent the typical range of expected onset for and infant activity in canonical babbling. These data were grouped into three age ranges for analysis and labeled with reference to the approximate mean age within each group: 6.5-8.49 months (7.5 months), 8.5-10.49 months (9.5 months), and 10.5-13 months (12 months). It should be noted that the 12-month age group included a slightly smaller age range (1.5 months) compared to the 7.5- and 9.5-months age groups (2 months).

Audio Recordings

Audio recordings were completed using LENA recording devices (Gilkerson et al., 2017; Zimmerman et al., 2009). These devices are battery powered and secured inside the pocket of a special vest or clothing item with button clasps and can record up to 16 hours of audio per charge. LENA devices have a 16 kHz sampling rate for adequate play-back of audio for human coding judgements of recorded material.

Recording Procedures

Families completed all-day recordings once a month starting from the first month of life through the third year of life. Once a month, parents were provided with a LENA recording device and were supplied regularly with appropriately sized clothing for their child to wear throughout the day, as well as full instructions on how to carry out recordings. They were asked to turn on the recorder when their child woke up in the morning on the day of the recording and leave it running until the child went to sleep at night, in order to obtain a representative naturalistic recording of the child's whole day. They were asked to remove the recorder and leave it running nearby during bath times, sleep, and any situation where the recorder would press on the child's chest or cause discomfort. They were also allowed to pause the recording in any situation that they felt would violate their right to privacy or confidentiality. Recordings were scheduled for the same calendar day each month, as far as possible, to rotate weekdays. The device was returned to the research project staff at the

Marcus Autism Center each month following recording days for data processing. Each family completed ~5 total recordings (range: 1-7) across the ages studied, with an average recording time of approximately 11 hours per day.

Coding Procedures

Twenty-one 5-minute segments were randomly extracted from each recording and coded in real-time for infant utterance counts by 16 trained graduate student coders[2] at the Origin of Language Laboratory (OLL). OLL staff were blinded to all diagnostic and demographic information associated with each infant recording throughout the coding process. From these 21, eight segments with the highest infant vocalization volubility and a range of infant-directed speech[3] were selected for further analysis from each recording, totaling 3799 segments. Fifteen of these segments were later excluded on the basis of having no infant vocalizations; therefore, final analyses were completed on a total of 3784 segments.

Canonical Babbling Ratios as a Measure of Advanced Prelinguistic Vocal Forms

In a second pass of coding, the 8 selected segments were coded in real-time for infant canonical and non-canonical syllable counts. Listeners identified a total of 30,263 canonical syllables, and 233,877 noncanonical syllables across all segments. To measure the emergence of advanced vocal forms, a canonical babbling ratio (CBR) was calculated as the total number of canonical syllables divided by the total number of syllables in each segment. Means and standard deviations of CBRs were calculated for each infant at each age. These data were then averaged within each age (7.5, 9.5, and 12 months) and risk group (HR and LR). Occasionally, families did not complete a monthly recording, and for those cases there was no data at the infant's age to include in the analysis. If there were multiple recordings per age and infant (occasionally two recordings were completed at a single age), the means and SDs of these recordings were averaged for analysis.

Turn Taking and Vocal Play as Measures of Infant Vocal Function

Following syllable coding of each 5-minute segment, coders answered a 17-item questionnaire regarding how often infants used vocalizations for various functions based on

[2] Graduate student coders were trained to differentiate canonical and non-canonical syllables during real-time coding and to rate the extent to which infants produced socially interactive (TT) and endogenous (VP) vocalizations during completion of the questionnaire that was filled out at the end of coding of each 5-minute segment.

[3] The amount of infant-directed speech (IDS) was rated using the questionnaire that followed each of the 21 segments coded in the first coding pass. The questionnaire was also used to indicate environmental contextual factors for each segment, including audibility, other-person activity level, and aloneness of the infant. As with the second coding pass, each questionnaire item required a 5-point Likert-scale response to the relevant question, e.g., for IDS, "*How often did someone talk to the infant?*"

the audible context of the infant's environment in each segment. See Long et al. (2020) for theoretical perspectives on making intuitive judgments of infant vocal functions. We used two items from the questionnaire to measure frequencies of naturalistic infant vocalizations that were judged to be inherently social and exploratory within each segment. Specifically:

1. *Turn Taking (TT)*: Were any of the infant's protophones used in vocal turn taking with another speaker?

2. *Vocal Play (VP)*: Were any of the infant's protophones purely vocal play or vocal exploration?

Coders were instructed to respond to each question using a Likert Scale which aligned to the following rating designations: *1 = Never, 2 = Less than half the time, 3 = About half the time, 4 = More than half the time, 5 = Close to the whole time*. For example, a TT rating of 5 was applied to segments where a caregiver was clearly speaking to the infant, and the infant was vocalizing in an apparent back and forth vocal interaction for essentially the whole segment. Segments with a VP rating of 5 would indicate the listener perceived the vast majority of infant vocalizations as playful and exploratory and not directed to another person in any way. TT and VP were not considered opposing vocal functions; in other words, a TT rating of 5 would not necessitate a VP rating of 1. In segments with very high infant vocal activity containing both interactive and non-interactive information, it is conceivable that a segment could be rated as having high TT (5) and high VP (5). Conversely, a segment with low infant vocal activity and limited interaction with the parent would have low TT and VP ratingsRatings for TT and VP were dissimilarly distributed across the Likert scale range, as shown in Table 2. In order to compare levels of TT and VP with maximally similar numbers of segments at two levels in both cases, we split TT ratings into "No Turn Taking" (Rating of 1) vs. "Any Turn Taking" (Ratings 2-5), and VP ratings into "Low Vocal Play" (Ratings 1-3) vs "High Vocal Play" (Ratings 4-5) levels. Even with this procedure, the TT split yielded a dramatic imbalance, with more than 80% of all segments pertaining to the No TT grouping. On the other hand, VP was very common, with only 8% rated as having no VP, and 55% rated as having VP occurring either "more than half the time" or "close to the whole time."

Table 2. Frequency distribution of segments for TT and VP

Following the coding of infant syllables in 3784 segments, coders rated each 5-minute segment on the frequency of vocal turn taking (TT) and Vocal Play (VP) for infants at low risk (LR) and high risk (HR) for autism. The distribution of segments along the rating scale for TT and VP was similar for both risk groups. Ratings for each variable were combined into two levels for maximally similar numbers within each category: "No TT" (TT rating: 1) vs "Any TT," (TT: 2-5) and "Low VP" (VP: 1-3) vs "High VP" (VP: 4-5).

Likert scale rating	**Interpretation (Level of occurrence)**	**TT Level**	**TT count**		**VP Level**	**VP count**	
			HR	LR		HR	LR
1	Never	No TT	1564	1482	Low VP	164	147
2	Less than half the time	Any TT	295	312		307	254
3	About half the time		42	55		431	383
4	More than half the time		18	14	High VP	506	526
5	Close to the whole time		2	0		513	553

Coder Agreement

Inter-rater agreement was examined for CBRs, TT level, and VP level using a secondary LENA recording dataset coded by 7 of the same graduate student coders following the coding protocol used in this study. The 5-minute segments that had already been coded—each by one of the 7 individuals—came from a set of over 1000 such segments randomly selected from the all-day recordings of eight infants at each of six ages across the first year of life. >380 of these segments had been coded in the very same way as in the present study, with determination of CBR, TT, and VP. A subset of 212 of these segments was semi-randomly selected to be assigned for a second pass of agreement coding, where the agreement coder would always be a different individual from the one who had provided the original coding. The number 212 was based on available coder time and the desire for a large enough sample to yield trustworthy agreement data.

Every one of the 7 agreement coders was assigned to at least 5 segments that had originally been coded by each of the other 6 coders. In addition, all agreement coders were assigned to at least 5 segments from each of the 8 infants. Finally, all the ages of infants were included for assignments to each of the agreement coders for at least 5 segments. The agreement coding was conducted blind, in the sense that no coder knew who had originally coded the segments assigned to them in the agreement phase, nor were they supplied with information about age or identity of the infants.

The agreement coding for canonical babbling ratios revealed high agreement for both the entire set, with ages ranging across the entire first year (r = .89), and for the subset that pertained only to the second half year (r = .87), a time period during which CBR varies substantially above 0 across the entire range of ages. Both the questionnaire items yielded far better than chance levels of agreement on the Likert-scale judgments categorized binarily as in the present work (No TT = 1, Any TT = 2-5; Low VP = 1-3, High VP = 4-5) based on Chi square analysis ($p < .001$). For VP there was agreement on 66% of pairings, while for TT there was agreement on 87%, with only fair agreement on kappa (TT = .40, VP = .33). This level of agreement should offer little surprise, given the subjective nature of the judgments. We have been surprised, however, by the power to significantly predict CBR that these blunt measures offer, as will be seen below.

