CAMBRIDGE
UNIVERSITY PRESS

BRIEF RESEARCH REPORT

The ecology of prelinguistic vocal learning: parents simplify the structure of their speech in response to babbling

Steven L. ELMLINGER, Jennifer A. SCHWADE, and Michael H. GOLDSTEIN*

Department of Psychology, Cornell University, USA
*Corresponding author: E-mail: mhg26@cornell.edu

Abstract

What is the function of babbling in language learning? We examined the structure of parental speech as a function of contingency on infants' non-cry prelinguistic vocalizations. We analyzed several acoustic and linguistic measures of caregivers' speech. Contingent speech was less lexically diverse and shorter in utterance length than non-contingent speech. We also found that the lexical diversity of contingent parental speech only predicted infant vocal maturity. These findings illustrate a new form of influence infants have over their ambient language in everyday learning environments. By vocalizing, infants catalyze the production of simplified, more easily learnable language from caregivers.

Introduction

In many species, parents' responses to their offspring's immature vocalizations play a key role in the development of communication. Results from vocal learning studies on humans (Goldstein & Schwade, 2008), songbirds (King, West, & Goldstein, 2005), and marmosets (Takahashi et al., 2015; Gultekin & Hage, 2018) indicate that adults who coordinate their vocalizations with those of their offspring create contingent social feedback that facilitates learning of more advanced vocal patterns. Human infants have a long period of vocal immaturity, during which vocal development seems particularly open to social input (e.g., Kuhl, Tsao, & Lui, 2003; Goldstein & Schwade, 2010; Ramírez-Esparza, García-Sierra, & Kuhl 2017). By five months, infants have come to expect that their babbling (i.e., all speech-related prelinguistic vocalizations, excluding cries and vegetative sounds; Oller, 2000; Warlaumont, Richards, Gilkerson, & Oller, 2014) will reliably elicit an adult's response (Goldstein, Schwade, & Bornstein, 2009). Social influences continue to facilitate improvements in vocal learning throughout the babbling phase. By nine months, infants produce more speech-like vocalizations when caregivers' responses are contingent on their

CrossMark

vocalizations (Goldstein, King, & West, 2003). Contingent vocal responses from parents predict the content of infants' vocal changes (e.g., vowel resonance and faster consonant–vowel transitions) (Goldstein & Schwade, 2008).

Findings from infant vocal learning studies clearly demonstrate the power of parents' contingent responses in facilitating vocal production learning, but the statistical and linguistic structure of contingent parental speech is unknown. In what ways does babbling influence the statistics of parental speech? We know that infants can detect many forms of auditory statistics. Prelinguistic infants can segment sequences of phonemes based on their co-occurrence regularities (Saffran, Aslin, & Newport, 1996). Infants also pool phoneme information to construct phonological categories found in their ambient language (e.g., Mattys, Jusczyk, Luce, & Morgan, 1999). Children use the co-occurrence of heard words and seen object statistics to rapidly learn object–referent mappings (Smith, Suanda, & Yu, 2014). Infants' sensitivity to the statistics of language input, and their ability to abstract structure from spoken language at multiple levels of linguistic organization, means that variation in parental speech plays a powerful role in infant language learning and development (Newman, Rowe, & Ratner, 2016).

Parental speech to infants is characterized by several acoustic and linguistic features that increase its signal value, organize infant attention, and simplify the language learning task. Infant-directed speech (IDS) generally consists of shorter utterances than are found in adult-directed speech (Newport, Gleitman, & Gleitman, 1977; Snow, 1977). While the linguistic content of IDS is simplified, the prosodic features (e.g., pitch contours) are elaborated. IDS has a higher overall pitch and larger dispersion of pitch compared to adult-directed speech (Fernald & Simon, 1984). Differences in the pitch of caregivers' speech may facilitate infants' attention to the co-occurrence of speech sounds over time (Thiessen, Hill, & Saffran, 2005).

Parents also tend to maintain a small set of consistent grammatical structures in their speech to infants, and those specific structures predict later child language (Cameron-Faulkner, Lieven, & Tomasello, 2003). Isolated words spoken to children are predictive of the words that are most likely to be produced by children later in development (Brent & Siskind, 2001). The simplicity of language spoken to children and infants may interact with the diversity or variability of words parents utter over extended time periods. Numerous findings from learning studies implicate the unique role of variability in the input, specifically the number of words and specific word types heard by children. Previous work demonstrates that exemplar variability promotes both language development (Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010) and generalization in learning auditory statistics (Vukatana, Graham, Curtin, & Zepeda, 2015).

