



Song Processing in the Zebra Finch Auditory Forebrain Reflects Asymmetric Sensitivity to Temporal and Spectral Structure

Lisbeth Van Ruijssevelt, Stuart D. Washington, Julie Hamaide, Marleen Verhoye, Georgios A. Keliris and Annemie Van der Linden*

Bio-Imaging Lab, Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

OPEN ACCESS

Edited by:

Einat Liebenthal,
Brigham & Women's Hospital/Harvard
Medical School, United States

Reviewed by:

Jon Sakata,
McGill University, Canada
Henning U. Voss,
Weill Cornell Medical College,
United States

*Correspondence:

Annemie Van der Linden
annemie.vanderlinden@
uantwerpen.be

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 17 July 2017

Accepted: 20 September 2017

Published: 05 October 2017

Citation:

Van Ruijssevelt L, Washington SD,
Hamaide J, Verhoye M, Keliris GA and
Van der Linden A (2017) Song
Processing in the Zebra Finch Auditory
Forebrain Reflects Asymmetric
Sensitivity to Temporal and Spectral
Structure. *Front. Neurosci.* 11:549.
doi: 10.3389/fnins.2017.00549

Despite being commonly referenced throughout neuroscientific research on songbirds, reports of hemispheric specialization in the processing of song remain controversial. The notion of such asymmetries in songbirds is further complicated by evidence that both cerebral hemispheres in humans may be specialized for different aspects of speech perception. Some studies suggest that the auditory neural substrates in the left and right hemispheres of humans process temporal and spectral elements within speech sounds, respectively. To determine whether songbirds process their conspecific songs in such a complementary, bilateral manner, we performed functional magnetic resonance imaging (fMRI) on 15 isoflurane anesthetized adult male zebra finches (*Taeniopygia guttata*) while presenting them with (1) non-manipulated, (2) spectrally-filtered (reduced spectral structure), and (3) temporally-filtered (reduced temporal structure) conspecific song. Our results revealed sensitivity of both primary (Field L) and secondary (caudomedial nidopallium, NCM) auditory regions to changes in spectral and temporal structure of song. On the one hand, temporally-filtered song elicited a bilateral decrease in neural responses compared to the other stimulus types. On the other hand, spectrally filtered song elicited significantly greater responses in left Field L and NCM than temporally filtered or non-manipulated song while concurrently reducing the response relative to non-manipulated song in the right auditory forebrain. The latter hemispheric difference in sensitivity to manipulations of spectral structure in song, suggests that there is an asymmetry in spectral and temporal domain processing in the zebra finch auditory forebrain bearing some resemblance to what has been observed in human auditory cortex.

Keywords: hemispheric lateralization, zebra finch, songbird, functional MRI, auditory processing, spectro-temporal

INTRODUCTION

The contributions of the left and right cerebral hemispheres to the processing of human speech remain controversial. The classical view is that the left hemisphere contains a uniquely human closed system (i.e., a speech module) for processing speech and other linguistic stimuli with little to no contribution from the right hemisphere (Liberman and Mattingly, 1989). However, increasing

evidence suggests that the left and right hemispheres provide differential yet complementary contributions to speech perception. One possibility is that the left and right auditory cortices are, respectively, specialized for temporal and spectral domain processing (Zatorre et al., 2002; Poeppel, 2003). Since time and frequency are canonically conjugate variables (Gabor, 1947; Joos, 1948), a necessary consequence of this specialization is that at the upper limit of resolution the refined temporal precision of the left auditory cortex (e.g., the ability to process rapid changes in frequency, such as the formant transitions found in consonants) must come at the expense of spectral precision (Robin et al., 1990; Zatorre and Belin, 2001). Likewise, the refined spectral precision (e.g., the ability to detect variations in prosody and speaker identity) of the right auditory cortex has been demonstrated by these same studies to come at the expense of temporal precision (Robin et al., 1990; Zatorre and Belin, 2001). Such a domain-general perspective implies that animals other than humans may possess comparable hemispheric specializations for processing their conspecific communication sounds. Other than the spectral vs. temporal hypothesis for left hemispheric specialization for speech perception described above, alternative possibilities include functional specialization of somatosensory and motor areas determining left hemispheric specialization for speech perception within auditory processing streams (Liebenthal et al., 2013) and speech-selective acoustic sensitivities in the left hemisphere that are indissociable from the articulatory systems that produce speech sounds (McGettigan and Scott, 2012).

Songbirds are the most common and best-known animal models of human speech, owing to their ability to learn their songs from tutors and the critical periods associated with this vocal learning (Doupe and Kuhl, 1999; Brainard and Doupe, 2002). The songs of certain bird species (e.g., zebra finches) are highly stereotyped and consist of a rich spectro-temporal structure, comparable to human speech sounds. It stands to reason that songbirds may process the temporal and spectral components of their complex conspecific songs in different cerebral hemispheres, as is postulated in humans. However, most studies of hemispheric differences in the processing of conspecific birdsong have used complete, non-manipulated songs rather than spectral and/or temporal song components. Results from these studies are less than consistent, with most of them implicating left hemispheric auditory forebrain nuclei as auditory (memory) centers supporting the acquisition, processing, and recognition of conspecific birdsong (Cynx et al., 1992; Avey et al., 2005; Hauber et al., 2007; Poirier et al., 2009; Moorman et al., 2012). Others show a right hemispheric specialization for conspecific birdsong perception (Voss et al., 2007; Phan and Vicario, 2010), suggest a hemispheric asymmetry in the processing of harmonic structure vs. familiarity of song (Cynx et al., 1992), or report no hemispheric differences in the perception of conspecific song at all (Chew et al., 1996). Showing hemispheric differences in temporal and spectral domain processing in the auditory forebrains of songbirds could shed light on this controversy and demonstrate that the observed asymmetry is a functional analog of the classically reported hemispheric specialization for speech perception in humans.

Here, we performed functional magnetic resonance imaging (fMRI) experiments on 15 adult male zebra finches while presenting them with three stimulus types: (a) non-manipulated, (b) spectrally filtered (spectral information reduced with temporal information retained), and (c) temporally filtered (temporal information reduced with spectral information retained) unfamiliar conspecific song. Voxel-based and region-of-interest (ROI) analyses revealed significantly differential sensitivity to manipulations of spectral and temporal structure of song in left vs. right hemispheric auditory forebrain nuclei.

