Research Article

Social Feedback to Infants' Babbling Facilitates Rapid Phonological Learning

Michael H. Goldstein and Jennifer A. Schwade

Cornell University

ABSTRACT—*Infants' prelinguistic vocalizations are rarely considered relevant for communicative development. As a result, there are few studies of mechanisms underlying developmental changes in prelinguistic vocal production. Here we report the first evidence that caregivers' speech to babbling infants provides crucial, real-time guidance to the development of prelinguistic vocalizations. Mothers of 9.5-month-old infants were instructed to provide models of vocal production timed to be either contingent or noncontingent on their infants' babbling. Infants given contingent feedback rapidly restructured their babbling, incorporating phonological patterns from caregivers' speech, but infants given noncontingent feedback did not. The new vocalizations of the infants in the contingent condition shared phonological form but not phonetic content with their mothers' speech. Thus, prelinguistic infants learned new vocal forms by discovering phonological patterns in their mothers' contingent speech and then generalizing from these patterns.*

The vocal abilities of infants change dramatically over the first year of life. After producing their earliest, immature vocalizations, infants make rapid progress, typically producing their first words by 12 months of age (Fenson et al., 1994). Along the way, they learn to produce speechlike syllables (Oller, 2000) and to structure their phonology in accordance with their language environment (Boysson-Bardies, 1993). Although the vocal achievements of the first year are well described, less is known about the mechanisms of change that drive vocal development. Most research in this area emphasizes internal causes of developmental change, such as anatomical changes in the vocal tract (e.g., Kent, 1981). However, the social environment of a vocalizing infant is a source of rich structure that can guide advances in vocal learning. Caregivers provide consistent responses to babbling (Goldstein & West, 1999), and infants use caregivers' reactions to learn new patterns of vocalizing (Goldstein, King, & West, 2003).

Although there is a paucity of data on mechanisms of vocal learning, a large body of research demonstrates the influence of the ambient language on speech perception. Adult speech exerts a canalizing effect on the development of phonological categories by the time infants are 8 to 10 months old, leading to the formation of perceptual categories that match the phonemic contrasts present in the language environment (Jusczyk, 1992; Werker & Curtin, 2005). The development of categorical speech perception is facilitated by social interaction. Social feedback may focus attention on relevant features of the speech signal or provide additional sources of information that specify perceptual categories (Kuhl, 2007). The role of social interaction in focusing attention on caregivers' speech may be particularly important in vocal learning, as infants show deficits in speech processing when attending to other speech functions, such as when learning word-object associations (Werker & Curtin, 2005).

If the experience of hearing other people speak influences speech perception and production, how is the phonology of the ambient language incorporated into babbling? Few attempts have been made to explore the role of social feedback on early phonological and articulatory development. Instead, infant sounds have been extensively studied from a taxonomic perspective, which describes vocal development in terms of universal and invariant stages (Oller, 2000; Stark, 1980). In Oller's (2000) taxonomic approach, infants' vocalizations are described in terms of their resemblance to the acoustics of mature syllable production, or *infraphonology*, so that vocal development can be compared across languages. When vocal learning has been

Address correspondence to Michael H. Goldstein, Department of Psychology, Cornell University, 242 Uris Hall, Ithaca, NY 14853-7601, e-mail: mhg26@cornell.edu.

observed, it has been ascribed to a mechanism of imitation (Kuhl & Meltzoff, 1996). However, recent studies of mother-infant interactions found little mimicry by prelinguistic infants; most matching results from caregivers imitating their infants, rather than from infants matching their caregivers (Jones, 2007).

The goal of the current study was to determine the role of contingent adult speech in providing real-time guidance to infants' prelinguistic vocal learning. The vocal behavior of infants is responsive to social contingencies (Bloom, Russell, & Wassenberg, 1987; Masataka, 2003), which creates opportunities for trial-and-error learning long before the onset of traditional cognitive milestones of communication (Locke, 2001). Infants can use social feedback to facilitate developmental transitions in vocal behavior (Goldstein et al., 2003; Masataka, 2003). For example, 3-month-old infants who engaged in vocal turn taking with adults who responded contingently produced sounds that were more speechlike (less nasalized and more fully resonant) than those of infants who vocalized to adults who responded on a random schedule (Bloom et al., 1987). In another study, when caregivers coordinated their behavior with their infants' babbling, they created contingent social feedback that resulted in increased production of more advanced forms of vocalizing, such as canonical syllables (Goldstein et al., 2003); only contingent social feedback created changes in infants' vocalizations. Thus, contingent social responses to babbling constitute a form of prelinguistic coordination in which caregivers provide structured feedback to their infants' vocalizations. These results show that the responses of caregivers may function as a social mechanism of vocal learning. The present study is the first to measure infants' vocal learning in real time by manipulating the phonological structure present in caregivers' speech.

METHOD

Participants

Sixty 9.5-month-old infants (mean age = 9 months 17 days, range: 8 months 27 days–10 months 14 days) participated with their mothers. Participants were recruited from birth announcements in local newspapers and through advertisements. An additional 31 infants were tested but excluded because the infants cried or fussed excessively ($n = 14$), the mothers did not follow directions for when and how to react to babbling ($n = 8$), there was an equipment failure ($n = 1$), or the infants produced fewer than five vocalizations, which was below the 10th percentile ($n = 8$). (Because measures of vocal quality were calculated as proportions, the latter infants were excluded so that they would not disproportionately increase the amount of vocal changes observed.) Participants received an infant-size T-shirt or bib.

