Received: 30 May 2016 | Accepted: 16 October 2017
DOI: 10.1111/desc.12641

PAPER

WILEY Developmental Science

The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness

Rachel R. Albert[1] | Jennifer A. Schwade[2] | Michael H. Goldstein[2]

[1]Department of Psychology, Lebanon Valley College, Annville, PA, USA

[2]Department of Psychology, Cornell University, Ithaca, NY, USA

Correspondence
Rachel R. Albert, Psychology Department, 101 N. College Avenue, Lebanon Valley College, Annville, PA 17003, USA.
Email: ralbert@lvc.edu

Funding information
NSF (0844015) and NICHD (5R03HD61524-2)

Abstract

What is the social function of babbling? An important function of prelinguistic vocalizing may be to elicit parental behavior in ways that facilitate the infant's own learning about speech and language. Infants use parental feedback to their babbling to learn new vocal forms, but the microstructure of parental responses to babbling has not been studied. To enable precise manipulation of the proximal infant cues that may influence maternal behavior, we used a playback paradigm to assess mothers' responsiveness to prerecorded audiovisual clips of unfamiliar infants' noncry prelinguistic vocalizations and actions. Acoustic characteristics and directedness of vocalizations were manipulated to test their efficacy in structuring social interactions. We also compared maternal responsiveness in the playback paradigm and in free play with their own infants. Maternal patterns of reactions to babbling were stable across both tasks. In the playback task, we found specific vocal cues, such as the degree of resonance and the transition timing of consonant-vowel syllables, predicted contingent maternal responding. Vocalizations directed at objects also facilitated increased responsiveness. The responses mothers exhibited, such as sensitive speech and vocal imitation, are known to facilitate vocal learning and development. Infants, by influencing the behavior of their caregivers with their babbling, create social interactions that facilitate their own communicative development.

RESEARCH HIGHLIGHTS

- Babbling contains acoustic features that have strong salience to mothers and organize mothers' responsiveness.
- Cues in prelinguistic vocalizations, such as vocal maturity and directedness, influence mothers' contingent, sensitive reactions and imitation of infant sounds.
- Given the stability and robustness of mothers' responses to babbling, babbling likely creates predictable social reactions that provide opportunities for infant learning.

1 | INTRODUCTION

Parents play a strong role in the development of communication. Most research on the role of parenting in communicative development examines the long-term effects of responsiveness to infant behavior on later language outcomes (e.g., Baumwell, Tamis-LeMonda, & Bornstein, 1997; Hart & Risley, 1995). For example, caregivers' responses to infants' prelinguistic noncry vocalizations (babbling) predict later receptive and productive vocabulary (Tamis-LeMonda, Bornstein, & Baumwell, 2001). Different forms of caregiver behavior have different impacts on infant language. Caregivers' sensitive responses such as providing object labels (Stevens, Blake, Vitale & MacDonald, 1998) or asking questions (Furrow, Nelson, & Benedict, 1979; Gleitman, Newport, & Gleitman, 1984) have positive relations with later language development. In contrast, redirective responses predict poorer vocabulary outcomes (Akhtar, Dunham, & Dunham, 1991; Tomasello & Farrar, 1986). The results of these studies clearly identify components of parenting behavior that predict infant language learning.

However, the specifics of *how* parenting matters for later language development are largely unknown because the ways in which the form

 © 2017 John Wiley & Sons Ltd

and timing of parental behavior impact infant learning mechanisms has received relatively little attention. Recent evidence indicates that infant vocal learning is embedded in a social feedback loop (Goldstein & Schwade, 2010; Warlaumont, Richards, Gilkerson, & Oller, 2014). Infants monitor the immediate social effects of their behaviors, while parental responses influence learning. Real-time manipulations of parental responsiveness demonstrate that infants rapidly learn from adult behavior that is contingent on their vocalizations (e.g., Goldstein & Schwade, 2008; Gros-Louis, West, & King, 2014). Caregivers' reactions to infant vocalizations are thus an important source of information for language learning. Mothers provide a wide range of contingent responses to infant vocalizations such as labeling objects, imitating, affirming, describing the infant's behaviors, asking questions, or redirecting the infant's attention (e.g., Paavola, Kunnari, Moilanen, & Lehtihalmes, 2005; Tamis-LeMonda et al., 2001). Infants modify their vocalizations in response to caregivers' contingent feedback to be more speech-like and incorporate new phonological patterns (Goldstein, King, & West, 2003; Goldstein & Schwade, 2008). Infants also show facilitated learning of word-object associations when, after babbling at an object (object-directed vocalizations: ODVs), adults immediately label that object (Goldstein, Schwade, Briesch, & Syal, 2010). Further, recent work using computational modeling has revealed evidence of neural mechanisms (e.g., spike-timing dependent plasticity) that could underlie changes in vocalizations as a result of contingent feedback (e.g., Takahashi et al., 2015; Takahaski, Liao, & Ghanzafar, 2017; Warlaumont & Finnegan, 2016). Not all forms of caregiver responsiveness, however, have a positive effect on infant learning. Labeling absent objects in response to 9-month-olds' ODVs is negatively correlated with later vocabulary (Goldstein & Schwade, 2010).

Recent work that closely measured the microstructure of parent–infant interactions has revealed forms of parental behavior that guide infant attention and facilitate learning in real-time. For example, parents coordinate infant attention during triadic interactions by manipulating and discussing objects. During these interactions, infants rarely fixate on mothers' faces (Yu & Smith, 2015), but preferentially attend to mothers' manipulations of objects (Deák, Krasno, Triesch, Lewis, & Sepeta, 2014). When infants do reference their mother's face, her gaze is typically fixated on the held object, which redirects the infant's gaze back to the hands (Deák et al., 2014). Through joint interactions, infants learn to interpret their mother's signals and react in ways that further promote aligned responses (Yu & Smith, 2015). These findings suggest that infants emit reliable signals of their attentional focus via gaze, which mothers could use to organize their responses and facilitate infant learning.

In addition to gaze, infant vocalizations are another reliable cue that attracts attention and elicits regular responses from mothers. How does babbling organize the real-time structure of parental behavior? To assess the influence of specific forms of infant sounds and actions, we designed a playback paradigm, a method widely used in studies of animal communication (e.g., Illman, Neuhauserova, Pokorna, Chaloupkova, & Simeckova, 2008; Shizawa, Nakamichi, Hinobayashi, & Minami, 2005; Smith, King, & West, 2000). We recorded and recombined infant vocalizations and actions to determine their specific effects on caregiver responses. Stimulus infants were unfamiliar to participants so that we could directly compare responsiveness across individuals. Stimuli were representative of the wide range of vocalizations that infants produce at 9 months of age (see Methods for details), that vary in the degree to which they incorporate mature speech acoustics. We also manipulated infants' apparent attentional focus. In our stimuli, vocalizations were either directed at objects or were undirected, with no clear focal point of attention. We compared mothers' responses to changes in speech quality and attentional focus to determine how each contributed to parental behavior.

