Title

Functional magnetic resonance imaging in zebra finch discerns the neural substrate involved in segregation of conspecific song from background noise.

Abbreviated title

fMRI in songbirds exposed to degraded song stimuli

Authors and author addresses

Tiny Boumans1*, Clémentine Vignal2,3, Alain Smolders4, Jan Sijbers4, Marleen Verhoye1,4, Johan Van Audekerke1, Nicolas Mathevon3, Annemie Van der Linden1

1. Bio-Imaging Lab, University of Antwerp, Belgium
2. TSI CNRS UMR 5516, Université Jean Monnet, Saint-Etienne, France
3. NAMC CNRS UMR 8620, Université Paris XI, Orsay & ENES EA3988, Université Jean Monnet, Saint-Etienne, France
4. Vision Lab, University of Antwerp, Belgium

* Corresponding author
Tiny Boumans
University of Antwerp
Bio-Imaging Lab
Groenenborgerlaan 171
B-2020 Antwerp
Tel.: ++ 32 (0)3 2653230
Fax: ++ 32 (0)3 2653233
Email: Tiny.Boumans@ua.ac.be

Includes: 5 figures and 1 table
Includes: 35 pages (including this one)
Includes: 0 equations

Total number of words in:

<table>
<thead>
<tr>
<th>Section</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>whole manuscript</td>
<td>7437</td>
</tr>
<tr>
<td>abstract</td>
<td>249</td>
</tr>
<tr>
<td>introduction</td>
<td>418</td>
</tr>
</tbody>
</table>
Abstract

Recently, fMRI was introduced in a well-documented animal model for vocal learning, the songbird. Using fMRI and conspecific signals mixed with different levels of broadband noise, we now demonstrate auditory induced activation representing discriminatory properties of auditory forebrain regions in anesthetized male zebra finches (*Taeniopygia guttata*). Preceding behavioral tests showed comparable calling responses to the original conspecific song stimulus heard outside and inside the magnet. A significant fMRI response was elicited by conspecific song in the primary auditory thalamo-recipient subfield L2a, in neighboring sub areas L2b, L3 and L, and in the rostral part of the higher order auditory area NCM. Temporal BOLD response clustering revealed rostral and caudal clusters that we defined as ‘cluster Field L’ and ‘cluster NCM’ respectively. However, since the actual border between caudal Field L subregions and NCM cannot be seen in the structural MR image and is not precisely reported elsewhere, the cluster NCM might also contain subregion L and the medial extremes of the subregions L2b and L3. Our results show that whereas in cluster Field L the response was not reduced by added noise, in cluster NCM the response was reduced and finally disappeared with increasing levels of noise added to the song stimulus. The activation in cluster NCM was only significant for two experimental stimuli that showed significantly more behavioral responses than the more degraded stimuli, suggesting that the first area within the auditory system where the ability to discern song from masking noise emerges is located in cluster NCM.
Introduction

Background noise can be an obstacle to the successful perception of significant information in acoustic signals. In songbirds (Passeriformes: Oscines) the acoustic signals that contain significant auditory information are the songs and calls, which are learned from an adult male by vocal imitation for the use in particularly individual recognition, mate attraction and territorial defense (Nowicki and Searcy, 2004). The communicability of auditory signals provided by bird vocalizations is dependent on many factors present in the natural environment, not least of which are wind noise and foliage density. The capacity to segregate auditory signals in unfavorable auditory environments requires peripheral filtering at the level of the cochlea (Evans, 1992), but auditory recognition and memorization are the product of more central structures, such as the caudomedial nidopallium (NCM), a higher order auditory region of the telencephalon (Bolhuis et al., 2001; Mello and Clayton, 1994; Stripling et al., 2001; Terpstra et al., 2004). Neurophysiologic recordings and immediately early genes (IEG) expression have shown that NCM is responsive to the playback of conspecific songs including the bird’s own song (BOS) and tutor song (Mello and Clayton, 1994; Mello and Ribeiro, 1998; Stripling et al., 2001; Terpstra et al., 2004; Velho et al., 2005) and exhibits stimulus dependent adaptation, which could serve as a mechanism for the memory of familiar conspecific songs (Chew et al., 1996; Mello et al., 1995; Stripling et al., 1997). Recently, stimulus-specific processing in the auditory region NCM was revealed by functional magnetic resonance imaging (fMRI) applied in anesthetized intact songbirds (Van Meir et al., 2005). Functional MRI that relies on blood oxygen level dependent (BOLD) contrast (Ogawa et al., 1990) is one of the commonly used techniques for imaging brain activity in humans. This technique allows the in-vivo non-invasive investigation of local hemodynamic changes during neural activation induced by various simple
but also complex tasks such as auditory scene analysis. Using fMRI in the zebra finch, we
wanted to investigate the effect of background noise on song-induced activation in auditory
regions of the telencephalon of songbirds. We investigated activation in the auditory forebrain of
anesthetized adult male zebra finches (*Taeniopygia guttata*) using fMRI during playback of a
conspecific signal mixed with different levels of broadband noise, and tested the behavioral
response of zebra finches exposed to the original intact conspecific song stimulus recorded
within the magnet, to provide comparable stimuli as perceived by the birds within the MR
scanner. Successful recognition of conspecific song in a noisy background was reflected in the
fMRI response of rostral NCM.
Materials and Methods

Experimental subjects

Seven adult male zebra finches (*Taeniopygia guttata*, 12-20 g b.w.) served as subjects for this experiment. The birds were obtained from local suppliers and were kept for about 5 months in the lab aviaries with food and water ad libitum, temperature between 20°C and 25°C, and natural light/dark rhythm. Since the birds came in August (summer time) and experiments were done in January (winter time), the natural light/dark rhythm ranged from 15 to 8 hours of daylight. Four birds underwent fMRI measurements and three birds were used for behavioral tests on stimulus recognition. Experimental procedures were in agreement with the Belgian laws on the protection and welfare of animals and had been approved by the ethical committee of the University of Antwerp (Belgium).

