ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 25 (2005) 1242–1255

Spatiotemporal properties of the BOLD response in the songbirds' auditory circuit during a variety of listening tasks

Vincent Van Meir,[a,*] Tiny Boumans,[a] Geert De Groof,[a] Johan Van Audekerke,[a] Alain Smolders,[b] Paul Scheunders,[b] Jan Sijbers,[b] Marleen Verhoye,[a] Jacques Balthazart,[c] and Annemie Van der Linden[a]

[a]*Bio-Imaging Laboratory, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*
[b]*Vision Laboratory, University of Antwerp, Belgium*
[c]*Research Group in Behavioral Neuroendocrinology, Center for Cellular and Molecular Neurobiology, University of Liège, Belgium*

Received 27 September 2004; revised 19 November 2004; accepted 5 December 2004
Available online 7 March 2005

Auditory fMRI in humans has recently received increasing attention from cognitive neuroscientists as a tool to understand mental processing of learned acoustic sequences and analyzing speech recognition and development of musical skills. The present study introduces this tool in a well-documented animal model for vocal learning, the songbird, and provides fundamental insight in the main technical issues associated with auditory fMRI in these songbirds. Stimulation protocols with various listening tasks lead to appropriate activation of successive relays in the songbirds' auditory pathway. The elicited BOLD response is also region and stimulus specific, and its temporal aspects provide accurate measures of the changes in brain physiology induced by the acoustic stimuli. Extensive repetition of an identical stimulus does not lead to habituation of the response in the primary or secondary telencephalic auditory regions of anesthetized subjects. The BOLD signal intensity changes during a stimulation and subsequent rest period have a very specific time course which shows a remarkable resemblance to auditory evoked BOLD responses commonly observed in human subjects. This observation indicates that auditory fMRI in the songbird may establish a link between auditory related neuro-imaging studies done in humans and the large body of neuro-ethological research on song learning and neuro-plasticity performed in songbirds.

Keywords: Imaging; Functional magnetic resonance imaging (fMRI); Blood oxygenation level dependent (BOLD) response; Auditory; Cognitive; Songbird; Starling (*Sturnus vulgaris*)

Introduction

Functional magnetic resonance imaging (fMRI) is a powerful tool for studying brain function in living animals and human. The fMRI technique most frequently used for imaging brain activity relies on blood oxygenation level dependent (BOLD) contrast (Ogawa et al., 1990). This technique allows repeated analysis of information processing by neuronal networks during various tasks ranging from simple sensory motor to highly cognitive tasks. Although fMRI has been increasingly used in human cognitive studies, its use in small animal models has been more limited due to the required immobilization during imaging, which limits the type of questions that can be addressed.

In humans, communication through a learned spoken language provides an opportunity to design auditory stimulation paradigms ranging from simple sound stimulation to word or sentence recognition, interpretation, and memorization. The elicited neural responses are a combination of complex hierarchical sensory processes possibly influenced by previous experience (Bernal and Altman, 2001; Seifritz et al., 2001). Introduction of an animal model that evolved a similar type of communication would provide new opportunities for the validation of BOLD responses observed during auditory stimulation in humans.

A few animal groups (songbirds, hummingbirds, and parrots, and possibly whales, dolphins, and bats) have independently developed a learned "vocalization" system used in social interactions. Extensive neuro-ethological research in songbirds has yielded significant insight into how their brain generates and perceives species-specific vocalizations in different social contexts and developmental stages (see references in Jarvis et al., 2002). Although vocal communication presumably emerged independently in songbirds and humans (Hauser et al., 2002), intriguing parallels between song and human speech have been identified (for recent reviews, see Doupe and Kuhl, 1999; Kuhl, 2003; Wilbrecht and Nottebohm, 2003), including the existence of critical periods for learning, dependence of this process on auditory experience and feedback, and lateralization of sound production. Interestingly, in songbirds, the auditory and sensori-

* Corresponding author. Fax: +32 3 265 32 33.
E-mail address: vincent.vanmeir@ua.ac.be (V. Van Meir).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.12.058

motor brain regions involved in song acquisition and production show reliable electrophysiological responses to acoustic stimuli even under anesthetized conditions (Capsius and Leppelsack, 1996; Grace et al., 2003; Schmidt and Konishi, 1998). This system is therefore ideally suited for fMRI studies and potentially provides a unique model to study complex cognitive processes associated with vocal communication in a system which, contrary to the human brain, can easily be experimentally manipulated.

The present study explores BOLD fMRI as a new tool for the visualization of sound perception and processing in auditory brain regions of a songbird, the European starling. Results are interpreted in light of previous anatomical, molecular, and electrophysiological studies on songbirds, and comparable fMRI data of sound and speech processing in humans.

Materials and methods

Scanner specifications and animal preparation

Twelve adult male European starlings (*Sturnus vulgaris*, 75–95 g) were imaged at 300 MHz in a 7-T magnetic resonance microscope with horizontal bore (SMIS, UK) and 8 cm aperture gradient coils (Magnex Scientific Ltd., Oxfordshire, UK) with a maximum gradient strength of 0.40 T/m. All experimental procedures were approved by the committee on animal care and use at the University of Antwerp, Belgium.

Birds were anesthetized with an intramuscular injection in the chest of 0.4 ml of a mixture containing 10 ml medetomidine (1 mg/ml, Domitor, Orion, Finland) and 0.5 ml ketamine (50 mg/ml, Ketalar, Parke-Davis, Belgium). To maintain anesthesia, this mixture was continuously infused at a rate of 0.15 ml/h through a catheter positioned in the chest muscle. The head was positioned in a custom made stereotaxic device (Figs. 1A–B), leaving the ears free for auditory stimulation by a non-magnetic dynamic speaker based on a headphone set (type: HK 150, FNAC, France) positioned next to each ear at a distance of approximately 1 cm (Baumgart et al., 1998; Fig. 1C). A Helmoltz coil (45 mm) used in combination with a circular RF surface antenna (24 mm) positioned around the head served respectively for transmission and acquisition of radio frequency pulses. The body was wrapped in a jacket and laid on a warm water blanket connected to an adjustable heating pump (Neslab Instruments, ex111, Newington, CT, USA). Temperature was continuously monitored with a rectal probe (SA-instruments, Inc., New York, USA) and maintained at 41.5 ± 0.5°C. Respiration rate and expired CO_2 concentration were monitored with a capnometer (Capstar-100, CWI Inc., Diss Norfolk, UK) to keep the starlings in optimal and stable physiological conditions during the experiment.

Imaging settings

A set of 3 sagittal, 3 horizontal, and 3 coronal GE scout images and a set of 20 coronal GE images were first acquired to determine the position of the brain (FOV = 40 mm, TE = 5 ms, TR = 500 ms, acquisition and reconstruction matrix = 128 × 128).

Functional brain imaging was then done using a T_2^*-weighted multi-slice gradient-echo fast low angle shot (GE-FLASH) sequence (TR = 80 ms, TE = 14 ms, flip angle = 11°, acquisition matrix = 128 × 64, FOV = 30 mm, slice thickness = 0.7–0.8 mm, gradient ramp time = 1000 μs). Long gradient ramp times reduced the gradient noise to 70–80 dB. Images were reconstructed to a matrix of 128 × 128 using a hanning filter of 25%. Two sagittal slices were acquired with a total acquisition time per image set of 5 s and a spatial resolution of 230 × 230 μm. High resolution images were also obtained at the same position with either a GE sequence (TR = 500 ms, TE = 7 ms, acquisition and reconstruction matrix = 256 × 256, FOV = 30 mm, slice thickness 0.7–0.8 mm, and 2 averages) or a SE sequence (TR = 500 ms, TE = 10 ms).

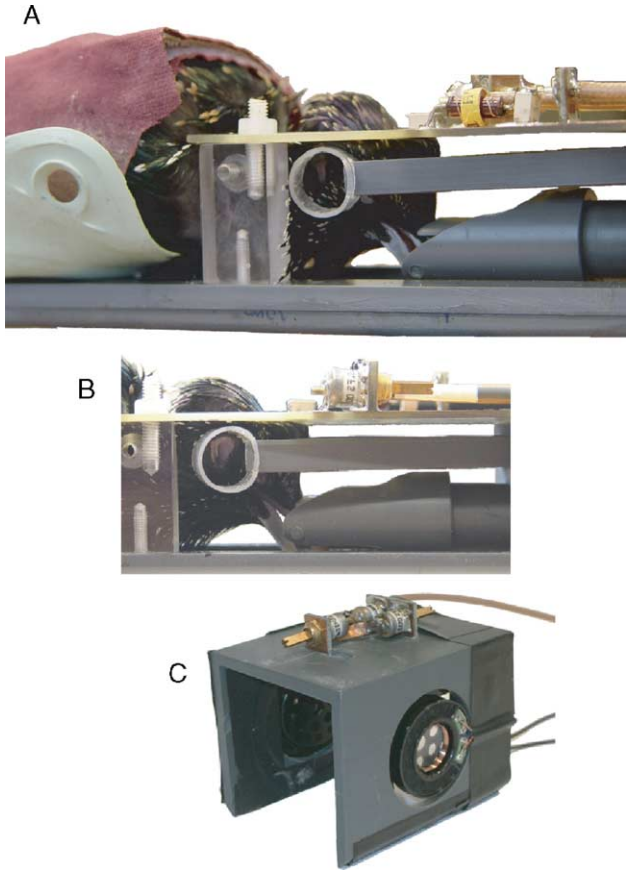


Fig. 1. Experimental setup in function of optimal auditory stimulation during fMRI. A stereotaxic device (A) was used to immobilize the head consisting of a beak holder and two hollow cylindrical tubes which left the ears free for auditory stimulation, see insert (B). Acoustic stimuli were applied using dynamic speakers that had their magnet removed (C).