Playback studies in which infant or child behavior has been presented to parents or adults have generally found a high degree of consistency in their responses. Computer simulations of parenting scenarios have been used to study parental cognitions, finding that parents' choices of reactions to presented scenarios accurately reflect their real-world behavior (Holden, 1985; Holden, Ritchie, & Coleman, 1992). Several studies found reliable patterns of responsiveness to prerecorded fusses and cries (e.g., Gustafson & Green, 1989; Wood & Gustafson, 2001; Zeskind, Klein, & Marshall, 1992). Some studies alternated live over video interaction with playbacks of prerecorded video to disrupt interaction timing, finding that mothers' infant-directed speech was reduced during and immediately after noncontingent interaction (e.g., Braarud & Stormark, 2008).

While a playback paradigm allows for precise control over the presentation of infant behavior, the responses of adults to the stimuli will necessarily be different in form from the behaviors they use in live interaction. The ways in which responses are assessed have varied across studies. However, most studies that examined caregivers' reactions to prerecorded babbling had adults rate infant characteristics, such as happiness, attractiveness, or communicative intent, rather than specify their own behavioral responses to the vocalizations (Bloom, D'Odorico, & Beaumont, 1993; Bloom & Masataka, 1996; Degotardi & Sweller, 2012; Goldstein & West, 1999; Papoušek, 1989).

In contrast to previous work, we asked participants to provide in-the-moment responses indicating their reactions to stimulus infants whose speech and gaze direction were systematically manipulated. Mothers imagined they were interacting with the infants and provided vocal reactions to each stimulus by talking to the infants on the display. It is not understood whether parents systematically react to babbling in ways that are a good fit with the social learning capacities that previous studies (discussed above) have revealed. Quantifying social responsiveness with measures that are based on infant learning capacities is critically important for connecting adult behavior to vocal learning and development.

We assessed the efficacy of specific acoustic features of babbling (e.g., vowel resonance and consonant-vowel transition timing), as well as the contextual cue of directedness to objects, in facilitating parental reactions. We predicted that mothers would respond more frequently to vocalizations that had mature speech characteristics or were directed at objects. Developmentally advanced prelinguistic vocalizations tend to receive interpretations of “wanting” from caregivers (Goldstein & West, 1999). If mothers interpret infant vocalizations as meaningful, then object-directed vocalizations may motivate them to provide labeling responses. We also assessed the validity of the playback paradigm as a measure of caregiver behavior by comparing mothers’ reactions to prerecorded stimuli to those given to their own infants during play. In addition to the real-time measures of responsiveness, we assessed mothers’ perceptions of the maturity of infant vocalizations by asking them to rate each vocalization on a 7-point Likert scale for how speech-like it sounded. We expected that mothers would perceive differences in acoustic quality, as previous research indicates that parents are able to identify mature syllables without training (Oller, Eilers, & Basinger, 2001).

2 | METHODS

2.1 | Participants

Forty mothers (*n* = 20 mothers of a single child (*M* = 31.80 years, range: 21–45 years); *n* = 20 mothers of multiple children (*M* = 34.73 years; range: 27–44 years; one mother did not provide her age)) participated with their 9-month-old infants (20 male; mean age 9 months 24 days; range 8 months 21 days–10 months 15 days). Participants were recruited from birth announcements printed in local newspapers. Sixteen additional dyads were tested but excluded because infants cried or fussed excessively (*n* = 7), caregivers did not follow directions (*n* = 4) or speak in English (*n* = 2), or equipment malfunctioned (*n* = 3). Mothers received an infant T-shirt or book.

2.2 | Apparatus

The session took place in a large (3.65 m × 4.57 m) playroom containing toys and a toy box. During the live play task, dyads were videorecorded via three remote-controlled cameras mounted on the walls. To obtain accurate and detailed recordings of infant vocalizations, infants wore a wireless microphone concealed inside the lining of adjustable denim overalls. To obtain accurate recordings of caregiver speech, mothers wore a wireless lapel microphone and transmitter. For the playback task, mothers sat in front of a 19″ LCD computer monitor and a digitizing tablet. They tapped a stylus on the tablet to advance to the next stimulus clip. Stimuli were presented in randomized order by Adobe Director 11.5 (Adobe, 2009). Vocalizations were played over headphones (Sony MDR-7506).

2.3 | Stimuli

Digital audio/video examples of infant behavior were obtained from previously recorded parent–infant play sessions. Stimuli were obtained from 20 9-month-old infants (half female) and included a wide range of prelinguistic vocal forms. Prelinguistic vocalizations can be categorized by their acoustic properties using an infraphonological coding system (Oller, 2000). Infraphonology captures the major changes in vocal production over the first year with four major syllable types. Quasi-resonant vowels (QR) are characterized as creaky or nasalized vocalizations, resulting from minimal breath support. Fully-resonant vowels (FR) are produced when the vocal tract is open, resulting in normal phonation. Marginal syllables (MS) are slow sequences of consonant–vowel articulation, with long transitions (> 200 ms) between the consonant and vowel. Canonical syllables (CS) have mature consonant–vowel transitions and incorporate fully-resonant vowels (e.g., [ba], [da]; Oller, 2000). By 9 months, most infants regularly produce all four infraphonological types (Oller, Eilers, Neal, & Schwartz, 1999). Stimuli also encompassed a range of vocal directedness. Object-directed vocalizations (ODVs) are produced while the infant is looking at an object that is held or within reach. Undirected vocalizations (UDVs) are produced when the infant is not looking at an object or a caregiver.

Each stimulus infant contributed two vocalizations, matched either on vowel resonance (QR or FR) or consonant–vowel transition timing (MS or CS). Each vocalization pair (QR/FR or MS/CS) was matched to two video clips of the same infant presented in an object-directed context and in an undirected context. Thus, each infant provided four clips. The object-directed vocalization videos showed the infant looking at an object and the undirected vocalization videos showed the infant looking off screen to the right or left into empty space. Infants in the undirected vocalization videos were next to or holding objects, but were not looking at them. The two vocalizations were each paired with both video clips to create four stimulus clips for each infant. Each clip lasted 7 s with the vocalization occurring 2.5 s into the clip. Recombining audio and video could create visible mismatches between articulatory movements and sounds that could influence participants’ stimulus perception (McGurk & MacDonald, 1976). Thus, each video clip paused with a freeze-frame during the vocalization. Mothers of stimulus infants were not visible in any clips. Clip presentation order was randomized. An additional 12 practice clips preceded the test stimulus set.

2.4 | Procedure

Mother–infant dyads came to the laboratory for a single one-hour session. During live play, mothers were asked to play with their infants for 15 minutes as they would at home. Next, mothers responded to playback stimuli while their infants played with an assistant in another room. Mothers were asked to imagine that each stimulus infant was actually in the room with them. Mothers were told that if they felt inclined to respond verbally to the infant after seeing the behavior they should speak that response out loud as if they were actually interacting with the infant.

