

Cambridge University Press  
978-0-521-86915-7 - Neuroscience of Birdsong  
Edited by H. Philip Zeigler and Peter Marler  
Excerpt  
[More information](#)

---

## PART I

### Foundations: singing and the brain

# 1 • Introduction

H. P. Zeigler

The chapters in this introductory section focus on two central topics: singing behavior itself and the songbird brain. As Williams points out, what was always a source of aesthetic pleasure for bird watchers has developed into a major area of neurobiological research. Researchers have come to recognize that this apparently simple, “innate” behavior is the product of an ontogenetic learning process; that it is limited to only a very few vertebrate species (including humans) and serves a critical communicative function. Indeed, the initial attraction of birdsong for neuroscientists was as a model system for the study of learned vocalizations, analogous to speech acquisition in humans. For this reason the section opens with a somewhat abridged version of a classic paper by Doupe and Kuhl comparing birdsong and human speech.

The Williams chapter provides an introduction to the essential features of birdsong as a learned vocalization, including those features that qualify it as a model system for studies of the development of human language. But her main focus is on the adult song, the remarkable degree of variability it exhibits and the importance of the social context in which it occurs. She notes that because song is linked to two adaptive behaviors – territory defense and mate attraction – it has evolved specifically to influence listeners’ behavior. Thus, while the stereotype of song is often viewed as one of its defining features, the presence and behavior of conspecific “listeners” can significantly affect both what is sung (song structure) and how (amplitude and tempo). She also calls attention to the role of the visual displays associated with song, an often neglected component of singing behavior. While the links between singing and breathing are already the focus of considerable research effort (chapters by Wild, Suthers and Zollinger, Goller and Cooper, Schmidt and Ashmore), little is known as to the neural circuitry which integrates vocal and visual signals. What is known, however, is that the responsiveness of these circuits is modulated by hormones, and the interplay between neural, hormonal and sensory factors in modulating singing behavior is one

of the central themes of this volume (see chapters by Harding, Ball *et al.*, and Brenowitz).

Importantly, Williams emphasizes that much of our knowledge of songbird neurobiology is based upon data from a small number of species selected primarily for their amenability to study in the laboratory (zebra finches, song sparrows, canaries). Singing behavior in these species has a clear reproductive function, is linked specifically to the breeding season and is performed exclusively by males. Moreover, differences in singing behavior between males and females of these species are correlated with differences in the brain circuits mediating singing behavior. As Williams points out, it is all too easy to generalize from male zebra finches (or canaries) to “birds.” Thus, our conceptualization of “song” has come to be defined by the song of a small number of avian species, and male brains and behavior have come to be the primary focus of research on brain mechanisms of song learning.

Kaplan’s chapter on the Australian magpie provides a thought-provoking challenge to this conceptualization. As described by Kaplan magpie song is complex and its song repertoire enormous, including a capacity for mimicry. However, since both males and females sing, their song circuit structures are of comparable size and song has no obvious reproductive function, magpie vocalizations do not fit our current conceptualization of song. Kaplan argues that the features we think of as defining song, especially sexual dimorphism and breeding-related song, are characteristic of northern hemisphere birds, while the singing behavior of the Australian magpie is, in fact, representative of a very large and diverse range of songbird species, including tropical birds. From this standpoint, she contends, it is certainly as appropriate a model as zebra finch or canary song. Moreover, since in neither humans nor magpies is vocal learning an *exclusively* reproductive behavior, it may even be a more appropriate model for comparative neurobiological studies of language development.

Both Williams and Kaplan would probably agree that the concept of a canonical stereotyped learned song, as in zebra finches, is useful as a tool for studying neural processes and that the ability of a species to thrive in captivity is a useful feature for a model. However, they would probably also agree that the field would certainly benefit from exposure to the widest possible array of vocalization types, including those, like the magpie, that function as exceptions which “prove” (test) the rule.