Statistical Approach

We used Generalized Estimating Equations (GEE) implemented in R to analyze main effects and interactions of Risk, Age, and TT and VP on infant CBRs and also tested independently for main effects of all four independent variables. GEE analyses are an advanced form of modeling providing a non-parametric alternative to generalized linear mixed models for estimating within-subject covariance and population-averaged model parameters (Liang & Zeger, 1986). GEE has advantages over other mixed models frameworks especially in cases where data across conditions and from participants are intercorrelated and where numbers of observations per participant or condition varies. Another advantage is that the GEE approach requires no normality assumption. A GEE analysis is appropriate here because this is a longitudinal dataset with an unequal number of observations on infants, number of recordings per Age and Risk group, and number of observations of TT and VP ratings within each level.

Results

We ran three GEE models evaluating interactions and main effects for 1) TT, Age, and Risk, 2) VP, Age, and Risk, 3) Age and Risk, and we ran a fourth GEE model on main effects only for 4) Age, Risk, TT, and VP.

Turn Taking, Age, and Risk

1. *Predicted interaction of Risk and TT*: Based on predictions derived from the social motivation theory, we predicted higher *CBRs in low-risk (LR) infants during segments with high turn taking but no such pattern in high-risk (HR) infants*. However, the results (Table 3) did not confirm the hypothesis (p = .144). The mean CBR for HR and LR infants was quite similar for segments with no TT, but somewhat (though not significantly) higher in LR infants for segments with any amount of TT.

In the full GEE model, we found no main effect of TT, that is, no significant difference in CBRs between segments rated as having No vs Any TT (p = .347). Differences in CBRs between Risk groups were also non-significant (p = .111), with somewhat higher CBRs in the LR group. In the same model, the main effect for CBR from 7.5 to 9.5 months of age was highly significant (p < .001, b = .06), but differences from 9.5 to 12 months were not (p = .121), reflecting the fact that CBRs went up more from 7.5 to 9.5 months than they did from 9.5 to 12 months.

Table 3. Turn Taking, Age, and Risk interaction model
Full interaction GEE model for CBR with Age group (7.5 to 9.5, and 9.5 to 12 mo.), Risk Group (HR vs LR), and Turn Taking as a factor (No TT vs Any TT).

Variable	Effect size (*b*)	SE	p
TT (No vs Any)	0.02	0.02	0.347
Risk (LR vs HR)	-0.02	0.01	0.111
Age (7.5 to 9.5 mo.)	0.06	0.01	< .001
Age (9.5 to 12 mo.)	0.02	0.02	0.121
TT * Risk	0.04	0.03	0.144
TT * Age (7.5 to 9.5 mo.)	0.01	0.02	0.812
TT * Age (9.5 to 12 mo.)	0.02	0.02	0.426
Risk * Age (7.5 to 9.5 mo.)	0.05	0.02	0.004
Risk * Age (9.5 to 12 mo.)	0.03	0.02	0.075
TT * Risk * Age (7.5 to 9.5 mo.)	-0.04	0.04	0.368
TT * Risk * Age (9.5 to 12 mo.)	-0.05	0.03	0.136

The results did not show significant two-way interactions between TT and either of the Age group comparisons: 7.5 to 9.5 months (p = .812) or 9.5 to 12 months (p = .426). There was, however, a significant interaction between Risk and Age for 7.5 to 9.5 months (p = .004, b = .05); CBRs increased in HR infants to a greater extent between 7.5 and 9.5 months compared to LR infants across these two ages. This difference was reversed from 9.5 to 12 months such that LR infants (p = .075) showed a greater increase than HR infants in that age interval, an interaction that approached statistical significance. No significant three-way interactions were observed in this model. Figure 1 provides graphic illustration of the results presented in the full model for Age, Risk, and TT level.

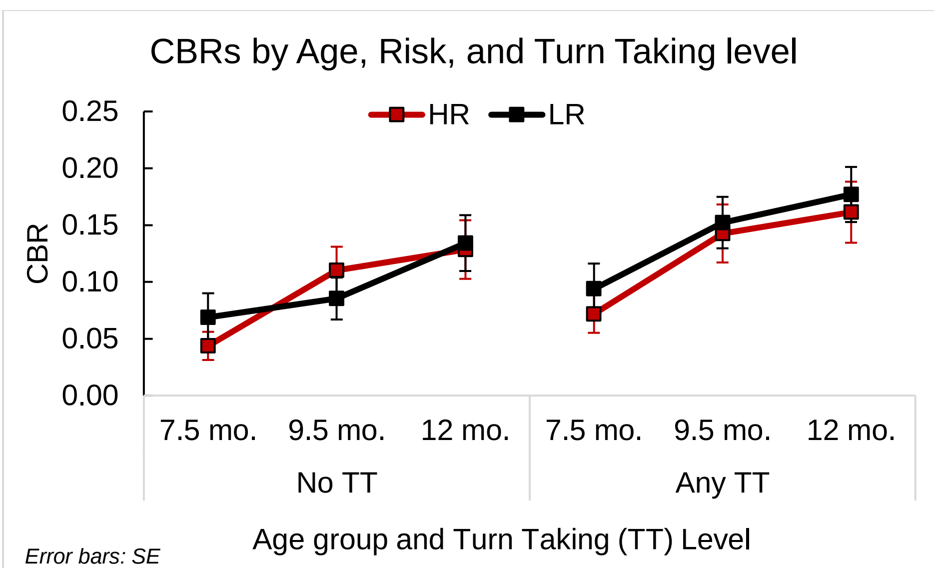

Figure 1. Canonical babbling by Age, Risk, and Turn Taking level

Canonical babbling ratios of infants at high-risk (HR) and low-risk (LR) for autism during segments with no vs any turn taking (TT) across three age ranges, 6.5-8.49 (7.5 months), 8.5-10.49 (9.5 months), and 10.5-13 (12 months) months. CBR was significantly higher from 7.5 to 9.5 months (p < .001, b = .06), and there was a significant two-way interaction of Risk and Age again between 7.5 and 9.5 months (p = .004, b = .05). There were no significant interactions including TT as a variable, including the three-way interactions of Age, Risk, and

TT level. The values presented in the figure were computed from the raw data with means and SEs weighted for the number of infants who contributed data in each Risk group at each Age.

Vocal Play, Age, and Risk

In the full GEE model for Age, Risk, and VP (Table 4) we found several significant effects not observed in the model for Age, Risk, and TT.

2. *Predicted interaction of Risk and VP*: Based on predictions derived from the social motivation theory, we predicted *an increase in CBRs in HR infants from segments with low to high VP, and a lesser increase or no increase from low to high VP for LR infants*. There was indeed a significant interaction between VP level and Risk group (p = .021, b = -.03), but the direction of the effect was the opposite of that predicted. CBRs of LR infants increased to a greater extent from low to high VP than CBRs of HR infants. Based on calculations for Figure 2, CBRs at low VP were comparable (HR = .079, LR = .080), while those at high VP differed more, favoring the LR group (HR = .119, LR = .124).

There was a highly significant main effect of VP, corresponding to a higher overall mean CBR produced by all infants during high VP compared to low VP (p < .001, b = .09). As with the full TT model, we observed no significant difference between Risk groups in the full VP model. There was, however, a significant effect of Age at both levels in the full VP model, with CBRs significantly increasing between ages 7.5 and 9.5 months, (p = .023, b = .03) and between ages 9.5 and 12 months, (p = .001, b = .05).

Table 4. Vocal Play, Age, and Risk interaction model

Full interaction GEE model for CBR with Age (7.5, 9.5, and 12 mo.), Risk (HR vs LR), and Vocal Play (Low VP vs High VP).