Parents' speech and responsive behaviors to infants during free play with objects have been shown to facilitate word learning in toddlers between 12 and 24 months of age (e.g., Tamis-LeMonda, Bornstein, & Baumwell, 2001; Weisberg, Zosh, Hirsh-Pasek, & Golinkoff, 2013). Might babbling also influence parental speech in ways that facilitate language learning? The present research examines whether the structural patterns of parents' infant-directed speech change in response to babbling. Past findings suggest that the complexity of parents' interactions with their infants reflects infants' current linguistic ability. When speaking to 4-month-olds, mothers' utterances are characterized by more exaggerated pitch contours and lexical repetitiveness than when infants are newborns, 12, or 24 months of age (Stern, Spieker, Barnett, & MacKain 1983). Parents also raise the pitch of their speech to

their 4-month-old infants as a function of their infants' responses to their speech (Smith & Trainor, 2008). Parents' speech to 6-month-old infants utilizes more nonsense sounds to attract infants' attention than when infants are 12 or 19 months (Fernald & Morikawa, 1993). The length of parents' utterances to children in everyday learning environments may decrease until children learn specific target words within the utterances (Roy, Frank, & Roy, 2009). Thus, parents' speech is sensitive to children's overall developmental progression, but the role of prelinguistic vocalizing in organizing parents' speech has received little attention.

Recent findings suggest that the window of time after infants vocalize may prove crucial for the promotion of learning. Infants' own vocalizations may serve to modulate their attention and afford learning the mapping of parents' speech to attended objects in the environment (Albert, Schwade, & Goldstein, 2017). When infants babble while looking at an object, they create a state of receptivity for learning at the same moment caregivers are likely to provide specific information about the target object (Albert *et al.*, 2017). These object-directed vocalizations (ODVs), coupled with an immediate label of the object by an adult, result in stronger associations between word–object referents (Goldstein, Schwade, Briesch, & Syal, 2010). These findings suggest that contingent parental speech capitalizes on infant attentional focus and facilitates the mapping of heard words to seen objects.

We first compared the linguistic structure of parental speech as a function of its contingency on infant babbling. Parental speech structure was quantified in terms of three measures. To measure lexical diversity, we counted the number of unique words (types) parents said to infants. Parents' speech was also analyzed with two measures of syntactic complexity. We determined the mean length of utterances in words (MLUw) (Parker & Brorson, 2005) and the proportion of utterances which contained only a single word.

Next, we examined the relations between prelinguistic vocalization quality and contingent parental speech. We used Oller's infraphonological framework in our classification of 9- and 10-month-old infants' vocalizations to quantify the maturity of infant vocalizations. Over the first year of life, infants' vocal production undergoes extensive changes (Oller, 2000). Infants between 1 and 4 months begin to produce immature vowel sounds characterized by minimal breath support and a closed vocal tract, which result in creaky and nasalized vocalizations called quasi-resonant vowels. Incrementally, between 3 and 6 months, infants begin to produce fully resonant vowels which have full breath support and are produced with an open vocal tract. By 9 months, infants begin to produce mature consonant–vowel alternations which resemble the sounds found in well-formed language (e.g., [ba], [da]; Oller, Eilers, & Basinger, 2001). Infants' prelinguistic vocal repertoires are most diverse between 9 and 10 months, a time when they are beginning to incorporate phonological characteristics of their ambient language into their vocal productions.

Parental speech is sensitive to these developmental changes in infants' vocalizations. Parents are more likely to respond to vocalizations which more closely resemble adult speech, such as consonant–vowel vocalizations (Abney, Warlaumont, Oller, Wallot, & Kello, 2016; Albert *et al.*, 2017). Parents can also identify when infants' babbling matures. Parents' identification of qualitative variance in infant's vocalizations has been documented (Oller, Eilers, & Basinger, 2001). In addition to responding differently based on the acoustic maturity (e.g., vowel resonance, consonant–vowel timing) of vocalizations, parents respond to the directedness of infants' vocalizations (Albert *et al.*, 2017). In response to infants' object-directed vocalizations, parents

responded more frequently, were more likely to respond sensitively to infants' current attentional focus, and more often provided information content about infants' current attentional focus. However, Albert *et al.* investigated only the content of parents' contingent speech to infant babbling. By conducting acoustic and linguistic analyses of both contingent and non-contingent speech to infants, we can better understand the role that babbling may play in influencing social interactions in ways that might facilitate the infant's own learning.

Methods

Participants

Thirty mother–infant dyads participated (infant mean age = 9.20 months, range: 9.4–10.14 months). Participants were recruited from birth announcements in local newspapers and through advertisements. Mothers received a T-shirt or a book as a gift for their participation. Participants were part of a larger corpus from a previously published study (Goldstein & Schwade, 2008).