All participants saw 12 practice stimuli followed by 80 test stimuli. Rate of stimulus presentation was participant controlled. Upon completion, mothers again heard the audio files for the complete stimulus set in a randomized order, and were asked to rate each vocalization on an ordinal 7-point speech maturity scale, defined as how speech-like the vocalization was (1 = least speech-like, 7 = most speech-like).

2.5 | Data coding and analysis

2.5.1 | Infant vocalizations in live play

To calculate the frequency of infant vocalizations produced in live play, vocalizations were divided into syllables. Each syllable comprised a single vowel (V) or a consonant and vowel (CV or VC). Each syllable was classified according to an infraphonological coding system (Oller, 2000). Fusses, vegetative sounds (e.g., coughs), and sounds with oral obstructions (e.g., toy in mouth) were excluded from analyses. The first author coded 100% of live periods and a second coder independently coded 20% of live periods. Reliability was $\kappa = .98$ for infant vocalizations.

2.5.2 | Adult responses in live play and playback

Responses to infant behavior were classified into one of six categories (cf. Tamis-LeMonda et al., 2001; Table 1).[1] The first author coded 100% of the mothers' responses to vocalizations during live play and a second coder independently coded 20% of live periods. Reliability was $\kappa = .89$ for sensitive, $\kappa = .85$ for comment narratives, $\kappa = .86$ for affirmations, $\kappa = .96$ for imitations, and $\kappa = .81$ for redirective responses (overall $\kappa = .91$).

Three coders independently categorized responding to playback stimuli and 20% of the data were recoded as a reliability check. Reliability was $\kappa = .95$ for sensitive, $\kappa = .95$ for comment narratives, $\kappa = .91$ for affirmations, $\kappa = .90$ for imitative responses (overall $\kappa = .96$). All coding was conducted using ELAN coding software created by the Language Archive at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands (https://tla.mpi.nl/tools/tla-tools/elan/; Sloetjes & Wittenburg, 2008).

3 | RESULTS

We analyzed mothers' responses in four ways. First, we compared mothers' response patterns during the playback task and live interaction to assess the validity of the playback paradigm. Second, we analyzed mothers' ratings of how speech-like the playback vocalizations sounded. Third, we tested whether mothers' speech-like ratings were related to their responsiveness to playback vocalizations. Fourth, we tested the effects of vocalization type and context on mothers' responsiveness to playback vocalizations.

3.1 | Validation of playback methodology

To validate the playback paradigm as a meaningful method for assessing caregiver responsiveness, we compared mothers' responses to their own infants during live play to their responses to playback stimuli. First, we calculated the proportion of vocalizations that received a verbal response. In the playback paradigm, each stimulus clip contained a single infant syllable (e.g., [ba]), thus caregivers could respond to a maximum of 80 vocalizations. However, in live play, mothers could not respond to all syllables because infants frequently produced several syllables in rapid succession (e.g., [bababa]); thus mothers would appear less responsive. To equate responsiveness during live play and playback, we counted the number of vocal phrases produced during live play. A phrase consisted of any vocalizations produced in a single breath group or within one second of each other (e.g., [bababa]; cf. Lynch, Oller, Steffens, & Buder, 1995; Gustafson & Green, 1989). Responsiveness during live play was calculated as a proportion of the number of maternal responses to the number of infant vocal phrases; playback responsiveness was calculated as the proportion of maternal responses to total number of syllables.

Verbal response type was compared across the live play and playback tasks to test for consistent maternal reactions. Mothers responded to similar proportions of infant behavior across tasks (playback: $M = .655$, $SD = .241$, live play: $M = .664$, $SD = .186$), $t(39) = -.219$, $p = .828$. Redirective and comment non-sequitur responses were rare in both tasks (< 3% and < .25%, respectively) and were excluded from further analysis.

Response type was then compared across tasks with a 2(task) × 4(response type) repeated-measures ANOVA (Figure 1). A main effect of response type, $F(3,117) = 12.604$, $p < .001$, $\eta_p^2 = .244$, was qualified by a task × response type interaction, $F(3, 117) = 2.794$, $p = .043$, $\eta_p^2 = .067$. In live play, mothers used more sensitive and affirmation responses than narrative and imitation responses, $F(3, 117) = 13.885$, $p < .001$, $\eta_p^2 = .263$, Tukey's HSD *ps* < .01. When responding to playback infants, mothers used more sensitive responses than narrative or imitation responses, $F(3, 117) = 6.556$, $p < .001$, $\eta_p^2 = .144$, Tukey's HSD *ps* < .01.

TABLE 1 Verbal response categories

Response Type	Definition	Example
Sensitive	Statements or actions directly related to the object the baby is focused on	*That's a ball*
Affirmation	Conversational turns that do not provide new information	*Uh-huh, I know*
Narrative	Statements related to baby's state or actions	*You're so big!*
Imitation	Duplications of baby's sound	Baby: [ba]; Mom: [ba]
Redirection	Attempts to move infant attention elsewhere	*Look at this toy instead*
Non-sequitur	Statements unrelated to the infant or current context of the infant's environment	*What should we have for dinner?*

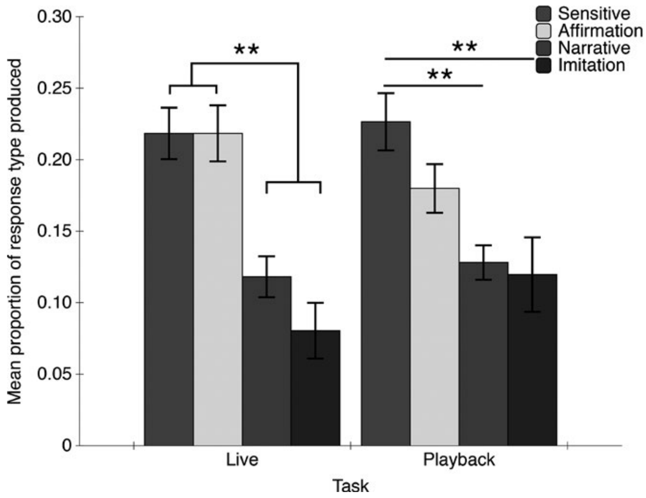

FIGURE 1 Mean proportion of responses to infant vocalizations by task and response type (± 1 *SE*). ** $p < .01$

