

Own-Song Recognition in the Songbird Auditory Pathway: Selectivity and Lateralization

Colline Poirier,¹ Tiny Boumans,¹ Marleen Verhoye,^{1,2} Jacques Balthazart,³ and Annemie Van der Linden¹

¹Bio-Imaging Laboratory, Department of Biomedical Sciences, and ²Vision Laboratory, University of Antwerp, B-2020 Antwerp, Belgium, and ³Research Unit in Behavioral Neuroendocrinology, Center for Cellular and Molecular Neurobiology, University of Liège, 4000 Liège, Belgium

The songbird brain is able to discriminate between the bird's own song and other conspecific songs. Determining where in the brain own-song selectivity emerges is of great importance because experience-dependent mechanisms are necessarily involved and because brain regions sensitive to self-generated vocalizations could mediate auditory feedback that is necessary for song learning and maintenance. Using functional MRI, here we show that this selectivity is present at the midbrain level. Surprisingly, the selectivity was found to be lateralized toward the right side, a finding reminiscent of the potential right lateralization of song production in zebra finches but also of own-face and own-voice recognition in human beings. These results indicate that a midbrain structure can process subtle information about the identity of a subject through experience-dependent mechanisms, challenging the classical perception of subcortical regions as primitive and nonplastic structures. They also open questions about the evolution of the cognitive skills and lateralization in vertebrates.

Key words: birdsong; auditory; lateralization; midbrain; functional MRI; cognition

Introduction

Like human speech, birdsong is a learned behavior that supports species and individual recognition. Behavioral studies indicate that songbirds perceive their own song differently from other conspecific songs (Cynx and Nottebohm, 1992; Cynx, 1993). At the neuronal level, the forebrain regions involved in song production and learning, the so-called song control system, are also auditory responsive and display a strong selectivity for the bird's own song (BOS) (Doupe and Konishi, 1991; Margoliash and Fortune, 1992; Janata and Margoliash, 1999). BOS selectivity is commonly expected to emerge at least at an intermediate level in the ascending auditory system (Prather and Mooney, 2004; Theunissen et al., 2004). Understanding how and where this exquisite selective response arises in the auditory system is critical for understanding auditory-vocal integration necessary for song learning and song maintenance. Previous electrophysiological studies have looked for this selectivity in the primary and secondary auditory telencephalic regions, but without success (Lewicki and Arthur, 1996; Janata and Margoliash, 1999; Amin et al., 2004; Shaevitz and Theunissen, 2007).

Electrophysiological techniques require having an a priori hypothesis about the localization of the neuronal substrate support-

ing the investigated process. They are also limited by the number of locations that can be sampled in one experiment. In contrast, functional MRI (fMRI), which was recently adapted to starlings (Van Meir et al., 2005) and zebra finches (Boumans et al., 2007, 2008a,b; Voss et al., 2007), allows a whole-brain approach and can be used to perform assumption-free experiments. Taking advantage of this technique, we measured the blood oxygen level-dependent (BOLD) neural responses in anesthetized male zebra finches that were exposed to BOS, a conspecific song (CON) and a heterospecific song (HET) to determine where BOS selectivity occurs in the ascending auditory pathway. Comparison of neural activity triggered by BOS versus CON and BOS versus HET revealed that BOS selectivity measured at the neural level (i.e., at the level of large pool of neurons) is present in the right auditory midbrain.

Materials and Methods

Subjects. Adult male zebra finches (*Taenopygia guttata*) purchased from local suppliers were used in this experiment. Birds were housed together in a big aviary, had access to food and water ad libitum and were maintained under a 12 h light/dark photoperiod throughout the experiment. Experimental procedures were in agreement with the Belgian laws on the protection and welfare of animals and were approved by the ethical committee of the University of Antwerp.

Stimuli and stimulation device. Undirected song of each bird was recorded in a soundproof chamber with a Sony ECM MS 907 microphone (Sony) and a Phonic MU 502 amplifier (Phonic). BOS/CON stimuli were constituted of several individual songs of each bird interleaved with silence periods of 0.5 s (see supplemental Fig. S1, available at www.jneurosci.org as supplemental material). This procedure conserved the natural intraindividual and interindividual variability of song length. The total length of each stimulus was 16 s. The intensity of each song was normalized (in term of matched root-mean-square) before being integrated into the complete stimulus (songs and silence periods). These manipulations were done using Praat software (University of Amster-

Received Sept. 29, 2008; revised Dec. 17, 2008; accepted Jan. 8, 2009.

This work was supported by a Francqui grant and a Bijzonder Onderzoeks Fonds—Klein Project grant from the University of Antwerp to C.P., a grant from the National Institute of Neurological Disorders and Stroke (NS035467) to J.B., grants from the Research Foundation—Flanders (FWO, project number G.0420.02), Concerted Research Actions (Geconcerte Onderzoeks Actie funding) from the University of Antwerp, by European Commission—Sixth Framework Program (EC—FP6) project DiMI, LSHB-CT-2005-512146, and EC—FP6 project EMIL LSHC-CT-2004-503569 to A.V.d.L. T.B. is doctoral fellow of the FWO. We thank G. F. Ball, M. Gahr, O. Güntürkün, P. Joris, and V. Van Meir for their comments on an earlier version of this manuscript.

Correspondence should be addressed to Colline Poirier, Bio-Imaging Laboratory, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium. E-mail: colline.poirier@ua.ac.be.