Apparatus

Infants were tested in a naturalistic environment that allowed for detailed audio and video recording. All sessions took place in a large, infant-safe playroom (12 ft. × 18 ft.) containing toys, a toy box, and posters of animals. The infants were free to move about and explore, so they did not have to engage in social interaction with their caregivers. Dyads were recorded with three remotely controlled digital cameras via a video mixer. So that we could obtain high-quality recordings of the infants' vocalizations, each infant wore a pair of denim overalls concealing a wireless microphone and transmitter. Each mother wore a wireless lapel microphone with a transmitter concealed in a pouch at her waist. During the second session (see Procedure), the experimenter gave mothers instructions via wireless headphones. A CD player routed to the yoked control (see Procedure) mothers' headphones cued their responses. Mothers' speech, the experimenter's instructions, and infants' vocalizations were recorded on separate audio channels.

Procedure

We manipulated the form and timing of caregivers' vocal responses to their 9.5-month-old infants in order to assess the infants' ability to learn new patterns of vocalizing from the phonological structure of their caregivers' speech. Participants came to the lab for two 30-min play sessions, spaced 24 hr apart. The first was a familiarization session that allowed dyads to get used to the playroom. The second session followed an ABA design and was divided into three 10-min periods: baseline recording (Baseline 1), social response, and baseline recording (Baseline 2). During the baseline periods, mothers were asked to play as they would at home. During the social-response period (the only time we manipulated mothers' behavior), mothers received instructions (over wireless headphones) as to when and how they should react to a babble.

Dyads participated in one of the four social-response conditions. In the contingent conditions, mothers were asked to respond to each babble by speaking while moving closer to, smiling at, and touching their infants. Half the mothers were asked to respond to babbles by speaking fully resonant vowels (contingent-resonant group), and half were asked to respond by speaking words so that the infants would be exposed to consonant-vowel (CV) alternation (contingent-CV group). Thus, each mother's utterances were consistent in phonological structure (either fully resonant vowels or CV alternation), but varied in phonetic features, as the specific phonemes used varied.

To separate the effect of contingency from that of arousal caused by mothers' responses, we ran yoked control conditions. Mothers in the yoked-resonance and yoked-CV groups responded using the same utterances as the contingent-condition mothers with whom they were paired. The timing of yoked control mothers' responses was governed by the contingent-condition mothers. Thus, the contingent groups were run first. To control the timing of yoked control mothers' responses, we created a CD track of each contingent-condition mother's contingent utterances. Extraneous sounds were removed while preserving the timing of speech. These CD tracks were played back to the yoked control mothers over the wireless headphones. These mothers were asked to repeat the vocalizations that they heard over their headphones, while getting closer to, touching, and smiling at their infants. The timing and form of yoked control mothers' responses were thus linked to those of the contingent-condition mothers so that yoked control infants received the same amount and type of social stimulation as the contingent-condition infants, but the stimulation was not synchronized with their vocalizations.

The first 30 infants were tested in the contingent condition; they were randomly assigned to either the resonance ($n = 15$) or the CV ($n = 15$) group. The remaining 30 infants participated in the yoked control conditions ($n = 15$ in each). Gender of the infants was balanced within each condition.

Data Coding and Analysis

Maternal Responses

A maternal response was coded when a mother changed her behavior toward her infant within 2 s of the offset of the infant's vocalization (e.g., by speaking to her infant or shaking a toy). Mothers' responses to fusses, raspberries, and vegetative sounds (e.g., coughs) were excluded from analyses. Seven coders were trained until they reached better than 90% reliability with the first author. All coders were blind to the hypotheses of the study and the experimental group to which each infant was assigned. Twenty percent of sessions were recoded by an additional trained coder. Mean reliability was .95 (range: .90–1.00).

Phonology of Infants' Vocalizations

Infants' vocalizations were coded into four types according to Oller's (2000) infraphonological acoustic classification system. This system incorporates acoustic parameters (e.g., fundamental frequency, or F0; formant transitions) and qualitative descriptors (e.g., phonetic categories). A *quasi-resonant nucleus* is produced with a closed throat and little breath support, and is characterized by creakiness, nasality, or both. A *fully resonant nucleus* is produced with an open throat and normal phonation, yielding a clear formant structure. A *marginal syllable* is a slow sequence of CV articulation, with a long (> 250 ms) transition between consonant and vowel. The vowel is often distorted because of slow movement of the articulators. A *canonical syllable* has a fast (< 250 ms) CV articulation with a fully resonant vowel. Quasi-resonant nuclei and marginal syllables are immature forms. Fully resonant nuclei and canonical syllables are mature forms typical of adultlike speech. We tested 9.5-month-olds because they typically have all four vocalization types in their repertoires.

Vocalizations were coded by the same coders described in Maternal Responses, following the same training procedure. Reliability for coding vocal types was calculated for 20% of the sample. Reliability was .94 (range: .87–1.0).

We tallied each infant's vocalizations in each experimental period. We also calculated three proportions to characterize each infant's vocalization quality: (a) ratio of fully resonant vocalizations to total number of vocalizations, (b) ratio of vocalizations with CV structure to total number of vocalizations, and (c) ratio of canonical syllables to total number of CV-structured syllables. To assess the effects of contingent and yoked interactions on infants' affective state, we coded amount of fussing or crying, looks to mother, and smiles to mother. Smiles were coded when infants raised the corners of their lips by moving the zygomatic muscle (Jones, Collins, & Hong, 1991).