Stimulation protocols

Auditory stimuli were presented with a maximum sound pressure level (SPL) of 100 dB in 28 (Experiment 1) or 14 (Experiment 2) repeated blocks of respectively 30 or 60 s followed by 30 or 60 s of rest (Fig. 2). Images were collected with a block-design paradigm consisting of either 28 cycles (here defined as trials) of 6 images collected during stimulation and 6 images collected during rest (Fig. 2A; Experiment 1), or 14 cycles of 12 images collected during stimulation and 12 images collected during rest (Fig. 2E; Experiment 2), resulting in sets of 336 functional images. Each experiment, which was preceded by the acquisition of 12 dummy images to allow the signal to reach a steady state, thus took approximately 30 min. Each paradigm was applied in 6 different birds.

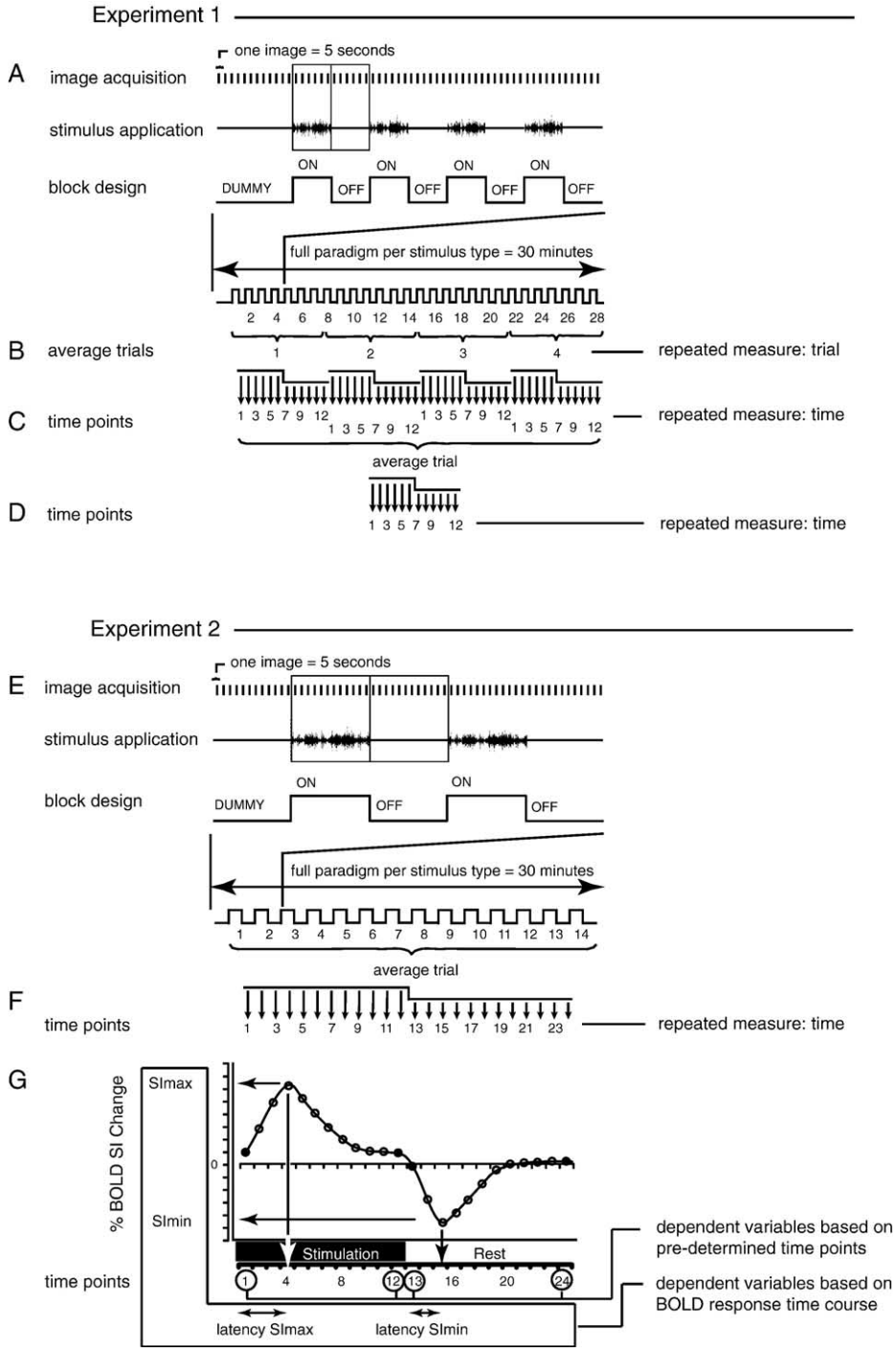


Fig. 2. Data acquisition and processing: experimental design, terminology, and variables. This figure describes the paradigms used in two experiments (see A and E) and illustrates the averaging of the trials and time points prior to the analysis of signal intensity changes in defined ROIs that occurred due to auditory activation (B–D and F). (B) The raw data obtained during the 28 trials of experiment 1 were divided in 4 groups of 7 subsequent trials and these 7 trials were averaged which resulted in 4 new average trials. (C) These trials each include 12 time points. The signal intensity in the subsequent time points represents the time course of the BOLD response (see Fig. 3E for corresponding results). (D) After changes in signal intensity between subsequent time points of the BOLD response time course and changes in the BOLD response time courses between subsequent trials were evaluated, the trials were further averaged to one mean trial for further analysis. See Results for detailed description of the statistical tests. (F) In experiment 2, one average time course was obtained from the 14 original trials (see Figs. 4C and 5 for corresponding results). (G) Subsequently, specific time points from the BOLD response time course were determined (see corresponding results in Fig. 6). See Results for detailed explanations of the statistical analysis of these data. The right column indicates some of the variables used in the analyses of variance (ANOVAs).

In all birds, three consecutive experiments were performed in random order during which the birds were exposed to white noise or music (Bach: Brandenburg Concerto No. 4, Allegro) or conspecific song (male starling song recorded during the breeding season). The stimuli span a wide range in complexity and social relevance to the bird. The birds used for experiment 2 (60 s stimulation/60 s rest) were additionally exposed to repeated sections of 5 s of the same small sequence of the conspecific song and of the introductory notes of the Brandenburg concerto. The stimuli were prepared and presented by Presentation software (Neurobehavioral Systems, Inc., version 0.55, Albany, CA, USA).

Image processing

Changes in local susceptibility properties resulting in changed T_2*-values following auditory stimulation were analyzed for each type of stimulus applied to each individual bird with MEDx software (version 3.41, Sensor Systems Inc, Sterling, KS, USA). Before *Z* score calculations, images were preprocessed by specific algorithms which included (1) motion detection between subsequent images by means of a center of intensity algorithm in three directions; (2) spatial smoothing with a 0.7 mm (3 pixels) full-width-at-half-maximum Gaussian kernel; (3) normalization to a whole image mode value of 1000. The activated regions were defined by comparison of the first 6 images obtained during stimulation and the first 6 images obtained during rest with a voxel-level paired *t* test. The resulting functional activation maps display the *Z* score of each voxel. The *Z* scores of significantly activated pixels ($P < 0.05$) were overlaid on a high-resolution image with *Z* scores color coded on a predetermined color scale (Figs. 3B,D and 4B).

Analysis of signal intensity in defined regions of interest

A region of interest (ROI) analysis was subsequently performed. Rectangular ROIs of 6 pixels were first delineated at the level of the primary telencephalic auditory region, Field L, and the secondary telencephalic auditory region, caudomedial nidopallium, NCM (see Fig. 3C). The ROI in Field L was located at the center of the area that was activated by white noise and kept constant for a given bird throughout the different analyses of various stimulations. We confirmed that this ROI was actually located in the area defined as Field L by the laboratories of Leppelsack and Hausberger (see for example George et al., 2004) by comparing its position relative to the location of various landmarks that are visible in the high-resolution images that had been collected.

The second ROI was placed caudal to the ROI of Field L with a gap of exactly 4 pixels (i.e., 4 × 230 μm = 920 μm). The distance between the two telencephalic ROIs corresponds to known distances between electrophysiologically defined functional subdivisions of the auditory telencephalon of starlings. This is illustrated by George et al. (2004) (see Figs. 7 and 8 of this reference). Their terminology, however, does not totally fit the revised avian brain nomenclature (Reiner et al., 2004; http://www.avianbrain.org/); to meet the recommendations of this new nomenclature we named as caudomedial nidopallium (NCM) the ROI caudal to Field L.

In the datasets of experiment 2, a third ROI was also delineated at the level of the thalamic auditory relay center, nucleus ovoidalis (Ov). Its delineation was also based on the activation map obtained with white noise as illustrated in slice 2 in Figs. 4B–C.

The mean signal intensity (SI) of the six pixels in these ROIs was then calculated for each acquired image. This mean SI for each image (time point) of a stimulation–rest trial was then expressed as the percent change relative to the mean SI during the last 2 (for experiment 1) or last 4 (for experiment 2) time points of the rest period of the respective trial. All further averaging and analysis were based on these percentage BOLD SI changes. For each ROI and stimulation type, the BOLD amplitude of the response was also defined as the percent change of the mean SI in the first 6 images of the stimulation period relative to the mean SI in the first 6 images of the subsequent rest period.

