Philipps-Universität Marburg

Fachbereich Medizin

Schwerpunkt Brainimaging

Laboratory for Multimodal Neuroimaging

The networks they are a-changin' - Investigating auditory system connectivities during cortical processing of voice, song and music in left-handers and musicians

Masterarbeit

zur Erlangung des *Master of Science* (M.Sc.)

im interdisziplinären Master-Studiengang

"Kognitive und Integrative Systemneurowissenschaften"

vorgelegt von Benedikt Weiß

aus München

im März 2019

Danksagung

Statutory declaration

I declare that I have authored the present thesis with the title

"The networks they are a-changin' - Investigating auditory system connectivities during cortical processing of voice, song and music in left-handers and musicians"

independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

The present master's thesis hasn't been submitted in the same or in a similar version to any university or used to achieve any other academic grading.

All experimental data have been protected in a methodical and traceable manner and committed to the supervisor.

Marburg, 11th March 2019 _____

Benedikt Weiß

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Titel

„The networks they are a-changin' - Investigating auditory system connectivities during cortical processing of voice, song and music in left-handers and musicians“

selbständig verfasst und keine anderen als die im Text angegebenen Hilfsmittel verwendet habe.

Sämtliche Textstellen, die im Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind mit einer Quellenangabe kenntlich gemacht. Die Masterarbeit wurde in der jetzigen oder in ähnlicher Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Experimentelle Daten wurden ordnungsgemäß und nachvollziehbar gesichert und dem Betreuer/der Betreuerin übergeben.

Marburg, 11. März 2019 _____

Benedikt Weiß

Abstract

Music and speech are alike universally existent among human cultures. As they both serve an important purpose in mediating social communication and are quite unique evolutionary achievements of human kind, it plays a major role to unravel their underlying neuronal processing. Hereof, the present study aimed to investigate the processing of these complex sounds within the early auditory cortex of the human brain. A set of natural stimuli being associated to the three sound categories of voice, singing, and non-vocal music was presented during a fMRI measurement and the gathered data was examined by means of the Bayesian framework Dynamic Causal Modeling. As to that, alterations of effective connectivities between the primary auditory cortex and surrounding non-primary regions during the processing of the respective sound categories were investigated. In this respect, no separable processing framework between the modulations via voice, singing or music was displayed hinting towards less separable underlying mechanisms.
In addition to that, the influences of handedness and musicianship on respective auditory system connectivities were examined. Hereof, both seemed to have an influence on the neuronal processing of the different sound categories, yet it was difficult to generally deduce other models of effective connectivity in left-handers or musicians from the observed patterns.

A distinct functional processing of music and speech is vastly reported among the scientific literature. Concerning that, the present study did not find further evidence for such pronounced specializations. As to that, our findings give rise to rethink current views concerning the functional role of non-primary auditory regions, situated at early time points of the sound processing hierarchy.

Table of Contents

1 Introduction **1**

1.1 Structure and function of the auditory cortex 1

1.1.1 The primary auditory cortex 1

1.1.2 Non-primary auditory areas 3

1.2 Neuronal representations of speech, song and music 4

1.2.1 Comparing music and speech – differences and similarities 5

1.2.2 Singing as a neuronal intermediary 6

1.2.3 Lateralization of acoustic features and auditory categories 7

1.3 Handedness and musical training 9

1.3.1 Influences of musicianship on the auditory network 9

1.3.2 Handedness and hemispheric lateralization 10

1.4 Effective connectivity – dynamic causal modeling 11

1.5 Placement of the present study 12

1.6 Hypotheses and aims of the study 17

2 Material and Methods **18**

2.1 Participants 18

2.2 Stimulus material 19

2.3 Experimental design 19

2.4 Experimental procedure 20

2.5 Data acquisition 21

2.5.1 Image acquisition 21

2.5.2 Musical ear test 22

2.5.3 Questionnaires 22

2.6 Data handling 23

2.6.1 Preprocessing 23

2.6.2 First-level analysis 28

2.6.3 Dynamic causal modeling 29

2.6.3.1 Time series extraction 30

2.6.3.2 Construction and estimation of model space 31

2.6.3.3 Bayesian model selection and -averaging 34

2.6.3.4 Statistical analysis of parameter estimates 35

3 Results **37**

3.1 Validation of musicianship and handedness 37

3.2 Dynamic causal modeling 38

3.2.1 Family- and model-wise random effects Bayesian model selection . . 38

3.2.2 Inferences on effective connectivities through estimated parameters 40

3.2.2.1 Analysis of endogenous parameter estimates 40

3.2.2.2 Analysis of modulatory parameter estimates 42

4 Discussion **48**

4.1 Small sample sizes and statistical adaption 49

4.2 Insights into effective connectivities across the auditory cortex 50

4.2.1 A similar pattern of context-independent connectivity between participant groups 50

4.2.2 Similarities in the neuronal processing of complex sounds 51

4.2.3 Influence of musicianship on the processing of complex sounds . . . 54

4.2.4 Influence of handedness on the processing of complex sounds . . . 55

4.3 Recap of the experimental paradigm 57

4.3.1 The stimulus set 57

4.3.2 Passive listening versus performance of active auditory tasks 58

4.4 The hypothesis-driven construction of model space 60

4.5 Limitations and further work 62

4.6 Conclusion 63

A Appendix - Additional Data **64**

B Appendix - Scanner Log **70**

C Appendix - Ethic Proposal **84**

References **85**

1 Introduction

This first chapter will give an insight into the current views on the structure and function of the auditory cortex in general (1.1) and the processing of complex sounds and underlying auditory features in particular (1.2). Afterwards, handedness- and musicianship-related influences on the auditory system will be depicted (1.3) and the theory behind the research on brain connectivities in neurosciences as well as the examination of effective connectivities by means of Dynamic Causal Modeling will be described (1.4). Finally, the specific aims and hypotheses of the study will be explained (1.5 and 1.6).

1.1 Structure and function of the auditory cortex

Besides its function as an organ of perception of ones natural environment, the probably most important purpose of the auditory system is its mediation of human social communication. Evolutionary developed to detect sound pressure waves in a frequency range of 20-20000 Hz with highest sensitivity at around 3 kHz through amplification properties of the auditory meatus, it is optimally tuned to perceive human speech sounds (Purves et al., 2004, p.287). After mechanoelectrical transduction of topographically represented sound frequencies within the cochlea, this tonotopy called organizational feature is preserved throughout auditory substrates in the brainstem and thalamic midbrain, and transferred to the auditory cortex (for a review, see Saenz & Langers, 2014). Mainly situated on the upper range of the temporal lobe, the area of the auditory cortex (AC) contains neurons being tuned to several auditory features and is therefore responsible for the early cortical processing of sounds (Robert J Zatorre, 2002). Anatomical and physiological distinctions further separate AC into primary and non-primary areas, proceeding from central compartments of Heschel's Gyrus (HG) and the superior temporal plane (STP) towards surrounding areas (e.g., see Langers, Backes, & van Dijk, 2007). The structure and function of AC based on this subdivisions shall be described in more detail.

1.1.1 The primary auditory cortex

Macroanatomically residing on HG lies a region with typical features of a primary sensory cortical area, further referred to as the primary auditory cortex (PAC). Due to a vast amount of investigations on structure and functionality (e.g., see Da Costa et al., 2011; Humphries, Liebenthal, & Binder, 2010; Morosan et al., 2001), PAC consists of broadly agreed properties. Detailed studies on its cytoarchitecture by Morosan et al. (2001)

revealed an tripartite area with clearly distinguishable borders to surrounding tissue. Along the mediolateral axis of HG cytoarchitectonic subareas of Te1.1, Te1.0 and Te1.2 receive thalamic input via a prominent layer with densest amount of thalamocortical fiber tracts terminating in Te1.0. Therefore, this on the middle part of HG centered layer Te1.0, anterolaterally and posteromedially bordered by Te1.2 and Te1.1, is often referred to as the core area of PAC.

The results published by Morosan et al. (2001) make clear that in general PAC resides on HG, but the limitation to surrounding tissues needs a more detailed cytoarchitectonic view which does not consistently match macroanatomical landmarks like gyral and sulcal borders. A reason for this is the existence of anatomical variations between subjects and even brain hemispheres concerning HG (e.g., Dorsaint-pierre et al., 2006; Moerel, Martino, & Formisano, 2014). Partial or even complete duplications of this gyrus are not uncommon so that an additional so called sulcus intermedius may divide HG into an anterior and posterior part (Da Costa et al., 2011). Nevertheless, multiple investigations acknowledge a clear anatomical-functional relationship between variants of HG and PAC. Imaging studies derived through Magnetic Resonance Imaging (MRI) at 3 Tesla like the one of Humphries and colleagues (2010), followed by better resolved 7 Tesla scans used by Da Costa et al. (2011), consistently showed two distinct tonotopic maps spanning over HG and its bordering sulci (anterior: first transverse sulcus, FTS; posterior: Heschel's sulcus, HS). Both experiments visualized divergent neuronal cell responses to separately presented sound frequencies in terms of a tonotopic gradient starting anterior at high frequencies, descending centrally to low frequencies and ascending back to a posterior high frequency cluster. Therefore, this "high-low-low-high" gradient moves along an anterior-posterior axis on the STP and is orthogonally oriented to HG. In corresponding literature, these tonotopic subfields of PAC are usually termed as hR (anterior) and hA1 (posterior gradient), with the intermediate bordering either lying on the gyral crown of single gyri or within an intermediate sulcus of duplicated HGs'(Da Costa et al., 2011; nomenclature derived from homologous studies in non-human primates; e.g., see Rauschecker, Tian, Pons, & Mishkin, 1997).

Taken together, PAC is well definable based on its distinct cytoarchitectonic structure and main feature of functional tonotopy, and is almost superimposable to the macrostructure of HG. While the spectral mapping of PAC can be consistently replicated along corresponding literature (Saenz & Langers, 2014), other possible properties concerning temporal modulated scaling (periodotopy) are still a matter of debate (Barton, Venezia,

Saberi, Hickok, & Brewer, 2012; Herdener et al., 2013; Leaver & Rauschecker, 2016). Therefore, actual state of affairs assign the primary auditory region the main function of a frequency analyzer.

1.1.2 Non-primary auditory areas

HG separates the STP into an anterior part, referred to as the planum polare (PP), and a posterior region, the planum temporale (PT; Morosan et al., 2001). Together with caudolaterally adjacent superior temporal gyrus (STG), these PAC surrounding areas contain sound sensitive neuronal cell populations and are attributed to the non-primary auditory cortex (non-PAC; e.g., Upadhyay et al., 2008; Robert J Zatorre, 2002).
Studies on structural white matter connections using variant methods from post-mortem Dil staining (Tardif & Clarke, 2001) to non-invasive Diffusion Tensor- (DTI; Upadhyay et al., 2008) and Diffusion Spectrum-MRI (DSI; Cammoun et al., 2015) reveal a pattern of AC-connectivity which is compatible with a parallel and hierarchical cascade-like processing stream.
Strong intrinsic connections within regions of PP, PAC and PT are evident, with more narrow fiber structures between neighboring neurons in PAC and more distant connections spreading over larger parts in surrounding non-PAC regions (Tardif & Clarke, 2001). Also, PAC seems to be highly connected to adjacent areas, but has nearly no links to more faraway sections with whom it shares no borders (Cammoun et al., 2015). This proposes a model of an early stage auditory signaling cascade originating in PAC, propagating towards immediately surrounding tissues and from this intermediate switch point to more distant regions. Furthermore, findings on effective connectivity show a simultaneous, parallel transfer from PAC towards PP and anterior STG, as well as towards PT and posterior STG via structurally verified rostral and caudal projections of white matter fibers (Upadhyay et al., 2008).
Besides fundamental agreements on the structure of non-PAC, procedures concerning higher order processing of received sound cues become less clear compared to the well established tonotopic features of PAC. From a reductionist point of view every sound arises from the composition of acoustic energy along the frequency and time domain, as can be visualized in every ordinary sound spectrogram. Based on this acoustic features, Santoro et al. (2014) propose a model of multiple representations of sound spectrograms within non-PAC to dissect different natural sounds. Presenting a range of commonly perceived auditory stimuli with varying spectrotemporal compositions visualized a pattern

of voxels being either tuned to low spectral and high temporal precision or towards fine-grained spectral modulations with coarse temporal resolution. Mapping these patterns along the superior temporal lobe suggests that neuronal populations encoding a joint low spectral/high temporal "view" reside more laterally-posterior to HG on PT and pSTG, while anterior regions on PP and aSTG seem to include neurons tuned to the opposite high spectral/low temporal perception.

In order to maintain a common thread concerning the so far mentioned intrahemispheric organization of AC, it seems to be obligatory to bring up the study of Norman-Haignere, Kanwisher, and McDermott (2015). By measuring responses of individual voxels to a collection of natural sounds and inferring canonical response profiles from this data, the authors were able to detect six components with distinct characteristics along whole AC. These components were either explainable through their embodied selectivity for particular acoustic features like frequency, spectrotemporal modulation and pitch, or through their responsiveness to sound categories labeling for music and speech. Taken together, this new computational strategy was capable of both visualizing the typical tonotopic subfields of PAC (high- and low frequency fields, components 2 and 1 respectively) and replicating the non-primary multi-resolution representations proposed by Santoro et al. (2014; components 3 and 4: rapid temporal/coarse spectral vs. slow temporal/fine spectral modulations, concentrated posteriorly and anteriorly to PAC). Likewise, the authors findings inspire an ongoing dispute regarding the auditory processing of music and speech in distinct regions of non-PAC and hint to more complex, still insufficiently understood processing strategies of human auditory cortex.

1.2 Neuronal representations of speech, song and music

Music and speech are alike universally existent among human cultures. They both share an important purpose in mediating social communication and seem to be quite unique and ancient within human evolution (for a review, see Peretz, 2006; R. Zatorre, Belin, & Penhune, 2002). On the level of auditory perception they represent complex sounds consisting of varying amounts of basic acoustic features, who define their distinct characteristics (Whitehead & Armony, 2018). The differentiation between the domains of music and speech based on their specific attributes requires a further separation of the vocal and non-vocal, respectively instrumental form of music. Hereof, an intermediate domain between those two complex sounds is introduced, representing a vocal form of

music or, vice versa, a musical form of the voice. This intermediate category is singing and combines the unique attributes of both instrumental music and speech in one domain (Schön et al., 2010). Studies which investigate the neuronal representations of music often refer to its instrumental form. therefore in the following, the term "music" labels the instrumental, non-vocal domain to avoid any confusions.

1.2.1 Comparing music and speech – differences and similarities

Since the development of modern brain imaging techniques, an ongoing debate has been intensified questioning to what extent the processing of music and speech accesses same or distinct neuronal resources. Recent studies applying various kinds of auditory stimuli (natural sounds, pseudowords, "jabberwocky" sentences, scrambled stimuli, diverse instruments, human vocalizations), computational methods (univariate and multivariate analyses) and experimental designs (e.g., inclusion of tasks or sole passive listening) tried to shed light on respective mechanisms. Despite a lack of overall comparability derived through this experimental differences, still some compliant findings can be extracted.
Initial fMRI experiments contrasting pseudowords/"jabberwockey" sentences against musical excerpts revealed more activations for speech sounds in ventrolateral areas of the superior temporal lobe (STG), while melodies seemed to recruit more dorsomedial parts along the STP (Rogalsky, Rong, Saberi, & Hickok, 2011; Tervaniemi et al., 2006). Following this, Angulo-Perkins et al. (2014) confirmed bilateral activations of lateral STG and underlying temporal lobe structures in response to human vocalizations (speech and non-linguistic vocalizations) and a prominent sensitivity to instrumental music in anterior regions of non-PAC, PP respectively.
These findings were based on univariate methods and also visualized substantial overlaps within AC for both sound domains. Yet, recent performances based on multi voxel pattern analyses (MVPA) suggest that even within these overlapping regions selectivity for either music or speech occurs within distinct neuronal populations (Norman-Haignere et al., 2015; Rogalsky et al., 2011).
Speech and music as complex sounds are differently composed concerning the amounts of underlying acoustic features. To accurately process vocalization arrays, a fine-grained temporal sampling rate is required, whereas detailed spectral tuning affects sensation to a lesser demand. Vice versa, because pitch and harmonics are important aspects of musical structure, while fast temporal variations are not that prominent ("rhythm"), the spectrotemporal representation of music needs a fine-grained spectral resolution

accompanied by low temporal precision (R. Zatorre et al., 2002; Robert J. Zatorre & Baum, 2012). Referring to this, attempts have been made to infer divergent neuronal pathways of music and speech on the basis of their acoustic structure (Abrams et al., 2011; Fedorenko, McDermott, Norman-Haignere, & Kanwisher, 2012). The conclusion can be drawn that indeed spectral and temporal structures of music and speech are encoded differently and may be therefore involved in the observed distinct processing of music and speech. There is also evidence that a pitch-sensitive region, concentrated in anterior AC (PP included; Norman-Haignere, Kanwisher, & McDermott, 2013), partially overlaps with predominantly music-sensitive areas, as can be seen for overlying components 4 (selectivity to slow temporal/fine spectral modulations, pitch) and 6 (selectivity for the category "music") in the mentioned study of Norman-Haignere et al. (2015).

1.2.2 Singing as a neuronal intermediary

Singing can be seen as an intermediate domain combining both speech and music (Schön et al., 2010; Whitehead & Armony, 2018). By imagining song as including sound characteristics of melodies as well as voice, it may be possible to describe a separable processing framework for complex sounds. Apart from elaborate scrambling approaches, the inclusion of a stimulus category "singing" could be a more accessible method to dissect essential features selectively accounting for music and voice along with potential intermediary overlaps.

As to that, a current study published by Whitehead and Armony (2018) applied such an approach by presenting a variety of auditory stimuli ranging from human vocalizations (including speech and non-linguistic vocalization) to song excerpts (with and without words) to novel melodies (played on various instruments), which was capable to dissect category-specific activation clusters within regions of non-PAC. More precisely, "pure" musical aspects (subtracted out through contrasts "instruments minus speech" and "singing minus speech") activated bilateral areas of PP and right PT, while their vocal shares (conjunction analysis of speech vs. instrumental music and song vs. instruments) elicited stronger responses within lateral STG and subjacent temporal areas.

In short, the utilization of this newish approach allowed a more precise visualization of cortical auditory procedures and may provide a hint to more general processing aspects of complex sounds.

1.2.3 Lateralization of acoustic features and auditory categories

So far, the mentioned findings cover aspects of early cortical sound processing without any further discrimination between bilateral brain regions. However, when it comes to the issue of how auditory perception is executed there is various evidence that the processing of speech tends to be predominantly shifted towards the left hemisphere, while the right hemisphere seems to play a more important role in the processing of music (for a review, see Tervaniemi & Hugdahl, 2003).

Apart from this general observation, questions on lateralization within the auditory domain remain to be answered in more detail. In this respect, it is quite unclear at which point of the processing stream lateral shifts are introduced. Does lateralization already occur within areas of the early AC with the possible consequence of a maintained separation of processing in downstream "higher order" regions? Or is the early sound processing equally distributed among hemispheres introducing lateralized analyses of complex categories at later time points within more multimodally aligned areas? As to that, a popular model on the cortical distribution of the speech network proposed by Hickok and Poeppel (2007) can be consulted. Based on this, one may assume that the respective lateralization manifests during "higher order" processing of speech aspects like semantics and syntactics or during sensorimotor integration apart from early AC. At the same time, it is highly debated, whether an anatomical and coincidently functional left lateralization of the PT is existent, and if a structure-to-function relationship can be addressed to the processing of speech within this early auditory area (Dorsaint-pierre et al., 2006). Hereof, opposing views exist at all levels: On structural aspects, specific findings are in favor of a larger PT-volume within the left hemisphere (e.g., Geschwind & Levitsky, 1968) while others report a high structural variability between individuals, irrespective of any bias that could be predominantly assigned to the left hemisphere (e.g., Westbury, Zatorre, & Evans, 1999). Concerning the functional role of PT, modern imaging techniques also confound restrictions to speech cues, but prove its sensitivity to various types of complex sounds and auditory features as well as non-acoustic signals (e.g., Angulo-Perkins et al., 2014; Binder, Frost, Hammeke, Rao, & Cox, 1996; Griffiths & Warren, 2002; Norman-Haignere et al., 2015). As to that, the PT seems to contribute to the hierarchical processing of complex sounds, speech included, at a fundamental stage and may hereof display a more dominant role within the left hemisphere (Obleser, Zimmermann, Meter, & Rauschecker, 2007). Following this, updated models on the functional role of the PT have been suggested, some

proposing a computational hub for the segregation and matching of spectrotemporal sound patterns (Griffiths & Warren, 2002), others seeing the PT as a sensorimotor interface preceding higher order multimodal processes (Hickok & Poeppel, 2007). Nevertheless, more fine-grained information must be gathered to be able to comprehensively integrate the role of the PT into the auditory network.

As mentioned in prior sections, attempts have been made to infer the neuronal pathways of music and speech by dissecting compounds of underlying acoustic features. In order to get more insights into proposed hemispheric emphases, R. J. Zatorre and Belin (2001) investigated responses of the AC to spectrotemporally varied sound stimuli using Positron Emission Tomography (PET). The result was, that increasing the rate of temporal changes seemed to conduct increasing activations towards the left AC, while varying the number of spectral elements engaged the right AC more strongly. Based on this insights, they propose that the left-lateralization of speech, respectively right-lateralization of music, relies on an early stage dominance of a more fine-grained resolution for temporal features in the left AC, respectively of higher resolved spectral informations within areas of the right AC (Zatorre et al., 2002). Recent findings, like the mentioned study of Norman-Haignere et al. (2015) seem to approve the lateralization of specific acoustic features, at least in parts. Hereof, the components which were more sensitive to fine spectral/slow temporal modulations were slightly more right-lateralized than the components being more sensitive to rapid temporal/coarse spectral modulations. Still, this between-hemisphere bias was quite small relative to the observed within-hemisphere differences. Also, the components being selective for music and speech were more explainable through their category labeling than basic acoustic features, leaving a knowledge gap concerning the relationship between the processing of complex sounds and acoustic attributes.

Functional asymmetries between hemispheres have also been reported at earliest processing stages of the auditory network. Hereof, neurons of the right auditory core (HG/PAC, respectively) seem to display a higher frequency selectivity via a more pronounced voxel-by-voxel intrinsic functional connectivity than their homotopic counterparts within the left hemisphere (Cha, Zatorre, & Schönwiesner, 2016). Also, disrupting the auditory network at the level of right HG via Transcranial Magnetic Stimulation (TMS) caused widespread functional connectivity decreases in the bilateral AC, which seemed not to be the case when the left HG was modulated (Andoh, Matsushita, & Zatorre, 2015). These findings hint to a generally more important role of the right PAC for the first cortical processing of acoustic inputs, suggesting an increased transcallosal transfer of spectral

informations from right PAC towards left then vice versa. As to that, an advanced investigation of interhemispheric coupling strengths between HGs' as well as intrinsic functional connectivies of respective regions seems to be plausible to shed more light on processing mechanics of the early AC.

1.3 Handedness and musical training

Musical training as well as handedness are both considered to have an influence on processing strategies of neuronal systems (for a review, see Pantev & Herholz, 2011; Willems, Der Haegen, Fisher, & Francks, 2014). For the auditory network, various structural as well as functional differences have been reported for musicians compared to people lacking of musical expertise. Handedness-related differences affect generally observed functional brain lateralizations of various cognitive networks. The following section will take a closer look on these findings.

1.3.1 Influences of musicianship on the auditory network

Musicianship is considered to be a relevant factor of influence in the processing of speech and music within non-primary regions of the AC. Hereof, Angulo-Perkins et al. (2014) were able to prove that the neuronal activation patterns in musicians differed from those of non-musicians in the PP and PT. In this respect, it was found that the stronger activation of the PP in response to music was more frequently bilaterally displayed within musicians (bilateral PP-activations for less then a half of the included non-musicians). Concerning the right PT, non-musicians displayed a considerably higher activity in response to speech as compared to music, while musicians showed more similar responses for both sound categories. Based on these findings, the authors suggest that the disappearing discrimination of the two category types in musicians might derive from a more similar processing of music and language within regions of non-PAC.

Concerning structural differences, musical expertise seems to especially affect the neuronal plasticity of the PT and the interhemispheric connection between its homologous. In addition to previous findings of an increased Corpus Callosum size in musicians (Schlaug, Jäncke, Huang, & Steinmetz, 1995), Elmer, Hänggi, and Jäncke (2016) reported an increased white matter connectivity between bilateral PTs', as indexed by a reduced transcallosal radial diffusivity in musicians. Furthermore, the same research group was also able to prove an increased cortical surface area of the left PT in musicians compared

to non-musicians (Elmer, Hänggi, Meyer, & Jäncke, 2013). Also, the authors claimed a structure-to-function relationship in terms of a better performance of musicians in the discrimination of fast-changing phonetic cues. Hereof, the left PT of musicians exhibited enhanced activity during the processing a consonant-vowel syllable categorization task (Elmer, Meyer, & Jäncke, 2012). To sum it up, musicianship seems to alter performances of the auditory network not solely in the music-, but also in the speech domain.