The function of a model, of course, is to provide access to critical features of a natural phenomenon for experimental analysis. Given the apparent parallels between language development in humans and vocal learning in birds, the extent to which these processes share similar neural substrates in the two groups has been a critical issue for the birdsong community. If the brains of the two groups were obviously different it would be hard to justify the use of birdsong as a model system that might teach us something about the brain mechanisms underlying human language. While much of the brain of birds (from spinal cord to midbrain) reflects an organization common to most vertebrates, the “higher” brain regions, including the forebrain, are clearly different. The avian forebrain is made of large nuclear cell masses and its morphology resembles that of the mammalian basal ganglia (striatum); the mammalian forebrain, while it contains obvious striatal structures, is defined by its unique possession of a laminated isocortex (“grey matter”) separated from the underlying basal ganglia by a band of myelinated axons (“white matter”). Moreover, the obvious increase in cortical development from primitive mammals to *homo sapiens* seemed to be correlated with increased possession of that concatenation of abilities grouped colloquially under the name of “intelligence” – something that birds were thought to

lack. The absence of a cortex thus became a defining negative taxonomic feature of the avian brain while the notion that the avian forebrain was essentially a “striatal” structure was reflected in the nomenclature assigned to forebrain structures (*hyperstriatum*, *neostriatum*, *archistriatum*, *palaeostriatum*). The widespread use of the term *birdbrain* as a derogatory epithet reflects the view widely held among early twentieth-century comparative anatomists that the avian forebrain represents an evolutionary dead end. Despite the fact that there is no evidence that the organization of forebrain cells into lamina provides any unique advantages for neural processing, this view persisted well into the middle of the twentieth century. It was impossible to directly refute because, in the absence of a fossil record for soft tissue such as brain, the identification of homologies between avian and mammalian brain structures was not feasible. To provide an overview of some of these issues we have reprinted excerpts from a chapter by Reiner and his colleagues which originally provided an introduction and rationale for the recent revision of songbird brain nomenclature. The goal of this revision was to develop a nomenclature both more neutral in its attribution of homologies and more systematic in its application of the latest available anatomical, neurochemical and developmental information.

The advent of techniques for experimental anatomy and genomic analysis has made it possible to carry out detailed comparative studies of the organization of the embryonic and adult avian and mammalian forebrains. These studies have demonstrated many similarities between the connection patterns of structures in the avian and mammalian forebrain and have clarified the relation between basal ganglia structures in the two groups. Farries and Perkel provide an overview of the implications of these findings for both avian and mammalian brain research.

## 2 • Birdsong and human speech: common themes and mechanisms

Allison J. Doupe and Patricia K. Kuhl

### INTRODUCTION

Experts in the fields of human speech and birdsong have often commented on the parallels between the two in terms of communication and its development (Marler, 1970a; Kuhl, 1989). Does the acquisition of song in birds provide insights regarding learning of speech in humans? This review provides a critical assessment of the hypothesis, examining whether the similarities between the two fields go beyond superficial analogy. The often cited commonalities provide the topics of comparison that structure this review.

First, learning is critical to both birdsong and speech. Birds do not learn to sing normally, nor infants to speak, if they are not exposed to the communicative signals of adults of the species. This is an exception among species: most animals do not have to be exposed to the communicative signals of their species to be able to reproduce them. The fact that babies and songbirds share this requirement has intrigued scientists.

Second, vocal learning requires both perception of sound and the capacity to produce sound. At birth, both human infants and songbirds have been hypothesized to have innate perceptual predispositions for the vocal behavior of their own species. We review the nature of the predispositions in the two cases and the issue of whether they are similar. Given that innate predispositions exist, another important question is how subsequent experience alters perception and production in each case. Moreover, vocal perception and production are tightly interwoven in the vocal learning process. We examine what is known about the

relationship between perception and production and whether in these different vocal learners it is similar.

In addition, neural substrates of vocal communication in humans and birds have often been compared. Human brains are asymmetric and language tends to be organized in the left hemisphere as opposed to the right. Birds are also often assumed to show similar hemispheric specialization for song. What are the real parallels between the neural substrates in the two cases?

Finally, critical (sensitive) periods are evidenced in both species. Neither birds nor babies appear to learn their communicative signals equally well at all phases of the life cycle. This raises the questions of what causes the change in the ability to learn over time and with experience, and whether the causes are the same in human infants and songbirds. And if the plasticity of the brain is altered over the life cycle, what neural mechanisms control this changing ability to learn?