Motion control

Immobilization of the animal’s head is critical to allow accurate fMRI measurements. By using anesthesia and a robust stereotaxic device, motion was reduced to a minimum. Zebra finches were initially anesthetized with an intramuscular injection in the pectoral muscles of 25 mg/kg ketamine (Ketalar, 50mg/ml, Parke-Davis, Belgium) and 2 mg/kg medetomidine (Domitor, 1mg/ml, Orion Pharma, Finland). After 30 minutes, medetomidine was continuously infused at a rate of 0.02 ml/h through a catheter positioned in the chest muscle. This allowed the birds to be steadily anesthetized for a minimum of 8 hours. The anesthetized zebra finches were immobilized in a non-magnetic, custom made head holder composed of a beak mask and a circular radio-frequency (RF) surface antenna (diameter 15 mm) tightly placed around the bird’s head above both ears and
eyes. This allowed accurate and reproducible positioning of the bird within the magnet while at the same time preventing motion. Whole brain and especially the auditory regions of interest were situated in the sensitive region of the RF receiving circular antenna.

**Monitoring physiology**

To maintain an optimal and stable physiological condition during functional brain research, body temperature and respiration were continuously monitored. Body temperature was monitored with a cloacal temperature probe (SA-Instruments, Inc., New York, USA) and maintained at 40 °C (range: 39.3 to 41.7 °C) by a cotton jacket and a water-heated pad connected to an adjustable heating pump (Neslab Instruments, ex111, Newington, CT, USA). Respiration rate and amplitude were monitored with a small pneumatic sensor (SA-Instruments, Inc., New York, USA) positioned under the bird. The respiration rate could not be standardized and showed substantial variation between birds (range: 60 to 158 cycles/min) but rather small variation within birds (average range is 33 cycles/min). The stability of the expired pCO₂ was monitored by a small tube fixed to the stereotaxic mask and connected to a CO₂ analyzer (Capstar-100, CWI Inc., Diss Norfolk, UK). The pCO₂ fluctuations measured during the experiments were almost non-existent.

**Auditory stimulation**

*Experimental stimuli*

The auditory stimuli consisting of conspecific song and noise were provided and have been described previously by the group of Mathevon (Vignal et al., 2004). The original signal (conspecific signal: CS) was a sequence of songs and calls recorded in the zebra finch aviary of the respective laboratory. Because our fMRI experiments were performed in zebra finches of our
own aviary, the birds of this study were exposed to unfamiliar song. In the recorded 20 seconds of CS stimulus, 6% were silence and 94% were songs and calls. Three stimuli were built by mixing CS with different levels of a continuous masking noise (white noise: WN) using Syntana software (Aubin, 1994). In each mixed signal, all frequencies of the masking noise had equal energy and ranged from 0 Hz to the maximum CS frequency, i.e. 10 000 Hz. The stimuli had different CS/WN intensity level ratios whereas an equal average sound intensity was maintained. These intensity level ratios were defined as \( E = 20 \log(A_{CS}/A_{WN}) \), where \( E \) represents the emergence level of the CS in dB, \( A_{CS} \) the absolute amplitude of the CS and \( A_{WN} \) the absolute amplitude of WN. The values of \( E \) were -3 dB (stimulus SN-3), -9 dB (stimulus SN-9), and -18 dB (stimulus SN-18) (Figure 1).

Neuroimaging studies in humans reveal larger auditory activation with increasing sound level (Jancke et al., 1998). A common intensity (loudness) analysis measures the power of a signal using the measure Root Mean Square (RMS). Calculating mean RMS-power is well known for its accuracy, and allows the discrimination of sound sections that are perceived as loud or weak. To control for loudness effects, we calculated the mean and maximum RMS-power of the equalized signals that were recorded in the MR scanner with the electret microphone described in the data supplements, and checked them to match the loudness of original signals. The ratios of these mean and maximum values between original and recorded equalized versions were similar for all stimuli, meaning that the equalization and presentation of stimuli with a different noise level didn’t affect the loudness of one stimulus more than another.

To quantify the differences between the signals CS, SN-3, SN-9, SN-18 and WN, the correlations between the amplitude envelope of each stimulus and the CS envelope, and between the frequency
spectrum of each stimulus and the CS spectrum were assessed. Signal emergence over background noise was also assessed by computing the entropy (Shannon and Weaver, 1949). Because the background noise is a constant white noise, a signal lost in the background noise does not significantly modify the distribution of energy over time. On the contrary, a signal that emerges strongly from the background noise modifies the time distribution of energy. To quantify these energy distribution modifications, we measured the standard deviation $SD$ of the envelope of each experimental signal (CS, SN-3, SN-9, SN-18, WN). The entropy $H$ was then calculated according to the method described in Beecher (1988, 1989) : $H = \log \left( \frac{SD_{\text{experimental signal}}}{SD_{CS}} \right)$. To obtain the normalized entropy $H'$ ranging between 0 and 1, $H$ was divided by its maximum value. So, a value of $H'$ near 1 characterizes a signal almost lost in the background noise. The entropy value of each experimental signal, as well as the comparisons of their amplitude envelopes and their frequency spectra to those of the CS, gives a good picture of the different degradations of the original signal obtained in each stimulus (Table I). It appeared that all three mixed stimuli differed greatly from the CS. However, the stimulus SN-18 was much degraded, whereas the stimulus SN-3 conserved the main characteristics of CS.