1.3.2 Handedness and hemispheric lateralization

Handedness relates to the topology of neuronal networks underlying various cognitive functions. Concerning left-handedness, it has been shown that general observations on neuronal processing strategies often deviate from the right-handed majority (for a review, see Willems et al., 2014). In this respect, the usually detected left-lateralization of the language network is not such pronounced in left-handers (Pujol, Deus, Losilla, & Capdevila, 1999; Steinmetz, 1996; Szaflarski et al., 2002). As to that, Knecht et al. (2000) reported that language processing seems to occur more bilaterally in left-handers. Additionally, a greater proportion of the left-handers (27%) displayed a right-hemispheric dominance of the language network, whereas only a small share of the right-handers (4%) exhibited this quite rare natural phenomenon. As to functional brain lateralization in general, handedness-related differences were found for various other networks and respective cognitive functions, including the face perception network (Frässle, Krach, Paulus, & Jansen, 2016), the motor system (Dassonville, Zhu, Ugurbil, Kim, & Ashe, 1997; Klöppel et al., 2007) and the processing of spatial attention (Jansen, Flöel, Menke, Kanowski, & Knecht, 2005). Altogether, these findings hint to a fundamental linkage between handedness and brain lateralization in general.

Because of such deviations concerning cognitive functions, left-handed participants are often excluded in studies of cognitive sciences. Though it is reasonable to avoid variances in the data by studying homogenous samples matched for handedness, the additional value coming from the inclusion of left-handers towards a broader understanding of cerebral functions and brain lateralization is nowadays acknowledged (Willems et al., 2014). Therefore, especially for studies dissecting functional differences between hemispheres, handedness should be a considered factor. Speaking of the auditory cortex and its proposed hemispheric differentiation, it seems quite reasonable to extend current findings in connection with handedness.

1.4 Effective connectivity – dynamic causal modeling

The examination of cognitive processes through experimental manipulations discriminates between functional and effective interactions of brain regions. In neuroimaging studies, the analysis of BOLD signal time series' functional connectivity is defined as the observed "temporal correlations between spatially remote neurophysiological events" (K. J. Friston, Frith, Liddle, & Frackowiak, 1993) whereas effective connectivity is denoted as the "influence one neuronal system exerts over another" (K J Friston, Frith, & Frackowiak, 1993). In other words, the experimental manipulation of a cognitive process may reveal functionally specialized areas, whose simultaneous appearance is in the first place simply a statement about observed correlations. This functional interaction does not provide any direct insight into how these correlations are mediated and does not contain any informations on directionality (K. J. Friston, 1994). Effective connectivity by contrast denotes the situation in which the functional activity in one region predicts the functional activity of an other area and therefore infers on directed (causal) influences between neurons or neuronal populations (Friston, 1994). Taken together, while functional connectivity solely describes statistical dependencies of regional (BOLD-) time series between populations of a neuronal network, effective connectivity goes a step further by showing its time-dependent causal relations (K. J. Friston, 2011).
Hence, the investigation of effective connectivities between brain regions requires some preconditions. As to that, the validity of effective connectivity depends on a model depicting the states of a neuronal system to be examined. The construct validity of such a model is again dependent on the fact if its construction conforms to existing structural white matter connections. In this respect, the validity of effective connectivity ascribes to the models' validity and is therefore model-dependent (Friston, 1994). An other aspect is that like functional connectivity, fMRI-based inferences on effective connectivity depend on measurements of hemodynamic changes over time. Therefore, a correlating hemodynamic response pattern (BOLD signal changes within voxels) across regions of interest completely constraints the estimation of effective connections, which are therefore also time- and experiment-dependent (Friston, 1994).

In due consideration of the mentioned preconditions concerning dependencies of effective connectivities, also experimentally modulated changes of effective connectivities across neuronal networks can be investigated. A method which computes hypothesis-driven effective connectivities among brain regions is the Bayesian approach of Dynamic Causal

Modeling (DCM; K. Friston, Harrison, & Penny, 2003). Hereof, time series of fMRI BOLD activations can be utilized as the dependent variable to infer on causal interactions between neuronal populations of a modeled brain network. Due to the fact that inferences on neuronal brain activities are derived from an indirect metabolism-based measurement, the DCM-approach for fMRI uses a hemodynamic forward model which integrates a neuronal state equation into a set of hemodynamic state equations. These state equations contain parameters describing biophysically reasonable mechanisms, which finally enable to make assumptions on the dynamics of interactions among hidden states of a neuronal system (Klaas E. Stephan, Penny, Marshall, Fink, & Friston, 2005).

As to the implementation, the parameters of a reasonable neuronal system model are estimated such that a predicted BOLD signal, resulting from the conversion of modeled neuronal dynamics into hemodynamic responses (via a hemodynamic forward model), corresponds as closely as possible to the observed experiment-dependent BOLD time series (Friston et al., 2003). For the practical testing of hypotheses, a set of such generative context-dependent connectivity models, which mirror causal interactions among cortical regions of interest (dynamic causal models, DCMs), is estimated and posterior probabilities are compared (model estimates, parameter estimates).

To sum it up, K. E. Stephan et al. (2010) defined five key features of DCMs unifying all above-mentioned aspects: A DCM is 1) dynamic, using differential equations to describe hidden neuronal states 2) causal, explaining how dynamics in one neuronal population cause dynamics in another and how these interactions are modulated by experimental manipulations or endogenous brain activity 3) striving for neurophysiological interpretability 4) utilizing a biophysically motivated and parameterized forward model to link the modeled neuronal dynamics to specific features of measured data, and 5) entirely motivated on Bayesian statistics to infer on posterior probabilities. As to that, the investigation of effective connectivities among brain regions using DCM seems to be a valid attempt.

1.5 Placement of the present study

The present study aims to add knowledge to the scientific research concerning the processing of complex sounds within the early auditory cortex. Hereof, a set of various auditory stimuli being associated to the three sound categories of voice, singing, and non-vocal music will be presented during a fMRI measurement. Such attempts have

already been performed in recent neuroimaging studies (e.g., Angulo-Perkins et al., 2014; Whitehead & Armony, 2018) revealing functionally specialized auditory regions for the mentioned sound categories. Hence, to our knowledge there are no existing studies that examined effective connectivities between areas of the auditory core network during the processing of voice, singing and non-vocal musical stimuli. Therefore, our attempt is to add a more detailed insight into the neuronal mechanisms underlying the processing of complex sounds by investigating effective connectivities between primary and non-primary auditory regions, within and between hemispheres. Additionally, we want to examine the influences of handedness and musicianship on auditory system connectivities, derived during the processing of respective sound categories. Musical training is associated with neuronal plasticity of the auditory system (e.g., Elmer et al., 2016, 2013, 2012) while handedness influences the spacial location of various neuronal networks including the language network (e.g., Knecht et al., 2000; Willems et al., 2014). Therefore, the inclusion of (right-handed) musicians and left-handers (without any musical training) may result in a broader understanding of processing strategies within as well as between auditory cortices, especially when it comes to proposed functional lateralizations (e.g., Tervaniemi & Hugdahl, 2003; Zatorre et al., 2002).

For the examination of our hypotheses we aimed to explore effective connectivities of auditory areas being induced at the earliest time points of the cortical processing hierarchy, right after the receiving of thalamic input. This means that our interests were restricted to essential, unimodal functions of the auditory domain. Neither emotional contents of music nor semantic or syntax of speech, nor any higher-ordered multimodal integration were a matter of debate. As to that, the choice of the auditory stimuli, experimental design and cortical regions to be investigated was conditioned by this restriction and based on prior findings concerning auditory network processes.

The stimulus set being utilized herein was originally assembled for the mentioned study of Whitehead and Armony (2018). There are designated reasons why this set was appropriate for our purposes. First of all, it is composed of a variety of auditory cuttings for each of the three different sound categories voice, singing and non-vocal music. More precisely, the "voice" category consists of phrases spoken in many different languages as well as non-verbal stimuli ("baby talk"). The favorable attempt hereof was to exclude any semantic biases, because it can be assumed that only a few of the stimuli were understood by the participants. By investigating early auditory cortical areas, we aimed for non-semantic

"voice", which includes "speech" in terms of social communication, but is not limited to it, because the term refers generally to the non-verbal features of language (Whitehead & Armony, 2018). For the same reasons, the stimuli of the category "singing" were suitable, because they contain song excerpts in different languages as well as non-verbal chants. Finally, due to the fact that the non-vocal music-stimuli of the set consist of melodies played on various instruments, we ensured to avoid any bias towards specific acoustic features like instrumental timbre.

By containing the three different sound categories, the stimulus set covers a broad range of the acoustic features underlying complex sounds. As to that, instrumental music and speech differ in terms of their basic acoustic structure. For example, speech is low spectrally and coincidently high temporally modulated, while non-vocal music depicts a more prominent spectral modulation at a simultaneously low temporal modulation rate (Zatorre & Baum, 2012). Singing as an intermediate domain combines the characteristics of both music and speech (Schön et al., 2010). Therefore, the presentation of voice, singing and non-vocal music introduces a gradient of underlying feature modulations, which combined display the unique characteristics of the respective categories. Concerning the early auditory cortex, it would be possible to describe a separable processing framework for complex sounds apart from elaborate scrambling approaches, visualized through different effective connectivity patterns and modulation strengths between regions of the neuronal network.

As mentioned, our investigation distinguished between non-semantic voice, "a capella" singing and instrumental (non-vocal) music. Therefore, to avoid any possibility of confusion on subsequently used terms, the terminology "speech" refers equally to "voice" and the label "music" will be purely restricted to the instrumental domain, while "singing" attributes to the vocal form of music (or vice versa musical form of voice). Studies investigating early auditory cortical areas (this one included) generally use "speech-stimuli" (excerpts of phrases), but aim for non-semantic voice, not being interested in respective semantics or syntactics of language (e.g., Whitehead & Armony, 2018).

Measuring BOLD signal time series via fMRI was the central methodological approach for acquiring data to examine our hypotheses. It is one of the most important techniques to spatially visualize cognitive processes in the human brain. The advantages hereof are the non-invasiveness during the measurement as well as its high spatial resolution (Huettel, Song, & McCarthy, 2004). Especially for investigations of the auditory domain,

neuroimaging has provided new insights into neuronal processes and is therefore being used in a vast amount of recent studies (e.g., Angulo-Perkins et al., 2014; Cha et al., 2016; Norman-Haignere et al., 2015; Rogalsky et al., 2011; Santoro et al., 2014).
Concerning the implementation of the experiment, a passive listening paradigm was preferred over the inclusion of any active task performances. The reason hereof is that various researches confirmed the influence of auditory tasks on activations within the AC besides stimulus-dependent modulations (e.g., Grady et al., 1997; Petkov et al., 2004; Woodruff et al., 1996; Woods et al., 2009). As to that, our subjects passively listened to the presented stimuli without any further active performances while watching a muted nature documentation to ensure that they were paying equal (non-selective) attention to the sounds.

The regions of interest (ROIs) we chose to dissect were located within primary and non-primary areas of AC. Macroanatomically, PAC resides on HG and can be further separated into the three cytoarchitectonically distinct areas Te1.0, Te1.1 and Te1.2 (Morosan et al., 2001). Due to the fact that the densest amount of thalamocortical fiber tracts terminates in Te1.0, it is often referred to as the core area of PAC. Therefore, the ROI which displayed PAC in this study was the area Te1.0. Hereinafter, when speaking of "PAC" or "HG" we refer to measurements within this area. For the non-primary regions, we focused on two areas on the superior temporal plane: the PP anterior to HG/PAC and the PT posterior to HG/PAC. Within both regions, activations are reported across many distinct studies dissecting the processing of complex sounds or designated auditory features (e.g., Angulo-Perkins et al., 2014; Norman-Haignere et al., 2013, 2015; Whitehead & Armony, 2018). PT is also vastly reported in terms of auditory lateralization (e.g., Dorsaint-pierre et al., 2006). Due to the fact that we were also interested in possible differences between hemispheres, the three ROIs (PP, PT, HG) were included bilaterally (see Fig.1).
The examination of effective connectivities between these six ROIs during the perception of complex sounds was conducted by means of the method DCM (Friston et al., 2003). The construction of the comprehensive model space, mapping the early auditory network, is depicted in detail in the materials and methods (subsection 2.6.3.2).

We further wanted to investigate the influences of handedness and musicianship on auditory system connectivities. Hereof, we constructed three participant groups: right-handed non-musicians, right-handed musicians and left-handed non-musicians. We a priori formulated own criteria, based on which we defined people as musicians or non-musicians: Musicians

had to have a minimum of 5 years of practical musical training without interruptions until now and had to practice currently for at least one hour per week. By contrast, for being defined as a non-musician, it was not allowed to currently play an instrument or sing in a choir, or to have any experience of musical training since the last five years. As for this definitions, we aimed for amateur musicians, e.g. students playing in the orchestra of the university and/or singing in a choir. Besides our own definitions, musicianship (and also handedness) of the participants were further validated by means of various tests and questionnaires (see subsection 2.5 of material and methods).

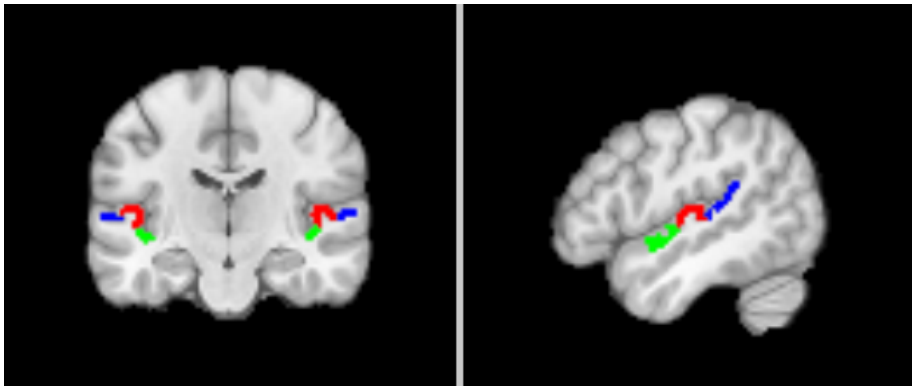

Fig. 1: ROIs, templates in MNI space and received from the research of Sam Norman-Haignere (http://web.mit.edu/svnh/www/homepage/Resources.html; Norman-Haignere, Kanwisher, & McDermott, 2013), plotted on the skullstripped anatomical MNI brain template (coordinates: x = 91, y = 109, z = 91). PT = blue; HG, respectively Te1.0 = red; PP = green.

1.6 Hypotheses and aims of the study

The aim of this study is to make a contribution to the recent research on processing strategies of the auditory cortex during the perception of complex sounds. Hereof, differences in effective connectivities between the bilateral primary auditory cortex and the non-primary areas PP and PT during the processing of voice, singing and music shall be investigated. The idea is that by including song as an intermediate category combining acoustic properties of both music and speech, it would be possible to describe a separable processing framework for complex sounds, visualized through modulation-dependent state changes of the modeled neuronal network. According to that, the question will be further expanded to the influence of handedness and musical training on the effective connectivity of inter- and intrahemispheric connections in terms of the presented sound categories. Taken together, the investigation was led by the following two hypotheses:

1. How does the application of different categories of complex sounds alter effective connectivities between areas of the early auditory cortex?

2. Do left-handers, respectively musicians show other models of neuronal effective connectivity than right-handed people without any musical training?

To our knowledge, no study has investigated processing strategies of AC based on effective connectivities. As to that, this study serves an exploratory purpose to confirm and extent current views. Besides such basic scientific attempts, adding more knowledge to that topic may provide the possibility for clinical applications in the future, i.e. in the therapy of hearing disorders and improvement of cochlea-implants.

2 Material and Methods

In this section, the participants (2.1), stimulus material (2.2), experimental design (2.3) as well as procedure (2.4) will be described, followed by a complete depiction of the acquired data (2.5) and the final experimental data handling (2.6) to examine the hypotheses.

2.1 Participants

Twenty-three subjects (22 native German speakers, 1 native Spanish speaker, 16 females; aged between 18-30 years, mean: 24.04 ± 3.21) with no history of psychiatric or neurological diseases participated in the experiment. All of them had normal or corrected to normal vision, no hearing impairments and fulfilled all of the inclusion criteria (see ethic proposal in appendix C for more details). Starting with the recruitment phase, the participants gave self-report on their handedness and musical education in context of the definitions for musicians and non-musicians (see subsection 1.5 of introduction). The respective questionnaires "Waterloo" (Elias, Bryden, & Bulman-Fleming, 1998) and "Oldfield" (Oldfield, 1971) for handedness, Musical Ear Test (Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010) and screening questionnaires on linguistic/musical aspects for musical education (more detailed information see subsections 2.5.2, 2.5.3) confirmed their membership to the three different groups of right-handed non-musicians (n=11), left-handed non-musicians (n=4) and right-handed musicians (n=8). The musicians played various instruments (piano, violin, contrabass, flute, clarinet, accordion, guitar, drums), four of them were additionally singing in choirs, and all of them maintained their musicianship on a regular basis (mean years of practicing without interruption: 14.25 ± 4.74, mean hours of practice per week: 4 ± 2.55).

The BOLD signal time series extraction within ROIs (see subsection 2.6.3.1) revealed that the datasets of three participants did not fulfill the threshold criteria ($p < 0.05$, uncorrected) and therefore had to be excluded, leaving a final sample size of 9 right-handed non-musicians (7 females, mean age: 23.56 ± 4.00), 4 left-handed non-musicians (2 females, mean: 24.00 ± 3.67) and 7 right-handed musicians (5 females, mean: 24.00 ± 1.70).

All subjects voluntarily participated without any monetary reward. Their recruitment took place via oral request and the central mail relay for students of the University of Marburg. Prior to the experiment, they signed the informed consent, an additional consent covering the new data protection ordinance of the European Union and the open brain consent

documentation, allowing to publish their anonymized data. The study was approved by the local ethics committee of the Medical Faculty of the University of Marburg (see ethic proposal in appendix C).

2.2 Stimulus material

The used auditory stimuli were already established for an other study (Whitehead & Armony, 2018) and were kindly provided by the authors after contacting them. In the following, the set will be shortly described.

The stimuli consisted of 180 short auditory excerpts of approximately 1.5 s duration and belonged to three different categories of complex sounds. 60 of them were recordings of novel compositions of melodies (duration: M=1.49 s, SD=0.13 s) played on various instruments (n=40) and were therefore assigned to the category of (non-vocal) music. Another 60 stimuli composed of phrases spoken in 45 different languages (stimulus duration: M=1.51 s, SD=0.22 s; speakers: 2 children, 33 male adults, 1 non-verbal stimulus "baby talk") and therefore belonged to the sound category "voice". The remaining 60 stimuli assigned to the category "singing" and consisted of new composed song excerpts (M=1.51 s, SD=0.23 s) sung by male (n=28) as well as female (n=32) individuals (children and adults) in 19 different languages or even without words (n=6). All of them were performed "a capella" and were either monophonic, homophonic or polyphonic.

According to the authors (see Whitehead & Armony, 2018), the stimuli were monaural, resampled to 32 bits at a sample rate of 44.1 kHz and adjusted for loudness by normalizing to the short-term loudness (STL) maximum using the Moore and Glasberg Loudness Model (Glasberg & Moore, 2002). After pilot measurements, the original loudness of the stimuli was moderately increased by 10 dB SPL using the free software program Audacity® (https://www.audacityteam.org/) to fit into the present environment of background scanner noise and application via pneumatic earphones.

2.3 Experimental design

The stimuli were presented binaurally on MRI- and EEG-compatible pneumatic headphones adopting a pseudo-randomized block design. Each block consisted of 5 pseudorandomly assigned, contiguous stimuli of the same category (voice, singing, music). Within one run all 180 stimuli were presented once in a corresponding sequence of 36 blocks with a proportion of 12 blocks for each modality. To avoid any type of adaptation, the

block sequences were applied in a counterbalanced manner with randomized intermediate inter-block-delays of varying length (no stimulus presentation for 4, 5, or 6 seconds). Altogether, the whole experiment consisted of 4 runs with a respective duration of about 9 and a half minutes, each of them with a unique order of stimuli, blocks and pause durations. Additionally, the chronology of the runs was also presented at random between the measured participants.

During the presentation of the auditory stimuli, sequences of a nature documentary film (Fothergill, A., Attenborough, D., Fenton, G., BBC Concert Orchestra., British Broadcasting Corporation., Discovery Channel (Firm), Nihon Hoso Kyokai., 2007) were presented soundless on a screen. The aim hereof was to keep the participants attentive during the whole measurement. At the same time, passive listening instead of active task performances were preferred, avoiding any task-dependent selective attention on specific stimuli or additional recruitment of cognitive resources within AC besides stimulus-dependent modulations (e.g., see Grady et al., 1997; Petkov et al., 2004; Woodruff et al., 1996; Woods et al., 2009). As to that, four different movie snippets (for 4 runs) were priorly cleaned from arousing scenes (fighting- and mating scenes, no spiders or snakes due to common phobias) and cut down to the length of the runs.

The whole experiment was programmed in Python (version 2.7; https://www.python.org/) and presented using PsychoPy2 (version 1.85.3; Peirce, 2007, 2008). The detailed code can be inspected on a related repository on Github® (https://github.com/weissbe92/MSc_-thesis_BenediktWeiss).

2.4 Experimental procedure

After signing the consent forms, the participants were prepared for the measurement (removal of any metal objects, dressing of a lab coat in order to avoid marks on clothing from electrode gel). A MRI-compatible 32-channel EEG cap (Brain Products®) was put on and electrode gel had to be applied to all contacts in order to reduce resistances between the electrodes and the participant's scalp. Hereof, the contact-resistances were checked simultaneously using the program Brain Vision® (https://brainvision.com/). The procedure continued, until all contacts displayed resistances below 5 kΩ. After that, participants were brought into the scanner room, where they put on MRI- and EEG-compatible pneumatic headphones above the EEG cap. Afterwards, the participants laid down on the scanner couch and were carefully moved into the scanner tube. All EEG-

associated devices (amplifier, etc.) were placed behind the head coil and the subjects were able to look at a screen behind the examination table through a projector-mirror-system. Participants were told not to move during the whole scanning session which initiated with a structural T1-measurement (5 min). Subsequently, after a short sound check to ensure that the stimuli were audible, the subjects passively listened to the auditory stimuli while watching the nature documentary film presented on the screen. There was no further active task during the functional scans (40 min). Between the 4 runs, participants were contacted through a speaker system. Together with a final structural T2-measurement (5 min), the whole experiment lasted about 50 minutes.

2.5 Data acquisition

Experimental data were collected by measuring fMRI-BOLD responses and EEG event related potentials (ERPs) in parallel. The crucial part for evaluating the hypotheses via the MRI measurement modality was the acquisition of functional T2*-images (see subsection 2.5.1). Aspects on EEG will be handled in an other scientific work and are no further discussed in this study. Besides the hypothesis-driven acquisition, subject specific data were gathered using a musical ear test (MET; Wallentin et al., 2010; see subsection 2.5.2) and various questionnaires (see subsection 2.5.3).

2.5.1 Image acquisition

Imaging data were acquired with a 12-channel head matrix receive coil on a 3 Tesla MRI scanner (Siemens Magnetom TrioTim syngo MR B17, Erlangen, Germany) at the Department of Psychiatry, University of Marburg. Following an initial localizer sequence (1 min), a T1-weighted volume covering the whole brain was obtained to get high resolved structural brain images (176 slices, time to repeat (TR) = 1900 ms, time to echo (TE) = 2.26 ms, voxel size = 1.0 x 1.0 x 1.0 mm^3, field of view (FoV) = 256 x 256 mm^2, flip angle = 9 degree). Subsequently, four runs of T2*-weighted functional images were provided using an echo-planar functional imaging sequence (EPI) covering the whole brain (30 slices, TR = 1450 ms, TE = 25 ms, voxel size = 3.0 x 3.0 x 4.0 mm^3, FoV = 192 x 192 mm^2, flip angle = 90 degree). This measurement made up the main part of the scanning session, imaging the BOLD contrast as the dependent variable during the presentation of the paradigm.

Finally, a second structural whole brain image was acquired (T2-weighted, 176 slices, TR

= 3200 ms, TE = 402 ms, voxel size = 1.0 x 1.0 x 1.0 mm^3, FoV = 256 x 256 mm^2). For more detailed information on scanning properties, see the measurement log in the appendix A.