The research reviewed here relates to ongoing work in developmental biology, ethology, linguistics, cognitive psychology, and computer science, as well as in neuroscience, and should be of interest to individuals in many of these fields. What our review reveals is that although the comparisons between birdsong and speech are not simple, there is a surprisingly large number of areas where it is fruitful to compare the two. Going beyond the superficial analogy, however, requires some caveats about what may be comparable and what clearly is not. In the end, understanding both the similarities and differences will provide a broader spectrum in which to view the acquisition of communication in animals and humans.

---

*Editors note:* Adapted and reprinted, with permission from the authors and from the *Annual Review of Neuroscience*, Volume 22, ©1999 by Annual Reviews [www.annualreviews.org](http://www.annualreviews.org). The sections on “Specialized neural substrates for song and speech learning (pp. 596–604)” on “lateralization” (pp. 606–609) and on “possible

neural mechanisms underlying the sensitive period and its closure” (pp. 618–619) have been omitted. The reader is referred to relevant chapters in this volume for reviews of more recent work on neural mechanisms.

*Neuroscience of Birdsong*, ed. H. Philip Zeigler and Peter Marler. Published by Cambridge University Press. © Cambridge University Press 2008.

SPEECH AND BIRDSONG :  
DEFINITIONS

Speech and song production

Both birdsong and human speech are complex acoustic signals. Figure 2.1 shows a spectrographic (frequency versus time) display of a spoken human phrase (“Did you hit it to Tom?”) and Figure 2.2 a similar display of

songs of two different songbird species. In both songbirds and humans, these sounds are produced by the flow of air during expiration through a vocal system. In humans, the process is relatively well understood: air from expiration generates a complex waveform at the vocal folds, and the components of this waveform are subsequently modified by the rest of the vocal tract (including the mouth, tongue, teeth, and lips) (Stevens, 1994).