Apparatus

Sessions were recorded in a naturalistic environment (a 12 ft. × 18 ft. playroom) with toys, a toy box, and posters of animals. Infants were free to roam around the room and explore. Interactions were video-recorded via three remote-controlled digital video cameras. To obtain detailed audio-recordings, each infant wore denim overalls concealing a wireless microphone (Telex FLM-22; Telex Communications, Inc., Burnsville, MN) and transmitter (Telex USR-100). Caregivers wore a wireless lapel microphone (Telex FMR-150) with a transmitter concealed in a pouch at their waist (Telex USR-100). Infant vocalizations and caregiver speech were recorded on distinct audio channels.

Procedure

Participants came to the lab for two 30-min play sessions, spaced approximately 24 hours apart. The first session and the first 10 min of the second session were unstructured, with parents instructed to play as they normally would at home.

Speech transcription

Parents' speech during session 1 was transcribed in full. Caregiver utterances were segmented if they were bounded by silence longer than 2 sec and/or if they exhibited terminal pitch contours (Venker, Bolt, Meyer, Sindberg, Weismer, & Tager-Flusberg, 2015). Utterances from the parents were categorized as contingent if they occurred within 2 sec of the offset of infants' vocalizations (including all vocalization categories as described by Oller, 2000, and defined below); all other parent utterances were categorized as non-contingent (Table 1). Caregiver responses to crying, fussing, and vegetative vocalizations (e.g., coughs) were excluded from our analysis.

Child-directed speech analyses

We computed three descriptive measures for caregivers' contingent and non-contingent speech during session 1: (1) the number of unique words in caregivers' speech, (2) the

Table 1. Caregiver speech descriptive statistics

Measure		Mean	SD
Mean utterances per caregiver			
	Contingent	43.30	24.78
	Non-contingent	106.80	40.39
	Total	150.10	44.23
Proportion of utterances			
	Contingent	0.29	0.15
	Non-contingent	0.71	0.15
Total utterances			
	Contingent	1,299	
	Non-contingent	3,204	
	Total	4,503	

mean length of caregivers' utterances in words (MLUw), and (3) the proportion of single-word utterances in caregivers' speech. To measure inter-rater reliability, a second coder transcribed 5 min of the sessions for a subset (20%, N = 23) of the dyads. The intra-class correlation (ICC) for the number of unique contingent words was good (.85) as was reliability for the number of non-contingent unique words (.79). The ICC for the total number of words spoken contingently was good (.76) and moderate for the total number of non-contingent words (.71). The ICC for the mean length of utterances was good for both contingent utterances (.87) and non-contingent utterances (.79).

To obtain acoustic parameters of parents' speech, we utilized a Praat (Boersma & Weenink, 2015) script that extracted a vector corresponding to the F_0 (fundamental frequency) over time from a subset (15%) of the total utterances in our dataset. Utterances used for analysis of acoustic differences across contingent and non-contingent speech were duration matched.

Phonology of infants' vocalizations

Infant vocalizations from the unstructured play period of session 2 were categorized into four different groups according to Oller's infraphonological acoustic classification system (Oller, 2000). This system considers both acoustic (e.g., timing of formant transitions) and qualitative (e.g., phonetic categories) features of infants' vocalizations. Vocalization boundaries were segmented according to breath groups (see Oller & Lynch, 1992; Oller, 2000). A QUASI-RESONANT NUCLEUS is produced with a relatively closed throat, has little breath support and is qualitatively creaky, nasal, or both. A FULLY RESONANT NUCLEUS is produced with an open vocal tract with normal phonation which yields a clear formant structure. A MARGINAL SYLLABLE is a slow movement between consonant and vowel (> 250 ms). The slow movement of the articulators often distorts the vowel. Following Oller's (2000) classification system, we included as consonants stops and fricatives, and excluded glides. Glides were included in the vowel categories (Stoel-Gammon 1989). A CANONICAL SYLLABLE has a fully resonant vowel and a quick (< 250 ms) transition between

consonant and vowel. We tested 9- to 10-month-olds because they typically have all four vocalization types in their repertoires. Infant vocalizations which contain consonant sounds are considered to be more developmentally advanced, as they are more speech-like. Reliability for infant vocalizations was calculated based on independent coding of 20% of the sample (ICC = .92).

We tallied infant vocalizations from session 2, after infants had become accustomed to the lab play room during session 1. We calculated the proportion of infant's vocalizations with CV structure to the total number of vocalizations to assess the maturity of infants' vocalizations. We included marginal and canonical syllables in our measure of CV vocalizations because over the course of the first year, the proportion of infant vocalizations containing consonants increases (Holmgren, Lindblom, Aurelius, Jalling, & Zetterström, 1986). We analyzed infant and caregiver speech from separate sessions to maximize our sampled data, as caregivers generated more utterances during session 1 and infants vocalized more during session 2. Analyzing caregiver and infant utterances from separate sessions also provided caregiver and infant variables that were more independent from each other.

