Functional MRI of the zebra finch brain during song stimulation suggests a lateralized response topography

Henning U. Voss*†, Karsten Tabelow‡, Jörg Polzehl‡, Ofer Tchernichovski§, Kristen K. Maul§, Delanthi Salgado-Commissariat¶, Douglas Ballon*, and Santosh A. Helekar¶

*Citigroup Biomedical Imaging Center, Weill Medical College of Cornell University, 516 East 72nd Street, New York, NY 10021; ‡Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany; §Department of Biology, City College of New York, 138th Street and Convent Avenue, New York, NY 10031; and ¶The Methodist Neurological Institute, 6560 Fannin Street, Suite 902, Houston, TX 77030

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved May 8, 2007 (received for review December 22, 2006)

Electrophysiological and activity-dependent gene expression studies of birdsong have contributed to the understanding of the neural representation of natural sounds. However, we have limited knowledge about the overall spatial topography of song representation in the avian brain. Here, we adapt the noninvasive functional MRI method in mildly sedated zebra finches (*Taeniopygia guttata*) to localize and characterize song driven brain activation. Based on the blood oxygenation level-dependent signal, we observed a differential topographic responsiveness to playback of bird's own song, tutor song, conspecific song, and a pure tone as a nonsong stimulus. The bird's own song caused a stronger response than the tutor song or tone in higher auditory areas. This effect was more pronounced in the medial parts of the forebrain. We found left–right hemispheric asymmetry in sensory responses to songs, with significant discrimination between stimuli observed only in the right hemisphere. This finding suggests that perceptual responses might be lateralized in zebra finches. In addition to establishing the feasibility of functional MRI in sedated songbirds, our results demonstrate spatial coding of song in the zebra finch forebrain, based on developmental familiarity and experience.

imaging | learning | memory

Birdsong is studied as a model of vocal learning, perception, production, and motor abnormalities of speech (1, 2). There are interesting parallels between song development and speech development (3) and between auditory and vocal pathways in the songbird and human brain (4). Therefore, insights from experiments on songbirds may contribute to the understanding of auditory and vocal function in humans. For example, minimal models of speech dyspraxia (5) and dysfluencies such as stuttering are being developed in zebra finches (6). Zebra finches are capable of learning, producing, perceiving, and discriminating complex sound patterns. Birdsong in zebra finches consists of a sequence of distinctive sounds produced by males and is characterized by a consistent and reproducible acoustic profile. Song is learned by imitating the song of an adult conspecific tutor during a sensitive period of development (7–10). Recently, it has been shown that songbirds are able to learn recursive syntactic patterns, presumably a simple form of grammar (11), thus extending the potential applicability of the birdsong model to our understanding of the biological basis of languages.

Several brain structures are required for learning, production, and perception of birdsong. It is known from electrophysiological studies that song learning nuclei, such as the lateral magnocellular nucleus of the anterior nidopallium (LMAN) and X (12), play an important role in song development. In parallel with song motor learning, auditory song selectivity gradually emerges during development (13–17). Robust sensory responses to auditory stimuli have been recorded in the primary auditory area in the caudal telencephalic region (field L), the caudomedial nidopallium (NCM), the caudal mesopallium (CM), and the caudomedial ventral hyperstriatum (18–21), as well as in the song nuclei HVC, LMAN, X, and nucleus interface of the nidopallium (22–28). Sensory representation of birdsong in the song nuclei and the secondary auditory areas NCM and CM is characterized by response selectivity to song ownership, familiarity, and species-specific features. For instance, neurons in these structures are more responsive to bird's own song (BOS) and tutor song (TUT) than to conspecific song (CON) and to CON compared with heterospecific song (26).

Electrophysiological recordings of multiple and single units have provided high temporal resolution regarding stimulus-specific sensory responses in functionally specific brain nuclei. Auditory-evoked responses recorded on the surface of the brain have also revealed temporal information about activation of the brain as a whole (29). However, because these methods do not provide 3D spatial resolution of relevant brain substrates, we have a poor understanding of the topography of sensory activity, both within each auditory area or song nucleus and more globally. Experiments examining the spatial patterns of song stimulation-induced upregulation of immediate early gene products such as ZENK have been successful in addressing this deficiency (30, 31). For example, Ribeiro *et al.* (32) have studied the topographic organization of song syllables in the canary NCM. ZENK expression also shows higher response of neurons in NCM to CON vs. heterospecific song (33).

In this study, we examine the spatial pattern of brain activation in response to auditory stimuli by adapting the blood oxygenation level-dependent (BOLD) functional MRI (fMRI) method (34) in mildly sedated zebra finches. This method enables us to obtain better spatial resolution and localization of neural representation of birdsong than in electrophysiological recordings and to investigate the effect of various stimuli on the same individual. The feasibility of fMRI on songbirds has been demonstrated recently in anesthetized European starlings (35). It was shown that a sensory BOLD response exists, is stable over time, and causes specific activation of auditory areas of the brain in response to auditory and song stimuli. In this experiment in awake, mildly sedated zebra finches we ask whether there are differences in the spatial distribution of stimulus-dependent activation based on species-specific stimulus saliency,

Author contributions: D.B. and S.A.H. contributed equally to this work; H.U.V., D.S.-C., D.B., and S.A.H. designed research; H.U.V., K.K.M., and S.A.H. performed research; H.U.V., K.T., and J.P. contributed new reagents/analytic tools; H.U.V., K.T., and J.P. analyzed data; and H.U.V., K.T., J.P., O.T., K.K.M., D.S.-C., D.B., and S.A.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: LMAN, lateral magnocellular nucleus of the anterior nidopallium; NCM, caudomedial nidopallium; CM, caudal mesopallium; CON, conspecific song; BOS, bird's own song; TUT, tutor song; BOLD, blood oxygenation level-dependent; fMRI, functional MRI; TONE, pure tone of 2-kHz frequency; EPI, echo-planar imaging; PS, propagation–separation.