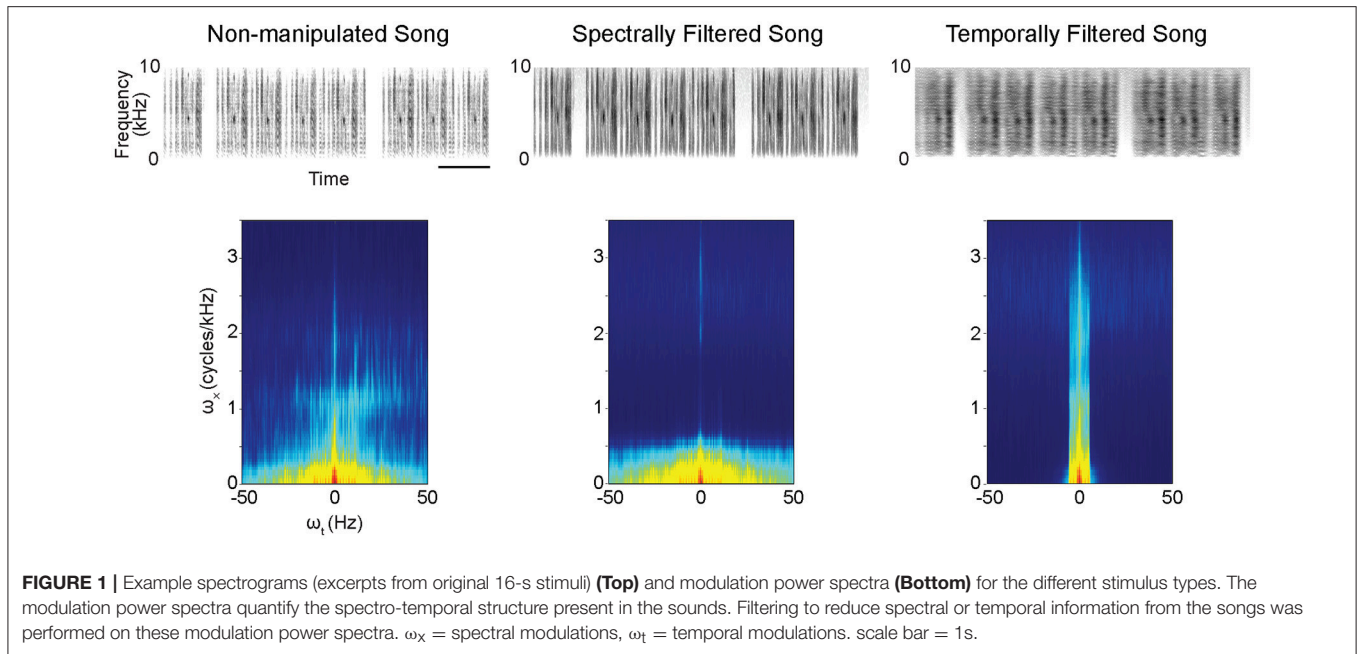
MATERIALS AND METHODS

Subjects

We obtained 15 adult male zebra finches (*Taeniopygia guttata*, >100 days old) from a local breeder for use in this study. We housed the birds together in a large, same-sex aviary maintained under a 12 h light/dark cycle with access to water and food *ad libitum*. The ethical committee of the University of Antwerp approved all experimental procedures (License number: 2016-32), which were in agreement with the Belgian laws on the protection and welfare of animals.

Auditory Stimuli

Conspecific songs were selected from previous recordings of four different male zebra finches. Our subjects were not housed with, nor did they have any contact with, the four birds that produced these songs. From the recordings, we selected and concatenated multiple song bouts to create 16 s long stimuli for each male's song separately. Applying a MATLAB-based algorithm for spectral and temporal filtering (Singh and Theunissen, 2003, <http://theunissen.berkeley.edu/Software.html>) to the song stimuli from the four males, we generated three stimulus types containing songs with varying levels of temporal and spectral structure (**Figure 1**, Figure S1): (a) non-manipulated song consisting of the song bouts as recorded (Audios S1, S4, S7, S10), (b) spectrally filtered song (Audios S2, S5, S8, S11), and (c) temporally filtered song (Audios S3, S6, S9, S12). For both spectral and temporal filtering, a low pass filtering procedure was performed in the space of the modulation spectrum (Singh and Theunissen, 2003; **Figure 1**-bottom) as previously described (Boumans et al., 2007; Elliott and Theunissen, 2009). Initial spectrograms were made with 125-Hz frequency width Gaussian filters. These spectrograms were then log transformed after which a 2D fast Fourier transform was taken to calculate the modulation spectrum (i.e., joint spectral and temporal amplitude spectrum). The applied filtering procedure on these modulation spectra generated songs with equal power and overall frequency power spectra but with a reduction of either the natural temporal modulations or the natural spectral modulations. For spectral filtering, the spectral modulation frequency cut-off was set at 5×10^{-4} cycles/Hz, which means that spectral structure within any 2-kHz band was filtered out (i.e., pitch below 2 kHz is filtered out). For temporal filtering, the temporal modulation frequency cut-off was set at 5 Hz, which means that all amplitude envelope changes, that are faster than 5 Hz were filtered out. This differential filtering method thus did not remove all temporal



and/or spectral information from a song but only reduced it. Note also that we cannot completely exclude the possibility that filtering of spectral features in a song did not influence temporal song features to some extent and vice-versa.

Data Acquisition and Stimulation Protocol

We acquired all of our MRI data on a horizontal 7 Tesla MR system (Pharmascan 70/16 US, Bruker Biospin, Germany) following established protocols (Van Ruijssevelt et al., 2013). In short, fMRI scans consisted of a time series of 298 T_2 -weighted rapid acquisition relaxation-enhanced (RARE) volumes of 15 slices covering the whole brain (slice thickness = 0.75 mm, interslice gap = 0.05 mm) with an in-plane resolution of $(0.25 \times 0.25) \text{ mm}^2$. We acquired a high-resolution anatomical 3D RARE image [resolution = $(70 \times 70 \times 70) \mu\text{m}^3$] for each bird in the same orientation as the fMRI scans to facilitate subsequent spatial registration of images. During the experiment, the animals were continuously anesthetized with 1.2% isoflurane in a mixture of oxygen and nitrogen (at flow rates of 100 and 200 cm^3/min , respectively).