Imitation

Because infants produced vocalizations with phonological patterns that resembled those of their mothers' speech, we tested whether changes in babbling could have resulted from imitation of mother's phonemes. A trained coder (the second author) used the International Phonetic Alphabet to phonetically transcribe contingent mothers' responses to infants' vocalizations during the social-response period. The infant vocalization following each maternal response was also transcribed. Additionally, a second trained coder transcribed 20% of the sessions (three from each contingent condition). Intercoder agreement was 92.6% (range: 86.9%–95.5%) for mothers' vocalizations and 89.5% (range: 83.8%–94.1%) for infants' vocalizations.

To calculate matching during the social-response period, we examined each pair of utterances (i.e., a maternal utterance and subsequent infant vocalization) and recorded the number of matches (infant's phonemes identical to mother's phonemes). The matching score for each pair was calculated as a proportion, with the number of matches divided by the number of phonemes that the infant produced. For example, if a mother said "Piglet," /pɪglɪt/, an infant's utterance /ledele/ would receive a matching score of 33% (one of three unique phonemes matched), and an utterance of /gɪde/ would receive a matching score of 50% (two of four phonemes matched). This method takes into account infants' partial matches of their mothers' utterances. Repetitions of a matching phoneme did not change the matching score.

To evaluate the likelihood that an infant would match a mother's sounds by chance, we created scrambled transcripts for each dyad by randomizing the sequence of infant vocalizations in the social-response period. Matching was then calculated for the scrambled transcripts. To increase the likelihood of detecting mimicry, we analyzed only mothers' one-word responses. Including longer, multiword utterances decreased the likelihood that infant matching exceeded chance levels.

RESULTS

We designed the experiment to assess learning within subjects. Thus, we used planned comparisons to measure changes across the three experimental periods.

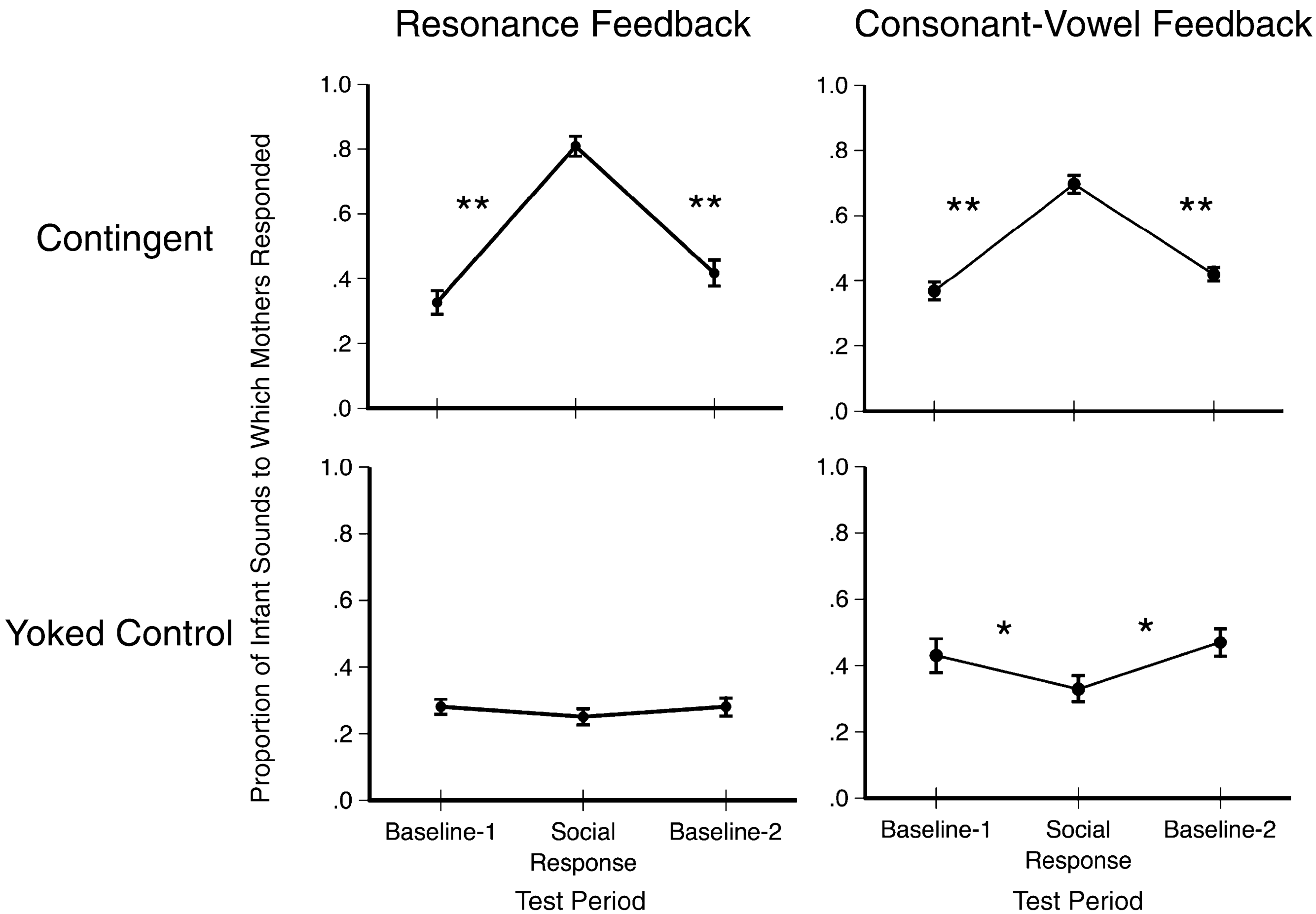

Fig. 1. Mean proportion (± 1 *SE*) of infants' sounds to which mothers responded in each 10-min test period. Data are reported for the contingent and yoked control conditions in both feedback groups. Asterisks indicate a significant change from one test period to the next $^{*}p < .05$, $^{**}p < .01$.

