Myeloarchitecture and Intrinsic Functional Connectivity of Auditory Cortex in Musicians with Absolute Pitch

Von der Fakultät für Biowissenschaften, Pharmazie und Psychologie
der Universität Leipzig
genehmigte

D I S S E R T A T I O N

zur Erlangung des akademischen Grades
Doctor rerum naturalium
Dr. rer. nat.

Vorgelegt von
Dr. rer. nat. Seung-Goo Kim

geboren am 06 April 1980 in Seoul, Südkorea.

Dekan: Prof. Dr. PD. Erich Schröger

Gutachter: Prof. Dr. Marc Schönwiesner

  – Prof. Dr. Jörg Jescheniak

  – Prof. Dr. Angela D. Friederici

  – PD. Dr. Thomas R. Knösche

Tag der Verteidigung: 5 January 2017

Acknowledgments

I am sincerely thankful for my privileged four years in Leipzig, the city of Wilhelm Wundt ("the father[1] of experimental psychology") and Johann Sebastian Bach ("the father of classical music"). It was a wonderful scientific experience and I want to express my gratitude for supports from many people.

My adviser Dr. PD. Thomas R. Knösche allowed me to seek satisfaction of my own personal curiosity and helped me to sublimate the exploration into scientific works. His scientific mind truly inspired me and experience of working with him will guide me in the future (for further exploration and its sublimation).

I thank two reviewers of the dissertation, Prof. Marc Schönwiesner and Prof. Peter Keller, and the Committee, Prof. Erich Schröger (chair), Dr. PD. Thomas R. Knösche, Prof. Marc Schönwiesner, and Prof. Jörg Jescheniak.

I also thank German taxpayers (federal tax in particular) and Max Planck Society, funded my PhD project *via* International Max Planck Research School on Neuroscience of Communication (IMPRS NeuroCom) with help from the coordinators, Dr. Antje Niven and Dr. Veronika Krieghoff.

My colleagues and friends helped and supported me in many ways: Daniela Sammler helped me with brilliantly inspiring discussion in developing the framework of the thesis. Jae-Hyun Cho has been a nice friend and a colleague. It was very lucky for me to have a Korean friend in the office in Germany, and he helped me with his expertise in connectivity analysis. Dariya Goranskaya encouraged (or forced) me to take initiative in my weak part (getting involved with strangers) for recruiting musicians and helped me to improve my defense presentation really a lot. Tim Kunze and Jan Schreiber help me substantially in my academic life in Germany. I also thank my Taiwanese friends, Pei-Chen Shih and Shih-Cheng Chien for being kind and sweet friends, who made my PhD days in Leipzig much more socially enjoyable.

Finally I thank my family who made it all possible with their trust and support. Although I am an atheist, I deeply thank your praying. I love you all!

---

[1] I know, where is the mother then?

감사의 말씀

지난 사년 간 라이프치히(실험적 심리학의 아버지인 빌헬름 분트와 고전음악의 아버지인 요한 세바스찬 바하의 도시, 그리고 제가 연구를 수행한 막스플랑크 인지뇌과학연구소의 도시!)에서 생활은 과학자로서 매우 값진 시간이었습니다. 많은 분들이 도움을 주셨습니다.

지도교수인 토마스 크노슈케 박사님은 제 개인적인 호기심을 과학적 연구로 승화시키는데 큰 도움을 주셨습니다. 특히 박사님은 과학자로서가져야 할 가장 중요한 소양은 바로 중요한 질문을 던지기 위해 집요하게 사고하는 것임을 보여주셨으며, 앞으로 저의 연구에 중요한 구심점이 되도록 마음에 세길 것입니다.

그리고 논문심사에 참가하셨던 교수님들, 에리히 슈로거, 마크 쉔비스너, 요르그 예쉔니에크께 감사를 드립니다. 논문 평가를 하셨던 피터 켈러 교수님과 마크 쉔비스너 교수님께도 감사를 드립니다.

독일의 납세자 여러분과 막스플랑크 그룹에도 감사를 드립니다. 세계적으로 최고 수준의 연구 설비(와 월급)을 지원해주신 덕에 마음껏 연구할 수 있었습니다. 외국인으로서 갖는 행정적 문제들에 큰 도움을 주신 안티에 니벤 박사님과 베로니카 크리그호프 박사님께도 감사를 드립니다.

동료들과 친구들에게도 큰 도움을 받았습니다. 다니엘라는 심도 깊은 토론을 통해 많은 과학적 영감을 얻을 수 있는 멋진 동료였으며, 재현이 낯선 독일에서 많이 의지할 수 있었던 좋은 친구였으며, 또 전문적인 지식으로 제게 큰 도움을 주었습니다. 다리야는 특히 논문심사 발표 연습에 훌륭한 조언을 해주었으며, 덕분에 논문심사를 성공적으로 마무리할 수 있었습니다. 연구실 동료인 팀 쿠르츠와 얀 슈크라이버는 독일에서의 생활에 도움을 주었습니다다. 대만에서 온 페이페이와 쉬첸 덕분에 라이프치히에서 연구실 밖의 생활도 매우 즐거울 수 있었습니다.

마지막으로 우리 가족들에게도 깊은 감사를 표합니다. 우리 지연이와 스서방, 형과 수자누나, 우리 귀여운 아림이랑 세림이, 그리고 우리 아빠와 엄마. 어쩌다보니 왠지 모르게 만학도가 된 둘째 아들/오빠/동생을 이해하고 도와주신 덕분에, 그리고 기도해주신 덕분에, 무사히 그리고 성공적으로 박사학위를 받을 수 있었습니다. 감사합니다. 그리고 사랑합니다!

이천십칠년 라이프치히에서 김승구

Table of Contents

I. Introduction

In the first chapter, I will briefly review literature on sound, pitch, and pitch chroma to introduce the concept of absolute pitch (AP) in the context of the recent auditory cognitive neuroscience. The AP is known as an ability to recognize pitch chroma (e.g., "C" or "A") in an absolute sense without external references. In my view, the most interesting component of AP processing is chromatic categorization based on the internal references. In order to extract the pitch chroma from a tonal sound in AP listeners, pitch extraction should precede or be simultaneously done with chroma extraction. Therefore, pitch extraction and the human auditory system will be discussed before a review of previous findings of behaviors and neural structures and functions related to AP.

I-1. Sound and Auditory Systems

I-1.1. Auditory Apparatus

From the perspective of 'physiological psychology' (Wundt, 1904), audition is a neural process acting on an oscillation that propagates, as a mechanical wave, through a medium such as air or water, from the vibrator (sound source) to the neural system that detects it. Waves can be described by a wave equation. For example, a one-dimensional wave is given as:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad\qquad \text{(I-1)}$$

where t is time, x is a location along the propagation direction, u is a displacement (e.g., of air particles) at time t and location x, and c is constant for a medium. The direction of displacement can be perpendicular (e.g., a transverse wave on a string), parallel (e.g., a longitudinal wave on an eardrum), or mixture of them with respect to the propagation direction. The wave equation describes that the curvature of the displacement with respect to time is determined by the curvature of the displacement with respect to location and a constant.

This traveling wave can be detected by the human auditory apparatus. An overview of human auditory apparatus is illustrated in Figure I-1. The mechanical vibration propagates across various media: air outside the outer ear, tympanic membrane, solid bone (i.e., ossicles) in the middle ear, oval window, fluid in the inner ear, and finally an elastic plane-like tissue called "basilar membrane" in the cochlea. This oscillation is picked up by tactile sensory cells called 'hair cells' and converted into electro-chemical neural signal. In evolutionary biology (Dudel et al., 2013), the vestibular-auditory system is believed to originate from the lateral-line of fish that detects water-flow change. This interesting idea provides an explanation about the common auditory and vestibular systems in the inner ear, which both detect the mechanical displacement of the membranes stimulating the hair cells.
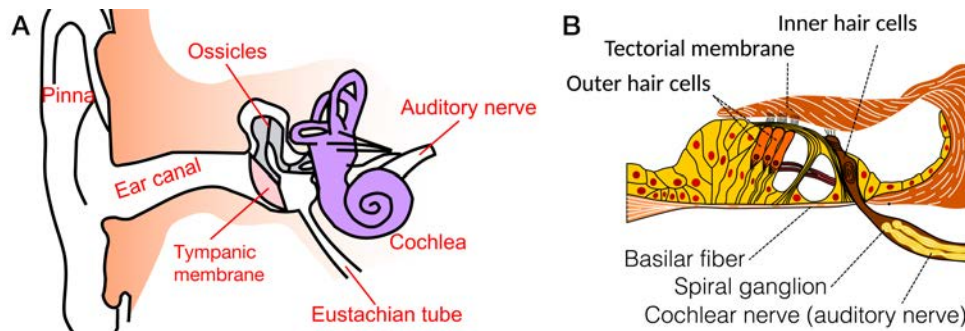
Figure I-1. **Human auditory apparatus**. Schematic drawings for (A) outer, middle, and inner ears and for (B) cochlea. Drawings are licensed under the terms of Creative-Commons-BY-SA 3.0: (A) created by Iain https://commons.wikimedia.org/w/index.php?curid=8759875; (B) created by Madhero88 https://commons.wikimedia.org/w/index.php?curid=6888273

An intriguing point about the basilar membrane in the cochlea is that its local elasticity differs across locations due to the gradual change of thickness. This mechanical feature results in spectral decomposition of the delivered wave. In other words, a group of hair cells around the basilar membrane works as a bank of band-pass filters (Patterson et al., 1995). Since hair cells in different places would encode different frequencies, it is called 'place coding' of frequency, along with the 'temporal coding', which is commonly observed from auditory brainstem response to sustained sound (Clark, 2013). Other than the local rigidity (or elasticity) of basilar membrane, it is also interesting that the global length of basilar membrane (loosely constrained by the size of head) also determines auditable range. For instance, the length of basilar membrane in a human is 30 mm with an audible range between 0.02-25 kHz whereas that in a bat is 16 mm with an audible range between 100-400 kHz (Neuweiler and Schmidt, 1993).

Then, the auditory information travels from the hair cells to the cochlear nucleus (CN) in the brainstem (pons, in particular) by auditory nerves (or 'cochlear nerves'), which are the axons of the spiral ganglion cells (i.e., uni-/bipolar cells that connect hair cells and cochlear nucleus).

I-1.2. Subcortical Auditory Systems

The auditory brainstem includes the cochlear nucleus (CN), the superior olivary complex (SOC), the lateral lemniscus (LL), and the inferior colliculus (ICC). The auditory thalamus comprises the medial geniculate body (MGB). The auditory subcortical structures are visualized in Figure I-2.

Remarkably, subcortical structures get binaural inputs from interhemispheric connections at multiple levels, which indicates that computations requiring binaural inputs (for instance, localization of the acoustic source) take place from very early stages of audition. Another interesting feature of the subcortical auditory system is its tonotopic organization, which is a

topological organization of neurons in terms of preferred frequency and is retained through the subcortical auditory stream from the CN to the MGB, and further to the cortex. From the MGB, finally the neural representation of sound goes into the primary auditory cortex (PAC).
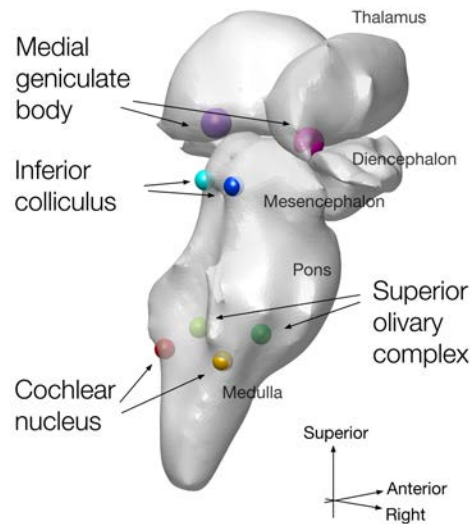
Figure I-2. Subcortical auditory systems. The isosurfaces of subcortical structures were created from automatic segmentation of a standard template in FreeSurfer. Spheres were placed for the coordinates determined as follows: CN and SOC are located based on histological criteria (Hawley et al., 2005). Stereotaxic coordinate of LL is unknown, thus no spheres were drawn for the LL. ICC is manually determined by its pronounced bump in the T1w image. MGB is chosen from the PickAtlas (http://fmri.wfubmc.edu/software/pickatlas).

I-1.3. Primary and Non-primary Auditory Cortex

Primary and non-primary human auditory cortices (PAC and non-PAC) are visualized in Figure I-3. Macroscopically, PAC corresponds to the medial two thirds of Heschl's gyrus (HG) (Morosan et al., 2001) whereas non-PAC regions (non-PAC) include planum temporale (PT; posterior to HG) and planum polare (PP; anterior to HG) on the supratemporal plane, and lateral superior temporal gyrus (LSTG) and superior temporal sulcus (STS) on the temporal convexity (Clarke and Morosan, 2012). Cytoarchitectonically, the layer IV (or inner granular layer) is well developed in PAC, which indicates strong thalamic projections from the medial geniculate body (MGB) to the layer IV (Clarke and Morosan, 2012). Compared to adjacent non-PAC regions, the PAC is easily distinguishable by "the predominance of small granular cells across all layers" (Clarke and Morosan, 2012). Myeloarchitectonically, dense cortical myelination from layer III to the white matter is typically found in the PAC (Clarke and Morosan, 2012), which is the heaviest myelination of all temporal regions (Nieuwenhuys, 2013a).

For non-PAC, the most well-studied non-PAC region is PT, due to its relevance for language comprehension based on very early neurological observations (i.e. Wernicke aphasia). PT is known for a cell density in layer V that is higher than in PAC but lower than in LSTG, and a myelin density that is smaller than in PAC but greater than in LSTG (Clarke and Morosan, 2012). An architectonical segregation of the PT into anterior and posterior halves was proposed by von Economo and Horn (1930) (TBa and TBp for anterior and posterior, respectively) and Hopf (1954) (tpartr and tsep.m), as recently replicated by Morosan et al. (2005) (Te2.1 and Te2.2). The PP is known for its relatively thin cortical sheet, compared to other auditory regions and for relatively prominent infragranular layers (Clarke and Morosan, 2012).

The non-PACs on the temporal and parietal convexities, which have formerly been collectively known as Brodmann area 22, were subdivided into Te3 (LSTG) and Te4 (STS) based on cyto-/receptor-architectonic data (Morosan et al., 2001; Morosan et al., 2005). Te3 was characterized by a dense distribution of pyramidal cells in layer III and thick (although thinner than PAC) granular layers II and IV. Receptor-architectonically, Te3 was characterized by lower densities of noradrenergic receptors as well as glutamatergic receptors compared to Te4 (Clarke and Morosan, 2012).

Segregation of auditory streams has been theorized based on the microarchitectonic heterogeneity of the non-PAC regions on the supratemporal plane and relevant behavioral and perceptual in-vivo imaging studies. The subregions of the non-PAC on the STP are AA (anterior), MA (medial), ALA (anteriolateral), LP (lateroposterior), PA (posterior), LA (lateral), and STA (superior temporal) auditory areas (Wallace et al., 2002). Human neuroimaging studies suggest that ALA and AA regions (i.e., PP) are involved in the ventral auditory stream, whereas LA, PA, and STA regions (i.e., PT) are engaged in the dorsal auditory stream (Clarke and Morosan, 2012). The dual-pathway hypothesis will be again discussed in later sections.
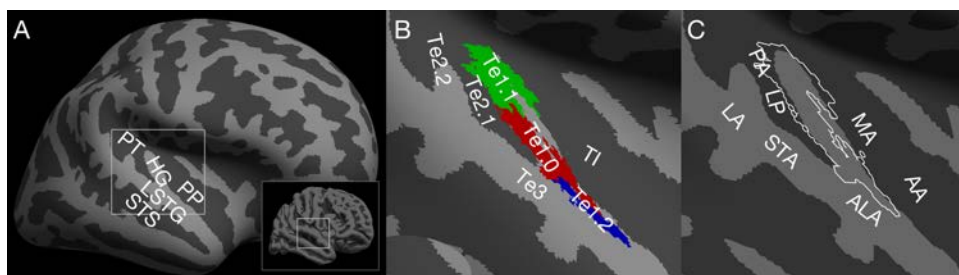
Figure I-3. Subregions of the human auditory cortex (AC). (A) Overview of the right hemisphere with macroscopic notations (i.e., HG, Heschl's gyrus; PP, planum polare; PT, planum temporale; LSTG, lateral superior temporal gyrus; STS, superior temporal sulcus). (B) Subregions of the primary auditory cortex (PAC; Te 1.0/1.1/1.2 in red, green, blue, respectively) in Jülich histological atlas based on cyto-/myeloarchitectonic data (Morosan et al., 2001) as included in FSL (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases) with non-PAC

regions (Te2/2.1/3 and TI) implicated in (Morosan et al., 2005). (C) Subregions of non-PAC outside of the PAC (white contour for Te1.0/1.2) based on Clarke and Morosan (2012); (AA, anterior auditory; MA, medial auditory; ALA, anteriolateral auditory; LP, lateroposterior); PA, posterior auditory; LA, lateral auditory; STA superior temporal auditory).