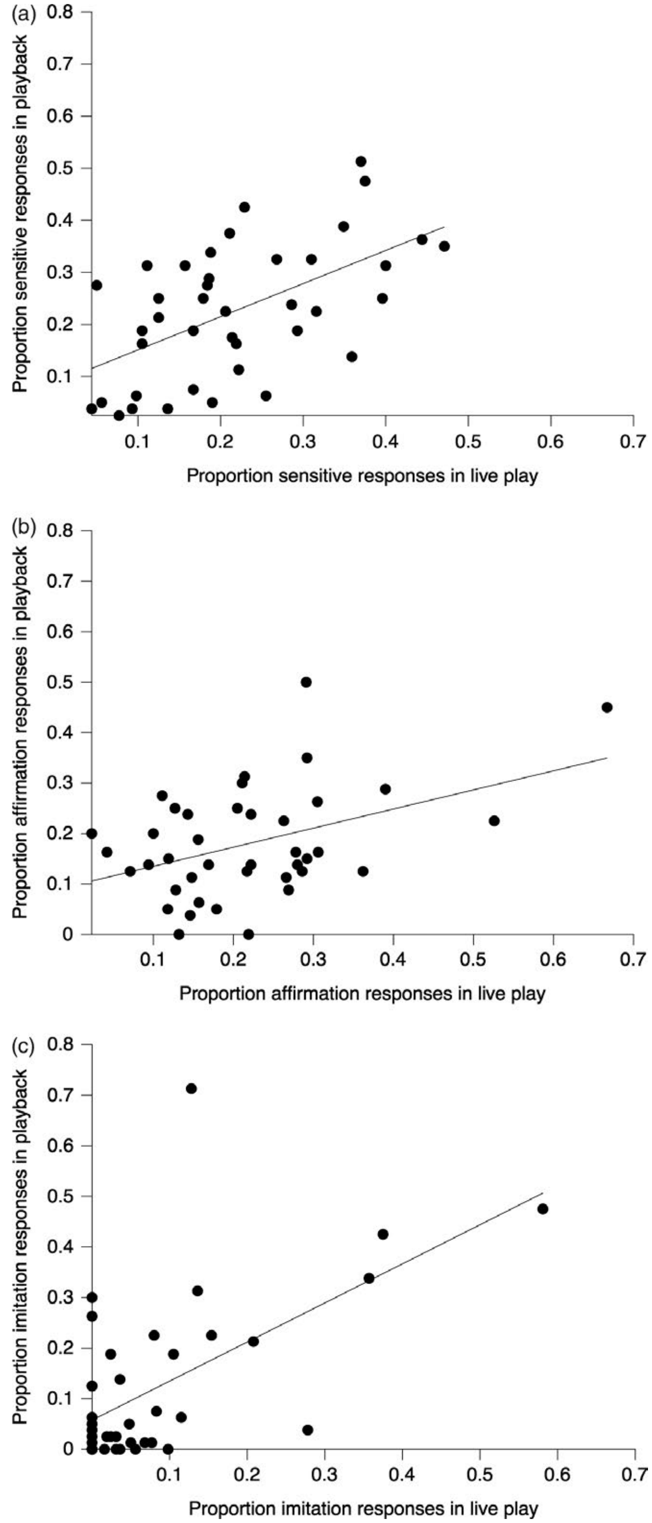

FIGURE 2 Mothers' responses in the live play and playback tasks. (2a) Sensitive responses. (2b) Affirmation responses. (2c) Imitative responses

To further confirm the validity of the playback paradigm, we correlated mothers' response types on the two tasks to examine the types of responses mothers provided when they chose to respond (Figure 2). Mothers' sensitive ($r(38) = .562$, $p < .001$), affirmation ($r(38) = .430$, $p = .006$), and imitative responses ($r(38) =. 596$, $p < .001$) were significantly correlated between the two tasks. Mothers' narrative responses were not significantly correlated between the two tasks ($r(38) = -.041$. $p = .80$). Thus, mothers gave similar patterns of verbal responses to their own infants and to the unfamiliar playback stimuli. These results support the validity of the playback paradigm as an indicator of natural maternal responding.

3.2 | Perception of vocal maturity

To test whether infraphonological type influenced perception of stimulus infant vocalizations, a Friedman test with repeated measures on infraphonological type (QRV, FRV, MS, CS) was conducted for mothers' ratings of the speech-like quality of infant vocalizations. There was a significant main effect of infraphonological type, χ^2 (3) = 79.42, $p < .001$ (Figure 3). Post-hoc tests (Siegel & Castellan, 1988) indicated that mothers rated canonical syllables as significantly more speech-like than the other three syllable types (CS vs. QR, $p < .001$; CS vs. MS, CS vs. FR, $ps < .01$). Quasi-resonant vowels were rated as significantly less speech-like than the other three syllable types ($ps < .001$). However, ratings of fully-resonant vowels and marginal syllables did not significantly differ ($p > .05$). Thus, the infraphonological syllable type influenced perception of infant vocal quality.

3.3 | Responses to playback stimuli

3.3.1 | Response rates by speech rating

Mothers used the majority of ratings on the speech maturity scale, with a majority of mothers using the ratings from 1 to 6 (33% of mothers did not use the highest rating on the speech-like scale; Figure 4). To test the relation between mothers' response rates and their ratings of the speech-like quality of sounds, we conducted a Friedman test with repeated measures on mothers' speech-like

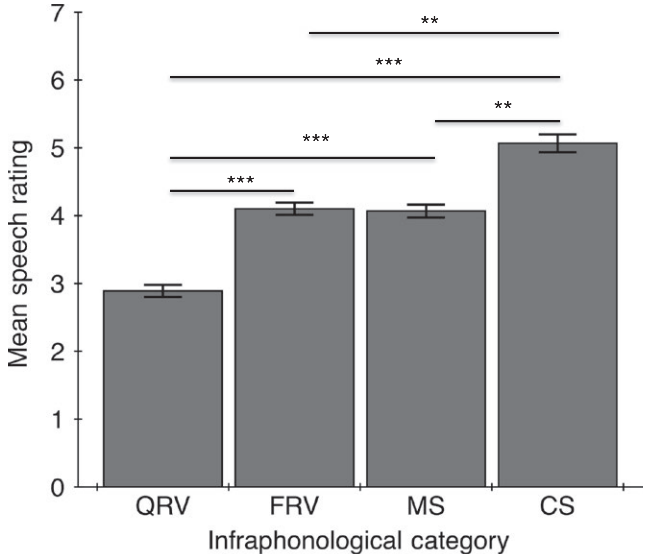

FIGURE 3 Mean speech maturity rating for each infraphonological type (±1 *SE*). *** $p < .001$

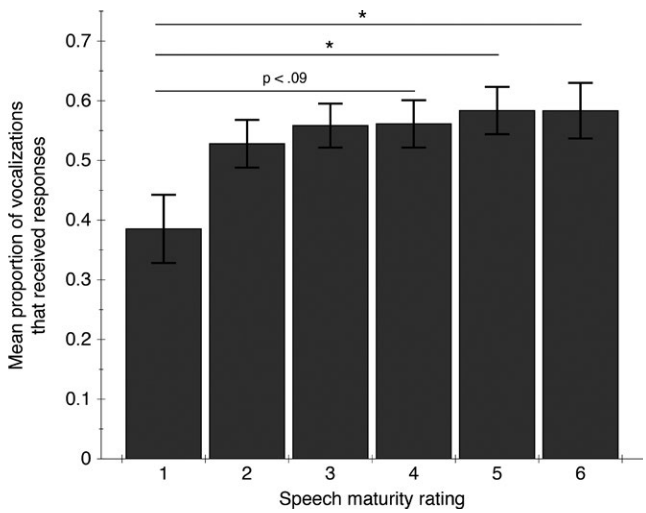

FIGURE 4 Proportion of stimuli eliciting a response by speech rating. * $p < .05$

rating scores (from 1 to 6) to determine whether mothers were more likely to give verbal responses to the sounds they later rated as more speech-like. There was a significant effect of speech rating on response rates, χ^2 (5) = 11.405, p = .044. Post-hoc tests (Siegel & Castellan, 1988) revealed that mothers were significantly more responsive to the sounds they rated as more speech-like (5 or 6) than to sounds that they rated as least speech-like (a 1 on the scale), *ps* < .05, with a trend toward more responsiveness for sounds given a rating of 4 than for sounds rated 1, $p < .09$. Speech ratings were significantly correlated with response rate, r = .311, p = .003 (within-subjects correlation calculated following the procedure of Bland & Altman, 1995). Mothers were more likely to respond to the vocalizations that they perceived as more speech-like.