**Behavioral response to auditory stimuli generated in the fMRI environment**

To test whether the magnetic field based speaker (see data supplements) affects sound perception by the birds, an ethological test was performed on three adult male zebra finches that did not take part in the fMRI study. They were exposed to the playback of three acoustic signals: WN (negative control), CS (positive control), and equalized CS that was recorded inside the MR scanner (test signal). The equalizer and recording settings used to obtain the test signal were the same as described in the data supplements for the determination of frequency response curves.
Each acoustic signal lasted for 5 seconds, and they were presented to the birds in random order, separated by 30 seconds of silence. The emission chain was composed of two high fidelity speakers (Triangle Comete 202) placed at either end of the experimental cage connected to a DAT recorder (Sony DTC-ZE 700) and an amplifier (Yamaha AX-396). During each test only one randomly chosen speaker emitted the playback stimulus (sound level: 60 dB at 1 meter). Each subject was put in the experimental cage (240 x 50 x 50 cm³) equipped with roosts and placed in an acoustic isolated chamber 24 hours prior to the start of the playback procedure (12L/12D photoperiod). During the playback test, the activity of the tested bird was recorded with a video recorder (Sony DCR-TRV33). Male zebra finches respond to the playback of conspecific calls by producing vocalizations from which the distance calls (or long calls (Vicario et al., 2001; Zann, 1996)) are the sounds most frequently emitted. To assess the bird’s response during the playback tests, the number of distance calls emitted during song presentation was determined in the recordings using Syntana software (Aubin, 1994) and Goldwave 4.26.

**Stimulation protocol**

A block-design was used in the fMRI experiment. Auditory stimuli were presented in 6 repeated blocks consisting of 20 seconds stimulation and 60 seconds of rest (no auditory stimulation). Images were collected with a block-design paradigm consisting of 6 cycles of 8 images collected during stimulation, and 24 images collected during rest, resulting in 192 functional images (Figure 2 A). Each experiment, which was preceded by the acquisition of 8 dummy images to allow the signal to reach a steady state, thus took approximately 8.5 minutes. In all birds, five consecutive experiments were performed in random order during which the birds were exposed to one of the
five different stimuli CS, SN-3, SN-9, SN-18 and WN. The average song power (average over an entire song) was set at 70 dB SPL (sound pressure level).

fMRI experiments

Imaging settings

An in-vivo 7T NMR Microscope was used with a console from MR Solutions (Guildford, UK) and a magnet from Magnex Scientific Ltd (Oxfordshire, UK). The horizontal bore of the magnet is 150 mm wide and the actively shielded gradient-insert (Magnex Scientific Ltd, Oxfordshire, UK) has an inner diameter of 100 mm and a maximum gradient strength of 400 mT/m. A Helmholtz (45 mm) antenna served for transmitting the RF pulses and a circular RF surface antenna (15 mm) was used for MR signal reception.

A set of 1 parasagittal, 1 horizontal and 1 coronal gradient-echo (GE) scout image and a set of 12 horizontal GE images were first acquired to determine the position of the brain in the magnet. Functional imaging was performed using a T2*-weighted single-slice GE fast low angle shot (FLASH) sequence (FOV = 25 mm, TE = 14 ms, TR = 40 ms, flip angle = 11°, gradient ramp time = 1000 µs, acquisition matrix = 128x64, reconstruction matrix 128x128, slice thickness = 0.5 mm). Long gradient ramp times (1000 µs instead of 100 µs) reduced the gradient noise to 63 dB. The total acquisition time per image was 2.56 sec and a spatial resolution of 195x195 µm² was obtained. As illustrated on Figure 3, the functional images were acquired on a parasagittal slice in the right hemisphere that was chosen to go through the auditory forebrain regions Field L and NCM. Since the fiber track that defines subregion L2a (Fortune and Margoliash, 1992; Vates et al., 1996) can be clearly seen on a structural MR image, and since NCM begins next to the midline
as a small circular area and becomes gradually larger in more lateral sections up to 1 mm lateral (Mello et al., 1992), the lateral position of our functional slice was chosen to cover a big part of NCM whereby subregion L2a at the rostral side can easily be seen. This was fulfilled for a lateral distance from 0.25 mm to 0.75 mm from the midline. Anatomical high resolution imaging was performed at the same position as the functional imaging slice with a T2-weighted spin-echo (SE) sequence (FOV = 25 mm, TE = 45 ms, TR = 2000 ms, acquisition matrix = 256x128, reconstruction matrix 256x256, slice thickness = 0.5 mm, and 8 averages).

**Image processing**

The fMRI data series were first preprocessed with MEDx software (version 3.41, Sensor Systems Inc, Sterling, KS, USA). The following algorithms were included: (1) motion detection between consecutive images by means of a center of intensity algorithm in three directions, (2) spatial smoothing with a 3x3 pixel Gaussian convolution filter, (3) intensity normalization with a resulting mean image intensity value of 1000. Functional activation maps that display the Z-score of each voxel were calculated by comparison of the 8 images obtained during stimulation and the 24 images obtained during rest with a voxel-level unpaired t-test. The Z-score is a statistical measure that quantifies the signal intensity (SI) difference (measured in standard deviations) between the images acquired with and without stimulation. The pixels with a significant Z-score ($P < 0.05$) were overlaid on the corresponding anatomical images, resulting in these high resolution images showing in color scale the regions with a significant SI difference between the images acquired with and without stimulation (Figure 2 B).

*Cluster analysis and brain structure assignment*
To distinguish adjacent activated regions, significant activated voxels ($P < 0.05$) were clustered by means of the SI time course during successive stimulation and rest periods. First, a time course was generated by averaging the 6 consecutive trials to obtain a 32-dimensional vector space, where each average voxel’s time course is represented as one point. Amplitude normalization of the signals ensured clustering was done on the shape of the time course, and not only on the amplitude that can be subject to physiological and/or imaging trends. After determination of the significantly activated voxels (t-test, $P < 0.05$), isolated voxels were removed. Finally, a reduction of the dimensionality of the feature space was obtained by applying principal components analysis (PCA), i.e. by projecting data points onto the most relevant components. Structure in the measurement data was detected by Fuzzy C-Means (FCM), a clustering technique which looks for similarities in the fMRI data feature space (Bezdek et al., 1984; Fadili et al., 2000) based on the Euclidean distance of the individual data points in the cluster space. As a result, a spatial map with two detected clusters and their corresponding time courses was obtained. Cluster analysis was exclusively done for the fMRI data series acquired during exposure to the stimuli CS and WN (Figure 2 C), since primary and secondary auditory processing regions could be discriminated by comparing undegraded biologically relevant signals (CS) with signals that did not contain any relevant information (WN).