DOI:10.1523/JNEUROSCI.4650-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/292252-07\$15.00/0

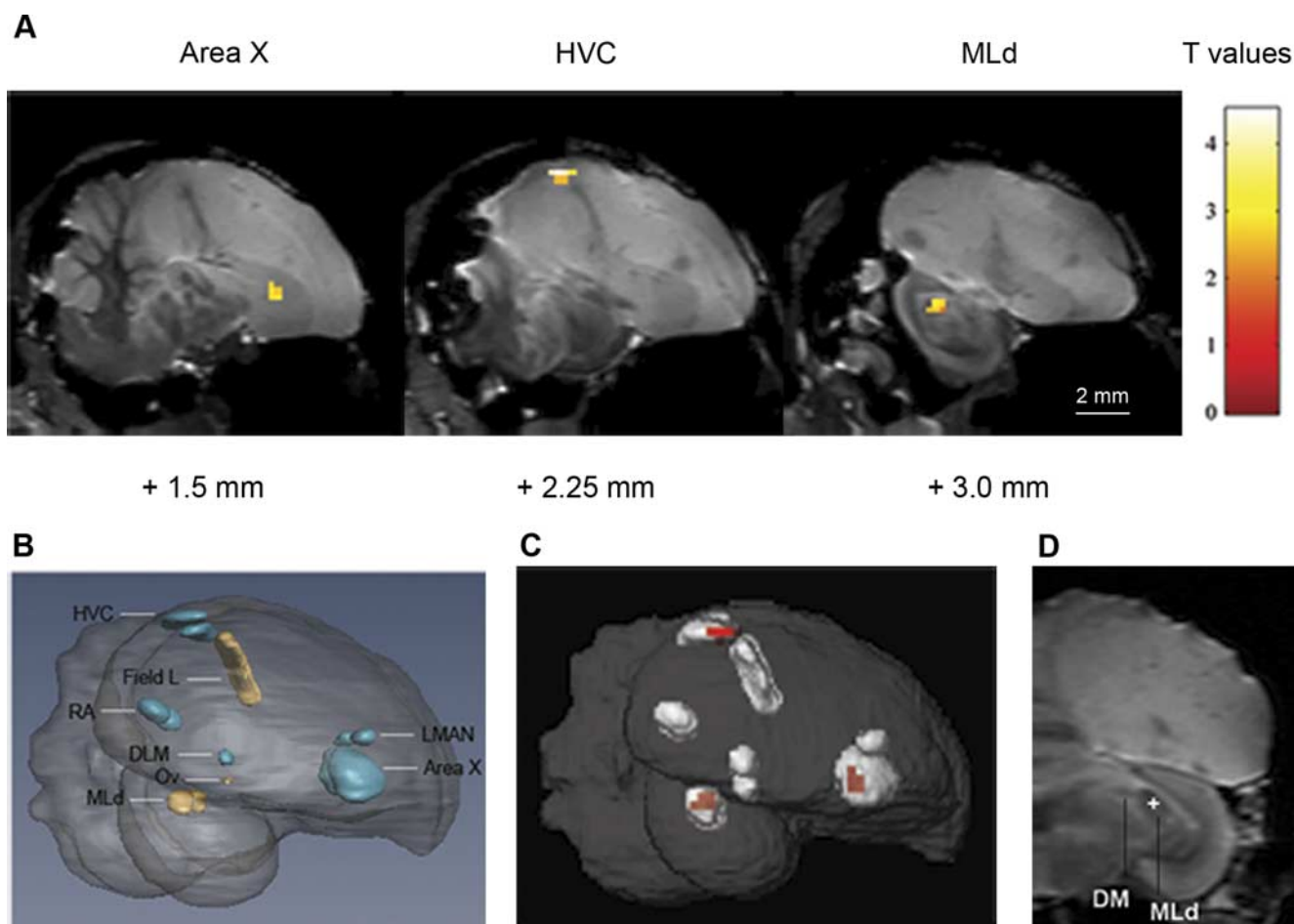


Figure 1. Neural substrates of own-song recognition (BOS vs CON) in the song control and auditory regions. **A**, Superimposition of the statistical results to anatomical sagittal images coming from a recently developed MRI zebra finch atlas (Poirier et al., 2008). Coordinates expressed in millimeters from the midline are indicated under each map. The + sign indicates that slices (and statistical results) are from the right hemisphere. *t* values are color coded according to the scale displayed on the right side of the panel. **B**, Rendered images of the whole zebra finch brain (gray), song control nuclei (blue), and auditory nuclei (yellow) (right sagittal view). For definitions of abbreviations, see text. **C**, Superimposition of the statistical results to rendered images of the same nuclei as in **B** (right sagittal view). **D**, Position of the right MLd activation (indicated by the + sign) relative to the adjacent motor dorsomedial nucleus of the mesencephalon (DM) in a coronal view, illustrating the clear separation between these two nuclei and specificity of the activation in the auditory MLd.

dam, The Netherlands). Heterospecific stimuli corresponded to songs recorded from two European starlings (*Sturnus vulgaris*) and two canaries (*Serinus canaria*), resulting in four different stimuli. These songs, in addition to other starling and canary songs, had been played to the experimental subjects on average 10 times per day during at least 4 weeks in an attempt to decrease their novelty. Songs were delivered by a D-NE240 CD player (Sony) and VS 2220 amplified speaker system (Altec). The duration of each heterospecific stimulus was 16 s, and its intensity was matched to the intensity of zebra finch songs (in term of root-mean-square).

During imaging experiments, auditory signals were presented to the birds with magnetless dynamic speakers connected to an amplifier as described by Van Meir et al. (2005). Stimulus application was controlled by Presentation software 0.76 (Neurobehavioral Systems). To test whether some sound frequencies were not absorbed or enhanced into the magnet, stimuli were first emitted with the magnetless dynamic speakers in the presence of a test bird and recorded with a nonmagnetic Electret microphone. Frequencies between 2500–5000 Hz were found to be enhanced. To compensate this artificial enhancement, an equalizer function was applied to each stimulus using WaveLab software. The function consisted of a Gaussian kernel with the following parameters: maximum amplitude: -20 dB, centered on 3750 Hz, width: 0.05 octaves (corresponding to the range 2500–5000 Hz). During the experiments, stimuli were delivered monophonically, with a sampling frequency of 22050 Hz and their global intensity was ~ 67 dB sound pressure level. By comparison, the magnet noise was ~ 65 dB.

Experimental design. The experiment consisted of an ON/OFF block design alternating auditory stimulation periods (ON blocks) with resting periods (OFF blocks). Each block (ON and OFF) lasted 16 s, which corresponds to the acquisition time of 2 images (see supplemental Fig. S2, available at www.jneurosci.org as supplemental material). Each stimulus type was presented 25 times, resulting in the acquisition of 50 images per stimulus and per subject. The presentation order of the conditions was randomized within and between subjects. Since the BOS is unique for each bird, each subject was tested with one BOS, one CON and one HET stimulus.