3.3.2 | Response rates by vocalization type

Stimulus infants provided clips either for vowel (QR/FR) comparisons or for consonant-vowel (MS/CS) comparisons. We conducted separate analyses on vowels and consonant-vowel syllables. We also compared responsiveness across infraphonological categories (see Supporting Material).

Vowels

To compare how changing vowel type and vocalization context influenced mothers' response type responsiveness to vowels was analyzed with a 2(directedness: ODV, UDV) × 2(vowel type: QRV, FRV) × 4(response type: sensitive, affirmation, narrative, imitation) repeated-measures ANOVA. There was a significant main effect of directedness, $F(1, 39) = 28.292$, $p < .001$, $\eta_p^{\ 2} = .420$. Mothers responded significantly more often to object-directed vocalizations (M = .177, SD =.153) than to undirected vocalizations (M = .146, SD = .142). A significant main effect of vowel type showed that mothers also responded significantly more often to fully-resonant vowels (M = .172, SD = .159) than to less-mature quasi-resonant vowels (M = .150, SD = .141), $F(1, 39) = 30.333$, $p < .001$, $\eta_p^{\ 2} = .437$. A main effect of response type showed that sensitive responses were more frequent than narratives and imitations, $F(3, 117) = 14.614$, $p < .001$, $\eta_p^{\ 2} = .273$, Tukey HSD *ps* < .05. Thus, mothers' responses were influenced by vocalization directedness and infraphonological category.

These main effects were qualified by a significant directedness × response type interaction, $F(3, 117) = 34.480$, $p < .001$, $\eta_p^{\ 2} = .469$ (Figure 5). Sensitive responses were given more often to ODVs than to UDVs, $F(1, 39) = 69.333$, $p < .001$, $\eta_p^{\ 2} = .640$. Affirmations were given more often to UDVs than to ODVs, $F(1, 39) = 5.861$, $p = .02$, $\eta_p^{\ 2} = .131$. There was a trend for narrative responses to be given more often to UDVs than ODVs, $F(1, 39) = 4.072$, $p = .051$, $\eta_p^{\ 2} = .095$. For imitative responses, there was no significant effect of directedness, $p = .156$. No other main effects or interactions were significant, *ps* > .2. In sum, mothers' response types were influenced by vocalization directedness.

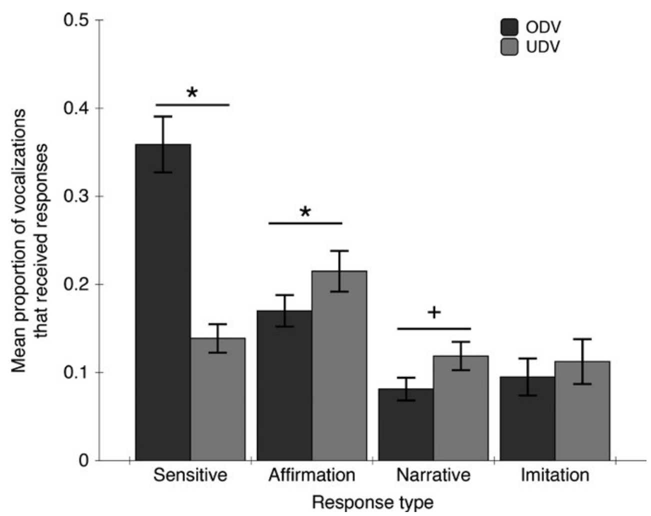

FIGURE 5 Mothers' mean verbal responses to vowel playback stimuli (±1 *SE*) by directedness and response type. + $p = .051$

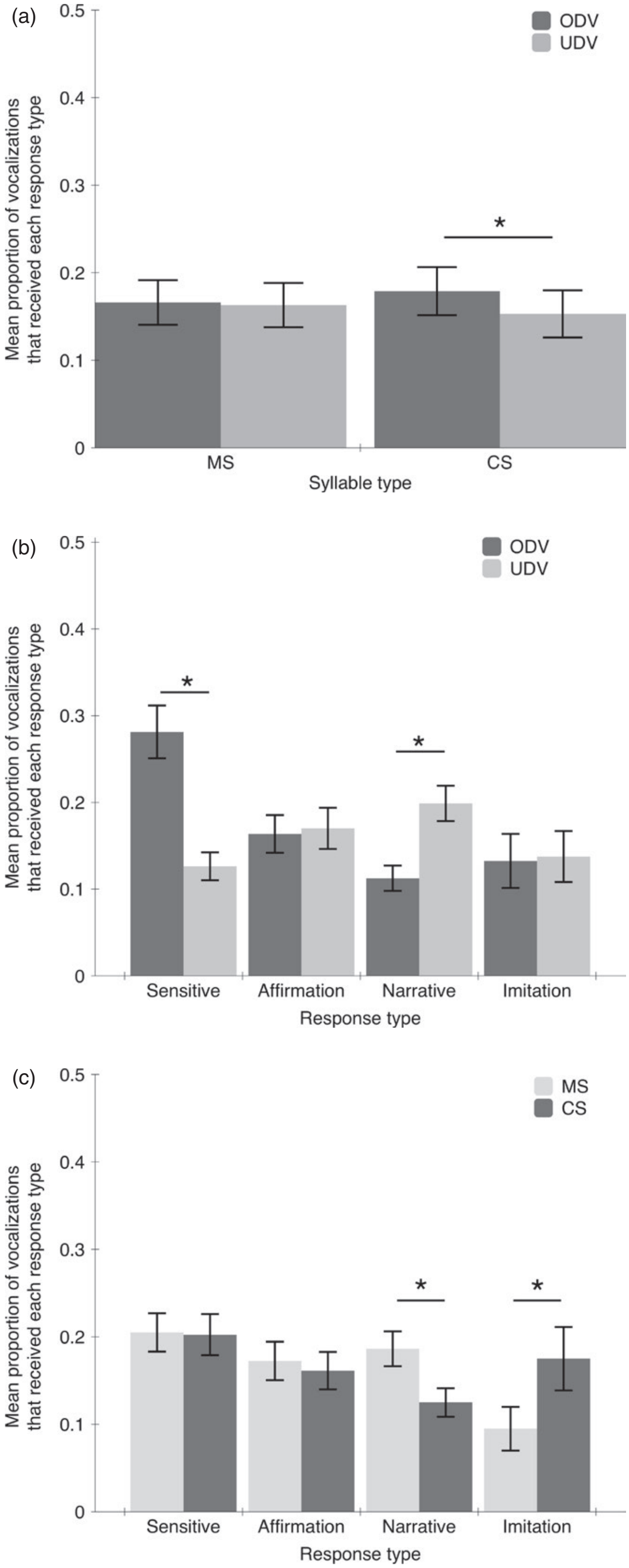

FIGURE 6 Mothers' mean verbal responses to consonant-vowel playback stimuli (±1 *SE*). (6a) Responses to consonant-vowel stimuli by directedness and infraphonological type. (6b) Responses to consonant-vowel stimuli by directedness and response type. (6c) Responses to consonant-vowel stimuli by infraphonological type and response type