Variable	Effect size (*b*)	SE	p
VP (Low vs High)	0.09	0.01	< .001
Risk (LR vs HR)	0.00	0.02	0.800
Age (7.5 to 9.5 mo.)	0.03	0.01	0.023
Age (9.5 to 12 mo.)	0.05	0.02	0.001
VP * Risk	-0.03	0.02	0.021
VP * Age (7.5 to 9.5 mo.)	-0.06	0.02	0.001
VP * Age (9.5 to 12 mo.)	-0.04	0.02	0.059
Risk * Age (7.5 to 9.5 mo.)	0.02	0.02	0.426
Risk * Age (9.5 to 12 mo.)	-0.01	0.02	0.679
VP * Risk * Age (7.5 to 9.5 mo.)	0.06	0.03	0.063
VP * Risk * Age (9.5 to 12 mo.)	0.06	0.03	0.039

There was a significant two-way interaction for CBR between VP level and Age for 7.5 to 9.5 months (p < .001, b = -.06); this interaction reflects the fact that CBRs differed more between high VP and low VP at 9.5 than at 7.5 months. The difference between VP level and Age for 9.5 to 12 months approached significance (p = .059), and the effect was in the opposite direction, namely CBRs differed less for high VP vs low VP at 12 than at 9.5 months. There were no differences between Risk and Age at either age comparison.

There was a significant three-way interaction between VP level, Risk, and Age for ages 9.5 and 12 months (p = .039, b = .06). The three-way interaction for Risk, VP level, and Age at 7.5 and 9.5 months approached significance (p = .063). Figure 2 provides a graphic display of the effects found with the second model and helps illustrate the nature of the three-way interactions. The data from segments rated as having high VP (right-hand panel) suggest a tendency of CBR to grow rapidly from 7.5 to 9.5 months in the HR infants, but to grow much less rapidly in the LR infants. The opposite growth pattern (LR more rapid, HR less rapid) is seen from 9.5 to 12 months. No such differentiation is observable in the left panel. Thus, the data suggest the LR and HR infants show very different patterns of growth in CBR with age, but only in cases of high VP.

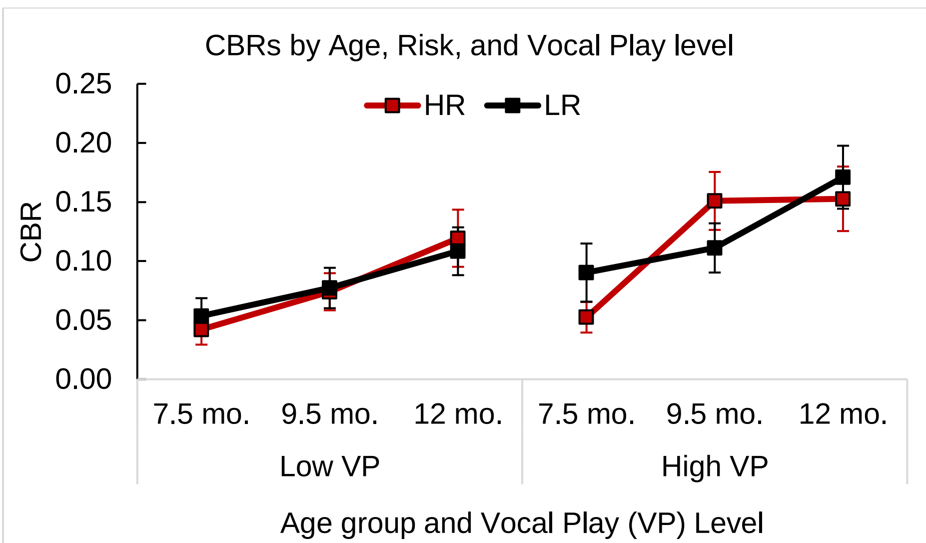

Figure 2. Canonical babbling by Age, Risk, and Vocal Play

Canonical babbling ratios (CBRs) of infants at high risk (HR) and low risk (LR) for autism in segments with low vs high vocal play (VP) across three age ranges, 6.5-8.49 (7.5 months), 8.5-10.49 (9.5 months), and 10.5-13 (12 months) months. In this model, there was a significant effect of Age for both 7.5 to 9.5 months (p = .023, b = .03) and 9.5 to 12 months (p = .001, b = .05). A significant interaction occurred between Risk and VP level (p = .021, b = -.03) and Age and VP level at 7.5 to 9.5 months (p < .001, b = -.06), with the interaction approaching significance for ages 9.5 to 12 months (p = .059). The three-way interaction among VP level, Age, and Risk was significant for ages 9.5 to 12 months (p = .039, b = .06), and approached significance for 7.5 to 9.5 months (p = .063). As in the case of Figure 1, the values presented here were computed from the raw data with means and SEs weighted for the number of infants who contributed data in each Risk group at each Age. Standard error (SE) bars are shown.

Age and Risk

1. Based on the preponderance of prior research in autism, we predicted that *CBRs of LR infants would increase to a greater extent across the three ages than CBRs of HR infants*. The results did not conform simply to the prediction. In fact CBRs for HR infants rose *more* in the first age interval (from 7.5 to 9.5 months, ~.067 CBR units) than for LR infants (~.015), while they rose *less* in the second interval for HR infants (~.010) than for LR infants (~.065). These patterns corresponded to a significant interaction of Risk by Age at the first interval (7.5 to 9.5 months, p = .017, b = .04), but a non-significant interaction of Risk by Age at the second interval (9.5 to 12 months, p = .192).

Table 5 presents the full GEE model comparing Age and Risk groups. There was a significant main effect for both Age intervals (7.5 to 9.5 months, p < .001, b = .06; 9.5 to 12 months, p = .047, b = .03), suggesting an overall increase in CBRs over time, as expected. As in the prior models, there was no significant difference between Risk groups (p = .319).

Table 5. Age and Risk interaction model

GEE interaction model for CBR with Age (7.5, 9.5, and 12 mo.) and Risk (HR and LR) only.

Variable	Effect size (*b*)	SE	p
Age 7.5 to 9.5 mo.	0.06	0.01	< .001
Age 9.5 to 12 mo.	0.03	0.01	0.047
Risk	-0.02	0.02	0.319
Risk * Age (7.5 – 9.5 mo.)	0.04	0.02	0.017
Risk * Age (9.5 – 12 mo.)	0.02	0.02	0.192

Figure 3 illustrates these data, showing CBRs of LR infants increased only slightly in the first age interval and a much larger increase in the second interval. Conversely, CBRs in the HR group increased much more in the first interval than in the second. Comparing this interaction with the data in Figures 1 and 2 offer perspective. In Figure 1 (TT model), Risk and Age interacted such that the greater growth of CBR for HR infants in the first age interval applied primarily to the circumstance of No TT, although the three-way interactions corresponding to this observation were not significant. In Figure 2 (VP model), Risk and Age interacted such that the greater growth of CBR for HR infants in the first age interval applied primarily to the circumstance of high VP, and the three-way interactions corresponding to this observation were significant for the first interval and approached significance for the second.

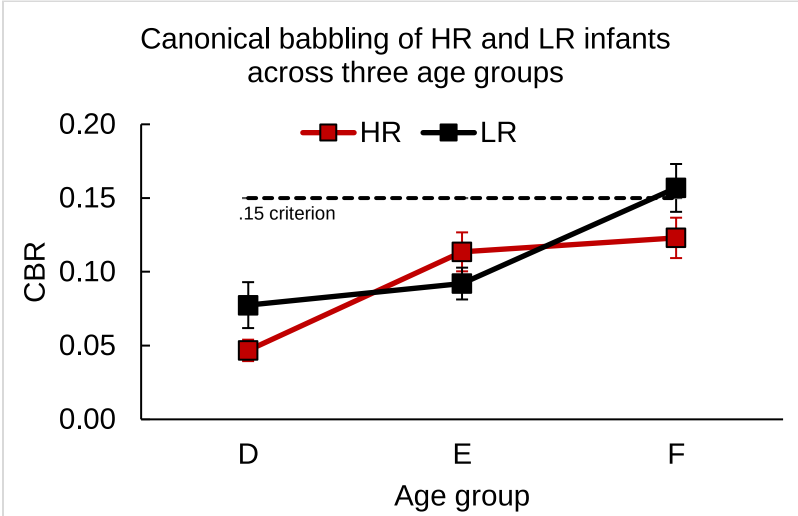

Figure 3. Canonical babbling ratios by Age and Risk

Canonical babbling ratios of infants at high-risk (HR) and low-risk (LR) for autism across three age ranges, 6.5-8.49 (7.5 mo.), 8.5-10.49 (9.5 mo.), and 10.5-13 (12 mo.). Overall, we found a significant interaction of Risk by Age for the first interval (7.5 to 9.5 months, p = .017), with CBRs rising much faster for HR infants than LR infants. The pattern was reversed, but not significantly in the second interval. Standard error (SE) bars shown.