Cluster analysis of SI time courses

A time course was generated by averaging the 28 consecutive trials to obtain a 12-dimentional vector space, where each average voxel's time course is represented as one point. A reduction of dimensions was first obtained by Principle Components Analysis, i.e., by projecting data points onto the most relevant components (3 to 7). The *K* means clustering schemes (Fadili et al., 2000; Goutte et al., 1999) were then applied to cluster voxels with similar time courses, based on all inter-distances of the individual data points in the cluster space.

Statistical analyses

All analyses were done by multi-factors analyses of variance (ANOVA) using SPSS (SPSS, Inc., Chicago, IL, USA). These analyses were followed, when appropriate, by post hoc comparisons with the Tukey Honest Significant Difference test. All data are presented in the figures as means and standard errors (SEM). Stimulus type (white noise, music and song, and their repeated versions), brain region (Field L and NCM), and brain side (left and right) were considered as independent factors while trials and successive time points within a trial were considered as repeated factors.

Results

Experiment 1

BOLD fMRI in songbirds reveals region-specific auditory discrimination: statistical brain maps

In the first experiment, we applied acoustic stimuli in 28 repeated blocks of 30 s stimulation interleaved with rest periods of 30 s resulting in consecutive stimulation–rest trials of 60 s (Fig. 2A). Comparison of the mean signal intensity (SI) of images obtained during the 30 s stimulation and rest periods revealed differences in activation of the two main subdivisions of the auditory telencephalon. *Z* score maps showed that stimulation with white noise activated a region at the level of Field L (Fig. 3B). More complex sounds such as music or conspecific song extended the activation in a caudal direction into NCM. Based on these findings, a region of interest (ROI) analysis was performed focusing on two ROIs of 6 pixels (Fig. 3C), one in the area activated by all three stimuli (Field L) and one in the more caudal region only activated by complex stimuli (NCM).

Detailed analysis of regions of interest

A 3-way ANOVA with the BOLD response amplitude as dependent variable (i.e., the average percent SI change between stimulation and rest periods over the 28 trials) and with the stimulus type, brain side, and brain region as independent factors demonstrated the existence of a significant effect for brain region ($F_{1,60} = 71.396$; $P < 0.001$) and stimulus type ($F_{2,60} = 3.824$; $P < 0.05$), but not for the brain side ($F_{1,60} = 0.736$; $P = 0.39$). A significant interaction between brain region and stimulus type ($F_{2,60} = 5.627$; $P < 0.01$) was also observed, indicating that the two areas responded differentially to each stimulus (Fig. 3C). Post hoc comparisons on the interaction between brain region and stimulus showed that all stimuli elicited the same response in Field L (Tukey HSD; all $P > 0.99$), whereas NCM showed differential responses to white noise and conspecific song ($P < 0.005$). Music induced an intermediate response that was not significantly different from white noise nor conspecific song (both $P > 0.22$). The amplitude of the response also differed significantly between Field L and NCM for white noise and music (all $P < 0.001$), but not for song ($P = 0.094$).

Stability of the BOLD response amplitude and SI time course over subsequent trials

Amplitude. To determine whether repetition of the same stimulus had any effect on the observed responses, the set of 28 repeated trials was divided in 4 groups of 7 successive trials each. In these groups of 7 trials, we calculated the *Z* score maps for each stimulus type in each bird (Fig. 3D). These maps showed no apparent change in the activation pattern due to stimulus repetition. To confirm this conclusion statistically, ROI data from the 6 birds were analyzed separately for each stimulus type, brain region, and brain side by 12 separate ANOVAs (3 stimuli × 2 brain regions × 2 brain sides) with one repeated factor constituted by the BOLD amplitude changes averaged over the 4 groups of 7 consecutive trials (see Fig. 2B). No significant effect of the repeated factor (trial repetition) was observed in these 12 ANOVAs except in the analysis of effects of white noise in the right (but not left) Field L (data not shown). It can therefore be considered that the differences in BOLD signal between stimulation and rest periods are stable over time.

Time course. Since 12 images were acquired during one stimulation–rest trial, a detailed reconstruction of the SI changes during and after the stimulus could also be made. We thus evaluated the stability of these SI changes at the level of individual time points within one trial. As for the amplitude analysis, results of the 28 repeated trials were divided in 4 groups of 7 successive trials each (see Fig. 2B) and the average SI time course in each group of 7 trials was calculated (Fig. 3E). The changes in BOLD SI in the 12 images collected during a trial (average SI time course in each group) and between trials (between groups) were analyzed separately for each stimulus, brain area, and brain side by ANOVAs with two repeated factors: the 4 groups of 7 successive stimulation–rest periods (coded as trials afterwards) and the average BOLD SI time course within these trials (coded as time; 12 time points in each case) (Fig. 2C).

The time effect (SI change during a trial) was highly significant for all stimuli in Field L and NCM on both brain sides, except for white noise in the left and right NCM. Out of twelve analyses (3 stimulus types × 2 brain regions × 2 sides), only one overall effect of trials was detected: for white noise in the left NCM. The interaction between time and trials was significant in only one case: in the right Field L during white noise presentation. Considering the very small number of significant effects of trial and of interactions between trial and time, it can be concluded that the temporal aspects of the BOLD SI time course are consistent over repeated trials.

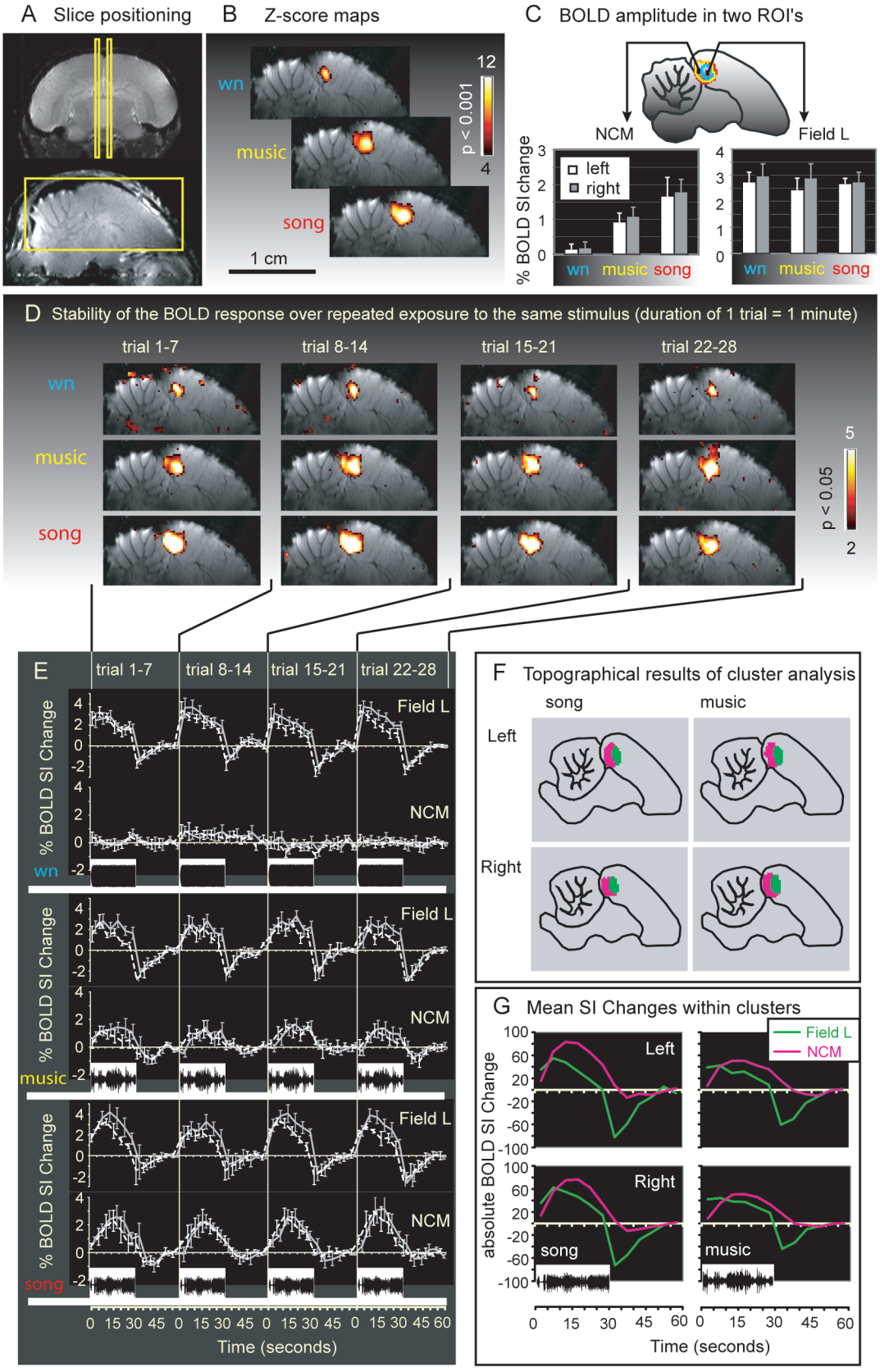
Region specificity and lateralization of the BOLD SI time course