The positions of the two clustered regions were compared relative to the location of various landmarks that are visible in the anatomical images. Most pronounced is the darker ellipsoid region that corresponds to the dense fiber track that defines subregion L2a (Fortune and Margoliash, 1992; Vates et al., 1996) (Figure 3). To verify the position of the two clusters of activated pixels, we delineated on the anatomical images the darker region (i.e. subfield L2a) and the border with the cerebellum, and pasted these marks on the clustering maps calculated for the
fMRI data series CS and WN (Figure 2 C). Given the wider spread of activity in the rostral-caudal dimension of all measured WN data series compared to the reported size of L2a in the rostral-caudal dimension in parasagittal slices (Fortune and Margoliash, 1992; Vates et al., 1996) and to the darker ellipsoid region discerned in the anatomical MR images, we conclude that the BOLD activation upon white noise also extends to the neighboring subregions L2b, L3 and subarea L within Field L (see Figure 3). Therefore all subsequent reportings of ‘cluster Field L’ in this paper include these different Field L subregions. If cluster analysis discriminated this cluster Field L from a more caudal region which was exclusively activated upon CS stimulation, we concluded that there was a second pole of BOLD activation that, based on its anatomical location, could be the secondary auditory region NCM. However, based on observations reported in more detailed histological studies (Fortune and Margoliash, 1992), this more caudal area defined here as ‘cluster NCM’ seems to correspond mostly to rostral NCM, and to largely overlap with the subregion L as defined in Fortune and Margoliash (1992). Also the very medial extremes of the subregions L2b and L3 might have been included (see Figure 3). The correspondence between the activated areas and the anatomical landmarks was established in each bird separately.

Regional analysis

SI changes and Z-scores were subsequently analyzed in defined regions of interest (ROIs). Regional analysis was performed for ROIs at the level of the primary telencephalic auditory region, Field L, and the secondary telencephalic auditory region, NCM. On the basis of the clustering maps with marks at the location of subfield L2a and cerebellum, ROIs were chosen to consist of 6 connected pixels each [i.e. $6 \cdot (195 \, \mu m)^2$] that could be placed centrally in cluster Field L and cluster NCM with a gap of minimum one row pixels between both ROIs. The ROI in
Field L was selected to overlap with the darker region representing subfield L2a, and with the region activated by WN that was discriminated from other activated regions by the CS. The ROI in NCM was located at the center of the caudal extension activated by CS, and rostral from the cerebellum. In all the fMRI data series we calculated for both ROIs the mean SI for the 192 time points and the Z-score. Figure 2D illustrates the % BOLD SI changes during the stimulation periods of 2 successive cycles in the ROIs Field L and NCM of one representative bird that was exposed to the original signal CS and WN. These percentage values were calculated relative to the mean SI of the last 16 time points of the rest period in the respective stimulation/rest block. Because discriminatory properties of auditory regions are better characterized by different levels of auditory activation, further analysis started with the Z-scores.

Previously we showed that Field L shows a positive BOLD response to presentation of several different kinds of stimulus (Van Meir et al., 2005). Therefore, we controlled for this positive response in cluster Field L during each stimulation period preventing possible underestimations of the NCM response. If the mean % BOLD SI change in ROI Field L during a stimulation period was negative, this respective stimulation/rest block was excluded from further data analysis and the Z-score was recalculated for the remaining time points. The occurrence of the negative mean % BOLD SI change in Field L seemed to be minimal. From a total of 120 stimulation/rest blocks (6 blocks x 5 stimuli x 4 birds), we excluded only seven blocks spread among the 4 different birds. Also the stimulus and number of period repetitions seemed to have no causal connection with the occurrence of the negative Field L response. We excluded two blocks for stimulus SN-18, two blocks for stimulus SN-9 and 3 blocks for WN.

Statistical analysis
Statistical analysis of the Z-scores was performed with SPSS (Statistical Package for Social Sciences, Chicago, IL, USA). Differences in Z-scores were statistically analyzed using an analysis of variance (ANOVA, $P = 0.05$) for repeated measures with the independent factors being the brain region (Field L and NCM) and stimulus type (CS, SN-3, SN-9, SN-18, WN). Correlations between Z-scores and auditory stimuli were statistically analyzed using a linear regression analysis ($P = 0.05$) with the dependent variables the Z-scores in cluster Field L and cluster NCM, and the independent factors the signal degradation values (the entropy value of each experimental signal, and the comparisons of their amplitude envelopes and their frequency spectra to those of the CS). All data are presented in the corresponding figures as means with standard errors (S.E.M.).
Results

The principal goal of this study was to investigate the effect of background noise on song-induced activation in auditory regions of the telencephalon of songbirds. For this purpose, we used recently developed BOLD imaging techniques for small animals and songbirds in particular. Our study is one of the first quantitative assessments of the distribution of the auditory evoked BOLD response in the zebra finch forebrain.

Behavioral response to auditory stimuli generated in the fMRI environment

Each tested bird answered to the original CS as well as to the equalized version recorded in the magnet by emitting distance calls [(mean number of calls ± SD) = (6 ± 2) and (5 ± 1), respectively], whereas WN did not provoke any vocal response. During WN as well as during the first 40 seconds of silence in the behavioral test, all birds remained completely silent. Thus, the acoustic stimulus significantly influences the number of distance calls emitted by the birds (Friedman ANOVA, $\chi^2(3,3) = 8.111, P < 0.0438$), indicating stimulated calling behavior rather than just spontaneous activity. Moreover, the number of distance calls emitted by the bird in response to the original CS and to the equalized version recorded in the magnet did not differ significantly (Wilcoxon matched pair test, $Z = 1.069, P = 0.285$). The fact that the recorded equalized song was always recognized reveals that the sounds generated by the magnetless dynamic speakers preserve all relevant information to maintain song recognition by the bird.

fMRI experiments

The mean Z-score ($n = 4$) for ROIs in cluster Field L and cluster NCM to presentation of CS, SN-3, SN-9, SN-18 and WN is displayed in Figure 4. One dataset from an individual bird with
stimulation SN-9 was excluded from data analysis as a result of imaging artifacts that caused considerable phasic SI changes that were not correlated with the functional stimulation protocol.