Results

Caregiver speech: linguistic comparisons

Parents produced less contingent than non-contingent speech, with significantly fewer unique words spoken as part of contingent utterances (Figure 1) ($t(29) = -7.92$, $p < .001$, $d = -1.71$). A binomial test showed that a significant number of caregivers (n = 28 of 30) uttered fewer unique words contingently on infant vocalizations than non-contingently ($z = -4.52$, $p < .001$).

To compare syntactic complexity across speech types, caregivers' MLUw was calculated for contingent and non-contingent utterances (Figure 2). Caregivers had significantly shorter contingent utterances than non-contingent utterances ($t(29) = -4.76$, $p < .001$, $d = -0.75$). A binomial test showed that a significant number of caregivers (n = 24) showed this pattern ($z = 3.97$, $p < .001$). Because tests of individual subject means violate non-independence assumptions, we used R (R Core Team, 2018) and R package *lme4* (Bates, Maechler, Bolker, & Walker, 2015) to perform a linear mixed effects analysis of the relationship between MLUw and speech contingency. We entered contingency of speech (without an interaction term) into the model as a fixed effect, and subject intercepts and by-subject random slopes for the effect of speech contingency as random effects. Visual inspection of qq plots of residuals did not reveal obvious deviations from homoscedasticity. *P*-values for the mixed effects model were obtained by a likelihood ratio test of the full model against a null model. The model estimated an average change in MLUw of –1.07 words from non-contingent to contingent speech ($p < .001$).

To further test syntactic complexity, the proportion of utterances which contained a single word were calculated for contingent and non-contingent utterances (Figure 3). A significantly higher proportion of contingent than non-contingent utterances were a single word ($t(29) = 5.45$, $p < .001$, $d = 1.20$). A binomial test showed that a significant number of caregivers (n = 26) showed this pattern ($z = 4.47$, $p < .001$).

All caregiver speech and language measures were significantly inter-correlated (Table 2). As expected, diversity of words was positively correlated with utterance length. Conversely, proportion of single-word utterances was negatively correlated with both utterance length and diversity of words.
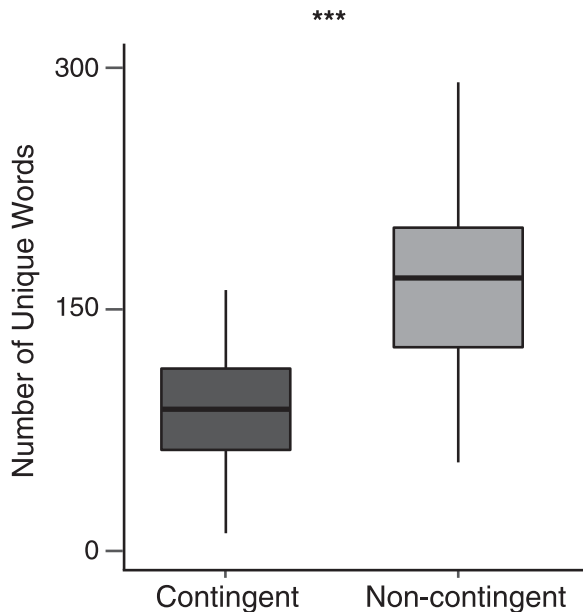
Figure 1. Boxplot of the number of unique words parents uttered as part of contingent and non-contingent utterances. *** *p* < .001.

Caregiver speech: acoustic comparisons

Analysis of caregiver fundamental frequency [F_0] during contingent and non-contingent speech reveals no significant differences between speech types. The mean F_0 for contingent speech (M = 275.54 Hz, SD = 58.26) was not significantly different than for non-contingent speech (M = 286.26 Hz, SD = 43.62) ($t(16) = 2.11$, $p = .42$). The F_0 range for contingent speech (M = 192.42 Hz, SD = 61.12) was likewise not significantly different than for non-contingent speech (M = 210.19 Hz, SD = 57.64) ($t(16) = 1.74$, $p = .31$). The mean and range for F_0 for both contingent and non-contingent speech are consistent with previously descriptions of naturally produced IDS (e.g., Fernald & Simon, 1984). To assess effects of individual differences in mean pitch on our analyses on pitch, we used mixed effects modeling as in the above MLUw analysis. The model estimated an average change of –12 Hz when changing from non-contingent to contingent speech ($p = .14$).

Infant vocalizations and caregiver speech

We analyzed the relation between parents' lexical diversity and infant vocal maturity (Figure 4). The number of unique words in contingent parental speech predicted vocal maturity in infants ($r(28) = .40$, $p = .02$). The number of unique words in non-contingent speech did not significantly predict vocal maturity ($r(28) = .04$, $p = .80$).