Since observed BOLD response amplitude changes and SI time course were not affected by trial repetition, we averaged the respective time points of all 28 trials and evaluated these average SI time course curves for differences between the two telencephalic regions (Field L and NCM) and the two brain sides. Given the significant differences between stimulus types and the significant interaction between stimulus type and brain region observed in the analysis of the average BOLD response amplitude described earlier, we compared the SI time course for each stimulus type separately. For each stimulus, the SI time course was analyzed by three-way repeated measures ANOVA

Fig. 3. Auditory activation in the songbird telencephalon. Summary of fMRI data obtained during experiment 1 using a paradigm of 30 s stimulation and 30 s rest. (A) One sagittal slice was positioned at 0.5 mm from the midline in each hemisphere as shown on a frontal brain section (top). The yellow rectangle (bottom) corresponds to the position of the *Z* score maps presented in panels (B and D). (B) *Z* score maps illustrating the localization of significant signal intensity changes during auditory stimulation: from top to bottom activation during white noise (wn), a concerto of Bach (music) and a stimulation with song from a male starling (song). The location of the activated areas is schematically presented in (C) (top; wn: blue; music: yellow; song: red). Two ROIs of six pixels each were selected based on these activation patterns to evaluate the discrimination properties of the two regions in more detail. The average BOLD response amplitude for each stimulus was determined in each ROI (bottom). Based on their position and auditory discrimination properties, these regions were identified as NCM and Field L. Values of both left (white bars) and right (gray bars) hemispheres were not significantly different in any of the ROIs. (D) To visualize possible effects of stimulus repetition, 4 maps were constructed based on the 28 successive activation trials. Each map represents the *Z* score obtained from 7 stimulation versus rest trials in the right brain hemisphere of one individual. The spatial extend of the BOLD response is similar in all successive maps within one stimulation type. (E) To further illustrate the stability of signal intensity changes across stimulus repetitions, the mean traces of all 6 animals corresponding to the 4 *Z* score maps made in (D) are illustrated. The shape of the BOLD response was stable in time and did not differ between the left (white dotted line) and right (grey line) hemisphere. The shape of the response, however, differed between Field L and NCM as a function of the stimulus type. (F) Cluster analysis (*K* means cluster algorithm) of pixels based on the temporal aspects of the signal intensity changes identifies two distinct brain regions. The green region on the figure overlaps with the region activated by white noise (see B and C) and is equivalent to Field L, while the region indicated in pink represents the more caudal extension of the auditory area only activated by complex stimuli and represents NCM. Clustering was based on images acquired during application of the complex stimuli only: song (left) and music (right). No obvious difference could be observed in the areas delineated based on exposure to these two stimuli or between the left (top) and right (bottom) hemisphere. (G) Mean signal intensity changes in the green (Field L) and pink (NCM) clusters. Notice the post-stimulus undershoot in Field L whereas NCM shows a delayed BOLD response onset and no post-stimulus undershoot. Shape of the trace does not differ between conspecific song (left) and music (right) nor between left (top) and right (bottom) hemisphere, although the response in NCM during music does not reach the same amplitude as with conspecific song stimulation.

with the successive time points as repeated measure (code: time; Fig. 2D) and the brain regions and brain side as independent factors (Table 1).

As expected, the overall effect of time was highly significant for all stimuli (all $P < 0.001$). The interaction between time and brain region was also significant for all stimuli (all $P < 0.001$), but



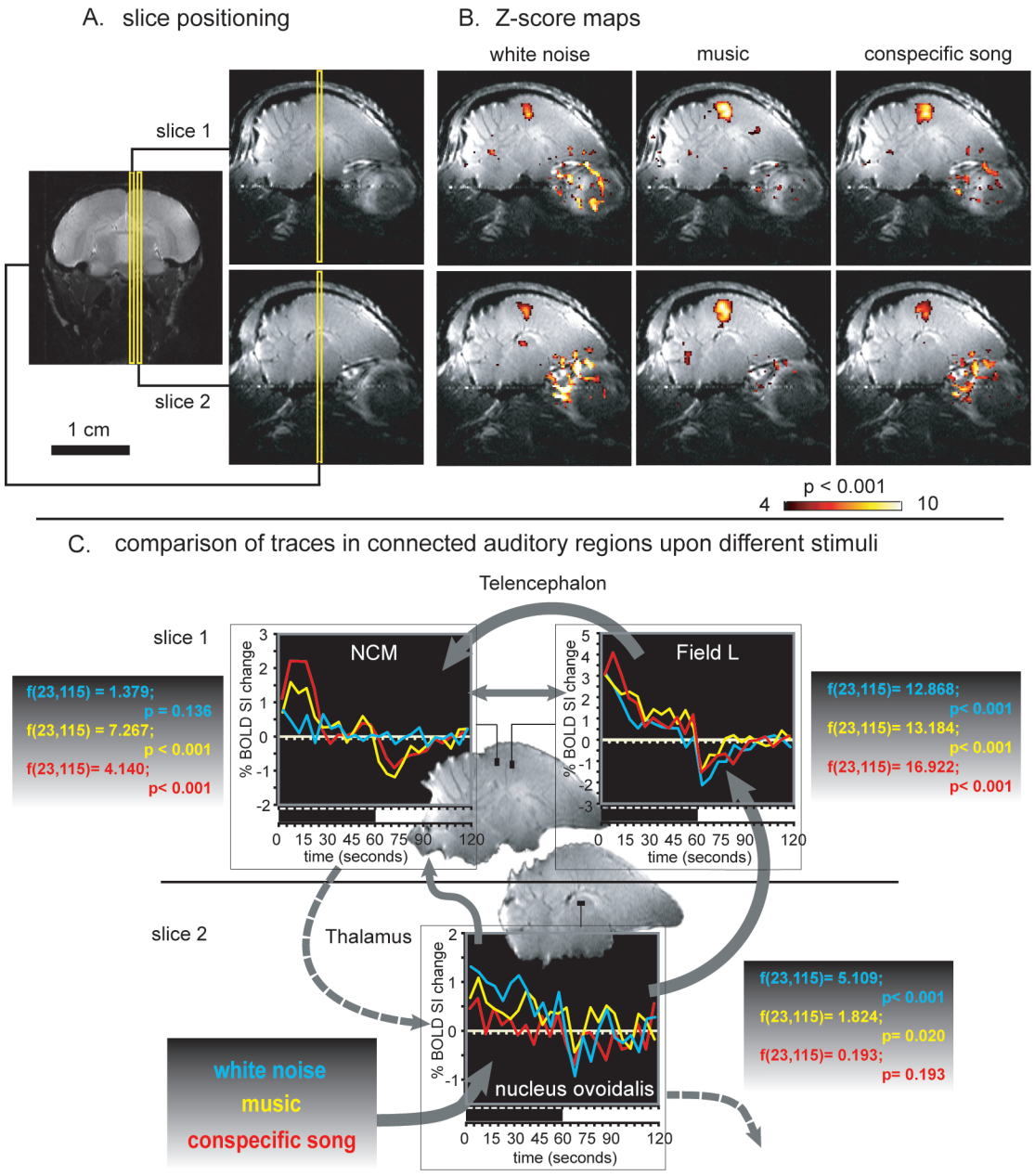


Fig. 4. Brain activation by auditory stimuli and functional connectivity of known and unknown auditory networks. Data presented in this figure were obtained with a paradigm of 60 s stimulation and 60 s rest. (A) Two sagittal slices were positioned in the right hemisphere as shown on the frontal section (left). The more medial and lateral slices were called slice 1 and 2, respectively. (B) On *Z* score maps corresponding to the three stimulus types (white noise, music, and song), activation is observed in the thalamus and cerebellum in addition to Field L and NCM. The highlighted pixels around the eye are due to eye movements. (C) The telencephalon receives input from the thalamic nucleus ovoidalis. Therefore, we evaluated by ROI analysis the BOLD response signal intensity changes observed in Field L, NCM, and nucleus ovoidalis as a function of stimulus complexity and position of these nuclei in the auditory network. Position of the ROIs is indicated on the high-resolution images (slice 1: NCM, left; and Field L, right; slice two: nucleus ovoidalis). Broad arrows between graphs indicate connectivity. The black horizontal bar under each graph indicates the stimulation period. In each graph, the trace of each stimulus is color coded (white noise, blue; music, yellow; song, red). The significance level of the changes in the trace is indicated next to each graph.

there was never a significant interaction between time and brain side nor between time, brain side, and brain region (all $P > 0.764$). The other main effects (brain region and brain side) and their interactions were not significant (all $P > 0.223$), with only one exception: white noise induced a significantly different response in the two brain regions ($P < 0.001$).

