

# Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds

Mimi L. Phan\*, Carolyn L. Pytte<sup>†</sup>, and David S. Vicario\*\*

\*Department of Psychology, Rutgers University, Piscataway, NJ 08854; and <sup>†</sup>Department of Biology, Wesleyan University, Middletown, CT 06459

Communicated by Charles R. Gallistel, Rutgers University, Piscataway, NJ, December 2, 2005 (received for review June 3, 2004)

In both humans and songbirds, infants learn vocalizations by imitating the sounds of adult tutors with whom they interact during an early sensitive period. Vocal learning occurs in few animal taxa; similarities in the imitation process between humans and songbirds make the songbird a unique system in which vocal learning mechanisms can be studied at the neurobiological level. One theory of vocal learning proposes that early auditory experience generates auditory memories that subsequently guide vocal imitation. We now present a combination of behavioral and neurophysiological results, obtained in a songbird, that support this theory. We show that neurons in a forebrain auditory area of adult male zebra finches are selectively tuned to the song of a tutor heard early in development. Furthermore, the strength of this selectivity shows a striking correlation with the fidelity of vocal imitation, suggesting that this auditory memory may have served as the model for song learning.

development | electrophysiology | memory | caudal medial nidopallium | zebra finch

Human infants exposed to adult speech in an appropriate social context (1) selectively attend to common vocal sounds in the speech stream and remember those sounds (2). This memory results in perceptual tuning to major phonological categories, typical of the language that the infant hears (3). Infants subsequently begin to produce differentiated sounds, shaping their vocal output toward sound targets, perhaps based on a comparison between auditory feedback from self-vocalizations and the acquired perceptual tuning (4). However, the neuronal mechanisms of this form of perceptual tuning and its contribution to speech acquisition are not well understood (5). Numerous similarities in vocal learning between songbirds and humans (5), together with the fact that vocal imitation occurs in very few animal taxa (6, 7), make the songbird a unique model system in which developmental theories of vocal learning can be subjected to neurophysiological investigation.

In songbirds, song learning and production are mediated by a set of interconnected brain regions termed the “song system” (8). Song learning is thought to involve a sensory phase during which a juvenile memorizes a model song (produced by an adult tutor), and a sensorimotor phase when that auditory memory is used as a “template” to compare auditory feedback from his own juvenile song with the stored memory of the tutor’s song (9, 10). The output of this comparison is thought to progressively modify subsequent motor commands, resulting in imitation of the tutor’s song (9). Imitation is not perfect, however, and each bird acquires a song similar to that of the tutor, but unique to the individual, called the bird’s own song (BOS).

Neurophysiological studies have looked for evidence of this template memory in the auditory responses of song system neurons in both the direct song motor pathway (11) and the anterior forebrain pathway (12–15), which is implicated in song learning (16, 17). In those reports, BOS is typically the preferred stimulus, although, during song development, responses selective for the tutor have also been observed (11–15). In the adult zebra

finch, responses to the tutor in the anterior forebrain pathway persist (14, 15), but are either very rare (14) or may be attributable to the similarity between the tutor and BOS (15). Overall, these findings suggest that a neural representation of the tutor’s song is apparently lost or largely overwritten during the juvenile’s acquisition of his own song (11–15).

Here we provide clear neurophysiological evidence for a long-lasting memory of the tutor’s song, located not in the traditional song system, but in an area of the songbird auditory forebrain, the caudal medial nidopallium (NCM). The NCM participates in discrimination and memory of the complex vocalizations of individual conspecifics, based on both physiological and immediate early gene studies (18–23). Furthermore, gene expression studies have implicated the NCM in tutor song memory (19, 20). For these reasons, we hypothesized that (i) the NCM contains an auditory memory of the tutor’s song (19–21, 24) that can guide vocal development and (ii) deficits in this memory should result in deficits in imitation. We tested our hypotheses by raising juvenile male zebra finches in a controlled song tutoring situation (25), assessing the fidelity of their imitation of the tutor’s song in adulthood, and measuring the strength of their memory for the tutor song by recording neuronal responses in the NCM to a variety of familiar and novel auditory stimuli in adulthood. Our study demonstrates an auditory memory that could be used in song learning that is located in an auditory area outside of the song system. Furthermore, the strength of this auditory memory is correlated with the degree to which the bird has imitated the tutor’s song, suggesting that this memory either is the memory used for imitation or is tightly coupled to it.