2.5.2 Musical ear test

The MET for musical competence after Wallentin et al (2010) is an approach to distinguish between musicians and non-musicians and was therefore used to further validate the self-reported membership of the participants to the corresponding groups. It contains 104 pairs of auditory stimuli and is divided into two parts. The first part tests for melody-based musical competence and contains 52 pairs of short melodies played on a piano. Participants have to listen to these melodies and decide afterwards if they were equal or different. Half of these melodic phrases were equal and half of them contained a pitch difference in the second melody. The second part of the MET aims to test for rhythm-based musical competence. 52 pairs of rhythmic phrases played with wood blocks are presented and the participants have to decide again if they were equal or not. Here again, half of the rhythm pairs were equal and the other half differed.
The test was adapted to a non paper-based version using PsychoPy2 (version 1.85.3; Peirce, 2007, 2008) and presented on a MacBook Pro via headphones (Sennheiser® HD 380 pro). Participants had to respond to the stimuli pairs via keypress (j for equal and n for non-equal). During the two test-sections no feedback was given, but there was an example with feedback before every part. There was also an audio instruction in German at the beginning of the MET explaining the tasks. The whole test lasted about 20 minutes.

2.5.3 Questionnaires

Respective questionnaires were used to validate the participants handedness, screen relevant linguistic (e.g., mother tongue/learned languages) as well as musical aspects (e.g., level of musical training to confirm the attribution to musicians or non-musicians based on own criteria), and evaluate the auditory stimuli based on their familiarity. All of them were filled in electronically on the mentioned MacBook Pro using PsychoPy2 (version 1.85.3; Peirce, 2007, 2008) and can be inspected in the ethic proposal (see appendix C).

Due to its importance concerning the hypotheses, the validation procedure on handedness will be described in more detail. A degree of handedness for each of the subjects was evaluated using two questionnaires: The *Edinburgh Handedness Inventory* (Oldfield, 1971)

and the *Waterloo Handedness Questionnaire* (Elias et al., 1998). Both, "Oldfield" and "Waterloo" resemble each other by interrogating hand preferences on various (daily) activities, which could be answered on a scale from "always preferred" (2 points) to "usually preferred" (1 point) to "no preference" (1 point for both scorings of left- and right-handedness) in context of the corresponding hand. A score is then calculated representing the degree of someones handedness (the more positive the score the more right-handed is a person, the more negative the more left-handed). An example for the evaluation of "Oldfield" can also be found online (http://zhanglab.wdfiles.com/local—files/survey/handedness.html). A third, less important questionnaire inquired the handedness of close relatives, if known by the participants. Altogether, the subjects had to complete five questionnaires (three on handedness, one on linguistic/musical aspects, one on stimulus familiarity).

2.6 Data handling

In the following chapter the processing and statistical analysis of the gathered MRI-data will be described. The chronology of the subsections is in conformity with the consecutive processing steps, starting with an anatomical and functional data preprocessing (see 2.6.1), followed by a first-level analysis to set up the general linear model and define contrasts (2.6.2), and concluding with the DCM-based examination of effective connectivities in context of different stimulus modalities and participant groups (2.6.3).

2.6.1 Preprocessing

First of all, in order to provide a comparable and comprehensible data set structure, the scanner derived raw data were reordered to the BIDS standard (Brain Imaging Data Structure, http://bids.neuroimaging.io/). This common standard has the aim to facilitate the sharing and reuse of MRI data between researchers of the neuroimaging community (Gorgolewski et al., 2016). The conversion was performed by using the open source software HeuDiConv (Heuristic DICOM Converter, https://github.com/nipy/heudiconv; detailed instructions on http://nipy.org/heudiconv/) via a docker image. The generated BIDS data set was further validated through an online BIDS-Validator (http://bids-standard.github.io/bids-validator/) and preserved throughout the subsequent processing steps.

The next step was to anonymize the data set of every participant to guarantee the protection of data privacy. Hereof, the anatomical images covering

the whole participants head were defaced using the python-based tool pydeface (https://github.com/poldracklab/pydeface, see Fig.2).

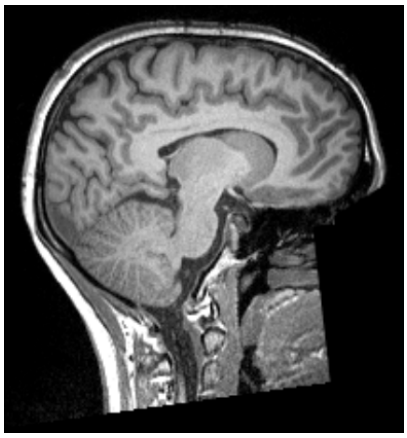

Fig. 2: Defaced anatomical T1 image of sub-02.

Following, a data quality measure was applied using the open source software MRIQC (https://github.com/poldracklab/mriqc; Esteban et al. (2017)) via a docker image. This software includes the identification of background noise, motion artifacts and outliers in the acquisitions. Manual inspection of the results for all structural and functional images of every participant confirmed that no data had to be excluded due to poor quality.

Next, to perform an anatomical preprocessing on the structural MRI data, the open source toolbox mindboggle was adopted (https://mindboggle.info; Klein et al. (2017)). This application provides accurate volume- as well as surface-based brain measures and segmentations, combining software from FreeSurfer (https://surfer.nmr.mgh.harvard.edu/; Dale et al., 2004; Fischl et al., 2002; Fischl et al., 2004; Han & Fischl, 2007) and ANTs (Advanced Normalization Tools, http://stnava.github.io/ANTs/; Avants et al. (2012)) with further morphometric extraction algorithms, and can be utilized via an existing docker image (https://mindboggle.readthedocs.io/en/latest/). Important outputs for subsequent preprocessing steps were skullstripped and intensity normalized whole brain volumes for later co-registration as well as cortical ribbon templates to generate specific brain masks.

Afterwards, functional preprocessing steps which are typically deployed on fMRI data were programmed in python using a Nipype pipeline (Neuroimaging in Python Pipelines and Interfaces; Gorgolewski et al. (2011)). Nipype provides python-based interfaces of several external neuroimaging software packages like SPM (Statistical Parametric Mapping,

Welcome Department of Cognitive Neurology, Institute of Neurology, London, United Kingdom), FSL (FMRIB Software Library; Jenkinson, Beckmann, Behrens, Woolrich, and Smith (2012)), FreeSurfer and ANTs, and has therefore the advantage of combining different tools of these packages into one environment. This enables to design an individual workflow for the data and has further improvements concerning automatization and parallelization of subsequent processing steps.

The array of the workflow initialized with a realignment of the functional images using the "Realign" function of the SPM interface (release 6685). The aim hereby is to correct for head motions along the gathered series of images, which can occur translational and rotational along the x-, y-, and z-axis. The motion correction was performed by realigning the first volume in each run to the first volume of the first run, and then registering each image in each run to the first volume of that run as a reference. The gained motion parameters (the transformation along the axes which provides the best match to the reference volume) were later included into the general linear model (GLM) as regressors (6 regressors for translation and rotation in each direction along the three axes). Also, the nipype algorithm ArtifactDetect (https://www.nitrc.org/projects/rapidart/) was used to detect potential outliers in the realigned images. Hereof, the intensity- and motion parameters were used to compute the maximum head movement of each subject (threshold to detect motion-related outliers for composite motion = 1; intensity Z-threshold to detect images that deviate from the mean = 3).

Subsequently, the workflow continued in two separated pipelines, one for generating specific brain masks in functional space and one for smoothing the realigned images. The detailed steps will be described hereinafter.

To limit the number of voxels being included in later analysis, a mask was created solely containing the brain tissue of the cortical ribbon (gray matter volume). For this, the FreeSurfer derived ribbon template had to be converted from MGZ- into NifTI format (Neuroimaging Informatics Technology Initiative, https://nifti.nimh.nih.gov/), binarized, transformed from structural into functional space (applying the realigned mean functional image as a source) and dilated. This steps were accomplished using various tools of the Freesurfer interface (Binarize, MRIConvert, ApplyVolTransform). Finally, the hereby created functional "ribbon mask" was combined with a skullstripped mask image of the realigned mean functional volume via FSL's ImageMaths function. Taken together, the skullstripping, performed via the BET module of FSL, subtracted non-brain parts from

the volume (skull and meninges) and the overlaid ribbon mask isolated the gray matter shares from sub-cortical areas and cerebrospinal fluid.
Before the realigned functional images entered the first-level analysis, a sparse smoothing was applied using an isotropic 3 mm full-width at half-maximum (FWHM) Gaussian kernel. Smoothing averages data points (in this case voxel volumes) with their neighbors towards a normal distribution. This procedure also suppresses, to some extent, high frequency signals and enhances low frequency signals.

Furthermore, various transformation steps were conducted in separated nipype workflows, starting with a co-registration of the anatomical image to a MNI-template, followed by the transformation of structural ROI-masks into each participant's native space, and terminating with a final conversion from anatomical into functional space.
The final goal of the following preprocessing steps was to be able to perform further analyses on corresponding ROIs within each participant's native, functional space. The challenge hereby is to implement an "inverse normalization". This means that at first the individual shape of each subject's brain volume has to be transformed into the MNI space to receive the standard coordinates of wanted ROIs (this is also called normalization). Secondly, this process has to be reversed using the same transformation parameters applied in the prior conversion in opposite direction towards native space again. In the following, this approach will be described in more detail.
First of all, each participant's anatomical image, derived from the FreeSurfer segmentation ("brain.mgz"), was co-registered to a MNI template (ICBM 2009c Nonlinear Asymmetric template, 3 mm resolution, http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009; Fonov, Evans, McKinstry, Almli, and Collins (2009)) using the Registration function of the ANTs interface. The MNI space is a standard space for brain volumes, introduced by the Montreal Neurological Institute and defined through a large series of MRI scans on normal controls. The anatomical image was already skullstripped, therefore the MNI template's skull was also removed (using again FSL's BET interface) prior to the conduction of the transformation.
The mentioned co-registration computed the transformation parameters between MNI space and the anatomical native space of the corresponding participant. In the next step these parameters were inversely used to transform the ROI templates, which subdivided the superior temporal plane into anatomical sub-regions (PP, PT, core region Te1.0 of

PAC/HG) and are based on the MNI coordinate system, back into native space. This was performed using the module ApplyTransforms of the ANTs interface. The native anatomical image served as the reference and the ROI templates were received from the research of Sam Norman-Haignere (http://web.mit.edu/svnh/www/homepage/Resources.html; Norman-Haignere, Kanwisher, & McDermott, 2013)

.

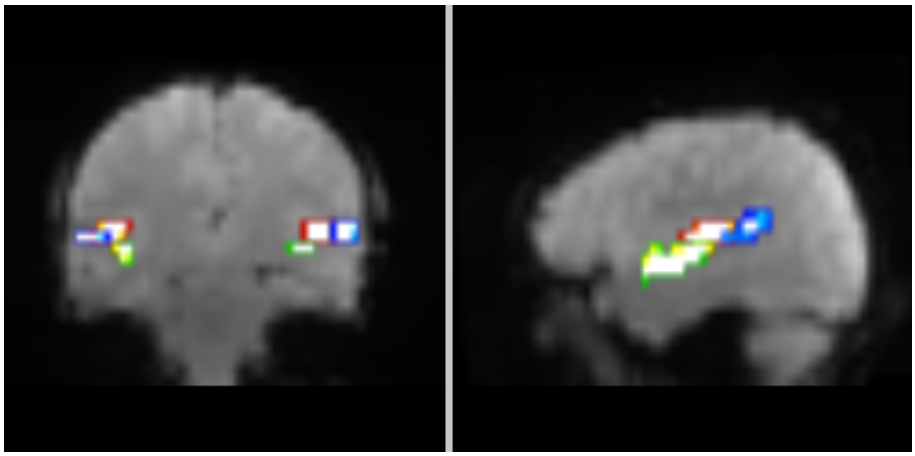

Fig. 3: Transformed ROI-masks, overlaid over the realigned mean functional image. Functional native space of sub-02 (color code: PP in green, HG in red, PT in blue).

The transformation resulted in masks of the ROIs (6 regions including PAC, PP and PT in both hemispheres; see subsection 1.5 of introduction) within a given individual participant's native, anatomical space. Finally, these masks were transformed into functional space by utilizing the FLIRT Linear Registration algorithm included in the FSL package (version 5.0.9). Hereof, the anatomical ROI-masks were priorly binarized via the Freesurfer function Binarize and overlaid across the subject-specific structural "brain.mgz" volume. Together with the ROI-mask overlays, this anatomical volume was then registered to the realigned mean functional image (see Fig.3). The hereby derived functional ROI-masks were later utilized to extract the BOLD signal time series in the respective regions (see subsection 2.6.3.1).

Detailed code concerning the whole procedure of functional preprocessing can be inspected on the mentioned repository on Github® (https://github.com/weissbe92/MSc_thesis_-BenediktWeiss).

2.6.2 First-level analysis

The implementation of a DCM-analysis requires the setup of a design matrix to define the corresponding regressors as its inputs. For this, a first-level analysis was priorly performed to compute a general linear model (GLM) containing the three conditions as regressors of interest. Hereof, the principle of a GLM is to explain the observed BOLD time course (dependent variable) for a given voxel as the sum of weighted explanatory variables (regressors of interest and regressors of no interest) plus a random error term.
As mentioned above, the experimental conditions were defined through a block by block presentation of the sound categories voice, singing or non-vocal music (see subsection 2.3). So as a first step, the time course of the experiment was split up into sections matching to the presented sequence of stimulus blocks. To achieve this, the onset, duration and containing stimulus modality of each block was extracted from the PsychoPy output csv-file and transcribed into a new tsv-file (BIDS standard) using the DataFrame class of the python-based Pandas library (http://pandas.pydata.org/, version 0.23). This file was further employed to define the presentation times of respective conditions. As to that, the timescales of the displayed stimulus blocks were sorted by category and concatenated into separated arrays. Altogether, the GLM design matrix specified 3 regressors of interest for the sound modalities voice, singing, music (each presented at 12 time points within one run and divided through times of no stimulus presentation with varying lengths) and 6 nuisance regressors for the motion parameters of the realignment (see preprocessing). Furthermore, one regressor unifying all three conditions and therefore depicting the nonspecific presentation of complex sounds had to be defined, which was subsequently utilized to specify the driving input for the DCM (subsection 2.6.3.2).
After deploying the design matrix, the GLM was estimated by convolving the time course of the respective conditions/regressors with the canonical hemodynamic response function (SPM algorithm). This computed an appropriate BOLD time course for the experimental procedure.
The whole first-level analysis was performed in a nipype workflow using various tools of the SPM interface. The smoothed functional images as well as the realignment-derived motion parameters from the preprocessing served as inputs to specify the model (SpecifySPM-Model; TR = 1.45 s, high-pass filter cut-off frequency = 1/128 Hz), generate the design matrix (Level1Design) and estimate the containing model parameters (EstimateModel). In addition, the previously generated brain mask (“functional ribbon mask”) was used to

restrict the GLM-analysis to voxels within the gray matter volume. Finally, contrasts were defined separately for the conditions voice, singing and music on the one hand, and for all conditions together on the other hand. These contrasts were further estimated using the EstimateContrast function of the SPM interface.
The corresponding output of the first-level analysis were 4 t-contrasted images of the respective conditions (3 modalities of complex sounds separated and one "all-conditions"-t-contrast) as well as one F-contrasted "effects of interest" image. Together with the obtained beta images (weight of regressors; multiplication with weighting factor), these results showed the degree to which the BOLD time course within each voxel could be explained by the different model regressors. The contrast- and beta images were further required to extract the BOLD signal time series within ROIs and the generated spm-matrix-file (SPM.mat) containing the time course of the experiment served to define the DCM-inputs via corresponding regressors. Detailed code of the first-level analysis can be inspected on Github® (https://github.com/weissbe92/MSc_thesis_BenediktWeiss).

2.6.3 Dynamic causal modeling

Effective connectivities between neuronal populations of the early auditory network were investigated by means of the method Dynamic Causal Modeling. This is a Bayesian framework, first introduced for fMRI data by Friston et al. (2003), which enables the examination of neuronal dynamics under the influence of experimentally controlled perturbations. Hereof, generative context-dependent connectivity models are constructed, which mirror causal interactions among cortical regions of interest (Friston et al., 2003). As to that, the modeling has to be strictly hypothesis-driven and depends on prior knowledge (i.e. on anatomical connections) to be able to infer on effective strengths of synaptic connections among neuronal populations and their context-dependent modulation (i.e. conditional changes of coupling strengths; Stephan et al., 2010). In this regard, the bilinear neural state equation (derived from the SPM12 manual; Ashburner et al. (2017)) describes the states of the examined brain network:

$$\dot{z} = Az + \sum_{j=1}^{m} u_j B_j z + Cu$$

The state changes $\dot{z}$ of a neuronal system are characterized through three parameter

matrices: A represents the endogenous connection strengths mediated by anatomical connections (context-independent, "resting state" connectivity), Bj shows the influences of the experimental manipulation uj on the endogenous connections (context-dependent, "modulatory" connectivity), and C quantifies direct external inputs (experimental manipulations u) into the system that drive regional activity (context-dependent, "driving" inputs). The distinction of external perturbations between "driving" and "modulatory" is neurobiologically relevant: driving inputs exert their effects through direct synaptic responses in the target region (i.e. a primary sensory area), whereas modulatory inputs change synaptic responses in the target region in response to inputs from another region (i.e. connection strengths between primary and non-primary areas; Friston et al., 2003). For this purpose, hypothesis-driven models for the dissection of neuronal state changes in primary and non-primary AC (effective connectivity between the bilateral ROIs HG, PP, PT) during the processing of complex sounds (experimental perturbations via the perception of voice-, song- and music-stimuli) were constructed. This approach was further used to investigate if the factors handedness and musical training evoke different connectivity patterns, meaning that the observed data in the recruited participant groups (right-handed non-musicians, right-handed musicians, left-handed non-musicians) fit to different models.

Prior of building the models, time- and context-dependent activation changes of voxels within ROIs had to be extracted to serve as a quantifiable measurement variable (BOLD signal time series). Furthermore, the regressors of the constructed design matrix (see 2.6.2) defined the external DCM-inputs ("driving" and "modulatory").

2.6.3.1 Time series extraction

As a pre-step for the DCM estimation, the BOLD signal time series within each of the six auditory ROIs (PP, HG, PT bilaterally) were extracted as the first eigenvariate of all activated voxels ($p < 0.05$, uncorrected). Hereof, the t-map contrasting all sound conditions together against the baseline served as input contrast to receive cortical voxel activations and the functional ROI-masks generated in preprocessing were used to match corresponding volumes of interest (VOIs). Also, the first-level derived SPM.mat along with the beta images defined the experimental time course and regressor weights. Finally, an "effects of interest" F-contrast was used for mean-correcting the time series and remove movement-related variance. An F-contrast contains several T-contrasts (in this case the 3 separately contrasted conditions voice, singing, music) and can be utilized for detecting

unspecific, stimulus dependent activation in general ("effects of interest": activation during stimulus presentation).

The time series extraction was performed using SPM12 (v5; Statistical Parametric Mapping, Welcome Center for Human Neuroimaging, London, UK; https://www.fil.ion.ucl.ac.uk/) and Matlab (Mathworks, Natick, MA, USA). The corresponding script can be inspected on the Github®-repository (https://github.com/weissbe92/MSc_thesis_BenediktWeiss).

The output was stored in VOI.mat-files for each region of interest and each run. The analysis also revealed, that the time series of 3 participants (sub-09, -10, -18; two right-handed non-musicians, one right-handed musician) couldn't be extracted, because they did not show strong enough voxel activations to reach the masking threshold ($p < 0.05$, uncorrected). Therefore, these subjects had to be excluded from the subsequent analyses.

2.6.3.2 Construction and estimation of model space

Analysing the effective connectivity between brain regions via DCM requires the setup of a hypothesis-driven model space. As to that, a balance must be found to construct a comprehensive depiction of the examined neuronal system and its reasonable states on the one hand, and to restrict the number and complexity of models to be tested on the other hand (Friston et al., 2003). In this respect, overly increasing the complexity of a model by expanding the amount of its free parameters would lead to a so called model "overfit" at a certain point. This means that an excessively complex model may fit best to the observed BOLD time series, but concurrently its explanatory power concerning hypotheses decreases. Instead, it will also increasingly fit to data set specific noise at the cost of the model's generalizability (Stephan et al., 2010). So the premise regarding to a model explaining the observed data best, is the optimal trade-off of accuracy and complexity to afford a maximal generalizability (Pitt & Myung, 2002). Taking this consideration into account and based on prior knowledge, a comprehensive model space was defined mapping the early auditory network to test the respective hypotheses.

Over all models, the endogenous connectivity and driving inputs of the neuronal system were identical (parameter-matrices A and C of the neural state equation; Fig.4). Within hemispheres, endogenous reciprocal connections between HG (core region Te1.0 of PAC) and adjacent regions of the superior temporal plane (PP, PT) were modeled, mirroring a parallel and hierarchical connectivity from the primary auditory core to its surrounding non-primary belt (Cammoun et al., 2015; Tardif Clarke, 2001; Upadhyay et al., 2008).

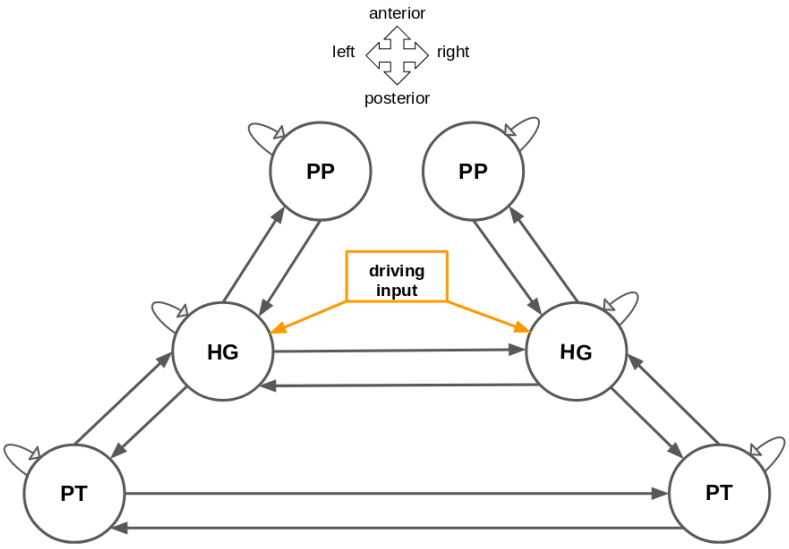

Fig. 4: Modeling of the neuronal network of interest. Endogenous connections (gray arrows; A-matrix) between ROIs (circles) and site of driving inputs (highlighted in orange; C-matrix) for all models.

Due to the fact that HG separates the anterior from the posterior STP (e.g., see Morosan et al., 2001) and that corresponding findings propose narrow fiber structures between neighboring areas contrary to connections among non-adjacent regions of non-PAC (e.g., see Cammoun et al., 2015), direct tracts between PP and PT were excluded. Furthermore, in consideration of imaging studies on transcallosal fiber tracts (Andoh, Matsushita, Zatorre, 2015; Elmer, Hänggi, Jäncke, 2016) reciprocal interhemispheric connections were set between homotopic regions of both hemispheres (HG ↔ HG, PT ↔ PT). No conjunction between bilateral PPs' was set due to lacking validation via corresponding literature data. Also, interhemispheric connections between heterotopic regions seem to be less pronounced in humans (Catani & Thiebaut de Schotten, 2008) and were therefore excluded. Finally, intrinsic functional connections for all ROIs in both hemispheres were comprised (Cha et al., 2016; Häkkinen & Rinne, 2018). Bilateral HG served as input region where the experimental manipulations entered the cortical network as driving inputs.

The previous setup of a GLM defined four regressors which were used to specify the external inputs for the DCM. Hereof, the primary sensory cortical area situated on HG is the site where all sound stimuli firstly enter the cortex (for a review, see Morosan et al., 2001; Saenz & Langers, 2014). Therefore, the regressor unifying all three conditions of complex sounds ("allcondition-regressor"; nonspecific presentation of the sounds voice, singing, music) was used to define the driving inputs. For the specification of modulatory

inputs (parameter matrix B of the neural state equation), the three conditions representing experimentally controlled manipulations via the presentation of music-, song-, or voice-sounds were allowed to modulate all intra- and interhemispheric endogenous connections as well as the intrinsic connections of bilateral HG.

The model space was then constructed by systematically varying the sites of modulatory perturbations, so that models differed with regard to their exact modulatory connectivity patterns. The differences between the competing models could be expressed by three factors: 1) the modulation of intrahemispheric connections could occur either in a parallel reciprocal manner from HG towards both plana (PP and PT) or in a non-parallel reciprocal way only in anterior direction towards PP or posterior direction towards PT 2) the perturbation on intrahemispheric connections could appear in context of a left-hemispheric unilateral, right-hemispheric unilateral or bilateral reciprocal modulation of endogenous connections 3) the interhemispheric structural connections between bilateral HGs' could be modulated in three ways: leftwards modulation from right HG to left HG, rightwards modulation from left HG to right HG, or bi-directional modulation between both HGs'. In consideration of these three factors, the model space finally consisted of 27 possible models differing with respect to their modulation sites. The third factor (interhemispheric effective connectivity) was also used to group the models into separated families (Penny et al., 2010), equally splitting up the model space into three model families á 9 models. The construction of families on the basis of interhemispheric modulatory effects aimed to enable a collective comparison to examine aspects of AC-lateralization. On that front, various assumptions exist concerning the processing of music and speech (e.g., see Cha et al., 2016; Tervaniemi & Hugdahl, 2003; Zatorre & Belin, 2001), musicianship (Angulo-Perkins et al., 2014; Elmer et al., 2016, 2013) and handedness (Knecht et al., 2000; Willems et al., 2014).The Figures 5 and 6 summarize the setup of the model space showing the three different model families with their crucial modulation sites (factor 3 respectively; Fig.5) and illustrating all 9 models included in family A as an example (based on intrahemispheric factors 1 and 2; Fig.6).