Maternal Responses

Mothers were able to follow our instructions to respond to their infants (Fig. 1). Contingent-resonance mothers changed their responsiveness across the three test periods, $F(2, 28) = 39.82$, $p < .001$, $p_{\mathrm{rep}} > .99$, $\eta_p^2 = .74$. They showed an increase in responses from Baseline 1 to the social-response period and a decrease in responses from the social-response period to Baseline 2 (Tukey's *HSD*, $p < .01$). Mothers in the contingent-CV group showed the same pattern of responding, $F(2, 28) = 64.19$, $p < .001$, $p_{\mathrm{rep}} > .99$, $\eta_p^2 = .82$ (Tukey's *HSD*, $p < .01$). When contingent-resonance mothers spoke to their infants during the social-response period, they used multiple vowel types ($M = 10.33$, $SD = 3.09$, range: 4–14). When contingent-CV mothers spoke, they used multiple consonant types ($M = 17.67$, $SD = 4.03$, range: 10–23) and vowel types ($M = 11.67$, $SD = 1.59$, range: 9–15). Most of their words contained CV alternation ($M = 93.5\%$, $SD = 5.4\%$, range: 83.6%–100%).

Mothers in the yoked-resonance group did not show a significant change in responses across the test periods, $F(2, 28) = 0.80$, $p = .46$ (Fig. 1). Yoked-CV mothers showed a significant decrease in responsiveness during the social-response period, $F(2, 28) = 7.51$, $p = .002$, $p_{\mathrm{rep}} = .98$, $\eta_p^2 = .35$ (Tukey's *HSD*, $p < .05$). Changes in their behavior from Baseline 1 to the social-response period did not have a negative impact on infants' behavior. Yoked control and contingent-condition infants did not differ in amount of fussing or crying, looks to mother, or smiles to mother (assessed by t tests, all $ps > .25$). In addition, attrition due to fussiness was comparable in these two groups (contingent condition: $n = 9$, yoked control: $n = 5$).

Phonology of Infants' Vocalizations

In the contingent-resonance group, infants' vocalizations became more fully resonant across the test periods, $F(2, 28) = 3.65$, $p = .039$, $p_{\mathrm{rep}} = .89$, $\eta_p^2 = .21$ (Fig. 2). Infants' vocalizations improved from Baseline 1 to the social-response period (Tukey's *HSD*, $p < .05$), but did not decrease in quality from the social-response period to Baseline 2. Vocalizations of infants in the contingent-resonance group did not change in the proportion of vocalizations with CV structure, $F(2, 28) = 1.22$, $p = .31$ (Fig. 2), or the proportion of CV syllables with canonical form, $F(2, 22) = 0.14$, $p = .87$.