We observed a significant ($P < 0.05$; $Z > 1.65$) Z-score in cluster Field L in response to the presentation of each acoustic stimulus (CS, SN-3, SN-9, SN-18, WN). The Z-score in cluster NCM varied gradually among the degraded stimuli. We observed a significant ($P < 0.05$; $Z > 1.65$) Z-score in this cluster to presentation of the stimuli CS, SN-3 and SN-9, but not to the stimuli SN-18 and WN. Repeated measures two-way ANOVA ($P = 0.05$) with the Z-score of the four birds as dependent variables and the brain region and stimulus as independent factors demonstrated no significant effect for acoustic stimulus ($P = 0.412$, $F(4,16) = 1.052$, $R^2 = 0.263$), but indeed a significant effect for brain region ($P = 0.022$, $F(1,4) = 13.341$, $R^2 = 0.074$) indicating that the two areas responded differentially to one or more stimuli. Nevertheless, no significant interaction between acoustic stimulus and brain region was observed ($P = 0.622$, $F(4,16) = 0.670$, $R^2 = 0.168$). One-way ANOVA ($P = 0.05$) with the stimulus type as independent factor revealed no significant effect for stimulus for both cluster Field L ($P = 0.84$, $F(4,14) = 0.35$, $R^2 = 0.09$, $\omega^2 = 0$) and cluster NCM ($P = 0.21$, $F(4,14) = 1.66$, $R^2 = 0.32$, $\omega^2 = 0.12$). However, because $\omega^2_{(NCM)} > 0$ and because we observed a clear linear trend in the NCM Z-scores as a function of the signal degradation, we explored the use of linear regression on non-nominal (scale) stimulus parameters. A linear regression analysis ($P = 0.05$) with the Z-scores of the four birds as dependent variables and the three quantitative signal degradation values of Table I as independent factors also revealed a different auditory response in cluster Field L and cluster NCM. We observed no significant correlation in Field L, but indeed a significant correlation in NCM between the Z-score and the signal degradation values representing comparisons between amplitude envelopes ($R^2 = 0.297$, $P = 0.016$), frequency spectra ($R^2 = 0.240$, $P = 0.033$) and entropy ($R^2 = 0.238$, $P = 0.034$).
These results demonstrate that the auditory response in Field L subregions, at least at the neuronal population level, is not influenced by the CS/WN ratio of the stimulus, whereas the response in more caudal regions NCM and potentially subregions L2b, L3 and L decreases with more degradation of the CS. The behavioral data also show that the recorded undegraded song was recognized, meaning that the sounds generated within the magnet preserve all relevant information to maintain song recognition by the bird. Moreover, the activation in cluster NCM was only significant for the two experimental stimuli SN-3 and SN-9 that showed significantly more behavioral responses than the more degraded stimuli. This means that only those two conditions are recognized as conspecific song and that the first area within the auditory system where the ability to discern song from masking noise emerges is located in cluster NCM.
Discussion

In the present paper, we used BOLD fMRI to investigate the effect of background noise on song-induced activation in the auditory forebrain regions of a songbird, the zebra finch (12-20 g b.w.). Auditory information in the avian brain travels from the cochlear nuclei through the midbrain to the thalamic nucleus ovoidalis (Ov) and from there to the telencephalic Field L. The main ovoidalis thalamo-recipient zone is subfield L2a. The subregions L1 and L3 are immediately adjacent to L2a, and receive L2a input as well as a smaller amount of thalamic input from the ovoidalis shell region. The presence of subfields is based on differences in cytoarchitecture and connectivity (Fortune and Margoliash, 1992; Vates et al., 1996). Field L projects to secondary auditory areas in the telencephalon, including the NCM. Besides these auditory inputs into NCM from medial subfield L2a and subfield L3, NCM also receives input from the secondary caudal medial mesopallium (CMM) and from Ov (Vates et al., 1996). The telencephalic regions Field L and NCM are part of the avian analogue of the mammalian auditory cortex.

The impact of anesthesia

Most MRI studies in animals are conducted under anesthesia to minimize motion artifacts in the imaging. Previous recordings have shown that auditory responsiveness in the forebrain song system nucleus HVC is affected by changes in the bird’s behavioral state (Cardin and Schmidt, 2003; Rauske et al., 2003; Schmidt and Konishi, 1998). In awake behaving zebra finches, HVC auditory responses are largely suppressed and less selective than responses in anesthetized and sleeping birds. These studies did not show similar effects in the auditory area Field L. Medetomidine, a nonnarcotic sedative and analgesic, is a potent α2-adrenoreceptor agonist that produces sedation and analgesia. It has been shown that noradrenergic terminals are found
throughout the avian auditory and vocal system (Mello et al., 1998) and that α-adrenergic receptor blockade abolishes song-induced ZENK induction in zebra finch NCM (Ribeiro and Mello, 2000). In medetomidine anesthetized starlings, however, stimulus specific fMRI responses were previously revealed in the NCM, but this drug was given at a lower dose (Van Meir et al., 2005) than the one applied in the current study on the zebra finch. Further studies are required to determine the effects that anesthetic agents may have on auditory processing, as compared to fully awake birds.