During the fMRI session, the birds were exposed to the three different stimulus types (non-manipulated song, spectrally filtered song, temporally filtered song) derived from a song from one of the four males (either song 1, 2, 3, or 4; respectively Audios S1, S4, S7, S10) as described above. Stimulus types were pseudo-randomly presented in an ON/OFF blocked paradigm that alternated of 16 s of stimulation (ON periods) and 16 s of rest (OFF period). Presentation of the three song types was always grouped together in consecutive ON blocks to ensure an equal number of presentations per stimulus type at the end of the scanning session. The order in which the three types were presented within each consecutive repetition of the group of stimuli was randomized (for example: unmanipulated (UN), spectrally filtered (SF), temporally filtered (TF), | SF, TF, UN, |

UN, TF, SE, | SF, TF, UN | etc.). In total, an fMRI session consisted of 72 ON blocks (24 per stimulus type) and 72 OFF blocks. Two T_2 -weighted RARE images were acquired during each block, resulting in 48 images per stimulus type, per subject. Auditory stimuli were played back at a mean intensity (root-mean-square) of 70 dB through dynamic loudspeakers (Visation, Germany; magnets removed) placed at both sides of the bird's head. To ensure the validity of any observed hemispheric differences, the orientation of the headphones was interchanged between consecutive experiments (subject 1: left speaker \rightarrow right ear; Subject 2: left speaker \rightarrow left ear; Subject 3: left speaker \rightarrow right ear; etc.).

Data Analysis

Our inclusion criteria for fMRI time series in the subsequent analysis were (1) limited head motion (<0.5 mm translation in either of the image's 3 directions), and (2) detection of bilateral positive BOLD responses in the primary auditory region Field L.

We performed both data preprocessing and statistical voxel-based analyses using the Statistical Parametric Mapping toolbox (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>). These analyses were similar to those previously described (Van Ruijssevelt et al., 2013). In short, we realigned and co-registered the fMRI time series to their corresponding 3D RARE. A population based template was generated from the anatomical 3D RARE images of all animals (ANTs; <http://stnava.github.io/ANTs/>). Spatial normalization of all scans to this template enabled between-subject comparisons (Avants et al., 2011). Finally, to achieve in-plane smoothing, we applied a Gaussian kernel of 0.5 mm full width at half maximum (FWHM).

High pass filtering (352-s cutoff period) removed low frequency drifts in the BOLD signal. For the first level analysis, we subsequently modeled for each subject the BOLD responses

as a box-car function convolved with a canonical hemodynamic response function within the framework of the general linear model (GLM) to analyze brain activation differences related to the onset of the different stimuli. We included the six estimated movement parameters derived from the realignment corrections as regressors in the model to account for the residual effect of head motion and restricted the analysis to voxels within the brain by using a whole brain mask. After the estimation of the GLM parameters (β), we calculated different t-contrast images (containing weighted parameter estimates) for different comparisons including non-manipulated song vs. rest, temporally filtered song vs. rest and spectrally filtered song vs. rest as well as comparisons of the activations of these different stimulus types between each other.

Next, to study the effect of song filtering on the activation of auditory regions in the zebra finch forebrain at the group (second) level, we entered the subjects' contrast images (non-manipulated song vs. rest, temporally filtered song vs. rest and spectrally filtered song vs. rest) into a flexible factorial analysis with subjects as random variable. To explore the main effect of stimulus class, we restricted the analysis to voxels that showed a positive BOLD response to any song (t -test all stimuli vs. rest). Within the main effect of stimulus ($p_{\text{uncorrected}} < 0.005$), we performed *post-hoc* t -tests to determine regions sensitive to temporal and/or spectral filtering of song. Since we performed statistical tests on a voxel-by-voxel basis, p -values were adjusted to the number of independent tests performed via Family Wise Error (FWE) correction. As this is a voxel based analysis, results are reported by the highest voxel t -value within each cluster (t_{max}) and the associated voxel p -value. Our threshold for significance was $p_{\text{FWE}} < 0.05$. For visualization, an explorative threshold of $p_{\text{uncorrected}} < 0.005$ was applied.

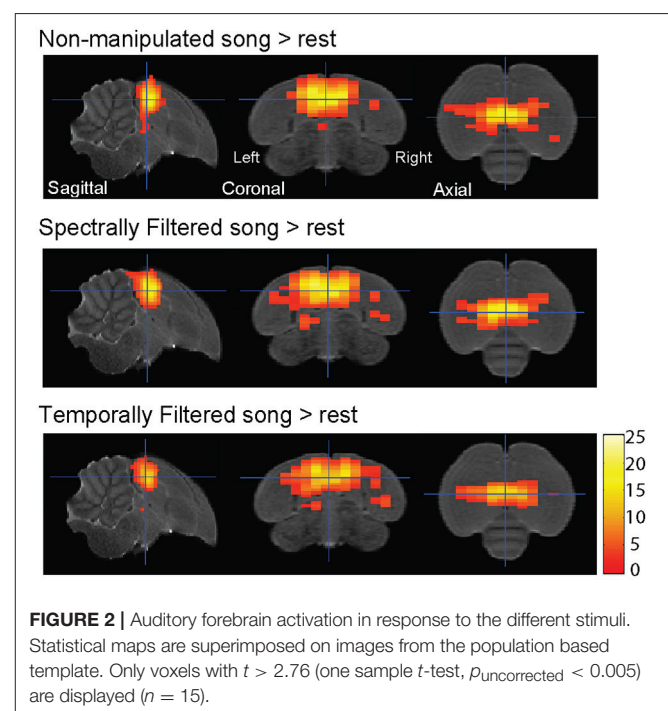
Further, to assess laterality of the observed sensitivity to spectral and temporal filtering in the auditory forebrain, we used the "AveLI" SPM add-on to compute lateralization indices (LIs) for the three song types for left vs. right regions of interest (ROIs) in a threshold-free manner (Matsuo et al., 2012; freely available at <http://aveli.web.fc2.com/>). The ROI for this analysis was determined by the ensemble of voxels (and their mirrored counterparts) in which a significant effect of song type was determined in the voxel-based group analysis (main effect of stimulus class; see results). The computed LI by the AveLI code represents the portion of the signal that is found either left (values toward -1) or right (values toward $+1$). Further, we verified the calculated AveLIs with alternative measures for calculating LI described in literature (Fernández et al., 2001; Jansen et al., 2006; Seghier, 2008), including LIs based on the number of significant voxels as well as on the magnitude of the effect in the left vs. the right hemisphere using either no threshold (Seghier, 2008) or a variable subject specific adjusted threshold (Fernández et al., 2001) to determine significant voxels. We found that the LIs obtained with the different methods correlated well (Figure S2) suggesting that our results were not significantly influenced by the method chosen to calculate LIs. We assessed LIs for different effects found to be significant in the voxel-based analysis: non-manipulated song > temporally filtered song, non-manipulated song > spectrally filtered song, spectrally filtered

song > temporally filtered song, spectrally filtered song > non-manipulated song. These comparisons are further referred to as "filtering effects". The LI indices were derived from t -contrast maps for these different filtering effects at the single subject level. Statistical analysis on the calculated LI-values was performed in JMP[®] (Version 13, SAS Institute Inc., Cary, NC, 1989–2007). First, we assessed whether for each filtering effect, the LI was significantly different from 0 using separate one-sample t -tests. Next, we compared the LIs of the different effects using a linear mixed model with filtering effect as a fixed, and subject as a random, variable. We used unbounded variance components to fit the data using the restricted maximum likelihood method. Tukey's HSD (honest significant difference) test was used for all *post-hoc* tests and results were considered significant for $p < 0.05$. Data are represented as mean \pm standard error of the mean (SEM).