†To whom correspondence should be addressed. E-mail: hev2006@med.cornell.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0611515104/DC1.

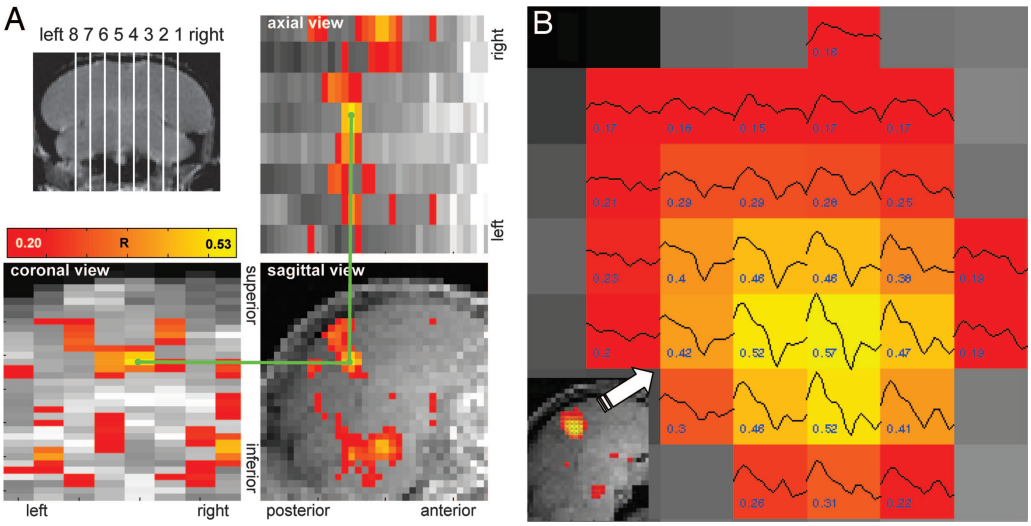

Fig. 1. Location and time trace of main BOLD response. (*A*) Slice midlines used to prescribe the fMRI scans (*Upper Left*) and maximum intensity projections of color-coded correlation coefficients $R > 0.2$ ($P < 0.001$), along three orthogonal views, of the brain of a male zebra finch stimulated with a CON (*Upper Right* and *Lower*). For demonstration, the green lines connect corresponding voxels in the three views. (*B*) Cluster of activated voxels in the forebrain area close to the sagittal midline (*Inset*) of a male zebra finch stimulated with his own song. Colors code the correlation coefficients of the response with the stimulation function, also given as numbers ($R > 0.16$, $P < 0.005$). Shown are averaged BOLD response time series with a characteristic shape as expected for this block design paradigm. The average was taken over the 16 stimulation blocks.

ownership of song stimulus, and experience-based familiarity of song stimulus. Accordingly, we image sensory BOLD responses to a BOS, TUT, CON, and a nonsong pure tone and determine their differential spatial patterns of functional activation in the zebra finch brain.

Our results provide insights into the 3D representation of birdsong in the zebra finch brain and clearly establish the feasibility of fMRI in awake songbirds.

Results

fMRI Scanning of the Awake Zebra Finch Brain. We performed fMRI in 16 awake, mildly sedated male adult zebra finches during auditory stimulation in a 3.0-T MRI scanner. The auditory stimuli were a pure tone of 2-kHz frequency (TONE), a CON, the BOS, and the TUT. Visual inspection of time traces averaged over all stimulation blocks immediately revealed BOLD responses to auditory stimuli in all birds. In most birds, clearly visible stimulus-evoked activations could also be seen by comparing the "on-off" stimulus indicator function with time traces in voxels with a large correlation coefficient between the time traces and the stimulus indicator function. The maximum positive correlation coefficient observed was $R = 0.78$ ($P < 10^{-16}$). Using the first modeling approach as described in *Methods*, all 16 birds showed significant and reproducible stimulus-evoked BOLD activation clusters within the forebrain.

Fig. 1*A* shows a representative maximum intensity projection of significantly active voxels for the whole brain in three orthogonal views. The BOLD response is seen at similar locations in both hemispheres, with a pronounced caudal bilateral cluster in the medial slice closest to the midline, and extending into the slice adjacent to it. This cluster was present in 63 of the 64 scans performed and presumably includes parts of field L, NCM, and CM. The BOLD response time series within that cluster (Fig. 1*B*) has a characteristic shape with a sharp rise at the beginning and a negative undershoot in the off part of the on–off stimulation block. Another significant bilateral cluster is seen in a more ventral position at the location of the midbrain. In addition to these consistently observed clusters, more variable activations were seen in other brain areas. Activation in slices 1 and 8 appeared to be most variable, presumably because of partial volume effects with areas outside the brain and are therefore not used in the following. Fig. 2*A* shows for all stimuli averaged activation clusters from the outer parasagittal slices 2 and 7 (lateral) to the inner parasagittal slices 4 and 5 (medial). Before averaging, all data were approximately geometrically normalized to a template brain. In the following, we first describe topographical properties of the BOLD response (location and extent) and then the amplitude of the BOLD response.