## Results

**Song Imitation Measured in the Songs of Adults.** In adulthood, all of the experimental birds underwent a combination of behavioral and neurophysiological assessments. To determine the degree of song imitation, the song similarity index (SI) between the tutor song and each adult BOS was measured (25) for each of the intact birds. The SI of experimentally tutored birds ( $n = 9$ ; mean SI = 61.4, SEM 6.2; Fig. 1A) was not significantly different from ( $t = 1.29$ ;  $P = 0.217$ ) the SI of birds raised by their parents in a family cage ( $n = 11$ ; mean SI = 71, SEM 4.0). Thus, this tutoring paradigm results in comparable song imitation to that which occurs in the normal social learning context. Furthermore the SI of the tutored birds was significantly different ( $t = 1.80$ ,  $P = 0.003$ ) from that of the untutored birds ( $n = 5$ ; mean SI = 30, SEM 5.5). Even untutored birds can show some degree of similarity to the tutor song used in this study (Fig. 1B), which

Conflict of interest statement: No conflicts declared.

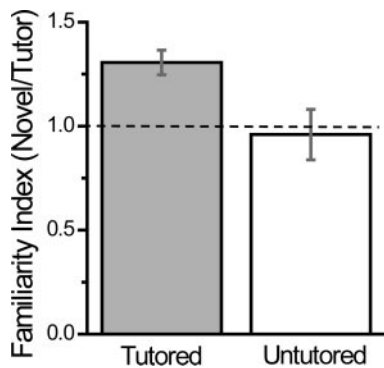
Freely available online through the PNAS open access option.

Abbreviations: BOS, bird’s own song; BOS<sub>juv</sub>, juvenile BOS; NCM, caudal medial nidopallium; SI, similarity index; FI, familiarity index; phd, posthatch day.

<sup>†</sup>To whom correspondence should be addressed. E-mail: vicario@rci.rutgers.edu.

© 2006 by The National Academy of Sciences of the USA





**Fig. 3.** Auditory selectivity for the tutor song is stronger in tutored birds. The FI (see *Methods*) for the tutor song, a measurement of song memory derived from neuronal responses in the NCM, differs significantly ( $t = 2.627$ ,  $P = 0.0392$ ) between birds that were tutored ( $n = 9$ , mean FI = 1.315, SEM = 0.0592; filled bar) and those that were untutored ( $n = 5$ , mean FI = 0.959; SEM = 0.122; open bar). An FI ratio near 1 (dashed line) indicates that the tutor song is not distinguished from novel songs; an FI ratio  $> 1$  indicates that the tutor song is more familiar than the novel songs.

makes it unlikely that song delivery conditions played the critical role in the very long-term auditory memory observed in the birds tutored as juveniles. Instead, it is likely that a different memory-inducing mechanism is engaged during the sensitive period, resulting in a much more durable song memory.

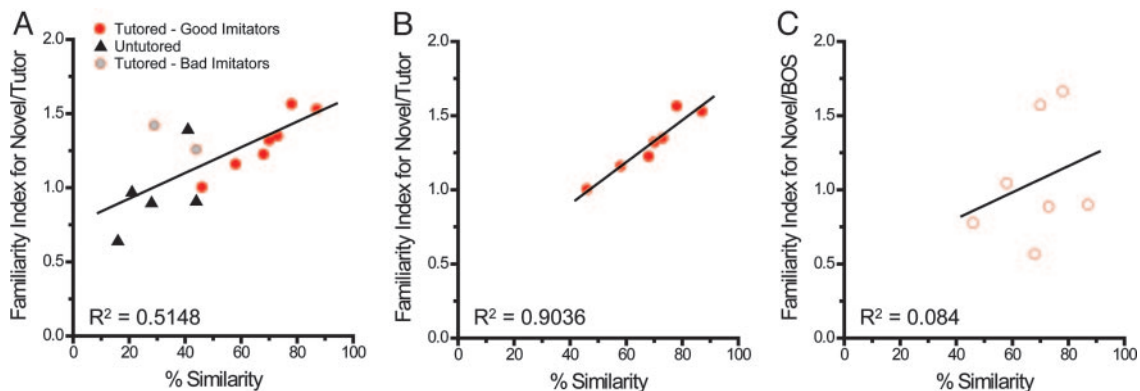
In our experimental paradigm, tutored birds were in isolation; they only heard the tutor song and themselves. As a result, we do not know whether this long-term memory mechanism is exclusive to the tutor song or might make a lasting record of other songs heard in the juvenile period, but not adopted by the bird as song models. A partial test of this idea comes from the results obtained when responses to the BOS before crystallization, juvenile BOS (BOS<sub>juv</sub>), were measured. Despite considerable exposure, the mean FI for BOS<sub>juv</sub> was not significantly different from 1 (based on confidence interval tests) in either intact tutored males (FI<sub>BOS<sub>juv</sub></sub> = 1.04; SEM = 0.081,) or in isolate males (FI<sub>BOS<sub>juv</sub></sub> = 0.89; SEM = 0.153), whose only song exposure was to their own vocalizations.