RESULTS

Voxel-based group analysis revealed extensive bilateral activation of the auditory lobule in response to all stimulus types vs. the rest periods (Figure 2).

The statistical map of the main effect of stimulus class (Figure 3A) shows that song filtering (either temporal or spectral) affects the activation of primary (Field L) and secondary (caudomedial nidopallium, NCM) auditory forebrain regions in both hemispheres with the most distinct effect in the left hemisphere. The following *post-hoc* tests clarify the nature of the observed effect (Figures 3B–E). An overview of the supra-threshold clusters for this analysis is given in Table 1. The effect of temporal filtering was most pronounced in right Field L where



A Main effect of stimulus class: Auditory regions sensitive to spectral and/or temporal filtering

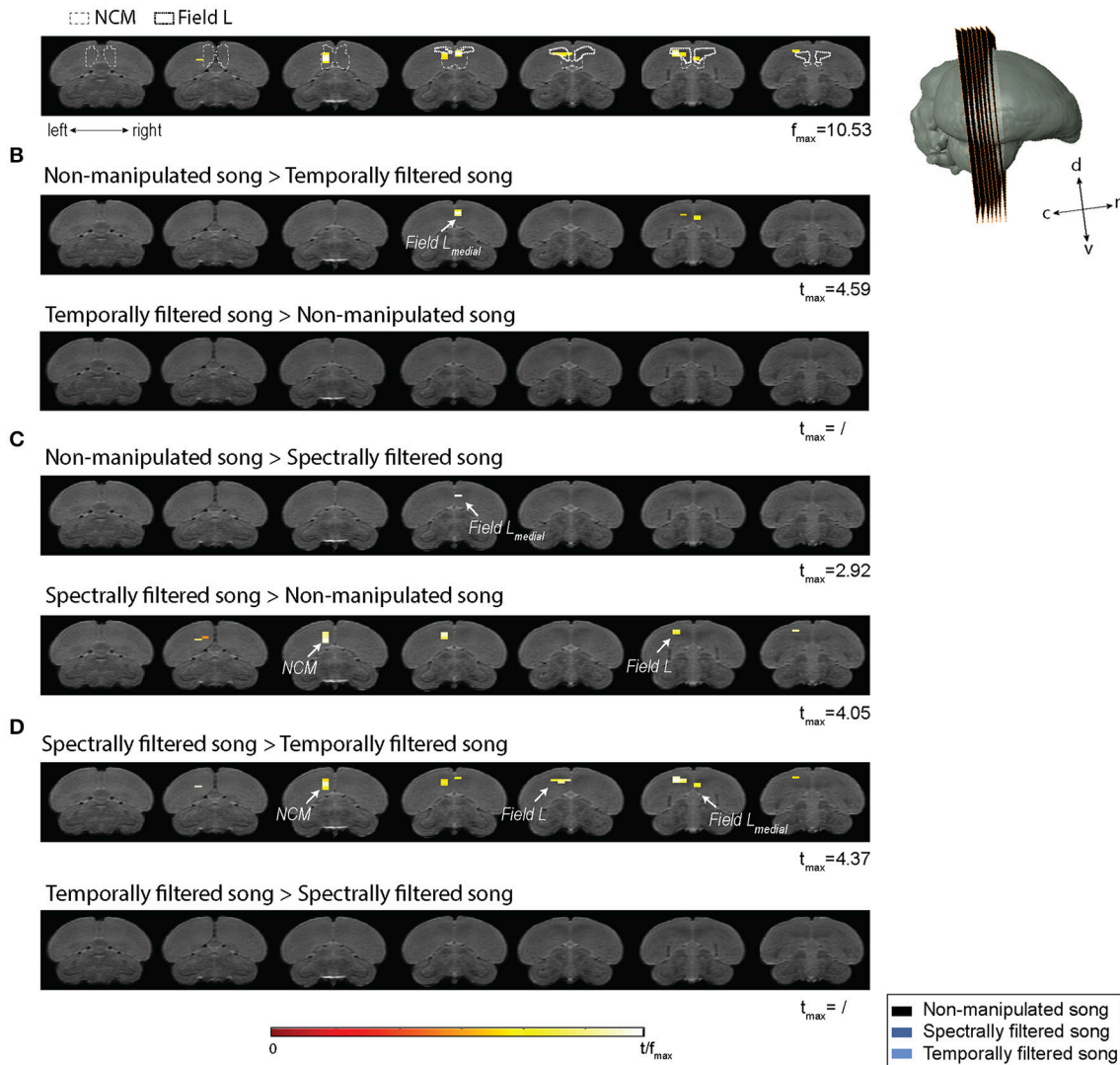


FIGURE 3 | Effect of temporal and spectral filtering of song on neural activation of the auditory forebrain. The image series represents adjacent coronal slices of the population based template with the left image corresponding to the most caudal and the right to the most rostral slice as illustrated by the 3D rendering on the right. Statistical maps are superimposed on the image series. *T*-values are color coded according to the scale displayed in the figure. All voxels with $t > 2.76$ ($p_{\text{uncorrected}} < 0.005$) are displayed. **(A)** Statistical map of the main effect of stimulus class (one way ANOVA—within subjects). Delineations of two auditory forebrain regions from the zebra finch MRI atlas dataset as published earlier are included as a guide for the localization of the effect (Poirier et al., 2008; van der Kant et al., 2013). **(B–D)** Statistical maps of all *post-hoc t*-tests. **(E)** Average of the estimated relative response amplitude (BOLD contrast estimates (β), expressed in non-dimensional units) elicited by playback of non-manipulated song and its corresponding temporally and spectrally filtered variants in the indicated clusters (clusters of significant voxels in different sub-regions as illustrated in the maps above and mirrored counterparts in the opposite hemisphere). The (sub)cluster field L-medial represents the ensemble of voxels within the most medial portion of the large Field L cluster. The zero level corresponds to the estimated mean during rest periods and the error bars to standard errors across voxels in the cluster. c, caudal; d, dorsal; r, rostral; v, ventral; NCM, caudomedial nidopallium ($^*p_{\text{FWE}} < 0.05$; $^{\circ}p_{\text{FWE}} < 0.10$; $n = 15$).