Discussion

We found that parents simplified the statistical and syntactic structure of their speech in response to babbling. Contingent speech contained fewer unique words and contained both shorter utterances and more single-word utterances. In combination, these

***
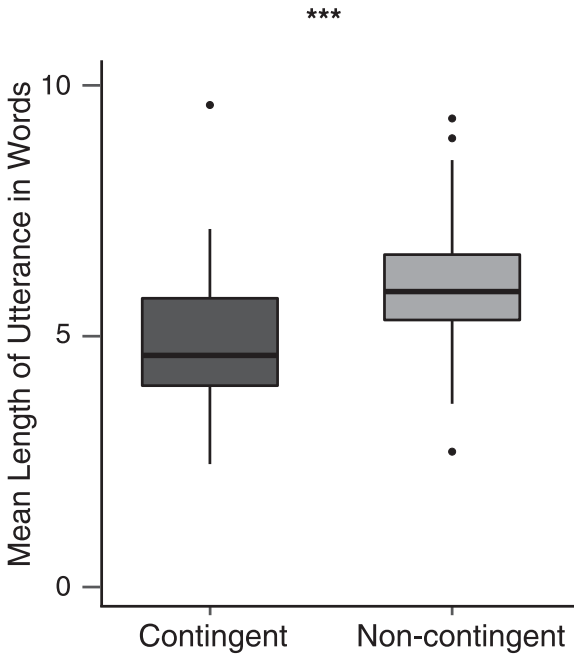
Figure 2. Boxplot of the mean number of words per contingent and non-contingent parent utterances. *** *p* < .001.

characteristics of parents' contingent speech suggest a new form of influence of infants' prelinguistic vocalizing on the ambient linguistic environment. Infants' immature vocalizations may create language learning opportunities by eliciting responses from parents that contain simplified, more learnable information.

Linguistic simplification of parents' contingent speech may provide particular benefits for infant learning, because infants often babble at times of focused attention and heightened arousal (Goldstein *et al.*, 2010b). Infants more accurately remember the features of objects to which they had babbled, as compared to objects that received similar looking and handling but no babbling (Goldstein *et al.*, 2010a). In addition, studies of infant vocal learning have shown that infants rapidly learn patterns of parental speech when they are uttered contingently on babbling (Goldstein & Schwade, 2008). When parents change their speech statistics in response to infant vocalizations, they provide more learnable input to infants at a time that infants seem best able to learn from it.

Parents simplified their speech in two characteristic ways that could facilitate language learning. First, contingent parental speech contained fewer unique words than non-contingent speech. Providing a narrower distribution of words in contingent speech might serve as simplified input that is more tuned to infants' developmental level. Second, contingent parental speech has a higher proportion of single-word utterances and shorter utterances than non-contingent speech.

The higher proportion of single-word utterances in contingent speech may benefit infants. Single-word utterances simplify the task of finding word boundaries, facilitating statistical learning (Lew-Williams, Pelucchi, & Saffran, 2011). Parents' production of single words predicted children's later production of those words (Brent & Siskind, 2001). Parents also used shorter utterances in contingent than
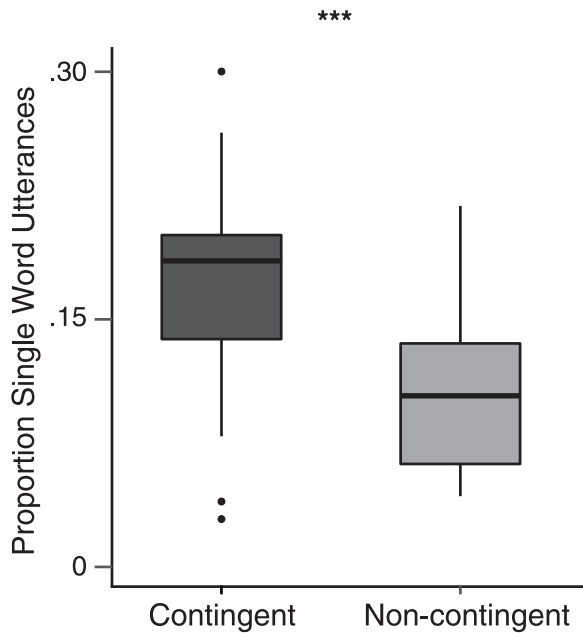
Figure 3. Boxplot of the proportion of contingent and non-contingent parent utterances that were a single word in length. *** *p* < .001.

Table 2. Correlations between linguistic measures across contingent and non-contingent utterances. * *p* < .05, ** *p* < .01, *** *p* < .001.

