

Human speech and birdsong: Communication and the social brain

Patricia K. Kuhl*

Center for Mind, Brain, and Learning, Department of Speech and Hearing Sciences, University of Washington, Mailstop 357920, Seattle, WA 98195

At first glance, communication in babies and birds appears to have little in common. Babies' babblings become words and sentences, whereas birds' initial notes become species-typical songs. However, a comparison of the ontogeny of communicative repertoires in human infants and avian song birds shows striking parallels (1) (Fig. 1). Both species show initial innate predispositions for species-typical signals and have to be exposed to species-typical vocalizations during a sensitive period to acquire them. Both show an initial phase of developmental learning that is primarily perceptual, during which babies and birds commit to long-term memory the detailed characteristics of the communicative repertoires they hear. Both species subsequently use the patterns stored in memory to guide motor production through the process of imitation (2–4). For babies, the task involves committing to memory the phonetic units and prosodic (pitch and intonation) characteristics that typify the mother tongue: Japanese is not French. Birds store the specific notes, syllables, and prosodic characteristics that typify their species. Storing the species-typical patterns is not the end of the task for either species. Both babies and birds must then rehearse and refine their communicative repertoires, actively comparing and gradually matching (through auditory feedback) their productions to the sound patterns stored in auditory memory (5, 6).

The article by Goldstein *et al.* (7) in a recent issue of PNAS provides support for yet another fascinating link between communicative development in babies and birds. Goldstein *et al.* show that social contingency strongly influences babbling in human infants. Although well established in certain songbirds, experimental evidence that social cues affect the development of speech in human infants is novel and important.

Goldstein *et al.* (7) show that social feedback makes a difference in both the quantity and quality of the utterances that young infants produce. In the study, mothers' responsiveness to their infants' vocalizations was manipulated. After a baseline period of normal interaction, half of the mothers were instructed to respond immediately to their infants' vocalizations by smiling at, moving closer to, and touching their infants; they were the "contingent

condition" (CC) mothers. The other half of the mothers were "yoked controls" (YC); their reactions were identical, but timed (by the experimenter's instructions) to coincide with vocalizations of infants in the CC group. Infants in the CC group therefore experienced contingent social reactions to their vocalizations, whereas infants in the YC group experienced an equal amount of social stimulation, but not timed in response to their vocalizations. The results demonstrated that infants in the CC group not only produced more vocalizations than infants in the YC group, but their vocalizations were more mature and adult-like when compared with those of the YC group. CC infants produced a greater proportion of fully resonant syllables with canonical consonant-vowel structure, both indicative of advanced speech production, when compared with the YC infants. The complexity of babbling was thus modulated by social cues.

As the authors note, social contexts can advance song production in birds; male cowbirds respond to the social gestures and displays of females, which affect the rate, quality, and retention of song elements in their repertoires, and white-crowned sparrow tutors provide acoustic feedback that affects the repertoires of young birds. Additional evidence of the impact of social interaction can be adduced from the sensory learning phase in birds. In young zebra finches, visual interaction with a tutor bird is typically required to learn (8); in fact, the impact of social interaction is so potent that young zebra finches will learn an alien song from a Bengalese finch foster father who feeds them (9). White-crown sparrows will also learn an alien song from a live tutor, even though they reject those songs when presented on audio tape. In barn owls (10) and in white-crowned sparrows (11), the duration of the sensitive period for learning is altered by the richness of their social environments.

Of interest in the Goldstein *et al.* (7) case will be studies that examine whether contingency itself, in the absence of a human being, plays a role in inducing greater frequency and complexity of vocalizations. Would a contingently delivered nonhuman stimulus show the same effect? In birds, effective interactions can take a variety of forms. If young zebra

finches are operantly conditioned to present conspecific song to themselves by pressing a key, song learning occurs (12, 13), suggesting that active participation and the attention it produces may be critical. In human infants, contingency itself may be an important component contributing to the effectiveness of social interaction.

What do we know about the impact of social interaction on infant language learning? We know, from the (thankfully few) instances in which children have been raised in social isolation, that social deprivation has a severe and negative impact on language development; normal language skill is never acquired (14). We also know that, in children with autism, language and social deficits are tightly coupled (15).