I-2. Pitch Extraction

One of the most important properties of a sound is pitch, which is the topic of this section. I will discuss about pitch models, in order to demonstrate that pitch extraction requires complex computations rather than just the simple spectral decomposition performed at the cochlea level. Amongst various issues in modeling pitch perception, I will briefly review the basic problems, biologically plausible models, and related empirical studies.

I-2.1. Pitch Model

The pitch of a sound is often misunderstood to have one-to-one correspondence with the principal frequency in the sound, thus considered to be straightforward. However, such a one-to-one relationship only exists for pure tones. Given that we perceive pitch from various natural sounds, which have different spectra, it is difficult to define pitch only based on physical properties. Instead, psychoacoustic aspects (i.e., not yet solved physiological mechanisms) are placed in formal definitions (Plack et al., 2006). For instance, the American National Standards Institute defines pitch as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high" (ANSI, 1973). If we try to mimic human pitch perception, this point becomes more explicit.

In a nutshell, there are two approaches to model how the auditory system extracts pitch from a sound (De Cheveigne, 2006): spectral and temporal approaches. The spectral approaches are mainly based on Fourier analysis to decompose frequency components whereas the temporal approaches are generally based on auto-correlation to detect periodicity (De Cheveigne, 2006).
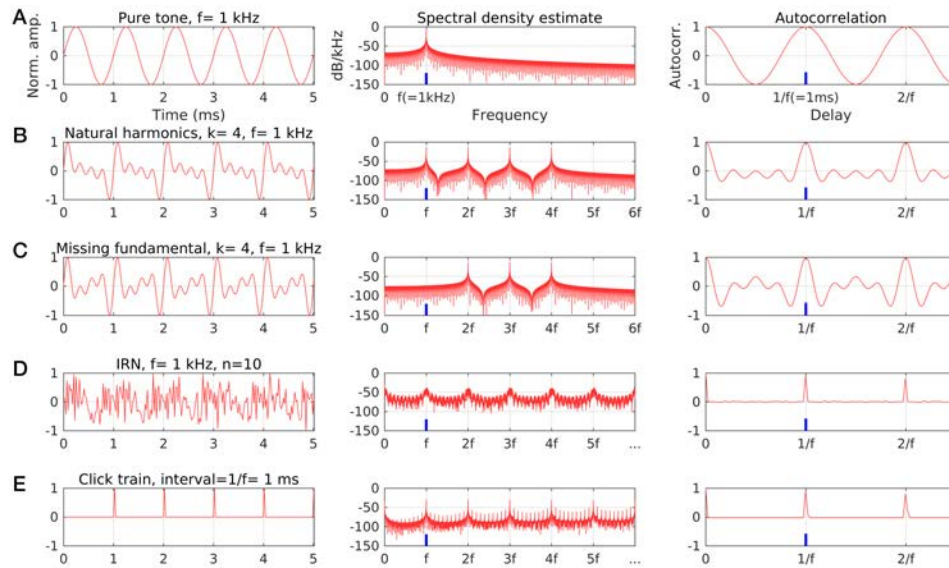
Figure I-4. Various waveforms that evoke the same perception of pitch of 1 kHz. The waveform (left column), spectrum (middle column), and auto-correlation (right column) of each waveform are presented at each row. The common perception of 1 kHz is marked by a blue vertical line at the frequency f in the spectra and the period (1/*f*) in the auto-correlograms. An audio file of the examples can be downloaded from: https://goo.gl/uSq1Xe

First, consider a sine wave over time $y(t)$ with the frequency f = 1 kHz in Figure I-4 (A), which can be written as:

$$y(t) = \sin(2\pi tf),\qquad (\text{I-2})$$

In this case, the single peak at the frequency f in the spectra directly maps to our pitch perception of f.

However, in case of a complex tone with natural harmonics up to the 4-th order in Figure I-4 (B) as:

$$y(t) = \sum_{k=1}^{4} \sin(2\pi tkf),\qquad (\text{I-3})$$

there are four multiple peaks at *f, 2f, 3f, 4f* with identical power. Thus we can modify the definition of pitch as the lowest frequency peak (i.e. *f*).

But then, with this definition, we cannot find the pitch from the famous 'missing fundamental' (Fletcher, 1924) in Figure I-4 (C) as:

$$y(t) = \sum_{k=2}^{4} \sin(2\pi t k f), \tag{I-4}$$

which does not contain any component with the frequency f but evokes the pitch perception of f. This would be a good example of the motivation of the introduction of the concept of 'periodicity' to describe temporal regularity in the sound besides frequency; there is no peak at the frequency f in the waveform, but it does have regularity (i.e., a complex waveform pattern) that repeats every millisecond ($1/f$).

Other than sine waves, white noise or a step function can also evoke the sensation of pitch with the periodicity. The 'iterated-rippled-noise' (IRN) (Yost, 1996) is created by iterations of summing delayed broadband white noise in Figure I-4 (D), which is given by:

$$y_i(t) = y_{i-1}(t) + y_{i-1}(t-d), \tag{I-5}$$

for the i-th iteration with the delay d ($=1/f$), and the time series at the 0-th iteration $y(0)$ is white noise from a uniform distribution. What is interesting about the IRN is that it provides a parameterization of 'tonality' (Griffiths and Hall, 2012), which is also another famous topic in psychoacoustics. Each iteration increases regularity (or periodicity) of the waveform and one can perceive 'a tone' appearing from the broadband noise, which already hints about the physical necessity for pitch perception. However, in terms of pitch extraction, it is not very surprising that an IRN after 10 iterations is perceived as pitch, because we can clearly see peaks at the natural multiples of f (i.e. natural harmonics of f).

This is also the case for click train (i.e., a point function) with a regular interval d of 1 ms ($=1/f$) in Figure I-4 (E) as:

$$y(t) = \begin{cases} 1 & \text{if } t \text{ is a multiple of } d \\ 0 & \text{otherwise} \end{cases}, \tag{I-6}$$

which also shows the natural harmonics of f from its spectra.

Other than the spectral approaches, the temporal approach uses autocorrelation seeking the maxima at the lags of zero and the multiples of its period (i/f) where i is an integer. Unlike the spectral approach, the temporal approach can apply the same rule also for the missing fundamental complex tone (I-4) to compute the pitch.

I-2.2. Pitch Process in Auditory Systems

A pitch extraction model based on autocorrelation generally works very well. For instance, a model that uses autocorrelation to model neural activities of a neural population ('chopper units') in the mammal ventral CN successfully predicted electrophysiological data of the auditory response suggesting its biological plausibility (Wiegrebe and Meddis, 2004).

A model of Langner (Langner, 1997) described the role of ICC as a 'coincidence detector', which provides neural integration of the outputs of CN that code frequency and periodicity separately. The idea originated from an electrophysiological study of the cat's auditory system using amplitude-modulated complex tones (Langner and Schreiner, 1988).

Griffith and colleagues discussed about the potential regions for 'pitch center' (Griffiths and Hall, 2012). As illustrated in the review (Griffiths and Hall, 2012), both in human and macaque monkey, BOLD activation in the PAC related to frequencies below the human audible range was weaker compared to that in the audible range. On the other hand, this effect of frequency was not found in the ICC, suggesting the sensation of pitch is in greater agreement with the PAC activation compared to the subcortical activities.

Another interesting fact about the IRN is that it can differentiate pitch-onset from sound-onset because both the IRN and the original white noise deliver equivalent energy. It was demonstrated that pitch onset is reflected by a response at around 170 ms after the onset of the IRN (Seither-Preisler et al., 2014), different from the common sound-onset responses such as P50, N100, and P200. Neuroimaging studies using IRN often found pitch-evoked responses in the lateral HG and anteriolateral region of the PT (Plack et al., 2014).

I-2.3. Perceptual Bias in Pitch Extraction of a Complex Tone without Low-order Harmonics

As seen in the example of the 'missing fundamental' (I-3), it appears that the pitch is extracted from the relationships amongst spectral components. If one perceives the lowest component of the complex tone (i.e., $2f$) as its pitch, the ratio of the components would be 1, 1.5, and 2. On the other hand, if one hears the non-existing component f as its pitch, then the ratio would be 2, 3, and 4. Given that most of people perceive the missing fundamental f as the pitch of the complex tone, pitch extraction seems to rely on not only the existing components but also the relationship between the components. Hypothetically, one can imagine a 'pitch template' cortical region that gets inputs from all natural harmonics of the frequency f. This group of neurons would fire when a missing fundamental complex tone lacking f is perceived, because of inputs from other components. Presumably this matching process might take place in PT because of the vast number of inputs and outputs from/to the posterior non-primary auditory cortex (Griffiths and Warren, 2002).

Then, how similar should the spectrum of a sound be to evoke the same pitch perception? This interestingly question has been previously studied in a large number of musicians (Schneider et al., 2005; Wengenroth et al., 2010). The aim of these studies was to parameterize the perceptual tendency (or 'preference') to 'recover' missing lower components and to find its neural correlates. The authors indeed found a correlation between the perceptual tendency and the functional and structural asymmetry indices based on auditory-evoked field and the volume of Heschl's gyrus

from anatomical MRI (Schneider et al., 2005). More recently, the perceptual bias was found to be related to the frequency-following response (FFR) from the brainstem and the cortical gamma-band activities recorded by EEG (Coffey et al., 2016b). More interestingly, in the study (Coffey et al., 2016b), the authors reported top-down switching between the auditory perceptual modes (i.e. either based on the missing fundamental component or the existing components) with accompanying neural activities within the same subject (Coffey et al., 2016b).

I-3. Pitch Chroma Perception

I-3.1. Pitch Chroma and Pitch Height Model

Pitch can be thought as a one-dimensional entity over the audible frequency band (i.e., approximately 20 Hz ~ 20 kHz). However, general human perception finds the multiples of a given frequency (i.e., *f*, *2f*, *3f*, …) very 'similar' compared to the other tones, even if the actual frequency difference would be smaller (e.g., *f* < *g* < *2f*) (Moerel et al., 2015). This perceptual phenomenon is reflected in the Western musical system, where two dimensions, although they are orthogonal only within an octave, called 'pitch chroma' and 'pitch height' (or in short, just 'chroma' and 'height') are used to describe a pitch (Krumhansl, 1995).

In the most commonly used Western scale called 'twelve-tone-equal temperament' (12-TET), one octave is equally divided into 12 tones. When f_1 denotes the pitch of the first (reference) tone, the pitch of the i-th tone f_i as:

$$f_i = \varphi(\frac{i-1}{12})f_1,\tag{I-7}$$

where $\varphi(k) = 2^k$ for the simplicity of notation. usical unit of an interval between two tones is called 'semitone' and one semitone corresponds to the ratio of $\varphi(1/12)$ in 12-TET. Thus a musical interval of the i-th and j-th tones (i.e., i –j semitones) between two pitches f_i and f_j can be written as:

$$i - j = 12\log_2\frac{f_i}{f_j}.\tag{I-8}$$

Alphabetical labels of the 12 tones (i.e., C, C#, D, D#, E, F, F#, G, G#, A, A#, B) are called 'pitch chroma'. Since pitch chroma recurs per each octave, there are multiple pitches with the same pitch chroma but in different octaves. A set of these tones is called 'pitch class'. In order to distinguish tones in a pitch class, an integer index for its octave is used such as "A4" (440 Hz) and "A3" (220 Hz). These indices are called 'pitch height'.

While changes in frequency would alter pitch chroma and/or pitch height (e.g., an increase of one semitone from B3 is C4), changes in pitch height do not alter pitch chroma (e.g., B3 to B4). Changes in pitch chroma (e.g., one counter-clock-wise step from B to C) may occur with change in pitch height (e.g., B3 to C4) but not necessarily (e.g., B3 to C3). These relationships have been often illustrated using a helix model as given in Figure I-5.
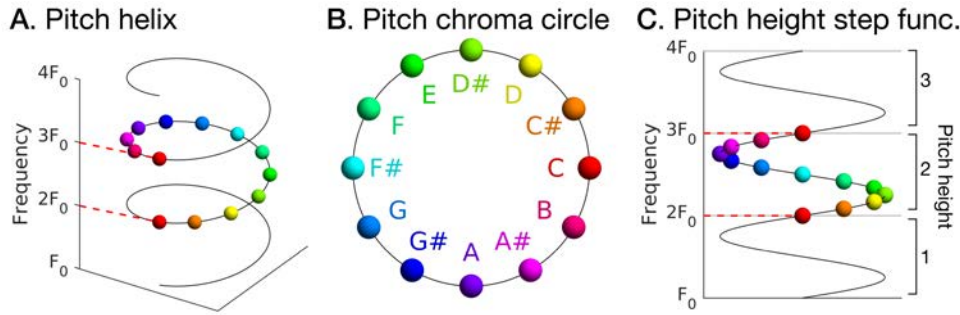
Figure I-5. Pitch helix model. (A) 3-D overview of the helix with the pitch is given along the Z-axis. (B) Pitch chroma is given as a polar coordinate on the X-Y plane. (C) Pitch height is given as an integer along the Z-axis.

Note that both pitch chroma (polar coordinates) and pitch height (step function) affect pitch as a whole, but chroma and height are orthogonal based on the definitions here. For a given pitch p (Hz), a pitch height h can be found as:

$$h = \left\lfloor \log_2 \frac{p}{f_0} \right\rfloor, \tag{I-9}$$

where $\lfloor \cdot \rfloor$ is a floor function, which returns the largest integer that is smaller than a given number. f_0 is a reference frequency of "C-zero" (approximately 16.3516 Hz) under the modern tuning convention (i.e., A4 = 440 Hz). Subsequently, for the given h and p, the pitch chroma c is determined using (I-8) as:

$$c = 1 + 12 \log_2 \frac{p}{f_0 \varphi(h)}, \tag{I-10}$$

where $c \in \{1,2,3,\cdots,12\}$ is an index for pitch chroma such that $c = 1$ for "C", $c = 2$ for "C#", ..., c = 12 for "B" in the 12-TET. As abovementioned, $\varphi(k) = 2^k$ for the simplicity of notation. To describe subsemitone deviation, c can be a non-integer, i.e., $c \in [1,13) \subset \mathbb{R}$ [2].

Inversely, pitch p can be found for given pitch chroma c and pitch height h using (I-7) as:

$$p = \varphi(\frac{c-1}{12} + h)f_0. \tag{I-11}$$

It should be noted that the choice of the reference pitch is purely arbitrary. Conventionally the circle of pitch chroma starts from C to B, thus one semitone increase from B3 to C4 might appear as an abrupt increase in pitch height, at least from the notation. However, because any choice of the reference pitch is equally possible, it does not have any meaning other than the change in notation.

Based on the current definition, it is clear that a system to compute h and c requires not only p but also f_0. In other words, any form of information about the reference f_0 is necessary to recognize pitch chroma. This reference f_0 can be given externally as an explicit reference or a tonal context. Another possibility is to have internal reference in the system. In the following subsections, first I will discuss how the external reference frequency works in perceiving pitch chroma.

I-3.2. Similarity within a Pitch Class

The simplest form of pitch chroma perception would be to determine whether two tones with different pitches have identical pitch chroma or not. If there is a group of neurons that responds to tones spaced by one or several octaves (which presumably could have developed by rich harmonic structures of natural sounds accordingly to the Hebbian rule), this neural group will get excited by the tones in a pitch class. This idea was recently tested in an fMRI study (Moerel et al., 2015), which identified a candidate for a neural basis of the pitch chroma perception by fitting auditory responses to piano tones onto a predicted tuning curve based on a helix model for harmonic series. Previously, tuning curves with multiple peaks in the AC were derived based on tonotopy mapping using fMRI (Moerel et al., 2013). The existence of pitch-class-selective neurons was reconfirmed using piano tones with natural harmonics, however the location of such neurons was found to be diffused over the bilateral STPs (Moerel et al., 2015).