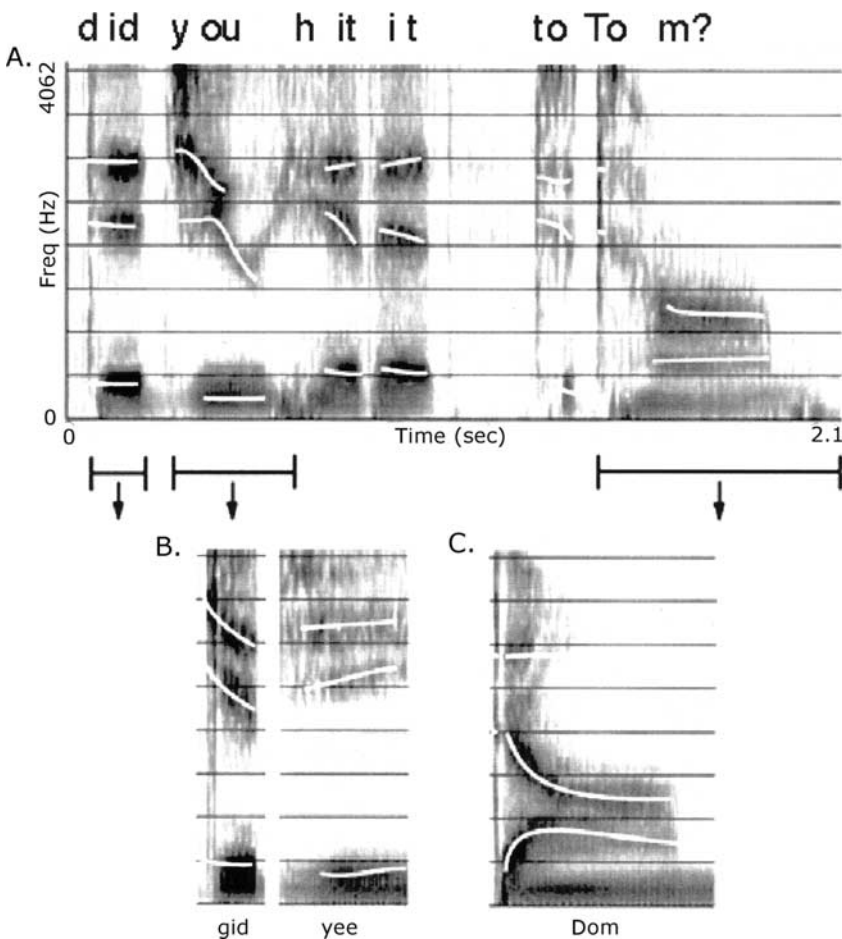


Figure 2.1 Human speech. Three dimensions of speech are shown in a spectrogram: time or duration along the horizontal axis; frequency along the vertical axis; and intensity, which is correlated with loudness, by the relative darkness of each frequency. This spectrogram shows the phrase “Did you hit it to Tom?” spoken by a female (A). White lines are the formants that characterize each individual phoneme. (B–C) Variations on words from the full sentence. (B) A place of articulation contrast using a spectrogram of the nonsense word “gid,” which differs from its rhyme “did” (in A) in that it has a decreasing frequency sweep in the second and third formants (between 2000 and 3000 Hz). This decreasing formant pattern defines the sound “g” and a pattern of flat formants defines the sound “d.” (C) The words “Tom” and “Dom” contrast in voice onset time (VOT). Notice the long, noisy gap in “Tom” (A), which has a long VOT, compared with the short gap in “Dom.”

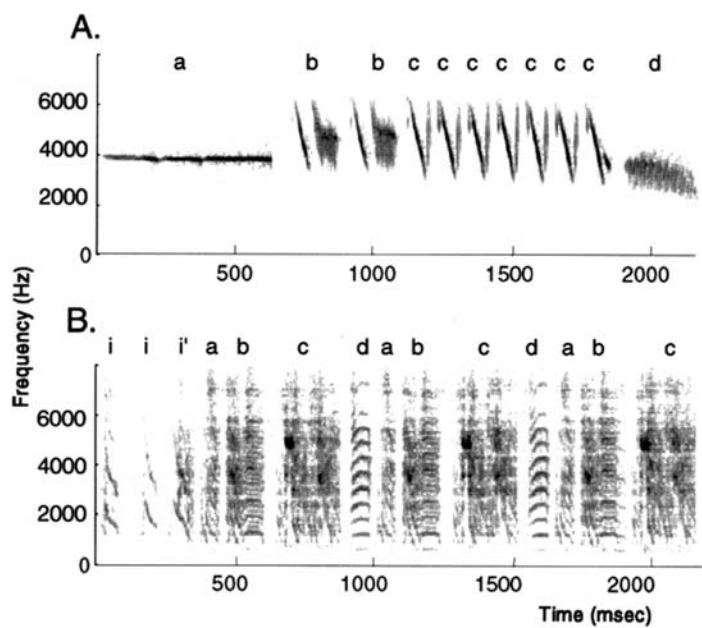


Figure 2.2 Examples of birdsongs from two species. (A) A typical song of a white-crowned sparrow. The smallest elements, the notes, are combined to form syllables (lower case letters), and these are repeated to form phrases. White-crowned sparrow songs typically begin with (a) a long whistle followed by (b, c) trills and (d) buzzes. (B) A typical song of a zebra finch. Note the noisy spectral quality (more like humans) that distinguishes it from more tonal species like the sparrows. Zebra finch songs start with a number of introductory syllables (marked with i), followed by a sequence of syllables (lower case letters), that can be either simple or more complex, with multiple notes (e.g. b, c). Particular sequences of syllables are organized into phrases called motifs, which are repeated.

The vocal tract acts as a filter, creating concentrations of energy at particular frequencies, called formant frequencies (Figure 2.1). Vowels are characterized by relatively constant formant frequencies over time (Figure 2.1A, C), whereas during consonant production the formant frequencies change rapidly (20–100 ms), resulting in formant transitions (Figure 2.1A, B, D).