In typically developing children, the effects of a social partner are positive and bidirectional. Reciprocity in adult–infant language can be seen in infants' very early vocal "turn-taking," their tendency to alternate their vocalizations with those of an adult (3, 16). Adults not only vocalize in front of babies but speak in a special vocal register, often called "motherese," which has a unique acoustic signature (17) and exhibits greater clarity in the individual phonetic units (18, 19). Infants prefer this kind of speech. Given a choice in an experimental setting, they listen longer to it (20), and there is recent evidence suggesting that it may assist language learning (21). Infants' early social awareness is itself a predictor of later language skill. Measured by infants' tendencies to follow the gaze of an adult in a communicative setting (22), early social awareness predicts advanced later word comprehension and word production (23). Our recent work verifies the profound importance of social interaction on human language learning. In a recent study, we exposed 9-month-old American infants to a foreign language (Mandarin Chinese) spoken by native speakers who read books and played with toys. The infants were later tested on phonetic units contained in Mandarin but not in English. Infants learned phonetic units from the foreign language with <5 h of experience, but,

See companion article on page 8030 in issue 13 of volume 100.

*E-mail: pkkuhl@u.washington.edu.

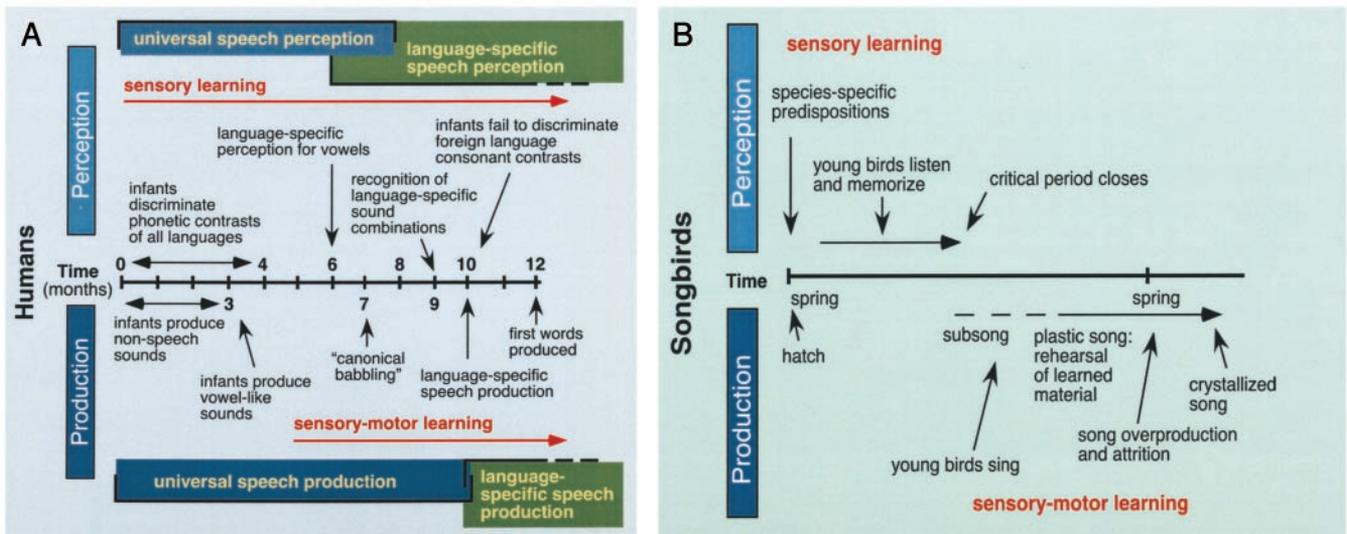


Fig. 1. Timelines for speech development in infants and song development in birds. Both species show innate predispositions for the perception of species-typical signals, and periods of sensory learning followed by periods of sensory-motor learning. [Reproduced with permission from ref. 1 (Copyright 1999, Annual Reviews).]

interestingly, they failed to learn when the same speakers using the same materials were prerecorded and presented from audiovisual or audio-only DVDs (24).

Taking all of the data into account, there is a very tight coupling between language and social cognition. The questions are why and how.

Why do social factors affect language acquisition? A general answer is that language evolved to address a social need: to communicate with a specific listener. There is ample evidence that, as speakers, we unconsciously make subtle adjustments in speech to take our audience into account. These are not the practiced tactics of speech-makers, but the adjustments all speakers unconsciously make when they, for example, automatically increase the loudness level of speech in a noisy environment (the well known “Lombard effect”) or adjust the prosody, clarity, and complexity of speech when addressing different people (motherese is one exam-

ple) (25). Speakers are exquisitely tuned to their listeners’ needs and make adjustments to accommodate them. Infants in the present experiments appear to be doing just that. When parents are attentive, infants produce the most sophisticated utterances in their repertoires. Perhaps this is an example of infants’ nascent recognition of an audience.