Consonant-vowels

To compare how changing syllable type and vocalization context influenced mothers' response type, responsiveness to CV syllables was analyzed with a 2(directedness) × 2(infraphonological type: MS, CS) × 4(response type) repeated-measures ANOVA. As when responding to vowels, mothers responded significantly more often to object-directed vocalizations (M = .173, SD = .167) than to undirected vocalizations (M = .158, SD = .165), $F(1, 39) = 12.058$, $p = .001$, $\eta_p^2 = .236$. No other main effects were significant, *ps* > .193.

A significant directedness × infraphonological type interaction [$F(1, 39) = 5.274$, $p = .027$, $\eta_p^2 = .119$] was decomposed by infraphonological type (Figure 6a). For canonical syllables, mothers gave more responses when the vocalization was object-directed than undirected, $F(1, 39) = 12.143$, $p = .001$, $\eta_p^2 = .237$. For marginal syllables, directedness was not significant, $p = .560$.

A significant directedness × response type interaction [$F(3, 117) = 22.051$, $p < .001$, $\eta_p^2 = .361$] was decomposed by response type (Figure 6b). As for vowels, mothers gave more sensitive responses to ODVs than to UDVs, $F(1, 39) = 44.145$, $p < .001$, $\eta_p^2 = .531$. Mothers gave more narrative responses to UDVs than to ODVs, $F(1, 39) = 23.213$, $p < .001$, $\eta_p^2 = .373$. No other effects were significant, *ps* > .70.

A significant infraphonological type × response type interaction [$F(3, 117) = 8.248$, $p < .001$, $\eta_p^2 = .175$ (Figure 6c)], decomposed by response type, revealed that mothers imitate canonical syllables more often than marginal syllables, $F(1, 39) = 16.129$, $p < .001$, $\eta_p^2 = .293$. In contrast, narrative responses were given more often to marginal syllables than to canonical syllables, $F(1, 39) = 9.635$, $p = .004$, $\eta_p^2 = .198$. No other main effects or interactions were significant, *ps* > .550. In sum, mothers' response types were influenced by vocalization directedness and infraphonological type.

4 | DISCUSSION

Investigating real-time reactions to controlled examples of infant behavior revealed new patterns of parental sensitivity to early vocalizations. We manipulated the acoustic qualities and directedness of infant vocalizations to systematically examine their effects on mothers. Mothers reliably ranked the maturity of the vocalizations based on the acoustic features of resonance and the timing of consonant–vowel transitions. Mothers responded more frequently and sensitively to object-directed vocalizations than to undirected vocalizations. Mothers also responded more frequently to more developmentally advanced vowels. We predicted that ODVs would elicit sensitive responses because infants were vocalizing while attending to an object that could be labeled. In contrast, UDVs elicited more narrative and affirmation responses because there was no obvious "referent" to the babble and thus nothing to label. The differential information content of responses to object-directed and undirected vocalizations may provide infants with different opportunities for learning. In sum, vowel *quality* influenced how likely mothers were to respond, but not how they responded; vowel *directedness* influenced both the frequency and type of caregiver response.

Mothers also differentiated their responses to more complex syllables that had a consonant-vowel structure. Canonical syllables, the most mature form, received the highest proportions of imitative speech, while less-mature marginal syllables elicited the highest proportion of narrative responses. Parental imitation is hypothesized to be helpful for infant speech development by providing infants with a model to target in future vocalizations (Kuhl & Meltzoff, 1996). Recent computational models of speech acquisition using robotic learners have demonstrated the importance of imitative input for facilitating learning (e.g., Yoshikawa, Asada, Hosoda, & Koga, 2003). Parental imitation of infant behavior may also act as "social glue", demonstrating similarity and social affiliation (Chartrand & Bargh, 1999; van Baaren, Holland, Kawakami, & van Knippenberg, 2004). Mothers may imitate the sounds they perceive as most speech-like in an attempt to affiliate with infants. Marginal syllables were speech-like enough to elicit responses, but did not evoke imitations. Mothers may have commented on infants' behaviors as a way of taking a conversational turn to continue the interaction. Taken together, these results indicate reliable proximal cues that facilitate maternal behavior, leading to new learning opportunities for infants.

Our findings suggest that babbling can organize the social interactions in which infants experience contingent feedback from caregivers. Such contingent feedback facilitates language development. Previous research examining infant vocal learning in moment-to-moment interactions demonstrates that infants learn phonological patterns and words from contingent feedback to their vocalizations (Goldstein & Schwade, 2008, 2010). We have shown differential parental responsiveness as a function of infant vocal quality and directedness. More developmentally advanced and object-directed vocalizations were most likely to obtain a response. Thus, infants may contribute to their own language development via the influence their early sounds have on caregivers.

The playback paradigm allowed us to explore interactions between acoustic and contextual characteristics of infant vocalizations. Mothers most frequently provided sensitive responses to object-directed vocalizations. Mothers also responded more to the most acoustically advanced sounds when they were in an object-directed context rather than undirected. These results support the idea that caregivers frequently respond to ODVs by commenting on the object, which aligns the caregiver's focus of attention with that of the infant. These findings complement previous correlational studies that examined long-term relations between the type of label provided after an ODV and later vocabulary (Goldstein & Schwade, 2010). Not only do ODVs appear to elicit more specific information from caregivers, but previous work also demonstrates that infants learn the labels for objects better after receiving that information following an ODV versus a silent look (Goldstein et al., 2010). Infants may modulate their own arousal level and attentional focus through the production of prelinguistic vocalizations, especially those that are directed at nearby objects (Albert, Schwade, & Goldstein, submitted). In our view, an infant's object-directed vocalization creates a state of receptivity for learning at the moment the caregiver is likely to label the object to which the infant is attending. Therefore, the infant's own vocalizations serve to structure social interactions in ways that facilitate learning.