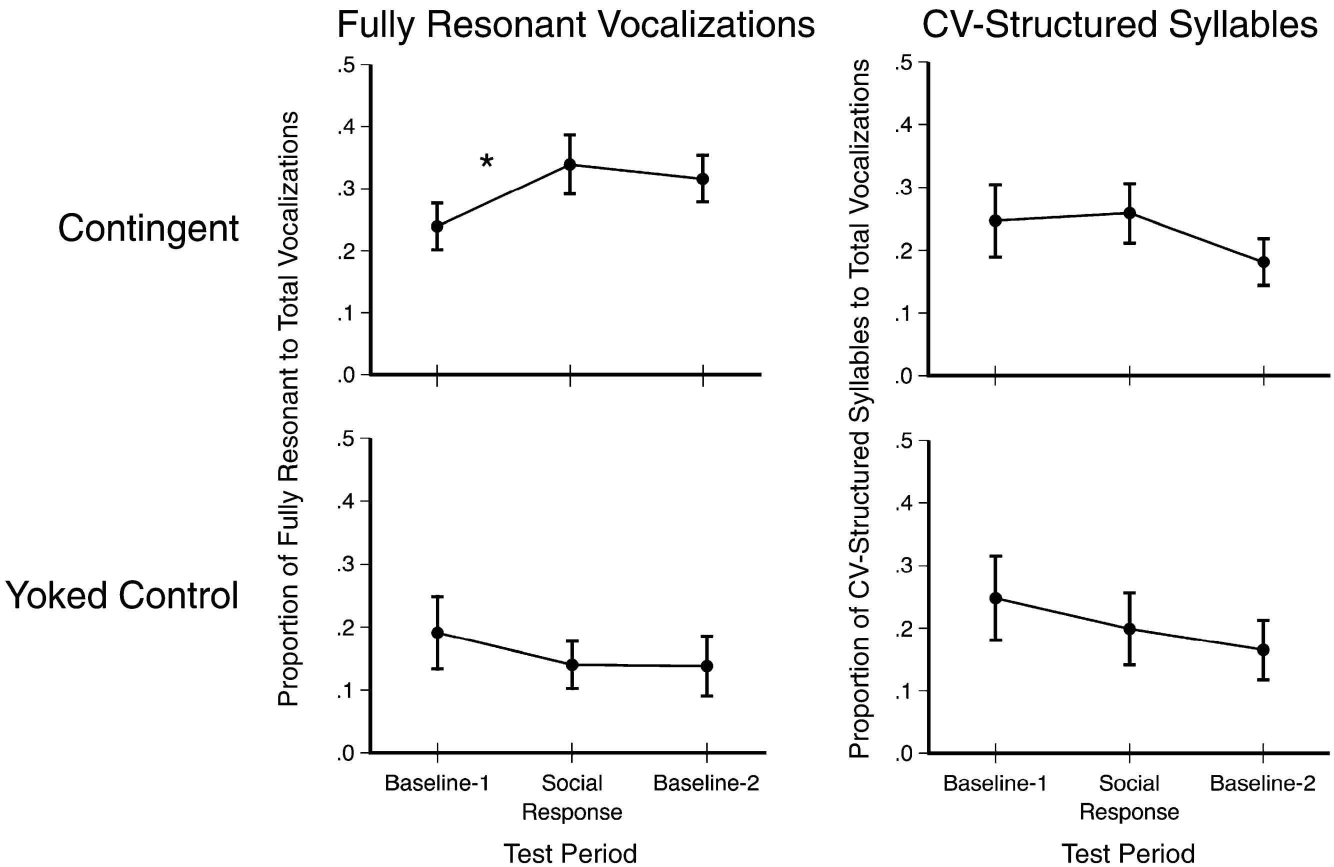

Fig. 2. Changes in vocalization quality for infants in the contingent and yoked control conditions who received feedback structured as fully resonant vowels (resonance feedback group). Graphs in the left column indicate the ratio (mean ±1 *SE*) of fully resonant to total vocalizations across the three 10-min test periods. Graphs in the right column indicate the ratio (mean ±1 *SE*) of syllables with consonant-vowel (CV) structure to total vocalizations across the three 10-min periods. The asterisk indicates a significant change between test periods, $p < .05$.

In the contingent-CV group, infants increased the proportion of vocalizations with CV structure across the test periods, $F(2, 28) = 4.60$, $p = .019$, $p_{\text{rep}} = .92$, $\eta_p^2 = .25$ (Fig. 3). Infants' vocalizations improved from Baseline 1 to the social-response period (Tukey's *HSD*, $p < .05$), but did not decrease in quality from the social-response period to Baseline 2. The proportion of vocalizations with full resonance did not change for the contingent-CV infants, $F(2, 28) = 0.49$, $p = .62$ (Fig. 3), nor did the proportion of CV syllables with canonical form, $F(2, 16) = 0.38$, $p = .69$.

Infants in the yoked control groups did not show changes in resonance, the proportion of syllables with CV structure, or the proportion of CV-structured syllables with canonical form, *F*s < 2.75, *p*s > .10 (see Figs. 2 and 3).

Imitation

Because we observed vocal learning in the contingent conditions, we examined vocal imitation within the dyads in these conditions. We analyzed 706 pairs of utterances. In the contingent-resonance group, infants matched, on average, 29.3% of the phonemes in their mothers' previous utterances (*SD* = 13.8%). This did not differ from the amount of matching in the scrambled transcripts for this group (*M* = 26.6%, *SD* = 13.3%), $t(14) = 1.67$, $p = .12$. In the contingent-CV group, infants matched, on average, 15.0% of their mothers' sounds (*SD* = 13.3%). This did not differ from the amount of matching in the scrambled transcripts for this group (*M* = 13.1%, *SD* = 10.3%), $t(14) = 0.89$, $p = .39$. Matching also did not differ between original and scrambled pairs when the contingent conditions were combined (original: *M* = 22.2%, *SD* = 15.2%; scrambled: *M* = 19.8%, *SD* = 13.5%), $t(29) = 1.74$, $p = .10$.

Number of Vocalizations by Infants

Infants in the contingent-resonance group showed a significant change in amount of vocalizing across the three test periods, $F(2, 28) = 8.00$, $p = .002$, $p_{\text{rep}} = .98$, $\eta_p^2 = .36$ (Fig. 4). They increased their number of vocalizations from Baseline 1 to the social-response period (Tukey's *HSD*, $p < .05$), then decreased their vocalizations during Baseline 2 (Tukey's *HSD*, $p < .01$). Infants in the contingent-CV group did not show a significant change in amount of vocalizing across the test periods, $F(2, 28) = 0.78$, $p = .47$ (Fig. 4).

Infants in the yoked-resonance group increased their vocalizations across the three test periods, $F(2, 28) = 3.90$, $p = .032$, $p_{\text{rep}} = .91$, $\eta_p^2 = .22$ (Fig. 4). Their amount of vocalizing