---

[2] If we define pitch chroma as $c \in [0,1) \subset \mathbb{R}$ and the reference frequency as $f_0 = 1$ Hz, chroma can be more concisely written as $c = \log_2 p - \lfloor \log_2 p \rfloor$ as given in Wakefield GH. Mathematical representation of joint time-chroma distributions; 1999. International Society for Optics and Photonics. p 637-645.

I-3.3. Pitch Chroma Recognition in Tonal Context

A more natural setting where we recognize pitch chroma would be 'tonal context' (Krumhansl and Shepard, 1979). By virtue of the established tonality, the basis of a musical scale can work as a temporary reference. A musical scale is a series of tones with specific intervals. That is, the intervals between the first tone and other 6 tones in a major key are (2, 4, 5, 7, 9, 11) semitones whereas such intervals of a minor key are (2, 3, 5, 7, 8, 11) semitones as shown in Figure I-6.
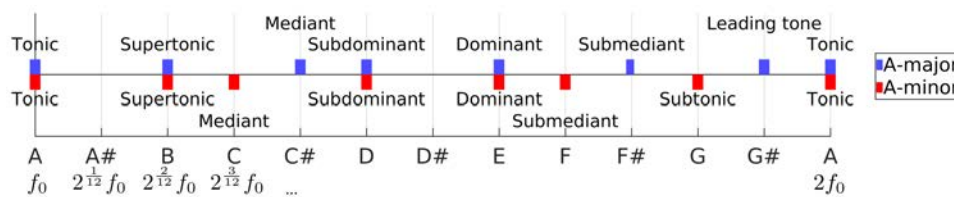
Figure I-6. Musical scales. Along the horizontal axis, pitch chroma (i.e., A, A#, …) and frequency based on a reference tone f_0 are noted. 'Tonal functions' (i.e., tonic, supertonic, …) are given for the A-major key (blue) and the A-minor key (red).

A 'tonal function'[3] of a tone in a key (e.g., A in A-major) describes its distance from the basis (the first tone called 'tonic') of the key and associated 'attraction' (or evoked expectation) on the following tone. Establishment of tonality requires presentation of the whole scale, frequent occurrence of cardinal tones (e.g., tonic and dominant), or canonical chord progressions (e.g., authentic cadence). Once the tonality is established, the general population (with and without musical background) would perceive the stability and attraction of tones differently; i.e., a certain tone (e.g., tonic) sounds 'more stable' than another one (e.g. subtonic) as found in behavioral studies using subjective rating (Krumhansl and Shepard, 1979).

In the context of a musical scale, educated listeners with musical experience can recognize the tonal function of a given tone with respect to the basis of the key (or equivalently the interval between the tone of question and the basis). This kind of perception is called 'relative pitch'[4] (Miyazaki, 1990). Using the perceived tonal function and additional information of the key, one

---

[3] More commonly, solfeggio is used to refer to a tonal function: e.g., *Do*, *Mi*, *Sol* is the first, third, and fifth note of a major scale, respectively. Such a system is called 'Movable-*Do*' whereas 'Fixed-*Do*' means using solfeggio of C major in all keys (e.g., reading the tonic of A-major as *La* instead of *Do*)

[4] Often, in neuroscientific and behavioral research, the term 'relative pitch' is used as a counterpart of the 'absolute pitch' to describe musicians without absolute pitch. However, the 'relative pitch' (or perception of intervals) is one of the basic skills that all musicians (with and without absolute pitch) develop through formal musical training such as sight-singing and hearing. In fact, for musicians with absolute pitch, perceiving tonal functions could be more difficult due to Stroop-like interference (Schulze K, Mueller K, Koelsch S. (2013): Auditory stroop and absolute pitch: an fMRI study. Hum. Brain Mapp. 34(7):1579-90.).

can calculate the pitch chroma (e.g., the dominant in A-major is "E" whereas that in C-major is "G"). That is, pitch chroma recognition using relative pitch perception is an effortful and explicit process (although it can be nearly automatized by intensive training) to link separate pieces of information together.

I-3.4. Pitch Chroma Recognition Based on Internal Reference

In contrast the to abovementioned case, some people effortlessly recognize pitch chroma without external reference in an absolute sense. This unique perceptual and cognitive ability is known as 'absolute pitch' (AP) (Miyazaki, 2004a), which is the main topic of the current thesis. Given that the recognition of absolute chroma requires a reference (i.e., f_0 in the equation (I-10)) in any form, the fact that AP listeners can recognize pitch chroma indicates there must exist an internal reference at some point during the auditory processing. I will closely review related studies with behavioral, electro-/magneto-encephalographic, neuroimaging, and genetic measurements about AP with an emphasis on the perceptual aspect of AP.

I-4. Absolute Perception of Pitch Chroma

I-4.1. Behavioral definition and characteristics of AP

As mentioned above, AP is essentially defined by the ability to recognize pitch without any external reference. Miyazaki, who has been working on AP for several decades, defines AP as "an ability to identify the pitch of isolated tones using musical pitch labels or to produce the pitch of any tones designated by note names without comparing with any reference pitch" (Miyazaki, 2004a). I argue that the emphasis should be placed on pitch chroma other than pitch itself (including pitch height). It has been consistently shown that AP perception does not enhance the determination of octave (i.e., pitch height), as shown in several behavioral experiments (Deutsch, 2013; Deutsch and Henthorn, 2004; Kim and Knösche, 2016; Miyazaki, 1988; Takeuchi and Hulse, 1993).

   The earliest scientific description of AP that can be found using Google Scholar[5] and Web of Science[6], is a report by Max Meyer on AP performance of himself and his colleague (Meyer, 1899)[7].

---

[5] https://scholar.google.com/

[6] https://webofknowledge.com/

[7] An earlier publication (Kries Jv. (1892): Über das absolute Gehör. Zeitschrift für Psychologie und Physiologie der Sinnesorgane 3:257-279.) was cited in the article (Meyer M. (1899): Is the Memory of Absolute Pitch Capable of Development by Training? Psychol. Rev. 6(5):514-516.). Thus Meyer (1899) is not the very first scientific publication on the absolute pitch but simply the first article of which the content is accessible online.

As well as in modern psychophysical studies, pure tones (using tuning forks) and piano tones (using an actual piano) were used for experimental stimuli. The author carried out behavioral experiments to find out how long the learning effect of AP lasts, in favor of the notion that AP can be obtained and retained by adult humans, in opposition to a previous publication (Kries, 1892), which suggested that acquiring AP after the developmental period is very unlikely. Although their experimental design bears multiple weaknesses, it marks the introduction of the scientific approach to the investigation of psychoacoustic problems in the late 19th century (Meyer, 1899). Given that Wilhelm M. Wundt opened the very first experimental psychology lab at Leipzig University in 1879, it seems that AP has consistently drawn the attention of experimental psychologists as a special kind of auditory perception.

Later in the late 20th century, Siegel carried out behavioral studies using computers to refine stimulus manipulation, control experimental protocols, and measure precise reaction times (Siegel, 1974; Siegel and Siegel, 1977; Siegel, 1972). The former study investigated the decay of the pitch memory of a given melody in AP listeners (Siegel, 1972). The author studied whether verbal labeling enhances the pitch memory of AP listeners. When the verbal labeling was hindered (e.g., when verbal fillers are presented between a reference tone and a target tone or when slightly deviated tones that are hard to verbally encode are presented), the performance of AP listeners decreased. The insightful report stated that "the subjects with AP categorize pitch on an absolute basis in the same way they categorize letters, words, or common objects" (Siegel, 1974).

Throughout the literature reporting the various aspects of behavioral characteristics of AP, the definition of AP becomes operational in terms of behavior. One simplest criterion of AP would be the 100% accuracy in recognizing pitch for all kinds of tonal sounds, which should be related to the layman's term 'perfect pitch' for AP. However, once a researcher starts quantifying the performance of AP by means of behavioral measurements, it becomes apparent that there is more variability in AP performance: some AP possessors need more time and make more mistakes than others, although their AP responses are still superior compared to ones without AP.

Back in the mid 20th century, Bachem (1955) deliberately expressed his 'concern' about calling a slight improvement of adults in pitch identification tasks after a short-term training as learning 'absolute pitch'. So he suggested referring to such ability as 'pseudo-absolute pitch'. In fact, it has been argued that this type of 'implicit absolute pitch' (Miyazaki, 2004a) of non-musicians (i.e., remembering very familiar tones such as line dial tone or the tones of one's favorite song) shares the same mechanism with the highly fluent AP of some musicians, differing only in its degree (Levitin, 1994; Smith and Schmuckler, 2008) and that it would be possible to increase that degree by training even for adults (Van Hedger et al., 2015). Miyazaki also suggested a term 'quasi-absolute pitch' to describe APs with a limited number of internal references. For example, a violist might know only one tone ("violin A") with confidence and in order to determine the pitch of any other tone, he/she requires some time to retrieve the internal reference tone and compute the

interval between the reference and the target to estimate its pitch chroma. In both cases of pseudo-/quasi-AP, Miyazaki pointed out that the accuracy is worse compared to the case of 'genuine absolute pitch'. He suggested some unique features of the 'genuine-AP' that distinguish it from the pseudo-/quasi-APs: (1) quick reaction time, (2) high pitch discrimination resolution (less than 10 cents), (3) no limit in timbres and range, (4) 'octave errors' with consistency (Bachem, 1955). The 'octave error' is defined by an incorrect pitch height with a correct pitch chroma. These features reported by Miyazaki are particularly interesting, because his ideas based on qualitative observations are in partial agreement with the later behavioral studies, as will be detailed hereunder:

(1) Rapid reaction time is indeed what many studies find consistently. For an extreme example (Miyazaki, 1990), some Japanese participants (majoring in music education) were able to respond within 0.6 s (by pressing a key on a muted digital piano) while maintaining accuracy at over 80 %. In an independent study (Bermudez and Zatorre, 2009b), the rapid reaction time of musicians with AP compared to musicians without AP was replicated (i.e., 2.5 vs. 6.3 s). The reaction time could be an useful index for further categorization within the musicians with AP (Bermudez and Zatorre, 2009b) although no reaction time difference (in verbally responding) was found between genuine-AP and quasi-AP listeners (i.e., 2.1 vs. 2.3 s) (Wilson et al., 2009).

(2) Apparently, subsemitone accuracy in pitch discrimination was found to be not different between musicians with and without AP in several studies (Siegel, 1974; Siegel, 1972; Takeuchi and Hulse, 1993). In one of the experiments (Siegel, 1974), AP listeners showed no better performance than a control group for a subsemitone interval of 10 cents (i.e., 0.1 semitone) compared to a larger interval of 75 cents, although the task was not purely perceptual but involved verbal encoding and decoding of tones for a memory test (see Study #1 in the present thesis for a perceptual experiment).

(3) It is observed that certain individuals with extreme proficiency of AP (i.e., 100 % accuracy without any octave errors and with short reaction times) can identify the pitch chroma with any timbre and any range without difficulty (Miyazaki, 1988). However, what is more often found is a limited generalizability of their internal reference in terms of timbre and octave range. This specific preference presumably due to familiarity (usually limited to their favorite/major musical instruments) has been a known feature of AP for long enough to call it 'absolute piano' or 'absolute violin' instead of AP (Ward, 1985). However, it should be also noted that unfamiliarity to the timbre of perceived tones only affects the pitch height decision, not the one on pitch chroma (Miyazaki, 1988).

(4) The occurrence of octave errors is one of the very well-known features of AP; not in the sense that APs are more prone to octave errors, as Bachem suggested, but because of the interesting dissociation between their precise recognition of pitch chroma and not better than normal recognition of pitch height (Deutsch, 2013; Deutsch and Henthorn, 2004;

Miyazaki, 1988; Takeuchi and Hulse, 1993). This dissociation is particularly interesting because it tells specifically to which aspects of pitch AP actually refers, i.e., pitch chroma. Moreover, it implies that chroma and height are separately processed, at least in AP musicians' auditory systems, which guided this thesis towards a hypothesis on distinct neural pathways for chroma and height.

(5) In addition to the points mentioned in Bachem (1955), another interesting aspect of AP is spontaneity and uncontrollability. This is impressively demonstrated by AP musicians' ability to recognize the pitch chroma of glasses clattering, the siren of an ambulance car, or people's speaking tone without efforts (Miyazaki, 2004a). The level of automaticity of AP is very high, thus it is usually difficult to control (although it is possible through training), as demonstrated by Stroop-like effects, namely slower and inaccurate responses to incongruently labeled tones (i.e., movable-Do solfeggio) (Akiva-Kabiri and Henik, 2012; Itoh et al., 2005; Schulze et al., 2013). Apart from in musical composition and conducting, AP is not generally regarded as being musically beneficial. Instead, it could even be disadvantageous for players or singers because of the uncontrollability. For instance, musicians with AP may have difficulties recognizing transposed melodies (Miyazaki, 2004b).

Even though a number of unique features of AP are known, so far there are multiple ways to classify musicians into non-AP, pseudo-AP, quasi-AP, and genuine-AP. As already mentioned, one study made boundaries at 20% and 90% hit-rates (Wilson et al., 2009), another study used 90 % threshold in tests with pure tones and piano tones for AP (Keenan et al., 2001), and yet another study drew the line between non-AP and AP at the saddle point of a bimodal distribution of the performance of AP of all subjects (at around greater than 51 % of scores) (Wengenroth et al., 2014).

Partial scoring is another issue. In order to declaim the possession of absolute pitch, some researchers simply give no score for errors even with a few semitones (Miyazaki, 1990), others give full credit for one semitone error (Ward and Burns, 1982), partial credit for the error of two semitones (Wengenroth et al., 2014), or the scores are proportional to the inverse of the errors (Bermudez and Zatorre, 2009b; Kim and Knösche, 2016).

It would be the best to find criteria of classification or a comprehensive scoring based on a large-scale dataset. In fact, a project to find genetic factors of AP published data obtained via a web-based test from 2,213 individuals (Athos et al., 2007). In this dataset, a clear bimodal distribution over a 2-D space of pure tone and piano tone scores was shown for genuine APs (very close to the maximal score, probably truncated variance due to the ease of the task) and non-APs (centered around the random choice level), but without any third mode for quasi-AP. Rather, some data points were spread over a large area between the non-APs and the genuine-APs. In the paper (Athos et al., 2007), the authors choose a cutoff point for genuine-AP at 24.5 points (out of 36 points; 68 %) only for pure tone questions with giving partial credit of 0.75 for one-semitone errors.

This cutoff point appears to be liberal compared to that of other studies (90 %) with full credit for one-semitone errors (Gaab et al., 2003; Keenan et al., 2001). However, the proportion of one-semitone errors was not reported, thus comparability of the criteria across studies is unclear. Open Science (Owens, 2016), which is a recent trend in scientific communities, might enable sharing such valuable data and thereby help to derive better criteria that are backed up by empirical evidence.

I-4.2. Acquisition and Loss of AP over Lifetime

I-4.2.1. Genetic Factors

There are studies that imply that genetic factors are important for the acquisition of AP. In a heritability study on AP (Baharloo et al., 1998), the likelihood to have at least another AP listener among relatives (e.g., siblings, parents, cousins) was 48 % for AP listeners whereas 14 % for non-AP listeners. This principal finding was significant even after controlling for the music training onset age (Baharloo et al., 2000), indicating a genetic predisposition of AP.

Furthermore, the link between the AP and synesthesia has been investigated for their rare association across modalities: e.g., pitch and verbal labels in AP listeners; pitch and color in people with synesthesia (known as 'chromesthesia'). In certain cases, people with chromesthesia perceive a specific color associated to a pitch chroma or a musical key thus they can recognize chroma without efforts (Ward et al., 2006). For instance, as famously quoted (Mahling, 1926), Hungarian virtuoso Franz Liszt requested his orchestra do play 'more bluish' not orange-like. Furthermore, a recent genetic study found phenotypic and genetic overlaps between synesthesia and AP (Gregersen et al., 2013) suggesting genetic candidates for the predisposition of AP.

Using personality trait questionnaires, an unexpected linkage to a disease with known genetic risk factors, namely autistic personality disorder was reported (Dohn et al., 2012). The idea is based on the 'cognitive style hypothesis', which assumes AP listeners focus more on the local information than on the global configuration thus recognize each tone in an absolute sense (Chin, 2003). Although Autism-like personality traits could be another phenotype of AP, the actual relationship remains largely unclear so far.