The estimation of the constructed model space was performed for each subject using the DCM12 toolbox implemented in SPM12 (v5). Hereof, the extracted BOLD signal time series stored in VOI.mat-files for each ROI and each run, and the SPM.mat containing the design matrix served as inputs. The detailed Matlab-script for the definition and estimation of the DCMs can be inspected on the Github®-repository (https://github.com/weissbe92/MSc_thesis_BenediktWeiss).

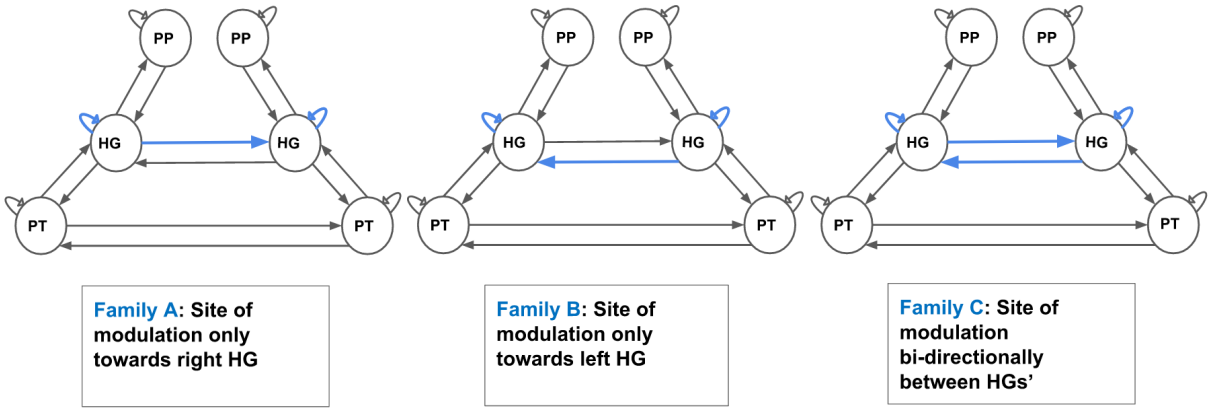

Fig. 5: Depiction of created model families. Crucial modulation sites (modulations of interhemispheric structural connections and intrinsic connections of HG) are illustrated as blue arrows.

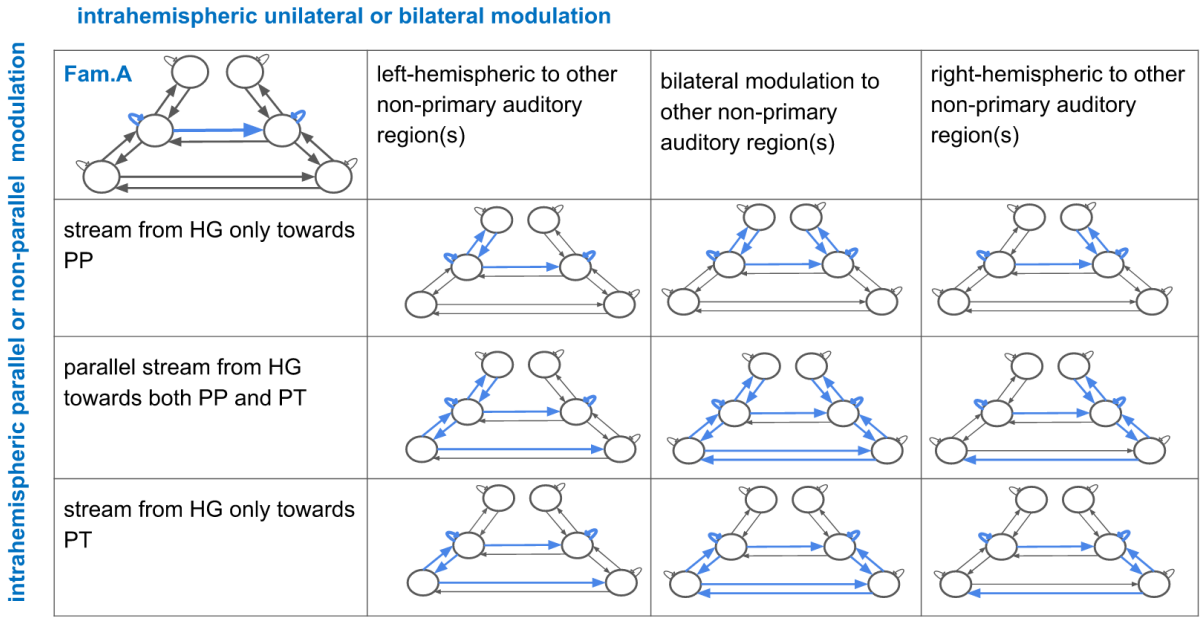

Fig. 6: Model space of family A with explanation of the respective modulatory connectivity patterns (factor 1 horizontally, factor 2 vertically).

2.6.3.3 Bayesian model selection and -averaging

As a next step, random effects Bayesian Model Selection (BMS) was used to compare models at the family level based on their negative free energy. Hereof, the output of the prior model estimation (a DCM.mat for each of the 27 models over each session and for each subject) was separately analysed for the three participant groups of right-handed non-musicians, right-handed musicians and left-handed non-musicians. Probabilities for each model and collectively compared for each of the three model families were received

for all participant groups in order to locate winning models and possible winning families (in terms of highest evidences; Penny et al., 2010).
Due to the fact that no obvious winning families were detected via BMS, the following random effects Bayesian Model Averaging (BMA) for the calculation of individual parameter estimates was performed across the whole space of 27 models.
The BMS/BMA was executed in SPM12 (v5) using the respective modules of the implemented DCM-toolbox and the output was automatically stored in a BMS.mat-file for each subject group. In order to perform the final analysis in python again, all parameter-matrices of interest (1 A-matrix and 3 B-matrices for each participant) were extracted from the BMS.mat by means of the freely available software GNU Octave (version 4.0.0, http://www.octave.org./).

2.6.3.4 Statistical analysis of parameter estimates

Inferences on effective connectivities between ROIs were conducted by the BMA-derived parameter estimates of endogenous- and modulatory connections. Hereof, the measured connectivity patterns were assessed across all participants, irrespective of their group membership, and separately for the 3 group-samples (right-handed non-musicians, right-handed musicians, left-handed non-musicians).
To determine whether parametric or non-parametric analyses were adequate on the observed data, the parameter estimates of the respective samples were priorly tested on normal distribution using the Shapiro-Wilk Test (Shapiro & Wilk, 1965). This normality test has a high statistical power, even for small sample sizes ($n < 50$). The test was conducted by utilizing the parameter-matrices of interest (1 A-matrix containing the parameter estimates of endogenous connections; 3 B-matrices including the parameter estimates of the modulated connections induced by the experimental perturbations "voice, singing, music") over all subjects and separately within respective groups.
Due to the fact of small, non-balanced sample sizes and non-normally distributed data, distribution-independent non-parametric tests were chosen for further analyses. First of all, the significance of each parameter concerning endogenous and modulated connections was tested by means of the one-sample Wilcoxon Test (Wilcoxon, 1945) at a statistical threshold of $p < 0.05$. Second, differences in the parameter estimates of respective connections were investigated between modulatory inputs and participant-groups using the two-sample Wilcoxon-Mann-Whitney Test (also "Mann-Whitney-U" Test, $p < 0.05$; Mann & Whitney, 1947; Wilcoxon, 1945). As to that, comparisons on parameter estimates of the

modulatory connections were performed regarding to the two hypotheses: 1) differences between effective connectivity patterns during the perception of either voice-, singing-, or music-stimuli were examined for the group of right-handed non-musicians 2) right-handed musicians and left-handed non-musicians we separately compared with the group of right-handed non-musicians in terms of effective connectivity patterns during the perception of respective sound categories.

All statistical analyses were performed in Python using the modules Pandas (http://pandas.pydata.org/; McKinney, 2010), Numpy (http://www.numpy.org/; Oliphant, 2006) and various statistical functions of SciPy (shapiro, wilcoxon, mannwhitneyu; https://www.scipy.org/; Oliphant, 2007). The detailed code can be inspected on the mentioned Github®-repository (https://github.com/weissbe92/MSc_thesis_BenediktWeiss).

3 Results

In this section, the evaluation of the MET and validation questionnaires is reported (3.1), followed by the description of the results from the implemented DCM analysis (3.2).

3.1 Validation of musicianship and handedness

The performance on the MET was assessed for participants, who were finally included in the DCM-analysis (exclusion of sub-09, -10, -18 after time series extraction; two right-handed non-musicians, one right-handed musician; see subsection 2.6.3.1 of materials and methods). For the first part of the MET, testing for melody-based musical competence, all subjects performed above chance level (right-handed non-musicians (n=9): M=35.56, SD=3.17; right-handed musicians (n=7): M=41.43, SD=3.46; left-handed non-musicians (n=4): M=30.75, SD=3.40; all values in absolute counts of correct answers on the 52 sound-pairs). Also the second part, testing for rhythm-based musical competence, fulfilled this criteria for all subjects (RH non-musicians: M=37.44, SD=3.97; RH musicians: M=41.29, SD=2.63; LH non-musicians: M=38.75, SD=5.32; all values in absolute counts of correct answers on the 52 rhythm-pairs). As a next step, accomplishment on the MET in total was evaluated (relative counts of correct answers on 104 stimulus-pairs; RH non-musicians: M=70.19%, SD=6.76%; RH musicians: M=79.53%, SD=5.47%; LH non-musicians: M=66.83%, SD=10.70%; see Fig.7) and two-sample t-test confirmed a significantly better performance of the right-handed musicians compared to the right-handed non-musicians (RH musicians vs. RH non-musicians: T=4.08, p=0.0003). Therefore, the MET validated a significant difference of musicianship between the recruited groups.

Concerning the validation of handedness, all of the self-reported right-handers achieved "Oldfield"-handedness-scores which confirmed them to the category of "pure right-handers" (reached minimum scores: +66.7; reached maximum scores: +100; for classification criteria, see http://zhanglab.wdfiles.com/local—files/survey/handedness.html). For the left-handers, the calculated scores labeled three subjects as "pure left-handers" (scores: -100, -89.5, -87.5) and one participant as a "mixed left-hander" (score: -11.1). "Waterloo" as a more detailed questionnaire on handedness was calculated the same way as "Oldfield" (sole difference is 34 questions on hand preferences instead of 10), which resulted in an identical classification of the participants, except that the "mixed left-hander" seemed tending more towards ambidexterity (reached score: +27.3). One right-handed participant

reported re-education to right-handedness, respective scores ("Oldfield":+66.7, "Waterloo": +57.9) classified him as a "pure right-hander".

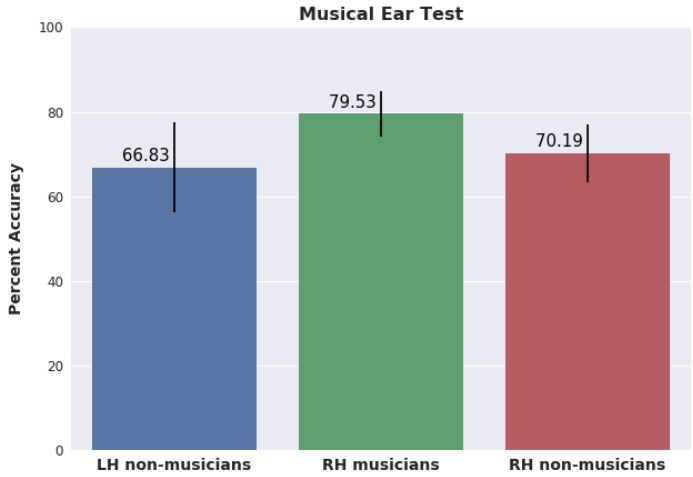

Fig. 7: Mean performance of participant groups on the MET in relative counts of accuracy. The respective standard deviations are visualized as vertical black lines.

3.2 Dynamic causal modeling

The following section describes the results of the DCM-analysis. Hereof, group-comparisons on model- and model-family estimates via random effects BMS are reported first (3.2.1), followed by the evaluation of effective connectivities based on the BMA-derived estimations of endogenous and modulatory parameters (3.2.2).

3.2.1 Family- and model-wise random effects Bayesian model selection

Comparing models at the family level revealed different winning families for the respective participant groups (see Fig.8), however none of the received exceedance probabilities gave rise to further analyses based on model families (winning family A for RH non-musicians: expected probability: 0.38, exceedance probability: 0.44; winning family C for RH musicians: expected probability: 0.50, exceedance probability: 0.72; winning family B for LH non-musicians: expected probability: 0.41, exceedance probability: 0.51). There seems to be a tendency for the musician group in preferring family A, which is at the same time the least likely family in the other two groups (exceedance probabilities 0.14 and 0.23 for RH- and LH non-musicians), suggesting a bi-directional modulation of the interhemispheric HG-connection in musicians as a response to the presentation of complex

sounds. Yet, the small sample sizes do not allow further conclusions so that the results of the model- and family-wise random effects BMS remain solely descriptive. For a more fine-grained inspection, also the particular estimates of all models are listed in Figure 8.

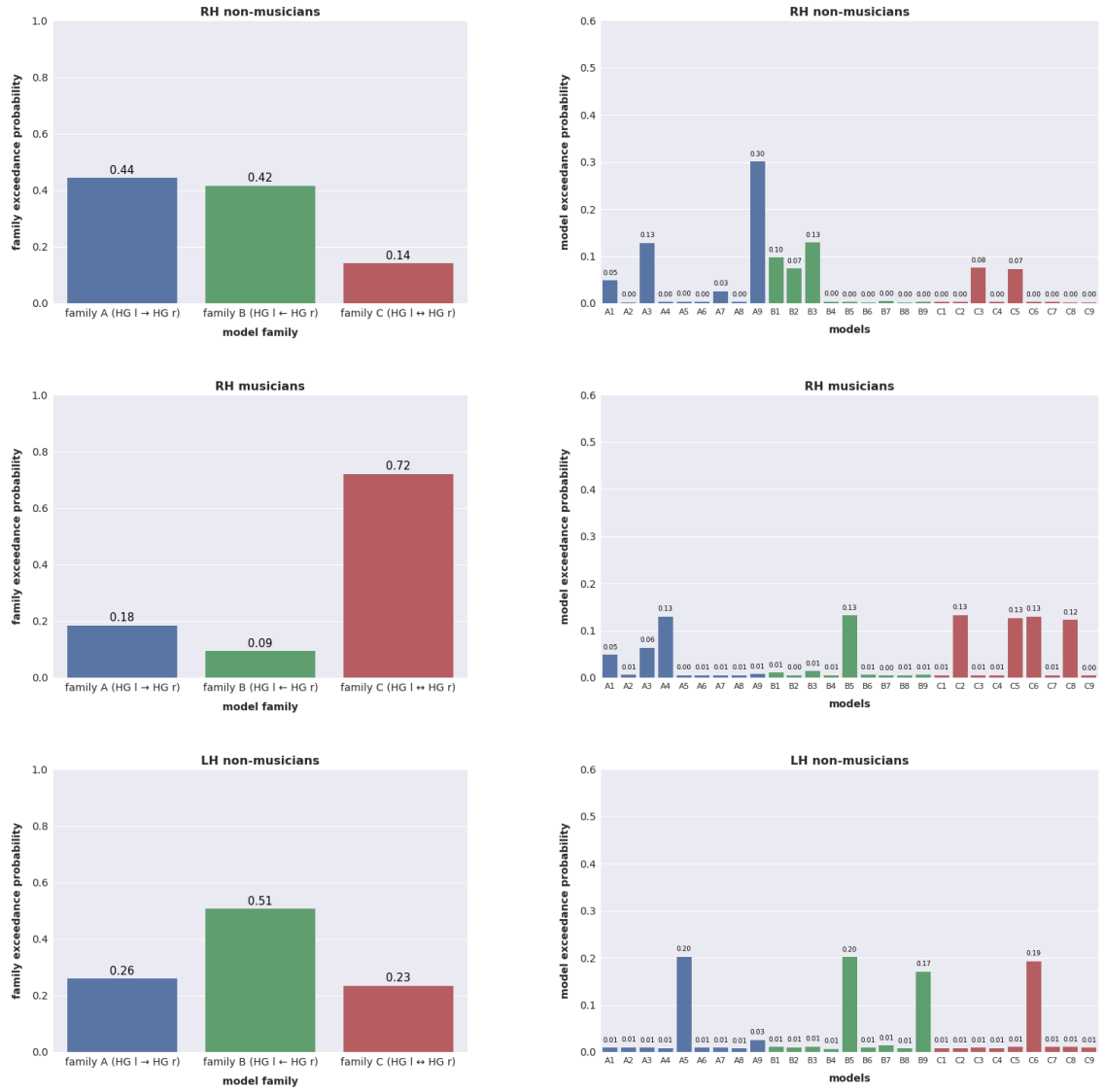

Fig. 8: Random effects Bayesian model selection, separately performed on the three participant groups of right-handed non-musicians (upper row), right-handed musicians (middle row) and left-handed non-musicians (bottom row). Received exceedance probabilities (entered on the y-axis) of the constructed model families (left column) and of all 27 models in detail for a more fine-grained inspection (right column).

Further analyses were restricted to the level of parameter estimates, calculated via random effects BMA across all 27 models, irrespective of any family affiliation.

3.2.2 Inferences on effective connectivities through estimated parameters

Effective connectivities between regions of the early AC were priorly examined on the basis of the estimated parameters of context-independent endogenous connections. Adding to the endogenous connectivity, the modulations of intra- and interhemispheric connections through the perception of the presented complex sounds (voice, singing, music) were analysed in order to answer formulated hypotheses.

3.2.2.1 Analysis of endogenous parameter estimates

Prior to further analyses, the measured parameter estimates of the endogenous connections were examined on normal distribution using the Shapiro-Wilk Test (Shapiro & Wilk, 1965). At subject-level (sample over all subjects, irrespective of group memberships), the parameter values of 2 out of 14 endogenous connections were normally distributed. Restricting the testing to group-level (separated samples of respective groups) revealed a tendency towards more connections with normally distributed parameter values (RH non-musicians: 7 out of 14 connections; RH musicians: 12/14; LH non-musicians: 12/14). The inconsistency of these results concerning data distributions and the fact that small sample sizes were used in the separated groups gave rise to further examinations via non-parametric test statistics.

In the following, summary statistics on endogenous parameter estimates were performed separately for groups and over all subjects together using the one-sample Wilcoxon Test (Wilcoxon, 1945) at a statistical threshold of $p < 0.05$. Estimations at the subject-level revealed significant excitatory connections from left HG towards the other auditory regions (intrahemispheric connections towards left PP/PT; interhemispheric connection towards right HG). The same connectivity pattern was observed for the right hemisphere (right HG towards right PP/PT and left HG; upper left network model of Fig.9; for a detailed inspection of parameter values, see also listing in Fig.10). None of the reciprocal connections from non-primary towards primary auditory cortex and none of the modeled intrinsic connections reached statistical significance. There was also no significance for the interhemispheric connection between PTs'. This pattern proposes a context-independent "resting-state"-connectivity of early AC without visible "feedback" couplings from non-PAC towards PAC, no visible intrinsic "self-regulation" of ROIs and no pronounced linkage between modeled regions of non-PAC (between PTs'). Analyses of separated participant groups displayed less significant connections compared to the examination over all subjects, hence the observed patterns were always restricted to connections that were also reported

at subject-level (3 out of 6 priorly reported connections in RH non-musicians: left HG towards left PP/PT, right HG towards left HG; 1/6 in RH musicians: right HG to right PT; no significance in LH non-musicians; for detailed parameter values on individual groups, see Fig.15,16,17 in appendix). It seems to be that the count of significant couplings decreases with decreasing sample sizes (n=20, n=9, n=7, n=4), but the pattern of coupled regions persists (e.g., the two next higher p-values of connections in RH-musicians did not reach significance, but were priorly reported as significant at subject-level: left HG $\rightarrow$ right HG, $p < 0.091$; right HG $\rightarrow$ right PT, $p < 0.091$).

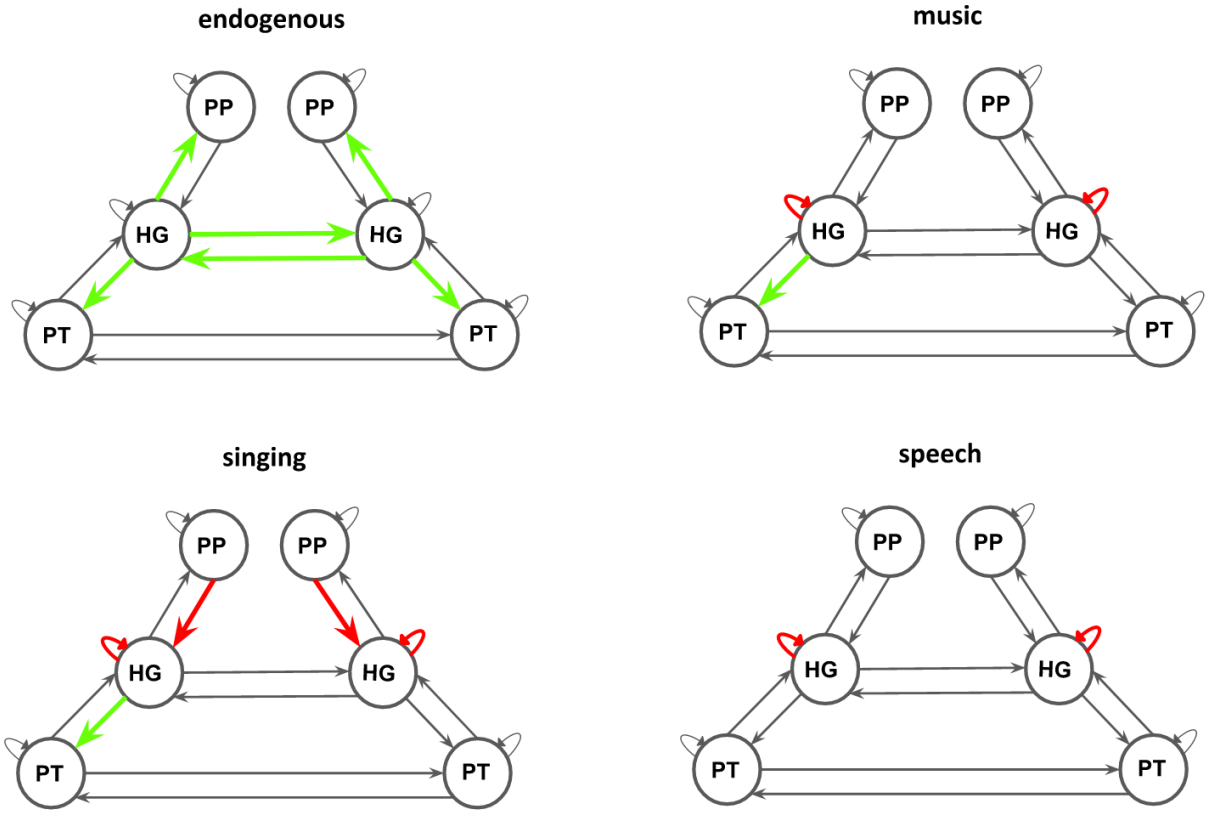

Fig. 9: Significant Context-independent endogenous- as well as experiment-dependent modulatory effective connectivities of the modeled auditory network, examined over all subjects, irrespective of musicianship or handedness. "Resting-state"of early AC in absence of experimental manipulations, shown in the upper left network model, and network states during the perception of music (upper right), singing (lower left), or speech (lower right). Green arrows depict an excitatory coupling, Red arrows an inhibitory linkage.