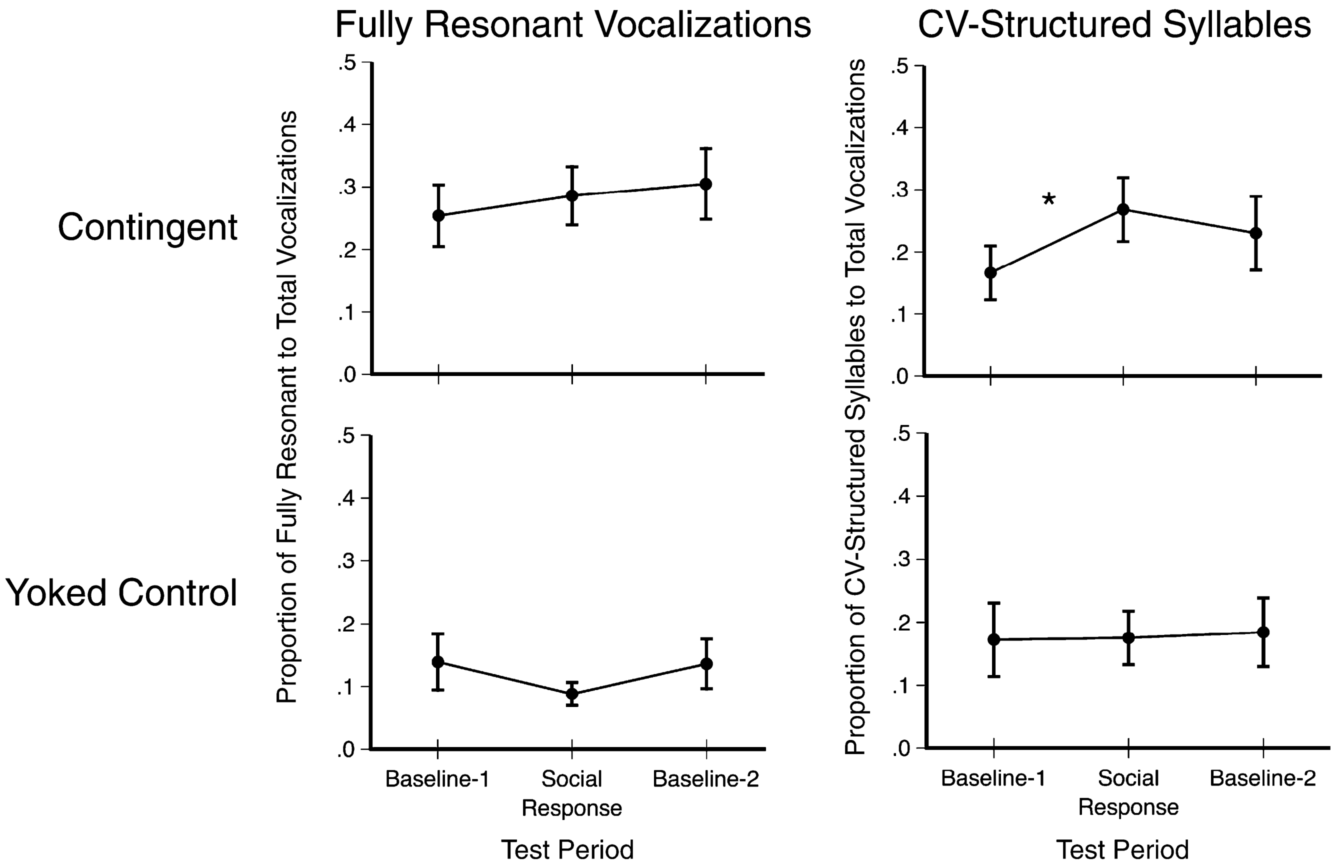

Fig. 3. Changes in vocalization quality for infants in the contingent and yoked control conditions who received feedback structured as consonant-vowel (CV) syllables (CV feedback group). Graphs in the left column indicate the ratio (mean ± 1 *SE*) of fully resonant to total vocalizations across the three 10-min test periods. Graphs in the right column indicate the ratio (mean ± 1 *SE*) of CV-structured syllables to total vocalizations across the three 10-min periods. The asterisk indicates a significant change between test periods, $p < .05$.

increased from Baseline 1 to the social-response period (Tukey's *HSD*, $p < .05$). Infants in the yoked-CV group showed the same pattern, $F(2, 28) = 3.85$, $p = .033$, $p_{rep} = .90$, $\eta_p^2 = .22$ (Tukey's *HSD*, $p < .05$).

DISCUSSION

In summary, infants modified their babbling in accordance with the phonological structure present in their caregivers' contingent utterances. Comparison with baseline levels showed that infants in the contingent-resonance condition increased their proportion of fully resonant vowels, and infants in the contingent-CV condition increased their proportion of CV-structured syllables. Yoked control infants, who received identical but noncontingent feedback, did not change the phonological characteristics of their babbling, though they increased their amount of vocalizing during the social-response period. The increase in their vocalizing was likely due to increased arousal from noncontingent reinforcement (Domjan, 1993). Only infants in the contingent-CV condition did not increase their number of vocalizations from Baseline 1 to the social-response period. These infants learned a more complex production rule; thus, their utterances were longer than those of the other groups.

In the contingent-resonance and contingent-CV conditions, changes in the infants' vocalizations occurred at the level of phonological patterns, rather than only at the phonetic level. Though these infants produced sounds with more resonant vowels or CV syllables, they did not do so by producing the same phonemes their mothers modeled. Thus, these infants did not learn new sounds by mimicking the surface features of their mothers' speech. We examined whether infants' immature articulatory skills prevented accurate mimicry. Nine-month-old infants have a large vocal repertoire (Oller, 2000), and almost all of the phonemes used by the mothers in our study were sounds that infants of this age can produce. In addition, infants of this age produce multisyllable utterances. Our calculation of imitation gave infants credit for even partial matches with their mothers' speech. Thus, infants' mimicry was not constrained by a lack of articulatory competence. The infants may have imitated at a more abstract level of speech, such as the category "fully resonant vowel" or "CV combination" (it has been observed that language learning requires infants to perceive linguistic patterns at levels higher than the phoneme level; Chomsky, 1957).