These results contribute to a growing understanding of the role of social feedback in infant vocal learning (e.g., Gros-Louis et al., 2014; Warlaumont et al., 2014), which stands in contrast to the historical view of prelinguistic vocalizations in which babbling was assumed to be motor practice, with no function in the development of communication and language (e.g., Jakobson, 1968). Although it is well known that social responsiveness to infant behavior has impacts on later language development (e.g., Rollins, 2003; Tamis-LeMonda et al., 2001; Tomasello & Farrar, 1986), identifying the specific infant behaviors that modulate responsiveness will move us closer to understanding the real-time dynamics of how and when vocal learning occurs. In toddlers, quantifying the dynamics of parent–infant interaction is proving useful for understanding word learning (e.g., Pereira, Smith, & Yu, 2014; Yu & Smith, 2012); we believe that a similar approach with younger infants will shed light on how lexical and syntactic form emerge from vocal precursors.

The length of infant utterances may also influence responding. Single syllables have historically been treated as the minimum rhythmic unit of speech (Lynch et al., 1995). As infants age, they begin producing disyllables (e.g., baba) and strings of repeated single syllables known as reduplicated babbles (Oller, 2000). Variegated babbles, utterances made up of varied syllables produced in a single breath, also become more common (Stoel-Gammon, 1992). In the current study we restricted our stimuli to investigate the influences of single-syllable utterances so as to isolate the effects of specific acoustic features. Future studies will investigate the impact of multi-syllable utterances on caregiver responding. In addition to vocal quality, infant gaze while vocalizing may influence caregiver responsiveness. We restricted our study to comparisons of object-directed and undirected vocalizations because previous research indicated that object-directed vocalizations were salient to caregivers and directly influenced infant language learning (Goldstein et al., 2010).

When we validated the playback method, we found that mothers' responses to audiovisual examples of unfamiliar infants' vocalizations paralleled their patterns of reactions to their own infants. Maternal responsiveness was stable across tasks. Thus, the playback paradigm is a valid and informative method for quantifying the social environment in which early communication and language learning are embedded. Overall, our findings support and extend earlier studies that found consistency in parents' reactions to immature infant behavior (e.g., Wood & Gustafson, 2001). However, responding to prerecorded stimuli even with naturalistic behaviors such as talking is not the same as having a conversation with a live infant. Mothers were not perfectly consistent in their behavior across tasks. Speech that was sensitive, affirmative, or imitative was significantly correlated across tasks, although these correlations were in the .43–.59 (moderate) range. In contrast, narrative speech was not consistent. These differences in consistency point to the need for future studies to include validity checks across playback and live interaction. The findings from the present study demonstrate that vocal quality influences caregiver responses to babbling. The differential feedback mothers provide to prelinguistic vocalizations may explain past correlations between early babbling and later language. For example, infants who produce more canonical syllables at 1 year of age have more advanced vocabulary and speech later in development (Stoel-Gammon, 1992). Further, children who have a delay in the onset of canonical babbling tend to be late talkers, with delayed productive vocabulary (Paul & Jennings, 1992; Rescorla & Ratner, 1996). Early vocalizations could contribute to a feedback loop in which parents' reactions reward more speech-like babbling (Goldstein & Schwade, 2010; Warlaumont et al., 2014). In particular, sensitive responses (e.g., object labels) given to more advanced vocalizations may promote word learning.

The results of our study, put into the context of a vocal feedback loop, may also be helpful for understanding vocal development in infants in at-risk populations such as those with hearing delays, Down syndrome, and autism spectrum disorder (ASD). For example, children with ASD produce atypical vocalizations (Oller et al., 2010), specifically showing a delayed onset in canonical syllable production (Patten et al., 2014). Children with ASD also produce fewer speech-like vocalizations in comparison to typical children (Warlaumont et al., 2014). Parents of children with ASD provide fewer contingent responses to speech-related vocalizations, which likely reduces the strength of the social feedback loop (Warlaumont et al., 2014). These small differences in moment-to-moment interactions between infants and caregivers may cascade into long-term differences in response expectancies, impacting language development over time as opportunities for learning from contingent parental responses are reduced. We suggest that rich, detailed datasets taken from playback experiments and real-time interaction be combined with longitudinal designs to reveal both sides of the feedback loop: how the microstructure of interaction supports and facilitates vocal development and communication, and how advances in vocalizations change parental behavior.

In summary, studying the microstructure of parental behavior via responses to playback of controlled infant stimuli revealed new patterns of maternal behavior that are likely to facilitate infant vocal learning and development. Future playback studies can be improved by leveraging imaging technology to create even more realistic stimuli that will include vocalizations directed to the caregiver (i.e., the viewer). Camera angle and infant gaze can be tightly aligned to create the illusion of direct eye contact between the stimulus infant and the participant. While we recognize that vocalizations directed at other people likely impact the development of communication, these technical concerns led us to focus first on understanding the impact of object-directed and undirected vocalizations on caregiver responding. Future studies will test the influences of person-directed vocalizations on caregiver behavior. While our focus in this paper was on understanding the social functions of babbling, an advantage of the playback paradigm is that it can be used to investigate a multitude of infant social cues. For example, previous work shows that infant gestures often elicit labeling responses from caregivers (Wu & Gros-Louis, 2014). Future playback studies could examine the differential effects of deictic gestures (e.g., showing, pointing) versus more conventional gestures (e.g., waving or headshaking) alone and in conjunction with vocalizing. Playback studies could also compare maternal and paternal behavior, as well as assess cross-cultural differences in patterns of parental responsiveness, as culture is a major source of variability in parenting (Bornstein, 2012).

Parental responsiveness is dynamic and changes as infants display more mature behaviors and vocalizations (Holden, 2010). As infants age, caregivers respond less by describing and encouraging exploration of objects and increasingly respond with vocal imitations and questions (Bornstein, Tamis-LeMonda, Hahn, & Haynes, 2008). Combining assessments of developmental changes in parental responsiveness to prelinguistic vocalizations with infant vocal learning and language measures will illuminate the specific aspects of parental behavior that facilitate speech and language development.

ACKNOWLEDGEMENTS

This research was supported by grants from NSF (0844015) and NICHD (5R03HD61524-2) to MHG. We thank the three anonymous reviewers for their helpful comments on earlier versions of this manuscript. We also thank BABY Lab research assistants for their assistance with data collection and coding and the families who participated in this research.

ENDNOTE

[1] Responses were coded using a coding scheme derived from Tamis-LeMonda et al. (2001). However, some of those categories (e.g., "exploratory prompts", in which the mother attempted to draw the infant's attention to an object she was holding) were not appropriate for both the live play and playback contexts. In addition to vocal responses, overt non-verbal behaviors such as gestures were also counted as responses. Smiles were not counted as a response in either task. In the live play task, non-verbal responses were typically accompanied by verbal responses. Non-verbal behaviors were rare in the playback task given the nature of the task (e.g., holding a stylus, inaccessibility of the toys present in the video).

REFERENCES

Adobe (2009). Adobe Director 11.5 [Software].

Akhtar, N., Dunham, F., & Dunham, P.J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language*, *18*, 41–49.

Albert, R.R., Schwade, J.A., & Goldstein, M.H. (submitted). Playback of prelinguistic object-directed vocalizations facilitates word learning.

Baumwell, L., Tamis-LeMonda, C.S., & Bornstein, M.H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant Behavior and Development*, *20*, 247–258.