In order to achieve a comprehensive view on endogenous effective connectivities, differences in the respective parameter estimates were investigated between the participant groups by means of the two-sample Wilcoxon-Mann-Whitney Test (Mann Whitney, 1947; Wilcoxon, 1945) at a statistical threshold of $p < 0.05$. Except of one significant connection when comparing RH non-musicians versus RH musicians (increased modulation of connection

right HG → right PT in musicians; $p < 0.03$), no group-related differences in endogenous connectivity could be observed. This result hints to a similar context-independent effective connectivity in human AC irrespective of handedness or musicianship.

connection	all_subjects			
	mean	std	median	p_value
HG_l → HG_l	0.108119378843	0.416900155842	0.000186007195	0.41146474044
HG_l → PP_l	0.136079936781	0.334897470651	0.05267488998	0.033340220251
HG_l → PT_l	0.12319035463	0.463046817013	0.043276774781	0.047857511396
HG_l → HG_r	0.237913412847	0.33167361664	0.018807427692	0.00573399095
HG_r → HG_r	-0.074284586744	0.425444321304	-0.047962464365	0.204330247339
HG_r → PP_r	0.159846387642	0.306447603952	0.023606383863	0.020633435106
HG_r → PT_r	0.272402089511	0.43282755138	0.23078679266	0.007189168391
HG_r → HG_l	0.205931890137	0.327495945211	0.081246970828	0.008034428856
PP_l → HG_l	0.042622017905	0.559576596861	0.007534573085	0.793839045542
PP_r → HG_r	-0.167249338417	0.528850963689	-0.012248586309	0.262722246841
PT_l → HG_l	-0.053673369346	0.425719287024	-0.045959895205	0.478125471196
PT_l → PT_r	-0.05839579165	0.390195841647	-0.017072925888	0.217957265588
PT_r → HG_r	-0.035709900707	0.427835563439	0.00813427849	0.793839045542
PT_r → PT_l	0.052707696459	0.245278092191	0.008165967273	0.370261270453

Fig. 10: Results on parameter estimates of the endogenous connections after performance of summary statistics over all subjects together by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

3.2.2.2 Analysis of modulatory parameter estimates

Similarly to the endogenous parameter estimates, the measured parameter estimates of the modulated connections, induced by the experimental perturbations voice, singing and music, were examined on normal distribution using the Shapiro-Wilk Test. Conducted over all subjects, the parameter values of the respective connections were mainly not normal distributed (connections which displayed a normal distribution: 2 out of 14 for the modulation via music; 1/14 for the modulation via singing; 1/14 for the the modulation via speech). Restriction of sample sizes to respective participant groups revealed mixed results: For the right-handed non-musicians, most of the data were not normally distributed (normally-distributed parameters: 2/14 connections for “music”, 2/14 for “singing”, 1/14 for “speech”), whereas the parameter values of the musicians and left-handers mostly displayed normal distributed connections (RH musicians: 13/14 for music-modulation, 13/14 for singing-modulation, 10/14 for speech-modulation; LH non-musicians: 11/14 for music-modulation, 12/14 for singing-modulation, 12/14 for speech-modulation). Therefore, non-parametric test statistics were used for the following evaluation of modulated effective

connectivities in the early AC.

Summary statistics on modulatory parameter estimates were performed separately for groups and over all subjects together using again the one-sample Wilcoxon Test at a statistical threshold of $p < 0.05$. Hereof, the respective results are reported consecutively for the three different modulatory inputs of complex sounds.

connection	music_allsubs			
	mean	std	median	p_value
HG_l → HG_l	-0.688672016528	1.096098808484	-0.334199111469	0.018674874007
HG_l → PP_l	0.115165060088	0.380043403817	0.0	0.172954917988
HG_l → PT_l	0.341279454873	0.593200151526	0.006344674015	0.006406490222
HG_l → HG_r	-0.193390254971	0.588107922205	0.0	0.172848349135
HG_r → HG_r	-1.055795810795	0.859740887013	-1.209649737328	0.00018901207
HG_r → PP_r	0.107645148943	0.244415561919	0.0	0.085830958444
HG_r → PT_r	0.118505134445	0.468102222938	0.0	0.480176889906
HG_r → HG_l	0.040274251225	0.614137224754	0.0	0.641660126605
PP_l → HG_l	-0.018935575724	0.962317104752	0.0	0.806766322623
PP_r → HG_r	0.176778015716	0.743200322163	0.0	0.173070920805
PT_l → HG_l	-0.191118336487	0.774520922894	-0.00058148922	0.078292294146
PT_l → PT_r	-0.094467709386	0.534032299742	0.0	0.7764250048
PT_r → HG_r	-0.07770933082	0.521263584363	0.0	0.432767580668
PT_r → PT_l	-0.131942557134	0.532117474259	0.0	0.272095182736

Fig. 11: Results on the music-modulated parameter estimates of connections after performance of summary statistics over all subjects together by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

At subject-level, the perception of music modulated the intrinsic connections of bilateral HGs' in an inhibitory manner (left HG → left HG: $p < 0.02$; right HG → right HG: $p < 0.0002$). Also, the connection from left HG towards left PT displayed a significant excitatory modulation ($p < 0.007$). No further connections reached statistical significance (see upper right network model of Fig.9 and Fig.11 for a detailed inspection on conducted values). Restricting the analysis to the group-level entailed less modulated connections. For RH non-musicians, only the intrinsic connection of right HG showed a statistically significant inhibitory modulation ($p < 0.02$). This connection was modulated in the same inhibitory way for the sample of RH musicians ($p < 0.03$). Neither the RH non-musicians, nor the groups of RH musicians or LH non-musicians revealed further significant modulations. Hence, for the group of left-handers, the self-regulatory intrinsic connections of bilateral HGs' almost reached the threshold level ($0.05 < p < 0.07$) and seemed to display an inhibitory modulation (for detailed parameter values on individual groups, see

Fig.18,19,20 in appendix A).

Starting with the examination at subject-level, the presentation of song-stimuli negatively modulated the bilateral intrinsic HG-connections (left HG: $p < 0.0003$; right HG: $p < 0.002$) as well as the feedback-connections from PP towards HG in both hemispheres (left PP→left HG: $p < 0.02$; right PP→right HG: $p < 0.04$). Additionally, singing positively affected the connection of left HG towards left PT ($p < 0.03$; for a summary, see lower left network model of Fig.9 and and Fig.12).
Summary statistics restricted to the group of RH non-musicians revealed one significant inhibitory modulation of the intrinsic connection of left HG ($p < 0.03$) and one nearly significant negative modulation of the self-regulatory connection of right HG ($p = 0.05$). For the RH musicians, the estimated pattern displayed similar negative modulations of bilateral intrinsic HG-connections (left HG: $p < 0.03$; right HG: $p < 0.02$), but additionally the reciprocal connection from left PP towards left HG was also negatively modulated ($p < 0.05$). No significant modulations were observed for the separated group of left-handed non-musicians (for detailed parameter values on individual groups, see Fig.21,22,23 in appendix A).

connection	singing_allsubs			
	mean	std	median	p_value
HG_l → HG_l	-0.910775322368	0.971796337615	-0.748139361049	0.000293161548
HG_l → PP_l	0.094915781498	0.465433596512	0.000168492981	0.27870738325
HG_l → PT_l	0.351877892375	0.587061562538	0.000349576003	0.026756553624
HG_l → HG_r	-0.167426275612	0.770484125205	0.0	0.363487946982
HG_r → HG_r	-0.952965627459	1.079765262251	-0.833772408388	0.001507286841
HG_r → PP_r	0.095275764667	0.234294522308	0.0	0.138640633813
HG_r → PT_r	0.102433042019	0.476884635022	0.0	0.432767580668
HG_r → HG_l	-0.164841192703	0.558761642113	-2.5769013e-05	0.325870024297
PP_l → HG_l	-0.581639129513	1.115396362635	-0.000769406441	0.015906444102
PP_r → HG_r	-0.178369489706	0.394964239976	0.0	0.038151710173
PT_l → HG_l	0.129883301807	0.704460336642	0.0	0.532129888417
PT_l → PT_r	-0.003737381252	0.456391020668	0.0	0.864704576038
PT_r → HG_r	0.041396894483	0.405105080802	0.0	0.875329249148
PT_r → PT_l	-0.023386272408	0.303454992995	0.0	0.637870179898

Fig. 12: Results on the song-modulated parameter estimates of connections after performance of summary statistics over all subjects together by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

The modulatory input via speech did not reveal divergent modulation patterns between the subject groups. A first dissection over all subjects revealed that the intrinsic connections

of bilateral HGs' were modulated in an inhibitory manner (left HG: $p < 0.0001$; right HG: $p < 0.005$; see lower right network model of Fig.9 and Fig.13). Further restriction of analysis to separated groups showed the same pattern of negative intrinsic modulations, hence the intrinsic connection of right HG did not reach significance in RH musicians, but displayed the next higher p-value ($0.05 < p < 0.065$). In the same way, the bilateral intrinsic HG-connections did not fulfill the threshold criteria in left-handers, but showed the next higher p-values (left HG: $0.05 < p < 0.07$; right HG: $0.05 < p < 0.11$). For detailed parameter values on individual groups, see Fig.24,25,26 in appendix A.

connection	speech_allsubs			
	mean	std	median	p_value
HG_l → HG_l	-0.734668009129	0.836729409351	-0.324681058281	8.8574577e-05
HG_l → PP_l	0.021178409536	0.297432924278	0.0	0.972125329732
HG_l → PT_l	0.241780406936	0.593827765625	0.0	0.172848349135
HG_l → HG_r	-0.044430124429	0.677795292893	0.0	0.460301556379
HG_r → HG_r	-1.046549701621	1.336520495766	-0.863136609854	0.00404519705
HG_r → PP_r	0.057426385945	0.239600750087	0.0	0.213524354036
HG_r → PT_r	0.132406992904	0.577449047618	0.0	0.209426756961
HG_r → HG_l	-0.011461991967	0.602198098614	0.0	0.958760749938
PP_l → HG_l	0.275809279248	0.909254075094	0.0	0.196050963481
PP_r → HG_r	0.059899076696	0.646660431737	0.0	0.76709686841
PT_l → HG_l	-0.453299421432	1.194987226474	-0.001920482989	0.060892661901
PT_l → PT_r	-0.015932825978	0.403307868905	0.0	0.864704576038
PT_r → HG_r	-0.317772861267	0.724218818936	0.0	0.209426756961
PT_r → PT_l	-0.032373753656	0.266269901667	0.0	0.753683717726

Fig. 13: Results on the speech-modulated parameter estimates of connections after performance of summary statistics over all subjects together by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

The final analysis on the measured modulatory parameter estimates conducted hypothesis-driven comparisons of modulation patterns, between the experimental inputs of complex sounds on the one hand, and between respective participant groups on the other hand.
As for the first hypothesis, differences in modulatory effective connectivities during the perception of voice, singing or music were investigated by means of the two-sample Wilcoxon-Mann-Whitney Test (Mann Whitney, 1947; Wilcoxon, 1945; $p < 0.05$). Hereof, the group of right-handed non-musicians was used as a fixed sample for all contrasts between the complex sounds (music vs. speech, music vs. singing, speech vs. singing). For all comparisons, no significant category-related alterations of connection strengths could be observed. Therefore, according to the present data set, the first hypothesis if

the processing of complex sounds alters effective connectivities of the early AC, must be rejected.

In order to answer the second hypothesis, if musicianship and handedness have a influence on the processing of voice, singing and music in primary and non-primary AC, various contrasts were examined via two-sample Wilcoxon-Mann-Whitney ($p < 0.05$). Hereof, the measured parameter estimates of connections during the perception of voices, songs, or musical excerpts were each investigated between the different participant groups (9 contrasts in summary, 3 for each modulation: "music-RH-non-musicians" vs. "music-RH-musicians", "music-RH-non-musicians" vs. "music-LH-non-musicians", "music-LH-non-musicians" vs. "music-RH-musicians", "singing-RH-non-musicians" vs. "singing-RH-musicians",...).
Comparing RH non-musicians with RH musicians revealed that the musicians displayed a stronger modulation of the intrinsic connection in the right HG and a stronger modulation of the intrahemispheric connection between the left PP and left HG during the perception of singing (see Fig.14). Also, music modulated the intrahemispheric linkage from the right PP towards the right HG more strongly in musicians as compared to non-musicians. Modulation via speech displayed no significant differences between the mentioned groups.

Next, RH non-musicians were compared with LH non-musicians in context of modulatory inputs. Modulation via music revealed enhanced coupling strengths between right HG and right PT for the right-handers compared to the left-handers. By contrast, singing modulated the connection from right HG towards right PP and the feedback connection from right PT to right HG more strongly in the left-handed participant group. Also, the feedback connection right PP→right HG turned out to be more strongly modulated in left-handers then in right-handers after speech-modulation (for an overview, see Fig.14).
In order to achieve a holistic view, also differences between the musicians and left-handers were considered. As to that, the modulation through singing seemed to adjust the feedback connection between left PT and left HG more strongly in the group of LH non-musicians then in the RH musicians. Modulations either through music or speech did not reveal any further differences in coupling strengths for the group-comparison between left-handers and musicians. Due to the fact that this contrast is off-topic concerning our research questions, it will not further debated.

Taken together, the comparison of the connections' coupling strengths between the participant groups in the context of the perceived complex sounds (modulation through

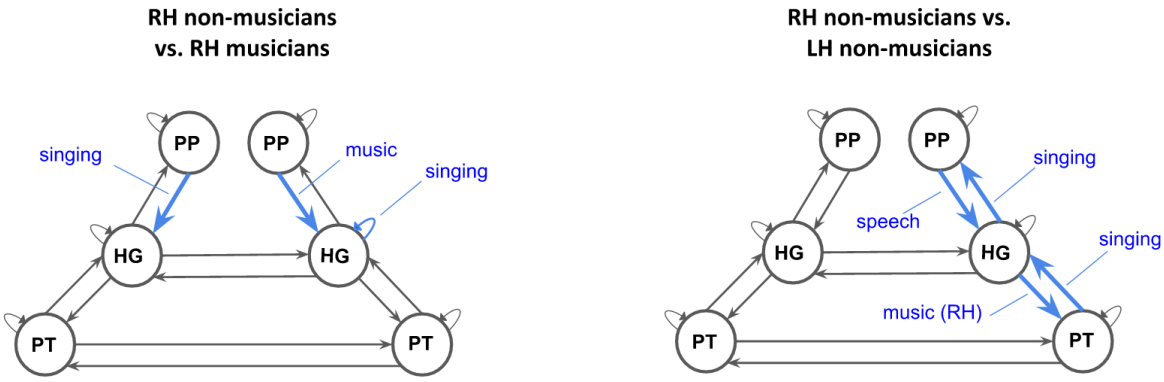

Fig. 14: Group comparisons of modulatory connections during the perception of complex sounds in order to examine influences of musicianship (left network model) and handedness (right network model). Blue arrows depict observed alterations of coupling strengths via respective sound-categories. For the left depiction, respective connections were stronger modulated in musicians. For the right depiction, respective connections were stronger modulated in left-handers, except for the HG → PT-connection, which was more strongly modulated in right-handed non-musicians (note the "RH" in brackets for "right-handers").

voices, songs, or melodies) displayed eight differences between designated regions in total. The results concerning modulation differences of specific modulatory connections between groups will be further discussed in the following section in terms of a general validity and assumptions of the current scientific literature.

4 Discussion

In the present study, we examined the neuronal mechanisms underlying the processing of complex sounds by investigating effective connectivities of primary and non-primary auditory cortex within and between hemispheres. Furthermore, the influences of handedness and musicianship on auditory system connectivities were investigated.

First of all, our results approve principle assumptions of an hierarchically organized AC, as is displayed by the observed pattern of effective connectivities between endogenous connections in absence of experimental perturbations. Hereof, "resting-state"-connectivity displayed excitatory interhemispheric connections between homotopic regions of PAC as well as an excitatory parallel stream from PAC towards anterior and posterior regions of surrounding non-PAC for both hemispheres. At the same time, no reciprocal coupling from non-PAC towards PAC could be observed.

The examination of modulated connections during the perception of music, singing or speech depicted rather ambiguous patterns. A general observation for the current data set was, that quite few connections were modulated by the presented experimental manipulations at all. In addition, the different sound categories caused similar coupling strengths of the modulatory connections. Therefore, our first hypothesis has to be answered such that no alterations in modulatory effective connectivities for the different auditory sound categories could be observed.

Handedness as well as musical training by contrast seemed to have an influence on the neuronal processing of the different sound stimuli. Hence, it is difficult to generally deduce other models of effective connectivity in left-handers or musicians from the observed patterns. As to that, the detailed results concerning modulation differences of specific connections between the groups must be further discussed in terms of a general validity and assumptions of the existing scientific literature.

The following section discusses the gathered results of the study and confronts them with respective findings of the current scientific literature. In respect of our research questions, the debate will be constructed around the following leading questions: What were the findings and do they confirm and extend current views? What could be possible reasons, why the present study was not able to reveal assumed operating principles of the auditory network? According to that, was the used experimental approach capable to contribute to the corresponding field of research?

In this respect, a general debate addressing the results on effective connectivities in terms of

sound categories and compared participant groups will be conducted (4.2). Following that, the experimental paradigm in respect of the stimuli and their kind of presentation will be recapitulated and compared with current paradigms in the field of auditory neuroimaging (4.3). Furthermore, the constructed model space and its capability to cover the respective research field and formulated hypotheses will be discussed (4.4). Finally, limitations of the study and starting points for further investigations will be depicted (4.5), followed by a final conclusion (4.6).
Prior to the general discussion, a short statement on the small sample size of the used data set and the statistical approach will be issued (4.1).

4.1 Small sample sizes and statistical adaption

Previous to further considerations, it should be mentioned that the gained results of the current study were established on data of a small sample. Especially when it comes to comparisons between the participant groups, one can criticize the insufficient and non-balanced sample sizes. For similar fMRI-based DCM studies, which infer on effective connectivities via comparisons on estimated parameters, it is common to recruit at least 20 volunteers per group (e.g., Frässle et al. (2016); also see recommendations of Goulden et al. (2012)). As to that, it was initially planned to include more than the current amount of subjects to reach an adequate sample size for our hypotheses. Hence, a long-lasting breakdown of the MRI scanner and the following replacement of a broken gradient coil prevented this ambitions. A new gradient coil entails that further measurements would have occurred on a different scanner device. This aspect has to be taken into account when it comes to the comparability of the gathered data before and after the replacement. Hereof, proceeding the recruitment under such circumstances would have meant to include even more volunteers to preserve data comparability. Concerning capabilities of time and money, it was decided to close the recruitment and proceed with the present data set. As a consequence, non-parametric test statistics were adapted in order to meet the requirements of small and unequal sample sizes and non-normal distributed data (http://blog.minitab.com/blog/adventures-in-statistics-2/choosing-between-a-nonparametric-test-and-a-parametric-test; Arnold, 1965; McCrum-Gardner, 2008; Zimmerman, 1987). That kind of statistical adjustment enabled us to still exploit scientifically interpretable findings from the available data. Hence, the gathered results have to be considered under the viewpoint that the statistically evaluated data were finally not corrected for multiple comparisons.

4.2 Insights into effective connectivities across the auditory cortex

The following subsection discusses the conducted results on modulatory effective connectivity patterns in terms of the sound categories in general (4.2.2) and the observed influences of musicianship and handedness in detail (4.2.3, 4.2.4). Prior to that, also the conducted results on the endogenous connections will be debated (4.2.1).

4.2.1 A similar pattern of context-independent connectivity between participant groups

A general observation was that when pooling the participant groups together and therefore conducting summary statistics across a greater sample of 20 subjects, more endogenous connections displayed significant couplings compared to separated group samples. Hereof, a connectivity pattern could be observed, where the primary auditory cortices displayed excitatory interhemispheric connections towards their homotopoic counterparts on the one hand, and excitatory intrahemispheric connections towards respective non-PAC areas (PP, PT) on the other hand (see Fig.9 of the results). This coupling pattern persisted for the analysis on separated groups. Hereof, less connections reached statistical significance, probably due to smaller sample sizes, but no other connections then the ones displayed for all groups together showed pronounced linkages. Additionally, non-significant connections of the supposed pattern mostly displayed the next higher p-values in the right-handed groups. Solely the analysis restricted to the sample of left-handed non-musicians seemed to be too small to derive any valid conclusions.

Comparing endogenous coupling strengths between the participant groups revealed no differences, except for the connection from right HG towards right PT. This linkage showed a stronger coupling in RH musicians compared to RH non-musicians. From a first point of view, one could deduce some kind of rightwards asymmetry from that finding. Hence, this would be an obstacle to the common position, that musical training causes more pronounced asymmetries towards the left hemisphere, especially the left PT (Elmer et al., 2012; Elmer et al., 2013; Meyer, Elmer, & Jäncke, 2012; Schlaug, Jäncke, Huang, & Steinmetz, 1995). Apart from this difference, which must be considered critically, the results hint more to a similar context-independent effective connectivity, irrespective of handedness or musicianship. Such steady coupling strengths point out that increased structural connectivities, which are for example found in musicians (Elmer et al., 2016;

Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995), do not automatically come along with an altered neuronal communication in absence of auditory cues.

The revealed pattern of endogenous effective connectivities reflects the underlying anatomy of the early auditory network and approves principle assumptions of an hierarchically organized AC (e.g., Cammoun et al., 2015; Häkkinen & Rinne, 2018; Tardif & Clarke, 2001). As to that, parallel streams from HG towards the anteriorly situated PP and posteriorly situated PT as well as an interhemispheric transfer between bilateral HGs' are shown (e.g., Andoh, Matsushita, & Zatorre, 2015; Upadhyay et al., 2008). Furthermore, the to our knowledge novel investigation of endogenous effective connectivities within AC extends previous views based on functional (resting-state) connectivity analyses. Hereof, the auditory regions were connected in an excitatory manner and it seemed to be that neither "feedback"-couplings from non-PAC towards PAC, nor intrinsic "self-regulation" of ROIs, nor linkages between regions of non-PAC (interhemispheric connection between PTs') were pronounced in absence of stimuli. Still, the findings on the modulatory connections relating to the experimental manipulations are of more interest in terms of our research questions and shall be further debated.

4.2.2 Similarities in the neuronal processing of complex sounds

Adding to the observed endogenous connectivity, only a limited number of intrahemispheric- as well as intrinsic connections were modulated at all. For a first overview, significant couplings between ROIs were evaluated by including the factors handedness and musicianship (summary statistics over a sample including all participants). This revealed inhibitory modulations of the self-regulatory connections within bilateral HG for all complex sound categories, excitatory unilateral modulations from left HG towards left PT during the perception of music as well as singing, and a bilateral inhibitory modulation from PP towards HG for the sound category singing (see Fig.9 of the results). As for the endogenous connections, the modulatory couplings were analysed within separated groups, with the consequence that less connections reached statistical significance. Also, no other than the mentioned connections were displayed. Again, this observation may be caused due to the fact of small sample sizes for the separated groups, especially for the LH non-musicians, who showed not a single one significant connection.
To investigate alterations of effective connectivities for the processing of complex sounds, the received modulation patterns during the perception of music, singing or speech were

compared using the group of RH non-musicians as a sample to control for handedness and musicianship. The result was that no differences in coupling strengths between the sound categories could be observed, which hints towards a similar neuronal processing of complex sounds within the modeled auditory network without any differences between hemispheres.

The aim of this study was to provide a mechanistic explanation of how complex sounds are being processed within the early auditory cortex and therefore extent current findings, who investigated functional activation patterns within the auditory domain. Hereof, activation differences were found for music and speech (Angulo-Perkins et al., 2014; Rogalsky et al., 2011; Whitehead & Armony, 2018) as well as for their underlying acoustic features (Norman-Haignere et al., 2013; Santoro et al., 2014), whose cortical representations partially overlapped with their respective categories (Norman-Haignere et al., 2015). As to that, it is compliantly reported that especially the bilateral PP (or a "medial anterior region of the STP", according to Rogalsky et al., 2011) preferably processes music and music-related features. Therefore, an altered recruitment of this region during the perception of music (and maybe singing as its vocal form) as compared to speech could give a mechanistic explanation of these functional findings on the basis of preceding effective couplings. Concerning such effective connectivities, this study did not find any evidence, if the PP is indeed the location of a music-specific categorial integration. Hence, it has to be taken into account that only few modulatory connections were significantly displayed in general.

Nevertheless, one could also reconsider the functional properties of the PP and PT. As "belt"-regions being adjacent to PAC and both receiving direct inputs from this auditory core, they clearly contribute to the hierarchical processing of complex sounds at a fundamental stage (Cammoun et al., 2015; Upadhyay et al., 2008). Following that, these early AC-regions could also be seen as computational hubs for the segregation and matching of spectrotemporal sound patterns, deriving acoustic structures from incoming sounds, which are then sent downstream to "higher order" regions, where a final integration towards a categorial perception proceeds (Griffiths & Warren, 2002). Hereof, findings which report multiple cortical representations of sound spectrograms with different degrees of spectral and temporal resolution, situated around PAC, support such a proposal (Santoro et al., 2014). Therefore, the PT and PP could contribute to an early non-primary network, both delivering different "views" on an incoming sound in terms of its spectrotemporal

modulations, what further results in a distributed sound-feature pattern across a network of hubs. According to that, thinking of both plana as intermediary nodes with a more distributional than integrative function could explain the observed sparse modulatory coupling between this regions and PAC without any differences between sound categories.