Results from another experimental group, consisting of birds that were devocalized before tutoring, suggest the need for an additional factor beyond exposure, e.g., some aspect of senso-

rimotor integration, in establishing or maintaining the memory of the tutor's song. These birds heard the tutor song in the same behavioral context as the intact tutored birds, but were unable to produce any audible vocalizations and thus could not imitate the tutor's song. The FI<sub>tutor</sub> in these birds ( $n = 6$ , mean FI<sub>tutor</sub> = 1.12, SEM 0.14) was intermediate between FI<sub>tutor</sub> for tutored intact birds and FI<sub>tutor</sub> for untutored intact birds and not significantly different from either group. Although we cannot draw strong conclusions from these data, they suggest that the memory for the tutor song is weaker when the auditory feedback from self-vocalizations is eliminated, and thus that sensory exposure alone is not sufficient to establish a strong auditory memory of the tutor song.

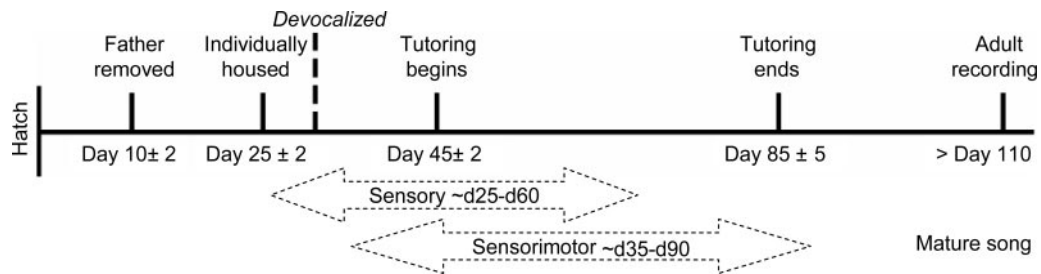
**Strength of Auditory Memory Correlates with Learning.** One important prediction for a song memory that is used as a template for vocal learning is that the auditory representation of the tutor song might determine or constrain imitation success. Across all tutored and untutored intact birds, there was a significant correlation between FI<sub>tutor</sub> in the NCM and the magnitude of SI (Fig. 4A,  $n = 14$ ,  $R^2 = 0.515$ ,  $F_{(1,13)} = 12.733$ ,  $P = 0.0039$ ). However, among the tutored birds, two showed no imitation: they had very low SIs (Fig. 4A, gray circles) that did not differ significantly from those of the untutored birds ( $n = 2$ ,  $P > 0.05$ ). We limited further analysis to the seven of nine birds that successfully imitated the tutor, as determined by SIs that were significantly greater than those of the untutored birds ( $n = 7$ , two-tailed  $t$  test, range of  $t$  values 2.306–2.776, all  $P$  values  $< 0.005$ ; mean SI = 69; SEM 5.1). The data from this set of tutored birds showed an even stronger correlation between FI<sub>tutor</sub> and SI (Fig. 4B,  $R^2 = 0.904$ ,  $F_{(1,6)} = 46.88$ ,  $P = 0.001$ ). This relationship suggests that, for birds that imitate, the quality of the tutor song memory in the NCM is highly correlated with the fidelity of song imitation.

We also explored the possibility that, when a bird makes a good copy of the tutor song, BOS is very similar to the tutor song, and thus FI<sub>tutor</sub> in the NCM indirectly measures the familiarity of BOS. When we directly computed the FI for BOS, the correlation between FI<sub>BOS</sub> and the SI was not significant (Fig. 4C;  $R^2 = 0.084$ ,  $F_{(1,6)} = 0.458$ ,  $P = 0.528$ ). Furthermore, a confidence interval test indicated no significant difference between FI<sub>BOS</sub> and 1, which indicates that the significant correlation between FI<sub>tutor</sub> in the NCM and the SI is not simply mediated by the relationship between tutor song and BOS. This