Bland, J.M., & Altman, D.G. (1995). Calculating correlation coefficients with repeated observations: Part 1—Correlation within subjects. *BMJ*, *310*, 446.

Bloom, K., D'Odorico, L., & Beaumont, S. (1993). Adult preferences for syllabic vocalizations: Generalizations to parity and native language. *Infant Behavior and Development*, *16*, 109–120.

Bloom, K., & Masataka, N. (1996). Japanese and Canadian impressions of vocalising infants. *International Journal of Behavioral Development*, *19*, 89–99.

Bornstein, M.H. (2012). Cultural approaches to parenting. *Parenting*, *12*, 212–221.

Bornstein, M.H., Tamis-LeMonda, C.S., Hahn, C.-S., & Haynes, O.M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, *44*, 867–874.

Braarud, H.C., & Stormark, K.M. (2008). Prosodic modification and vocal adjustments in mothers' speech during face-to-face interaction with their two- to four-month-old infants: A double video study. *Social Development*, *17*, 1074–1084.

Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Deák, G.O., Krasno, A.M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science*, *17*, 270–281.

Degotardi, S., & Sweller, N. (2012). Mind-mindedness in infant child-care: Associations with early childhood practitioner sensitivity and stimulation. *Early Childhood Research Quarterly*, *27*, 253–265.

Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, *6*, 423–443.

Gleitman, L.R., Newport, E.L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*, 43–79.

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, USA*, *100*, 8030–8035.

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

Goldstein, M.H., & Schwade, J.A. (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 and comparative neuroscience* (pp. 708–729). New York: Oxford University Press.

Goldstein, M.H., Schwade, J.A., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy*, *15*, 362–391.

Goldstein, M.H., & West, M.J. (1999). Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size. *Journal of Comparative Psychology*, *113*, 52–57.

Gros-Louis, J., West, M.J., & King, A.P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, *19*, 385–408.

Gustafson, G.E., & Green, J.A. (1989). On the importance of fundamental frequency and other acoustic features in cry perception and infant development. *Child Development*, *4*, 772–780.

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

Holden, G.W. (1985). Analyzing parental reasoning with microcomputer presented problems. *Simulation Gaming*, *16*, 203–210.

Holden, G.W. (2010). *Parenting: A dynamic perspective*. Thousand Oaks, CA: Sage.

Holden, G.W., Ritchie, K.L., & Coleman, S.D. (1992). The accuracy of maternal self-reports: Agreement between reports on a computer simulation compared with observed behavior in the supermarket. *Early Development and Parenting*, *1*, 109–119.

Illman, G., Neuhauserova, K., Pokorna, Z., Chaloupkova, H., & Simeckova, M. (2008). Maternal responsiveness of sows towards piglet's screams during the first 24 hours postpartum. *Applied Animal Behaviour Science*, *112*, 248–259.

Jakobson, R. (1968). *Child language, aphasia, and phonological universals*. The Hague: Mouton.

Kuhl, P.K., & Meltzoff, A.N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, *100*, 2425–2438.

Lynch, M.P., Oller, D.K., Steffens, M.L., & Buder, E.H. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychobiology*, *28*, 3–25.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *21*, 746–748.

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

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., Eilers, R.E., Neal, A.R., & Schwartz, H.K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, *32*, 223–246.

Oller, D.K., Niyogi, P., Gray, S., Richards, J.A., Gilkerson, J., Xu, D., … Warren, S.F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences, USA*, *107*, 13354–13359.

Paavola, L., Kunnari, S., Moilanen, I., & Lehtihalmes, M. (2005). The functions of maternal verbal responses to prelinguistic infants as predictors of early communicative and linguistic development. *First Language*, *25*, 173–195.

Papoušek, M. (1989). Determinants of responsiveness to infant vocal expression of emotional state. *Infant Behavior and Development*, *12*, 507–524.

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*, *44*, 2413–2428.

Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech and Hearing Research*, *35*, 99–107.

Pereira, A.F., Smith, L.B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin and Review*, *21*, 178–185.

Rescorla, L., & Ratner, N.B. (1996). Phonetic profiles of toddlers with specific expressive impairment (SLI-E). *Journal of Speech and Hearing Research*, *39*, 153–165.

Rollins, P.R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied Psycholinguistics*, *24*, 221–234.

Shizawa, Y., Nakamichi, M., Hinobayashi, T., & Minami, T. (2005). Playback experiment to test maternal responses of Japanese macaques (*Macaca fuscata*) to their own infant's call when the infants were four to six months old. *Behavioural Processes*, *68*, 41–46.

Siegel, S., & Castellan, N.J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd edn.). New York: McGraw-Hill.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category—ELAN and ISO DCR. In: Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).

Smith, V.A., King, A.P., & West, M.J. (2000). A role of her own: Female cowbirds, *Molothrus Ater*, influence the development and outcome of song learning. *Animal Behaviour*, *60*, 599–609.

Stevens, E., Blake, J., Vitale, G., & MacDonald, S. (1998). Mother-infant object involvement at 9 and 15 months: Relation to infant cognition and early vocabulary. *First Language*, *18*, 203–222.

Stoel-Gammon, C. (1992). Prelinguistic vocal development: Measurement and predictions. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 439–456). Timonium, MD: York Press.

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*, 734–738.

Takahashi, D.Y., Liao, D.A., & Ghazanfar, A.A. (2017). Vocal learning via social reinforcement by infant marmoset monkeys. *Current Biology*, *27*, 1844–1852.

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

Tomasello, M., & Farrar, M.J. (1986). Joint attention and early language. *Child Development*, *57*, 1454–1463.

van Baaren, R.B., Holland, R.W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science*, *15*, 71–74.

Warlaumont, A.S., & Finnegan, M.K. (2016). Learning to produce syllabic speech sounds via reward-modulated neural plasticity. *PLoS ONE*, *11*, e0145096.

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*, 1314–1324.

Wood, R.M., & Gustafson, G.E. (2001). Infant crying and adults' anticipated caregiving responses: Acoustic and contextual influences. *Child Development*, *72*, 1287–1300.

Wu, Z., & Gros-Louis, J. (2014). Caregivers provide more labeling responses to infants' pointing than to infants' object-directed vocalizations. *Journal of Child Language*, *42*, 538–561.

Yoshikawa, Y., Asada, M., Hosoda, K., & Koga, J. (2003). A constructivist approach to infants' vowel acquisition through mother–infant interaction. *Connection Science*, *15*, 245–258.

Yu, C., & Smith, L.B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*, 244–262.

Yu, C., & Smith, L.B. (2015). Linking joint attention with hand-eye coordination: A sensorimotor approach to understanding child–parent social interaction. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 2763–2768.

Zeskind, P.S., Klein, L., & Marshall, T.R. (1992). Adults' perceptions of experimental modifications of durations and pauses and expiratory sounds in infant crying. *Developmental Psychology*, *28*, 1153–1162.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Albert RR, Schwade JA, Goldstein MH. The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness. *Dev Sci*. 2017;e12641. https://doi.org/10.1111/desc.12641