Anesthesia and physiology monitoring. During the experiment, birds were continuously anesthetized with isoflurane (anesthesia induction: 3%, maintenance: 2%). The anesthetic gas was inhaled with a mixture of oxygen and nitrogen (at a flow rate of 100 and 200 cc/min, respectively) through a small mask over the bird's beak. Isoflurane is the most common anesthetic in clinical applications and it has also been used in most of the imaging studies with monkeys and cats (Logothetis et al., 1999; Kim et al., 2000; Duong et al., 2001). Isoflurane has the great advantage of having relatively minor side effects, and, indeed, 100% of our tested birds recovered within a few minutes after the end of the fMRI experiment. Body temperature was continuously monitored with a cloacal temperature probe and maintained at 40°C by a feedback controlled heating system (SA-Instruments). Respiration rate and amplitude were constantly monitored with a small pneumatic sensor (SA-Instruments) positioned under the bird.

Image acquisition. Specifications of our 7T scanner and the coils used

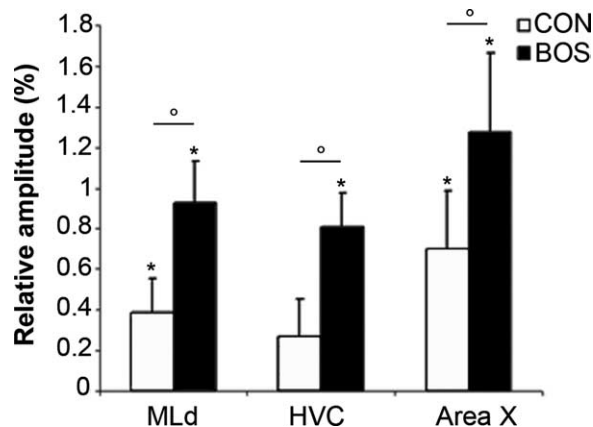


Figure 2. Relative amplitudes (in percentage) of neural activations elicited by CON (open bars) and BOS (closed bars) stimuli. Amplitudes have been averaged across the voxels constituting each cluster illustrated in Figure 1A. The zero level corresponds to the mean activation level during rest periods (exposure to scanner noise only). The error bars correspond to SEs across subjects (and not across voxels of each cluster). Asterisks indicate statistical significance of CON versus Rest and BOS versus Rest comparisons, whereas circles indicate statistical significance of BOS versus CON comparisons (corrected $p < 0.05$).

for the experiment can be found in Boumans et al. (2007). Functional MRI data were acquired using a T2-weighted spin-echo sequence [echo time (TE)/repetition time (TR): 60/2000 ms]. Fifteen continuous sagittal slices of 0.75 mm thickness covering nearly the entire brain were acquired. Voxel size was $0.25 \times 0.25 \times 0.75 \text{ mm}^3$. Anatomical 3D images required for localization of the functional data (see below) were obtained for each bird using a RARE T2-weighted sequence (TE/TR: 60/2000 ms, voxel size: $0.0625 \times 0.0625 \times 0.0625 \text{ mm}$).

Image processing. Intraindividual head motion was corrected using a six-parameter rigid body spatial transformation. In parallel, each individual anatomical three-dimensional (3D) dataset was coregistered with the zebra finch brain atlas that we developed in our lab (Poirier et al., 2008), using mutual information implemented in home-written Matlab code. The transformation matrix of this coregistration was then applied to the realigned functional data, resulting in functional data precisely coregistered to the atlas dataset. Finally, functional data were smoothed with a 0.5 mm width Gaussian kernel.

Statistical analysis. Statistical voxel-based analyses were performed using a mass-univariate approach based on General Linear Models, implemented in SPM5 (<http://www.fil.ion.ucl.ac.uk/spm/>). Because the shape of the hemodynamic response is not known in the zebra finch and because the temporal resolution of the spin-echo fMRI used in the present experiment was weak (two time points per ON block), the data were simply modeled as a box-car and filtered with a high-pass filter of 360 s. Model parameters were then estimated using a classical restricted maximum likelihood algorithm.

The statistical analysis was restricted to some a priori defined regions of interest (ROI), the song control nuclei: HVC, used as a proper name (Reiner et al., 2004), the nucleus robustus of the arcopallium (RA), area X and the LMAN (lateral magnocellular nucleus of the anterior nidopallium); and the auditory regions: the dorsal part of the lateral mesencephalic nucleus (MLd), Field L, the caudal mesopallium (CM), and NCM (caudomedial nidopallium). Because they were too small to be sampled by at least one sagittal slice, DLM (dorsolateral nucleus of the medial thalamus), Ov (nucleus ovoidalis), and the nucleus interface of the nidopallium (Nif) were not investigated. We first computed the mean effect of each stimulus (compared with the rest period) in each voxel of the ROI, for each subject (fixed-effect analysis). In a second step, a group analysis was computed on the effects identified by the previous analysis with a repeated measure ANOVA (one factor with three levels representing the three stimuli). The main “stimulus” effect was computed, then *post hoc* one-tailed paired t tests were performed only in the voxels found to be significant in the main effect. This second step corresponds to a mixed-effect analysis and allows establishing the invariant behavior of

Table 1. Statistics of BOS versus CON comparison ($N = 15$)

	Coordinates (in mm)			t value	Corrected p value	Cluster Size
	x	y	z			
Right MLd	3.00	−0.50	−0.50	2.9	0.009	5
Right HVC	2.25	0.00	4.25	4.5	< 0.001	6
Right area X	1.50	3.75	0.00	2.9	0.011	5

Coordinates, t , and p values correspond to the voxel with the maximum t value. Cluster size corresponds to the number of voxels found to be significantly more activated by BOS than by CON (corrected, $p < 0.05$). Coordinates are expressed according to the reference space defined in Poirier et al. (2008). The x , coordinate corresponds to the left/right axis (with positive values for the right side), the y , coordinate corresponds to the anterior/posterior axis (with positive values for the anterior side), and the z , coordinate corresponds to the dorsal/ventral axis (with positive values for dorsal side). The zero point corresponds to the intersection between the midsagittal plane and the posterior commissure.

the population from which the group was sampled. Correction for multiple comparisons was then applied on these voxels using the Family Wise Error method based on random field theory. The combination of voxel-based statistics and mixed-effect analysis, implemented for the first time on nonhuman fMRI data, represents a major improvement on small animal fMRI data analysis.