I-4.2.2. Early Experience during a Critical Period

The musical education during development (especially before the age of 6 years) has been strongly related to the acquisition of AP (Baharloo et al., 1998; Miyazaki and Ogawa, 2006) implicating that there is a critical period for the realization of AP given genetic 'predisposition' (Hallam, 2015), whereas the acquisition of AP for adults is known to be very difficult. However, some psychologists do believe that training AP in adulthood is certainly possible, as mentioned earlier (Schellenberg and Trehub, 2003; Van Hedger et al., 2015)

Theoretically, the experience during the critical period does not have to be in the Western musical system, but may involve any kind of musical system (or even non-musical systems that has corresponding labels for certain frequencies). The important feature would be a lexical system that corresponds to consistent segmentation of an octave (assuming the similarity within a pitch class is not culturally determined). In fact, a Japanese private music education program uses this idea to promote acquisition of AP of young children in the kindergarten (Miyazaki and Ogawa, 2006). Recently, a longitudinal study showed improvement in mapping triads via an educational program over three years (Sakakibara, 2014), but whether the recognition of triad helps the recognition of pitch chroma is yet to be discovered.

An intriguing demonstration of the importance of neuroplasticity in the acquisition of AP was given using a histone deacetylase inhibitor named valproate, which is known to promote neuroplasticity in animal models (Gervain et al., 2013). This study is very impressive in the sense that it directly showed the impact of neuroplasticity in acquiring AP, but the degree of improvement was still low (5/18 = 27% correct vs. 3/18 16 % by chance) as well as other learning studies with longer training sessions without medication.

It is known that AP is more commonly found in East Asians than in Europeans (30 % of Japanese musicians and 7 % of Polish musicians have AP according to Miyazaki et al. (2012)), which might be due to genetic and/or cultural factors. In an intercultural study (Miyazaki et al., 2012), the authors discussed about the possibility that the difference might be simply due to different training onset age (Asians started musical practice two years earlier than Europeans). Alternatively, it has been tried to explain the ethnicity effect by the use of tonal language (i.e., Chinese) (Deutsch et al., 2006), however this argument neglects the fact that Chinese use pitch contours as tonal cues (i.e., rapid ascending, slow descending, and so on), instead of absolute pitch.

I-4.2.3. Loss of AP in Late Phase of Life

On the one hand, it has been known that hearing loss in general (non-AP) population occurs from high to low frequencies, due to aging via various mechanisms including loss of ganglion cells, degeneration of hair cells, and basilar membrane stiffness (Liu and Yan, 2007). On the other hand, the loss of AP due to aging has been rarely discussed. An autobiographical thesis of an elderly AP listener described personal experiences of losing AP including that of herself (Bianco, 2015). This qualitative essay is especially important because non-AP elders cannot notice such mild alteration in pitch perception. For instance, suppose a frequency A stimulates a neuron A' at young age but the frequency A stimulates a neuron B' (of which characteristic frequency was B) at old age due to stiffness of basilar membrane. Because non-AP elders cannot recognize frequency A or B in absolute sense, this whole shift of all auditory input cannot be recognized in non-AP elders.

In her thesis (Bianco, 2015), a couple of interviewees described common distortion (i.e., overall shift by 1 or 2 semitones higher) of the AP. For instance, an aged AP listener, who suffers from

distorted AP, hears a song in A-major as if it was A# or B-major. This heightened perception of pitch chroma in elderly AP listeners might be due to an overall decrease of elasticity of the basilar membrane that shifts vibration loci for a certain frequency towards the thinner end. This could result in activations of hair cells at the thinner location. With otherwise intact auditory streams and cortical networks for AP, an elderly AP listener could recognize heightened pitch chroma confusing him/her with prior knowledge about how it should sound. The existence of the shifted AP perception may indicate that the degeneration starts from the peripheral apparatus rather than cortical networks.

I-4.3. Previously Found Neural Correlates of AP

Motivated by the arrival of in-vivo neuroimaging techniques, a line of neuroimaging studies has been published since the late 1990's. In this section, I will review morphological, functional, connectional, and neurological findings with an emphasis on the temporal structures. A commonly accepted intuitive segregation of the AP recognition separates the 'perceptual' component (i.e., chromatic categorization of pitch; related to the temporal lobes) and 'associative' components (i.e., link between a representation of pitch chroma and a verbal coding (Zatorre, 2003) and/or a motor sequence (Wong and Wong, 2014); related to frontal lobes). I believe a truly interesting aspect of AP is the perceptual component with its connectivity within the frontotemporal network as it demonstrates a very special (and more common than synesthesia) processing of pitch and can contribute to the understanding of human auditory perception in general (Zatorre, 2003).

I-4.3.1. Morphological Findings

The first finding of neuroanatomical correlates of AP was the increased leftward asymmetry of the area of the PT based on manual segmentation on MRI slices (Schlaug et al., 1995). The group difference in asymmetry was due to a smaller area of the right PT (not a larger area of the left PT) in musicians with AP in comparison with musicians without AP (Keenan et al., 2001; Schlaug et al., 1995). In contrast, another morphometric study found that the volumes of the left and right PTs were not significantly different between the musicians with and without AP (Zatorre et al., 1998). These studies were based on manual delineation of regions-of-interest (ROIs), which has inherent issues of validity and reliability: different studies may use different anatomical criteria, an application of criteria could be different across raters, and even there could be intra-rater variability. In a later study (Wengenroth et al., 2014), the authors compared the volumes of the left and right HG and found that the right HG was larger in AP musicians than in non-AP musicians. It was noted that a posterior shift of the Sylvian fissure and the parietal-temporal convexity in the right hemispheres of the AP musicians (Wengenroth et al., 2014), thus an enlarged right HG, could have resulted in the smaller segmentation of the right PT, which also provides an anatomical explanation for the "heightened leftward" asymmetry in PT area in previous ROI-based studies (Keenan et al., 2001; Schlaug et al., 1995).

Even if the ROI definition is accurate and reliable, the ROI-based morphometry has the inherent limitation of disregarding local differences within a structure. If the alteration in question is locally confined within a specific small sub-region of cortex, an investigation on the aggregated measurement cannot detect such variability. Therefore, voxel-/vertex-wise whole brain analyses followed ROI-based morphometric studies. With this approach, new challenges arise from the difficulty of establishing perfect one-to-one correspondence across individuals, due to inherent intersubject-variability in brain structures as well as computational limitations. A voxel-based morphometry (VBM) study on the asymmetry of tissue probability of gray matter (GM) reported increased leftward asymmetry in the HG and the anterior portion of the PT in all AP musicians as well as its agreement with the ROI-based results (Luders et al., 2004).

In contrast to the previous volumetric studies, later studies using cortical thickness analysis (CTA) did not find a significant leftward asymmetry of the posterior STG (Bermudez et al., 2009; Dohn et al., 2015). Instead, a significant correlation between the AP score and the cortical thickness in the left inferior frontal gyrus (IFG) and the postcentral gyrus was found, but no significant group difference or AP performance effect in the STG (Bermudez et al., 2009). In a recent paper (Dohn et al., 2015), in both STGs (including Heschl's gyrus and lateral STG), greater cortical thickness was found in APs for inferior temporal regions and frontal structures. The apparent inconsistency between the previous VBM asymmetry study (Luders et al., 2004) and the CTA group studies could originate from the different measure (GM probability vs. cortical thickness), different analysis (group difference in asymmetry vs. group difference in the measure itself), and different subject selection procedures (with and without behavioral tests for musical aptitude; although all studies were matched for musical training onset and duration). Besides temporal regions, frontal regions including superior frontal gyrus (SFG) (Bermudez et al., 2009) and IFG (Dohn et al., 2015) were found to be structurally different between the musicians with and without AP.

I-4.3.2. Functional Findings

An early functional imaging study reported sensational finding that the left frontal region (i.e., dorsolateral prefrontal cortex, DLPFC) was constantly active in AP musicians, regardless of the explicit requirement of the experimental condition (Zatorre et al., 1998). This was especially exciting because the constant neural activity concurred very well with the behaviorally known spontaneity of AP (Ward and Burns, 1982). In subsequent functional studies, the left DLPFC has been often found to respond differently in APs compared to non-APs.

Besides the frontal regions, the left posterior STG (Ohnishi et al., 2001) and the bilateral PTs (Wengenroth et al., 2014) showed stronger BOLD signals in APs. A negative peak at around 150 ms after stimulus onset (so-called 'AP negativity') in event-related potential (ERP) recordings was found during both of passive listening and naming conditions from a left posterior-temporal electrode (Itoh et al., 2005). In the multimodal imaging study (Wengenroth et al., 2014), the

authors also reported an effect of AP score in the auditory evoked potential. From the equivalent current dipole (ECD) localized in the right anterior PT, the synchronized neural activities at around 200 ms after stimulus onset correlated with the behavioral index of AP acuity whereas neither any HG nor the left PT showed such effects (Wengenroth et al., 2014).

What does the increased activation in the auditory cortex actually indicate? Considering the source of the signals, it can be interpreted as an indication of greater coherent neural activities in local systems, especially postsynaptic membrane potentials, if detected by M/EEG, or neurovascular coupling related to presynaptic membrane potential, which can be achieved by any changes in neural density (due to genetic/developmental factors), number of synapses, or even local synchronization. In order to differentiate such possibilities and to understand the nature of such in-vivo findings, further information on the microarchitecture of neural systems is required.

I-4.3.3. Connectional Findings

As the dual components framework implies, most of the connectional findings were focused on the fronto-temporal network. Focusing on the superior longitudinal fasciculus (SLF) in the left hemisphere, a combined tractography and voxel-based analysis reported correlation between the fractional anisotropy (FA; a directionality index of a tensor that models water diffusion in white matter) and AP performance in a couple of locations along the SLF (Oechslin et al., 2009). The abovementioned anatomical study (Dohn et al., 2015) also reported white matter alteration in the right temporal lobe nearby STP and insula using another voxel-based analysis called tract-based spatial statistics (TBSS) (Smith et al., 2006). The white matter regions that showed greater FA in AP musicians overlapped with the inferior longitudinal fasciculus (ILF), which is another important fronto-temporal pathway (Dohn et al., 2015). Another study compared "volume of tracts" and "number of fibers" between the posterior STG and the posterior middle temporal gyrus (MTG) using deterministic tractography (Loui et al., 2011). Note, however that this is not an entirely recommendable practice in the analysis of diffusion-weighted imaging (DWI) data; see Jones et al. (2013) for what is encouraged and discouraged in DWI analysis. Presumably, the authors found group difference in the number of voxels that the tractography reached and the number of streamlines but not in FA values averaged over such voxels (Loui et al., 2011).

For functional connectivity (FC), there have been multiple reports for resting and listening conditions. Based on the reconstructed electrical neural activities from EEG data, greater phase synchrony between the left STG and the left DLPFC during rest in favor of AP was reported (Elmer et al., 2015). It supports AP-specific fronto-temporal interaction, which is related to the associative part of AP recognition. Although without statistical inferences, widely spread FC (over the anterior and posterior STG, insula, and IFG) during passive listening was also reported (Wengenroth et al., 2014).

Whole-brain topological studies also have been published comparing graphs from musicians with and without AP using resting-state fMRI (Loui et al., 2012) and inter-subject covariance of cortical thickness (Jäncke et al., 2012). The authors of both papers reported alteration in degree centrality averaged across all nodes in the graph in opposite directions. Unfortunately, both studies used correlation matrices and coefficient-based thresholding to compare graphs, which can be methodologically problematic given that topological analysis seeks invariant features. For an undirected binary graph with N nodes, K edges[8], and no self-connection, the average of degree centralities of all nodes is $E(\deg) = 2K/N$, while the density of the graph is $d = K / \{N(N-1)/2\}$, therefore $E(\deg) = d(N-1)$. That is, the averaged degree across all nodes is simply a product of the density and the number of nodes subtracted by one. Although the density (or inversely, sparsity) of a graph might be meaningful to some questions, it only describes the global level of correlation, which can be easily compromised in resting-state fMRI data due to head motion artifacts and also signal-to-noise level in general. Therefore, in many topological studies on the cortical networks, the density of an individual graph is usually matched across graphs in order to quantify topological characteristics (Bullmore and Sporns, 2009; Hagmann et al., 2008; Hagmann et al., 2007; He et al., 2007; Kim et al., 2013).

I-5. Contents of the Current Thesis

I-5.1. Motivation and Framework

In the previous sections, I have briefly reviewed our current understanding of the perception of pitch chroma and the absolute recognition of it. Throughout the review, it became apparent that closer investigation focusing on the perceptual component of AP is relatively scarce compared to the research on the cognitive (associative) component of AP (Bermudez et al., 2009; Elmer et al., 2013; Zatorre et al., 1998) and its interaction with other high-level cognitive functions such as tonal and verbal working memory or Stroop-like effects (Gaab et al., 2006; Itoh et al., 2005; Koelsch et al., 2009; Schulze et al., 2009; Schulze et al., 2013; Wong and Wong, 2014), possibly influenced by the view that there is only a quantitative distinction between the 'implicit AP' (suggested based on 'absolute memory', which is remembering the first tone of one's favorite song) and the 'genuine AP' (Levitin, 1994; Smith and Schmuckler, 2008). The idea is not illogical: some people have good long-term memory of a few pitches due to frequent exposures. We know that people can learn skills of very complex motor sequences and cognitive functions after long and hard training, and later perform them almost effortlessly (at least apparently). Thus if we train ourselves to access the long-term memory and compute the interval between the reference and target tone very fast, we may be able to recognize pitch like the 'genuine-AP' in terms of accuracy

---

[8] For a weighted graph, K is the sum of all weights instead of the number of all edges.

and speed. However, it should be noted that most of quasi-AP listeners (usually professional musicians with one or two reference tones) perform slower and less accurate at AP test, even with the highly trained relative pitch. In an analogy to learning foreign language, it has been acknowledged that obtaining AP perception with the same proficiency as genuine-APs requires early experience around the age of 7 years (Levitin and Rogers, 2005). Therefore, even with the similar level of AP recognition in terms of behavior, the perceptual processing of the quasi-APs should be different from that of the genuine-APs. Especially the perceptual component, i.e., chromatic categorization, would be the most distinguished process. Therefore, in this thesis, I focus on the perceptual component of AP in terms of structures and functions of the auditory cortex.

Another important motivation of this thesis is to investigate the neural underpinning of macroscopic findings that are already known. As already mentioned in Section I-4.3, a neurobiological interpretation of findings based on non-invasive imaging/recording entails technical challenges, which render developing quantitative models that describe how AP can be neurally implemented more challenging. For instance, if the smaller right PT (Keenan et al., 2001), or the greater right HG (Wengenroth et al., 2014), is a necessary condition for acquisition of AP, how is the morphology related to AP processing? To answer this question quantitatively, i.e., to build and test computational models of AP, we need to learn further detailed information of the auditory systems in AP listeners. One of the recent advances is ultra-high-field (i.e., 7-Telsa) imaging with a novel MR sequence that is optimized for high-field anatomical imaging (Marques et al., 2010) allowing quantitative estimation of physical MR properties such as longitudinal relaxation time (T1). The application of T1 estimates in quantifying myelination of local circuits (i.e., intracortical axons) has been proposed (Geyer et al., 2011). Some studies used its reciprocal R1 (1/T1, called 'longitudinal relaxation rate') as a positive index of myelination (Lutti et al., 2014). Along with another approach to map cortical myelin (Glasser and Van Essen, 2011), the interest in 'quantitative morphometry' is currently increasing. A recent study (Stuber et al., 2014) directly demonstrated that it is not only myelin but also iron that affects the MR contrast considerably (albeit weaker than myelin). Therefore another recent study used multiple imaging contrasts to differentiate different processes of aging (Callaghan et al., 2014). In the current thesis, in-vivo cortical myelin mapping was used to study the microarchitecture of auditory cortex in musicians with AP.