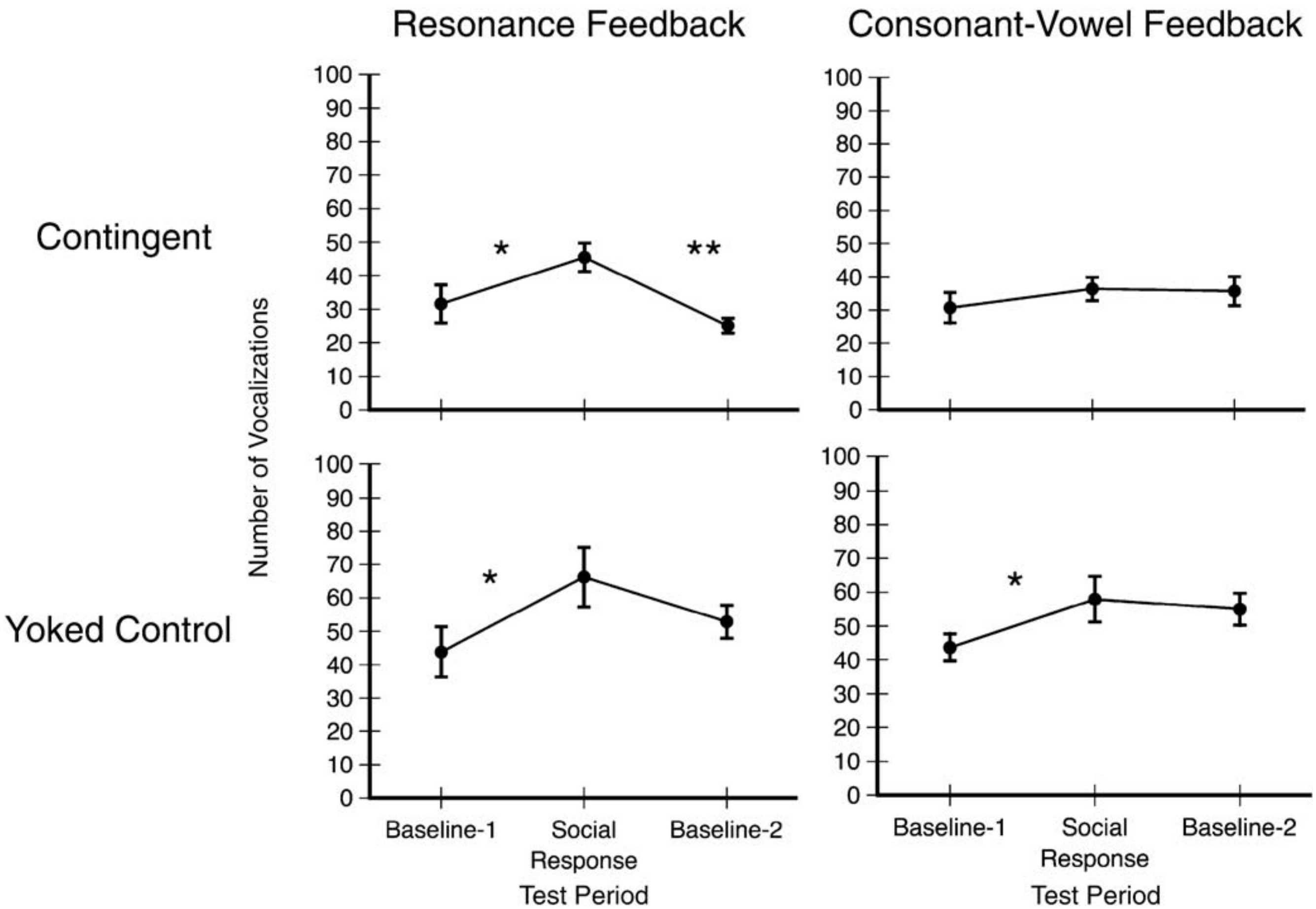

Fig. 4. Mean number (±1 *SE*) of infant vocalizations produced in each 10-min test period. Data are reported for the contingent and yoked control conditions in both feedback groups. Asterisks indicate a significant change from one test period to the next, $^*p < .05$, $^{**}p < .01$.

A prerequisite for such imitation, however, is ability to learn phonological patterns, an ability demonstrated in the present study.

What mechanisms allow such rapid vocal learning from social feedback? Our results indicate that the infants acquired the phonological patterns in their mothers' speech via socially guided statistical learning. Previous research on the perception of statistical structure in language has shown that 9-month-old infants can learn phonological patterns when given only minutes of exposure (Saffran & Thiessen, 2003). Infants' perception of an underlying phonological pattern is influenced by the variation of the elements presented as instances of the pattern. For example, when learning an artificial grammar containing nonadjacent dependencies between the first and third elements, infants require a high level of variability of the middle, adjacent element in order to extract the higher-order rule linking the first and last elements (Gómez, 2002). In another study, infants exposed to variable input acquired a rule that governed an underlying pattern in the utterances, rather than the surface features (Marcus, Vijayan, Bandi Rao, & Vishton, 1999). In general, infants show greater generalization of learning when given more variability in training (Rovee-Collier & DuFault, 1991). Infants given input without any variation (e.g., repeated presentations of a single vowel) produced vocalizations that closely resembled those of an adult model (Kuhl & Meltzoff, 1996). In the present study, mothers produced a diversity of phonemes when providing examples of fully resonant vowels or CV alternation, and in doing so likely made the underlying phonological patterns more salient to their infants. However, the mere presence of phonological patterns was not enough to facilitate vocal learning, as demonstrated by the lack of learning in the yoked control condition.

Why did learning require contingent feedback in this study? All previous studies of auditory statistical learning used perception tasks, but our study used vocal production as a measure of learning of phonological patterns. A production task is an inherently more difficult task than a perception task, because the learner must create new utterances that obey phonological rules, rather than recognize an existing legal utterance. Kuhl (2007) has argued that social contingency benefits infants' learning because it helps them focus attention and provides richer sources of information. In the contingent conditions of our study, infants' and mothers' utterances were close together in time, which may have made disparities between the infants' productions and the following adults' productions more salient. The present study demonstrates the influence of contingent social feedback on infants' learning of specific phonological patterns. The contingent structure of social feedback served to highlight statistical regularities in mothers' speech. Infants recognized these phonological regularities and used them to guide the production of new vocal forms.