Even if we accept that social interaction matters because language evolved in a social setting, the mechanism that controls the interface between language and social cognition remains a mystery. One possibility is that the effects of social environments are broad, general, and “top-down.” People (adults as well as infants) engaged in social interaction are highly aroused and attentive; general arousal mechanisms could enhance our abilities to store and remember information, as well as prompt the most sophisticated output we have in our action repertoires. Hormones could be the mediators, and they

have been implicated in learning and song production in birds (26). On the other hand, more specific “bottom-up” mechanisms could also be at work, mechanisms attuned to the particular form and content of contingent social cues and/or feedback.

Language is often viewed as an isolated “modular” ability, one that is separated from other, more general human systems (27, 28). Goldstein *et al.* (7) and the additional evidence cited here suggest that language emerges in infants by relying on a broader set of perceptual, cognitive, and social skills. Further studies will be needed to advance our understanding of the genetic, neurobiological, and anatomic components that link our language brains to our social brains. The current data suggest that this will be a fruitful line of research. Human language provides an opportunity to study the interface between systems that control the acquisition of complex behavioral repertoires in natural social settings.

- Doupe, A. J. & Kuhl, P. K. (1999) *Annu. Rev. Neurosci.* **22**, 567–631.
- Konishi, M. (1985) *Annu. Rev. Neurosci.* **8**, 125–170.
- Kuhl, P. K. & Meltzoff, A. N. (1996) *J. Acoust. Soc. Am.* **100**, 2425–2438.
- Marler, P. (1970) *Am. Sci.* **58**, 669–673.
- Konishi, M. (1965) *Z. Tierpsychol.* **22**, 770–783.
- Kuhl, P. K. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11850–11857.
- Goldstein, M. H., King, A. P. & West, M. J. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 8030–8035.
- Eales, L. A. (1989) *Anim. Behav.* **37**, 507–508.
- Immelmann, K. (1969) in *Bird Vocalizations*, ed. Hinde, R. A. (Cambridge Univ. Press, London), pp. 61–74.
- Brainard, M. S. & Knudsen, E. I. (1998) *J. Neurosci.* **18**, 3929–3942.
- Baptista, L. F. & Petrinovich, L. (1986) *Anim. Behav.* **34**, 1359–1371.
- Adret, P. (1993) *Anim. Behav.* **46**, 149–159.
- Tchernichovski, O., Mitra, P. P., Lints, T. & Nottebohm, F. (2001) *Science* **291**, 2564–2569.
- Fromkin, V., Krashen, S., Curtiss, S., Rigler, D. & Rigler, M. (1974) *Brain Lang.* **1**, 81–107.
- Dawson, G., Webb, S., Schellenberg, G., Dager, S., Friedman, S., Aylward, E. & Richards, T. (2002) *Dev. Psychopathol.* **14**, 581–611.
- Kuhl, P. K. & Meltzoff, A. N. (1982) *Science*, **218**, 1138–1141.
- Fernald, A. & Simon, T. (1984) *Dev. Psych.* **20**, 104–113.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U. & Lacerda, F. (1997) *Science* **277**, 684–686.
- Burnham, D., Kitamura, C. & Vollmer-Conna, U. (2002) *Science* **296**, 1435–1435.
- Fernald, A. & Kuhl, P. (1987) *Inf. Behav. Dev.* **10**, 279–293.
- Liu, H.-M., Kuhl, P. K. & Tsao, F.-M. (2003) *Dev. Sci.* **6**, F1–F10.
- Brooks, R. & Meltzoff, A. (2002) *Dev. Psych.* **38**, 958–966.
- Baldwin, D. A. (1995) in *Joint Attention: Its Origins and Role in Development*, ed. Moore, C. & Dunham, P. J. (Lawrence Erlbaum Associates, Hillsdale, NJ), pp. 131–158.
- Kuhl, P. K., Tsao, F.-M. & Liu, H.-M. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 9096–9101.
- Kuhl, P. K., Tsao, F.-M., Liu, H.-M., Zang, Y. & de Boer, B. (2001) *Ann. N.Y. Acad. Sci.* **935**, 136–174.
- Nottebohm, F. (1999) in *The Design of Animal Communication*, eds. Hauser, M. D. & Konishi, M. (MIT Press, Cambridge, MA), pp. 63–110.
- Fodor, J. A. (1983) *Modularity of Mind: An Essay on Faculty Psychology* (MIT Press, Cambridge, MA).
- Lieberman, A. M. & Mattingly, I. (1985) *Cognition* **21**, 1–36.