As a general framework, this thesis is based on the 'dual pathway' hypothesis (Rauschecker, 2015; Rauschecker et al., 1995). A line of evidence consistently found the anterior STP (forming the ventral pathway) to be sensitive to spatially invariant properties of auditory objects such as pitch, timbre, and phoneme (Altmann et al., 2007; Barrett et al., 2005; Barrett and Hall, 2006; DeWitt and Rauschecker, 2012; Hart et al., 2004). Conversely, there are a number of fMRI and MEG experiments (Arnott et al., 2004a; Brunetti et al., 2005; Warren and Griffiths, 2003b; Zimmer et al., 2006) reporting the posterior STP (initiating the dorsal pathway) to be sensitive to spatial

information of auditory stimuli. The localization of sound source has a crucial role for orienting the visual system (Arnott and Alain, 2011; Stecker et al., 2005). The spatially invariant, or intrinsic, features of an auditory object, such as weight or stiffness of the vibrator derived from the spectral profile of the sound (which renders timbre) are directly related to the identification of the sound source. Given that an adjacent region, i.e., the anterior aspect of the superior temporal sulcus (STS), has been spotlighted for its selectivity to human voice and its sensitivity to a specific person (speaker's identity) (Belin et al., 2004), the processing of such intrinsic features in the anterior STP supports the dual pathway hypothesis. Therefore, the involvement of the ventral auditory pathway in processing pitch chroma is highly likely to be also relevant in the absolute recognition of the pitch chroma in populations with AP.

I-5.2. Research Questions

In the current thesis, I will investigate the following research questions:

(1)   Whether and to what extent does the myeloarchitecture of the cortex have relevance to AP? Theoretically, increased electric insulation of local circuits (i.e., intracortical myelination) would result in (1) increase in information transfer efficiency, (2) decrease in transfer delay, thus increasing synchronization between neuronal groups, (3) reduction of ephaptic crosstalk between adjacent axons, thus increasing specificity of information transfer. In most of primary cortices, dense cortical myelination is typically found due to vast inputs and outputs from/to thalamus. If AP mechanism relies on preservation of fine-grained auditory information, heavily myelinated auditory cortices should be found.

(2)   How do the myeloarchitectonic correlates of AP, if they exist, affect functional connectivity? Considering the relationship between the cortical myelin and neural activity (Grydeland et al., 2015) and the widespread AP-related functional connectivity of the auditory cortex suggested by previous works (Jäncke et al., 2012; Wengenroth et al., 2014), it is expected to see connectional correlates of AP from the region with cortical myelin differences.

(3)   Does the engaging of the ventral auditory pathway correlate with the acuity of AP? Given the dual-pathway framework, pitch chroma information is likely to be processed through the ventral auditory pathway. In particular, given the supposed similarity between recognition of daily objects and pitch chroma in AP listeners, pitch chroma can be conceptualized as invariant property of the auditory object (i.e. any tonal sound). Therefore I expect to find neural correlates of AP in subregions of the auditory cortex that belongs to the ventral auditory pathway.

(4)   Are the altered auditory pathways lateralized in the musicians with AP? From the early studies (Keenan et al., 2001; Luders et al., 2004; Schlaug et al., 1995), an increased leftward asymmetry in the areas of PT due to a smaller right PT has been consistently reported. However, recent studies also reported bilateral alteration of cortices (Bermudez et al., 2009; Dohn et al., 2015). Moreover, it has been suggested that the smaller right PT might be explained by larger HG pushing the anterior boundary of PT towards posterior (Wengenroth et al., 2014). Hemispheric dominance has been found in many studies and suggested that the left AC is involved in language processing with higher temporal resolution, while the right AC is specialized in music-related processing with higher spectral resolution (Poeppel, 2003; Zatorre, 1998; Zatorre and Belin, 2001). Therefore a crucial role of the right hemisphere is expected in our studies.

I-5.3. Summary of Studies

In the following chapters, I will present individual studies that I worked on together with colleagues, in order to find answers to the questions listed above.

In the first study, we used MP2RAGE sequence with a 7-T MRI scanner for in-vivo mapping of intracortical myelination. We found a positive relationship between the acuity of AP and the cortical myelin in the medial region of the right planum polare (PP). We discussed possible roles of the cortical myelin in the non-primary auditory cortex that forms the ventral auditory pathway. This work has been published as a research article (Kim and Knösche, 2016).

In the follow-up study, we used resting-state fMRI data acquired with a 3-T MRI scanner[9]. Seeding from the right PP, which was highly myelinated in individuals with high acuity of AP, we found increase in cross-correlation and cross-coherence in bilateral auditory cortices, as well as frontal regions in the left hemisphere. We argued that such interhemispheric connectivities reflect integration of two components (perceptual and associative) of AP. The manuscript of this work is in the process of review for publication.

---

[9] The current (March, 2016) implementation of the head coil system in the institute (Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Saxony, Germany) suffers from signal loss in the ventral regions of the brain (a gradual decrease of signal starts from the middle temporal gyrus along the ventral direction). Therefore, for whole-brain analysis of connectivity, we used a 3-T MRI scanner for this study.

II. Methods and Materials

In this thesis, I used psychoacoustic behavioral tests as well as neuroimaging, i.e., magnetic resonance imaging (MRI) for in-vivo measurement of the anatomy and functional activation of human brains. In this chapter, I will briefly review the procedure of behavioral testing, the principles of MRI and its neurophysiological relevance, neuroimage processing, and statistical inference.

II-1. Behavioral Testing and Auditory Stimuli

For the current thesis, I employed three different behavioral tests for quantification of the acuity of absolute pitch (AP), pitch discrimination, and musical aptitude, respectively. I have implemented the first two tests based on previous literature, which I will explain in this section. For the last test, I used an already published test called 'Musical Ear Test' (Wallentin et al., 2010). The corresponding author kindly provided the original test material (audio files and answer sheet) upon my request. More information on the test stimuli and the reliability and validity of the test can be found in the original paper (Wallentin et al., 2010).

II-1.1. Absolute Pitch Test

As discussed in Section I-4.1, 'Behavioral definition and characteristics of AP', there have been multiple experimental protocols, stimuli, and response methods in order to quantify the AP perception based on behavior. Proposed and applied experimental protocols in terms of the participant's response include (1) reporting verbally or writing down the name of a given tone (Wilson et al., 2009; Zatorre et al., 1998), (2) pressing a corresponding key on a muted digital piano keyboard (Miyazaki, 1988), (3) clicking on a key on the graphically presented keyboard on a computer screen (Athos et al., 2007), and (4) clicking on a musical score on a computer screen (Bermudez and Zatorre, 2009b). In all cases, auditory feedback of the chosen input and a feedback on whether the answer is correct are not given, because otherwise the participants could use relative pitch based on the feedback.

The method (1) is the traditional way used in music schools to test one's AP listening. It does not require any special equipment or programming and it would be the most familiar setting to musicians with formal education in music. However, it is difficult to find the reaction time (RT) in this method: One can record behavioral sessions with a video camera or a microphone and compute RT manually or with the help of parsing software. But it would be very time-/labor-consuming, thus the RT is usually neglected in this method. The method (2) is still somewhat familiar to musicians and can measure RT (and even the velocity of key strokes) but it requires prior knowledge on how to play the piano. If the participant does not have any experience in

playing the piano (e.g., a non-musician without any musical experience would not even know whether the left side keys are lower or higher than the other side), the results of the AP test would be meaningless even with respect to pitch height. The methods (3) and (4) are based on a graphical user interface (GUI) on a computer or a mobile device such as a smartphone, which has a great advantage in acquiring large scale data online. Also it is possible to measure the RT, although the RT might strongly reflect the participant's familiarity with the GUI. An additional restriction of (3) is that it could be difficult to present full-size keyboard of 88 keys over 7 octaves (while each key is not too small to click/tab on) for low-resolution/small screens, thus participant's judgment on pitch height is discarded. To overcome this aspect, (4) presents a circle of 12-pitch-chroma for pitch chroma response and a grand staff (a combined two-row musical score usually for piano) for pitch height response. This seems to be a good idea, but musical knowledge and familiarity with the piano (as well as method (2) and (3)) would affect the RT and accuracy.

The choice of the testing material and protocols (timbre, pitch range, response method) is not a trivial issue because it could suffer from biases in favor of certain musical instruments (Wong and Wong, 2014). Wong and Wong found that the AP performance score is affected by the similarity between the experiment and one's daily musical experience. The effect was significant in both violinists and pianists with their own familiar context (Wong and Wong, 2014), which implies that the common choice of the piano timbre and piano keyboards is preferential to pianists. Given that the majority of the musicians in most of the experiments are pianists and most musicians know the general mapping between the piano key and the pitch name (e.g., "C" is the first white key in an octave), a setting based on the piano may not be a bad idea. However, for extra caution, in order to calibrate the motor response of non-pianists, or just to measure their baseline, a simple procedure (such as pressing a piano key that is visually presented in either graphical representation of piano keyboard or musical scores) might be useful.

In the current thesis, I used the timbre of pure tones and piano tones with a muted digital piano with 88 keys (CLP-150; Yamaha, Hamamatsu, Shizuoka, Japan) as an input device. The target tones spanned 3 octaves: C3 (130.81 Hz) to B5 (987.77 Hz) for sine waves and E3 (164.81 Hz) to D#6 (1,244.50 Hz) for piano timbre. The pure tones had a length of 1 sec and were prepared with the sine function of MATLAB (version 8.2; Mathworks Inc., Natick, MA, USA) with linear ramps of 10 msec at the beginning and the end. The piano tones were also 1 sec long, created using music creation software called GarageBand (version 10.1.0; Apple Inc., Cupertino, CA, USA) with a virtual musical instrument named "Steinway Grand Piano." Mono-channel audio files were sampled at 44,100 Hz with a precision of 16 bits per sample. Each target tone was presented once, resulting in 72 trials (36 pure + 36 piano tones). Stimuli were binaurally delivered through headphones (HS-800; A4tech, Taipei, Taiwan) at the approximate intensity of 70 dB sound pressure level (SPL) to a participant sitting in front of an 88-key digital piano (CLP-150; Yamaha, Hamamatsu, Shizuoka, Japan), which was linked to the experiment software package Presentation (version 14.9 build 20110719; Neurobehavioral Systems Inc., Berkeley, CA, USA) on the Windows

XP system (version 5.1.2600; Microsoft, Redmond, WA, USA) via musical instrument digital interface (MIDI).

To differentiate inter-individual variability in AP performance, we introduced a limited response time window from the stimulus onset to 4 seconds after the onset. Participants were instructed to respond as quickly as possible while maintaining accuracy within the given time window of 4 seconds. They were warned about time-out after 4 seconds. A short training session was used to familiarize the participants with the experimental procedure.

When comparing pianists with AP to non-pianists with AP, there was no significant difference in the AP score ($T(6) = 1.5$, $p = 0.18$), but there was a significant difference in the RT (mean RT of pianists = 1.26 ± 0.09 s; non-pianists = 2.27 ± 0.81 s; $T(6) = 2.89$, $p = 0.03$). This indicates that narrowing the response window to less than 2 seconds would decrease the AP scores of the non-pianists even if their AP performance in familiar context equals that of pianists.

The AP performance was analyzed in terms of (1) absolute error (AE), (2) absolute octave-error-corrected error (ACE), (3) difference between mean AE and mean ACE, as a measure of octave error, (4) AP score (APS), (5) hit rate, and (6) reaction time (RT) for pure and piano tones. "Octave-error" is defined by an answer with correct pitch chroma but incorrect pitch height, which is known to be frequently observable in AP musicians (Miyazaki, 1988; Miyazaki, 2004a). Octave error correction was achieved by disregarding errors with respect to octaves: for example, if the target is "C4" and the participant's answer is "C3", the error would be 12 semitones for AE, but 0 semitones for ACE. Arithmetically, the pitch chroma in the 12-tone key scale can be noted by a remainder after division of a natural index of a pitch in the unit of semitone (e.g., MIDI codes) by 12 as $c \in \{0, 1, 2, \cdots, 11\}$. An octave-corrected error E between two pitch-chromas c_1 and c_2 ($c_1 > c_2$) is given as:

$$E = \begin{cases} c_1 - c_2 & \text{if } c_1 - c_2 \leq 6 \\ -\{c_2 - (c_1 - 12)\} & \text{if } c_1 - c_2 > 6 \end{cases} \tag{II-1}$$

A positive, scaled score of AP (APS) was computed from ACE as $APS_j = 1 - m_j / m_{max}$ for the j-th timbre, where m is the mean ACE, and m_{max} is the maximum of possible corrected errors, which is 6 semitones. Thus the APS is confined between 0 and 1, positively coding the AP ability with the chance level of 0.5.

II-1.2. Pitch Discrimination Test

A behavioral test to measure one's frequency resolution in pitch perception, or frequency discrimination threshold (FDT) as suggested in (Micheyl et al., 2006), is more or less straightforward:

(1) Present a pair of sine waves (unlike in the AP test, only pure tones are used) with a certain interval sequentially;

(2) The participant is asked to determine which of the tones is higher than the other;

(3) If the answer is:

   a. correct, decrease the interval.

   b. incorrect, increase the interval.

      i. If it is the second trial and the current incorrect answer is followed by a correct answer in the previous trial, define the current trial and a 'reverse' and save the current interval.

(4) If the number of identified reverses is:

   a. bigger than a predetermined number (e.g., 15), terminate the loop and compute the geometric mean of the thresholds excluding the first 8 reverses to discard the familiarization of the participant to the test procedure, which is not relevant to the pitch discrimination ability.

   b. Otherwise go back to (1).

The pair of tones consisted of a reference tone (either 440 or 370 Hz) and a target tone (spaced at variable intervals from the reference during the test). The length was 200 ms and there was a delay between the tones to measure low-level auditory processing without confounding due to AP. All experimental routines were performed with my own code under MATLAB environment (version 8.2; Mathworks Inc., Natick, MA, USA).

The paper (Micheyl et al., 2006) carried out a large number of trials (6000 trials) to detect the effect of musical training in the FDT, which might have taken more than 1.5 hours. Because I carried out other psychoacoustic tests as well (i.e., musical aptitude test called 'Musical Ear Test' (Wallentin et al., 2010) and absolute pitch test), I restricted the number of required reverses to 15, which took less than 5 minutes. The effect of musical training was significant when I compared 19 musicians and 15 non-musicians (T(32) = 3.28, p = 0.0025) indicating that the number of used trials was sufficient.

II-2. Magnetic Resonance Imaging

Magnetic resonance imaging (MRI), which is a non-invasive in-vivo imaging technique, can be separated into (1) nuclear magnetic resonance (NMR) and (2) spatial encoding of the NMR source. In this section, I will shortly discuss the main principles of NMR and spatial encoding, examples of

MR sequences used in the thesis, and possible imaging artifacts with explanation of the source of such artifacts.

II-2.1. Nuclear Magnetic Resonance: Source of Signal

All elementary particles have 'spin', which is an intrinsic property in quantum physics. For instance, the nucleus of hydrogen, comprising a single proton, has a spin of ½. These protons behave like small magnets. Usually spins are orientated in random directions thus its average magnetization is zero. In a strong static magnetic field B0, spins are aligned in parallel or antiparallel to the direction of the external magnetic field (called longitudinal direction) with precession around the direction of B0 due to its spin (Rajagopalan et al., 2005). Because there are slightly more spins aligned in parallel (i.e., 0 degrees) than anti-parallel (i.e., 180 degrees) to the longitudinal direction, the net magnetization of spins is positive along the longitudinal direction.

The spins aligned in parallel in a low energy state can be pushed up to a high energy state if an electromagnetic particle (i.e., photon) hits the spin with an energy that matches to the energy difference between the low and high energy states. This energy difference v is given by Larmor frequency as:

$$v = \gamma \cdot B_0 \qquad (\text{II-2})$$

where γ is gyromagnetic ratio, which is constant for each nucleus and B_0 is strength of the external magnetic field. For proton in hydrogen nucleus of water molecule, which is the main source of MRI signal, the Larmor frequency is 42 GHz/Tesla (Buxton, 2009).

In NMR, radio frequency (RF) pulses at the Larmor frequency are used to lift the spins to the high energy-state. When an RF pulse is applied, the rotating axes of spins are synchronized and deflected from the direction of B0 towards a perpendicular plane to it (Rajagopalan et al., 2005). The angle of this deflection is called 'flip angle', which depends on the combination of the RF pulse strength and duration. The longitudinal component of the net magnetization (Mz, when the direction of B0 is z) is minimized when spins are flipped by 180 degrees and the transverse component of the magnetization (Mx and My, or combined as Mxy) is maximized when spins are flipped by 90 degree.