**Fig. 4.** The FI for tutor song observed in the NCM is a strong predictor of imitative success. (A) The habituation rate in the NCM is correlated with the degree of similarity between BOS and the tutor song ( $R^2 = 0.515$ ,  $F_{(1,13)} = 12.733$ ,  $P = 0.0039$ ) for all of the tutored intact and untutored intact male birds ( $n = 14$ ). The correlation for the group of untutored intact birds was not significant, as expected ( $n = 5$ ,  $R^2 = 0.407$ ,  $F_{(1,4)} = 2.062$ ,  $P = 0.247$ ). Further analysis showed that the measurements of the percent similarity for two of the nine tutored intact birds did not differ significantly from the untutored intact birds, indicating that they did not learn the tutor song. (B) If the regression is limited to the birds that showed significant learning, the correlation is significant ( $n = 7$ ,  $R^2 = 0.904$ ,  $F_{(1,6)} = 46.88$ ,  $P = 0.001$ ). (C) In these same birds, the FI for BOS was not significantly related to the similarity measurements of the percent ( $n = 7$ ,  $R^2 = 0.084$ ,  $F_{(1,6)} = 0.458$ ,  $P = 0.528$ ), indicating that the strong correlation seen for the FI of tutor song is not simply mediated by the similarity between tutor song and BOS.





**Fig. 5.** Timeline of key points in the song tutoring paradigm. Birds were isolated from their parents, tutored during both the sensory and sensorimotor phases of song learning (see text), and then kept in isolation for >30 days before neurophysiological recordings were made in adulthood.

finding is consistent with earlier results showing that BOS, despite being the most frequently heard song for an individual male, is not responded to as familiar in the NCM (22).

## Discussion

Our data are relevant to the template theory of vocal learning that proposes that an auditory memory of the tutor song is acquired during the sensory phase of song learning, which then guides vocal imitation in the sensorimotor stage of song learning. The results show that birds with a tutor song memory that is more precisely differentiated from the set of novel songs produce a proportionally better copy of the tutor song. This is what would be expected of an auditory memory that serves as an auditory model for song matching, although, of course, this correlation does not show causation. In addition, we cannot definitively exclude alternative interpretations that suggest that preexisting individual variability in some aspect of song processing might contribute to the observed correlation.

Furthermore, we cannot exclude that the tutor song memory in the NCM could be a readout that merely reflects the true template occurring elsewhere in the brain. The tutor song memory seen in the NCM may be one sensory component of a matching mechanism, or template system (10), with other components distributed in the song motor system (11) and/or the anterior forebrain pathway (12–15, 26). Because the song motor pathway and the anterior forebrain pathway are involved in the sensorimotor process of song learning, recent changes in what the bird produces during learning may dynamically affect the memory of the tutor song in those areas (11, 13, 15). However, studies that describe adult song changes in response to reversible deafening, auditory feedback distortions, or reversible manipulations of the vocal periphery, show recovery of the original song after removal of these perturbations (27–30). These findings suggest that the bird could retain a copy of the original tutor song during the period of aberrant feedback (cf. ref. 30), which then assists in BOS recovery. The NCM is primarily a sensory area, and it could thus be an effective storage site for an auditory memory of the tutor song that is archival.

The multiple song memories observed in different brain regions [i.e., in the NCM, high vocal center (11), lateral magnocellular nucleus of the anterior nidopallium (12–15), and area X (12–14)] may not exist in isolation from each other, but could interact. Results from our third experimental group, birds that were devocalized before tutoring, suggest that this kind of interaction might occur. These birds had a  $FI_{\text{tutor}}$  that was intermediate between the intact tutored birds and the untutored birds. This finding suggests that, in intact birds, the feedback from self-vocalizations, apparently involved in creating a sensorimotor memory for BOS in the high vocal center and the anterior forebrain pathway, may play some role in the establishment of a strong auditory memory of the tutor song in the NCM. The nature of this influence could be to strengthen, but not alter, the auditory memory, perhaps through use. Thus the memory

would remain true to the tutor song while its familiarity would be enhanced by the degree to which the bird's own vocalizations matched the tutor.

Although the FI was determined for adult birds, we infer that the auditory memory was originally laid down through exposure to the tutor song during the sensory phase of song learning and was used during the sensorimotor phase (although the present data do not reveal the precise time course). This possibility is consistent with results that show immediate early gene expression in the NCM can be induced by song playback at 30 phd, suggesting that auditory memories for songs may be formed as early as 30 days of age (18, 31).