A comment on the magnitude of the CBRs reported here seems warranted. The present data are based on all-day recordings sampled randomly; the CBRs are considerably lower than in prior reports based largely on short recordings usually conducted in laboratories and often selected for high infant volubility and/or interactivity. The Figure displays the criterion level of CBR that has sometimes been suggested to determine whether an infant is in the canonical stage based on a recorded sample (Lewedag, 1995; Oller, 2000; Oller et al., 2001; Patten et al., 2014). The mean CBR reached this .15 criterion for the LR infants only, and they reached it at 12 months only. The data suggest that the criterion level CBR for onset of the canonical stage should be considerably lower for all-day recordings sampled randomly than for laboratory recordings.

Main Effects

In a separate GEE model analyzing main effects only (Table 6), we found a significant effect of Age at both intervals (7.5 to 9.5 months, $p < .001$, $b = .04$; 9.5 to 12 months, $p < .001$, $b = .04$), evidencing a strong and near linear increase of CBRs over time for data amalgamated across the Risk groups and independent of TT and VP. There was also a significant effect for both TT ($p < .001$, $b = .04$) and VP ($p < .001$, $b = .06$). The effect sizes, reflected in the b values from the GEE analysis, can be placed in perspective by considering that TT had an effect roughly of the same magnitude as 2-3 months of growth in CBR, and that VP had an even larger effect.

Table 6. Main effects for Age, Risk, TT, and VP

Main effects model for Age (7.5, 9.5, and 12 months), Risk (LR and HR), Turn Taking (TT) level (No TT vs Any TT), and Vocal Play (VP) level (Low VP vs High VP).

Variable	Effect size (*b*)	SE	p
Age 7.5 to 9.5 mo.	0.04	0.01	< .001
Age 9.5 to 12 mo.	0.04	0.01	< .001
Risk	0.004	0.01	0.742
TT	0.04	0.01	< .001
VP	0.06	0.01	< .001

The magnitude of the significant effects by Cohen's *d*, computed from the raw data—with means and SEs weighted for the number of infants who contributed data in each Risk group at each Age—was 0.29 (small) for both TT and VP. The Age effect size was 0.36 (small) for the first interval, 0.21 (small) for the second, and 0.55 (medium) for a comparison of 7.5 months with 12 months. There was no main effect of Risk (p = .742). Figure 4 displays these main effects, including significantly higher CBRs during both any TT and high VP compared to periods of no TT and low VP, respectively.

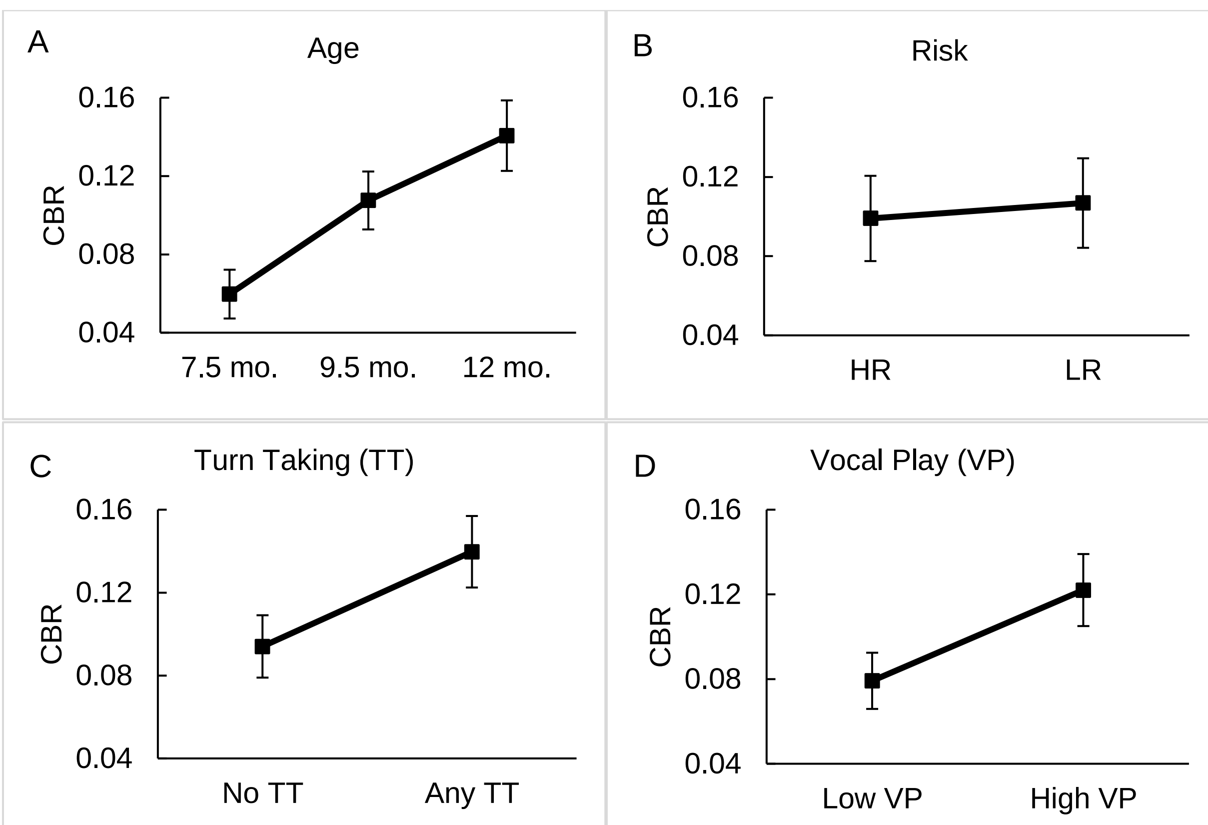

Figure 4. Main effects for Age, Risk, Turn Taking, and Vocal Play

Figure 4A illustrates the significant main effects of Age between 7.5 and 9.5 months ($p < .001$, $b = .04$) and 9.5 and 12 months ($p < .001$, $b = .04$). 4B shows the non-significant main effect of Risk group ($p = .742$). 4C *presents the significant main effect of Turn Taking, with higher CBRs during segments rated as having any TT compared to those rated as having no TT (p < .001, b = .04). Finally, 4D shows the significant main effect of Vocal Play, with higher CBRs present during segments with high VP compared to segments with low VP (p < .001, b = .06). Standard error (SE) bars are shown.*

Discussion

The present work evaluated the emergence of canonical babbling by comparing canonical babbling ratios (CBRs) in 98 infants either at low or high risk for autism across 3784 five-minute segments, selected from all-day recordings in the infants homes across the second half-year of life. The segments were coded by a team of trained listeners, who determined both CBRs and frequencies of vocal turn taking (TT) as well as vocal play (VP) in each segment. We addressed these data with expectations derived in part from the social motivation theory of autism, assuming that infants at high risk (HR) for autism may show lower social motivation than infants at low risk (LR). We also considered the data in light of evo-devo, a biological perspective in which it has been posited that early language development is driven by interplay between social motivation (presumably reflected in infant interest in caregiver vocalizations and in protoconversation) and an endogenous inclination in infants to produce copious amounts of vocalization, one that appears to have been naturally selected as a signal of fitness. These theoretical views led us to propose ways that risk for autism might play an important role in the emergence of canonical babbling, the sort of infant vocalization that is required in order for word learning to be launched in earnest, since words are overwhelmingly composed of canonical syllables.

Our particular predictions about effects of Risk did not, however, play out in the data. We observed no main effect of autism risk on CBRs. The finding adds further uncertainty to the already mixed evidence on canonical babbling emergence in autism and autism risk. The results support the argument that canonical babbling may be a robust developmental phenomenon and is more resistant to autism or autism risk than may have been previously assumed. Furthermore, and in contradiction to our initial expectation, we did not observe an overall tendency for CBRs to grow faster across Age in LR than HR infants. Instead, we found a tendency for CBRs of HR infants to grow faster in the first age interval (7.5 to 9.5 months) while CBRs of LR infants grew faster in the second (9.5 to 12 months). This pattern proved to be especially associated with segments where infants engaged in high VP, that is, when they were not vocalizing to other people, but vocalizing endogenously.

Overall main effects revealed, of course, the expected strong effect of Age on CBRs, a finding consistent with all prior longitudinal studies of canonical babbling. The present data do, on the other hand, provide new findings: we observed high CBRs in both Risk groups during segments with TT and during segments with high VP. The effect of TT was considerable, being equivalent to 2-3 months of growth in CBR, and the effect was even larger for high VP.