TABLE 1 | Summary of all supra-threshold clusters ($p_{\text{uncorrected}} < 0.005$) in the voxel-based analysis comparing BOLD responses to the different stimulus types.

Contrast	Region	Hemisphere	$F_{\text{max}}/t_{\text{max}}$	Extent (mm ³)
Main effect stimulus class	Field L + NCM	Left and Right	10.53	1.10
	Field L	Right	7.02	0.10
Post-hoc t-tests				
Non-manipulated song > Temporally filtered song	Field L (medial)	Right	3.59*	0.15
	Intersection Field L/NCM (medial)	Right	3.45*	0.10
	Field L	Left	2.64°	0.05
Temporally filtered song > Non-manipulated song	No supra-threshold clusters			
Non-manipulated song > Spectrally filtered song	Field L (medial)	Right	2.92°	0.05
Spectrally filtered song > Non-manipulated song	NCM	Left	4.05*	0.40
	Field L	Left	3.14*	0.15
Spectrally filtered song > Temporally filtered song	Field L + NCM	Left	4.37*	0.95
	Field L (medial)	Right	3.35*	0.10
	Intersection Field L/NCM (medial)	Right	3.42*	0.05
Temporally filtered song > Spectrally filtered song	No supra-threshold clusters			

One way ANOVA—within subjects ($n = 15$); For post-hoc t-tests: * $p_{\text{FWE}} < 0.05$; ° $p_{\text{FWE}} < 0.10$.

a decrease in BOLD response to the filtered song compared to the response to non-manipulated song was observed (**Figure 3B**). In contrast, the effect of spectral filtering was most prominent in the left hemisphere with an increased response in both Field L and NCM after filtering (**Figure 3C**). In the right hemisphere, we observed a small cluster within Field L with decreased responses to spectrally filtered song vs. non-manipulated song. The strongest (in terms of t_{max}) and most extended effect was observed when comparing the responses to the manipulated songs against each other. This effect was related to a significantly higher response to spectrally filtered song vs. temporally filtered song. Similar to the increased response of spectrally filtered as compared to non-manipulated song, this effect was mainly observed in left Field L and NCM.

Lateralization

Although the effect of spectral and temporal filtering was observed in both hemispheres, the relative pattern of activations showed a clear dissociation between hemispheres depending on the type of filtering. To study this further, we calculated LIs for each of the observed filtering effects (i.e., increased or decreased response to one stimulus type relative to responses elicited by the other two stimulus types) in each subject. The ROI for this analysis was composed of the ensemble of voxels (and their mirrored counterparts) in which a significant effect of song filtering was found in the above described voxel-based group analysis (**Figure 4, Top**).

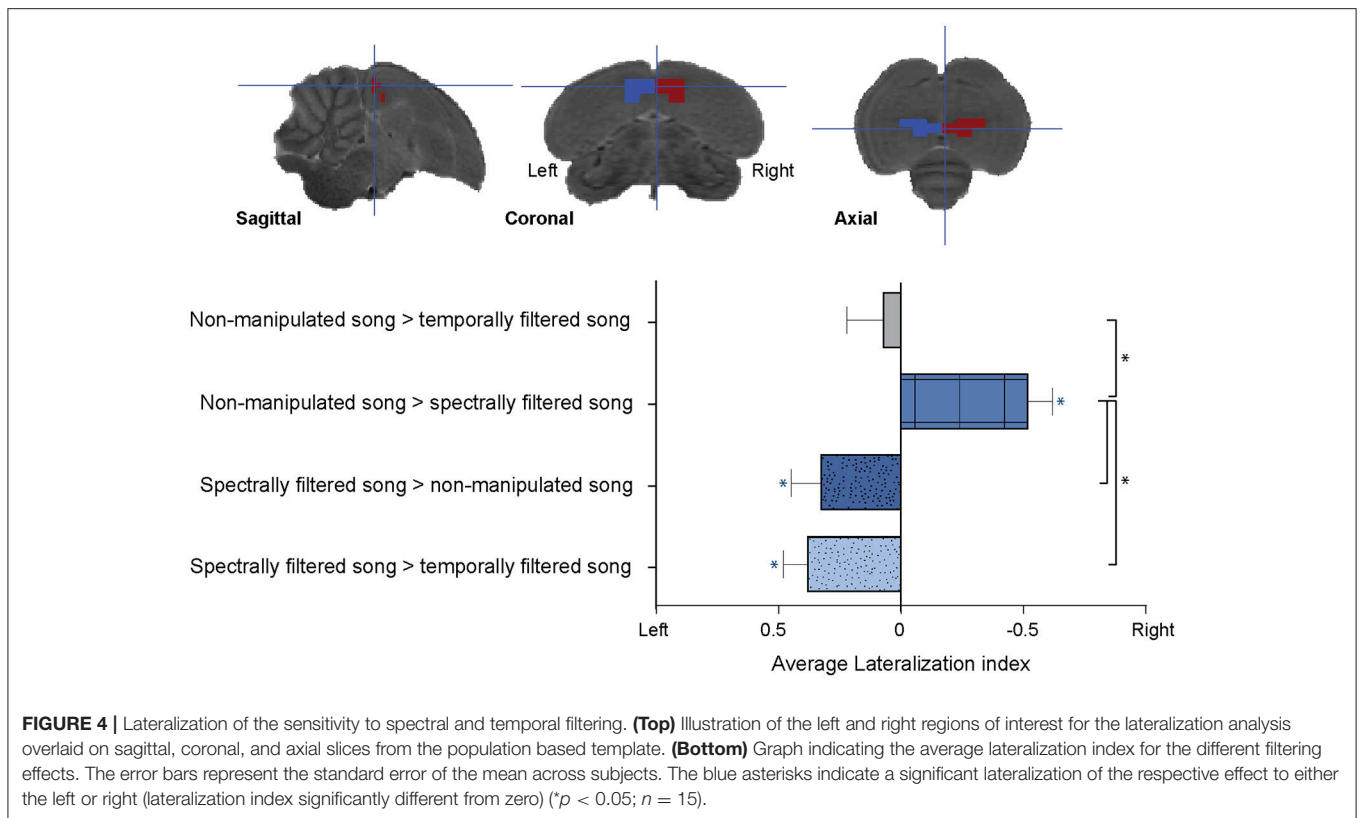
We found a significant right lateralization for the decrease in BOLD response after spectral filtering of song (AveLI = -0.5218 ± 0.09818 ; $p < 0.0001$). In contrast, we observed a significant left sided bias for the increase in BOLD response to the spectrally filtered song relative to non-manipulated (AveLI = 0.3245 ± 0.1261 ; $p = 0.0111$) and temporally filtered song (AveLI = 0.3783 ± 0.1011 ; $p = 0.0011$). This observed left lateralization corroborated the increase in BOLD response revealed by the voxel-based analysis. In

contrast to the lateralization observed for the effects of spectral filtering on the BOLD response, no significant lateralization was observed for the decrease in response after temporal filtering (AveLI = 0.0688 ± 0.1497 ; $p = 0.3263$; **Figure 4, Bottom**).