In songbirds, sounds are produced by the flow of air during expiration through an organ called the syrinx, a bilateral structure surrounded by specialized muscles, which sits at the junction of the bronchi with the trachea. A number of aspects of syringeal function are understood, although the exact mechanism of sound generation is controversial and is under active investigation (Gaunt, 1987; Goller and Larsen, 1997a; Suthers, 1997; Fee *et al.*, 1998). Also, there are indications that the upper vocal tract in bird structures sound in a manner like the upper vocal tract in humans. Recent research suggests that the width of beak opening (known as beak gape) affects sound frequency

(Westneat *et al.*, 1993; Suthers, 1997), and there may be some degree of coupling between the syrinx and the vocal tract (Nowicki, 1987). Regardless of differences in component structures, for both birdsong and speech the production of highly structured and rapidly changing vocalizations requires elaborate neural control and coordination of respiration with a variety of vocal motor structures.

The structure of speech and song

It is useful to define the basic terms used in each field, and the various ways in which vocal behavior is described, in order to assess what aspects of each of the signals are comparable. Human speech can be described at many different levels. It can be written, spoken, or signed (using a manual language such as American Sign Language). In all these forms, language consists of a string of words ordered by the rules of grammar to convey meaning. Structurally,



language can be analyzed from the standpoint of semantics (conceptual representation), syntax (word order), prosody (the pitch, rhythm, and tempo of an utterance), the lexicon (words), or phonology (the elementary building blocks, phonemes, that are combined to make up words).

Speech, and especially its development, has been intensively studied at the phonological level. Phonetic units are the smallest elements that can alter the meaning of a word in any language, for example the difference between /r/ and /l/ in the words “rid” and “lid” in American English. Phonemes refer to the phonetic units critical for meaning in a particular language. The phonetic difference between /r/ and /l/ is phonemic in English, for example, but not in Japanese. Each phonetic unit can be described as a bundle of phonetic features that indicate the manner in which the sound was produced and the place in the mouth where the articulators (tongue, lips, teeth) were placed to create the sound (Jakobson *et al.*, 1969). The acoustic cues that signal phonetic units have been well documented and include both spectral and temporal features of sound (Figure 2.1) (Stevens, 1994). For instance, the distinction between /d/ and /g/ depends primarily on the frequency content of the initial burst in energy at the beginning of the sound and the direction of formant transition change (Figure 2.1A, B). An example of a temporal acoustic dimension of speech is voice-onset time (VOT), which refers to the timing of periodic laryngeal vibration (voicing) in relation to the beginning of the syllable (Figure 2.1A, D). This timing difference provides the critical cue used to identify whether a speech sound is voiced or voiceless (e.g. /b/ versus /p/, /do/ versus /to/) and is a classic distinction used in many speech studies.

Which aspects of birdsong can be usefully compared with speech? Birdsongs are distinct from bird calls (which are brief and generally not learned), last from a few seconds to many tens of seconds, and, like speech, consist of ordered strings of sounds separated by brief silent intervals (Figure 2.2). The smallest level of song usually identified is the note or element, defined as a continuous marking on a sound spectrogram; these may be analogous to the smallest units of speech, or phonetic units. Notes can be grouped together to form syllables, which are units of sound separated by silent intervals. When singing birds are interrupted by an abrupt light flash or sound, they

complete the syllable before stopping (Cynx, 1990); thus, syllables may represent a basic processing unit in birdsong, as posited for speech.

Another feature that birdsong and language share is the conspicuous timing and ordering of components on a timescale longer than that of the syllable. Song syllables are usually grouped together to form phrases or motifs (Figure 2.2), which can be a series of identical or different syllables. Many songbirds sing several phrases in a fixed order as a unit, which constitutes the song, whereas other species such as mockingbirds and warblers produce groups of syllables in fixed or variable sequences. The timing and sequencing of syllables and phrases are rarely random but instead follow a set of rules particular to a species. In the songbird literature, the ordering of syllables and phrases in song is often called song syntax. The same word applied to human speech, however, implies grammar, i.e. rules of ordering words from various grammatical classes to convey meaning. Therefore, in this review, we avoid using the word syntax for song and simply use “order.” Thus, language and song share a dependence on timing on several timescales: a shorter timescale (on the order of tens of milliseconds), as in phonemes and syllables, and a longer one, up to many hundreds of milliseconds (as in syllable, phrase, and word ordering).