Measures	Contingent	Non-contingent
Lexical diversity × proportion of single-word utterances	−0.36*	−0.56***
Lexical diversity × Mean Length of Utterance in Words	0.49**	0.61***
Proportion of single-word utterances × Mean Length of Utterance in Words	−0.58***	−0.78***

non-contingent speech. Previous studies have shown that parents may make modifications to their utterance complexity over the course of infant language development that are tied to their infants' learning progress. Evidence from large-scale recordings of a single family suggest that caregivers' MLUw may decrease until children begin to combine words around 16 months (Roy *et al.*, 2009). These changes in parent MLUw may correspond to a shift in children's language comprehension. This interpretation is consistent with previous hypotheses of parents adjusting their speech to their infants' developmental level (Snow, 1995). The fine-tuning hypothesis (Snow, 1995) suggests that adults adapt the complexity of their speech to infants and children in response to properties of their immature speech and language. Taken together, these speech simplifications indicate that infants' immature vocalizing serves to elicit more learnable speech from their caregiver. Because parents are sensitive to the maturity of infants' vocalizations, future research will compare differences of parent responses to infants' cries and non-cry vocalizations and responsiveness to precanonical and canonical vocalizations.

(a)

(b)
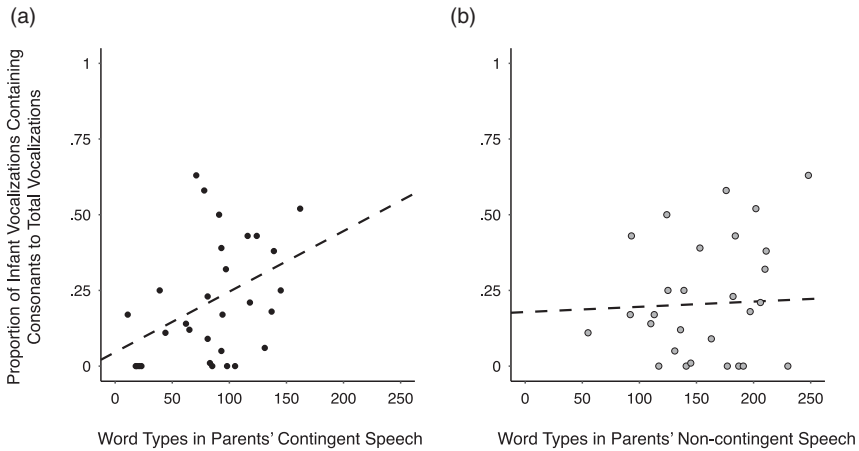
Figure 4. *a.* Parents who used more lexically diverse contingent speech had infants with more advanced vocalizations ($r(28) = .40$, $p = .02$). The best outcomes for infant vocal development were associated with parents producing approximately 75–150 word types. *b.* The correlation between lexical diversity in non-contingent speech and infant vocalization quality was not significant ($r(28) = .04$, $p = .80$).

In contrast to these linguistic changes, the prosodic features (e.g., pitch contour) of parental speech did not differ significantly across contingent and non-contingent utterances. These findings are consistent with past work demonstrating the ubiquity of pitch changes in infant-directed speech (Fernald & Morikawa, 1993). Stability of prosodic features may highlight changes in the underlying linguistic structure of parents' contingent utterances that may be relevant for speech and language development.

We next examined relations between parents' speech and their infants' vocal development. We found that the lexical diversity of contingent parental speech predicted infants' vocal maturity. In contrast, the lexical diversity of non-contingent speech did not predict vocal development. More studies are required to ascertain causality, but there are several reasons why the structure of contingent parental speech may influence infant vocal learning. Although more lexically diverse contingent speech was associated with more advanced infant vocal productions, the amount of lexical diversity in contingent speech was generally less than was present in non-contingent speech. Recent research has established that infants seem to seek information streams characterized by intermediate levels of complexity (e.g., Kidd & Hayden, 2015). Such information streams might be optimal for learning, as they are neither overly simple nor insurmountably complex. Infant learning might thus be optimized when parents provide an intermediate level of variability in their speech. Other research has established the crucial contribution of variability to learning. For example, infants' construction of cross-modal associations is facilitated by variability in exemplars of category membership (Vukatana *et al.*, 2015). Infants exposed to single exemplars of a novel animal were unable to learn pairings of the novel animal with a novel animal sound. In contrast, when infants were exposed to multiple exemplars of the animal category during training, they both learned the animal–sound pairing and generalized the animal–sound pairing to new category exemplars.