**Comparison of data obtained with fMRI, electrophysiology and gene expression**

Electrophysiological studies in zebra finches have shown that the selectivity for song-like sounds increases in a hierarchical manner along ascending processing stages in the auditory system (Hsu et al., 2004). Woolley et al. (2005) showed that the discrimination of conspecific vocalizations from other sounds results from tuning properties to temporal modulations that differ most across sounds. Most Field L auditory neurons are more selective to either conspecific song or white noise with a much smaller number of neurons showing weak or no selectivity (Grace et al., 2003). These electrophysiological studies suggest that complex natural sounds, such as conspecific song, are preferentially represented in the neural activity of the auditory forebrain relative to other background sounds that are commonly present in the bird’s environment.

Analysis of immediate early gene (IEG) expression has been very useful in generating high-resolution maps of brain activation associated with perceptual and motor aspects of vocal communication in songbirds (Mello, 2002; Mello et al., 2004; Mello, 2004; Ribeiro and Mello, 2000; Velho et al., 2005). By contrast to the electrophysiological results, IEG induction is absent in subfield L2 while robust in the NCM following song presentation (Mello et al., 1992; Mello and
Clayton, 1994; Velho et al., 2005; Gobes and Bolhuis, 2007; Mello et al., 2004; Theunissen and Shaevitz, 2006). Vignal et al. (2004) investigated behavioral responses and IEG expression to the same stimuli (SN-3, SN-9, SN-27 and WN) used in this fMRI study (with the exception of SN-27). The stimuli SN-3 and SN-9 that elicited significant behavioral responses and gene activation in the NCM also elicited a significant fMRI response in the region that appears to comprise at least the rostral part of NCM. These results suggest that successful recognition of relevant information in noise may be reflected in the fMRI response of cluster NCM. In addition, fMRI allowed us to quantify the activity modulations in Field L subregions (see below), a goal which cannot be achieved by IEG analysis because of the absence of ZENK expression in subfield L2.

**Discriminatory properties of Field L and NCM**

Vocal communication in songbirds involves the recognition of individuals based on their vocal performance and segregation of these vocalizations in a noisy environment. A behavioral approach in adult male canaries demonstrated that the accuracy of the discrimination between two conspecific song segments progressively declines as a function of the number of masking distractors (Appeltants et al., 2005). Noisy signals have to contain sufficient information to allow successful recognition against noise (Vignal et al., 2004). Several lines of evidence indicate that the regions NCM and CMM mediate auditory processing, with a more specific role for the NCM in facilitating recognition of species-specific song (Mello et al., 2004; Bailey et al., 2002; Terpstra et al., 2004; Gentner et al., 2004; Gobes and Bolhuis, 2007). The acoustical features underlying song recognition and discrimination in birds are not well understood, but most suggestions rely on the difference in spectral and/or temporal modulations of sounds.
For all auditory stimuli here presented, cluster Field L showed a substantially larger Z-score than cluster NCM, indicating a larger SI difference in Field L between stimulation and rest periods. At the neuronal population level, Field L was shown to be responsive to any auditory input with no significant response modulation between the different degraded stimuli. On the other hand, the fMRI response in cluster NCM varied gradually among the degraded stimuli and showed a significant regression with degradation of the conspecific signal. The lack of a significant effect in the ANOVA might have been due to the small sample size. Nevertheless, the significant regression found in cluster NCM implies that in the songbird the capacity to segregate meaningful auditory signals in unfavorable auditory environments may be an emergent property of NCM and potentially subregions L2b, L3 and L. In combination with the findings of behavioral studies (Appeltants et al., 2005; Vignal et al., 2004), our results suggest that successful recognition of relevant information against noise is reflected in the auditory NCM response. Moreover, data from Van Meir et al. (2005) showed that BOLD responses in Field L and NCM were significantly different with white noise stimuli. Together with our significant correlation result in cluster NCM (Table 1), we suggest that the first area within the auditory system where the ability to discern song from masking noise emerges is located in cluster NCM.

Electrophysiology in canary brain revealed that Field L and NCM differ in their electrophysiological response properties to pure tone stimuli, suggesting differential roles in auditory processing (Terleph et al., 2006). NCM properties, in particular, may allow for response integration across multiple spectrally varying stimulus elements, such as those that occur during birdsong. Particularly, information from multiple Field L2 sites (each tuned to a narrow frequency range) may converge onto a single site in NCM. Unlike electrophysiology, fMRI reveals neural activity by detecting hemodynamic changes induced by groups of cells. It is important to note,
however, that the measured regional heterogeneity between clusters Field L and NCM may be dependent on the resolution of the fMR image. Our image voxels of 195x195x500 µm³ correspond to hundreds of cells or more, and at this resolution it may be impractical to distinguish between the separate subfields in Field L. Because cell groups, belonging either to different subregions or to one subregion, may express substantially different response properties to conspecific song, the capacities ascribed to Field L may cover several such cell populations. Likewise, the potential differences between regions may reflect quantitative shifts in the proportions of cells coding separate stimulus features rather than qualitative differences in the information coded in Field L (at large) and NCM.
Abbreviations

BOLD, blood oxygen level dependent; CMM, caudomedial mesopallium; CS, conspecific song; fMRI, functional magnetic resonance imaging; FOV, field of view; GE, gradient-echo; IEG, immediate early gene; NCM, caudomedial nidopallium; RF, radio-frequency; RMS, root mean square; ROI, region of interest; SE, spin-echo; SI, signal intensity; SPL, sound pressure level; TE, echo time; TR, repetition time; WN, white noise

Acknowledgments

We are grateful to Joris Dirckx who helped with the construction of the magnetless dynamic speaker, and to Martin Wild, Frederic Theunissen and Jacques Balthazart who provided useful comments on the manuscript.

Grants

The fMRI experiments were supported by grants from the Research Foundation - Flanders (FWO, project Nr G.0420.02) and by Concerted Research Actions (GOA funding) from the University of Antwerp to A.V.d.L. T.B. is research assistant of the Foundation for Scientific Research in Flanders (FWO–Flanders, Belgium). The behavioral experiments were funded by a grant from the French National Research Agency (ANR, project ‘Birds’ voices’). NM is supported by the Institut universitaire de France (IUF).
References


**Figures**

**Figure 1. Experimental stimuli.** Oscillograms (above) and spectrograms (below) of the different auditory stimuli: CS (conspecific signal), SN-3 (signal-to-noise ratio = -3 dB), SN-9 (signal-to-noise ratio = -9 dB), SN-18 (signal-to-noise ratio = -18 dB), and WN (white noise).