Language is also characterized by a boundless and flexible capacity to convey meaning, but this property is not shared with birdsong. The whole set of different songs of a bird is known as its song repertoire and can vary from one (in species such as the zebra finch or white-crowned sparrow) to several hundreds (for review see Konishi, 1985). Numerous behavioral studies, usually using the receiver’s response, suggest that songs communicate species and individual identity (including “neighbor” and “stranger”), an advertisement for mating, ownership of territory, and fitness. Some birds with multiple song types use different songs for territorial advertisement and for mate attraction (Catchpole, 1983; Searcy and Nowicki, 1998). Nonetheless, large song repertoires do not seem to convey many different meanings, nor does song have the complex semantics of human speech. The definitions above suggest that the phonology (sound structure), the rules for ordering sounds, and perhaps the prosody (in the sense that it involves control of frequency, timing, and amplitude) are the levels at which birdsong can be most usefully compared with

language, and more specifically with spoken speech, and are thus the focus of this review.

## VOCAL LEARNING IN HUMANS AND SONGBIRDS

### Which animals are vocal learners?

Many animals produce complex communication sounds but few of them can and must learn these vocal signals. Humans are consummate vocal learners. Although there is emerging evidence that social factors can influence acoustic variability among nonhuman primates (Sugiura, 1998), no other primates have yet been shown to learn their vocalizations. Among the mammals, cetaceans are well known to acquire their vocal repertoire and to show vocal mimicry (McCowan and Reiss, 1997); there are also some bats whose vocalizations may be learned (Boughman, 1998). Among avian species, songbirds, the parrot family, and some hummingbirds meet the criteria for vocal learning, but the term birdsong is usually reserved for the vocalizations of passerine (perching) songbirds and that is the focus of this review. The many thousands of songbird species, as well as the parrots and hummingbirds, stand in striking contrast to the paucity of mammalian vocal learners.

Nonhuman primates can, however, make meaningful use of vocalizations: for instance, vervets use different calls to indicate different categories of predators. Production of these calls is relatively normal even in young vervets and does not appear to go through a period of gradual vocal development, but these animals must develop the correct associations of calls to predators during early ontogeny (Seyfarth and Cheney, 1997). What songbirds and humans share is not this development of associations of vocalizations with objects or actions, but the basic experience-dependent memorization of sensory inputs and the shaping of vocal outputs.

### Evidence for vocal learning

The basic phenomenology of learning of song or speech is strikingly similar in songbirds and humans. Initial vocalizations are immature and unlike those of adults: babies babble, producing consonant–vowel syllables that are strung together (e.g. bababa or mamama), and

young songbirds produce subsong, soft and rambling strings of sound. Early sounds are then gradually molded to resemble adult vocalizations. The result of this vocal development is that adults produce a stereotyped repertoire of acoustic elements: these are relatively fixed for a given individual, but they vary between individuals and groups (as in languages and dialects, and the individually distinct songs and dialects of songbirds within a particular species). This variability is a reflection of the fact that vocal production by individuals is limited to a subset of all sounds that can be produced by that species. Layered on top of the developing capacity to produce particular acoustic elements is the development of sequencing of these elements: for humans this means ordering sounds to create words and, at a higher level, sentences and grammar; in birds this means sequencing of elements and phrases of song in the appropriate order. An important difference to remember when making comparisons is that the numerous languages of humans are not equivalent to the songs of different species, but rather to the individual and geographical variations of songs within a species.

## LEARNED DIFFERENCES IN VOCAL BEHAVIOR

That the development of a mature vocal repertoire reflects learning rather than simply the expression of innate programs is apparent from a number of observations. Most important, for both birds and humans, there exist group differences in vocal production that clearly depend on experience. Obviously, people learn the language to which they are exposed. Moreover, even within a specific language, dialects can identify the specific region of the country in which a person was raised. Likewise, songbirds learn the songs sung by adults to which they are exposed during development: this can be clearly demonstrated by showing that birds taken from the wild as eggs or nestlings and exposed to unrelated conspecific adults, or even simply to tape recordings of the song of these adults, ultimately produce normal songs that match those that were heard (Marler, 1970b; Thorpe, 1958, 1961). Even more compelling are cross-fostering experiments, in which birds of one species being raised by another will learn the song, or aspects thereof, of the fostering species (Immelmann, 1969). In addition, many songbirds have song “dialects,”



particular constellations of acoustic features that are well defined and restricted to local geographic areas. Just as with human dialects, these song dialects are culturally transmitted (Marler and Tamura, 1962).