Social Motivation in Early Infancy

The social motivation reasoning behind our predictions is based in the assumption that HR infants may present with a reduced experience of social reward compared to LR infants and thus demonstrate early differences (presumably reductions) in vocal performance during social interaction. The findings for CBRs during TT, however, suggest similar levels of social motivation in both groups, with both showing the tendency to produce higher CBRs during segments rated as having any TT compared to those rated as having no TT whatsoever. These findings suggest robustness of social motivations for infant vocalization. Our hypotheses were based on an expectation of anomalous development in HR infants, assuming social motivation for vocalization may break down in the presence of neurodevelopmental differences affecting social cognition. The results suggest a stronger mechanism where human infant vocal tendencies may have been selected to withstand the neurodevelopmental differences associated with autism risk.

There can be no doubt that humans are highly social beings. Clearly, early hominins' relatively large living groups necessitated a high level of social bonding, which created a need for an efficient communication method, and resulted in positive selection pressures on the evolution of language (Dunbar, 1993, 1996, 2004). Chevallier (2012) noted that "social motivation constitutes an evolutionary adaptation geared to enhance the individual's fitness in collaborative environments" (p. 2). Thus, it is reasonable to assume that precursors to language such as canonical babbling must be robust during development. Although often delayed in developmental disorders, including autism (Chericoni et al., 2016; Iverson & Wozniak, 2007; Patten et al., 2014), canonical babbling is well-established as a robust stage of development, known to emerge eventually even when infants cannot hear sounds produced in their environment, as in the case of deafness or severe hearing impairment (Eilers & Oller, 1994; Oller & Eilers, 1988). Our results indicated no overall difference between CBRs of HR and LR infants—only the patterns of growth in CBR appeared to differ—suggesting the quality of prelinguistic vocal forms (i.e., CBR) produced during early face-to-face interactions may be robust with respect to these evolutionary pressures.

One important consideration and potential limitation in this evaluation of social motivations in early infancy relates to the measures we used to assess the sociality of vocalizations. To measure infant TT, coders were asked to estimate on a Likert scale how often infants engaged in TT for each segment. This subjective measure, obtained immediately after coding for CBR for each segment, can be portrayed as a blunt instrument, subject to only fair inter-observer agreement, but it is founded in the notion that human judgments are the gold standard for any such measure, and our method of obtaining the judgments was convenient and workable. A perhaps more reliable measure would require labeling the social or exploratory function of each utterance individually with repeat-observation (and especially with both audio and video), a measure that requires at least tenfold more time to obtain (see Long et al. (2020) for an analysis using this method). Future studies using this more expensive measure of the role of TT in infant vocalization are planned. An additional consideration includes examining infant-directed speech using similar methods employed in this study, as briefly discussed in Appendix A.

TT occurred, according to the coders, in only about 20% of the segments, a pattern that applied roughly equally to both Risk groups. This low rate of TT surprised us, given that so much of the literature on early language development focuses on protoconversation and its presumable importance in development. The low rate of TT may also have imposed a power limitation on the statistical analyses of the effects of TT and its interactions with the other variables in the present work.

Endogenous Motivation and Canonical Babbling

The VP measure was also based on a Likert scale, where coders were asked to judge each segment on how much of the time the infant had engaged in independent, not socially-directed vocalization (presumably endogenously motivated). Unlike TT, VP was found by the coders to be present in the vast majority of segments, and again this was true of both Risk groups—the plurality of segments having been rated 5 (VP present in close to the whole segment) by the coders for both Risk groups. Our surprise at low rates of TT in the all-day recordings is matched by our surprise at the near omnipresence of VP.

Again, however, the instrument measuring VP can be portrayed as blunt, having been obtained as a quick judgment from coders right after having completed listening to each segment and lacking high inter-coder agreement. The subjectivity of the judgments can be viewed in the same way as the TT judgments—human coding must be the gold standard in spite of its limitations. However, as with TT, more time-consuming judgments with audio and video and with repeat-observation coding are desirable.

Our hypotheses regarding VP were also based in part on the social motivation theory. We expected HR infants to show vocal behaviors similar to motoric behaviors that are characteristic of autism, such as frequent isolated play, stereotypic repetition of motoric behaviors, and preference for physical properties of objects (and thus acoustic-perceptual properties of sounds). Therefore, we anticipated HR infants would show a tendency to produce more canonical syllables than LR infants during high VP.

Overall, both the LR and HR groups produced more canonical syllables during high VP compared to low VP, but perhaps the most interesting outcome was the three-way interaction in the full VP model. The interaction suggests different rates of growth in CBR for HR and LR infants during the first and second age intervals (HR infants progressing faster in the first interval, LR infants faster in the second), but only for high VP segments. Low VP segments showed no such differentiation of Risk groups.

The social motivation theory posits that reduced early social reward processing affects later social cognitive functioning; however, Bottini (2018) described alternative hypotheses that have also been proposed to describe differences observed in autism, including general reward processing deficits in both social *and* non-social domains (Dichter et al., 2012; Kohls et al., 2013), and greater reward processing for non-social stimuli (Benning et al., 2016; Kohls et al., 2014; Sasson et al., 2012). Our findings hint at the possibility that whatever the social motivation or reward systems are, they may function differentially at different points in time for infants at risk for autism and for infants not at risk. One might propose that HR infants may experience greater intrinsic reward when producing canonical syllables during bouts of vocal play (i.e., as non-social stimuli) compared to LR infants; yet this greater reward applied from 7.5 to 9.5 months, while dropping substantially from 9.5 to 12 months.

As previously mentioned, one of the primary diagnostic characteristics of autism is the presence of restricted interests and repetitive behaviors (RRBs), including repetitive movements with objects, repeated body movements, ritualistic behavior, restricted interests, and sensory sensitivities (American Psychiatric Association, 2013). RRBs are present in typically developing infants (Arnott et al., 2010), but occur more frequently in infants with autism than in neurotypical controls as young as 6 months of age (Richler et al., 2007; Rogers, 2009). High rates of canonical syllables observed during bouts of VP in these HR infants may represent manifestations of vocal stereotypies, similar to those seen in autism. It is thought that autistic infants may prefer playing with the sensorimotor characteristics of a syllable through repetition, while their neurotypical counterparts tend to play with varying aspects pertaining to individual syllables, modifying duration, placement, and various articulatory patterns from utterance to utterance. Thus, attending to the repetitive physical and acoustic properties of sounds during bouts of VP may be more intrinsically rewarding to infants with autism compared to typical development, who may be vocally exploring phonetic nuances. This idea is supported by the speech attunement framework (Shriberg et al., 2011), which proposes that autistic children process acoustic-perceptual characteristics more easily than semantic-linguistic information (Heaton et al., 2008; Järvinen-Pasley et al., 2008; Mottron et al., 2006).

Yet, the higher CBRs in HR infants compared to LR infants during high VP applied only for the first age interval, with an opposite pattern occurring thereafter (LR infants showing greater growth of CBR). Thus, if HR infants' increased CBRs in the first age interval are the result of autism-like repetition and stereotypy, there must be some other force at stake in the second age interval. Perhaps the robust tendency for canonical babbling to develop—based on the critical requirement for command of canonical syllables—drives all infants to reach a minimal level of canonical babbling control by the time word learning begins to take off at the end of the first year. Delays in the emergence rate of advanced vocal forms in infants at risk may become more evident at later ages as greater social and linguistic demands are placed on children who will show effects of autism. Such later delays may be foreshadowed in our finding of a plateauing of CBRs in HR infants by 12 months.

On the Criterion for Canonical Babbling Onset

Canonical babbling onset has often been often suggested in the vocal developmental literature as requiring a .15 CBR based on a coded sample (i.e., 15% of syllables in a sample are canonical) (Lewedag, 1995; Nathani et al., 2006; Oller, 2000). However, this level appears to be untenable for the kinds of recordings and methods reported on in the present work. Our findings reveal that even the LR infants (who were presumably all typically developing) did not reach this criterion level until 12 months of age, despite expected mastery by 7-10 months. Previous research supporting the .15 criterion has primarily been completed in laboratory conditions (Molemans et al., 2012; Oller et al., 1994), settings in which parents have been expected to intentionally (or unintentionally) elicit and induce infant vocalization. Our findings and those discussed by Oller et al. (2020) support the notion that the average CBR in all-day recordings is lower than in laboratory settings.