**Figure 2. Data acquisition and analysis.** (A) Schematical representation of the auditory stimulation design. The entire paradigm was repeated 5 times with alternate presentation of one of the five different stimuli CS, SN-3, SN-9, SN-18, WN. (B) Z-score map illustrating in color the localization of significant SI changes during auditory stimulation (CS in this figure). Voxels display a significant activation ($P < 0.05$) when $Z > 1.65$. (C) Cluster analysis on temporal BOLD responses of these significant activated voxels identifies two distinct clusters with a different SI time course during successive stimulation and rest periods. The brown region in the figure overlaps with the region activated by WN and was defined in this study as ‘cluster Field L’ comprising subregions L2a, L2b, L3 and L while the region indicated in yellow represents the more caudal extension of the auditory area only activated by complex stimuli and defined as ‘cluster NCM’ comprising the most rostral part of the higher auditory NCM but perhaps also subregions L2b, L3 and/or L. Since the actual border between caudal Field L subregions and NCM cannot be seen in the structural MR image and is not precisely reported in more detailed histological studies, the subregions L2b, L3 and/or L potentially belong to cluster NCM instead of cluster Field L. (D) % BOLD SI changes for a total of 2 stimulation periods in the ROIs in cluster Field L and cluster NCM (6 pixels each) of a bird exposed to the control stimuli CS and WN.
**Figure 3. Visualization of Field L2a on high resolution T2-weighted SE images.** The figure displays how the subfield L2a in the study of Vates et al. (1996) compares to the darker ellipsoid region of our anatomical high resolution MR images that corresponds to the dense fiber track that defines subregion L2a. (Schematic illustration adapted from Vates et al., 1996).

**ABBREVIATIONS**, Ch. O. = Optic Chiasm; CMM = caudomedial mesopallium; DLM = medial nucleus of the dorsolateral thalamus; FPL = lateral forebrain bundle; L2a, L2b, L3 = subregions of Field L; NCM = caudomedial nidopallium; Ov = nucleus ovoidalis; tOM = tractus occipitomesencephalicus; X = area X.

**Figure 4. Auditory fMRI response in cluster Field L and cluster NCM.** Statistical significant Z-score calculated in the clusters Field L and NCM for the stimuli CS (conspecific signal), SN-3 (signal-to-noise ratio = -3 dB), SN-9 (signal-to-noise ratio = -9 dB), SN-18 (signal-to-noise ratio = -18 dB), and WN (white noise). The region displays a significant activation ($P < 0.05$) when $Z > 1.65$. The data are presented as means ($n = 4$) with corresponding standard errors (S.E.M.).

**Supplemental Figure I. Magnetless dynamic speaker and frequency response curve.** (A) Auditory stimuli were applied using a magnetless dynamic speaker system integrated in the RF transmitting antenna. When placed around the birds head, sound was delivered to both ears from a distance of approximately 1 cm. The orientation of the magnetic field $B_0$ is indicated by the arrows. (B) Frequency response of the magnetless dynamic speaker in the strong magnetic field (7
(Tesla) of the MR scanner without (black line) and after (grey line) convolution of the sounds with an equalizer function.
Tables

<table>
<thead>
<tr>
<th>STIMULUS</th>
<th>CS</th>
<th>SN-3</th>
<th>SN-9</th>
<th>SN-18</th>
<th>WN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation between the amplitude envelope of the stimulus and the amplitude envelope of CS</td>
<td>1.00</td>
<td>0.52</td>
<td>0.20</td>
<td>0.11</td>
<td>0.00</td>
</tr>
<tr>
<td>Correlation between the frequency spectrum of the stimulus and the frequency spectrum of CS</td>
<td>1.00</td>
<td>0.77</td>
<td>0.60</td>
<td>0.53</td>
<td>0.00</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.00</td>
<td>0.81</td>
<td>0.96</td>
<td>0.98</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Table 1. Signal degradation values.** Comparisons showing the degradation of the original CS (conspecific signal) obtained in the different stimuli. The stimulus SN-18 (signal-to-noise ratio = -18 dB) is much degraded, whereas the stimulus SN-3 (signal-to-noise ratio = -3 dB) conserves the main characteristics of the CS. (for more details see materials and methods section)
Data supplements

Sound generation

Auditory stimulation during an fMRI experiment poses several challenges. Sound has to be generated and conducted to the subject within the high static magnetic field of the MRI machine. No distortion of the sound and no disruption of the imaging process can be tolerated. Firstly, we used a dynamic speaker from which the permanent magnet had been removed (as described by Baumgart et al., 1998). This magnetless speaker was placed in the permanent magnetic field, which in turn moved the speaker coil. Secondly, because the stereotaxic device and the antennas for transmitting and receiving RF pulses require most of the space within the narrow bore of the magnet, it was necessary to integrate the speaker in one or both devices. The use of a capacitively coupled Helmholtz RF transmission antenna gave the possibility of fixing the two parts of the dynamic speaker to both sides of the antenna (Figure I). As a result, a space-saving solution was obtained that delivered sound to both ears from a distance of approximately 1 cm. Thirdly, the auditory stimuli must be presented according to the functional ON/OFF stimulation block paradigm during which images were acquired. Stimulus application was controlled by Presentation software (Neurobehavioral Systems, Inc., version 0.55, Albany, CA, USA). On the basis of gating signals, this program made it possible to play specific sound fragments during acquisition of specific images and to develop complicated stimulation paradigms composed of different auditory stimuli presented with variable duration and order.