In summary, early exposure to a tutor produces a song memory in the auditory forebrain that lasts into adulthood and is specific to the tutor song, not BOS. This is the clearest physiological evidence to date for a tutor song memory that could function as a template for song imitation and song maintenance. This memory is located in an auditory area, not in the traditional song motor nuclei or the anterior forebrain pathway. Further, the strength of this memory shows a striking correlation with the fidelity of the bird's vocal imitation. Our data are consistent with findings that have shown a relationship between immediate early gene expression in the NCM and the strength of tutor song learning (19, 20).

Together, these results suggest that the NCM is a possible component of the proposed template mechanism that guides vocal learning. We can speculate that, once the tutor song memory is acquired in a juvenile bird, NCM neurons will respond differently to self-produced vocalizations that match the tutor song versus nonmatching vocalizations. If these differential responses bias future song attempts, the tutor song memory will tend to shape vocal development toward the acquired acoustic categories that are the components of the tutor song. This speculation is consistent with current conceptions of how human infants may use the acquired phonological categories of the language to which they are exposed to guide speech development, and suggests, at least in principle, that the basic sensory mechanisms that underlie vocal learning may be shared across species.

## Methods

**Animals and Behavioral Training.** As outlined in Fig. 5, zebra finches (*Taeniopygia guttata*) were raised in family breeding cages with their parents and siblings in a common room until approximately phd 10. At this age, which is before the onset of the sensory period of song learning, juveniles and mothers (females do not sing) were placed in isolation chambers in family groups without adult males to prevent exposure to adult song. Mothers were retained to provide supplemental feeding, needed until  $\approx 25$  phd. The sensory period during which songs may become memorized occurs between  $\approx 25$  and 60 phd, and the sensorimotor period of song production occurs between  $\approx 35$  and 90 phd (see Fig. 5). Juvenile males ( $n = 20$ ) were then separated from their mothers

and siblings and placed into individual isolation chambers at the ages of 23–27 phd. Isolation chambers contained a plastic replica of an adult male, a speaker, a microphone, and two behavioral keys. Continuous recordings of all vocalizations, from the time of individual isolation to neurophysiological assessment, were made with SOUND ANALYSIS PRO V1.04 (25).

These juvenile males were then either tutored (playback of 20 presentations daily of two motifs of the tutor song in response to key pecks;  $n = 9$ ) or untutored (maintained in auditory isolation;  $n = 5$ ) throughout the normal period of song development consisting of the sensory and sensorimotor stages of song imitation (start of tutoring 43–47 phd; end of tutoring 80–90 phd, Fig. 5). All birds had access to the keys and their auditory environment was limited to their own vocalizations (for all intact birds), the tutor song (for the tutored birds), and environmental noises, e.g., bird hops, air pump hiss, cage bumps. Posttutoring, all birds were then held in isolation for >30 days pending neurophysiological study in adulthood. Previous experiments have shown that memories in the NCM for conspecific songs typically last up to 2 days, and up to 4 days under special conditions (22). The delay of 30+ days between tutor song exposure and NCM electrophysiology was designed to provide an “auditory washout” period of recent memory for the tutor song. All rearing, training, and experimental procedures followed the guidelines approved by the Animal Care and Use Committee at Rutgers University.

Data were also collected from three additional groups of males. To assess whether acquisition of BOS influences the memory of the tutor’s song, one group of birds was permanently devoiced before the sensorimotor period (surgery at 25–37 phd;  $n = 6$ ). This group was then exposed to the same tutoring regime as the intact juveniles. To test the contributions of the bird’s age versus tutoring conditions to the observed results, another group of adult males, ( $\geq 110$  phd, raised in individual family cages in the general aviary,  $n = 3$ ) was isolated and exposed to the tutor song for 40–50 days under the same conditions as the juvenile birds, followed by an isolation period of 30 days before neuronal recording. To test the efficacy of our experimental tutoring conditions versus normal social tutoring, a third group of intact males ( $n = 11$ ) was raised in individual family cages (with father, mother, and siblings) in the general aviary. When they reached adulthood, their songs were recorded and compared with their father’s songs.

**Devoicing Surgeries.** Birds were anesthetized with Nembutal (50 mg/kg), and the syrinx was exposed through an incision in the interclavicular (i.c.) air sac. An opening was made in the trachea rostral to the syrinx, continuing midsagittally between the paired ventralis muscles (modified from ref. 32). This exposed subsyringeal and supersyringeal airways to the same respiratory pressure within the i.c. air sac, presumably eliminating the pressure differential and thus air flow through the syrinx. After this procedure, males engaged in song-like beak movements and did not phonate, but on rare occasions produced short breathy sounds.