Stimulus-Dependent Differentiation of Sensory BOLD Response Topography in Auditory Areas. Focusing on the medial brain slices 4 and 5 in Fig. 2*A*, the region of the brain that shows most pronounced, consistent, and reproducible patterns of activation, we see distinct differences in the distribution of the voxels activated in response to different stimuli. In the averages over all birds, the largest contiguous area of activation is seen with TUT and the smallest with the unfamiliar CON stimulus. The two developmentally familiar songs, namely TUT and BOS, show a widening of the caudal-most extension of the activated area, which presumably corresponds to the secondary auditory areas CM and NCM, and their input field L subregions L1 and L3.

Fig. 2*B* shows differential profiles for all combinations of stimuli in the two medial slices, averaged over the two hemispheres. Most evident is a shift of the activation toward more caudal regions from TONE to BOS (red area, positive change vs. blue area, negative change from TONE to BOS); TUT shows much more pronounced activation throughout the activated region when compared with TONE, and greater amount of activation in the central and rostral field L portion when compared with BOS. In the latter comparison, TUT and BOS show nearly equal amounts of activation in the wider posterior caudal area that corresponds to NCM.

A simple cluster analysis (Fig. 2*C*) of the caudomedial region, based on the general linear model coefficients and their estimation errors of all 16 birds and all four stimuli, yielded no clusters 6 times, one cluster 23 times, two clusters 34 times, and four clusters once. The shape of the clusters was variable. There was no dependence between the number of clusters and the stimulus used. Of the 35 cases where two or more clusters were found, in 23 cases the cluster with the stronger activation covered more likely field L than NCM (five cases showed the opposite behavior, i.e., weaker activation in field L; the remaining seven cases were inconclusive). All field L

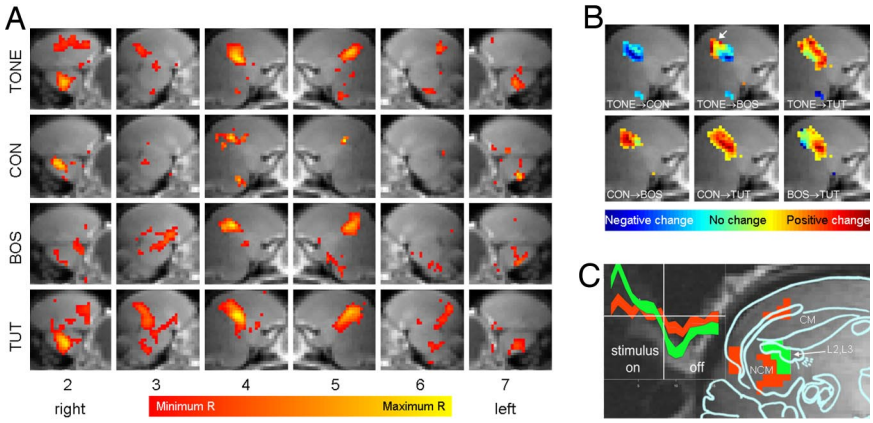

Fig. 2. Differential topography of activation of auditory areas in response to different stimuli. (*A*) Averaged functional activations depending on stimulus and sagittal slice position. The slice numbers as corresponding to Fig. 1 *A* are given at the bottom. Functional activations are shown as average correlation coefficients. The grayscale background consists of a representative anatomical volume (smoothed and averaged over both hemispheres). (*B*) Comparison of averaged activation for different stimuli. Positive changes of the correlation coefficient from stimulus 1 to stimulus 2 are shown in red, and negative changes are in blue, as indicated with the color bar. For example, the stimulation with BOS yields more posterior (red area, arrow) and less anterior (blue) activation than the stimulation with TONE. (*C*) Example of differences in the shape of the mean BOLD response between primary and secondary auditory areas for stimulation with TUT. Using a simple clustering mechanism based on the strength of activation, the activated area segments into two (noncontiguous) clusters shown in red and green, presumably corresponding to activations in field L (green) and other areas including NCM (red). The corresponding averaged time series within the two clusters, shown in *Inset* in the same color as the clusters and with 66% confidence intervals, show a distinctively different time course. The anatomical overlay was redrawn from ref. 42.

activations (green area in Fig. 2*C*) had a stronger undershoot of the BOLD response in the off part of the stimulation than NCM (posterior red areas in Fig. 2*C*) and CM (anterior red area in Fig. 2*C*) activations. The onset of all but six BOLD response time series occurred earlier in field L than in NCM.

Discrimination Between Stimuli in the Forebrain as a Whole. A quantification of the response by measuring the time-averaged BOLD response amplitude relative to the mean signal intensity for significantly activated voxels ($P < 0.005$) gave the following values: the maximum time-averaged BOLD response amplitude found in the whole brain was similar across all stimuli: 4.4% (TONE, CON, BOS) and 4.5% (TUT). The average of the strongest activated voxel in each bird (average over all 16 birds), however, was very interesting: 2.7% (TONE), 2.8% (CON), 2.6% (TUT), but 3.7% for BOS. Not surprisingly, this effect stems mostly from high auditory nuclei, so that in the area containing field L, NCM, and CM, the average of the strongest activated voxel in each bird (average over all 16 birds) was 1.8% (TONE), 1.9% (CON), 1.9% (TUT), and again very high, 2.7% to BOS. The maximum time-averaged individual BOLD response amplitude in this area was 3.3% (TONE), 2.7% (CON), 3.5% (TUT), and 4.4% (BOS). The standard deviations of these measures ranged from 0.2% to 0.5% of the mean signal intensity.