In our view, intermediate levels of variability in parents' words could direct infants' attention to shared features of the words (e.g., mature consonant–vowel alternations) and away from features irrelevant to the learning task (e.g., pitch fluctuation). A longitudinal study of children aged one to four years found that caregivers' production of a larger variety of words over that time period was positively related to their children's vocabulary (Huttenlocher *et al.*, 2010). To investigate longitudinal changes in how parents respond to infants' vocalizations, we are currently investigating parents' speech in response to vocalizations of infants younger and older than those in the present study. Partial variation in language content over successive utterances may perceptually highlight changes in language content and make them easier to learn (Goldstein *et al.*, 2010b). The presence of intermediate levels of variation in contingent feedback from parents should make the structure of language more salient to the learner. To test the causal influence of exemplar variability on infant vocal learning, current studies in our lab are experimentally manipulating the level of speech variability infants hear and measuring subsequent in-the-moment changes in infant vocal production. If manipulations of the distributional properties of parents' speech lead to changes in infants' vocalizations, the findings would indicate that parental speech variability plays a causal role in infant vocal learning. Past findings regarding intermediate complexity and infant attention come from non-social stimuli (Kidd, Piantadosi, & Aslin, 2014; Kidd & Hayden, 2015). However, if, as posited, the effect of complexity on learning is a general principle of learning, then it should also hold for other contexts. Future analyses will also consider whether CV vocalizations (compared to vowel-only vocalizations) differentially influence parents' contingent speech, as infant vocalization type influences other forms of parental responses (Albert *et al.*, 2017).

In addition to the structural modifications in parents' contingent speech, the contingent timing of parents' speech on infant vocalizations may facilitate learning for multiple reasons. First, as found in previous studies of social influences on prelinguistic vocal learning, infants learn new phonological patterns and word–object associations better when information is presented contingently on babbling (Goldstein & Schwade, 2008; Goldstein *et al.*, 2010a). Second, infants may find contingent responses THEMSELVES rewarding, which may facilitate learning. The affiliative bonds present in parent–infant relationships may have deep connections with reward and memory formation (Depue & Morrone-Strupinsky, 2005). Reward might also play a role in combination with activation of social motivation regions of the brain during learning (Syal & Finlay, 2010) which are highly interconnected with the motor circuitry of vocalizing (Theofanopoulou, Boeckx, & Jarvis, 2017). In addition, learning itself is hypothesized to be rewarding to infants (Kidd & Hayden, 2015). Thus, social interaction with mature vocalizers which provides information not available in the absence of social interaction is a primary contender for facilitating learning. Third, infants have limited memory and may need to capitalize on linguistic content within short time-windows when their attention is heightened and they are expecting a response from their parents (Kareev, 1995; Goldstein & Schwade, 2010).

In our view, immature vocalizations create learning opportunities by eliciting social responses that contain simplified, learnable information. These findings have important implications for current large-scale data collection and intervention studies on language development. Sophisticated and useful data on changes in linguistic structure as a function of contingent timing can be gleaned from home recording efforts that are

currently focused on turn-taking and other forms of parent–infant interaction (Romeo *et al.*, 2018). Several interventions for at-risk families currently focus on increasing the number of words parents say (e.g., Providence Talks; <http://www.providencetalks. org>) or turn-taking interactions with infants (Leffel & Suskind, 2013) but have not, to the best of our knowledge, focused on the relevance of learning distributional and temporal properties of parents' speech to infants.

Our results suggest that infants, via their immature vocalizing, play an important role in shaping their own language environment. Infants bring curiosity about their environment to the continuous series of new situations they are exposed to (Moulin-Frier, Nguyen, & Oudeyer, 2014; Kidd & Hayden, 2015). Accurate prediction of environmental changes, an underlying learning mechanism in computational models of vocal learning, may also support infant learning in social contexts (Moulin-Frier *et al.*, 2014). Such models choose to learn from data over which they can minimize the error of their own predictions at the highest rate. By vocalizing, infants have the opportunity to observe the effects of their vocalizations on parents. Over their first year, infants quickly come to associate their immature vocalizations with responses from their parents (Goldstein *et al.*, 2009). Eliciting mature speech sounds from caregivers may become the target of infants' curiosity and subsequent guidance of vocal development. Parents infant-directed behaviors are multimodal in nature. Here we report the first indication (to our knowledge) that parents' speech is sensitive to infant vocal behavior in real time; however, future research could investigate whether parents simplify their behaviors in other modalities. For more advanced understanding of early infant learning, future experimental, large-scale observational, and computational research should incorporate the affects infants have on the temporal and distributional properties of parents' speech.

Author ORCIDs. 🆔 Michael H. Goldstein, 0000-0001-6672-3752.

References

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2016). Multiple coordination patterns in infant and adult vocalizations. *Infancy*, 22(4), 514–39.

Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2017). The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science*, 18, e12641.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Boersma, P., & Weenink, D. (2015). *Praat*: Doing phonetics by computer. retrieved from <http://www. praat.org>.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–B44.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27(6), 843–73.