For lateralization analyses, we first calculated for each subject the differential effect between stimuli (BOS vs CON and CON vs HET) in the two MLd clusters found to be significantly activated in the group analysis. Then, these differential effects were compared between each hemisphere using two-tailed paired t tests. Differences were considered as statistically significant when $p < 0.05$.

Results

Neural substrates of bird’s own song recognition

We first compared the BOLD response elicited by BOS versus CON stimuli. BOS stimuli were undirected songs produced by the tested bird whereas CON stimuli were undirected songs produced by a cage mate. BOS perception induced a greater activity in parts of two telencephalic song control regions, the right HVC and the right area X, and in one auditory midbrain region, the right MLd (Figs. 1, 2) ($t \geq 2.85$, $p \leq 0.02$, $N = 15$) (for detailed statistics and coordinates of activations, see Table 1). The anatomical resolution of these fMRI data did not allow determining whether the HVC cluster was located in HVC itself and/or in the shelf region around HVC. For simplicity, this cluster will be called HVC in the text that follows.

To exclude the possibility that BOS selectivity observed in the right HVC, right X and right MLd was due to auditory characteristics specific to the BOS stimuli, we repeated the comparison BOS versus CON in a subgroup of 10 out of the 15 birds where BOS and CON stimuli were physically identical (each song was used as BOS in one bird and as CON in another bird). This analysis confirmed the greater recruitment of the right HVC, right X and right MLd during BOS perception compared with CON ($t \geq 2.52$, $p \leq 0.025$, $N = 10$) (Fig. 3).

To confirm the specific involvement of these three regions in own-song recognition, we also compared BOLD responses elicited by BOS versus HET. The heterospecific stimulus, a familiar starling or canary song, was used as a natural complex auditory control stimulus. The difference between these stimuli was found significant in the right HVC, right X and right MLd ($t \geq 2.25$, $p \leq 0.037$, $N = 15$; data not shown).

Neuronal activity can decrease in intensity with stimulus repetition, a phenomenon called habituation in birdsong literature (Chew et al., 1995). This phenomenon could potentially occur at the neural (i.e., neuronal population) level and differential activations, as described so far, could merely result from differential habituation rates between stimuli. To test for the potential effect of this confounding factor, we used two complementary methods. First, we fitted the data with a linear regression and tested

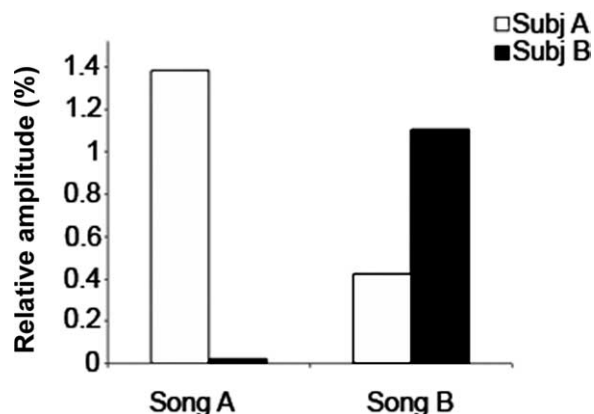


Figure 3. Relative amplitudes (in percentage) of the MLD activation in two representative subjects. The zero level corresponds to the mean activation level during rest periods. This figure illustrates that the intensity of the neural activity induced by one given song depends on the receiver of the signal (i.e., the experimental bird) and not on the song itself. The activation elicited by song A (left side of the graph) is greater when the receiver is the owner of the song (BOS, Subj A, open bar) and weaker when the receiver is a cage mate of A (CON, Subj B, closed bar). A similar relationship applies to song B (right side of the graph).

whether the slope of this regression significantly differed between the three stimuli. In the three regions, the slopes were not significantly different (right MLd: $F = 0.3$, $p = 0.62$; right HVC: $F = 1.2$, $p = 0.44$; right Area X: $F = 2.9$, $p = 0.14$, all $N = 15$). Because this approach requires defining an a priori linear shape of the trend, we also compared the activations between the first and the second half of the experiment by a two-way repeated measures ANOVA (within factors: stimuli and period of imaging). This second approach, although less powerful, has the advantage of being assumption-free regarding the shape of the habituation curve. Interaction between the two halves of the experiment and the three stimuli was not significant in the three regions ($F \leq 1.25$, $p \geq 0.39$, $N = 15$). The differential response to BOS versus CON and BOS versus HET thus cannot be solely explained by a differential habituation rate to the stimuli.

Neural substrates of conspecific song recognition

To test whether BOS and conspecific song recognition are supported by the same or different neural substrates, we identified the neural substrates of conspecific song recognition by comparing the neural activation elicited by CON and HET. This comparison revealed the specific involvement in CON perception of the left MLd, left Field L and of a region at the limit of left area X (Figs. 4, 5) ($t \geq 2.6$, $p \leq 0.003$, $N = 15$) (for detailed statistics and coordinates of activations, see Table 2).

To confirm the specific involvement of the left MLd, left Field L and left area X in conspecific song recognition, we also tested the comparison BOS versus HET, where BOS was considered as a particular conspecific song. The difference between these stimuli was also found significant in these three regions ($t \geq 2.0$, $p \leq 0.035$, $N = 15$).

Potential differences in habituation rates between stimuli were controlled with the two methods described previously. The habituation rates to these stimuli were not significantly different in left MLd ($F \leq 2.1$, $p \geq 0.27$, $N = 15$), left Field L ($F \leq 0.1$, $p \geq 0.61$, $N = 15$) and left Area X ($F \leq 3.14$, $p \geq 0.14$, $N = 15$).