Based on their findings on cortical speech-, respectively music-selectivity, Norman-Haignere et al. (2015) also point out that speech-selectivity is more anatomically segregated while music-selective locations seem to be more distributed across the neuronal network. This could also be a reason, why we were not able to detect more pronounced couplings towards a specific non-primary region during the perception of music. As to that, one must also keep in mind that that the PP as well as PT are based on macro-anatomical landmarks (for a review, see Moerel, Martino, Formisano, 2014). Therefore, when choosing ROIs to be examined, a more fine-grained, possibly cytoarchitectonic subdivision of the anterior and posterior STP (similar to PAC concerning subregion Te1.0) could give a more detailed view into the early non-primary processing of sound categories and underlying acoustic features.

Our results did also not find any evidence for a lateralized music-, respectively speech-processing within the modeled early auditory network. Speaking of asymmetries regarding the processing of music and speech in general basically refers to more widespread activation differences between the hemispheres (for a review, see Hickok Poeppel, 2007; Tervaniemi & Hugdahl, 2003). On a more detailed inspection of the AC, Zatorre and Belin (2001) reported a left-lateralization for the fine-grained processing of temporal features and a right-lateralization for the detailed processing of tonal attributes. They further propose that this early lateralization of basic acoustic features causes the observed left-lateralization of speech, respectively right-lateralization of music at higher levels (Zatorre et al., 2002). Hence, it has to be mentioned that the usage of PET for their study did not achieve a detailed spatial resolution. Recent higher resolved fMRI-based investigations did mainly not find any evidence for their proposal, neither for music and speech in general (Rogalsky et al., 2011; Whitehead & Armony, 2018) nor for respective acoustic features (Norman-Haignere et al., 2013; Santoro et al., 2014), at least not at this early stage of the processing hierarchy. Therefore, the missing of any hemispheric bias in our results suits compliant findings across various scientific studies.

As a last point, it is also possible that the modeled neuronal network did not represent regions for a pronounced speech integration. As to that, likewise to the mentioned

selectivity of music in bilateral PP, the abovementioned studies also compliantly report a speech-selectivity for the lateral STG as well as underlying structures of the STS and MTG (Angulo-Perkins et al., 2014; Norman-Haignere et al., 2015; Rogalsky et al., 2011; Whitehead Armony, 2018). These regions were not integrated in the construction of the comprehensive model space. Reasons for that were the restriction of model numbers and their complexity. A detailed debate concerning the construction of the hypothesis-driven model space is conducted in subsection 4.4.

4.2.3 Influence of musicianship on the processing of complex sounds

The analysis of modulatory connectivities during the perception of voice, singing or music revealed significant differences between the right-handed musicians and non-musicians. Hereof, the musicians displayed a stronger modulation of the intrinsic connection in the right HG as well as a stronger modulation of the intrahemispheric connection between the left PP and left HG during the perception of singing. Also, music modulated the intrahemispheric linkage from the right PP towards the right HG more strongly in musicians as compared to non-musicians (see Fig.14 of the results). No further differences were shown, especially not for the modeled interhemispheric connections between the homotopic PT- or HG-regions. It was also striking that none of the observed differences in coupling strengths included the PT. As to that, the results do not support an altered recruitment of the PT in musicians as compared to non-musicians, for none of the sound categories.

Structural and functional differences in the AC of musicians are reported for the PT, especially in the left hemisphere. Hereof, the left PT exhibits an structurally increased surface and functionally enhanced activity during the perception of phonetic cues (Elmer et al., 2013, 2012). Also, the interhemispheric transcallosal connection between homotopic PT-regions and the size of the Corpus Callosum in general seems to be increased in musicians (Elmer et al., 2016; Schlaug, Jäncke, Huang, Staiger, et al., 1995).
According to such findings, it is deduced that musicianship alters performances of the auditory network not solely in the music-, but also in the speech-domain and that music may be processed more similarly to speech (e.g., Angulo-Perkins et al., 2014). Once the PT contains neuronal clusters with a preference for fast temporal modulations (e.g., Santoro et al., 2014)and coincidentally, this region depicts an increased structure and enhanced activity in musicians, it can be considered that such neuroplasticity relates to

musical training in order to achieve a better temporal resolution of incoming musical cues.

The reason why we did not find any evidence for an enhanced recruitment of the PT may rely on the fact that this study examined amateur musicians. All of the abovementioned investigations included only professional musicians (Elmer et al., 2016, 2013, 2012), respectively defined more conservative criteria for musicianship (e.g., daily practicing and "formal studies"; Angulo-Perkins et al., 2014). On the other hand, performances on the MET clearly showed a difference of musicianship between the participant groups (see Fig.7 of results). The decision to perform respective investigations using amateur musicians was made, because it highlights a more universal picture of musicianship across the common population. Still, the obtained results leave the question open, if the mentioned findings are still found in terms of a more liberal definition criteria on musicianship.

4.2.4 Influence of handedness on the processing of complex sounds

Our analyses revealed significant differences between the right- and left-handed non-musicians for a couple of modulatory parameters. From a first point of view, it was striking that only the right hemisphere showed differences in connection strengths, while no alterations were displayed in the left hemisphere or for the interhemispheric connections during the perception of the complex sounds (see Fig.14 of the results). As to that, the left-handers displayed stronger modulations for the connections from the right HG towards the right PP as well as from the right PT towards the right HG during the perception of singing. Furthermore, speech modulated the connection from the right PP to the right HG more strongly in left-handers. By contrast, the right-hemispheric connection from HG to PT was more strongly modulated in the right-handers for the experimental manipulation via music.

Apart from the observation of a stronger modulation in right-handers for the right HG→PT-connection, these findings could hint towards a more pronounced right-hemispheric processing of sounds in left-handers. Admittedly, the respective differences in modulatory connection strengths were never triggered by more than one sound category, but singing combines acoustic properties of both instrumental music and speech (Schön et al., 2010; Whitehead & Armony, 2018), and therefore one could see the displayed pattern as a more general modulation via incoming "sound". Still, the observed differences were only sparse and not consistent over multiple sound categories to draw any valid conclusions. Also, the group of left-handers comprised just four people, what must be taken into account

when it comes to a more general evaluation regarding the validity of the reported results.

The examination of handedness in terms of strategies for the sound processing within early AC served an exploratory manner. It was initiated due to the fact that various handedness-related influences are reported concerning the lateralization of cognitive functions and their neuronal networks, which hints towards a fundamental linkage between handedness and brain lateralization (Willems et al., 2014). Regarding the auditory domain, the usually observed left-lateralization of the language network seems to be increasingly shifted towards a bilateral or even right-lateral processing (e.g., Knecht et al., 2000). Hereof, the term "language network" generally serves a holistic view on the brain and respective lateralization-studies do often not further dissect brain regions in detail, but refer on widespread activation differences between the hemispheres. As to that, The examined ROIs of such studies were coarse subdivisions of lobes (e.g., Pujol, Deus, Losilla, & Capdevila, 1999; Szaflarski et al., 2002), or even no detailed regions at all, but only investigations on BOLD signal differences between the middle cerebral arteries of hemispheres (Knecht et al., 2000). Therefore, current views lack of detailed knowledge at which point of the auditory network lateralization is introduced, and if divergent observations in left-handers are reasoned by the fact that hereof no hemispheric shifts or even opposite shifts occur. The non-verbal "sound-processing network", which is examined in this study, precedes a widespread, multimodal "language network" and can be seen as the earliest step of the processing hierarchy towards semantic language. Concerning this early cortical network, also hemispheric differences have been reported in terms of speech-, respectively music-related acoustic features, which in the following could cause a "whole brain" left-lateralization of speech, respectively right-lateralization of music (Zatorre et al., 2002). While our results did not find any evidence for a categorial lateralization at such early time points of the processing hierarchy in right-handers, it could still be that left-handers adopt an other strategy (e.g., a pronounced right-lateralization of early sound processing). Such proposals are quite newish and are jet not further examined. Our general aim was to add knowledge towards cognitive strategies in left-handers that differ from the right-handed majority by looking at the early AC. As to that, more scientific knowledge must be gathered to draw further conclusions.

4.3 Recap of the experimental paradigm

The following section discusses the experimental paradigm of this study concerning its capability for the research questions and performs a comparison with designs used in current investigations of the same research field. Hereof, the focus lies on the used auditory stimuli and their presentation via passive listening.

4.3.1 The stimulus set

As mentioned, the used stimulus set was originally assembled for the study of Whitehead and Armony (2018). Based on their findings using these stimuli, the authors claimed that the detected music-, respectively voice-selective regions were in fact encoding "higher-order" features towards a categorial perception of "music" (instrumental as well as vocal) and "voice" (including verbal speech and non-verbal vocalizations). As to that, enhanced activations to melodies (played on various instruments) and vocal "a capella" singing were displayed by the same areas. In the same manner, speech (non-verbal and verbal in various languages) and the vocal shares of "a capella" singing showed selective activations in consistent regions. Hereinafter, reasons are depicted why the composition of this set serves such research questions on the AC-processing of complex sounds as was also interrogated in this study.

First, the stimulus set includes an intermediary category "singing" which enables to confirm categorial music- respectively voice-selectivity. By solely using instrumental melodies for music-stimuli, one cannot exclude the possibility that eventually displayed regions just encode a basic acoustic feature like instrumental timbre (Leaver & Rauschecker, 2010). For the same reasons, the sole presentation of verbal speech does not completely cover a superimposed category "voice", which also includes non-verbal aspects of sound-processing in speech and the vocal shares of singing. In line with that, a second advantage of this set is the recording of phrases and songs in various languages and the inclusion of completely non-verbal chants and "baby talk". In this respect, our investigation of early AC-areas aimed for non-semantic sound processing preceding any semantic or syntactic analyses of language at higher levels of the processing hierarchy. In assurance that most of the stimuli were not linguistically understood by the participants, overlapping semantic biases could be excluded. Also, the recordings of melodies consist of excerpts being played on various instruments to make sure that diverse instrumental timbres are included and no specific bias towards a certain instrument is introduced.

The third advantage of the stimulus set directly concerns the aim of our study to investigate possible alterations of connectivity patterns across regions of the early AC during the processing of complex sounds. By containing voice-, singing-, and music-stimuli, it covers a broad range of acoustic features underlying complex sounds. As to that, it would have been possible to describe a separable processing framework via modulation changes of composed auditory features across the respective sound categories. Furthermore, this would have been achieved apart from the use of elaborate scrambling approaches (e.g., Rogalsky et al., 2011; Tervaniemi et al., 2006) inducing additional biases, or even more diverse sets using a broad range of "everyday" natural sounds (e.g., Norman-Haignere et al., 2015; Santoro et al., 2014), which would have induced various other factors to be considered. To sum it up, the chosen set was generally capable of our research questions, even if we were not able to visualize such a processing framework on the basis of altered effective connectivities.

4.3.2 Passive listening versus performance of active auditory tasks

In order to minimize possible confounding factors during the presentation of the stimuli, volunteers of our study didn't have to perform any additional active task, but solely listened in a passive manner. In this respect, it is not uncommon to include such context-independent active performances to ensure that subjects are paying equal attention over the whole experimental presentation. For example, in the study of Norman-Haignere et al. (2015) participants had to detect sporadic changes in sound levels by pressing a button when they heard a quieter sound. For the same reasons, the study of Angulo-Perkins et al. (2014) asked to press a button every time a pure tone was heard, which was presented at random during their paradigm.

Other studies by contrast (this one included) resigned from such additional active task performances (e.g., Rogalsky et al., 2011; Whitehead & Armony, 2018). The reason hereof is that various researches confirmed the influence of auditory tasks on activations within the AC besides stimulus-dependent modulations during auditory perception (e.g., Grady et al., 1997; Petkov et al., 2004; Woodruff et al., 1996; Woods et al., 2009). Such active tasks increase cognitive demands, resulting in additional recruitments of neuronal populations within and around the auditory cortex. Hereof, it has been shown that auditory task performances influence the magnitude of modulations in medial regions along the STP (HG included) and STG. Also, larger areas of activation were displayed during the sound processing, especially in lateral regions along the STG. As to that,

Petkov et al. (2004) identified specific "attention-independent" auditory fields situated on medial regions of AC (equally distributed attention) and "attention-dependent" fields in lateral regions of AC, showing enhanced activity when an additional performance on received stimuli was required (selective attention, driven by task).

As mentioned, the inclusion of an context-independent active task results from the fact that one has to assure that participants pay attention to the presented auditory stimuli. In this respect, studies preferring a passive listening paradigm still have to consider this aspect. Therefore, while passively listening to the stimuli, the subjects in our study watched a muted nature documentation to ensure that they were paying equal (non-selective) attention to the sounds. The fact that auditory and visual stimuli were hereof presented at the same time, was considered to be negligible, because attentional modulations of the early auditory-, respectively early visual cortices are considered to be modality-dependent (e.g., Woodruff et al., 1996). So, in this case we do not assume that the visual presentation had any further influence on the auditory domain of interest, even if selective shifts of spatial attention were initiated during the proceeding scenes. Also, the presented visual material was priorly cleaned from too arousing fighting- and mating scenes to minimize emotional content. Still, we can not fully exclude the possibility that any covered selective attention towards a certain stimulus or stimulus-category in the auditory domain occurred. Nonetheless, it is unlikely that such covered shifts of attention could have had too much influence within our chosen ROIs. As to that, Petkov et al., (2004) reported that medial AC-regions, which also cover our ROIs, are attention-independent during auditory perception.

As a last point concerning the capability of our paradigm it has to be mentioned that a recent meta-analysis across 115 fMRI-studies, which examined the role of attention on the processing of auditory stimuli (including voice), concluded that no additional areas in AC were recruited in active, compared to passive listening conditions (Alho, Rinne, Herron, & Woods, 2014). This finding can be further confirmed, when already mentioned studies are consulted. As to that, irrespective of the inclusion of an active task (e.g., Angulo-Perkins et al., 2014; Norman-Haignere et al., 2015) or sole passive listening (e.g., Rogalsky et al., 2011; Whitehead & Armony, 2018), all of these researches conducted overall comparable results. In this respect, we believe that the used paradigm was capable of our research questions.

4.4 The hypothesis-driven construction of model space

The results of the present study have to be evaluated under the condition that investigating effective connectivities always requires some model depicting the states of a neuronal system to be examined (Friston, 1994). In this sense, the visualization of effective connectivity patterns and their validity depend on the constructed model space covering anatomical as well as functional assumptions relating to the neuronal network (Friston et al., 2003). As to that, the setup of such a model space in turn must be considered carefully around prior scientific findings on structural and functional relations. Concurrently, a balance must be found to construct a comprehensive depiction of the examined neuronal system and its reasonable states on the one hand, and to restrict the number and complexity of models to be tested on the other hand (Stephan et al., 2010).

Due to the fact that DCM-based investigations of effective AC-connectivities were, to our knowledge, not conducted yet, the hypothesis-driven construction of the current model space was founded on a vast amount of prior functional findings. Hereof, especially the bilateral PP and, less consistently, the bilateral PT depicted a categorial music-, respectively music-feature selectivity (Angulo-Perkins et al., 2014; Norman-Haignere et al., 2013, 2015; Rogalsky et al., 2011; Santoro et al., 2014; Whitehead & Armony, 2018). Also, the same non-primary regions, here again particularly the (left) PT, but also the PP, are repeatedly mentioned when it comes to alterations within AC through musical training (Angulo-Perkins et al., 2014; Elmer et al., 2016, 2013, 2012). These spatial accordances enabled us to examine our two research questions restricted to a manageable number of ROIs. Therefore our constructed model space is obviously "hypothesis-driven".
Furthermore, the actual model space is, in our opinion, valid as we carefully considered the setup of intra- and interhemispheric connections between the ROIs based on various prior anatomical assumptions (Andoh et al., 2015; Cammoun et al., 2015; Catani Thiebaut de Schotten, 2008; Cha et al., 2016; Elmer et al., 2016; Häkkinen Rinne, 2018; Morosan et al., 2001; Tardif & Clarke, 2001; Upadhyay et al., 2008). With respect to the model-dependency of effective connectivity-based examinations, hereof conducted results would also have had a certain validity.

One critical point concerning the selection of ROIs and their integration into a comprehensive model space should be mentioned. As to that, we did not include a ROI along the ventrolateral STG, underlying our more dorsomedially located ROIs on the STP. As we also examined the effective processing of voice within AC, one could bring up

that this region is vastly reported with respect to categorial voice-, respectively speech-selectivity (Angulo-Perkins et al., 2014; Norman-Haignere et al., 2015; Rogalsky et al., 2011; Whitehead & Armony, 2018). Yet concurrently, voice-feature-, respectively speech-feature selectivity is also reported for the PT (e.g., Elmer et al., 2012; Meyer et al., 2012; Norman-Haignere et al., 2015; Santoro et al., 2014). In addition, the functional role of the PT is still not quite well understood and it is highly debated, whether an anatomical and coincidently functional left lateralization of the PT is existent (e.g., Dorsaint-pierre et al., 2006; Obleser, Zimmermann, Meter, Rauschecker, 2007). By keeping such open questions in respect of the PT in mind, it was, in our opinion, still reasonable to examine the effective recruitment of this region during the processing of complex sounds, voice included.

Another point why we decided to not include an additional ROI on the lateral STG is that we wanted to restrict the number and complexity of the analysed models. In this respect, overly increasing the complexity of a model by expanding the amount of its free parameters would lead to a so called model "overfit" at a certain point. This means that an excessively complex model may fit best to the observed BOLD time series, but concurrently its explanatory power concerning hypotheses decreases. Instead, it will also increasingly fit to data set specific noise at the cost of the model's generalizability (Stephan et al., 2010). So the premise regarding to a model explaining the observed data best, is the optimal trade-off of accuracy and complexity to afford a maximal generalizability (Pitt & Myung, 2002).

As a last point it has to be mentioned, that the conducted results on effective connectivities are also time- and experiment-dependent (Friston, 1994). This means that the estimation of effective connections across the ROIs is based on the fMRI measurement of BOLD signal time series and therefore completely constrained to the current data set. As to that, possibly sparse BOLD responses in general could have resulted in the few observations of significant modulatory connections of the current experiment. The exclusion of three participants after the BOLD signal time series extraction may hint towards such a general issue (see subsection 2.6.3.1 of the material and methods). Nevertheless this last aspect remains unanswered.

4.5 Limitations and further work

The present study comes with some limitations in terms of its experimental design and the used methodological approach. As to that, the method DCM, respectively all approaches which aim for the analysis of effective connectivities, depends on time series during which the state changes of neuronal systems are observed (Friston, 2011; Friston et al., 2003). As for fMRI, the hereof measured dependent variable is the BOLD signal from which the respective time series within the defined ROIs are extracted. In this respect, two disadvantages concerning fMRI in general and fMRI-based DCM in particular have to be mentioned. First, by using fMRI, one always infers from an indirect metabolism-based measurement how brain activities at the neuronal level proceed. This goes along with limitations concerning spacial resolution (one voxel contains thousands of neurons), but more important, temporal resolution (Huettel et al., 2004). While the BOLD signal is established over seconds, neurons fire within milliseconds. And this brings us to the second disadvantage: while functional connectivity measurements average voxel activations over time, effective connectivity goes a step further trying to reveal time-dependent hidden neuronal mechanisms underlying functional processes. This requires a temporal resolution being as near at the neuronal level as possible, which fMRI does not meet. To overcome such limitations, combined EEG-fMRI measurements could be promising to achieve a good spatial resolution via the fMRI-modality with a concurrently higher temporal resolution delivered through EEG (Mullinger, Castellone, & Bowtell, 2013). Currently, the same data set is evaluated for EEG as part of an other study which is remaining at the moment. A further limitation of this study is the construction of the model space for the DCM-analysis. As abovementioned, the current model design did possibly not cover an overall view on the processing of complex sounds within the early AC. Hereof, future attempts could model the respective neuronal network by including the lateral STG as ROI. According to that, non-primary "belt" and "parabelt" ROIs could be defined on the basis of cytoarchitectonic subdivisions instead of the used macroanatomical landmarks to achieve a more fine-grained view on the processing hierarchy.

Furthermore, a similar experiment on auditory effective connectivities could be conducted by using a set of natural sounds instead of the present three categories of voice, singing and music to possibly achieve an even more general view on the processing of complex sounds. On the other hand, one could criticise that the used stimuli are too widespread and could therefore implement our approach with a limited number of the existing stimuli.

Also, our criteria concerning musicianship were fitted for amateur musicians to conduct a more generalizable view across the population. Still, one could examine, if stricter definitions or the sole inclusion of professional musicians would have distinct effects on effective connectivities within the modeled neuronal network.

4.6 Conclusion

By investigating DCM-derived effective connectivities within the early auditory system for the first time, the present study tried to reveal the neuronal mechanisms underlying the processing of complex sounds. Hereof, our findings give rise to rethink current views concerning the functional role of non-primary auditory regions, situated at early time points of the sound processing hierarchy. As to that, it seems to be more likely that the lateralization of music and speech, generally observed on widespread networks across the brain hemispheres, is introduced at later time points in a possibly more multimodal manner. Also, the observed connectivity patterns during the perception of complex sounds did not reveal a straightforwardly separable mechanistic framework which hints towards a more distributed processing of the respective categories. Concerning this, it may be not advisable to assign categorial differentiations to macroanatomical regions like the PP and PT, but thinking of the non-primary "belt" as a more distributed, intermediary network. The investigation of musicianship and handedness revealed that these two factors have an influence on the neuronal processing of complex sounds within early AC. Though we were not able to generally deduce other models of effective connectivity in left-handers or musicians from the observed patterns, the current findings could be a good starting point towards a more detailed investigation of these two factors within the auditory domain. On that front, investigating lateralization of neuronal networks by means of effective connectivities seems to be promising.

Taken together, the present study utilized an advanced methodological approach in order to reveal more generalizable functionalities of the early auditory cortex. As this study served an exploratory manner, our findings together with the elaborated paradigm may be a good starting point towards further investigations on this field of research.