Frequency response of the magnetless dynamic speaker
The magnetless dynamic speaker was built of small speakers from a low cost headset (HK 150, FNAC, France). We removed its small permanent magnets, and mounted the speaker system in the MR scanner with the plane of the coil windings parallel to the magnetic field lines (Van Meir et al., 2005) (Figure I A). Because we had no experience with the behavior of dynamic capsules in a strong magnetic field, we first determined the frequency response in situ of the magnetless dynamic speaker. To measure the actual sound pressure level (SPL) in the magnet, a small electret microphone was used because no large microphones with ferro-electric materials can be brought into the magnetic field. The electret microphone has a frequency response which is flat within 3 dB between 40 Hz and 20 kHz, and it was first carefully calibrated using a quarter-inch microphone (Brüel & Kjaer, Denmark, model 4165). Leaving all stereotaxic and monitoring devices in place, the electret microphone was positioned perpendicular to the left speaker at a distance of 1 cm, while the other speaker was switched off. Using an audio signal generator (Wavelab 4.0g, Steinberg Media Technologies), pure sine tones were generated with frequencies between 100 Hz and 10 000 Hz. For each of these waveforms, the output of the microphone was visualized by an oscilloscope. SPL (dB) was determined from peak voltage which is defined as 

$$20\log\left(\frac{V_{RMS}}{V_{ref}}\right) + 94\,\text{dB},$$

where $V_{RMS}$ represents the root mean square voltage ($= V_p * 1/\sqrt{2}$) and $V_{ref}$ the reference voltage, which was determined after calibration of the microphone.

Figure I B represents the frequency response of the magnetless dynamic speaker when placed in the strong magnetic field of the MR scanner. The black line shows the original frequency response to the pure sine tones. It shows a less favorable frequency response (SPL) in the lower frequency band: below 3000 Hz the response dropped sharply. Between 100 Hz and 1600 Hz, the response showed bumps of about 5 dB, and a gradual increase with higher frequencies. The grey line displays the frequency response of the same speaker in the magnet when an equalizer function
with a compromising shape as the black line was used to adjust for the drop in SPL. This means that the sine tones were first convolved with an equalizer function (parameters: low shelf = 1500 Hz (gain: +6 dB) and high shelf = 2000 Hz (gain: -6 dB)), after which the SPL in the magnet was measured. This resulted in a decrease of the SPL drop from about 25 dB (black line) to about 15 dB (grey line). Consequently, all auditory stimuli were first convolved with this equalizer function before presentation.
Figure 1. Experimental stimuli. Oscillograms (above) and spectrograms (below) of the different auditory stimuli: CS (conspecific signal), SN-3 (signal-to-noise ratio = -3 dB), SN-9 (signal-to-noise ratio = -9 dB), SN-18 (signal-to-noise ratio = -18 dB), and WN (white noise).

160x101mm (400 x 400 DPI)
Figure 2. Data acquisition and analysis. (A) Schematical representation of the auditory stimulation design. The entire paradigm was repeated 5 times with alternate presentation of one of the five different stimuli CS, SN-3, SN-9, SN-18, WN. (B) Z-score map illustrating in color the localization of significant SI changes during auditory stimulation (CS in this figure). Voxels display a significant activation ($P < 0.05$) when $Z > 1.65$. (C) Cluster analysis on temporal BOLD responses of these significant activated voxels identifies two distinct clusters with a different SI time course during successive stimulation and rest periods. The brown region in the figure overlaps with the region activated by WN and was defined in this study as ‘cluster Field L’ comprising subregions L2a, L2b, L3 and L while the region indicated in yellow represents the more caudal extension of the auditory area only activated by complex stimuli and defined as ‘cluster NCM’ comprising the most rostral part of the higher auditory NCM but perhaps also subregions L2b, L3 and/or L. Since the actual border between caudal Field L subregions and NCM cannot be seen in the structural MR image and is not precisely reported in more detailed histological studies, the subregions L2b, L3 and/or L potentially belong to cluster
NCM instead of cluster Field L. (D) % BOLD SI changes for a total of 2 stimulation periods in the ROIs in cluster Field L and cluster NCM (6 pixels each) of a bird exposed to the control stimuli CS and WN.

160x177mm (400 x 400 DPI)
Figure 3. Visualization of Field L2a on high resolution T2-weighted SE images. The figure displays how the subfield L2a in the study of Vates et al. (1996) compares to the darker ellipsoid region of our anatomical high resolution MR images that corresponds to the dense fiber track that defines subregion L2a. (Schematic illustration adapted from Vates et al., 1996). ABBREVIATIONS, Ch. O. = Optic Chiasm; CMM = caudomedial mesopallium; DLM = medial nucleus of the dorsolateral thalamus; FPL = lateral forebrain bundle; L2a, L2b, L3 = subregions of Field L; NCM = caudomedial nidopallium; Ov = nucleus ovoidalis; tOM = tractus occipitomesencephalicus; X = area X.

160x86mm (400 x 400 DPI)
Figure 4. Auditory fMRI response in cluster Field L and cluster NCM. Statistical significant Z-score calculated in the clusters Field L and NCM for the stimuli CS (conspecific signal), SN-3 (signal-to-noise ratio = -3 dB), SN-9 (signal-to-noise ratio = -9 dB), SN-18 (signal-to-noise ratio = -18 dB), and WN (white noise). The region displays a significant activation ($P < 0.05$) when $Z > 1.65$. The data are presented as means ($n = 4$) with corresponding standard errors (S.E.M.).
Supplemental Figure I. Magnetless dynamic speaker and frequency response curve. (A) Auditory stimuli were applied using a magnetless dynamic speaker system integrated in the RF transmitting antenna. When placed around the birds head, sound was delivered to both ears from a distance of approximately 1 cm. The orientation of the magnetic field $B_0$ is indicated by the arrows. (B) Frequency response of the magnetless dynamic speaker in the strong magnetic field (7 Tesla) of the MR scanner without (black line) and after (grey line) convolution of the sounds with an equalizer function.