Lateralization of selectivity in MLd

In the analyses presented above, BOS selectivity was significant in the right MLd only but was not present on the left side (BOS vs

Table 2. Statistics of CON versus HET comparison ($N = 15$)

	Coordinates (in mm)			<i>t</i> value	Corrected <i>p</i> value	Cluster size
	<i>x</i>	<i>y</i>	<i>z</i>			
Left MLd	−3.00	−0.50	−1.00	3.1	0.009	7
Left field L	−1.50	1.00	3.00	2.6	0.009	2
Left area X	−1.50	4.00	0.75	3.5	0.003	6

For details, see Table 1.

CON on the left side, $p = 0.51$). In contrast, conspecific song selectivity was significant in left MLd only (CON vs HET on the right side, $p = 0.31$). To further test for the potential lateralization of these differential activations, we compared them directly in the left and right MLd (as defined functionally by previous comparisons) (Fig. 6). The differential activations elicited by BOS versus CON and by CON versus HET were both significantly lateralized (BOS minus CON, left vs right: $t = -2.7$, $p = 0.02$, $N = 15$; CON minus HET, left vs right: $t = 2.4$, $p = 0.03$, $N = 15$).

Discussion

Consistently with the electrophysiological (neuronal) literature (Margoliash and Fortune, 1992; Janata and Margoliash, 1999), the present fMRI study revealed BOS selectivity in HVC and area X at the neuron population (neural) level. Additionally, a similar selectivity was discovered in the ascending auditory pathway at the MLd level. Because BOS and CON stimuli were physically identical in a subgroup of subjects, BOS selectivity could in no way be biased by acoustic peculiarities in the signals. The absence of significant differential neural activation in the other auditory and song control regions that were investigated cannot be interpreted as an absence of neuronal BOS selectivity. To be detectable by fMRI, neuronal BOS selectivity has to be expressed by a substantial number of neurons concentrated at one location. BOLD fMRI signal is best correlated to local field potentials (Logothetis et al., 2001) and does not allow distinguishing between excitatory and inhibitory neurons. In a recent electrophysiological study investigating neuronal BOS selectivity in the secondary auditory region CM, Bauer et al. (2008) found that some neurons were inhibitory and not selective to BOS whereas other neurons were excitatory and that among these excitatory neurons, some of them were BOS selective. This neuronal heterogeneity could explain why we did not observe any significant selectivity in CM at the neural level with fMRI.

Auditory feedback in song learning and production

Juvenile songbirds learn their song by listening to an adult conspecific male, the tutor (sensory phase), and by progressively matching their own vocalizations to this memorized model (sensory-motor phase). The first phase depends on auditory perception of the tutor song while the second one depends on auditory perception of the own song via auditory feedback. Both types of auditory experiences are necessary for normal song development. Auditory feedback is also important for adult song maintenance in zebra finches. Adult deafening (Nordeen and Nordeen, 1992) or artificially distorted auditory feedback imposed during singing (Leonardo and Konishi, 1999) induce alterations of the song structure.

For now more than 25 years, playback experiments have revealed that some neurons in telencephalic nuclei of the song control system fire selectively in response to playbacks of the bird's own song. These BOS-selective responses are thought to support the auditory feedback and/or the integration of this feedback with the motor circuitry for vocal learning and maintenance (Prather

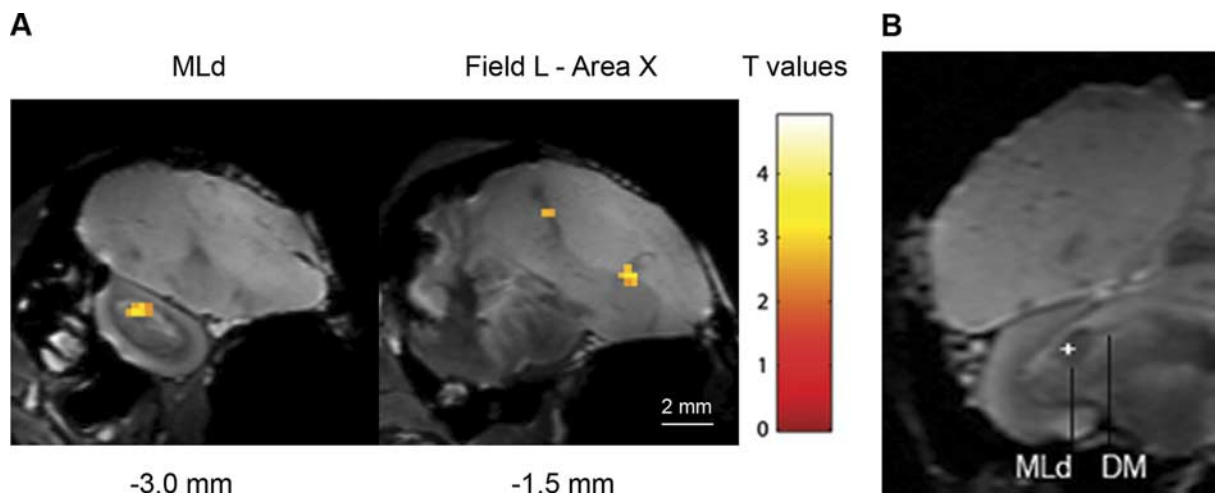


Figure 4. Neural substrates of conspecific song recognition (CON vs HET) in the song control and auditory regions. **A**, Superimposition of the statistical results to anatomical sagittal images coming from a recently developed MRI zebra finch atlas (Poirier et al., 2008). Coordinates expressed in millimeters from the midline are indicated under each map. The — sign indicates that slices (and statistical results) are from the left hemisphere. *t* values are color coded according to the scale displayed on the right side of the panel. **B**, Position in a coronal view of the left MLd activation (indicated by the + sign) in relation to the adjacent motor DM.

and Mooney, 2004; Theunissen et al., 2004). Perception of BOS playback through loudspeakers is to some extent different from the natural auditory feedback that the bird experiences when it is spontaneously singing (e.g., pitch distortion due to bone conduction, ...). However, numerous experiments have demonstrated the specificity of BOS-selective responses. These auditory responses have been shown to be sensitive to specific temporal and spectral features found in the BOS. It is therefore quite unlikely if not impossible that the differential response observed here in birds exposed to BOS compared with CON was induced by these limited differences between sounds associated with auditory feedback and with BOS playback through loudspeakers.