Models of unsupervised language learning (e.g., Solan, Horn, Ruppin, & Edelman, 2005) discover underlying structure by examining input for utterances containing partial redundancy. Over time, patterns are distilled from this redundancy. The discovery of structure in the input would be greatly facilitated if the relevant utterances had reliable temporal contiguity. We propose that infants use a similar mechanism of correlation between their own and other speakers' utterances to structure their vocalizations along the most salient dimensions of adult speech. Current studies of vocal learning are testing this hypothesis in our laboratory.

Findings from the present study have important implications for understanding the neural substrates of language learning. The strong effect of social feedback on vocal learning suggests that brain structures that mediate feedback-driven learning play a role in the acquisition of speech. In studies of phonological learning in adults, feedback on performance is related to activation of the caudate nucleus, part of the basal ganglia (Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006). Studies of songbirds' vocal development have found that the anterior frontal pathway, which is similar to the basal ganglia in humans (Doupe & Kuhl, 1999), is a crucial component for vocal learning (Brainard & Doupe, 2000). Thus, in humans, there may be specific physiological mechanisms that relate social feedback to facilitation of speech and language learning. The finding of feedback-driven vocal learning indicates that the basal ganglia should be a focus of neurobiological investigation of speech and language development.

Prerequisites for language come from multiple sources, such as speech perception and social development (Kuhl, 2003, 2007; Locke, 2001; Snow, 1988). Data from this study, along with earlier data from our laboratory (Goldstein et al., 2003), show that prelinguistic vocal production is sensitive to social feedback. In addition, caregivers provide contingent responses to babbling that are sensitive to the acoustic quality of infants' sounds and the context in which the sounds are occurring (Goldstein & West, 1999; Gros-Louis, West, Goldstein, & King, 2006). In our view, infants' prelinguistic vocalizations, and caregivers' reactions to those immature sounds, create opportunities for social learning that afford infants knowledge of phonology. Socially guided learning is thus an important mechanism in early vocal development, laying the foundation for advances in communication and language. We conclude that immature, prelinguistic vocalizations have important functional significance. The way caregivers talk when reacting to an infant's babbling has immediate consequences for vocal learning. Babbling, when studied in social context, constitutes a crucial and formative phase in language development.

Acknowledgments—We thank Shimon Edelman, James Cutting, Steven Robertson, Susan Jones, Martin Pickering, and two anonymous reviewers for their helpful comments on earlier versions of this manuscript. Jacquelyn Briesch, Erica Moran, Angela Narayan, Jessie Northrup, Jennifer Stone, Supriya Syal, Laura Vollmer, Anya Wachterhauser, and Veronika Zeppenfeld provided invaluable assistance with testing infants and coding vocalizations. We also thank the families who participated in the study.

REFERENCES

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, *14*, 211–227.

Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In B. Boysson-Bardies, S. Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 353–363). Dordrecht, The Netherlands: Kluwer.

Brainard, M.S., & Doupe, A.J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature*, *404*, 762–766.

Chomsky, N. (1957). *Syntactic structures*. The Hague, The Netherlands: Mouton.

Domjan, M. (1993). *The principles of learning and behavior* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Doupe, A.J., & Kuhl, P.K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, *22*, 567–631.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5, Serial No. 242), 1–173.

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

Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

Gros-Louis, J.G., West, M.J., Goldstein, M.H., & King, A.P. (2006). Mothers provide differential feedback for infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*, 509–516.

Jones, S.S. (2007). Imitation in infancy: The development of mimicry. *Psychological Science*, *18*, 593–599.

Jones, S.S., Collins, K., & Hong, H. (1991). An audience effect on smile production in 10-month-old infants. *Psychological Science*, *2*, 45–49.

Jusczyk, P. (1992). Developing phonological categories from the speech signal. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17–64). Timonium, MD: York Press.

Kent, R.D. (1981). Articulatory-acoustic perspectives on speech development. In R.E. Stark (Ed.), *Language behavior and early childhood* (pp. 105–126). New York: Elsevier.

Kuhl, P.K. (2003). Human speech and birdsong: Communication and the social brain. *Proceedings of the National Academy of Sciences, USA*, *100*, 9645–9646.

Kuhl, P.K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, *10*, 110–120.

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.

Locke, J.L. (2001). First communion: The emergence of vocal relationships. *Social Development*, *10*, 294–308.

Marcus, G.F., Vijayan, S., Bandi Rao, S., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.

Masataka, N. (2003). *The onset of language*. Cambridge, England: Cambridge University Press.

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

Rovee-Collier, C., & DuFault, D. (1991). Multiple contexts and memory retrieval at three months. *Developmental Psychobiology*, *24*, 39–49.

Saffran, J.R., & Thiessen, E.D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*, 484–494.

Snow, C.E. (1988). The last word: Questions about the emerging lexicon. In M.D. Smith & J.L. Locke (Eds.), *The emergent lexicon* (pp. 341–353). San Diego, CA: Academic Press.

Solan, Z., Horn, D., Ruppin, E., & Edelman, S. (2005). Unsupervised learning of natural languages. *Proceedings of the National Academy of Sciences, USA*, *102*, 11629–11634.

Stark, R.E. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology* (pp. 73–90). New York: Academic Press.

Tricomi, E., Delgado, M.R., McCandliss, B.D., McClelland, J.L., & Fiez, J.A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience*, *18*, 1029–1043.

Werker, J.F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*, 197–234.

(Received 6/28/07; Revision accepted 11/5/07)