**Electrophysiology and Stimulus Characteristics.** We measured multiunit habituation rates of auditory responses in the NCM for all groups in adulthood ( $\geq 110$  phd). Neurophysiology procedures were performed and recording sites verified as described (21, 22, 33). Briefly, 2 days before each acute recording session, under Nembutal anesthesia (50 mg/kg), a small craniotomy was made in the skull over the bifurcation of the midsagittal sinus. Dental cement was used to form an access chamber and to affix a stainless-steel post to the skull. For recording sessions, birds were awake but gently restrained in a custom holding tube with the steel post fixed to a stereotaxic

frame. Before recording data, the NCM was located physiologically, by using a white noise stimulus, shaped with the amplitude envelope of zebra finch song, as a search stimulus. A multiple microdrive advanced 7 tungsten electrodes (1–4 M $\Omega$  impedance; Thomas Recording, Ekhorn Design, Giessen, Germany) independently into the brain to record extracellular neuronal activity (see examples in Fig. 2A). The neuronal signals were then amplified, filtered below 500 Hz and above 5 kHz (Amplifier 440; Brownlee Precision, San Jose, CA), and digitized (Power1401 signal processor, SPIKE2 software; Cambridge Electronic Design, Cambridge, U.K.) for further analysis. All experiments were conducted in a double-walled sound booth (Industrial Acoustics, Bronx, NY).

Auditory stimuli included the tutor song, BOS (adult song), and BOS<sub>juv</sub> in nondevoiced birds, and the songs of three novel conspecifics. BOS stimuli were constructed from motifs, stereotyped syllable sequences characteristic of zebra finch songs, selected as the most representative from recordings of the bird’s song, made before crystallization (BOS<sub>juv</sub> at day 75) and in adulthood during the week before electrophysiological assessment (BOS). All stimuli were digitally equated for loudness and similar in duration to the tutor song. Stimuli were presented at 65 dB sound pressure level (model 712, A scale; Larson-Davis, Provo, UT) via a speaker located 0.5 m in front of the bird. Stimuli were presented in a pseudorandom order (25 repetitions of each stimulus, two motifs per stimulus, 8-s interstimulus interval, 44-kHz playback rate; all stimuli were presented in each recording session).

**Data Analysis.** SOUND ANALYSIS PRO (SAP) was used to assess the degree of imitation for each intact bird by calculating the SI between one motif of the tutor song and one motif of BOS (25). SAP calculates the SI between two songs based on a comparison of multiple features, producing an objective quantification of the similarity between the tutor’s song and BOS.

Based on histology and neuronal response properties, only data from electrodes determined to be in the NCM were used for further analysis. We briefly describe our data analysis methods here (details in refs. 21 and 33). The rms was used to calculate the amplitude of the multiunit activity during two time periods: a baseline period (500 ms before each stimulus), and a response period (from stimulus onset to offset plus 100 ms). Response magnitude was calculated by subtracting the baseline rms from the response rms.

Auditory responses to a novel stimulus in the NCM are initially vigorous, but habituate rapidly when the stimulus is repeated (Fig. 2A) (21). When responses are plotted as a function of repetition number (Fig. 2B), the normalized slope of the linear portion of the resulting function is the habituation rate, which is higher for novel stimuli than for familiar stimuli, thus enabling the familiarity of a stimulus to be quantitatively assessed. We calculated the habituation rate at each electrode for each stimulus. This metric provided a quantification of the degree a stimulus is remembered (the steeper the rate the less familiar; the shallower the rate, the more familiar).

Using these habituation rates, we calculated a FI at each electrode for each test song (tutor song, BOS or BOS<sub>juv</sub>) by dividing the habituation rate for that stimulus into the mean habituation rate for a set of novel stimuli (three different conspecific songs) tested at the same time. This ratio quantified the familiarity of the tutor stimulus in contrast to a novel stimulus. If FI  $\approx 1$ , this implies equal novelty for the tutor song and the novel songs; if FI > 1, this implies that the tutor song is more familiar than the novel songs. For each bird, we averaged the FI across electrodes to achieve a mean FI for each stimulus (FI<sub>tutor</sub> or FI<sub>BOS</sub>) in each bird that was used for further analyses.

Two-tailed *t* tests assuming unequal variances were used to