#### VOCALIZATIONS IN THE ABSENCE OF EXPOSURE TO OTHERS

Another line of evidence supporting vocal learning is the development of abnormal vocalizations when humans or birds with normal hearing are socially isolated and therefore not exposed to the vocalizations of others. The need for auditory experience of others in humans is evident in the (fortunately rare) studies of children raised either in abnormal social settings, as in the case of the California girl, Genie, who was raised with almost no social contact (Fromkin *et al.*, 1974), or in cases in which abandoned children were raised quite literally in the wild (Lane, 1976). These and other documented instances in which infants with normal hearing were not exposed to human speech provide dramatic evidence that in the absence of hearing speech from others, speech does not develop normally. Similarly, songbirds collected as nestlings and raised in isolation from adult song produce very abnormal songs (called “isolate” songs) (Marler, 1970b; Thorpe, 1958). This need for early auditory tutoring has been demonstrated in a wide variety of songbirds (for reviews see Catchpole and Slater, 1995; Kroodsmma and Miller, 1996). Strikingly, although isolate songs are simplified compared with normal, learned song, they still show some features of species-specific song (Marler and Sherman, 1985).

One caveat about studies of isolated songbirds or humans is that many aspects of development are altered or delayed in such abnormal rearing conditions. Nonetheless, the results of isolation in humans and songbirds are in striking contrast to those seen with members of closely related species, such as nonhuman primates and nonsongbirds such as chickens, in whom vocalizations develop relatively normally even when animals are raised in complete acoustic isolation (Konishi, 1963; Kroodsmma, 1985; Seyfarth and Cheney, 1997). In combination with the potent effects of particular acoustic inputs on the type of vocal output produced, these results demonstrate how critically both birdsong and speech learning depend on the auditory experience provided by hearing others vocalize.

#### THE IMPORTANCE OF AUDITION IN SPEECH AND SONG

##### The importance of hearing one's own vocalizations

Vocal learning, shared with few other animals, is also evident in the fact that both humans and songbirds are acutely dependent on the ability to hear themselves in order to develop normal vocalizations. Human infants born congenitally deaf do not acquire spoken language, although they will, of course, learn a natural sign language if exposed to it (Petitto, 1993). Deaf infants show abnormalities very early in babbling, which is an important milestone of early language acquisition. At about 7 months of age, typically developing infants across all cultures will produce this form of speech. The babbling of deaf infants, however, is maturationally delayed and lacks the temporal structure and the full range of consonant sounds of normal-hearing infants (Oller and Eilers, 1988; Stoel-Gammon and Otomo, 1986). The strong dependence of speech on hearing early in life contrasts with that of humans who become deaf as adults: their speech shows gradual deterioration but is well preserved relative to that of deaf children (Cowie and Douglas-Cowie, 1992; Waldstein, 1990).

Songbirds are also critically dependent on hearing early in life for successful vocal learning. Although birds other than songbirds, e.g. chickens, produce normal vocalizations even when deafened as juveniles, songbirds must be able to hear themselves in order to develop normal song (Konishi, 1963, 1965b; Nottebohm, 1968). Songbirds still sing when deafened young, but they produce a very abnormal, indistinct series of sounds that are much less songlike than are isolate songs; although it varies from species to species, often only a few features of normal songs are maintained, primarily their approximate duration (Marler and Sherman, 1983). As with humans, once adult vocalizations have stabilized, most songbird species show decreased dependence on hearing (Konishi, 1965b; but see below).