After the RF pulse, the spins in the high energy state (i.e., flipped spins) relax back to the low energy state (thus the process is called 'relaxation'). During the relaxation, the longitudinal magnetization increases up to the state before excitation whereas the synchronization of phases in the transverse magnetization is desynchronized over time. These processes result in recovery of Mz and decaying of Mxy. The net magnetization can be picked up by transversely orientated coils, which are often also used to generate the RF pulses. Mz and Mxy during the relaxation are

modeled by Bloch (Bloch, 1946). The solution of the Bloch equation is simply given (Buxton, 2009) as:

$$M_z(t) = M_0 - [M_0 - M_z(0)]\exp(-t/T_1)$$
$$M_{xy}(t) = M_{xy}(0)\exp(-t/T_2)$$

(II-3)

where t is time after RF pulse given, Mo is Mz before excitation, T1 is the longitudinal relaxation time, and T2 is the transverse relaxation time. T1 and T2 differ across tissue types (i.e., concentration and embedding of water within microstructures of tissues) and the strength of the Bo field. In MRI, the main source of signal is protons in the water molecules in biological tissues. At 1.5 T, approximate T1 and T2 values of the gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) are given in Table II-1. An example of relaxation process from the flip angel of 90 degrees is illustrated in Figure II-1.

Table II-1. Approximate T1 and T2 values from (Buxton, 2009).

B0 at 1.5 Tesla	GM	WM	CSF
T1 (ms)	950	700	2500
T2 (ms)	95	80	250
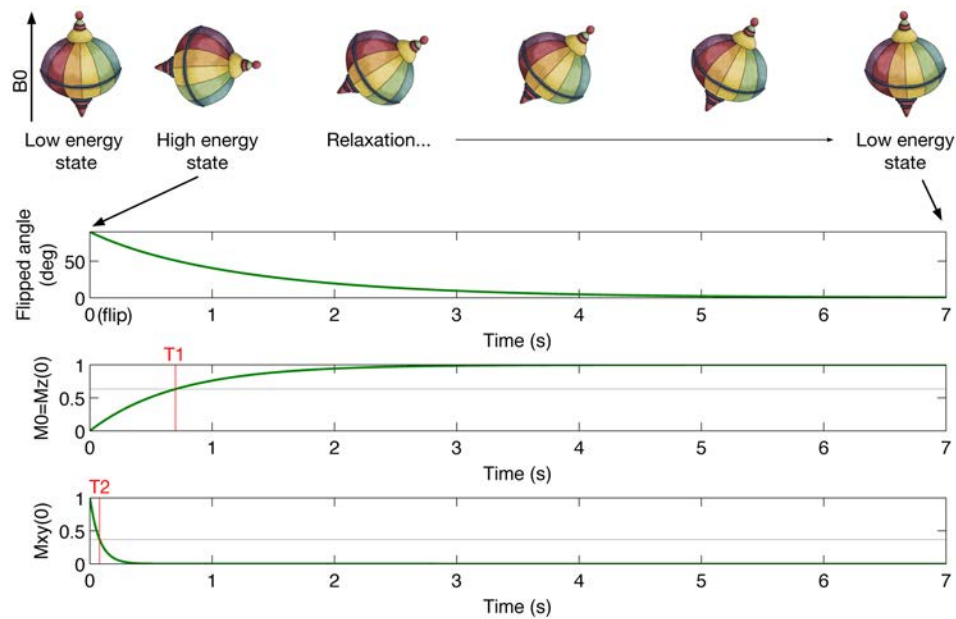
Figure II-1. Illustration of spinning state of proton spin (as a top). Flipped angle (after a 90-degree RF pulse), Mz, and Mxy are plotted during the relaxation of proton spins in white matter in a Bo static field at 1.5-Tesla. T1 (700 ms) and T2 (80 ms) are marked in red. Magnetization level that defines T1 or T2 time is marked as a horizontal line in gray.

The illustration of a top was provided from http://www.educationscotland.gov.uk/ under the Creative Commons license.

The primary motivation of using MRI is to acquire geometrical information of neural tissues in-vivo. For that purpose, it is desirable to have an image where various tissues of brain (i.e., GM, WM, and CSF) are easily distinguishable. This can be achieved by specifying readout time as illustrated in Figure II-2. Although the NMR signal is not magnetization itself, it should be mentioned that the main motivation of designing MR pulse sequences is to maximize contrast between certain pairs of brain tissues, which is based on the different magnetization in different tissues. Much more complicated issues exist for the design of unbiased, efficient, robust, and optimized MRI sequences. Besides the tissue contrast (i.e., relative difference between tissues), accurate estimation of the MR property (e.g., T_1 or T_2 of a voxel) has been recently proposed to be used for 'quantitative' analysis of brain structures (Geyer et al., 2011), which will be shortly reviewed in the Section II-2.3.
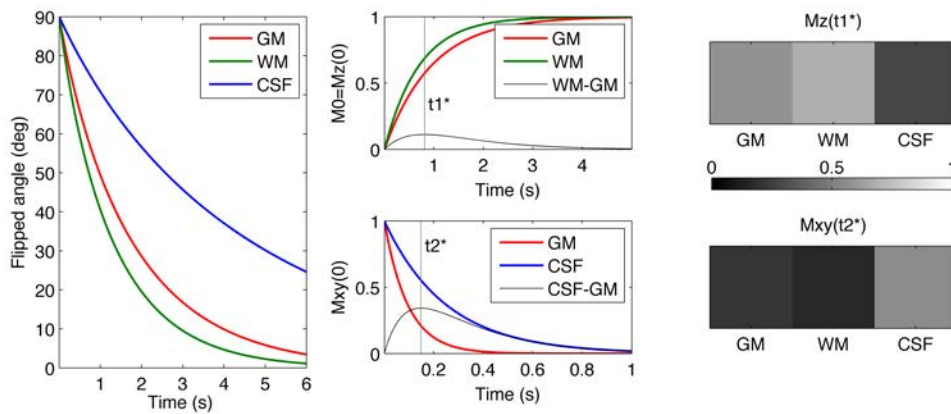
Figure II-2. Tissue contrast change over time due to different Mz recovery (or Mxy decay) trajectories of each tissue. The flipped angle recovers differently in each tissue (left), thus contrasts in magnetization (either Mxy, or Mz) between GM/WM or CSF/GM also differently evolve. The contrasts can be maximized at different time points (i.e., t_1^* and t_2^* marked by gray vertical lines) as shown in three pixels ("Mz(t_1^*)" or "Mxy(t_2^*)").

II-2.2. Spatial Encoding: Localization of the Signal Source

Spatial encoding, i.e., localization of the source of NMR signal, is necessary for 3-D reconstruction of brains from the non-invasive NMR signals. This is achieved by applying various magnetic field gradients, which can be generated by the gradient system in the scanner. Using specific configurations of coils in the cylinder shape, orthogonal gradient of magnetic field can be created as demonstrated in Figure II-3.
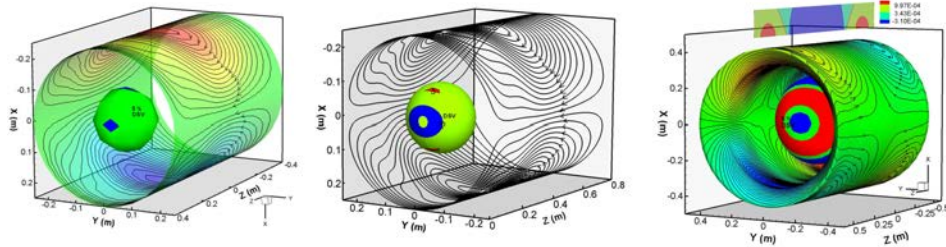
Figure II-3. Gradient system. Various gradients in magnetic field can be generated by different coil configurations and application of currents. The figures was reproduced from Lopez et al. (2009), copyright reserved by IEEE.

The gradient can be applied during excitation, readout, and other timing for an efficient and effective imaging. For an example, an MRI sequence to obtain an axial slice at each readout is a combination of (1) slice selection by Z-gradient during the RF pulse, (2) phase encoding by Y-gradient during the inverse X-gradient, and (3) frequency encoding by X-gradient during readout to specify 3-D coordinate of the signal as in Figure II-4.
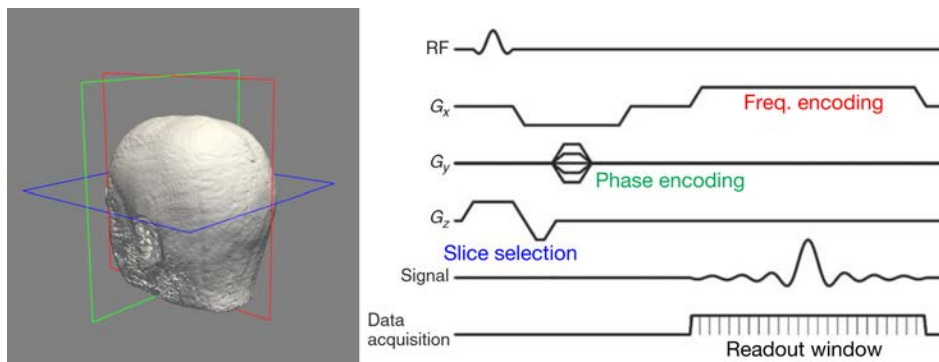
Figure II-4. An example of spatial encoding. For an axial slice (z-coordinate in blue), the x-coordinate is encoded by frequency (red) and the y-coordinate is encoded by phase (green). A head model is shown for orientation (example data of 'Bert' included in FreeSurfer). The MRI sequence scheme was reproduced and modified from Buxton (2009), copyright reserved by Cambridge University Press.

Because the geometrical space is encoded by frequency and phase, the actual acquisition of a 2-D slice is done in the parametric space called k-space instead of the space in the scanner. The k-space is formed by Fourier transformation of a 2-D image based on orthonormal bases (sine and cosine functions along the X-/Y-direction). The coordinate in the k-space corresponds to a specific combination of basis functions and its intensity is the coefficient of that combination of bases. An example of an axial slice in a X-Y space (Cartesian space) and corresponding coefficient image in the frequency-phase space (k-space) is given in Figure II-5.
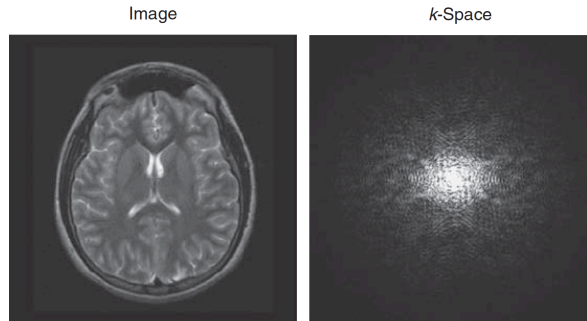
Figure II-5. A T2-weighted image in the Cartesian space (left) and the k-space (right). The figure was reproduced from Buxton (2009), copyright reserved by Cambridge University Press.

The correspondence can be more clearly shown from the relationship between value ranges in the k-space and spatial frequency in Figure II-6.
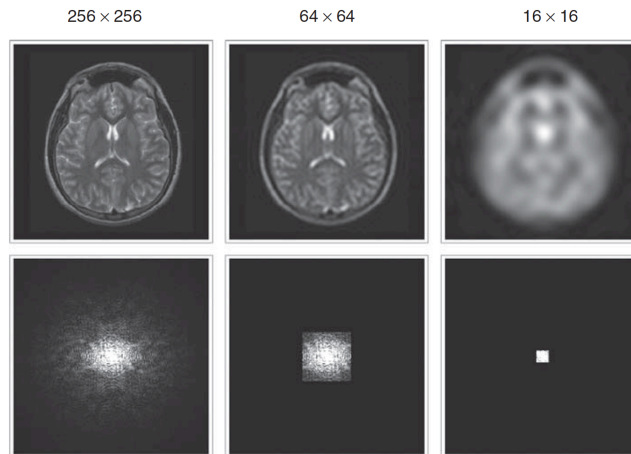
Figure II-6. A T2-weighted images in the Cartesian space (upper row) and the coefficient images in the k-space (lower row) for various spatial bandwidths in Cartesian space (upper) and coefficient value ranges in the k-space (lower). The figure was reproduced from Buxton (2009), copyright reserved by Cambridge University Press.

As another demonstration, the effect of coefficient values in the k-space to the image in the Cartesian space is shown in Figure II-7. The marked k-space pixels resulted in grating patterns in the Cartesian images, which are in fact combinations of two sine functions along the x-/y-directions for the coordinates of the pixels in the k-space.
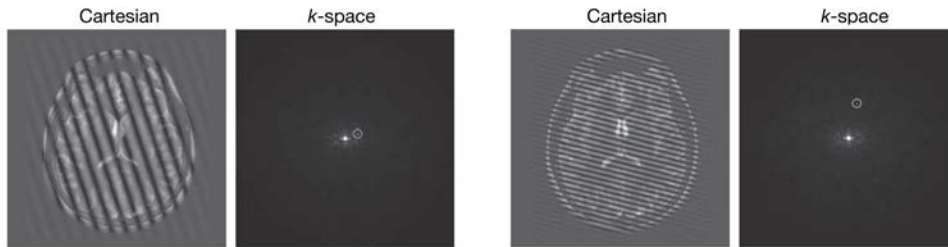
Figure II-7. Examples for the effect of artifacts in the k-space to the image in Cartesian image. The figure was reproduced from Buxton (2009), copyright reserved by Cambridge University Press.

II-2.3. MR sequences in Use

For the current thesis, we used quantitative T1 mapping (qT1), T1-weighted imaging (T1w), and T2*-weighted imaging (T2*w) with two MR systems at 3-T and 7-T, respectively. Using a 3-T whole-body MR system Magnetom Prisma (Siemens, Erlangen, Germany) with a 32-chanel head coil system, a *magnetization-prepared pulses and rapid gradient-echo* (MPRAGE) (Mugler and Brookeman, 1990) protocol based on the ADNI sequence[10] was used for T1w images, whereas echo-planar imaging (Stehling et al., 1991) was used for T2*w images to measure BOLD signals. Using a 7-T whole-body MR system (Siemens, Erlangen, Germany) with an 8-channel head coil system (RAPID MR International, Ohio, USA), *magnetization-prepared two rapid gradient echo* (MP2RAGE) (Marques et al., 2010) imaging was used for qT1 images. The parameters of the MR sequences used in this thesis are listed in Table II-2. Echo and inversion times are used for the spin echo pulse sequence and the inversion recovery pulse sequence for optimization of the imaging sequence. Examples of images are shown in Figure II-8.

One main source of the contrast in T1w is known to be myelin content (Eickhoff et al., 2005), which also strongly affects T1 values in the qT1 maps (Marques and Gruetter, 2013). The T1w contrast or qT1 values are not only sensitive to myelin but also iron (Stuber et al., 2014). However, myelin is the most influential component, thus supporting the usage of qT1 mapping as an in-vivo myelin mapping, especially for the axons within the GM (Geyer et al., 2011).

On the other hand, one main source of the contrast in T2*w is the blood-oxygen-level-dependent (BOLD) effect: i.e., the different magnetic susceptibility of oxygenated and deoxygenated hemoglobin. The BOLD effect is related to the cerebral blood flow/volume and the blood oxygenation level that are affected by neurovascular coupling phenomenon, which was observed from optogenetic neurons in rodent brains (Lee et al., 2010). Taken together, the

---

[10] http://adni.loni.usc.edu/methods/documents/mri-protocols/

relationship between the BOLD signal and neural activities are generally accepted in the field of neuroscience.

Table II-2. MR sequence parameters used in the current thesis.

	Tesla	FA (deg)	AF	TA1 (ms)	TR (ms)	TE (ms)	TI (ms)	Matrix ($\phi \times f$ *)	Number of slices	Voxel (mm)[#]
MPRAGE	3	9	4	300,000	2,300	2.98	900	240 × 256	176	1.00
EPI	3	60	1	1400	87.5	30	-	88 × 88	64 (4)[+]	2.30
MP2RAGE	7	5/3	3	450,000	5,000	2.45	900/2,750[@]	320 × 320	240	0.70

* The number of phase and frequency encoding; [#]Isotropic voxel resolution. [@]First and second inversions for MP2RAGE. Abbreviations: FA, flip angle; AF, acceleration factor; TA1, time of acquisition of one volume; TR, time of repetition; TE, time of echo; TI, time of inversion; deg, degree; ms, millisecond; mm, millimeter.
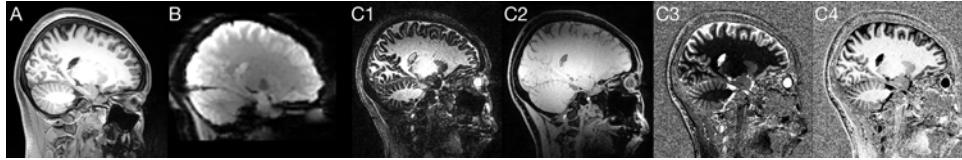
Figure II-8. Examples of MR images used in the present thesis. (A) T1w from MPRAGE sequence and (B) T2*w from EPI sequence at 3-T, (C1) the first inversion, (C2) the second inversion, (C3) quantitative T1 mapping, and (C4) isotropic tissue contrast images from MP2RAGE sequence at 7-T.