A separation of the brain into medial and lateral parts reveals differential selectivity of the medial parts with respect to the stimulus. Fig. 3*A* shows a comparison of BOLD response amplitudes of different stimuli for the two medial slices 4 and 5 and the lateral slices 2 and 7. A single-factor ANOVA test across the stimuli yielded *P* values of 0.0006 and 0.5 for the medial and lateral slices, respectively. In particular, the response to TONE stimulation is significantly smaller than to CON or BOS stimulation, and response to TUT is weaker than to BOS (Table 1; Wilcoxon signed rank test, $P < 0.05$). In contrast, in the two more lateral slices 2 and 7, the only significant discrimination found was between TUT and CON stimulation. In these slices, CON yielded the strongest response amplitude but more variable topography. To specify the medial response topography, slices 4 and 5 were subdivided into three mutually exclusive regions: region 1 containing the caudal areas with NCM and field L, region 2 containing the cerebellum and ventral parts of the brain, and region 3 containing the rostral forebrain. Only region 1 showed a stimulus-dependent BOLD response ($P = 0.05$, 0.36, and 0.30 for regions 1, 2, and 3, respectively; single-factor ANOVA). A two-factor ANOVA was significant with respect to the brain region ($P = 6 \times 10^{-6}$) and stimulus type ($P = 0.02$) but not with respect to the interaction between brain region and stimulus type ($P = 0.7$).

Finally, a comparison between the left and right hemispheres showed that the average BOLD response is of comparable ampli-

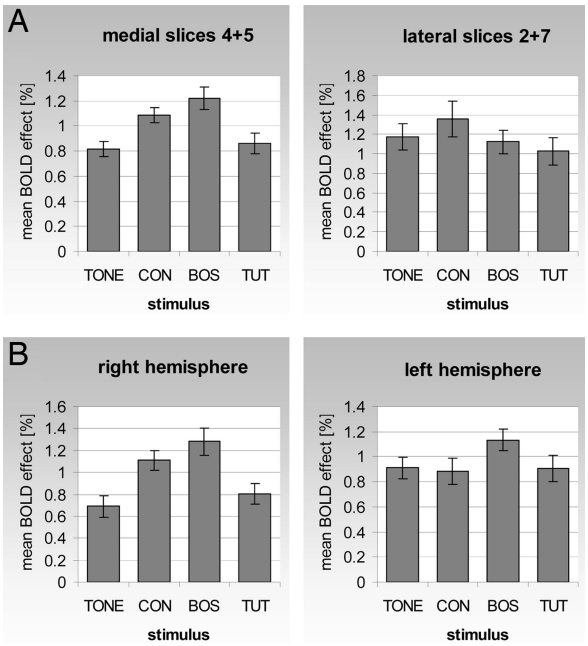

Fig. 3. Larger mean BOLD response amplitude to BOS compared with TUT and right hemispheric bias in stimulus-dependent topographic differences. (*A*) Mean and standard error of the BOLD response amplitude over all activated voxels ($P < 0.005$) in medial slices 4 and 5 and lateral slices 2 and 7, for all 16 birds. The *P* values of a single-factor ANOVA test across the stimuli are $P = 0.0006$ for the medial and $P = 0.5$ for the lateral slices. (*B*) The same for the right medial slice vs. the left medial slices; $P = 0.0005$ and 0.2, respectively.

Table 1. Statistical significances for pairwise differences in the medians of the BOLD response amplitude for all stimuli pairs and medial slices 4 and 5, lateral slices 2 and 7, right medial slice 4, and left medial slice 5 (Wilcoxon signed rank test)

Stimulus pair	Medial	Lateral	Right medial	Left medial
CON–TONE	0.01	0.3	0.02	0.7
BOS–TONE	0.0008	0.6	0.002	0.02
TUT–TONE	0.4	0.6	0.5	0.8
BOS–CON	0.2	0.2	0.4	0.09
TUT–CON	0.07	0.04	0.06	0.7
TUT–BOS	0.005	0.4	0.007	0.09

tude. However, a test of significant differences in dependence of the stimulus showed that only the right hemisphere discriminates between different stimuli, in the aforementioned sense (Fig. 3*B*). The BOLD response in the right hemisphere represented by slice 4 discriminates TONE and TUT from BOS (single-factor ANOVA across stimuli, right hemisphere $P = 0.0005$, left hemisphere $P = 0.2$). A two-factor ANOVA with stimuli and hemisphere (left/right) as independent factors, however, was only significant with respect to stimulus ($P = 0.0003$) but not with respect to hemisphere ($P = 0.9$) nor interaction between hemisphere and stimulus ($P = 0.09$). The BOLD responses in the left and right lateral slices, 2 and 7, again do not discriminate between the different stimuli (single-factor ANOVA across stimuli, left hemisphere $P = 0.7$, right hemisphere $P = 0.08$). As far as topographic distribution between the left and right hemisphere is concerned, the distribution is variable, but on average there is a larger spatial pattern of activation on the right (Fig. 2*A*).

In summary, areas in the two medial slices containing predominantly NCM, CM, and field L show greater sensitivity to BOS/CON compared with TONE and to BOS compared with TUT. There is no significant difference in mean BOLD response amplitudes between stimuli in the lateral slices. Activations in the right brain hemisphere discriminate better between stimuli than those in the left hemisphere, and there appears to be a more variable but wider representation of stimuli on the right side.