Direct comparison of the calculated LIs revealed a strong main effect of filtering effect [$F_{(3, 42)} = 13.3502$, $p < 0.0001$; **Figure 4**]. *Post-hoc* analysis indicated that the right sided bias observed for the decrease in BOLD response after spectral filtering was significantly different ($p = 0.0033$) from the decrease in neuronal response after temporal filtering, which was as stated above found to be more bilateral. Second, this right hemispheric bias for the decrease in BOLD response after spectral filtering was in strong contrast to the left hemispheric bias observed for the increase in BOLD response after similar (spectral) filtering of the song (LI spectrally filtered > non-manipulated song vs. LI non-manipulated song > spectrally filtered song: $p < 0.0001$; LI non-manipulated > spectrally filtered song vs. LI spectrally song > temporally filtered song: $p < 0.0001$). Note that an additional analysis was performed in which the large ROI was split into two separate ROIs including only Field L or NCM voxels, respectively. For both the Field L and NCM ROIs, this analysis yielded similar results as those reported for the large ROI with more pronounced effects in Field L as compared to NCM (**Figure S3**).

DISCUSSION

The auditory forebrain in songbirds is specialized to selectively process conspecific song (Grace et al., 2003; Hsu et al., 2004). Specific neuronal tuning to natural spectral and temporal modulations in song is thought to drive this selectivity and as such enable discrimination of natural sounds in the auditory system (Woolley et al., 2005). In the present study, we used fMRI to investigate the influence of filtering the natural spectral or temporal modulations from song, on BOLD responses in the auditory forebrain. It has been shown that responses at the population level in forebrain regions are indicative



of response selectivity for song over other complex sounds (Grace et al., 2003). In line with this, we detected robust BOLD responses to song across auditory forebrain regions and found that these responses were influenced by our filtering procedures. Analyzing differences between the responses to the non-manipulated songs vs. the filtered songs revealed sensitivity to manipulations of the natural spectral and/or temporal song structure in both field L and NCM. Although the classical view on auditory processing dictates that neural coding of song becomes more specialized and complex in higher order auditory regions (Theunissen and Shaevitz, 2006), our results indicate equivalent population response selectivity in primary field L and secondary NCM. These findings confirm earlier observations by Grace and colleagues who found that, when considering average firing rate, selectivity for conspecific song vs. synthetic songs with changed spectro-temporal statistics, was similar in Field L and NCM (Grace et al., 2003).

The specific selectivity patterns for natural song vs. filtered song observed in the present study appeared particularly interesting. Spectral filtering induced the most pronounced changes in BOLD responses and showed distinct effects in the two hemispheres. We found that reducing natural spectral structure in conspecific song elicits an increase in BOLD signal in left hemispheric auditory forebrain structures and a subtle decrease in the right hemisphere. In contrast, a reduction of temporal information in conspecific song induced a small

bilateral decrease in BOLD signal in Field L. Although we cannot exclude that filtering in one dimension (i.e., frequency or time) did not influence the other dimension to some extent, the fact that filtering of spectral modulations reduced right hemispheric neural activity suggests a right hemispheric selectivity to process spectral song information. This right hemispheric selectivity for spectral elements of song was also previously demonstrated when the lesioning of right hemispheric auditory circuitry reduced the ability of zebra finches to process harmonic structure in song (Cynx et al., 1992) and might be analogous to right hemispheric spectral domain processing observed in the human auditory cortex (Obleser et al., 2008). The strong increase in left hemispheric activity after spectral filtering is a novel finding which is discussed in greater detail below.

An earlier study from our group similarly analyzed fMRI data obtained by presenting zebra finches with natural conspecific song and spectrally and temporally filtered versions of the song (Boumans et al., 2007). In this study, the strongest effects were observed after modulation of temporal structure and no hemispheric asymmetries were detected (Boumans et al., 2007). We believe that the difference in relative response patterns for the different stimulus types compared to the present results may be due to considerable methodological differences between the two studies. This previous study employed a different pulse sequence and acquired only a single coronal slice through the songbird brain. This protocol limited the studied region to only the ventral

part of the NCM, central field L, and the dorsal part of CMM. The present study assessed responses in the entire forebrain using an SE RARE sequence. The latter allows very accurate between-subject voxel-by-voxel comparisons of the BOLD responses to the different stimulus types. Such voxel-based statistics were not performed in the earlier study and responses in larger sub-regions were averaged to interpret results. Additionally, temporal filtering in this previous study was greater than in the present study such that the syllabic structure was removed altogether. Such stringent temporal filtering, which retained spectral content, may have produced spectrally refined stimuli that elicited stronger activity from a tonotopically organized area, such as Field L (Zaretsky and Konishi, 1976). Future studies of the hemispheric specialization in the zebra finch may be well-served to incorporate such stringent temporal filtering into their experimental designs to investigate this further. Further, although responses to filtered song under isoflurane are reported by Boumans and point to limited effects of filtering on the BOLD response, the final interpretation of the effects described was based on data acquired under medetomidine anesthesia. The discrepancy between results over the different studies might point to an important effect of anesthesia on response selectivity and lateralization of auditory processing as was observed earlier in starling (George et al., 2004, 2005). This finding implies that data on lateralization should be carefully interpreted and verified with data obtained in different states of wakefulness before being able to relate the selectivity of auditory forebrain regions found in this study to perceptual behavior.

The fact that the birds were anesthetized during fMRI in this study whereas humans are seldom anesthetized during studies on hemispheric lateralization of speech processing complicates any comparison between the present study and similar experiments in humans. However, studies reporting left hemispheric specialization for birdsong employed a variety of techniques including ZENK expression (Avey et al., 2005; Moorman et al., 2012), behavioral training (Cynx et al., 1992), and single unit recording (Hauber et al., 2007), only the last of which employed anesthesia of any kind. Likewise, a neuroimaging study of humans reported preserved left hemispheric dominance for receptive language (i.e., speech sounds) during general anesthesia (Rezaie et al., 2014). Nonetheless, employing neuroimaging protocols that require no sedation of the songbirds would be a worthy goal for future studies as this approach would simplify comparisons between human and avian fMRI studies.