II-2.4. Artifacts in MRI

II-2.4.1. Gibbs Ringing and N/2 Ghost

Some artifacts in MR images are related to the spatial encoding scheme, which uses spatial Fourier series to describe a 2-D image. For instance, Gibbs ringing artifacts (i.e., wiggling intensity) are caused by overshooting sums of finite Fourier series at the high contrast boundary, which can be perfectly expressed by infinite Fourier series. This artifact appears by an insufficient number of encoding steps in frequency and/or phase encoding directions as can be seen in Figure II-9 (A).

Another artifact is called 'N/2 ghost' as shown in Figure II-9 (B), where N is the number of pixels along the axis. It is related to the alternating direction of phase sampling. The 'ghost' appears as shift by N/2 pixels along the phase-encoding axis due to the disruption of phase encoding.
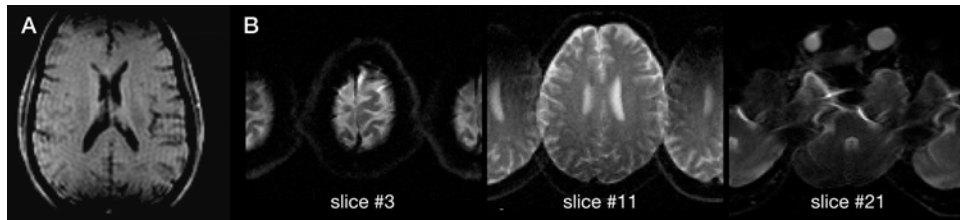
Figure II-9. Examples of (A) Gibbs artifact and (B) the N/2 ghost image. The image A was reproduced from Buxton (2009), copyright reserved by Cambridge University Press. The image B was reproduced from Lee et al. (2002), copyright reserved by John Wiley & Sons.

II-2.4.2. Spatial Distortion (Warping) due to Inhomogeneous Susceptibility

Spatial encoding assumes the static field is homogeneous and any difference in MRI signal (spatial encoding with phase and frequency) is due to the location of the source. However, in reality, the static field is heterogeneous, especially in the head. This leads to wrong inference on the location of sources resulting in non-linearly deformed images. What creates such local distortions of the magnetic field is called the 'magnetic susceptibility effect'. 'Magnetic susceptibility' is the degree to which a material becomes magnetized when placed in a magnetic field. Different head tissues have different susceptibility. However, the distortion mainly comes from the head's non-spherical geometry. Even if the head would consist of only one material, local distortions would arise when the additional magnetic field from the magnetized object (the head) is not fully cancelled, which is only the case for a perfect sphere (Buxton, 2009).

The distortion is more severe at boundaries between different materials (e.g., water and air in frontal sinus), and also for certain MRI sequences (e.g., EPI). In order to reduce distortions arising from the susceptibility effect, there are additional coils called 'shim coils', which perform a 'shimming' or flattening of the non-uniform static field. The susceptibility effect is actually the main source of contrast in BOLD imaging: Oxygenated hemoglobin is diamagnetic (zero or negative susceptibility) whereas deoxygenated hemoglobin is paramagnetic (positive susceptibility).

An example of a whole-brain image measured at 7-Tesla (1-mm-isovoxel) is given in Figure II-10. The alignment between the EPI image and T1w image (red contour) is generally acceptable, but there are misalignments, especially around the frontal pole and orbitofrontal cortex (marked by arrows) in Figure II-10 (A). These distortions can be corrected using the B0-field map that is related to local variation in susceptibility (which takes less than 2 minutes for a low resolution image). After correction ('unwarping'), matching between the EPI and T1w images was considerably improved (green and cyan arrows in Figure II-10 (B)) but certain regions remained imperfectly matched (e.g., medial orbitofrontal cortex marked by blue arrow).
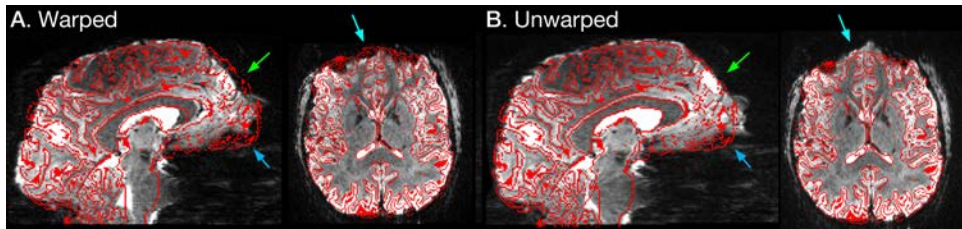
Figure II-10. Example of distortions in an EPI measured at 7 Tesla at 1-mm-isovoxel resolution. Sagittal and axial slices of (A) spatially distorted ('warped') EPI and its (B) corrected ('unwarped') image. The contour of the anatomical image is overlaid in red.

II-3. Image processing

II-3.1. Overview of Neuroimage Processing

There exist a number of different approaches to access neural structures and activities from the acquired MR images. I will start with the morphological studies, where usually one subject has one image and we are interested in group differences or any group-wise correlation with other (behavioral or clinical) measures. Based on the element (i.e., a measurement to compare between groups or to correlate with another measure), the between-subject analyses can be divided into (1) region-of-interest (ROI)-based analysis, (2) voxel-based analysis (VBA) on structural measures, and (3) VBA on functional measures and the dimensions of the measurement space.

II-3.1.1. Between-subject Analysis
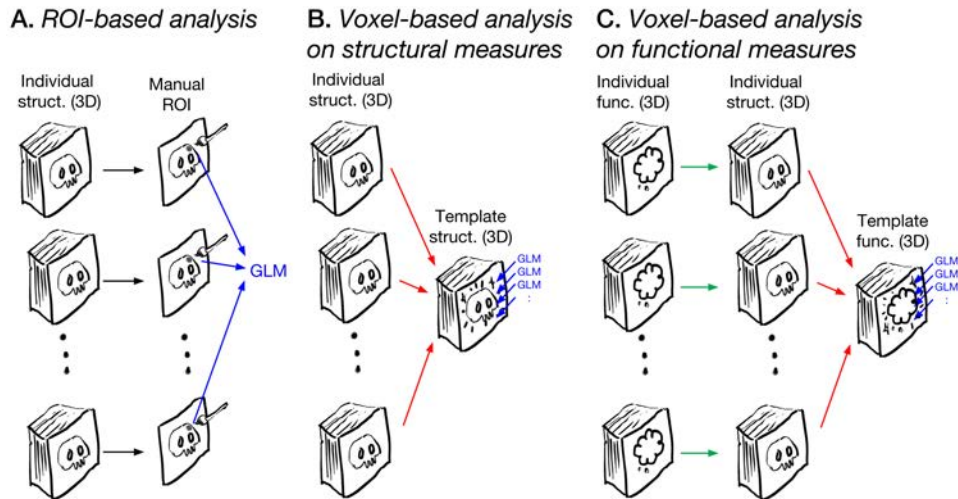
Figure II-11. Overview of neuroimage processing. (A) Manual region-of-interest (ROI)-based analysis, (B) voxel-based analysis (VBA) on local morphometry, (C) VBA on local functional measures. Process flows are marked by arrows: manual segmentation in black, registration in red, intrasubject (inter-imaging-modality) registration in green, and statistical inference using general linear model (GLM) in blue. Abbreviations: struct., structural images; func., functional images.

An overview of the between-subject neuroimage analyses is given in Figure II-11. The analysis pipelines are as follows:

(A)   *ROI-based analysis* is the most intuitive way to compare gross morphology of a certain anatomical structure that is well identifiable from the 3-D MRI such as T1w or T2w images. Commonly, manual work based on macroscopic guides is the main part of the processing, then statistical inference (such as GLM) is done for gross measures such as the volume of the structure.

(B)   *Voxel-based analysis of morphological measures* is to perform GLM on each spatial element (i.e., a voxel in 3-D space) after registration of individual images into a stereotaxic space. Thus construction of intersubject correspondence is an important issue of this approach, which usually relies on the non-linear registration. That is, the quality of the registration (or deformation between individual images and a template) should be state-of-art, in order not to violate the assumption of the correspondence. Proposed local morphometry metrics include GM tissue probability (called voxel-based morphometry or VBM (Ashburner and Friston, 2000)), deformation vector as a multivariate measure (called deformation-based morphometry, or DBM (Ashburner et

al., 1998; Chung et al., 2001a), and Jacobian determinant of the inverse (template to individuals) deformation field (called tensor-based morphometry or TBM because the Jacobian matrix of the deformation field is a tensor (Chung et al., 2001a; Hua et al., 2008)).

(C)   *Voxel-based analysis of functional measures* is very similar to pipeline (B), except that the GLM is performed on 'functional' measures that are registered into the individual structural images. This type of analysis is used to compare images other than the individual structural images, which is often a contrast image from a within-subject analysis on functional data but also can be any other scalar[11] images (e.g., an intracortical myelination mapping image, diffusion-weighted image). Because these images can be in different spaces (i.e., images were obtained from different MRI sessions, or a participant changed the position of the head between images), the quality of intra-subject registration is crucial. Once the 'functional' image is registered onto the structural image, non-linear registration from the individual structural image to the template structural image is applied so that the functional images can be compared in the template space. In order to reduce numerical errors and image degradation from multiple resampling, two consecutive registrations (functional to structural, and individual to template) can be combined and the image resampled only once. After registration, the rest of procedures are identical to pipeline (B).

II-3.1.2. Within-subject Analysis

As briefly mentioned above, there are studies that require within-subject analysis. Normally, functional experiments are design to investigate an effect of certain psychological events in physiological measures (such as neural activities) in general human beings. For that, there are two (at least) levels of statistical inferences (unless mixed-effect model is used). The first level is on whether a condition X evokes a neural activity Y in this particular individual A. This is done by carrying out image processing and statistical inference on the multiple images from the participant A, who is only one sample of humanity. Only with the data from A, we cannot tell whether the result from A is universal or only specific to A. Thus, we collect more data from participant B, participant C, and so on. Then, we test the effect of the condition X (fixed effect) controlling

---

[11] Theoretically, higher dimensional data such as displacement vector field and water diffusion tensor field, or even cross-correlation field can be also compared using multivariate GLM using random field theory (Worsley KJ, Taylor JE, Tomaiuolo F, Lerch J. (2004): Unified univariate and multivariate random field theory. Neuroimage 23 Suppl 1:S189-95.), but appropriate implementation of multiple comparison correction for high dimensional data is currently limited up to 3-D vector field (e.g., SurfStat). One can simply use permutation test for data in any arbitrary dimensions, although the computational cost will be even greater than conventional data.

individual variance (random effect). This can be properly done using mixed-effect GLM for low-dimensional, small-scale data. However, the mixed model for functional neuroimaging data often requires huge computational resources. For instance, for 10 subjects, 6 runs of 200 volumes result in a large design matrix (in the dimension of 12,000 by the number of regressors for fixed and mixed effects) to fit a mixed effect GLM for one out of hundreds thousands voxels. Thus mixed-effect model is not tractable in many cases. Therefore, the most common approach is to first perform image processing and within-subject (intrasubject) statistical inference at the first level. A summary value for the effect of condition X is computed (commonly a contrast image, which is a linear combination of estimated effect sizes), which is then used in the between-subject (intersubject) inference at the second level analysis.

From the perspective of dimensions, the first level analysis deals with data in space-by-time dimensions whereas the second level analysis solves problems in space-by-subject dimensions. Thus, the general pipelines of within-subject analyses are similar to the between-subject analysis as illustrated in Figure II-11. Theoretically there is no reason to spatial normalization at the first-level analysis, but in practice, due to numerical errors via optimal transformation estimation and resampling, the first and second level analyses are performed following spatial normalization of functional data.

II-3.2. Temporal Processing

II-3.2.1. Preprocess: Slice Timing Correction, Realignment, and Temporal Filtering

Four-dimensional functional MRI (fMRI) data is a series of 3-D EPI images. Commonly, about every 2 sec an EPI at around 3-mm-isovxel resolution is acquired. The 3-D EPI image is also a collection of 2-D images: either one or several slices at every 80 msec. Because of this, the last slice could be taken 2 sec later than the first slice. Although BOLD is a very slow process with an average time lag of about 6 sec, the difference of 2 sec could result in different brain states in a single volume. This is not a trivial issue but does affects fMRI experiment results (Sladky et al., 2011). To address this, one can temporally resample timeseries into referential temporal lattice (i.e., 1-D temporal interpolation) as in Statistic Parametric Mapping (SPM[12]) (Ashburner, 2012). On the other way around, one can modify regressors for each slice as in FMRIB Software Library (FSL[13]) (Jenkinson et al., 2012).

Also, it would be ideal if a subject does not move during imaging (like ex-vivo samples) for the quality of MRI. However, to investigate living human brains, there should be always certain degrees of head motion in the data. Correction of this artifact could begin with rigid-body

---

[12] http://www.fil.ion.ucl.ac.uk/spm/

[13] http://fsl.fmrib.ox.ac.uk/

registration of each time point to a reference image (either the first time point or average of all time point), which is commonly called 'realignment' (i.e., 3-D spatial interpolation).

As one can notice, slice timing correction and realignment assumes the other one is relatively fine. On one hand, in slice timing correction, although the time point may not be correct but spatial agreement between adjacent slices is assumed to be fine. On the other hand, in realignment, timepoints are sufficiently good to perform rigid-body transformation of the whole volume (all slices together). This dilemma was addressed by 4-D registration algorithm that solves the two problems simultaneously (Roche, 2011). Although it is an interesting approach, it has not become popular yet, probably because of low spatial (3-mm-isovoxel) and low temporal (2-sec) resolution of fMRI data.

Other than abrupt changes in position (e.g., head motion), very slow trends in the fMRI data often appear. It had been often attributed to physiological noise, the drift was observed in EPIs from a normal subject, a cadaver, and a non-homogeneous phantom, suggesting the slow drift may be due to slight changes in the local magnetic field affecting image intensity (Smith et al., 1999). To discard influence from such slow trends, polynomial functions (usually from the 0-th to 3-rd order) can be used in the within-subject level fMRI analysis. A high-pass filtering at around > 8 mHz (= 1/128 Hz) is also common choice.

II-3.2.2. Head Motion Artifacts Correction

Head motion of a subject during scanning violates one principal assumption of the NMR signal measurement and spatial encoding/decoding, namely that the sample is stationary. When this assumption is violated, the degree of excitation, and spatial encoding in frequency and phase are disrupted. For EPI images (with longer TR), the violation in NMR is more severe (than T1w images with shorter TR) even if the localization is more or less normal. That is, head motion could result in not only rigid-body motion of the brain but also in abrupt intensity changes in many voxels (Power et al., 2012). This is a very important point, because it tells that rigid-body transformation of 3D images is never enough to correct for the artifacts due to head motion. More critically, many of the commonly used functional connectivity measures such as cross-correlation or cross-coherence are vulnerable to such synchronized abrupt changes even by a slight head movement (Power et al., 2015).