A Appendix - Additional Data

connection	RH non-musicians			
	mean	std	median	p_value
HG_l → HG_l	0.065942929505	0.12693929974	0.000229075662	0.213524354036
HG_l → PP_l	0.181448681047	0.349207706153	0.047462920991	0.038151710173
HG_l → PT_l	0.188929608333	0.200927739586	0.109895865608	0.007685794055
HG_l → HG_r	0.197703535086	0.314017354873	0.008454019563	0.109744638747
HG_r → HG_r	0.001784535053	0.441217511566	-8.1448677e-05	0.76709686841
HG_r → PP_r	0.178341231883	0.338521797117	0.008803878079	0.138640633813
HG_r → PT_r	0.131540029799	0.472787403631	0.009116871277	0.37425931928
HG_r → HG_l	0.296424453904	0.373903140996	0.090229133947	0.007685794055
PP_l → HG_l	-0.076862800087	0.322870849483	0.008238876246	0.441268133329
PP_r → HG_r	-0.21926607935	0.689689534495	-0.023130911951	0.26039294361
PT_l → HG_l	-0.115860365571	0.383744564392	0.007455605196	0.593954675327
PT_l → PT_r	0.072649102143	0.410112161279	-0.013951782217	0.37425931928
PT_r → HG_r	0.033717793783	0.490974351302	0.008613132204	0.313938093775
PT_r → PT_l	-0.022551867526	0.155425603828	0.007573352352	0.76709686841

Fig. 15: Results on the endogenous parameter estimates of connections after performance of summary statistics over the group of RH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	RH musicians			
	mean	std	median	p_value
HG_l → HG_l	0.071010977693	0.318158913085	-2.721405e-05	0.735316690637
HG_l → PP_l	0.186779285674	0.318700174019	0.197595979275	0.176296374441
HG_l → PT_l	-0.008881142808	0.455738564599	0.04877912406	0.735316690637
HG_l → HG_r	0.292868286412	0.341001859363	0.332906142937	0.090968947975
HG_r → HG_r	-0.085892473717	0.497669804832	-0.137011164928	0.310494434317
HG_r → PP_r	0.226734273809	0.330594445847	0.028299177052	0.090968947975
HG_r → PT_r	0.501790657125	0.37091426631	0.458989893254	0.017960477526
HG_r → HG_l	0.205946889557	0.328530450878	0.222152557287	0.128190174345
PP_l → HG_l	-0.131193311432	0.325946078726	-0.022786073371	0.236723570638
PP_r → HG_r	-0.162435042679	0.338178427467	-0.054908429201	0.236723570638
PT_l → HG_l	-0.119846363268	0.398577306812	-0.216036251022	0.498962298604
PT_l → PT_r	-0.20824406683	0.419640536066	-0.280749462316	0.310494434317
PT_r → HG_r	-0.085221411131	0.187314367528	-0.009137610842	0.310494434317
PT_r → PT_l	0.015721465525	0.230578108603	0.007438128327	0.865772374993

Fig. 16: Results on the endogenous parameter estimates of connections after performance of summary statistics over the group of RH musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	LH non-musicians			
	mean	std	median	p_value
HG_l → HG_l	0.267956091868	0.901519051232	-0.026291119334	1.0
HG_l → PP_l	-0.054723598378	0.348253350659	0.03374844109	0.715000654688
HG_l → PT_l	0.206402154315	0.878631782856	0.022832519354	0.465208818452
HG_l → HG_r	0.232214609068	0.435036354387	0.018476309056	0.067889154862
HG_r → HG_r	-0.225126308582	0.289588955262	-0.189553071559	0.465208818452
HG_r → PP_r	0.001179187306	0.159857000913	0.039270119002	0.715000654688
HG_r → PT_r	0.187911730538	0.35133029994	0.105508262315	0.273321678292
HG_r → HG_l	0.002297372675	0.103646948021	0.032469374039	0.715000654688
PP_l → HG_l	0.615639684728	0.970328924269	0.397224315096	0.273321678292
PP_r → HG_r	-0.05863668886	0.500240483711	0.046775116374	0.715000654688
PT_l → HG_l	0.202050111525	0.573063091308	-0.051907643451	1.0
PT_l → PT_r	-0.091012321117	0.241306428666	-0.055712031745	0.715000654688
PT_r → HG_r	-0.105277070068	0.647805433757	0.065753179873	0.715000654688
PT_r → PT_l	0.286767619562	0.342112505354	0.208199196592	0.067889154862

Fig. 17: Results on the endogenous parameter estimates of connections after performance of summary statistics over the group of LH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	music_RH-nM			
	mean	std	median	p_value
HG_l → HG_l	-0.435686952573	0.959053032349	-0.011593943355	0.173070920805
HG_l → PP_l	0.050708013314	0.158726389694	0.0	0.600179487141
HG_l → PT_l	0.11806378241	0.365730024341	0.001604914066	0.176296374441
HG_l → HG_r	-0.309402104148	0.499474434611	0.0	0.345447530469
HG_r → HG_r	-0.963816427016	0.99850562164	-0.88258244136	0.015155973881
HG_r → PP_r	0.044121126533	0.132747903869	0.0	0.592980098017
HG_r → PT_r	0.220146972027	0.436588089443	0.0	0.144127034816
HG_r → HG_l	-0.232714346743	0.70700670763	-1.5221166e-05	0.398024719507
PP_l → HG_l	0.21402270823	0.820371020799	0.0	0.916511907864
PP_r → HG_r	-0.020593449729	0.0624569074	0.0	1.0
PT_l → HG_l	0.139443307777	0.790766894207	0.0	0.865772374993
PT_l → PT_r	-0.199043387943	0.566089381427	-0.002025675905	0.236723570638
PT_r → HG_r	-0.00114400571	0.405469685559	0.0	0.465208818452
PT_r → PT_l	0.047100055096	0.381651584378	0.0	0.715000654688

Fig. 18: Results on the music-modulated parameter estimates of connections after performance of summary statistics over the group of RH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	music_RH-M			
	mean	std	median	p_value
HG_l → HG_l	-0.570084692674	1.234437850832	-0.562956289361	0.310494434317
HG_l → PP_l	0.027311688339	0.257996566133	0.0	0.685830434452
HG_l → PT_l	0.400932165578	0.58073742013	0.076125833711	0.07961580146
HG_l → HG_r	-0.219012514046	0.807569785793	-0.192323976281	0.345447530469
HG_r → HG_r	-1.366635114637	0.758690481411	-1.534601497964	0.027991815486
HG_r → PP_r	0.093945712739	0.279236599507	0.0	0.465208818452
HG_r → PT_r	0.00174129034	0.598947067558	0.0	1.0
HG_r → HG_l	0.405238668611	0.502290532634	0.253569065021	0.115851497526
PP_l → HG_l	0.236211541715	0.763482472683	0.0	0.500184257071
PP_r → HG_r	0.214465866813	0.230349662214	0.229351722582	0.067889154862
PT_l → HG_l	-0.407680784861	0.441022276879	-0.39838110115	0.07961580146
PT_l → PT_r	0.080141572375	0.48564406497	0.0	0.500184257071
PT_r → HG_r	-0.243756001523	0.475451750264	0.0	0.273321678292
PT_r → PT_l	-0.202867397457	0.457738941255	0.0	0.273321678292

Fig. 19: Results on the music-modulated parameter estimates of connections after performance of summary statistics over the group of RH musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	music_LH-nM			
	mean	std	median	p_value
HG_l → HG_l	-1.465416227169	1.038934844443	-1.764576415485	0.067889154862
HG_l → PP_l	0.41393681589	0.751625512506	0.058717168532	0.179712494879
HG_l → PT_l	0.739122474182	0.910272146477	0.54025310365	0.108809430041
HG_l → HG_r	0.112475359056	0.232949820061	-0.000272806796	1.0
HG_r → HG_r	-0.718780642574	0.68492138349	-0.678743079388	0.067889154862
HG_r → PP_r	0.274548212725	0.35457386955	0.177297889806	0.179712494879
HG_r → PT_r	0.094147727071	0.313680328553	-0.019990659061	0.715000654688
HG_r → HG_l	0.015810866227	0.126479974068	-0.002730303787	1.0
PP_l → HG_l	-0.989599170135	1.154794427272	-0.887481322987	0.179712494879
PP_r → HG_r	0.554910073548	1.752382896231	0.0	0.654720846019
PT_l → HG_l	-0.555897751425	1.0621171074	-0.113388310915	0.285049407403
PT_l → PT_r	-0.164738675713	0.607800817249	0.017803318015	1.0
PT_r → HG_r	0.040600361412	0.85425948105	0.103995781063	1.0
PT_r → PT_l	-0.410669964088	0.87995142358	-0.517969365614	0.465208818452

Fig. 20: Results on the music-modulated parameter estimates of connections after performance of summary statistics over the group of LH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	singing_RH-nM			
	mean	std	median	p_value
HG_l → HG_l	-0.484987578035	0.563311948111	-0.333890265692	0.028401837044
HG_l → PP_l	0.056973271414	0.199500304908	0.000336985962	0.172954917988
HG_l → PT_l	0.240523411868	0.331985785099	0.000699152007	0.090968947975
HG_l → HG_r	-0.183571920009	0.899227421904	0.0	0.463071015015
HG_r → HG_r	-0.627796830272	0.837737891428	-0.251598802817	0.050612432239
HG_r → PP_r	0.027854267812	0.083844436965	0.0	1.0
HG_r → PT_r	0.222205664998	0.448705977037	0.0	0.465208818452
HG_r → HG_l	-0.260186078169	0.763157550468	0.0	0.498962298604
PP_l → HG_l	-0.05540710839	0.538080141073	0.0	0.600179487141
PP_r → HG_r	-0.039659351746	0.118623367069	0.0	0.285049407403
PT_l → HG_l	0.141150067669	0.719495357139	0.0	1.0
PT_l → PT_r	-0.109397567868	0.434033222634	-0.001163702719	0.398024719507
PT_r → HG_r	-0.093418056392	0.184555185546	0.0	0.144127034816
PT_r → PT_l	-0.032407223572	0.086514477686	0.0	0.465208818452

Fig. 21: Results on the song-modulated parameter estimates of connections after performance of summary statistics over the group of RH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	singing_RH-M			
	mean	std	median	p_value
HG_l → HG_l	-1.126489072284	1.188167139096	-0.806008561769	0.027991815486
HG_l → PP_l	-0.08597024879	0.26547900719	0.0	0.685830434452
HG_l → PT_l	0.363225185341	0.527188225356	0.0	0.224915884016
HG_l → HG_r	-0.281791412378	0.728002009074	-0.004717879324	0.345447530469
HG_r → HG_r	-1.369606680822	0.861534689742	-1.640834888923	0.017960477526
HG_r → PP_r	0.055397228427	0.188238351189	0.0	0.465208818452
HG_r → PT_r	-0.030278549474	0.531981128944	0.0	1.0
HG_r → HG_l	0.016703014841	0.159521490948	-5.1538026e-05	0.753152364766
PP_l → HG_l	-0.621126559792	0.612478942664	-0.824629834702	0.043114446783
PP_r → HG_r	-0.128008435889	0.271771729617	0.0	0.273321678292
PT_l → HG_l	-0.193695712418	0.634408822347	0.0	0.345231071772
PT_l → PT_r	0.044308926062	0.596842918919	0.0	0.685830434452
PT_r → HG_r	0.095476784579	0.454385880095	0.0	0.465208818452
PT_r → PT_l	0.097990129758	0.313586702351	0.0	0.465208818452

Fig. 22: Results on the song-modulated parameter estimates of connections after performance of summary statistics over the group of RH musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	singing_LH-nM			
	mean	std	median	p_value
HG_l → HG_l	-1.491298684762	1.099841098391	-1.689658574327	0.067889154862
HG_l → PP_l	0.49683698219	0.910403761176	0.063951737927	0.179712494879
HG_l → PT_l	0.582567710824	1.106576155607	0.491913018366	0.285049407403
HG_l → HG_r	0.069040413619	0.661538849051	0.000804086401	0.592980098017
HG_r → HG_r	-0.955473577746	1.818177735298	-0.529228305646	0.273321678292
HG_r → PP_r	0.31676157101	0.418679075625	0.191999330316	0.179712494879
HG_r → PT_r	0.065189925431	0.501429985756	-0.097549899454	1.0
HG_r → HG_l	-0.268017563606	0.504979635345	-0.025629694701	0.285049407403
PP_l → HG_l	-1.696558174053	1.96601740148	-1.595047880863	0.179712494879
PP_r → HG_r	-0.578599144298	0.724926139985	-0.406318179676	0.179712494879
PT_l → HG_l	0.670796353514	0.568147061383	0.720619877607	0.108809430041
PT_l → PT_r	0.149917000835	0.184463636717	0.141370050324	0.285049407403
PT_r → HG_r	0.250090726285	0.645122148881	0.127930373612	0.465208818452
PT_r → PT_l	-0.215497836083	0.531736318711	-0.309150882276	0.465208818452

Fig. 23: Results on the song-modulated parameter estimates of connections after performance of summary statistics over the group of LH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	speech_RH-nM			
	mean	std	median	p_value
HG_l → HG_l	-0.647109382341	0.879548655897	-0.299960108113	0.007685794055
HG_l → PP_l	0.003013444537	0.129084400237	0.0	0.753152364766
HG_l → PT_l	0.183950067505	0.325017492497	0.0	0.498962298604
HG_l → HG_r	0.056811557641	0.476723526754	0.0	0.600179487141
HG_r → HG_r	-0.509399268875	0.681461588852	-0.102653608316	0.038151710173
HG_r → PP_r	0.035891636344	0.107921419123	0.0	1.0
HG_r → PT_r	0.232399261308	0.483356433492	0.0	0.465208818452
HG_r → HG_l	0.006430794119	0.833081053476	0.0	0.612089880089
PP_l → HG_l	0.252713263802	0.816319825821	0.0	0.600179487141
PP_r → HG_r	-0.036375141953	0.10839053034	0.0	0.285049407403
PT_l → HG_l	-0.324118123306	0.768343706078	-0.003517558101	0.090968947975
PT_l → PT_r	-0.010240341463	0.470788302802	0.0	0.735316690637
PT_r → HG_r	-0.109295421067	0.72221343715	0.0	0.715000654688
PT_r → PT_l	0.061439705565	0.166258551034	0.0	0.144127034816

Fig. 24: Results on the speech-modulated parameter estimates of connections after performance of summary statistics over the group of RH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	speech_RH-M			
	mean	std	median	p_value
HG_l → HG_l	-0.877983761765	0.951557605541	-0.645177608408	0.017960477526
HG_l → PP_l	-0.0606171651	0.363734594589	0.0	0.685830434452
HG_l → PT_l	0.396463801477	0.730971395024	0.0	0.224915884016
HG_l → HG_r	0.043578809974	0.896958997737	-0.007329667944	0.753152364766
HG_r → HG_r	-1.764021378671	1.734907256923	-1.871470885267	0.062979051214
HG_r → PP_r	0.00119805195	0.340666876933	0.0	0.715000654688
HG_r → PT_r	0.089436976499	0.776737407158	0.0	0.715000654688
HG_r → HG_l	0.023392043588	0.448511184556	0.0	0.916511907864
PP_l → HG_l	0.365601689288	0.613858142795	0.005633156831	0.138010737569
PP_r → HG_r	-0.196174540782	0.751558673305	0.0	0.715000654688
PT_l → HG_l	-0.218753331625	0.352152182792	0.0	0.224915884016
PT_l → PT_r	-0.087835718372	0.38989093933	-0.009239908125	0.345231071772
PT_r → HG_r	-0.364148159154	0.80759381071	0.0	0.273321678292
PT_r → PT_l	-0.066760751878	0.310565539732	0.0	0.715000654688

Fig. 25: Results on the speech-modulated parameter estimates of connections after performance of summary statistics over the group of RH musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

connection	speech_LH-nM			
	mean	std	median	p_value
HG_l → HG_l	-0.68087235229	0.696715550929	-0.715389267963	0.067889154862
HG_l → PP_l	0.205191836396	0.435313711675	0.0	0.654720846019
HG_l → PT_l	0.10120273021	0.889260874696	0.222924705299	1.0
HG_l → HG_r	-0.426239544293	0.671765298563	-0.146374181154	0.285049407403
HG_r → HG_r	-0.999562740465	1.409012623885	-1.006246084386	0.465208818452
HG_r → PP_r	0.204279157036	0.250794768851	0.152108580736	0.179712494879
HG_r → PT_r	-0.017378082296	0.466852511326	0.128185323385	0.715000654688
HG_r → HG_l	-0.112715322883	0.16191405065	-0.053412882859	0.108809430041
PP_l → HG_l	0.170638596433	1.6340826219	0.0	0.654720846019
PP_r → HG_r	0.724644899245	0.846996935944	0.644200424227	0.179712494879
PT_l → HG_l	-1.154412999378	2.527416154124	-0.034244040916	0.592980098017
PT_l → PT_r	0.097089145554	0.325416217722	0.097015492116	0.592980098017
PT_r → HG_r	-0.705690330414	0.538900228995	-0.76001217557	0.144127034816
PT_r → PT_l	-0.183276790017	0.350561324045	-0.235249966955	0.273321678292

Fig. 26: Results on the speech-modulated parameter estimates of connections after performance of summary statistics over the group of LH non-musicians by means of the one-sample Wilcoxon Test (Wilcoxon, 1945). Unit of values in Herz.

B Appendix - Scanner Log

SIEMENS MAGNETOM TrioTim syngo MR B17

\\USER\AGB\Peer\TNAC\AAHScout

TA: 0:14 PAT: 3 Voxel size: 1.6×1.6×1.6 mm Rel. SNR: 1.00 SIEMENS: AALScout

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	On
Load images to graphic segments	On
Auto open inline display	Off
Start measurement without further preparation	Off
Wait for user to start	Off
Start measurements	single

Routine

Slab group 1	
Slabs	1
Dist. factor	20 %
Position	Isocenter
Orientation	Sagittal
Phase enc. dir.	A >> P
Rotation	0 deg
AutoAlign	Head
Phase oversampling	0 %
Slice oversampling	0.0 %
Slices per slab	128
FoV read	260 mm
FoV phase	100.0 %
Slice thickness	1.6 mm
TR	3.15 ms
TE	1.37 ms
Averages	1
Concatenations	1
Filter	Prescan Normalize
Coil elements	HEA;HEP

Contrast

Flip angle	8 deg
Averaging mode	Short term
Reconstruction	Magnitude
Measurements	1

Resolution

Base resolution	160
Phase resolution	100 %
Slice resolution	69 %
Phase partial Fourier	6/8
Slice partial Fourier	6/8
PAT mode	GRAPPA
Accel. factor PE	3
Ref. lines PE	24
Accel. factor 3D	1
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Integrated
Image Filter	Off
Distortion Corr.	Off
Unfiltered images	Off
Prescan Normalize	On
Normalize	Off
B1 filter	Off
Raw filter	Off
Elliptical filter	Off

Geometry

Multi-slice mode	Sequential
Series	Ascending

System

Body	Off
HEP	On
HEA	On
Positioning mode	REF
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Save uncombined	Off
Coil Combine Mode	Adaptive Combine
Auto Coil Select	Off
Shim mode	Tune up
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	Isocenter
Orientation	Transversal
Rotation	0.00 deg
R >> L	350 mm
A >> P	263 mm
F >> H	350 mm

Inline

Time to center	6.2 s

Sequence

Introduction	On
Dimension	3D
Asymmetric echo	Weak
Contrasts	1
Bandwidth	550 Hz/Px
RF pulse type	Fast
Gradient mode	Normal
Excitation	Non-sel.
RF spoiling	On

\\USER\AGB\Peer\TNAC\t1_mpr_sag_p2_iso

TA: 4:26 PAT: 2 Voxel size: 1.0×1.0×1.0 mm Rel. SNR: 1.00 SIEMENS: tfl

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	Off
Start measurements	single

Routine

Slab group 1	
Slabs	1
Dist. factor	50 %
Position	L0.0 A12.2 F1.9
Orientation	Sagittal
Phase enc. dir.	A >> P
Rotation	7.18 deg
Phase oversampling	0 %
Slice oversampling	18.2 %
Slices per slab	176
FoV read	256 mm
FoV phase	100.0 %
Slice thickness	1.00 mm
TR	1900 ms
TE	2.26 ms
Averages	1
Concatenations	1
Filter	Prescan Normalize
Coil elements	HEA;HEP

Contrast

Magn. preparation	Non-sel. IR
TI	900 ms
Flip angle	9 deg
Fat suppr.	None
Water suppr.	None
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	1
Multiple series	Each measurement

Resolution

Base resolution	256
Phase resolution	100 %
Slice resolution	100 %
Phase partial Fourier	Off
Slice partial Fourier	Off
Interpolation	Off
PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	24
Accel. factor 3D	1
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Integrated
Image Filter	Off
Distortion Corr.	Off
Unfiltered images	On
Prescan Normalize	On
Normalize	Off
B1 filter	Off
Raw filter	Off
Elliptical filter	Off

Geometry

Multi-slice mode	Single shot
Series	Ascending

System

Body	Off
HEP	On
HEA	On
Positioning mode	REF
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Save uncombined	Off
Coil Combine Mode	Adaptive Combine
AutoAlign	Head > Brain
Auto Coil Select	Off
Shim mode	Tune up
Adjust with body coil	On
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	Isocenter
Orientation	Transversal
Rotation	0.00 deg
R >> L	350 mm
A >> P	263 mm
F >> H	350 mm

Physio

1st Signal/Mode	None
Dark blood	Off
Resp. control	Off

Inline

Subtract	Off
Std-Dev-Sag	Off
Std-Dev-Cor	Off
Std-Dev-Tra	Off
Std-Dev-Time	Off
MIP-Sag	Off
MIP-Cor	Off
MIP-Tra	Off
MIP-Time	Off
Save original images	On

Sequence

Introduction	On
Dimension	3D
Elliptical scanning	Off
Asymmetric echo	Allowed
Bandwidth	200 Hz/Px
Flow comp.	No

Echo spacing	6.8 ms
RF pulse type	Normal
Gradient mode	Normal
Excitation	Non-sel.
RF spoiling	On

SIEMENS MAGNETOM TrioTim syngo MR B17

\\USER\AGB\Peer\TNAC\ep2d_bold_TNAC_run1_1450_TE25_p2_des_3x3x4

TA: 9:18 PAT: 2 Voxel size: 3.0×3.0×4.0 mm Rel. SNR: 1.00 SIEMENS: ep2d_bold

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	Off
Start measurements	single

Routine

Slice group 1	
Slices	30
Dist. factor	15 %
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Phase enc. dir.	A >> P
Rotation	-0.11 deg
Phase oversampling	0 %
FoV read	192 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
TR	1450 ms
TE	25 ms
Averages	1
Concatenations	1
Filter	None
Coil elements	HEA;HEP

Contrast

MTC	Off
Flip angle	90 deg
Fat suppr.	Fat sat.
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	380
Delay in TR	0 ms
Multiple series	Off

Resolution

Base resolution	64
Phase resolution	100 %
Phase partial Fourier	Off
Interpolation	Off
PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	32
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Separate
Distortion Corr.	Off
Prescan Normalize	Off
Raw filter	On
Elliptical filter	Off
Hamming	Off

Geometry

Multi-slice mode	Interleaved
Series	Descending
Special sat.	None

System

Body	Off
HEP	On
HEA	On
Positioning mode	FIX
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Sum of Squares
AutoAlign	Head > Brain
Auto Coil Select	Default
Shim mode	Standard
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Rotation	-0.11 deg
R >> L	192 mm
A >> P	192 mm
F >> H	138 mm

Physio

1st Signal/Mode	None

BOLD

GLM Statistics	Off
Dynamic t-maps	Off
Starting ignore meas	0
Ignore after transition	0
Model transition states	On
Temp. highpass filter	On
Threshold	4.00
Paradigm size	20
Meas[1]	Baseline
Meas[2]	Baseline
Meas[3]	Baseline
Meas[4]	Baseline
Meas[5]	Baseline
Meas[6]	Baseline
Meas[7]	Baseline
Meas[8]	Baseline
Meas[9]	Baseline
Meas[10]	Baseline
Meas[11]	Active
Meas[12]	Active
Meas[13]	Active
Meas[14]	Active
Meas[15]	Active
Meas[16]	Active
Meas[17]	Active
Meas[18]	Active
Meas[19]	Active
Meas[20]	Active
Motion correction	Off
Spatial filter	Off

Sequence

Introduction	On
Bandwidth	2232 Hz/Px
Free echo spacing	Off
Echo spacing	0.53 ms
EPI factor	64
RF pulse type	Normal
Gradient mode	Fast*

\\USER\AGB\Peer\TNAC\ep2d_bold_TNAC_run2_1450_TE25_p2_des_3x3x4

TA: 9:18 PAT: 2 Voxel size: 3.0×3.0×4.0 mm Rel. SNR: 1.00 SIEMENS: ep2d_bold

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	Off
Start measurements	single

Routine

Slice group 1	
Slices	30
Dist. factor	15 %
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Phase enc. dir.	A >> P
Rotation	-0.11 deg
Phase oversampling	0 %
FoV read	192 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
TR	1450 ms
TE	25 ms
Averages	1
Concatenations	1
Filter	None
Coil elements	HEA;HEP

Contrast

MTC	Off
Flip angle	90 deg
Fat suppr.	Fat sat.
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	380
Delay in TR	0 ms
Multiple series	Off

Resolution

Base resolution	64
Phase resolution	100 %
Phase partial Fourier	Off
Interpolation	Off
PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	32
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Separate
Distortion Corr.	Off
Prescan Normalize	Off
Raw filter	On
Elliptical filter	Off
Hamming	Off

Geometry

Multi-slice mode	Interleaved
Series	Descending
Special sat.	None

System

Body	Off
HEP	On
HEA	On
Positioning mode	FIX
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Sum of Squares
AutoAlign	Head > Brain
Auto Coil Select	Default
Shim mode	Standard
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Rotation	-0.11 deg
R >> L	192 mm
A >> P	192 mm
F >> H	138 mm

Physio

1st Signal/Mode	None

BOLD

GLM Statistics	Off
Dynamic t-maps	Off
Starting ignore meas	0
Ignore after transition	0
Model transition states	On
Temp. highpass filter	On
Threshold	4.00
Paradigm size	20
Meas[1]	Baseline
Meas[2]	Baseline
Meas[3]	Baseline
Meas[4]	Baseline
Meas[5]	Baseline
Meas[6]	Baseline
Meas[7]	Baseline
Meas[8]	Baseline
Meas[9]	Baseline
Meas[10]	Baseline
Meas[11]	Active
Meas[12]	Active
Meas[13]	Active
Meas[14]	Active
Meas[15]	Active
Meas[16]	Active
Meas[17]	Active
Meas[18]	Active
Meas[19]	Active
Meas[20]	Active
Motion correction	Off
Spatial filter	Off

Sequence

Introduction	On
Bandwidth	2232 Hz/Px
Free echo spacing	Off
Echo spacing	0.53 ms
EPI factor	64
RF pulse type	Normal
Gradient mode	Fast*

SIEMENS MAGNETOM TrioTim syngo MR B17

\\USER\AGB\Peer\TNAC\ep2d_bold_TNAC_run3_1450_TE25_p2_des_3x3x4

TA: 9:18 PAT: 2 Voxel size: 3.0×3.0×4.0 mm Rel. SNR: 1.00 SIEMENS: ep2d_bold

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	Off
Start measurements	single

Routine

Slice group 1	
Slices	30
Dist. factor	15 %
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Phase enc. dir.	A >> P
Rotation	-0.11 deg
Phase oversampling	0 %
FoV read	192 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
TR	1450 ms
TE	25 ms
Averages	1
Concatenations	1
Filter	None
Coil elements	HEA;HEP

Contrast

MTC	Off
Flip angle	90 deg
Fat suppr.	Fat sat.
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	380
Delay in TR	0 ms
Multiple series	Off

Resolution

Base resolution	64
Phase resolution	100 %
Phase partial Fourier	Off
Interpolation	Off
PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	32
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Separate
Distortion Corr.	Off
Prescan Normalize	Off
Raw filter	On
Elliptical filter	Off
Hamming	Off