The most pronounced effect observed in the current study, which has to our knowledge not yet been described earlier in any other songbird study, was the increase in left hemispheric activity after spectral filtering. Although such an effect has not been described in human speech processing, this finding is highly suggestive of a left hemispheric specialization for temporal domain processing that parallels what is commonly reported in humans (Efron, 1963; Schwartz and Tallal, 1980). Our spectral filtering retained much of the temporal content from the conspecific song, so the increase in left hemispheric activity likely reflects a form of temporal domain processing that

is routinely suppressed by the spectral content of natural sounds. This specialization for temporal domain processing would appear to differ somewhat from that which is reported in humans. In humans, increasing mean formant transition rate, increasing temporal complexity, or removing spectral content tends to cause a drop in right hemispheric activity relative to activity in the more consistent left hemisphere (Belin et al., 1998; Zatorre and Belin, 2001; Schonwiesner et al., 2005; Obleser et al., 2008). Thus, a left hemispheric specialization for temporal domain processing that is analogous, but not homologous, to what is reported in humans might have developed in the zebra finch auditory forebrain. With this interpretation of the data, one would expect to also see effects of temporal filtering in the left hemisphere which was not observed in the present study. It is possible that this lack of a lateralized response to the reduction of temporal information, and the preservation of spectral information, may be due to the relatively conservative level of temporal filtering used to maintain a degree of syllabic structure and to avoid confounding serial order as outlined above. Thus, temporal sensitivity might be mainly driven by temporal information between syllables rather than by the fine temporal structure within the syllables. This should be verified by exposing zebra finches to song stimuli with different degrees of temporal and spectral filtering. Such analysis would also help in differentiating between effects on the BOLD response induced by the nature of filtering (temporal vs. spectral) vs. the magnitude of the filtering.

In summary, our results in zebra finches demonstrate that both primary (Field L) and secondary (NCM) auditory forebrain regions in the left and right hemispheres of zebra finches are sensitive to artificial manipulations of temporal and/or spectral structure in conspecific song. Additionally, we found that the hemispheres differ substantially in how they process these different types of information, pointing to an asymmetric sensitivity to spectral and temporal song structure in the auditory forebrain. The possible left hemispheric sensitivity to temporal structure, together with the observed right hemispheric sensitivity to spectral structure, suggest parallels between our asymmetry findings in the zebra finch forebrain and the classical view on spectral vs. temporal domain processing in the human auditory cortex (Schwartz and Tallal, 1980; Sidtis et al., 1981; Belin et al., 1998; Obleser et al., 2008). Although more studies are necessary to confirm this, our results already provide a strong indication that songbirds can possibly represent a valuable animal model for studying the neural basis of hemispheric asymmetry in auditory processing of complex, learned conspecific communication sounds.

AUTHOR CONTRIBUTIONS

LV and SW contributed equally to this work. Conceived and designed the experiments: LV, SW, GK, and AV; Performed the experiments: LV and SW; Performed final data analysis: LV; Advised on data analysis: JH, MV, and GK; Contributed reagents, materials and/or analysis tools: AV; Wrote the paper: LV and SW.

All authors critically reviewed the manuscript and approved the final version.

FUNDING

This research was supported by grants from the Research Foundation—Flanders (FWO, project Nr G030213N, G044311N, and G037813N), the Hercules Foundation (Grant Nr AUHA0012), Concerted Research Actions (GOA funding) from the University of Antwerp and Interuniversity Attraction Poles (IAP) initiated by the Belgian Science Policy Office (“PLASTOSCINE”: P7/17) to AV, and by a grant from the Human Science Frontiers Program (HFSP—RGP0006/2015: “Wildcog”) to MV. LV is a Ph.D. fellow of the Research Foundation—Flanders (FWO). SW is a postdoctoral fellow of

the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience.

ACKNOWLEDGMENTS

We thank Dr. J. Steckel and Dr. H. Peremans for insightful discussions on stimulus design and data interpretation, and we thank Dr. F. Theunissen for kindly providing us his code for sound filtering.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2017.00549/full#supplementary-material>