Cluster analysis of SI time course

The pattern of changes in time of the BOLD signal during exposure to complex stimuli (music or songs) could additionally be used to delineate Field L and NCM (Figs. 3F–G). Field L showed a very transient response during stimulus presentation associated with a sharp negative undershoot at the end of the stimulation,

Table 1
Analysis of the BOLD response time course

Stimulus types→	*df*	White noise		Music		Song	
		F	*P*	*F*	*P*	*F*	*P*
Time	11,220	51.870	**<0.001**	42.614	**<0.001**	55.343	**<0.001**
Time × brain region	11,220	41.023	**<0.001**	13.288	**<0.001**	7.842	**<0.001**
Time × brain side	11,220	0.269	0.991	0.672	0.764	0.568	0.853
Time × brain side × brain region	11,220	0.154	0.999	0.138	1.000	0.251	0.993
Brain region	1,20	29.868	**<0.001**	0.777	0.389	0.137	0.715
Brain side	1,20	1.211	0.284	1.581	0.223	1.131	0.300
Brain side × brain region	1,20	0.519	0.480	0.883	0.359	0.882	0.359

Result of three 3-way repeated measures ANOVA with the signal intensity of 12 consecutive time points in the average trial as repeated factor (code: time) and brain region (Field L and NCM) and brain side (left and right hemisphere) as fixed factor. The difference in signal intensity time course between brain regions and brain sides was evaluated for each stimulus type separately. Effects were considered significant for $P < 0.05$ (indicated in bold).

whereas NCM displayed a more sustained response to the stimuli followed by a far less pronounced negative undershoot (Figs. 3E and G). Cluster analysis of the images acquired during exposure to either music or conspecific song was thus used to pool pixels with a similar temporal BOLD response. This analysis clearly discriminated two regions that correspond in shape and position to Field L and NCM (Fig. 3F). In addition to functional maps demonstrating average differences between stimulation and rest periods (*Z* score maps; Figs. 3B and D), functional maps based on a clustering algorithm analyzing the time course of the response were thus observed to discriminate functional subregions during exposure to complex sounds.

Experiment 2

Data described above raised several practical issues concerning the design and scientific potential of fMRI measurements in songbirds. The first concerned the timing of stimulus presentation: was the decrease in BOLD response at the end of the 30 s of stimulation representing a form of saturation of the response and was the observed post-stimulus undershoot due to the termination of the stimulus or just a passive extension of the saturation effect? If stimulation was prolonged, would the signal just remain at the baseline and only show a negative undershoot at stimulus termination or would the negative undershoot occur during the stimulation period?

Temporal aspects of the BOLD response indicate differential neuronal processing in thalamic and telencephalic areas

To address these questions, we designed a second experiment in which the duration of both the stimulation and rest periods was prolonged to 60 s (Fig. 2E) and two parasagittal slices were positioned in the right hemisphere (instead of one in each hemisphere) as shown in Fig. 4A. The *Z* score maps obtained in these conditions with different stimuli clearly identified an activation pattern in the telencephalon similar to that observed in the first experiment.

In this prolonged stimulation paradigm, there was also no evidence of habituation of the BOLD response to repeated stimulation (data not shown). All analyses presented below are thus based on the percentage of BOLD SI change averaged over the 14 trials (Fig. 2F).

The shape of the BOLD response in two ROIs located in Field L and NCM clearly showed that, with 60 s stimulation periods, SI reaches a maximum after 5–20 s but declines thereafter and reaches a stable level throughout stimulation. When stimulation was switched off, the SI showed a pronounced negative undershoot in both Field L and NCM, indicating that this negative response is linked to stimulus termination (Fig. 4C).

Depending on the stimulus type, other structures were also activated such as parts of the cerebellum (see white noise top panel and music bottom panel in Fig. 4B) and a few small spots in the frontal telencephalon (see music top panel Fig. 4B). Activation was also observed in the auditory relay center of the thalamus, nucleus ovoidalis (Ov) (see white noise bottom panel, Fig. 4B). However, based on the *Z* score maps (bottom panels, Fig. 4B) and SI changes observed during stimulation (Fig. 4C), activation in this region was less prominent during the presentation of complex stimuli contrary to what is observed in NCM. White noise led to a significant change in BOLD SI during stimulation in Ov, consistent with a significant response in Field L. However, presentation of complex stimuli such as music and song, which lead to a very specific BOLD response in both Field L and NCM, produced almost no responses in Ov.

Quantitative analysis of the BOLD response time course in Field L and NCM: effects of stimulus alterations

To evaluate the impact of stimulus characteristics on the BOLD SI time course, we manipulated the complex stimuli, song, and music so that they were now presented during 60 s as either their authentic version (continuous version) or as a repetition of the same small sequence (12 × 5 s) (repeated version). The average BOLD response observed during the 14 repeated presentations of these stimuli is illustrated in Fig. 6 separately for each stimulus type and brain region.

Quantitative analyses were performed on these curves by focusing, first, on specific time points during stimulus presentation and during the rest period and, second, on specific aspects of the BOLD response observed (Fig. 2G). During the first set of analyses, we determined by four separate two-way ANOVAs, with stimulus types (5 stimuli described above) and brain regions (Field L and NCM) as independent factors, the changes in SI observed in each bird at four pre-determined points in the protocol namely the first and the last point of the stimulation period (i.e., time points 1 and 12) and the first and the last point of the rest period (i.e., time points 13 and 24; see Figs. 2G and 6A, Table 2). These analyses identified an overall significant difference between Field L and NCM for the first time point of both the stimulation and rest period (i.e., points 1 and 13, $P < 0.001$ in each case). For all stimuli, a higher percentage increase in SI was observed at time point 1 in

Table 2
Analysis of pre-determined features of the BOLD response time course

Position of selected time points from the average trial →	*df*	1		12		13		24	
		F	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Brain region	1,50	73.992	**<0.001**	1.515	0.224	20.207	**<0.001**	0.537	0.467
Stimulus type	4,50	2.132	0.090	1.788	0.146	0.408	0.802	0.314	0.868
Brain region × stimulus type	4,50	0.479	0.751	0.681	0.609	1.565	0.198	0.954	0.441
BOLD response parameters →	*df*	SI_{max}		Latency SI_{max}		SI_{min}		Latency SI_{min}	
		F	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Brain region	1,50	44.983	**<0.001**	14.969	**<0.001**	20.752	**<0.001**	12.250	0.001
Stimulus type	4,50	4.177	**0.005**	0.518	0.722	0.909	0.466	1.109	0.363
Brain region × stimulus type	4,50	1.072	0.380	0.364	0.833	4.384	**0.004**	0.609	0.658

Results of eight 2-way ANOVAs analyzing pre-determined time points during a trial (top half of the table, points 1, 12, 13 and 14) or points selected on the individual curves as a function of the observed response (SI_{max} and SI_{min} and their latencies; bottom, see results section for additional explanations). Brain region (Field L and NCM) and stimulus type (white noise, music, repeated music, song, and repeated song) were considered as fixed factors; the dependent variable was the average signal intensity (or latency) of the respective time points over 14 trials. Effects were considered significant for $P < 0.05$ (indicated in bold).

Field L than in NCM and a higher decrease was observed at time point 13 (see Figs. 5 and 6A). No difference was in contrast observed at the end of these two periods, i.e., at points 12 and 24 (see Table 2). No overall significant difference was observed between stimuli and there was also no significant interaction between stimulus type and brain region (all $P > 0.090$).

A second set of ANOVAs was then performed on data selected based on the BOLD response observed in each subject. For this purpose, we measured in each individual curve (i.e., for each bird, stimulus type and brain region) four parameters describing key features of the BOLD response focusing on its most transient aspects (see Fig. 2G). These parameters are (1) the maximal SI reached during the stimulation period (SI_{max}) and (2) the time point when this maximum was reached relative to the onset of the stimulus (latency SI_{max}), (3) the minimal SI reached during the rest period (SI_{min}), and (4) the time when minimum was reached relative to the start of the rest period (latency SI_{min}) (Fig. 6B; Table 2). Separate two-way ANOVAs of each parameter with stimulus type and brain region as independent factors identified for all parameters a significant difference between Field L and NCM (all $P < 0.001$). Significantly higher maximal and minimal values were observed in Field L as compared to NCM, but these maxima/minima were invariably observed significantly earlier in Field L than in NCM. In other words, Field L reacted more intensely and more rapidly than NCM to all stimulus types (see Figs. 5 and 6B).

An overall effect of the stimulus type was found for SI_{max} only ($P < 0.01$). The corresponding data were thus further analyzed by Tukey HSD post hoc tests, which indicated the presence of overall significant differences between the SI_{max} induced by song versus white noise and song versus repeated music. Other post hoc comparisons were not significant. In particular, there was no significant overall difference between the SI_{max} induced by complex stimuli (song or music) and their repeated version, despite the fact that the continuous versions induced each time a numerically higher response.

No interaction was found between brain region and stimulus type for SI_{max}, its latency, and the SI_{min} latency, but a significant interaction was detected for SI_{min} ($P < 0.005$). Inspection of the corresponding data in Fig. 6B indicates that this interaction results largely from the dramatic post-stimulus undershoot observed at the end of the white noise stimulus in Field L but not in NCM.

In conclusion, these analyses indicate that the maximum and minimum intensities reached by the BOLD response are important parameters for the evaluation by fMRI of differences in auditory processing. The baseline level of the BOLD response reached at the end of the stimulation period (represented by time point 12) is, in contrast, not a useful indicator for differences in response to different acoustic stimuli. The analyses confirmed statistically that the BOLD responses have a later onset (indicated by the analysis of time point 1 and of SImax latencies) and a lower overall magnitude (as indicated by the SImax) in NCM than in Field L, indicating differences in the mechanisms that are driving the BOLD activation observed in both regions.

Discussion

To our knowledge, this is the first time that brain activity is imaged in a small songbird (75–95 g) with a non-invasive method based on BOLD fMRI. Complex auditory stimulation was successfully applied in the noisy environment of the MRI scanner and revealed SI changes in the auditory telencephalon that confirm and extend conclusions previously obtained by electrophysiology or analysis of immediate early genes expression. We here present a proof of concept for the design of an adequate imaging paradigm for auditory responses. Four issues had to be considered.