Song learning and maintenance is thought to depend on matching motor efference copies with a reafferent auditory feedback. Whether MLd activation as measured in the present study represents a signal from the auditory input, the motor efference copies, the reafferent auditory feedback, or an integration of these signals, remains difficult to determine.

MLd belongs to the ascending auditory pathway and sends ascending auditory information to nucleus ovoidalis, which is suspected to project to Nif (Wild, 2004), a BOS-selective song control nucleus (Janata and Margoliash, 1999) representing a major auditory input to HVC (Fortune and Margoliash, 1995). MLd activation could thus reflect only these ascending inputs. However the area surrounding the premotor nucleus RA, called the RA cup projects to MLd and these projections are thought to convey auditory information (Mello et al., 1998). MLd activation might thus as well reflect activity in these descending projections.

Finally, MLd activation might also reflect a potential motor efference copy signal. Indeed the relatively weak responses elicited by BOS in awake zebra finches were found, quite surprisingly, to be enhanced in anesthetized or sleeping birds (Dave et al., 1998; Cardin and Schmidt, 2004). Additionally, it has been shown that spontaneous ongoing premotor activity can still be recorded when the birds are sleeping, a kind of song replay (Dave and Margoliash, 2000). Since BOS selectivity in MLd was detected in the present study in anesthetized birds, this selectivity possibly reflects, at least in part, mechanisms occurring when the bird is sleeping, i.e., the premotor activity that was previously identified. Activation of auditory areas has indeed already been

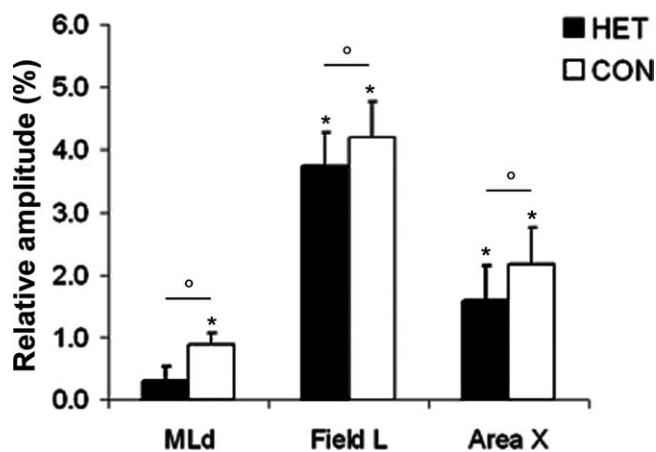


Figure 5. Relative amplitudes (in percentage) of neural activations elicited by HET (closed bars) and CON (open bars) stimuli. Amplitudes have been averaged across the voxels constituting each cluster illustrated in Figure 4A. The zero level corresponds to the mean activation level during rest periods (exposure to scanner noise only). The error bars correspond to SEs across subjects. Asterisks indicate statistical significance of HET versus Rest and CON versus Rest comparisons, whereas circles indicate statistical significance of CON versus HET comparisons (corrected $p < 0.05$).

found to be triggered by somato-sensory (human and nonhuman primates: Schroeder et al., 2001; Dhanjal et al., 2008) or premotor (zebra finches: Keller and Hahnloser, 2009) signals. However, to be responsible of BOS selectivity, the premotor activity would need to occur only during BOS playback periods. If we cannot exclude this hypothesis, there is, to our knowledge, no experimental data supporting it.

Lateralization in song and speech control

In humans, speech learning and maintenance are supposed to be supported by a feedforward and a feedback control. Under feedback control, speech production is monitored during execution and deviations from the expected signal are corrected according to auditory information. Under feedforward control, speech is produced based on previously learned commands, without reliance on incoming sensory information. Speech learning and maintenance are thought to result from the tuning of the feed-

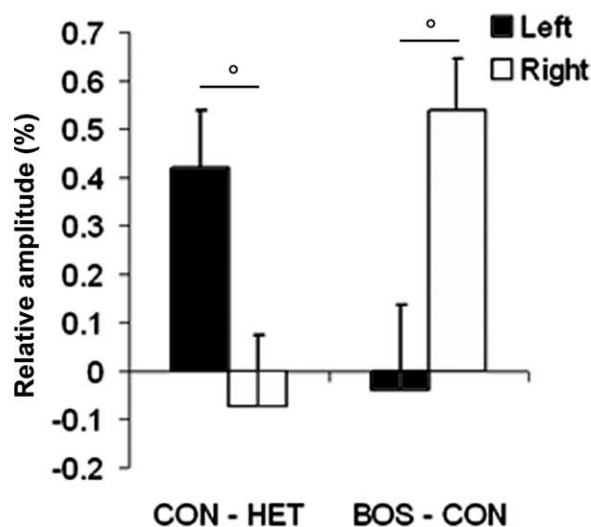


Figure 6. Relative amplitudes (in percentage) of conspecific and own-song recognition in left (closed bars) and right (open bars) MLD. Positive values indicate that the region was more activated by CON than by HET (or by BOS than by CON). Negative values indicate that the region was more activated by HET than by CON (or by CON than by BOS). The error bars correspond to SEs across subjects. Circles indicate that the difference between left and right MLD clusters is significant (corrected $p < 0.05$).

forward motor commands over time by the auditory feedback (Guenther, 2006). Recent studies suggest that the feedforward control is lateralized to the left hemisphere and the auditory feedback control is, at least partially, lateralized to the right (Toyomura et al., 2007; Tourville et al., 2008). Making the link between these potential hemispheric dominances and speech lateralization, Tourville et al. (2008) suggest that while the left lateralization of speech production deduced from lesions studies would reflect the left lateralization of the feedforward control system, the right lateralization of the auditory feedback control would explain the importance of the right hemisphere observed in numerous aspects of speech production. This auditory feedback could participate to self-recognition processes (Fu et al., 2006) that seem to be lateralized to the right hemisphere in the auditory (Nakamura et al., 2001; Rosa et al., 2008) as in the visual modality (Keenan et al., 2001; Platek et al., 2006; Uddin et al., 2007). Song lateralization in birds varies according to the species (Schmidt et al., 2004). In zebra finches, lesion studies suggest a moderate right lateralization of song production (Williams et al., 1992; Floody and Arnold, 1997). The right lateralization of BOS selectivity observed in the present study could thus reflect this potential motor lateralization. However, the rapid effect of lesions on song production compared with the slow effects induced by auditory feedback disturbance indicate two distinct phenomena. The more pronounced effects of peripheral and central lesions of the right motor system could thus rather reflect the right lateralization of the feedforward control system, whereas the right lateralization of BOS selectivity could be due to a right lateralization of the auditory feedback control system, suggesting an anatomofunctional convergence between birds and humans. If this second hypothesis is true, BOS selectivity should be found lateralized to the right side even in bird species where song production is lateralized to the left.