REFERENCES

- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54, 2033–2044. doi: 10.1016/j.neuroimage.2010.09.025
- Avey, M. T., Phillmore, L. S., and MacDougall-Shackleton, S. A. (2005). Immediate early gene expression following exposure to acoustic and visual components of courtship in zebra finches. *Behav. Brain Res.* 165, 247–253. doi: 10.1016/j.bbr.2005.07.002
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., et al. (1998). Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10, 536–540. doi: 10.1162/089892998562834
- Boumans, T., Theunissen, F. E., Poirier, C., and Van Der Linden, A. (2007). Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional MRI. *Eur. J. Neurosci.* 26, 2613–2626. doi: 10.1111/j.1460-9568.2007.05865.x
- Brainard, M. S., and Doupe, A. J. (2002). What songbirds teach us about learning. *Nature* 417, 351–358. doi: 10.1038/417351a
- Chew, S. J., Vicario, D. S., and Nottebohm, F. (1996). A large-capacity memory system that recognizes the calls and songs of individual birds. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1950–1955. doi: 10.1073/pnas.93.5.1950
- Cynx, J., Williams, H., and Nottebohm, F. (1992). Hemispheric differences in avian song discrimination. *Proc. Natl. Acad. Sci. U.S.A.* 89, 1372–1375. doi: 10.1073/pnas.89.4.1372
- Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631. doi: 10.1146/annurev.neuro.22.1.567
- Efron, R. (1963). Temporal perception, aphasia and D’ej’a Vu. *Brain* 86, 403–424. doi: 10.1093/brain/86.3.403
- Elliott, T. M., and Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Comput. Biol.* 5:e1000302. doi: 10.1371/journal.pcbi.1000302
- Fernández, G., de Greiff, A., von Oertzen, J., Reuber, M., Lun, S., Klaver, P., et al. (2001). Language mapping in less than 15 minutes: real-time functional MRI during routine clinical investigation. *Neuroimage* 14, 585–594. doi: 10.1006/nimg.2001.0854
- Gabor, D. (1947). Theory of communication. *J. Institut. Electr. Eng.* 94, 429–457.
- George, I., Cousillas, H. J., Richard, P., and Hausberger, M. (2005). State-dependent hemispheric specialization in the songbird brain. *J. Comp. Neurol.* 488, 48–60. doi: 10.1002/cne.20584
- George, I., Vernier, B., Richard, J. P., Hausberger, M., and Cousillas, H. (2004). Hemispheric specialization in the primary auditory area of awake and anesthetized starlings (*Sturnus vulgaris*). *Behav. Neurosci.* 118, 597–610. doi: 10.1037/0735-7044.118.3.597
- Grace, J. A., Amin, N., Singh, N. C., and Theunissen, F. E. (2003). Selectivity for conspecific song in the zebra finch auditory forebrain. *J. Neurophysiol.* 89, 472–487. doi: 10.1152/jn.00088.2002
- Hauber, M. E., Cassey, P. S., Woolley, M. N., and Theunissen, F. E. (2007). Neurophysiological response selectivity for conspecific songs over synthetic sounds in the auditory forebrain of non-singing female songbirds. *J. Comp. Physiol. A* 193, 765–774. doi: 10.1007/s00359-007-0231-0
- Hsu, A., Woolley, S. M., Fremouw, T. E., and Theunissen, F. E. (2004). Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *J. Neurosci.* 24, 9201–9211. doi: 10.1523/JNEUROSCI.2449-04.2004
- Jansen, A., Menke, R., Sommer, J., Forster, A. F., Bruchmann, S., Hempleman, J., et al. (2006). The assessment of hemispheric lateralization in functional MRI—robustness and reproducibility. *Neuroimage* 33, 204–217. doi: 10.1016/j.neuroimage.2006.06.019
- Joos, M. (1948). *Acoustic Phonetics*. Baltimore, MD: Linguistic Society of America.
- Lieberman, A. M., and Mattingly, I. G. (1989). A specialization for speech perception. *Science* 243, 489–494. doi: 10.1126/science.2643163
- Liebenthal, E., Sabri, M., Beardsley, S. A., Mangalathu-Arumana, J., and Desai, A. (2013). Neural dynamics of phonological processing in the dorsal auditory stream. *J. Neurosci.* 33, 15414–15424. doi: 10.1523/JNEUROSCI.1511-13.2013
- Matsuo, K., Chen, S. H., and Tseng, W. Y. (2012). AveLI: a robust lateralization index in functional magnetic resonance imaging using unbiased threshold-free computation. *J. Neurosci. Methods* 205, 119–129. doi: 10.1016/j.jneumeth.2011.12.020
- McGettigan, C., and Scott, S. K. (2012). Cortical asymmetries in speech perception: what’s wrong, what’s right and what’s left? *Trends Cogn. Sci.* 16, 269–276. doi: 10.1016/j.tics.2012.04.006
- Moorman, S., Gobes, S. M., Kuijpers, M., Kerkhofs, A., Zandbergen, M. A., and Bolhuis, J. J. (2012). Human-like brain hemispheric dominance in birdsong learning. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12782–12787. doi: 10.1073/pnas.1207207109
- Obleser, J., Eisner, F., and Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J. Neurosci.* 28, 8116–8123. doi: 10.1523/JNEUROSCI.1290-08.2008
- Phan, M. L., and Vicario, D. S. (2010). Hemispheric differences in processing of vocalizations depend on early experience. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2301–2306. doi: 10.1073/pnas.0900091107
- Poeppl, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3
- Poirier, C., Boumans, T., Verhoye, M., Balthazart, J., and Van der Linden, A. (2009). Own-song recognition in the songbird auditory pathway: selectivity and lateralization. *J. Neurosci.* 29, 2252–2258. doi: 10.1523/JNEUROSCI.4650-08.2009

- Poirier, C., Vellema, M., Verhoye, M., Van Meir, V., Wild, J. M., Balthazart, J., et al. (2008). A three-dimensional MRI atlas of the zebra finch brain in stereotaxic coordinates. *Neuroimage* 41, 1–6. doi: 10.1016/j.neuroimage.2008.01.069
- Rezaie, R., Narayana, S., Schiller, K., Birg, L., Wheless, J. W., Boop, F. A., et al. (2014). Assessment of hemispheric dominance for receptive language in pediatric patients under sedation using magnetoencephalography. *Front. Hum. Neurosci.* 8:657. doi: 10.3389/fnhum.2014.00657
- Robin, D. A., Tranel, D., and Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang.* 39, 539–555. doi: 10.1016/0093-934X(90)90161-9
- Schonwiesner, M., Rubsamen, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528. doi: 10.1111/j.1460-9568.2005.04315.x
- Schwartz, J., and Tallal, P. (1980). Rate of acoustic change may underlie hemispheric specialization for speech perception. *Science* 207, 1380–1381. doi: 10.1126/science.7355297
- Seghier, M. L. (2008). Laterality index in functional MRI: methodological issues. *Magn. Reson. Imaging* 26, 594–601. doi: 10.1016/j.mri.2007.10.010
- Sidtis, J. J., Volpe, B. T., and Holtzman, J. D. (1981). Cognitive interaction after staged callosal section: evidence for transfer of semantic activation. *Science* 212, 344–346. doi: 10.1126/science.6782673
- Singh, N. C., and Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am.* 114, 3394–3411. doi: 10.1121/1.1624067
- Theunissen, F. E., and Shaevitz, S. S. (2006). Auditory processing of vocal sounds in birds. *Curr. Opin. Neurobiol.* 16, 400–407. doi: 10.1016/j.conb.2006.07.003
- van der Kant, A., Derégnaucourt, S., Gahr, M., Van der Linden, A., and Poirier, C. (2013). Representation of early sensory experience in the adult auditory midbrain: implications for vocal learning. *PLoS ONE* 8:e61764. doi: 10.1371/journal.pone.0061764
- Van Ruijssevelt, L., De Groof, G., Van der Kant, A., Poirier, C., Van Audekerke, J., Verhoye, M., et al. (2013). Functional magnetic resonance imaging (fMRI) with auditory stimulation in songbirds. *J. Vis. Exp.* 3:76. doi: 10.3791/4369
- Voss, H. U., Tabelow, K., Polzehl, J., Tchernichovski, O., Maul, K. K., Salgado-Commissariat, D., et al. (2007). Functional MRI of the zebra finch brain during song stimulation suggests a lateralized response topography. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10667–10672. doi: 10.1073/pnas.0611515104
- Woolley, S. M., Fremouw, T. E., Hsu, A., and Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat. Neurosci.* 8, 1371–1379. doi: 10.1038/nn1536
- Zaretsky, M. D., and Konishi, M. (1976). Tonotopic organization in the avian telencephalon. *Brain Res.* 111, 167–171. doi: 10.1016/0006-8993(76)91058-1
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Van Ruijssevelt, Washington, Hamaide, Verhoye, Keliris and Van der Linden. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.