1. *Image quality.* Earlier MR imaging of songbird brains produced high-resolution images by application of T1- and T2-weighted spin echo (SE) sequences at 7 T (Van der Linden et al., 1998, 2002; Van Meir et al., 2004; Verhoye et al., 1998). Imaging with classical T2*-weighted gradient echo (GE) sequences (required to detect hemodynamic susceptibility changes in BOLD fMRI) is however affected in birds (severe susceptibility artifacts) by the air cavities that developed in the skull in relation with the flying life style of most species.
2. *Gradient sound.* The MRI scanner produces sounds of high intensity that might interfere with the detection of auditory stimuli. Uncoupling the auditory stimulus from image acquisition often solves this problem in human research because the maximal amplitude of the BOLD response is only reached 3–10

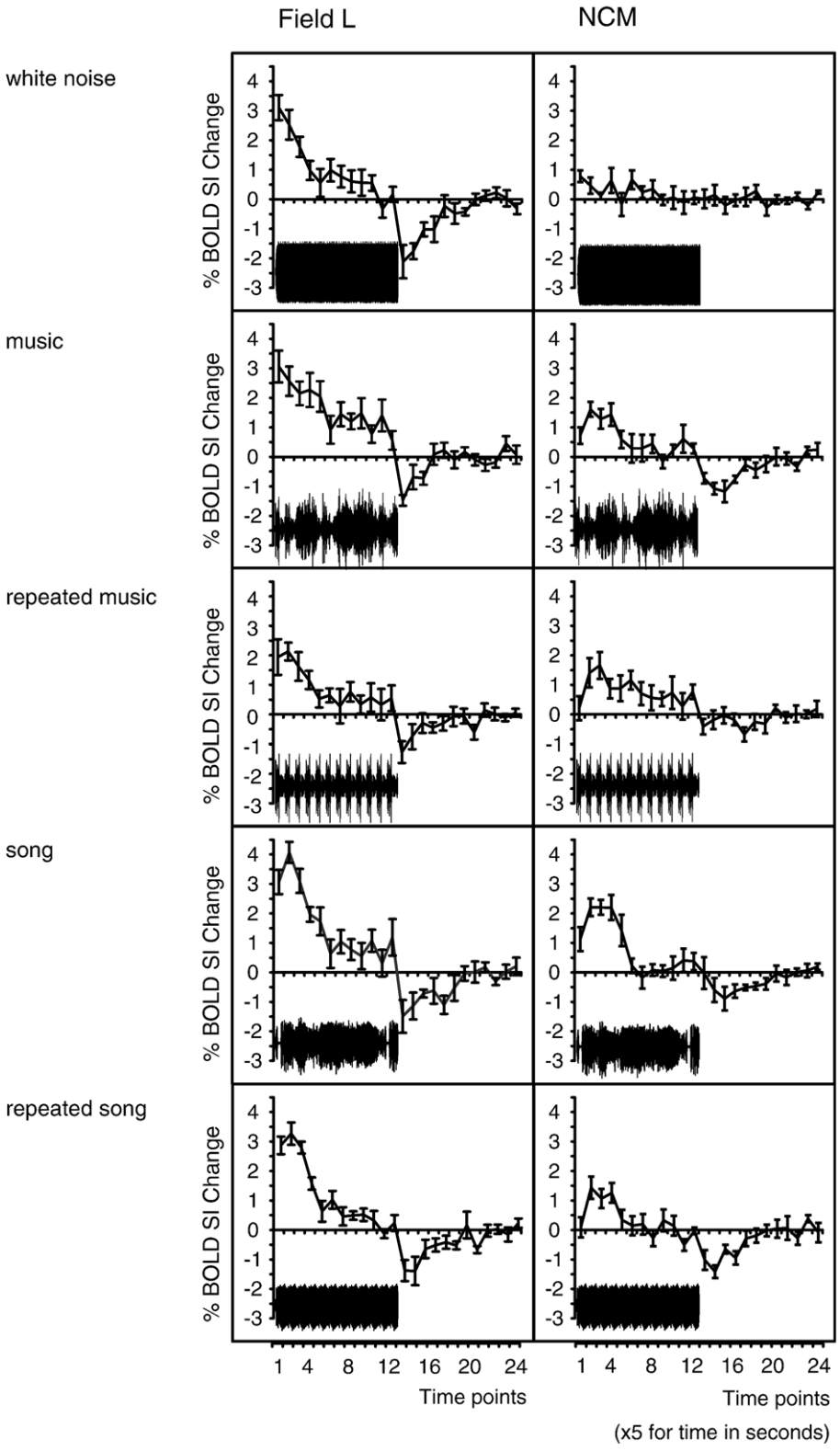


Fig. 5. Stimulus-specific processing in Field L and NCM during a variety of listening tasks. Average time course of the BOLD signal intensity changes obtained from 14 consecutive trials in six birds. Changes in signal intensity during 5 different stimuli are shown (top-down) for Field L (left) and NCM (right).

s after stimulus onset. Fast GE planar images (EPIs) with very long repetition times are often acquired using a sparse imaging paradigm to visualize activated regions (for recent reviews, see Amaro et al., 2002; Moelker and Pattynama, 2003). Since temporal aspects and optimal stimulus duration to obtain maximal BOLD responses were obviously unknown in song-

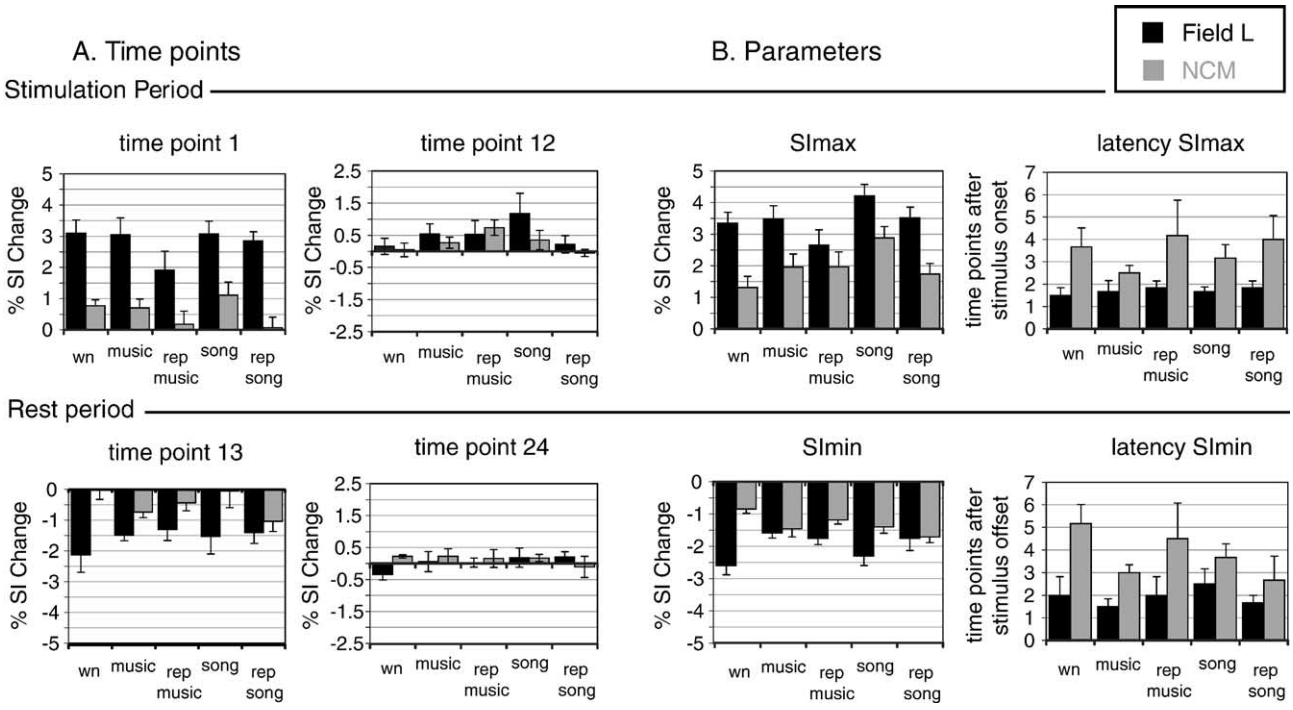


Fig. 6. Signal intensity and latency of representative time points during the BOLD response. (A) Signal intensity at time points defined by their position in the paradigm, namely the first (1) and last (12) time points of the stimulation period (top) and first (13) and last (24) points of the rest period (bottom). (B) Signal intensity at time points defined by the highest intensity (SI_{max}) during stimulation period (top) and the lowest intensity (SI_{min}) during rest period (bottom) with their respective latencies (right). Presented data are average data of six birds. In each graph, data from 5 different stimuli are shown; from left to right: white noise (wn), music and its repeated version (rep music), and song and its repeated version (rep song). For each stimulus data from ROIs, Field L (black bar) and NCM (gray bar) are shown.

birds and EPI is very loud and very prone to susceptibility, this approach was not an option.

3. *Temporal resolution.* Most of the BOLD fMRI experiments were conducted so far in mammals, it was thus hard to predict what would be the blood oxygenation response to a stimulus in a non-mammalian species with a different metabolic rate (birds have a body temperature of 41–42°C versus 36–37°C in mammals) and a substantially different pattern of telencephalic organization (Reiner et al., 2004).
4. *Paradigm design.* BOLD activations are usually subtle with SI changes in the range of 3–5%. Hence, averaging over subsequent stimulus trials is necessary for detection. Repeated exposure to a conspecific song is known to lead to habituation of molecular and electrophysiological events within NCM (Chew et al., 1995; Mello et al., 1995; Stripling et al., 1997) and thus might obscure interpretation of the results.