Depue, R. A., & Morrone-Strupinsky, J. V. (2005). A neurobehavioral model of affiliative bonding: implications for conceptualizing a human trait of affiliation. *Behavioral and Brain Sciences*, 28, 313–50.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variation in Japanese and American mothers' speech to infants. *Child Development*, 64, 637–56.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104–13.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Science*, *100*(13), 8030–5.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–23.

Goldstein, M. H., & Schwade, J. (2010) From birds to words: perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford handbook of developmental behavioral neuroscience* (pp. 708–29). Oxford University Press.

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), 636–44.

Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010a). Learning while babbling: prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy*, *15*(4), 362–91.

Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010b). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences*, *14*(6), 249–58.

Gultekin, Y. B., & Hage, S. R. (2018). Limiting parental interaction during vocal development affects acoustic call structure in marmoset monkeys. *Science Advances*, *4*, eaar4012.

Holmgren, K., Lindblom, B., Aurelius, G., Jalling, B., & Zetterström, R. (1986). On the phonetics of infant vocalization. In B. Lindblom & R. Zetterström (Eds.), *Precursors of early speech* (pp. 51–63) (Wenner-Gren Center International Symposium Series). London: Palgrave Macmillan.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–65.

Kareev, Y. (1995) Through a narrow window: working memory capacity and the detection of covariation. *Cognition*, *56*, 263–9.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*(3), 449–60.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*, 1795–804.

King, A. P., West, M. J., & Goldstein, M. H. (2005). Non-vocal shaping of avian song development: parallels to human speech development. *Ethology*, *111*(1), 101–17.

Kuhl, P. K., Tsao, F.-M., & Lui, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Science*, *100*, 9096–101.

Leffel, K., & Suskind, D. (2013). Parent-directed approaches to enrich the early language environments of children living in poverty. *Seminars in Speech and Language*, *34*, 267–78.

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, *14*, 1323–9.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465–94.

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*, 1–20. doi:10.3389/fpsyg.2013.01006

Newman, R. S., Rowe, M. L., & Ratner, N. B. (2016). Input and uptake at 7 months predicts toddler vocabulary: the role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*, 1158–73.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to Children: language input and acquisition* (pp. 109–49). Cambridge University Press.

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

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

Oller, D. K., & Lynch, M. P. (1992). Infant vocalizations and innovations in infraphonology: toward a broader theory of development and disorders. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications* (pp. 509–536). Timonium, MD: York Press.

Parker, M. D., & Brorson, K. (2005). A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). *First Language*, 25(3), 365–76.

R Core Team (2018). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online <https://www.R-project.org/>.

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2017). The impact of early social interactions on later language development in Spanish–English bilingual infants. *Child Development*, 88(4), 1216–34.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: children's conversational exposure is associated with language-related brain function. *Psychological Science*, 29(5), 700–10. Online <https://dspace.mit.edu/handle/1721.1/66701?show=full>.

Roy, B. C., Frank, M. C., & Roy, D. K. (2009). Exploring word learning in a high-density longitudinal corpus. *Proceedings of Cognitive Science Society*, 1–7.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–8.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word–referent learning. *Trends in Cognitive Sciences*, 18(5), 251–8.

Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13, 410–20.

Snow, C. E. (1977). Mothers' speech research: from input to interaction In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: language input and acquisition* (pp. 31–49). Cambridge University Press.

Snow, C. E. (1995). Issues in the study of input: finetuning, universality, individual and developmental differences, and necessary causes. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 180–93). Oxford: Blackwell.

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: infant age and context related changes. *Journal of Child Language*, 10, 1–15.

Stoel-Gammon, C. (1989). Prespeech and early speech development of two late talkers. *First Language*, 9, 207–24.

Syal, S., & Finlay, B. L. (2010). Thinking outside the cortex: social motivation in the evolution and development of language. *Developmental Science*, 14(2), 417–30.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, 349(6249), 734–8.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones, *Child Development*, 72, 748–67.

Theofanopoulou, C., Boeckx, C., & Jarvis, E. D. (2017), A hypothesis on a role of oxytocin in the social mechanisms of speech and vocal learning. *Proceedings of the Royal Society B: Biological Sciences*, 284 (1861). doi:10.1098/rspb.2017.0988.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53–71.

Venker, C. E., Bolt, D. M., Meyer, A., Sindberg, H., Weismer, S. E., & Tager-Flusberg, H. (2015). Parent telegraphic speech use and spoken language in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research*, 58(6), 1733–46.

Vukatana, E., Graham, S. A., Curtin, S., & Zepeda, M. S. (2015). One is not enough: multiple exemplars facilitate infants' generalizations of novel properties. *Infancy*, 20(5), 548–75.

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, 25(7), 1314–24.

Weisberg, D. S., Zosh, J. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Talking it up: play, language, and the role of adult support. *American Journal of Play*, 6(1), 39–54.