Discussion

Our findings in mildly sedated zebra finches further extend the recent observations in anesthetized European starlings (35). The similarities between the BOLD response in these two species of songbirds include: (*i*) a consistent and reproducible activation of NCM and field L, (*ii*) a higher responsiveness of NCM to song than to perceptually nonsalient auditory stimuli, and (*iii*) similar characteristics of the BOLD response time course within these two areas, namely an earlier onset and stronger undershoot in field L as compared with NCM. In our experiment we have been able to study responses in the awake, mildly sedated state, to a wider range of natural song stimuli, and in a larger brain volume in zebra finches.

fMRI of the Adult Zebra Finch Reveals Strong Auditory Stimulus-Evoked Activation Patterns. We have demonstrated that it is feasible to obtain localized sensory BOLD responses in awake, mildly sedated zebra finches by using a 3.0-T MRI scanner. Widespread functional activity is seen across the forebrain and midbrain. In particular, we could identify BOLD responses in forebrain loci that presumably correspond to the auditory areas NCM, CM, and field L. In addition, significant activations were found in lateral slices that contain parts of the midbrain, HVC/HVC shelf, and robust nucleus of the arcopallium (RA)/RA cup. The latter activations, however, were spatially more variable and diffuse, and we were not able at this point to make clear assignments to specific regions. It is worth mentioning that methods of immediate early gene expression, which have a considerable higher resolution than the method used here, reveal activity after song playback in some of these areas (in particular, HVC shelf, RA cup, and nucleus mesencephalicus lateralis pars dorsalis in the midbrain) (30, 36, 37).

The BOLD response amplitudes in the caudomedial brain region depended on stimulus type used. In particular, there were significant differences in the BOLD response amplitude for a pure tone versus CON and BOS activation, and the TUT response was different from the BOS response. We could also demonstrate that the auditory responses in the forebrain show high sensitivity to BOS [as was found in electrophysiological recordings in anesthetized birds (22, 26, 28, 38, 39)] and TUT in areas corresponding to the higher auditory area NCM [as was found in awake birds (40)], but that other stimuli activate the brain as well.

Familiar Song Stimuli Show Selective Differential Topography and Lateralization in Auditory Areas. The observation that stimuli with a high degree of developmental experience-based familiarity, namely TUT and BOS, show a wider extent of BOLD responsiveness in areas corresponding to the NCM and CM, and their input field L subregions L3 and L1, is highly consistent with the functional circuitry and electrophysiological data. The neuroanatomy of the songbird brain indicates that NCM and CM are secondary auditory areas with immediate bidirectional connections with L3 and L1, respectively (41, 42). These areas are likely to be modulated by experience and learning. Indeed, unit recordings show that neuronal responses are more selectively tuned to learned vocal sounds in NCM (20, 21) and CM (43–45), whereas the primary auditory subregions L2a and L2b are responsive to sounds within the wider species-specific spectrotemporal range (24, 46, 47). Measurement of long-term response habituation in NCM, by both electrophysiology and the study of the song stimulation-induced up-regulation of ZENK, has suggested that this area might encode the long-lasting sensory memory of the TUT (31, 40). Other experiments point to the fact that NCM and CM might be involved in short-term plasticity related to song discrimination (18, 19, 48). The greater representation of TUT and BOS revealed by fMRI in our experiments therefore might reflect an important aspect of the sensory memory for these developmentally salient familiar stimuli.

The observation that better discrimination between stimuli (measured as the mean amplitude of the BOLD response in significantly activated areas) is seen in the right hemisphere suggests a possible lateralization of the mechanisms underlying song perception. Bilateral asymmetry with either left-sided or right-sided dominance, or differential dual specialization, has been previously observed in songbirds in relation to both central and peripheral control of song production (49–52). As far as song perception is concerned, increased neuronal responsiveness to behaviorally relevant song stimuli has been observed in the field L complex and HVC on the right side in starlings (53, 54). Right hemispheric specialization is seen only in awake birds. These observations therefore are in close agreement with our fMRI findings in awake but mildly sedated zebra finches. Functional lateralization of this type is of obvious significance from the perspective of birdsong as a model of speech, a strongly lateralized human behavior.

fMRI in Songbirds as a Reliable Research Tool. Any fMRI experiment on animals differs from natural conditions. The background noise cannot be shielded completely and may affect the overall functioning of auditory pathways. However, by using a block design paradigm in which only responses with stimulus on and stimulus off are compared, it is assumed that a sufficient amount of the background noise effect is subtracted out, as in auditory fMRI studies in humans that use continuous scanning paradigms. An indication that this is true to some degree is that even pure tone stimulation at a frequency within the spectrum of the scanner noise still caused substantial BOLD activation. A caveat to our approach could be that sedation with Diazepam may enhance inhibitory activity in

some parts of the brain, and we cannot exclude that additional areas may have been activated without the use of Diazepam. However, this effect of Diazepam, albeit at doses higher than those that we used, has been shown to make electrophysiological neuronal responses in HVC less variable and more selective to BOS as compared with wakefulness (55). Further, as described in supporting information (SI) *Text*, SI Fig. 5, and SI Fig. 6, surface-evoked potential responses and field potential responses in NCM to the stimuli are not significantly altered by Diazepam at the dose used. The restraint of the birds may cause stress and possible changes in neuromodulator levels, which again may affect the BOLD response. There are remaining technical challenges and room for improvements to make fMRI in songbirds a reliable research tool. The details of the coupling of the BOLD response to neuronal activation in songbirds are unknown, and therefore, optimal timing and delivery of stimuli should be further investigated. The anatomical mapping of active areas could be further improved by the development of automatic registration methods suitable for the avian brain, which does not have anatomical landmarks as significant as the human brain. This would make the assignment of observed activations to anatomical structures more reliable, for example, in our case field L and NCM, which were based to some degree on function itself.

In conclusion, we have demonstrated that auditory fMRI of the awake zebra finch can reveal details about the 3D topography of neural correlates of song perception. It suggests a differential encoding underlying experience-based discrimination between song stimuli and a right hemispheric bias in auditory processing. Even though there is still room for improvement, the noninvasive nature of fMRI holds the promise of conducting within-subject longitudinal studies of the development of neural correlates of song perception.