Relationship between BOS and CON recognition

In addition to its role in song learning and song maintenance, BOS selectivity has been suggested to underlie conspecific song

recognition and discrimination via motor encoding (Williams and Nottebohm, 1985). A conspecific song could also be discriminated from a heterospecific song based on its closer perceptual proximity to the BOS. Our results suggest, however, the existence of distinct neural substrates for BOS and conspecific song recognition but with a possible facilitating influence of the former on the latter. BOS recognition was found to involve the right MLD and species recognition to involve the left MLD. Whereas both hemispheres are anatomically independent at the telencephalic level (due to an absence of a corpus callosum in birds), midbrain and thalamic regions possess interhemispheric connections. Such lateral connections could thus drive a modulatory input from right to left MLD, mediating a facilitating effect of BOS recognition on conspecific song recognition.

Role of subcortical structures in learning processes

Subcortical structures are usually considered to be primitive and are known to be responsible for simple decoding of physical parameters of the stimuli, but not for signal integration. Experience-dependent plasticity is usually considered limited to cortical structures. For instance, in songbirds, innate call production is supported by midbrain structures whereas learned call and song production are thought to rely on forebrain structures (Margoliash, 1997). This probably explains why previous investigations about BOS selectivity in the ascending auditory system focused on the telencephalic auditory regions (Lewicki and Arthur, 1996; Janata and Margoliash, 1999; Amin et al., 2004; Shaevitz and Theunissen, 2007; Bauer et al., 2008). However, more and more experimental results challenge this view. In barn owls, auditory experience has been shown to modify tuning properties of inferior colliculus neurons (Gold and Knudsen, 2000). In humans, plasticity has been observed in the auditory brainstem after linguistic training (Song et al., 2008). While the capacity to process complex information about self and others was thought to be computed by the recently evolved frontoparietal regions (Gallup, 1997; Uddin et al., 2007), the midbrain has been shown to be involved in own-face recognition (Sugiura et al., 2005). Previous electrophysiological studies in zebra finches revealed that MLD neurons respond to a wide range of complex sounds (Woolley and Casseday, 2005) and are tuned to auditory features that enhance the acoustic differences between conspecific songs (Woolley et al., 2005), suggesting already that these neurons are more complex than expected. Our results indicate that a subcortical structure can process subtle information about the identity of a subject through experience-dependent mechanisms.

References

- Amin N, Grace JA, Theunissen FE (2004) Neural response to bird's own song and tutor song in the zebra finch field L and caudal mesopallium. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 190:469–489.
- Bauer EE, Coleman MJ, Roberts TF, Roy A, Prather JF, Mooney R (2008) A synaptic basis for auditory–vocal integration in the songbird. *J Neurosci* 28:1509–1522.
- Boumans T, Theunissen FE, Poirier C, Van Der Linden A (2007) Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional MRI. *Eur J Neurosci* 26:2613–2626.
- Boumans T, Gobes SM, Poirier C, Theunissen FE, Vandersmissen L, Pintjens W, Verhoye M, Bolhuis JJ, Van der Linden A (2008a) A functional MRI of auditory responses in the zebra finch forebrain reveals a hierarchical organization based on signal strength but not selectivity. *PLoS ONE* 3:e3184.
- Boumans T, Vignal C, Smolders A, Sijbers J, Verhoye M, Van Audekerke J, Mathevon N, Van der Linden A (2008b) Functional magnetic resonance