Conclusions

The findings observed in the present study offer perspective on the ability to detect developmental differences in infant vocal turn taking and independent vocal production as potential indicators of autism. We observed a similar emergence of canonical babbling in infants at low and high risk for autism, and higher rates of canonical babbling overall during segments rated as having any turn taking and high vocal play. Our findings offer support for a potentially robust social motivation in infancy to produce higher rates of canonical syllables during interaction, even in the presence of possible social communication deficits. Differences observed between groups did occur when comparing low and high levels of independent vocal play. Evolutionary pressures may play a role in high-risk infants' increased rate of canonical syllables during vocal play early in the canonical babbling stage as a result of the need to signal fitness prior to vocal delays at later ages. These differences may also support an age-varying heightened intrinsic reward mechanism for producing and attending to acoustic-perceptual characteristics of vocal sounds potentially linked to genes associated with autism.

Acknowledgements

We wish to thank the families in Atlanta, GA who participated in this research and graduate student coders of this research.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this research was funded by grants [DC015108] from the National Institute of Deafness and Communication Disorders and [MH100029] from the National Institute on Mental Health. Additional funding support provided by the Plough Foundation, the Holly Lane Foundation, the Marcus Foundation, the Woodruff-Whitehead Foundation, and the Georgia Research Alliance.

Ethics Statement

This study was carried out in accordance with the recommendations of the Institutional Review Board (IRB) guidelines for the University of Memphis and Emory University with written informed consent from all parents of infants in the study in accordance with the Declaration of Helsinki. The protocol was approved by the University of Memphis (No. 2143) and Emory University (No. IRB0000059383) IRB committees.

Author Contributions

GR collected recorded data used for coding and analysis. HLL, DKO, MMR, and DDB designed and analyzed the results of the study. DDB completed formal statistical analyses. All authors assisted in interpretation of results, writing, and revising the manuscript prior to submission.

Disclosures

None of the authors have any financial or non-financial interests to declare.

Data Accessibility Statement

The analyzed data and code that support the findings of this study will be openly available on the Figshare repository platform and on the Open Science Framework at https://osf.io/ery6b/. Due to the nature of this research, participants of this study did not agree for audio recordings to be shared publicly, so raw recording data is not available.

References

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2018). The social functions of babbling: Acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science*, *21*(5), 1–11. https://doi.org/10.1111/desc.12641

American Psychiatric Association. (2013). Autism Spectrum Disorder. In *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*.

Arnott, B., McConachie, H., Meins, E., Fernyhough, C., Couteur, A. Le, Turner, M., Parkinson, K., Vittorini, L., & Leekam, S. (2010). The frequency of restricted and repetitive behaviors in a community sample of 15-month-old infants. *Journal of Developmental and Behavioral Pediatrics*, *31*(3), 223–229. https://doi.org/10.1097/DBP.0b013e3181d5a2ad

Atlas, J. A. (1990). Play in assessment and intervention in the childhood psychoses. *Child Psychiatry and Human Development*, *21*(2), 199–133.

Baranek, G. T., Watson, L. R., Boyd, B. A., Poe, M. D., David, F. J., & McGuire, L. (2013). Hyporesponsiveness to social and nonsocial sensory stimuli in children with autism, children with developmental delays, and typically developing children. *Development and Psychopathology*. https://doi.org/10.1017/S0954579412001071

Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, *263*(1), 101–113. https://doi.org/10.1111/j.1749-6632.1975.tb41575.x

Belardi, K., Watson, L. R., Faldowski, R. A., Hazlett, H., Crais, E., Baranek, G. T., McComish, C., Patten, E., & Oller, D. K. (2017). A retrospective video analysis of canonical babbling and volubility in infants with Fragile X syndrome at 9 – 12 months of mge. *Journal of Autism and Developmental Disorders*, *47*(4), 1193–1206. https://doi.org/10.1177/0885066614530659.The

Benning, S. D., Kovac, M., Campbell, A., Miller, S., Hanna, E. K., Damiano, C. R., Sabatino-DiCriscio, A., Turner-Brown, L., Sasson, N. J., Aaron, R. V., Kinard, J., & Dichter, G. S. (2016). Late positive potential ERP Responses to social and nonsocial stimuli in youth with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(9), 3068–3077. https://doi.org/10.1007/s10803-016-2845-y

Berk, L. E. (1994). Vygotsky's theory: The importance of make-believe play. *Young Children*, *50*(1), 30–39.

Bertossa, R. C. (2011). Theme issue: Evolutionary developmental biology (evo-devo) and behaviour: Papers of a Theme issue compiled and edited by Rinaldo C. Bertossa. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 2055–2180. https://doi.org/10.1098/rstb.2011.0035

Bodnarchuk, J. L., & Eaton, W. O. (2004). Can parent reports be trusted? Validity of daily checklists of gross motor milestone attainment. *Applied Developmental Psychology*, *25*, 481–490. https://doi.org/10.1016/j.appdev.2004.06.005

Bottini, S. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders*, *45*, 9–26. https://doi.org/10.1016/j.rasd.2017.10.001

Bråten, S. (1988). Dialogic mind: The infant and the adult in protoconversation. In M. E. Carvallo (Ed.), *Nature, Cognition and System I. Theory and Decision Library (Series D: System Theory, Knowledge Engineering and Problem Solving), vol 2* (pp. 187–205). Springer. https://doi.org/10.1007/978-94-009-2991-3_9

Bruner, J. S. (1974). From communication to language- A psychological perspective. *Cognition*, *3*(3), 255–287. https://doi.org/10.1016/0010-0277(74)90012-2

Camarata, S., & Yoder, P. (2002). Language transactions during development and intervention: Theoretical implications for developmental neuroscience. In *International Journal of Developmental Neuroscience* (Vol. 20, Issues 3–5, pp. 459–465). Elsevier Ltd. https://doi.org/10.1016/S0736-5748(02)00044-8

Carroll, S. B. (2005). *Endless forms most beautiful: The new science of evo-devo and the making of the animal kingdom*. W. W. Norton & Co. https://doi.org/10.1086/503946

Chericoni, N., de Brito Wanderley, D., Costanzo, V., Diniz-Gonçalves, A., Gille, M. L., Parlato, E., Cohen, D., Apicella, F., Calderoni, S., & Muratori, F. (2016). Pre-linguistic vocal trajectories at 6-18 months of age as early markers of autism. *Frontiers in Psychology*, *7*(OCT), 1595. https://doi.org/10.3389/fpsyg.2016.01595

Chevallier, C., Huguet, P., Happé, F., George, N., & Conty, L. (2013). Salient social cues are prioritized in autism spectrum disorders despite overall decrease in social attention. *Journal of Autism and Developmental Disorders*, *43*(7), 1642–1651. https://doi.org/10.1007/s10803-012-1710-x

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Davis, K. L., & Panksepp, J. (2018). *The emotional foundations of personality: A neurobiological and evolutionary approach*. W.W. Norton & Company.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.40.2.271

Delgado, C. E. F., Messinger, D. S., & Yale, M. E. (2002). Infant responses to direction of parental gaze: A comparison of two still-face conditions. *Infant Behavior and Development*, *25*(3), 311–318. https://doi.org/10.1016/S0163-6383(02)00096-6

Dichter, G. S., Felder, J. N., Green, S. R., Rittenberg, A. M., Sasson, N. J., & Bodfish, J. W. (2012). Reward circuitry function in autism spectrum disorders. *Social Cognitive and Affective Neuroscience*, *7*(2), 160–172. https://doi.org/10.1093/scan/nsq095

Dunbar, R. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*, 681–735. https://doi.org/10.1017/s0140525x00032325

Dunbar, R. M. (1996). *Gossiping, grooming and the evolution of language*. Harvard University Press.