Methods

Preparation of Birds. Sixteen male juvenile zebra finches were live-tutored by male adult zebra finches. Each group of two to six birds was raised with one tutor from posthatch day ≈15 through posthatch day ≈100. At the time of scanning the birds were ≈24–48 months old. They were sedated with 40 microl Diazepam (Abbott Labs, Abbott Park, IL) i.m. (1.66 mg/ml Diazepam in normal saline solution) 10 min before MRI scanning. The dose of Diazepam used for sedation was ≈5 mg/kg body weight (see also *SI Text*). The fMRI experiments lasted for <2 h after Diazepam injection. After the experiments, the birds appeared to be still sedated. To minimize the influence of the sedation level, we randomized the order of the stimulus application. After sedation, the birds were immobilized in a restraining device made of soft transparent Tygon plastic tubes (Saint-Gobaine Performance Plastics, Beaverton, MI) and a solid plastic tube (Kendall, Mansfield, MA) (Fig. 4). The birds were placed in a foam/rubber compound sound isolation box, and auditory paradigms were delivered by using a flash memory music player (Samsung, Seoul, Korea), a headphone volume booster (RadioShack, Fort Worth, TX), and a pair of stereo headphones (CV-200; COBY, Maspeth, NY) with the magnets removed. The two headphone parts were randomly exchanged between the right and left side. The distance to the bird's head was 4 cm. The sound pressure level of the auditory stimuli at the head position was ≈100 dB, and the background noise during the echo-planar imaging (EPI) sequence was ≈83 dB. The experiment was approved by the Institutional Animal Use and Care Committees of Cornell University and The Methodist Hospital Research Institute/Texas A&M Institute of Biotechnology.

MRI Parameters. Images were acquired on an Excite 3.0 T scanner (GE, Waukesha, WI) with an in-house built solenoid transmit/receive coil of 20-mm length and 15-mm inner diameter (Fig. 1*B*). BOLD-sensitive fMRI sequences of images were acquired by using a four-shot 2D gradient echo EPI sequence with repetition time/echo time = 1,000/25 ms. The effective repeat time per volume was 4 s. Eight sagittal slices of 1.0-mm thickness, 4 cm field of view (FOV) (phase FOV = 0.75), flip angle 70°, and a matrix size of 128 × 64 (zero filled to 128 × 128) were acquired with gradient ramp sampling. Slices were prescribed from right to left (Fig. 1*A*). The scan time per experiment was 1,024 s (256 repeats). Additionally, localizers, in-plane anatomical images, and field correction maps were acquired.

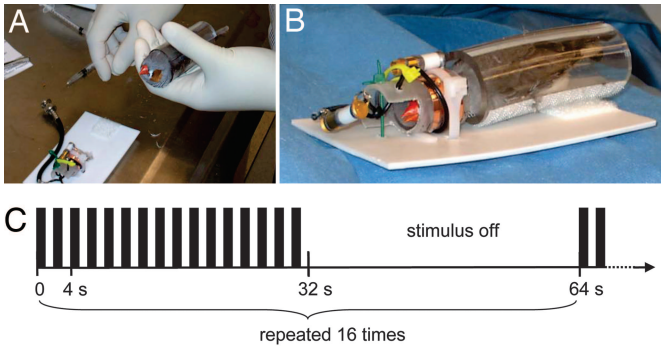

Fig. 4. Experimental setup and stimulation paradigms. (*A*) Immobilization of mildly sedated zebra finch in a flexible Tygon plastic tube. (*B*) Bird in the MRI radiofrequency coil. (*C*) Timing of stimuli. Here the pure tone stimulus TONE is shown. It consists of 1-s on and 1-s off tones, repeated twice within a sampling interval of 4 s. The sampling interval is repeated eight times (32 s), followed by a silent block of 32-s duration. The whole on–off block of 64 s is repeated 16 times, yielding 17:04 min stimulation time per experiment. For song stimulation (CON, BOS, TUT), the song is played twice within a sampling interval of 4 s, and the gaps between the repeated song playouts are adjusted accordingly.

Paradigms. All stimuli were delivered in 16 blocks, each consisting of a 32-s on and a 32-s off segment, totaling 1,024 s (Fig. 4*C*). All stimuli where normalized with respect to peak amplitude and played out 16 times (twice per sampling interval) during the on segment of each block, i.e., the number of stimuli per time was kept constant despite the different stimuli lengths. The auditory stimuli were TONE, CON, BOS, and TUT. The birds used for our experiments were obtained from the laboratories of O.T. and S.A.H. The BOSs and TUTs were highly dissimilar with the CON [similarity <20%, Sound Analysis program (55)], which was an unfamiliar song recorded from a bird in a different colony. Two CONs were tested before the experiment, yielding similar results, and one of these songs was used in all birds. All songs were female-directed songs recorded from birds that were >150 days old. All birds used in this study were tutored by a single tutor in groups of two to four. Seven different TUTs were used, depending on the exact tutoring history. The durations of the TONE and CON stimuli were 1,000 and 730 ms, respectively. The mean durations of BOS and TUT were 1,233 ± 406 and 1,519 ± 454 ms, respectively.

Postprocessing. BOLD-sensitive EPI images were corrected for distortions by using field correction maps and in-house software written in IDL (Research Systems, Boulder, CO) (56, 57). The images were then despiked and motion was corrected by using AFNI (58). After statistical modeling, as described below, the statistical parametric maps (SPMs) were registered to a brain template that consisted of the least distorted and most symmetric (with respect to the sagittal midline) EPI scan. The 2D registration was based on a locally affine but globally smooth transformation (59) estimated from the EPI data and applied to the SPMs and averaged BOLD response time series. All averaging of activities, as described for the two modeling approaches below, was done after this registration process. Activations in areas with an EPI intensity baseline <20% of the maximum slice intensity were discarded. It was not necessary to discard scans due to bulk motion which was

always small [unlike in our first experiment using mild anesthesia (60)]. Eye components were removed. Distortion corrected images were mapped to anatomical drawings (ref. 42 and B. Nixdorf and H.-J. Bischof, personal communication).