The effects of deafness in early life do not differentiate between the need for hearing others and a requirement for hearing oneself while learning to vocalize. In birds, however, there is often a separation between the period of hearing adult song and the onset of vocalizations, and this provided the opportunity to demonstrate that song is abnormal in birds even

when they have had adequate tutor experience prior to being deafened (Konishi, 1965b). This revealed that during song learning hearing functions in two ways, in two largely nonoverlapping phases (Figure 2.3B). During an initial sensory phase, the bird listens to and learns the tutor song. After this sensory learning, however, the memorized song, called the template, cannot be simply internally translated into the correct vocal motor pattern. Instead, a second, sensorimotor learning or vocal practice phase is necessary. The bird must actively compare and gradually match its own vocalizations to the memorized template, using auditory feedback. The need for the bird to hear itself is also evident in birds first raised in isolation and then deafened prior to sensorimotor learning. These birds sing abnormal songs indistinguishable from those of deafened tutored birds, demonstrating that the innate information about song that exists in isolate birds also requires auditory feedback from the birds' own vocalizations and motor learning in order to be turned into motor output (Konishi, 1965b). Thus, learning to produce song is crucially dependent on auditory experience of self as well as of others.

Humans likely also have to hear themselves in order to develop normal speech. This issue is more difficult to study in human infants than in songbirds, however, because the need for auditory input from others overlaps substantially in time with when children are learning to speak (Figure 2.3A). Studies of children becoming deaf later in childhood, however, indicate that speech still deteriorates markedly if deafness occurs prior to puberty (Plant and Hammarberg, 1983). Thus, even though language production is well developed by late preadolescence, it cannot be well maintained without the ability to hear, which suggests that feedback from the sound of the speaker's own voice is also crucial to the development and stabilization of speech production. In addition, special cases in which infants hear normally but cannot vocalize provide relevant data. Studies of speech development in children who prior to language development had tracheostomies for periods lasting from 6 months to several years indicate severe speech and language delays as a result (Locke and Pearson, 1990; Kamen and Watson, 1991). Although these studies cannot rule out motor deficits due to lack of practice or motor damage, the speech of these children, who have normal hearing, is similar in its structure to that produced by deaf children. These

studies, and the effects of deafness on older children, provide evidence that, just as in songbirds, both the sounds produced by the individuals themselves and those produced by others are essential for normal speech development.

### THE FUNCTION OF AUDITORY FEEDBACK IN ADULTHOOD

In both humans and songbirds, the strong dependence of vocal behavior on hearing early in life lessens in adulthood. Postlingually deaf adults do show speech deterioration (Cowie and Douglas-Cowie, 1992; Waldstein, 1990), but it is less than that of deaf children, and it can be rapidly ameliorated even by the limited hearing provided by cochlear implants (Tyler, 1993). In some songbird species, song deteriorates very little in deafened adults, which suggests song is maintained by nonauditory feedback and/or by a central pattern generator that emerged during learning. In other species, song deteriorates more markedly after deafness in adulthood, both in phonology and in syllable ordering (Nordeen and Nordeen, 1993; Woolley and Rubel, 1997; Okanoya and Yamaguchi, 1997). Even in these cases, in many species song deterioration is often slower in adults than in birds actively learning song and may depend on how long the bird has been singing mature, adult ("crystallized") song. Some birds are "open" learners: that is, their capacity to learn to produce new song remains open in adulthood (e.g. canaries) (Nottebohm *et al.*, 1986). Consistent with how critical hearing is to the learning of song, these species remain acutely dependent on auditory feedback for normal song production as adults.

Moreover, for both human speech and birdsong, incorrect or delayed auditory feedback in adults is more disruptive than the complete absence of auditory feedback. For instance, delayed auditory playback of a person's voice causes slowing, pauses, and syllable repetitions in that subject (Howell and Archer, 1984; Lee, 1950). In addition, presentation to adult humans of altered versions of the vowels in their own speech, at a very short time delay, causes the subjects unconsciously to produce appropriately altered speech (Houde and Jordan, 1998). In songbirds as well, recent results suggest that delayed or altered auditory feedback can cause syllable repetitions or song deterioration (Leonardo and Konishi, 1999; J. Cynx, personal