To deal with the 4 types of potential problems, we used a GE-FLASH sequence with long gradient ramp times to reduce the background noise levels (70–80 dB). The spatial (230 × 230 × 700 μm) and temporal (5 s) resolution respectively avoided susceptibility artifacts and were sufficient to evaluate the time course of the BOLD response. White noise was applied as a continuous noise at maximum SPL of 100 dB. Other more complex stimuli had a maximum peak level of 100 dB and had as a consequence a lower mean SPL as white noise. Despite this contrast in SPL, stimulation with white noise led to a focused activation that did not exceed the borders of Field L, whereas more complex stimuli like music or song activated brain areas extending into the more caudal region of NCM. It has been hypothesized that spreading of excitation is associated with the encoding of increasing stimulus intensity levels (Jäncke et al., 1998). During preliminary experiments, we varied the SPL of a continuous multi-tone stimulus which confirmed that broader areas are progressively activated as the SPL increases, but this activation always remained within the boundaries of Field L, even with increasing SPL up to 100 dB. Hence, the high SPL did not contribute to the activation of NCM during complex stimuli. The noisy environment of the MRI scanner did not obviously disrupt the activation of NCM by complex stimuli in the present experiments. This is consistent with the observation of a persistent induction of expression of the immediate early gene *zenk* in NCM in response to socially relevant auditory stimuli, even in the presence of background noise as long as the signal can be recognized (Vignal et al., 2004).

In experiment one, we found no evidence of habituation of the BOLD response amplitude or shape over different trials in anesthetized subjects (Figs. 3D–E). In experiment two, we subsequently focused on the BOLD response shape within the average trial. This average response was similar to the type of response observed during stimulation of the human auditory cortex (Giraud et al., 2000; Harms and Melcher, 2002; Jäncke et al., 1999; Seifritz et al., 2002, 2003), which might suggest similar neuro-vascular response properties (e.g., Buxton et al., 1998; Heeger and Ress, 2002; Logothetis, 2003). The initial peak (Field L = 3.5–4.1% after 5–10 s; $NCM_{white\ noise\ excluded}$ = 1.6–2.9% after 10–20 s) can be explained by the relatively fast increase in blood flow and a concomitant increase in oxygenated hemoglobin, both inducing an increase in SI. These events are followed by an increase in local blood volume which has a longer time lag to stimulus onset and termination and which has a negative effect on the SI. The presence

of such a blood volume change in our dataset, probably causing the habituation of the BOLD response during stimulation (Field L = reduced to 0.2–1.1% of rest baseline level; $NCM_{white\ noise\ excluded}$ = reduced to 0–0.6% of rest baseline level), is supported by the observation of a pronounced post-stimulus undershoot (Field L = 1.5–2.5% after 5–10 s; $NCM_{white\ noise\ excluded}$ = 1.2–1.7% after 15–20 s) which is likely due to the time lag in relaxation of the blood volume to resting baseline level whereas the inflow of fresh oxygenated blood has already ceased shortly after stimulus termination. A second indicator that the habituation of the BOLD response shape is related to an adjustment in the local blood volume could be found in the fact that the post-stimulus undershoot is reduced when the stimulation period is reduced to 30 s (thus before the baseline level is reached, and hence the adjustment of the blood volume has not been fully established) (Fig. 3E versus Fig. 5).

Not only was the observed BOLD response shape similar to the response observed in human experiments, but also the modulations of the response due to stimulus type were similar (Giraud et al., 2000; Harms and Melcher, 2002). Field L showed the most transient response when exposed to continuous white noise compared to the more temporally complex sounds including music or socially relevant sounds such as song (Figs. 4C and 5). This effect coincides with a significant effect of the stimulus type on the SI_{max} (Fig. 6 and Table 2) or the BOLD amplitude. Similar observations in humans have been related to differences in blood oxygenation levels and thus differences in neuronal activity between stimuli rather than to differences in blood volume (Seifritz et al., 2003). Our results in the songbird brain are thus consistent with the higher neuronal activity recorded in Field L of anesthetized birds when exposed to conspecific songs as compared to synthetic sounds or white noise (Grace et al., 2003).

The temporal aspects of the BOLD response were not only modulated by the stimulus type. During complex sound stimulation, there was also a significant difference in shape between field L and NCM (Figs. 3F–G). Differential responses between the two regions could be due to morphological components such as differences in vascularization or neural density. These morphological differences could have repercussions on the dynamics of the changes in blood flow and blood volume, which in part determine the BOLD response shape. However, the different dynamics of these regions in terms of oxygenated blood supply could also result from a differential underlying neuronal activity during acoustic stimulation.

Single unit recordings in the auditory telencephalon of songbirds demonstrate transient (phasic) excitation with the onset of continuous tone stimulation, while sounds with more complex acoustic features, such as modulations in temporal or spectral properties, elicit a sustained (tonic) excitation (Grace et al., 2003; Muller and Leppelsack, 1985). In NCM, the selectivity for stimuli with a certain complexity and the more sustained BOLD responses might be related to the underlying sustained neuronal activity during stimulus presentation. Field L contains both phasic neurons responding to stimulus onset and tonic neurons responding to complex acoustic features (Hausberger et al., 2000). The presence of phasic neurons in Field L might clarify the observation of faster changes in BOLD signal at stimulus onset and termination relative to NCM, which is the presumed site for processing of species-specific sounds (Mello et al., 1992). Electrophysiological and molecular studies have also revealed the implication of this area in functions concerned with song memorization (Bolhuis et al., 2000).

Auditory fMRI in the songbird: a missing link in cognitive neuroscience

The present data demonstrate the potential of fMRI for imaging auditory induced brain activity in songbirds. This in vivo technique permits repeated measures of brain activity in response to a variety of stimulations within one individual over a reasonably short time scale. Modulations in the activity of multiple brain regions can be identified and quantified by whole brain imaging, a goal which cannot easily be achieved by electrophysiological studies that must, by nature, focus on only a few neurons and/or brain regions at a time.

The learned vocal communication in songbirds constitutes a very attractive analogous model to study basic mechanisms that might underlie speech recognition and its subsequent reproduction, which can be regarded as the substrate for human language. In songbirds, auditory and cognitive processes can be directly controlled by short- and long-term endocrine (Ball et al., 2002) or pharmacological manipulations (Appeltants et al., 2002), or transitions from awake to anesthetized status (Capsius and Leppelsack, 1996; George et al., 2004; Schmidt and Konishi, 1998). This experimental approach combined with fMRI can highlight the involvement of previously ignored circuits that in other species proved to be important for auditory related learning processes. Such circuits include the cerebellum (Mathiak et al., 2002, 2004) in which we observed auditory-induced activation, and descending auditory pathways (Mello et al., 1998) that might be involved in the modulation of the BOLD signal in nucleus ovoidalis (He, 2003; Suga et al., 2003) (Fig. 5).

Whereas fMRI has never been used in the songbird model before, it is the non-invasive imaging method of choice for evaluating the function of the human brain. For ethological and practical reasons, experimental manipulations in humans are limited (Portas et al., 2000; Thiel et al., 2002), and similarly, pragmatic considerations limit long-term follow-up studies of neuronal changes occurring during speech acquisition in children. On these topics, fMRI research in the songbird should be able to provide complementary information by analysis of hemodynamic data.

Apart from the functional mapping of brain activity, quantitative information obtained from the BOLD response time course cannot usually be easily translated into underlying neuronal events (Heeger and Ress., 2002; Logothetis, 2003). Analysis of the BOLD response time course in a well-studied animal model will allow in parallel additional electrophysiological or molecular characterization of the corresponding neural substrate, with potential implications for the understanding of physiological events underlying the hemodynamic response to complex, and even learned, sensory inputs.

The successful application of BOLD fMRI in songbirds exposed to acoustic stimuli spanning a broad range of sensory complexities provides a new path for the analysis in small animal models of complex cognitive capacities. This initial fMRI study of songbirds should form the backbone for future research leading to increased interactions between researchers in the fields of birdsong and human verbal/acoustic learning.

Acknowledgments

Supported by grants from the National Science Foundation (FWO, project Nr G.0075.98N), and BOF-NOI and RAFO funds from the University of Antwerp to AVdL, grants from

the NINDS (NS 35467), the Belgian FRFC (Nbr. 2.4555.01), and the Government of the French Community of Belgium (ARC #99/044-241) to JB. JS is postdoctoral researcher and TB an aspirant with the National Science Foundation (FWO) Flanders. VVM is grant holder from the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT-V).

References

Amaro Jr., E., Williams, S.C., Shergill, S.S., Fu, C.H., MacSweeney, M., Picchioni, M.M., Brammer, M.J., McGuire, P.K., 2002. Acoustic noise and functional magnetic resonance imaging: current strategies and future prospects. J. Magn. Reson. Imaging 16, 497–510.

Appeltants, D., Del Negro, C., Balthazart, J., 2002. Noradrenergic control of auditory information processing in female canaries. Behav. Brain Res. 133, 221–235.

Baumgart, F., Kaulisch, T., Tempelmann, C., Gaschler-Markefski, B., Tegeler, C., Schindler, F., Stiller, D., Scheich, H., 1998. Electrodynamic headphones and woofers for application in magnetic resonance imaging scanners. Med. Phys. 25, 2068–2070.