Statistical Modeling. Statistical modeling was done following two approaches: correlation analysis and general linear modeling with spatial adaptive smoothing. Correlation analysis is easier to reproduce and turned out to be more suitable for obtaining averaged statistical parametric maps, whereas general linear modeling with spatial adaptive smoothing yields a better effective resolution and rendering of individual activation areas, as well as a higher statistical accuracy.

Correlation analysis. As a first and easily reproducible modeling approach, data were smoothed slicewise with a 2D Gaussian filter (half-width 1.5 voxels), voxelwise detrended by subtracting a linear fit, and temporally smoothed by convolution with a binomial filter over three time points. Statistical significance of activation was defined voxelwise by correlating the signal intensity with the on–off block stimulation function. In the BOLD response time series (Fig. 1*B*), all 16 repeated blocks were averaged. For each voxel in the average plot in Fig. 2*A*, a sufficient condition for a voxel to be displayed was that at least 4 of the 16 birds had a correlation coefficient of $R > 0.16$, corresponding to $P < 0.005$ for each bird.

General linear model. As a statistically more advanced approach that was used for quantification and statistics, we also fitted general linear models to the data that take local trends and the expected hemodynamic response function (HDR) into account. The statistical parametric map for this response has been smoothed by using the propagation–separation (PS) approach (61–63) to achieve noise reduction without blurring the shape of the activation areas. The PS method naturally adapts to different shapes of activation areas by generating a spatial structure corresponding to similarities and differences between time series in adjacent locations. The general linear model was given by $y_t = \beta_0 + \beta_1 h(t) + \beta_2 t + \beta_3 t^2 + \varepsilon_t$, with y_t the signal in one voxel at time t, ε_t the residual, $\beta_{(.)}$ coefficient vectors, and $h(t)$ the expected BOLD response function, defined as the convolution of the block stimulation function with an idealized HDR. The HDR was modeled as:

$$h(t) = \left(\frac{t}{d_1}\right)^{a_1} \exp\left(-\frac{t-d_1}{b_1}\right) - c\left(\frac{t}{d_2}\right)^{a_2} \exp\left(-\frac{t-d_2}{b_2}\right),$$

with $a_1 = 6, a_2 = 12, b_1 = b_2 = 0.9$, and $d_i = a_i b_i, i = 1, 2, c = 0.35$, and t the time in seconds (64). Subsequently, PS was applied as described (63). From the smoothed map of coefficients β_1 and its estimated standard deviation, voxelwise t scores were computed. Discrimination of the response to the four stimuli was performed by measuring the mean BOLD response amplitude in percent of the mean signal amplitude within significant voxels for $t > 2.60$ ($P < 0.005$) for the regions as described in *Results* and subsequent application of a Wilcoxon signed rank test for zero median of the differences (Table 1). To account for the multiple testing problem, these results were verified by a single-factor ANOVA test with the stimulus as parameter, as described in *Results*. All ANOVA tests were repeated-measure tests with the individual 16 birds as repeated measure.

Cluster analysis. The PS approach contains a simple clustering mechanism by segmenting regions of activations by the fit coefficient β_1 and its standard deviation. Clusters were computed only within activations with a global, i.e., multiple test corrected, threshold of $P = 0.05$, by including voxels with a voxelwise threshold of $P = 0.005$. Clusters within the activated region presumably containing field L and NCM were counted. Clusters were evaluated with respect to their exact position relative to each other; clusters were grouped by their position as described as rostral-dorsal (presumably field L) vs. caudal-ventral (presumably NCM), and these groups were compared with amplitudes and onset latencies of the corresponding averaged time series of activation.

All computations were performed with in-house software written in R (The R Project for Statistical Computing, www.r-project.org), MATLAB (Mathworks, Natick, MA), and Excel 2003 (Microsoft, Redmond, WA) on personal computers and Unix workstations.

We thank the three reviewers for valuable comments. This work was supported by National Institutes of Health Grants DC04778-01A1 and MH073900-01 (to S.A.H.), the Deutsche Forschungsgemeinschaft Research Center Matheon (K.T.), and a Weill Medical College–The Methodist Hospital Research Institute collaboration grant (to H.U.V. and S.A.H.).