Geometry

Multi-slice mode	Interleaved
Series	Descending
Special sat.	None

System

Body	Off
HEP	On
HEA	On
Positioning mode	FIX
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Sum of Squares
AutoAlign	Head > Brain
Auto Coil Select	Default
Shim mode	Standard
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Rotation	-0.11 deg
R >> L	192 mm
A >> P	192 mm
F >> H	138 mm

Physio

1st Signal/Mode	None

BOLD

GLM Statistics	Off
Dynamic t-maps	Off
Starting ignore meas	0
Ignore after transition	0
Model transition states	On
Temp. highpass filter	On
Threshold	4.00
Paradigm size	20
Meas[1]	Baseline
Meas[2]	Baseline
Meas[3]	Baseline
Meas[4]	Baseline
Meas[5]	Baseline
Meas[6]	Baseline
Meas[7]	Baseline
Meas[8]	Baseline
Meas[9]	Baseline
Meas[10]	Baseline
Meas[11]	Active
Meas[12]	Active
Meas[13]	Active
Meas[14]	Active
Meas[15]	Active
Meas[16]	Active
Meas[17]	Active
Meas[18]	Active
Meas[19]	Active
Meas[20]	Active
Motion correction	Off
Spatial filter	Off

Sequence

Introduction	On
Bandwidth	2232 Hz/Px
Free echo spacing	Off
Echo spacing	0.53 ms
EPI factor	64
RF pulse type	Normal
Gradient mode	Fast*

\\USER\AGB\Peer\TNAC\ep2d_bold_TNAC_run4_1450_TE25_p2_des_3x3x4

TA: 9:18 PAT: 2 Voxel size: 3.0×3.0×4.0 mm Rel. SNR: 1.00 SIEMENS: ep2d_bold

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	Off
Start measurements	single

Routine

Slice group 1	
Slices	30
Dist. factor	15 %
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Phase enc. dir.	A >> P
Rotation	-0.11 deg
Phase oversampling	0 %
FoV read	192 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
TR	1450 ms
TE	25 ms
Averages	1
Concatenations	1
Filter	None
Coil elements	HEA;HEP

Contrast

MTC	Off
Flip angle	90 deg
Fat suppr.	Fat sat.
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	380
Delay in TR	0 ms
Multiple series	Off

Resolution

Base resolution	64
Phase resolution	100 %
Phase partial Fourier	Off
Interpolation	Off
PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	32
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Separate
Distortion Corr.	Off
Prescan Normalize	Off
Raw filter	On
Elliptical filter	Off
Hamming	Off

Geometry

Multi-slice mode	Interleaved
Series	Descending
Special sat.	None

System

Body	Off
HEP	On
HEA	On
Positioning mode	FIX
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Sum of Squares
AutoAlign	Head > Brain
Auto Coil Select	Default
Shim mode	Standard
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	R0.1 P6.8 H11.3
Orientation	T > C5.0
Rotation	-0.11 deg
R >> L	192 mm
A >> P	192 mm
F >> H	138 mm

Physio

1st Signal/Mode	None

BOLD

GLM Statistics	Off
Dynamic t-maps	Off
Starting ignore meas	0
Ignore after transition	0
Model transition states	On
Temp. highpass filter	On
Threshold	4.00
Paradigm size	20
Meas[1]	Baseline
Meas[2]	Baseline
Meas[3]	Baseline
Meas[4]	Baseline
Meas[5]	Baseline
Meas[6]	Baseline
Meas[7]	Baseline
Meas[8]	Baseline
Meas[9]	Baseline
Meas[10]	Baseline
Meas[11]	Active
Meas[12]	Active
Meas[13]	Active
Meas[14]	Active
Meas[15]	Active
Meas[16]	Active
Meas[17]	Active
Meas[18]	Active
Meas[19]	Active
Meas[20]	Active
Motion correction	Off
Spatial filter	Off

Sequence

Introduction	On
Bandwidth	2232 Hz/Px
Free echo spacing	Off
Echo spacing	0.53 ms
EPI factor	64
RF pulse type	Normal
Gradient mode	Fast*

\\USER\AGB\Peer\TNAC\t2_spc_sag_p2_iso1mm

TA: 4:43 PAT: 2 Voxel size: 1.0×1.0×1.0 mm Rel. SNR: 1.00 SIEMENS: tse_vfl

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	Off
Start measurements	single

Routine

Slab group 1	
Slabs	1
Position	L0.0 A12.2 F1.9
Orientation	Sagittal
Phase enc. dir.	A >> P
Rotation	7.18 deg
Phase oversampling	0 %
Slice oversampling	0.0 %
Slices per slab	176
FoV read	256 mm
FoV phase	100.0 %
Slice thickness	1.00 mm
TR	3200 ms
TE	402 ms
Averages	1.0
Concatenations	1
Filter	Raw filter, Prescan Normalize
Coil elements	HEA;HEP

Contrast

MTC	Off
Magn. preparation	None
Fat suppr.	None
Water suppr.	None
Restore magn.	Off
Reconstruction	Magnitude
Measurements	1
Multiple series	Each measurement

Resolution

Base resolution	256
Phase resolution	101 %
Slice resolution	100 %
Phase partial Fourier	Allowed
Slice partial Fourier	Off
Interpolation	On
PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	24
Accel. factor 3D	1
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Integrated
Image Filter	Off
Distortion Corr.	Off
Unfiltered images	On
Prescan Normalize	On
Normalize	Off
B1 filter	Off
Raw filter	On
Intensity	Weak
Slope	25
Elliptical filter	Off

Geometry

Special sat.	None

System

Body	Off
HEP	On
HEA	On
Positioning mode	FIX
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Save uncombined	Off
Coil Combine Mode	Adaptive Combine
AutoAlign	Head > Brain
Auto Coil Select	Off
Shim mode	Tune up
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	Isocenter
Orientation	Transversal
Rotation	0.00 deg
R >> L	350 mm
A >> P	263 mm
F >> H	350 mm

Physio

1st Signal/Mode	None
Dark blood	Off
Resp. control	Off

Inline

Subtract	Off
Std-Dev-Sag	Off
Std-Dev-Cor	Off
Std-Dev-Tra	Off
Std-Dev-Time	Off
MIP-Sag	Off
MIP-Cor	Off
MIP-Tra	Off
MIP-Time	Off
Save original images	On

Sequence

Introduction	On
Dimension	3D
Bandwidth	751 Hz/Px
Flow comp.	No
Allowed delay	30 s
Echo spacing	3.36 ms

Adiabatic-mode	Off
Define	Echo trains
Turbo factor	141
Slice turbo factor	2
Echo trains per slice	1
Echo train duration	870
RF pulse type	Normal
Gradient mode	Fast
Excitation	Non-sel.
Flip angle mode	T2 var

C Appendix - Ethic Proposal

References

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, *21*(7), 1507–1518. doi:10.1093/cercor/bhq198

Alho, K., Rinne, T., Herron, T. J., & Woods, D. L. (2014). Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies. *Hearing Research*, *307*, 29–41. doi:10.1016/j.heares.2013.08.001

Andoh, J., Matsushita, R., & Zatorre, R. J. (2015). Asymmetric Interhemispheric Transfer in the Auditory Network: Evidence from TMS, Resting-State fMRI, and Diffusion Imaging. *Journal of Neuroscience*, *35*(43), 14602–14611. doi:10.1523/JNEUROSCI.2333-15.2015

Angulo-Perkins, A., Aubé, W., Peretz, I., Barrios, F. A., Armony, J. L., & Concha, L. (2014). Music listening engages specific cortical regions within the temporal lobes: Differences between musicians and non-musicians. *Cortex*, *59*, 126–137. doi:10.1016/j.cortex.2014.07.013

Arnold, H. J. (1965). Small Sample Power of the One Sample Wilcoxon Test for Non-Normal Shift Alternatives. *The Annals of Mathematical Statistics*, *36*(6), 1767–1778.

Ashburner, J., Barnes, G., Chen, C.-c., Daunizeau, J., Moran, R., Henson, R., ... Phillips, C. (2017). *SPM12 Manual*. doi:10.1111/j.1365-294X.2006.02813.x

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, C. (2012). A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. *NeuroImage*, *54*(3), 2033–2044. doi:10.1016/j.neuroimage.2010.09.025.A

Barton, B., Venezia, J. H., Saberi, K., Hickok, G., & Brewer, A. A. (2012). Orthogonal acoustic dimensions define auditory field maps in human cortex. *Proceedings of the National Academy of Sciences*, *109*(50), 20738–20743. doi:10.1073/pnas.1213381109

Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119*, 1239–1247.

Cammoun, L., Thiran, J. P., Griffa, A., Meuli, R., Hagmann, P., & Clarke, S. (2015). Intrahemispheric cortico-cortical connections of the human auditory cortex. *Brain Structure and Function*, *220*(6), 3537–3553. doi:10.1007/s00429-014-0872-z

Catani, M. & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132. doi:10.1016/j.cortex.2008.05.004

Cha, K., Zatorre, R. J., & Schönwiesner, M. (2016). Frequency Selectivity of Voxel-by-Voxel Functional Connectivity in Human Auditory Cortex. *Cerebral Cortex*, *26*(1), 211–224. doi:10.1093/cercor/bhu193

Da Costa, S., van der Zwaag, W., Marques, J. P., Frackowiak, R. S. J., Clarke, S., & Saenz, M. (2011). Human Primary Auditory Cortex Follows the Shape of Heschl's Gyrus. *Journal of Neuroscience*, *31*(40), 14067–14075. doi:10.1523/JNEUROSCI.2000-11.2011

Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., Fischl, B., & Se, F. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*, 1060–1075. doi:10.1016/j.neuroimage.2004.03.032

Dassonville, P., Zhu, X.-H., Ugurbil, K., Kim, S.-G., & Ashe, J. (1997). Functional activation in motor cortex reflects the direction and the degree of handedness. *Proceedings of the National Academy of Sciences*, *94*(25), 14015–14018. doi:10.1073/pnas.94.25.14015

Dorsaint-pierre, R., Penhune, V. B., Watkins, K. E., Neelin, P., Lerch, J. P., Bouffard, M., & Zatorre, R. J. (2006). Asymmetries of the planum temporale and Heschl ' s gyrus : relationship to language lateralization. *Brain*, *129*, 1164–1176. doi:10.1093/brain/awl055

Elias, L. J., Bryden, M., & Bulman-Fleming, M. (1998). Footedness is a better predictor than is handedness of emotional lateralization. *Neuropsychologia*, *36*(1), 37–43. doi:10.1016/S0028-3932(97)00107-3

Elmer, S., Hänggi, J., & Jäncke, L. (2016). Interhemispheric transcallosal connectivity between the left and right planum temporale predicts musicianship, performance in temporal speech processing, and functional specialization. *Brain Structure and Function*, *221*(1), 331–344. doi:10.1007/s00429-014-0910-x

Elmer, S., Hänggi, J., Meyer, M., & Jäncke, L. (2013). Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic

and temporal speech sounds. *Cortex*, *49*(10), 2812–2821. doi:10.1016/j.cortex.2013.03.007

Elmer, S., Meyer, M., & Jäncke, L. (2012). Neurofunctional and Behavioral Correlates of Phonetic and Temporal Categorization in Musically Trained and Untrained Subjects. *Cerebral Cortex*, *22*(3), 650–658. doi:10.1093/cercor/bhr142

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE*, *12*(9), e0184661. doi:10.1371/journal.pone.0184661

Fedorenko, E., McDermott, J. H., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology*, *108*(12), 3289–3300. doi:10.1152/jn.00209.2012

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... Dale, A. M. (2002). Whole Brain Segmentation : Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron*, *33*, 341–355.

Fischl, B., Salat, D. H., Kouwe, J. W. V. D., Makris, N., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23*, 69–84. doi:10.1016/j.neuroimage.2004.07.016

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. doi:10.1016/S1053-8119(09)70884-5

Frässle, S., Krach, S., Paulus, F. M., & Jansen, A. (2016). Handedness is related to neural mechanisms underlying hemispheric lateralization of face processing. *Scientific Reports*, *6*(June), 1–17. doi:10.1038/srep27153

Friston, K. J. [K. J.]. (1994). Functional and Effective Connectivity in Neuroimaging: A Synthesis. *Human Brain Mapping*, *2*, 56–78.

Friston, K. J. [K. J.]. (2011). Functional and Effective Connectivity: A Review. *Brain Connectivity*, *1*(1), 13–36. doi:10.1089/brain.2011.0008

Friston, K. J. [K J], Frith, C. D., & Frackowiak, R. S. J. (1993). Time-Dependent Changes in Effective Connectivity Measured With PET. *Human Brain Mapping*, *1*, 69–79.

Friston, K. J. [K. J.], Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J. (1993). Functional Connectivity : The Principal-Component Analysis of Large (PET) Data Sets. *Journal of Cerebral Blood Flow and Metabolism*, *13*, 5–14.

Friston, K., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. doi:10.1016/S1053-8119(03)00202-7

Geschwind, N. & Levitsky, W. (1968). Human Brain: Left-Right Asymmetries in Temporal Speech Region. *Science*, *161*(3837), 186–187. doi:10.1126/science.161.3837.186

Glasberg, B. R. & Moore, B. C. J. (2002). A Model of Loudness Applicable to Time-Varying Sounds. *Journal of Audio Engineering Society*, *50*(5), 331–342. doi:10.1016/S0378-5955(03)00347-2

Gorgolewski, K., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*, 160044. doi:10.1038/sdata.2016.44

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*(August). doi:10.3389/fninf.2011.00013

Goulden, N., Elliott, R., Suckling, J., Williams, S. R., Deakin, J. F. W., & McKie, S. (2012). Sample size estimation for comparing parameters using dynamic causal modeling. *Brain Connectivity*, *2*(2), 80–90. doi:10.1089/brain.2011.0057

Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., & Rauschecker, J. P. (1997). Attention-related modulation of activity in primary and secondary auditory cortex. *NeuroReport*, *8*(11), 2511–2516. doi:10.1097/00001756-199707280-00019

Griffiths, T. D. & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*(7), 348–353. doi:10.1016/S0166-2236(02)02191-4

Häkkinen, S. & Rinne, T. (2018). Intrinsic, stimulus-driven and task-dependent connectivity in human auditory cortex. *Brain Structure and Function*, *223*(5), 2113–2127. doi:10.1007/s00429-018-1612-6

Han, X. & Fischl, B. (2007). Atlas Renormalization for Improved Brain MR Image Segmentation Across Scanner Platforms. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, *26*(4), 479–486.

Herdener, M., Esposito, F., Scheffler, K., Schneider, P., Logothetis, N. K., Uludag, K., & Kayser, C. (2013). Spatial representations of temporal and spectral sound cues in human auditory cortex. *Cortex*, *49*(10), 2822–2833. doi:10.1016/j.cortex.2013.04.003

Hickok, G. & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Neuroscience*, *8*, 393–402.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. doi:10.1088/1475-7516/2003/08/005

Humphries, C., Liebenthal, E., & Binder, J. R. (2010). Tonotopic organization of human auditory cortex. *NeuroImage*, *50*(3), 1202–1211. doi:10.1016/j.neuroimage.2010.01.046

Jansen, A., Flöel, A., Menke, R., Kanowski, M., & Knecht, S. (2005). Dominance for language and spatial processing: Limited capacity of a single hemisphere. *NeuroReport*, *16*(9), 1017–1021. doi:10.1097/00001756-200506210-00027

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. doi:10.1016/j.neuroimage.2011.09.015

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., ... Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. doi:10.1371/journal.pcbi.1005350

Klöppel, S., van Eimeren, T., Glauche, V., Vongerichten, A., Münchau, A., Frackowiak, R. S. J., ... Siebner, H. R. (2007). The effect of handedness on cortical motor activation during simple bilateral movements. *NeuroImage*, *34*(1), 274–280. doi:10.1016/j.neuroimage.2006.08.038

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., ... Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, *123*(12), 2512–2518. doi:10.1093/brain/123.12.2512

Langers, D. R., Backes, W. H., & van Dijk, P. (2007). Representation of lateralization and tonotopy in primary versus secondary human auditory cortex. *NeuroImage*, *34*(1), 264–273. doi:10.1016/j.neuroimage.2006.09.002

Leaver, A. M. & Rauschecker, J. P. [J. P.]. (2010). Cortical Representation of Natural Complex Sounds: Effects of Acoustic Features and Auditory Object Category. *Journal of Neuroscience*, *30*(22), 7604–7612. doi:10.1523/jneurosci.0296-10.2010

Leaver, A. M. & Rauschecker, J. P. [J. P.]. (2016). Functional Topography of Human Auditory Cortex. *Journal of Neuroscience*, *36*(4), 1416–1428. doi:10.1523/JNEUROSCI.0226-15.2016

Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60.

McCrum-Gardner, E. (2008). Which is the correct statistical test to use ? *British Journal of Oral and Maxillofacial Surgery*, *46*, 38–41. doi:10.1016/j.bjoms.2007.09.002

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *PROC. OF THE 9th PYTHON IN SCIENCE CONF. 9*, 51–56.

Meyer, M., Elmer, S., & Jäncke, L. (2012). Musical expertise induces neuroplasticity of the planum temporale. *Annals of the New York Academy of Sciences*, *1252*(1), 116–123. doi:10.1111/j.1749-6632.2012.06450.x

Moerel, M., Martino, F. D., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, *8*(225), 1–14. doi:10.3389/fnins.2014.00225

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. *NeuroImage*, *13*(4), 684–701. doi:10.1006/nimg.2000.0715

Mullinger, K. J., Castellone, P., & Bowtell, R. (2013). Best Current Practice for Obtaining High Quality EEG Data During Simultaneous fMRI. *Journal of Visualized Experiments*, (76), 1–10. doi:10.3791/50283

Norman-Haignere, S., Kanwisher, N., & McDermott, J. H. (2013). Cortical Pitch Regions in Humans Respond Primarily to Resolved Harmonics and Are Located in Specific Tonotopic Regions of Anterior Auditory Cortex. *Journal of Neuroscience*, *33*(50), 19451–19469. doi:10.1523/JNEUROSCI.2880-13.2013

Norman-Haignere, S., Kanwisher, N., & McDermott, J. H. (2015). Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*, *88*(6), 1281–1296. doi:10.1016/j.neuron.2015.11.035

Obleser, J., Zimmermann, J., Meter, J. V., & Rauschecker, J. P. (2007). Multiple Stages of Auditory Speech Perception Reflected in Event-Related fMRI. *Cerebral Cortex*, *17*, 2251–2257. doi:10.1093/cercor/bhl133

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. doi:10.1016/0028-3932(71)90067-4

Oliphant, T. E. (2006). *Guide to NumPy*.

Oliphant, T. E. (2007). Python for Scientific Computing. *Computing in Science & Engineering*, *9*(10), 9–20. doi:10.1109/MCSE.2007.58

Pantev, C. & Herholz, S. C. (2011). Plasticity of the human auditory cortex related to musical training. *Neuroscience and Biobehavioral Reviews*, *35*(10), 2140–2154. doi:10.1016/j.neubiorev.2011.06.010

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. doi:10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*(January), 1–8. doi:10.3389/neuro.11.010.2008

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., & Leff, A. P. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology*, *6*(3). doi:10.1371/journal.pcbi.1000709

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*. doi:10.1016/j.cognition.2005.11.004

Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience*, *7*(6), 658–663. doi:10.1038/nn1256

Pitt, M. A. & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*(10), 421–425.

Pujol, J., Deus, J., Losilla, J. M., & Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology*, *52*(5), 1038–1038. doi:10.1212/WNL.52.5.1038

Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., LaMantia, A.-S., McNamara, J., & Williams, M. (2004). Neuroscience. *Sinauer Associates, Inc. Sunderland, Massachusetts U.S.A. 3*, 287. doi:10.1212/01.WNL.0000154473.43364.47

Rauschecker, J. P. [J P], Tian, B., Pons, T., & Mishkin, M. (1997). Serial and parallel processing in macaque auditory cortex. *103*(December 1996), 89–103.

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional Anatomy of Language and Music Perception: Temporal and Structural Factors Investigated Using Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *31*(10), 3843–3852. doi:10.1523/JNEUROSCI.4515-10.2011

Saenz, M. & Langers, D. R. (2014). Tonotopic mapping of human auditory cortex. *Hearing Research*, *307*, 42–52. doi:10.1016/j.heares.2013.07.016

Santoro, R., Moerel, M., De Martino, F., Goebel, R., Ugurbil, K., Yacoub, E., & Formisano, E. (2014). Encoding of Natural Sounds at Multiple Spectral and Temporal Resolu-

tions in the Human Auditory Cortex. *PLoS Computational Biology*, *10*(1), e1003412. doi:10.1371/journal.pcbi.1003412

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*(8), 1047–1055.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In Vivo Evidence of Structural Brain Asymmetry in Musicians. *Science*, *267*(5198), 699–701.

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J. L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage*, *51*(1), 450–461. doi:10.1016/j.neuroimage.2010.02.023

Shapiro, S. & Wilk, M. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3), 591–611.

Steinmetz, H. (1996). Structure , Function and Cerebral Asymmetry: In Vivo Morphometry of the Planum Temporale. *Neuroscience and Biobehavioral Reviews*, *20*(4), 587–591.

Stephan, K. E. [K. E.], Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage*, *49*(4), 3099–3109. doi:10.1016/j.neuroimage.2009.11.015

Stephan, K. E. [Klaas E.], Penny, W. D., Marshall, J. C., Fink, G. R., & Friston, K. J. (2005). Investigating the functional role of callosal connections with dynamic causal models. *Annals of the New York Academy of Sciences*, *1064*, 16–36. doi:10.1196/annals.1340.008

Szaflarski, J. P., Binder, J. R., Possing, E. T., McKiernan, K. A., Ward, B. D., & Hammeke, T. A. (2002). Language lateralization in left-handed and ambidextrous people. *Neurology*, *59*(2), 238–244. doi:10.1212/WNL.59.2.238

Tardif, E. & Clarke, S. (2001). Intrinsic connectivity of human auditory areas: a tracing study with DiI. *European Journal of Neuroscience*, *13*, 1045–1050.

Tervaniemi, M. & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Research Reviews*, *43*(3), 231–246. doi:10.1016/j.brainresrev.2003.08.004

Tervaniemi, M., Szameitat, A. J., Kruck, S., Schroger, E., Alter, K., De Baene, W., & Friederici, A. D. (2006). From Air Oscillations to Music and Speech: Functional Magnetic Resonance Imaging Evidence for Fine-Tuned Neural Networks in Audition. *Journal of Neuroscience*, *26*(34), 8647–8652. doi:10.1523/JNEUROSCI.0995-06.2006

Upadhyay, J., Silver, A., Knaus, T. A., Lindgren, K. A., Ducros, M., Kim, D.-S., & Tager-Flusberg, H. (2008). Effective and Structural Connectivity in the Human Auditory

Cortex. *Journal of Neuroscience*, *28*(13), 3341–3349. doi:10.1523/JNEUROSCI.4434-07.2008

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188–196. doi:10.1016/j.lindif.2010.02.004

Westbury, C. F., Zatorre, R. J., & Evans, A. C. (1999). Quantifying Variability in the Planum Temporale : A Probability Map. *Cerebral Cortex*, *9*(4), 392–405.

Whitehead, J. C. & Armony, J. L. (2018). Singing in the brain: Neural representation of music and voice as revealed by fMRI. *Human Brain Mapping*, (February), 1–12. doi:10.1002/hbm.24333

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, *1*(6), 80–83.

Willems, R. M., Der Haegen, L. V., Fisher, S. E., & Francks, C. (2014). On the other hand: Including left-handers in cognitive neuroscience and neurogenetics. *Nature Reviews Neuroscience*, *15*(3), 193–201. doi:10.1038/nrn3679

Woodruff, P. W. R., Benson, R. R., Bandettini, P. A., Kwong, K. K., Howard, R. J., Talavage, T., ... Rosen, B. R. (1996). Modulation of auditory and visual cortex by selective attention is modality-dependent. *NeuroReport*, *7*(12), 1909–1913. doi:10.1097/00001756-199608120-00007

Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., ... Kang, X. (2009). Functional maps of human auditory cortex: Effects of acoustic features and attention. *PLoS ONE*, *4*(4). doi:10.1371/journal.pone.0005183

Zatorre, R. J. [R. J.] & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral cortex (New York, N.Y. : 1991)*, *11*(10), 946–53. doi:10.1093/cercor/11.10.946

Zatorre, R. J. [Robert J]. (2002). Auditory Cortex. In *Encyclopedia of the human brain* (Vol. 1, pp. 289–301). Elsevier.

Zatorre, R. J. [Robert J.] & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, *10*(7), 5. doi:10.1371/journal.pbio.1001372

Zatorre, R., Belin, P., & Penhune, V. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46. doi:10.1016/S1364-6613(00)01816-7

Zimmerman, D. W. (1987). Comparative Power of Student T Test and Mann-Whitney U Test for Unequal Sample Sizes and Variances. *The Journal of Experimental Education*, *55*(3), 171–174.