Dunbar, R. M. (2004). Language, music, and laughter in evolutionary perspective. In D. K. Oller & U. Griebel (Eds.), *The Evolution of Communication Systems: A Comparative Approach* (pp. 257–

274). MIT Press. https://doi.org/10.7551/mitpress/2879.003.0021

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

Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019). The ecology of prelinguistic vocal learning: Parents simplify the structure of their speech in response to babbling. *Journal of Child Language*, *46*, 998–1011. https://doi.org/10.1017/S0305000919000291

Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C.-C., Lambert, B., & Oller, D. K. (2013). Effects of parental interaction on infant vocalization rate, variability and vocal type. *Language Learning and Development*, *10*(3), 279–296. https://doi.org/10.1080/15475441.2013.849176

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the USA*, *100*(13), 8030–8035. https://doi.org/10.1073/pnas.1332441100

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development*, *80*(3), 636–644. https://doi.org/10.1111/j.1467-8624.2009.01287.x

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*(1), 99–108. https://doi.org/10.1037/0012-1649.28.1.99

Gray, K. L. H., Haffey, A., Mihaylova, H. L., & Chakrabarti, B. (2018). Lack of privileged access to awareness for rewarding social scenes in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *48*(10), 3311–3318. https://doi.org/10.1007/s10803-018-3595-9

Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, *19*(4), 385–408. https://doi.org/10.1111/infa.12054

Harold, M. P., & Barlow, S. M. (2013). Effects of environmental stimulation on infant vocalizations and orofacial dynamics at the onset of canonical babbling. *Infant Behavior and Development*, *36*(1), 84–93. https://doi.org/10.1016/j.infbeh.2012.10.001

Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology*, *25*(6), 771–782. https://doi.org/10.1080/02643290802336277

Hsu, H. C., & Fogel, A. (2001). Infant vocal development in a dynamic mother-infant communication system. *Infancy*, *2*(1), 87–109. https://doi.org/10.1207/S15327078IN0201_6

Iverson, J. M., & Wozniak, R. H. (2007). Variation in vocal-motor development in infant siblings of children with autism. *Journal of Autism and Developmental Disorders*, *37*(1), 158–170. https://doi.org/10.1007/s10803-006-0339-z

Iyer, S. N., Denson, H., Lazar, N., & Oller, D. K. (2016). Volubility of the human infant: Effects of parental interaction (or lack of it). *Clinical Linguistics and Phonetics*, *30*(6), 470–788. https://doi.org/10.3109/02699206.2016.1147082

Järvinen-Pasley, A., Wallace, G. L., Ramus, F., Happé, F., & Heaton, P. (2008). Enhanced perceptual processing of speech in autism. *Developmental Science*, *11*(1), 109–121. https://doi.org/10.1111/j.1467-7687.2007.00644.x

Kellerman, A. M., Schwichtenberg, A. J., Tonnsen, B. L., Posada, G., & Lane, S. P. (2019). Dyadic interactions in children exhibiting the broader autism phenotype: Is the broader autism phenotype distinguishable from typical development? *Autism Research*, *12*(3), 469–481. https://doi.org/10.1002/aur.2062

Kohls, G., Schulte-Rüther, M., Nehrkorn, B., Müller, K., Fink, G. R., Kamp-Becker, I., Herpertz-Dahlmann, B., Schultz, R. T., & Konrad, K. (2013). Reward system dysfunction in autism spectrum disorders. *Social Cognitive and Affective Neuroscience*, *8*(5), 565–572. https://doi.org/10.1093/scan/nss033

Kohls, G., Thönessen, H., Bartley, G. K., Grossheinrich, N., Fink, G. R., Herpertz-Dahlmann, B., & Konrad, K. (2014). Differentiating neural reward responsiveness in autism versus ADHD. *Developmental Cognitive Neuroscience*, *10*, 104–116. https://doi.org/10.1016/j.dcn.2014.08.003

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In *Precursors of Early Speech* (pp. 37–50). Palgrave Macmillan UK.

Kuhl, P. K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, *10*(1), 110–120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Lang, S., Bartl-Pokorny, K. D., Pokorny, F. B., Garrido, D., Mani, N., Fox-Boyer, A. V., Zhang, D., & Marschik, P. B. (2019). Canonical babbling: A marker for earlier identification of late detected developmental disorders? *Current Developmental Disorders Reports*, *6*(3), 111–118. https://doi.org/10.1007/s40474-019-00166-w

LeBarton, E. S., & Iverson, J. M. (2016). Associations between gross motor and communicative development in at-risk infants. *Infant Behavior and Development*, *44*, 59–67. https://doi.org/10.1016/j.infbeh.2016.05.003

Lee, C.-C., Jhang, Y., Relyea, G., Chen, L.-M., & Oller, D. K. (2018). Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings. *Infant Behavior and Development*, *50*, 140–153. https://doi.org/10.1016/j.infbeh.2017.12.002

Levin, K. (1999). Babbling in infants with cerebral palsy. *Clinical Linguistics & Phonetics*, *13*(4), 249–267. https://doi.org/10.1080/026992099299077

Lewedag, V. L. (1995). *Patterns of onset of canonical babbling among typically developing infants*. University of Miami.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22.

Locke, J. L. (2017). Emancipation of the voice: Vocal complexity as a fitness indicator. *Psychonomic Bulletin & Review*, *24*(1), 232–237.

Lohmander, A., Holm, K., Eriksson, S., & Lieberman, M. (2017). Observation method identifies that a lack of canonical babbling can indicate future speech and language problems. *Acta Paediatrica, International Journal of Paediatrics*, *106*(6), 935–943. https://doi.org/10.1111/apa.13816

Long, H. L., Bowman, D. D., Yoo, H., Burkhardt-Reed, M. M., Bene, E. R., & Oller, D. K. (2020). Social and endogenous infant vocalizations. *PLoS ONE*, *15*(8), e0224956. https://doi.org/10.1371/journal.pone.0224956

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lynch, M. P., Oller, D. K., Steffens, M. L., Levine, S. L., Basinger, D. L., & Umbel, V. (1995). Onset of speech-like vocalizations in infants with Down syndrome. *American Journal of Mental Retardation*, *100*(1), 68–86.

Lyytinen, P., Poikkeus, A. M., Leiwo, M., Ahonen, T., & Lyytinen, H. (1996). Parents as informants of their child's vocal and early language development. *Early Child Development and Care*, *126*(1), 15–25. https://doi.org/10.1080/0300443961260102

Masataka, N. (2001). 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*. https://doi.org/10.1111/1467-7687.00161

Molemans, I., Van den Verg, R., Van Severen, L., & Gillis, S. (2012). How to measure the onset of babbling reliably? *Journal of Child Language*, *39*(3), 523–552. https://doi.org/10.1017/S0305000911000171

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27–43. https://doi.org/10.1007/s10803-005-0040-7

Moulin-Frier, C., Nguyen, S. M., & Oudeyer, P. Y. (2014). Self-organization of early vocal development in infants and machines: The role of intrinsic motivation. *Frontiers in Psychology*, *4*, 1006. https://doi.org/10.3389/fpsyg.2013.01006

Moulin-Frier, C., & Oudeyer, P. Y. (2013). The role of intrinsic motivations in learning sensorimotor vocal mappings: A developmental robotics study. *INTERSPEECH, ISCA*.

Mundy, P. (2017). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, 1–18. https://doi.org/10.1111/ejn.13720

Naber, F. B. A., Bakermans-Kranenburg, M. J., Van Ijzendoorn, M. H., Swinkels, S. H. N., Buitelaar, J. K., Dietz, C., Van Daalen, E., & Van Engeland, H. (2008). Play behavior and attachment in toddlers with autism. *Journal of Autism and Developmental Disorders*, *38*(5), 857–866.

https://doi.org/10.1007/s10803-007-0454-5

Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics and Phonetics*, *20*(5), 351–369. https://doi.org/10.1080/02699200500211451

Newman, S. A. (2000). The role of genetic reductionism in biocolonialism. *Peace Review: A Journal of Social Justice*, *12*(4), 517–524. https://doi.org/10.1080/10402650020014592

Newman, S. A. (2012). Physico-genetic determinants in the evolution of development. *Science*, *338*, 217–219. https://doi.org/10.1126/science.1224311

Nyman, A., & Lohmander, A. (2018). Babbling in children with neurodevelopmental disability and validity of a simplified way of measuring canonical babbling ratio. *Clinical Linguistics and Phonetics*, *32*(2), 114–127. https://doi.org/10.1080/02699206.2017.1320588

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology, Volume 1, Production* (pp. 93–1123). Academic Press.

Oller, D. K. (2000). *The emergence of the speech capacity*. Psychology Press.