1. Marler P (1970) *Am Sci* 58:669–673.
2. Konishi M (1965) *Z Tierpsychol* 22:770–783.
3. Doupe AJ, Kuhl PK (1999) *Annu Rev Neurosci* 22:567–631.
4. Jarvis ED (2004) *Ann NY Acad Sci* 1016:749–777.
5. Scharff C, White SA (2004) *Ann NY Acad Sci* 1016:325–347.
6. Helekar SA, Espino GG, Botas A, Rosenfield DB (2003) *Behav Neurosci* 117:961–969.
7. Immelmann K (1969) in *Bird Vocalizations*, ed Hinde RA (Cambridge Univ Press, Cambridge, UK), pp 61–77.
8. Konishi M, Nottebohm F (1969) in *Bird Vocalizations*, ed Hinde RA (Cambridge Univ Press, Cambridge, UK), pp 29–48.
9. Price P (1979) *J Comp Physiol Psych* 93:260–277.
10. Clayton NS (1987) *Anim Behav* 35:714–721.
11. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) *Nature* 440:1204–1207.
12. Brainard MS, Doupe AJ (2002) *Nature* 417:351–358.
13. Doupe A (1997) *J Neurosci* 17:1147–1167.
14. Doupe A, Solis M (1997) *J Neurobiol* 33:694–709.
15. Solis M, Doupe AJ (1997) *J Neurosci* 17:6447–6462.
16. Volman S (1993) *J Neurosci* 13:4737–4747.
17. Nick TA, Konishi MJ (2005) *J Neurobiol* 62:469–481.
18. Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS (1995) *Proc Natl Acad Sci USA* 92:3406–3410.
19. Gentner TQ, Margoliash D (2003) *Nature* 424:669–674.
20. Stripling R, Volman SF, Clayton DF (1997) *J Neurosci* 17:3883–3893.
21. Terleph TA, Mello CV, Vicario DS (2006) *J Neurobiol* 66:281–292.
22. Doupe A, Konishi M (1991) *Proc Natl Acad Sci USA* 88:11339–11343.
23. Bonke B, Bonke D, Scheich H (1979) *Cell Tissue Res* 200:101–121.
24. Müller C, Leppelsack H (1985) *Exp Brain Res* 59:587–599.
25. Williams H, Nottebohm F (1985) *Science* 229:279–282.
26. Margoliash D (1986) *J Neurosci* 6:1643–1661.
27. Margoliash D, Fortune E (1992) *J Neurosci* 12:4309–4326.
28. Vicario D, Yohay K (1993) *J Neurobiol* 24:488–505.
29. Espino GG, Lewis C Rosenfield DB, Helekar SA (2003) *Neuroscience* 122:521–529.
30. Mello CV, Clayton DF (1994) *J Neurosci* 14:6652–6666.
31. Bolhuis JJ, Zijlstra GO, den Boer-Visser AM, Van der Zee EA (2000) *Proc Natl Acad Sci USA* 97:2282–2285.
32. Ribeiro S, Cecchi GA, Magnasco MO, Mello CV (1998) *Neuron* 21:359–371.
33. Mello CV, Vicario DS, Clayton DF (1992) *Proc Natl Acad Sci USA* 89:6818–6822.
34. Ogawa S, Lee T, Kay A, Tank D (1990) *Proc Natl Acad Sci USA* 87:9868–9872.
35. Van Meir V, Boumans T, De Groof G, Van Audekerke J, Smolders A, Scheunders P, Sijbers J, Verhoye M, Balthazart J, Van der Linden A (2005) *NeuroImage* 25:1242–1255.
36. Mello CV, Ribeiro S (1998) *J Comp Neurol* 393:426–438.
37. Velho TA, Pinaud R, Rodrigues PV, Mello CV (2005) *Eur J Neurosci* 22:1667–1678.
38. Margoliash D (1983) *J Neurosci* 3:1039–1057.
39. Volman SF (1996) *J Comp Physiol A* 178:849–862.
40. Phan ML, Pytte CL, Vicario DS (2006) *Proc Natl Acad Sci USA* 103:1088–1093.
41. Fortune ES, Margoliash D (1992) *J Comp Neurol* 325:388–404.
42. Vates GE, Broome BM, Mello CV, Nottebohm F (1996) *J Comp Neurol* 366:613–642.
43. Sen K, Theunissen FE, Doupe AJ (2001) *J Neurophysiol* 86:1445–1458.
44. Hsu A, Woolley SM, Fremouw TE, Theunissen FE (2004) *J Neurosci* 24:9201–9211.
45. Gill P, Zhang J, Woolley SM, Fremouw T, Theunissen FE (2006) *J Comput Neurosci* 21:5–20.
46. Grace JA, Amin N, Singh NC, Theunissen FE (2003) *J Neurophysiol* 89:472–487.
47. Cousillas H, Leppelsack HJ, Leppelsack E, Richard JP, Mathelier M, Hausberger M (2005) *Hear Res* 207:10–21.
48. Mello C, Nottebohm F, Clayton D (1995) *J Neurosci* 15:6919–6925.
49. Nottebohm F (1971) *J Exp Zool* 177:229–261.
50. Hartley RS, Suthers RA (1990) *J Neurobiol* 21:1236–1248.
51. Floody OR, Arnold AP (1997) *Horm Behav* 31:25–34.
52. Halle F, Gahr M, Kreutzer M (2003) *J Neurobiol* 56:303–314.
53. George I, Cousillas H, Richard JP, Hausberger M (2005) *J Comp Neurol* 488:48–60.
54. George I, Vernier B, Richard JP, Hausberger M, Cousillas H (2004) *Behav Neurosci* 118:597–610.
55. Cardin JA, Schmidt MF (2003) *J Neurophysiol* 90:2884–2899.
56. Chang H, Fitzpatrick JM (1992) *IEEE Trans Med Im* 11:319–329.
57. Voss HU, Watts R Ulug AM, Ballon D (2006) *MRI* 24:231–239.
58. Cox RW (1996) *Computers Biomed Res* 29:162–173.
59. Periaswamy S, Weaver JB, Healy DM, Jr, Rockmore D, Kostelec PJ, Farid H (2000) *Proc SPIE Int Soc Opt Engineer* 4119:1066–1075.
60. Voss HU, Maul K, Ballon D, Tchernichovski OT, Helekar SA (2006) *Proc Int Soc Magn Reson Med* 14:2132.
61. Polzehl J, Spokoiny V (2001) *J R Statist Soc Ser C* 50:485–501.
62. Polzehl J, Spokoiny V (2006) *Probab Theory Relat Fields* 135:335–362.
63. Tabelow K, Polzehl J, Voss HU, Spokoiny V (2006) *NeuroImage* 33:55–62.
64. Glover GH (1999) *NeuroImage* 9:416–429.