- imaging in zebra finch discerns the neural substrate involved in segregation of conspecific song from background noise. *J Neurophysiol* 99:931–938.
- Cardin JA, Schmidt MF (2004) Auditory responses in multiple sensorimotor system nuclei are co-modulated by behavioral state. *J Neurophysiol* 91:2148–2163.
- Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. *Proc Natl Acad Sci U S A* 92:3406–3410.
- Cynx J (1993) Conspecific song perception in zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 107:395–402.
- Cynx J, Nottebohm F (1992) Role of gender, season, and familiarity in discrimination of conspecific song by zebra finches (*Taeniopygia guttata*). *Proc Natl Acad Sci U S A* 89:1368–1371.
- Dave AS, Margoliash D (2000) Song replay during sleep and computational rules for sensorimotor vocal learning. *Science* 290:812–816.
- Dave AS, Yu AC, Margoliash D (1998) Behavioral state modulation of auditory activity in a vocal motor system. *Science* 282:2250–2254.
- Dhanjal NS, Handunnetthi L, Patel MC, Wise RJ (2008) Perceptual systems controlling speech production. *J Neurosci* 28:9969–9975.
- Doupe AJ, Konishi M (1991) Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad Sci U S A* 88:11339–11343.
- Duong TQ, Kim DS, Uđurbil K, Kim SG (2001) Localized cerebral blood flow response at submillimeter columnar resolution. *Proc Natl Acad Sci U S A* 98:10904–10909.
- Floody OR, Arnold AP (1997) Song lateralization in the zebra finch. *Horm Behav* 31:25–34.
- Fortune ES, Margoliash D (1995) Parallel pathways and convergence onto HVC and adjacent neostriatum of adult zebra finches (*Taeniopygia guttata*). *J Comp Neurol* 360:413–441.
- Fu CH, Vythelingum GN, Brammer MJ, Williams SC, Amaro E Jr, Andrew CM, Yáguez L, van Haren NE, Matsumoto K, McGuire PK (2006) An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. *Cereb Cortex* 16:969–977.
- Gallup GG Jr (1997) On the rise and fall of self-conception in primates. *Ann N Y Acad Sci* 818:72–82.
- Gold JI, Knudsen EI (2000) A site of auditory experience-dependent plasticity in the neural representation of auditory space in the barn owl's inferior colliculus. *J Neurosci* 20:3469–3486.
- Guenther FH (2006) Cortical interactions underlying the production of speech sounds. *J Commun Disord* 39:350–365.
- Janata P, Margoliash D (1999) Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *J Neurosci* 19:5108–5118.
- Keenan JP, Nelson A, O'Connor M, Pascual-Leone A (2001) Self-recognition and the right hemisphere. *Nature* 409:305.
- Keller GB, Hahnloser RH (2009) Neural processing of auditory feedback during vocal practice in a songbird. *Nature* 457:187–190.
- Kim DS, Duong TQ, Kim SG (2000) High-resolution mapping of iso-orientation columns by fMRI. *Nat Neurosci* 3:164–169.
- Leonardo A, Konishi M (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* 399:466–470.
- Lewicki MS, Arthur BJ (1996) Hierarchical organization of auditory temporal context sensitivity. *J Neurosci* 16:6987–6998.
- Logothetis NK, Guggenberger H, Peled S, Pauls J (1999) Functional imaging of the monkey brain. *Nat Neurosci* 2:555–562.
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157.
- Margoliash D (1997) Functional organization of forebrain pathways for song production and perception. *J Neurobiol* 33:671–693.
- Margoliash D, Fortune ES (1992) Temporal and harmonic combination-sensitive neurons in the zebra finch's HVC. *J Neurosci* 12:4309–4326.
- Mello CV, Vates GE, Okuhata S, Nottebohm F (1998) Descending auditory pathways in the adult male zebra finch (*Taeniopygia guttata*). *J Comp Neurol* 395:137–160.
- Nakamura K, Kawashima R, Sugiura M, Kato T, Nakamura A, Hatano K, Nagumo S, Kubota K, Fukuda H, Ito K, Kojima S (2001) Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia* 39:1047–1054.
- Nordeen KW, Nordeen EJ (1992) Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav Neural Biol* 57:58–66.
- Platek SM, Loughead JW, Gur RC, Busch S, Ruparel K, Phend N, Panyavin IS, Langleben DD (2006) Neural substrates for functionally discriminating self-face from personally familiar faces. *Hum Brain Mapp* 27:91–98.
- Poirier C, Vellema M, Verhoye M, Van Meir V, Wild JM, Balthazart J, Van Der Linden A (2008) A three-dimensional MRI atlas of the zebra finch brain in stereotaxic coordinates. *Neuroimage* 41:1–6.
- Prather JF, Mooney R (2004) Neural correlates of learned song in the avian forebrain: simultaneous representation of self and others. *Curr Opin Neurobiol* 14:496–502.
- Reiner A, Perkel DJ, Bruce LL, Butler AB, Csillag A, Kuenzel W, Medina L, Paxinos G, Shimizu T, Striedter G, Wild M, Ball GF, Durand S, Gütürkün O, Lee DW, Mello CV, Powers A, White SA, Hough G, Kubikova L, et al. (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. *J Comp Neurol* 473:377–414.
- Rosa C, Lassonde M, Pinard C, Keenan JP, Belin P (2008) Investigations of hemispheric specialization of self-voice recognition. *Brain Cogn* 68:204–214.
- Schmidt MF, Ashmore RC, Vu ET (2004) Bilateral control and interhemispheric coordination in the avian song motor system. *Ann N Y Acad Sci* 1016:171–186.
- Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC (2001) Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 85:1322–1327.
- Shaevitz SS, Theunissen FE (2007) Functional connectivity between auditory areas field L and CLM and song system nucleus HVC in anesthetized zebra finches. *J Neurophysiol* 98:2747–2764.
- Song JH, Skoe E, Wong PC, Kraus N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci* 20:1892–1902.
- Sugiura M, Watanabe J, Maeda Y, Matsue Y, Fukuda H, Kawashima R (2005) Cortical mechanisms of visual self-recognition. *Neuroimage* 24:143–149.
- Theunissen FE, Amin N, Shaevitz SS, Woolley SM, Fremouw T, Hauber ME (2004) Song selectivity in the song system and in the auditory forebrain. *Ann N Y Acad Sci* 1016:222–245.
- Tourville JA, Reilly KJ, Guenther FH (2008) Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39:1429–1443.
- Toyomura A, Koyama S, Miyamaoto T, Terao A, Omori T, Murohashi H, Kuriki S (2007) Neural correlates of auditory feedback control in human. *Neuroscience* 146:499–503.
- Uddin LQ, Iacoboni M, Lange C, Keenan JP (2007) The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn Sci* 11:153–157.
- 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) Spatiotemporal properties of the BOLD response in the songbirds' auditory circuit during a variety of listening tasks. *Neuroimage* 25:1242–1255.
- Voss HU, Tabelow K, Polzehl J, Tchernichovski O, Maul KK, Salgado-Commissariat D, Ballon D, Helekar SA (2007) Functional MRI of the zebra finch brain during song stimulation suggests a lateralized response topography. *Proc Natl Acad Sci U S A* 104:10667–10672.
- Wild JM (2004) Functional neuroanatomy of the sensorimotor control of singing. *Ann N Y Acad Sci* 1016:438–462.
- Williams H, Nottebohm F (1985) Auditory responses in avian vocal motor neurons: a motor theory for song perception in birds. *Science* 229:279–282.
- Williams H, Crane LA, Hale TK, Esposito MA, Nottebohm F (1992) Right-side dominance for song control in the zebra finch. *J Neurobiol* 23:1006–1020.
- Woolley SM, Casseday JH (2005) Processing of modulated sounds in the zebra finch auditory midbrain: responses to noise, frequency sweeps, and sinusoidal amplitude modulations. *J Neurophysiol* 94:1143–1157.
- Woolley SM, Fremouw TE, Hsu A, Theunissen FE (2005) Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat Neurosci* 8:1371–1379.