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L. B., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6318–6323. https://doi.org/10.1073/pnas.1300337110

Oller, D. K., Caskey, M., Yoo, H., Bene, E. R., Jhang, Y., Lee, C.-C., Bowman, D. D., Long, H. L., Buder, E. H., & Vohr, B. (2019). Preterm and full term infant vocalization and the origin of language. *Scientific Reports*, *9*, 14734. https://doi.org/10.1038/s41598-019-51352-0

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development*, *59*(2), 441–449. https://doi.org/10.1111/j.1467-8624.1988.tb01479.x

Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science*, *4*, 49–60. https://doi.org/10.1111/1467-7687.00148

Oller, D. K., Eilers, R. E., Neal, A. R., & Cobo-Lewis, A. B. (1998). Late onset canonical babbling: A possible early marker of abnormal development. *American Journal on Mental Retardation*, *103*(3), 249. https://doi.org/10.1352/0895-8017(1998)103<0249:LOCBAP>2.0.CO;2

Oller, D. K., Eilers, R. E., Steffens, M. L., Urbano, R., & Lynch, M. P. (1994). Speech-Like Vocalizations in Infancy: An Evaluation of Potential Risk Factors. *Journal of Child Language*, *21*(1), 33–58. https://doi.org/10.1017/S0305000900008667

Oller, D. K., & Griebel, U. (2005). Contextual freedom in human infant vocalization and the evolution of language. In R. L. Burgess & K. MacDonald (Eds.), *Evolutionary Perspectives on Human Development* (pp. 135–166). SAGE Publications. https://doi.org/10.4135/9781452233574.n5

Oller, D. K., & Griebel, U. (2008). Contextual flexibility in infant vocal development and the earliest steps in the evolution of language. In D. K. Oller & U. Griebel (Eds.), *Evolution of Communicative Flexibility: Complexity, Creativity and Adaptability in Human and Animal Communication* (pp. 141–168). MIT Press. https://doi.org/10.7551/mitpress/9780262151214.003.0007

Oller, D. K., Griebel, U., Bowman, D. D., Bene, E. R., Long, H. L., Yoo, H., & Ramsay, G. (2020). Infant boys are more vocal than infant girls. *Current Biology*, *30*, R417-429. https://doi.org/10.1016/j.cub.2020.03.049

Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language. *Topics in Cognitive Science*, *8*(2), 382–392. https://doi.org/10.1111/tops.12198.Vocal

Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., Rogers, S. J., Rozga, A., Sangha, S., Sigman, M., Steinfeld, M. B., & Young, G. S. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(3), 256-66.e1-2.

Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition*, *14*(1), 30–80. https://doi.org/10.1016/j.concog.2004.10.004

Panksepp, J., & Biven, L. (2012). *The archaeology of mind: Neuroevolutionary origins of human emotions*. W.V. Norton & Company. https://doi.org/10.5860/choice.50-3555

Panksepp, J., Siviy, S., & Normansell, L. (1984). The psychobiology of play: Theoretical and methodological perspectives. *Neuroscience and Biobehavioral Reviews*, *8*(4), 465–492. https://doi.org/10.1016/0149-7634(84)90005-8

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders*, 1–16. https://doi.org/10.1007/s10803-014-2047-4

Paul, R., Fuerst, Y., Ramsay, G., Chawarska, K., & Klin, A. (2011). Out of the mouths of babes: Vocal production in infant siblings of children with ASD. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *52*(5), 588–598. https://doi.org/10.1111/j.1469-7610.2010.02332.x

Piaget, J. (1952). *Play, dreams and imitation in childhood*. W. W. Norton & Co. https://doi.org/10.4324/9781315009698

Pokorny, F. B., Schuller, B. W., Marschik, P. B., Brueckner, R., Nyström, P., Cummins, N., Bölte, S., Einspieler, C., & Falck-Ytter, T. (2017). Earlier identification of children with autism spectrum disorder: An automatic vocalisation-based approach. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2017-Augus*, 309–313. https://doi.org/10.21437/Interspeech.2017-1007

Richler, J., Bishop, S. L., Kleinke, J. R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(1), 73–85. https://doi.org/10.1007/s10803-006-0332-6

Rochat, P., Querido, J. G. Q., & Striano, T. (1999). Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Developmental Psychology*, *35*(4), 950–957.

Rogers, S. J. (2009). What are infant siblings teaching us about Autism in infancy? *Autism Research*, *2*(3), 125–137. https://doi.org/10.1002/aur.81

Sasson, N. J., Dichter, G. S., & Bodfish, J. W. (2012). Affective responses by adults with autism are reduced to social images but elevated to images related to circumscribed interests. *PLoS ONE*, *7*(8), e42457. https://doi.org/10.1371/journal.pone.0042457

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., Skudlarski, P., Lacadie, C., Cohen, D. J., & Gore, J. C. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archives of General Psychiatry*, *57*, 331–340. https://doi.org/10.1001/archpsyc.57.4.331

Scott-Van Zeeland, A. A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., & Bookheimer, S. Y. (2010). Reward Processing in Autism. *Autism Research*, *3*(2), 53–67. https://doi.org/10.1002/aur.122

Sepeta, L., Tsuchiya, N., Davies, M. S., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2012). Abnormal social reward processing in autism as indexed by pupillary responses to happy faces. *Journal of Neurodevelopmental Disorders*, *4*(1), 1–9. https://doi.org/10.1186/1866-1955-4-17

Sheinkopf, S. J., Iverson, J. M., Rinaldi, M. L., & Lester, B. M. (2012). Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorder. *Autism Research*, *5*(5), 331–339. https://doi.org/10.1002/aur.1244

Shriberg, L. D., Paul, R., Black, L. M., & Van Santen, J. P. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *41*(4), 405–426. https://doi.org/10.1007/s10803-010-1117-5

Sigman, M., & Ungerer, J. A. (1984). Attachment behaviors in autistic children. *Journal of Autism and Developmental Disorders*, *14*(3), 231–244. https://doi.org/10.1007/BF02409576

Stark, R. E. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanaugh, & C. Ferguson (Eds.), *Child Phonology* (Vol. 1, pp. 73–90). Academic Press.

Stark, R. E. (1981). Infant vocalization: A comprehensive view. *Infant Mental Health Journal*, *2*(2), 118–128. https://doi.org/10.1002/1097-0355(198122)2:2<118::AID-IMHJ2280020208>3.0.CO;2-5

Su, P. L., Rogers, S. J., Estes, A. M., & Yoder, P. J. (2020). The role of early social motivation in explaining variability in functional language in toddlers with ASD. *Autism: The International Journal of Research & Practice*. https://doi.org/10.1177/1362361320953260

Syal, S. (2011). *Socially motivated vocal learning.* [Doctoral dissertation, Cornell University]. https://doi.org/10.16194/j.cnki.31-1059/g4.2011.07.016

Vygotsky, L. S. (1978). Interaction between learning and development. In *Mind in Society* (pp. 79–91). Harvard University Press.

Weeks, S. J., & Hobson, R. P. (1987). The salience of facial expression for autistic children. *Journal of Child Psychology and Psychiatry*, *28*(1), 137–152. https://doi.org/10.1111/j.1469-7610.1987.tb00658.x

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief report: Recognition of autism spectrum disorder before one year of age: A retrospective study based on home videotapes. *Journal of Autism and Developmental Disorders*, *30*(2), 157–162. https://doi.org/10.1023/A:1005463707029

Williams, E., Reddy, V., & Costall, A. (2001). Taking a closer look at functional play in children with autism. *Journal of Autism and Developmental Disorders*, *31*(1), 67–77. https://doi.org/10.1023/A:1005665714197

Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, *124*, 342–349.

Zwaigenbaum, L., Bauman, M. L., Stone, W. L., Yirmiya, N., Estes, A., Hansen, R. L., McPartland, J. C. Natowicz, M. R., Choueiri, R., Fein, D., Kasari, C., Pierce, K., Buie, T., Carter, A., Davis, P. A., Granpeesheh, D., Mailloux, Z., Newschaffer, C., Robins, D., Roley, S. S., … Wetherby, A. (2015). Early identification of autism spectrum disorder: Recommendations for practice and research. *Pediatrics*, *136*, (Supplement 1) S10-S40. https://doi.org/10.1542/peds.2014-3667C

Appendix A

Considerations Regarding Infant-Directed Speech

The literature on early language suggests infant-directed speech (IDS) may also influence the emergence of canonical babbling, as previous research has highlighted the effects of social interaction on babbling (Albert et al., 2018; Goldstein et al., 2003; Goldstein & Schwade, 2008) and conversely, the effects of babbling on caregiver speech during interaction (Elmlinger et al., 2019). During our analyses, we ran a secondary main effects model including IDS as a variable. Our coding protocol also included counts of both infant- and other-directed speech (i.e., speech between two adults) in each segment, affording the opportunity to compare counts or proportions of IDS to CBRs. We found a significant effect of IDS on CBRs (p = .034, b = -.0004), but notably, this effect was extremely small and negative. Furthermore, the correlation between canonical babbling ratios and total IDS showed a weak, negative correlation (r = -.02). Because we continue to believe IDS is a variable worth exploring further as an influence on canonical babbling—both as a continuous variable based on the coded amount of IDS or as a categorical factor (i.e., Low vs High IDS or No vs Any IDS) based on questionnaire judgments of the same type as explored for TT and VP in the present research—we plan to examine IDS effects on canonical babbling more explicitly in a future paper.