Ball, G.F., Riters, L.V., Balthazart, J., 2002. Neuroendocrinology of song behavior and avian brain plasticity: multiple sites of action of sex steroid hormones. Front. Neuroendocrinol. 23, 137–178.

Bernal, B., Altman, N.R., 2001. Auditory functional MR imaging. Am. J. Roentgenol. 176, 1009–1015.

Bolhuis, J.J., Zijlstra, G.G., den Boer-Visser, A.M., Van Der Zee, E.A., 2000. Localized neuronal activation in the zebra finch brain is related to the strength of song learning. Proc. Natl. Acad. Sci. U. S. A. 97, 2282–2285.

Buxton, R.B., Wong, E.C., Frank, L.R., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med. 39, 855–864.

Capsius, B., Leppelsack, H.-J., 1996. Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. Hear. Res. 96, 59–70.

Chew, S.J., Mello, C., Nottebohm, F., Jarvis, E., Vicario, D.S., 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.

Doupe, A.J., Kuhl, P.K., 1999. Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. 22, 567–631.

Fadili, M.J., Ruan, S., Bloyet, D., Mazoyer, B., 2000. A multistep unsupervised fuzzy clustering analysis of fMRI time series. Hum. Brain Mapp. 10, 160–178.

George, I., Vernier, B., Richard, J.P., Hausberger, M., Cousillas, H., 2004. Hemispheric specialization in the primary auditory area of awake and anesthetized starlings (*Sturnus vulgaris*). Behav. Neurosci. 118, 597–610.

Giraud, A.L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., Kleinschmidt, A., 2000. Representation of the temporal envelope of sounds in the human brain. J. Neurophysiol. 84, 1588–1598.

Goutte, C., Toft, P., Rostrup, E., Nielsen, F., Hansen, L.K., 1999. On clustering fMRI time series. NeuroImage 9, 298–310.

Grace, J.A., Amin, N., Singh, N.C., Theunissen, F.E., 2003. Selectivity for conspecific song in the zebra finch auditory forebrain. J. Neurophysiol. 89, 472–487.

Harms, M.P., Melcher, J.R., 2002. Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. J. Neurophysiol. 88, 1433–1450.

Hausberger, M., Leppelsack, E., Richard, J., Leppelsack, H.J., 2000. Neuronal bases of categorization in starling song. Behav. Brain Res. 114, 89–95.

Hauser, M.D., Chomsky, N., Fitch, W.T., 2002. The faculty of language: what is it, who has it and how did it evolve? Science 298, 1569–1579.

He, J., 2003. Corticofugal modulation of the auditory thalamus. Exp. Brain Res. 153, 579–590.

Heeger, D.J., Ress, D., 2002. What does fMRI tell us about neuronal activity? Nat. Rev., Neurosci. 3, 142–151.

Jäncke, L., Shah, N.J., Posse, S., Grosse-Ryuken, M., Muller-Gartner, H.W., 1998. Intensity coding of auditory stimuli: an fMRI study. Neuropsychologia 36, 875–883.

Jäncke, L., Buchanan, T., Lutz, K., Specht, K., Mirzazade, S., Shah, N.J., 1999. The time course of the BOLD response in the human auditory cortex to acoustic stimuli of different duration. Brain Res. Cogn. Brain Res. 8, 117–124.

Jarvis, E.D., Smith, V.A., Wada, K., Rivas, M.V., McElroy, M., Smulders, T.V., Carninci, P., Hayashizaki, Y., Dietrich, F., Wu, X., McConnell, P., Yu, J., Wang, P.P., Hartemink, A.J., Lin, S., 2002. A framework for integrating the songbird brain. J. Comp. Physiol., A 188, 961–980.

Kuhl, P.K., 2003. Human speech and birdsong: communication and the social brain. Proc. Natl. Acad. Sci. U. S. A. 100, 9645–9646.

Logothetis, N.K., 2003. The underpinnings of the BOLD functional magnetic resonance imaging signal. J. Neurosci. 23, 3963–3971.

Mathiak, K., Hertrich, I., Grodd, W., Ackermann, H., 2002. Cerebellum and speech perception: a functional magnetic resonance imaging study. J. Cogn. Neurosci. 14, 902–912.

Mathiak, K., Hertrich, I., Grodd, W., Ackermann, H., 2004. Discrimination of temporal information at the cerebellum: functional magnetic resonance imaging of nonverbal auditory memory. NeuroImage 21, 154–162.

Mello, C.V., Vicario, D.S., Clayton, D.F., 1992. Song presentation induces gene expression in the songbird forebrain. Proc. Natl. Acad. Sci. U. S. A. 89, 6818–6822.

Mello, C., Nottebohm, F., Clayton, D., 1995. Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. J. Neurosci. 15, 6919–6925.

Mello, C.V., Vates, G.E., Okuhata, S., Nottebohm, F., 1998. Descending auditory pathways in the adult male zebra finch (*Taeniopygia guttata*). J. Comp. Neurol. 395, 137–160.

Moelker, A., Pattynama, P.M., 2003. Acoustic noise concerns in functional magnetic resonance imaging. Hum. Brain Mapp. 20, 123–141.

Muller, C.M., Leppelsack, H.J., 1985. Feature extraction and tonotopic organization in the avian auditory forebrain. Exp. Brain Res. 59, 587–599.

Ogawa, S., Lee, T.M., Kay, A.R., Tank, D.W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc. Natl. Acad. Sci. U. S. A. 87, 9868–9872.

Portas, C.M., Krakow, K., Allen, P., Josephs, O., Armony, J.L., Frith, C.D., 2000. Auditory processing across the sleep-wake cycle: simultaneous EEG and fMRI monitoring in humans. Neuron 28, 991–999.

Reiner, A., Perkel, D.J., Bruce, L.L., Butler, A.B., Csillag, A., Kuenzel, W., Medina, L., Paxinos, G., Shimizu, T., Striedter, G., Wild, M., Ball, G.F., Durand, S., Guturkun, O., Lee, D.W., Mello, C.V., Powers, A., White, S.A., Hough, G., Kubikova, L., Smulders, T.V., Wada, K., Dugas-Ford, J., Husband, S., Yamamoto, K., Yu, J., Siang, C., Jarvis, E.D., Avian Brain Nomenclature Forum, 2004. Revised nomenclature for avian telencephalon and some related brainstem nuclei. J. Comp. Neurol. 473, 377–414.

Schmidt, M.F., Konishi, M., 1998. Gating of auditory responses in the vocal control system of awake songbirds. Nat. Neurosci. 1, 513–518.

Seifritz, E., Di Salle, F., Bilecen, D., Radu, E.W., Scheffler, K., 2001. Auditory system: functional magnetic resonance imaging. Neuroimaging Clin. N. Am. 11, 275–296.

Seifritz, E., Esposito, F., Hennel, F., Mustovic, H., Neuhoff, J.G., Bilecen, D., Tedeschi, G., Scheffler, K., Di Salle, F., 2002. Spatiotemporal pattern of neural processing in the human auditory cortex. Science 297, 1706–1708.

Seifritz, E., Di Salle, F., Esposito, F., Bilecen, D., Neuhoff, J.G., Scheffler, K., 2003. Sustained blood oxygenation and volume response to repetition rate-modulated sound in human auditory cortex. NeuroImage 20, 1365–1370.

Stripling, R., Volman, S.F., Clayton, D.F., 1997. Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. J. Neurosci. 17, 3883–3893.

Suga, N., Ma, X.F., Gao, E.Q., Sakai, M., Chawdhury, S.A., 2003. Descending system and plasticity for auditory signaling processing: neuroethological data for speech scientists. Speech Commun. 41, 189–200.

Thiel, C.M., Friston, K.J., Dolan, R.J., 2002. Cholinergic modulation of experience-dependent plasticity in human auditory cortex. Neuron 35, 567–574.

Van der Linden, A., Verhoye, M., Van Auderkerke, J., Peeters, R., Eens, M., Newman, S.W., Smulders, T., Balthazart, J., DeVoogd, T.J., 1998. Non invasive in vivo anatomical studies of the oscine brain by high resolution MRI microscopy. J. Neurosci. Methods 81, 45–52.

Van der Linden, A., Verhoye, M., Van Meir, V., Tindemans, I., Eens, M., Absil, P., Balthazart, J., 2002. In vivo manganese-enhanced magnetic resonance imaging reveals connections and functional properties of the songbird vocal control system. Neuroscience 112, 467–474.

Van Meir, V., Verhoye, M., Absil, P., Eens, M., Balthazart, J., Van der Linden, A., 2004. Differential effects of testosterone on neuronal populations and their connections in a sensorimotor brain nucleus controlling song production in songbirds: a manganese enhanced-magnetic resonance imaging study. NeuroImage 21, 914–923.

Verhoye, M., Van der Linden, A., Van Audekerke, J., Sijbers, J., Eens, M., Balthazart, J., 1998. Imaging birds in a bird cage: in-vivo FSE 3D MRI of bird brain. MAGMA 6, 22–27.

Vignal, C., Attia, J., Mathevon, N., Beauchaud, M., 2004. Background noise does not modify song-induced genic activation in the bird brain. Behav. Brain Res. 153, 241–248.

Wilbrecht, L., Nottebohm, F., 2003. Vocal learning in birds and humans. Ment. Retard. Dev. Disabil. Res. Rev. 9, 135–148.