A common practice to deal with the motion artifacts is to regress out non-neuronal fluctuation in the BOLD timeseries from anatomically defined white matter (WM) and cerebrospinal fluid (CSF) voxels using principal component analysis (PCA). This approach was introduced as 'anatomical CompCor' (Behzadi et al., 2007). As an example, Figure II-12 shows estimated piecewise head motion (the 1st row) and uncorrected BOLD timeseries (the 2nd row). There are clear motion artifacts between 250 and 300 TRs, and also between 350 and 400 TRs (the 2nd row). Detrending and rigid-motion parameters could not remove those effects (Figure II-12, the 3rd

row). In contrast, CompCor regressors effectively removed those artifacts (Figure II-12, the 4th row). Additional regression of average timeseries from all GM voxels (namely 'global signal') did not result in apparent changes in this example (Figure II-12, the 5th row). Scrubbing regressors (for volumes with the Z-score greater than 3 or the motion parameter greater than 0.5 mm or 0.5 deg). In Figure II-13, the distribution of the correlation coefficients is shown that were sampled from a subset of GM voxels over the brain regions. The statistics of the distribution shows that adding CompCor regressors reduces skewness as well as kurtosis. Spatial dependency of correlation was proposed to identify spurious correlation in rs-fMRI (Power et al., 2015), but no obvious relationship between the change in correlation and the Euclidian distance between voxels is found as shown in the second column of Figure II-13. Finally, goodness-of-fit (GOF) values of the Kolmogorov–Smirnov (K-S) test and quantile-quantile (QQ) plots show that CompCor highly increased Gaussianity (from 10.3% to 23.6%), and the global signal (42.9 %) and scrubbing regressors (78.1 %) increased Gaussianity in this particular data. Figure II-14 shows how the geometrical pattern of correlation changes by adding regressors. There are widely spread correlations that are positive in the timeseries before regression (Figure II-14, top row), which could not be removed by rigid-motion parameters (Figure II-14, middle row), but could be successfully corrected by the CompCor regressors, thus changing the anatomical patterns of correlation accordingly. The global signal regression did not seem to be a very good idea in this particular example. Nonetheless, it has been proposed to reduce falsely heightened correlation levels (Power et al., 2013; Power et al., 2015; Smith et al., 2013; Sylvester et al., 2013). Although others have argued that the anticorrelation in resting-state fMRI connectivity is a simple artifact due to the 'invasive' correction (e.g., global signal or scrubbing) using overall GM timeseries (Saad et al., 2012), the existence of anticorrelation has been also reported without the global signal regression (Chai et al., 2012; Fox et al., 2009).
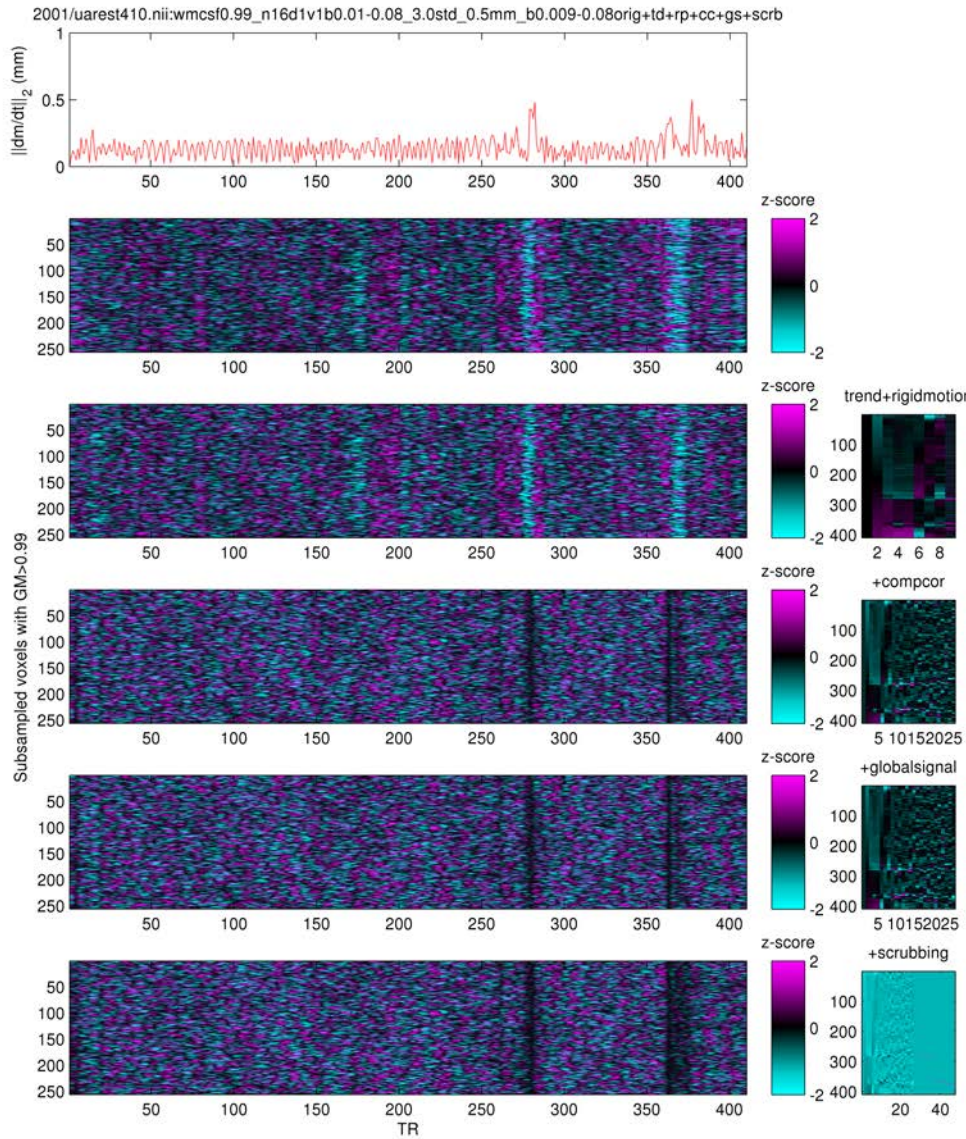
Figure II-12. Head motion and denoised BOLD timeseries with different regressors. The piecewise head motion (top, red) is calculated from rigid-motion parameters, i.e., L2-norm of the temporal deviation of 6 parameters. The BOLD timeseries of subsampled GM (tissue probability > 0.99) voxels, which are spread over all volume) are shown below the head motion. The timeseries values are standardized (z-score) in order to see global fluctuation. The first row shows unprocessed timeseries and the following rows show the residuals after additive inclusions of different denoising regressors: 2 polynomials and 7 motion parameters + 16 CompCor regressors + global mean of GM signal + 21 scrubbing regressors. The rightmost column shows the regressors in total.
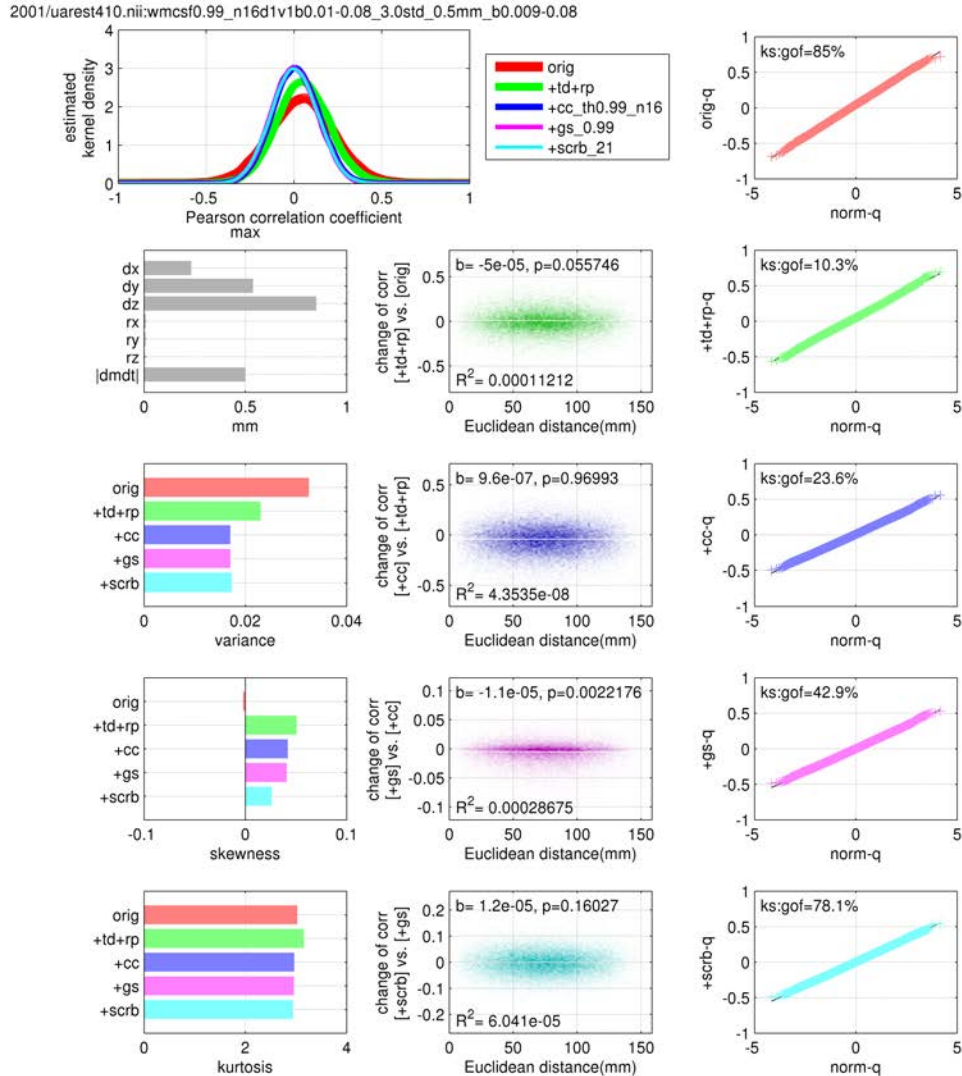
Figure II-13. Distribution of correlation coefficients for different denoising regressors. The distributions are overlaid in the upper-left panel. Colors code different sets of regressors (red, before regression; green, 2 polynomials and 7 motion parameters; blue, + 16 CompCor regressors; magenta, + global mean of GM signal, cyan, + 21 scrubbing regressions). Below that, from the leftmost column, maximal values of head motion, variance, skewness and kurtosis of the correlation coefficients are plotted. In the middle column, changes of coefficients (by adding more regressors) are plotted over Euclidian distance between the GM voxels. Finally in the right column, quantile-quantile (QQ) plot with goodness-of-fit values from Kolmogorov–Smirnov test.
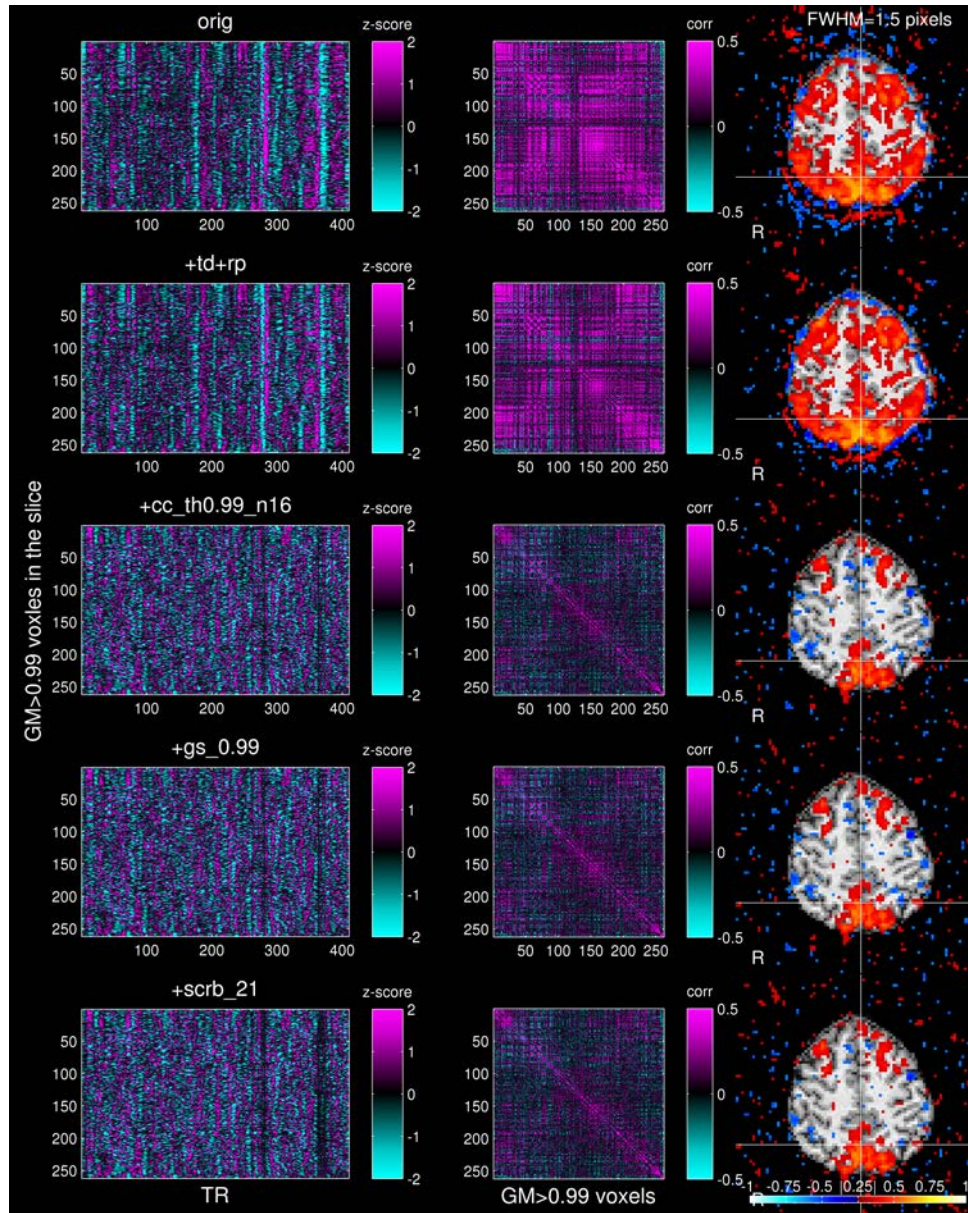
Figure II-14. Correlation matrices and maps for different denoising regressors. The standardized timeseries are shown in the left column, correlation matrices in the middle column, and correlation maps overlaid on axial slices in the right column. Here the timeseries were sampled from the axial slice, which is determined by the location of the precuneous, in order to see the default-mode network.

II-3.3. Spatial processing

II-3.3.1. Surface-based analysis of neuroimaging data

Geometrically, the surface of the cerebral cortex can be approximated by a convoluted 2-D sheet (or multiple layers of sheets) embedded in 3-D space (Chung et al., 2001b) and it has been proposed, and supported by several pieces of evidence, that the functional organization of the cerebral cortex of human and non-human primates is topological in principle (van Essen and Maunsell, 1980; van Essen and Zeki, 1978). One example is the central sulcus. From the crown of the precentral gyrus, the Euclidian distance to the crown of the postcentral gyrus is shorter than the Euclidian distance to the sulcal wall of the precentral gyrus even though the similarity of known functions are the opposite to what the Euclidian distance suggests. Although it is true that there is no simple correlation between functional similarity and geodesic distance everywhere in the cortex either, geodesic distance is still a far better predictor of functional similarity than Euclidian distance. Before the advent of MRI scanning and computational morphometry, neuroscientists desired a method to flatten the cerebral cortex onto a 2-D plane to investigate functional architecture of the cortex (van Essen and Maunsell, 1980), which later motivated the development of computational methods for 'virtual flattening' (Van Essen et al., 2001).

In this thesis, intracortical myelination and functional connectivity over the auditory cortex were carefully examined. Intracortical myelination of the local connections could be parallel to the cortex boundary. Functional connectivity of auditory cortex is also likely to follow the topological organization. Therefore, surface-based analysis could be more beneficial than volume-based analysis for this thesis.

Many algorithms and implementations have been developed to properly extract cortical surfaces from T1w images; for example, FreeSurfer (Fischl, 2012b), Brain voyager (Goebel, 2012), BrainSuite (Shattuck and Leahy, 2002), CRUISE (Han et al., 2004), and so on. The common goal is to create triangular meshes that model the interface between the gray matter (GM) and the white matter (WM) ('inner surface') and between the GM and the cerebrospinal fluid (CSF) ('outer surface') with a same topology as a sphere without 'topological deficits' (i.e., no holes, no self-intersection). The order of indexing of the vertices for all facets should be also 'correct' (i.e., counter-clockwise rotation when its seen from outside). To achieve this, some approaches starts with a sphere model with sound topology and facet orientation (MacDonald et al., 2000) or a crude isosurface (e.g., the face of a voxel is modeled by two triangles) of white matter (Dale et al., 1999). Then the surfaces are deformed to fit the GM/WM interface in the 3-D image as much as possible without disrupting topology or fixing the deficits afterwards.

Such cortical surface models offer a number of benefits for subsequent analysis: (1) one-to-one correspondence between the vertices of cortical surfaces at different cortical depths, (2) smoothing along the surface using mathematical/physical models such as the heat diffusion kernel (Chung et

al., 2005), (3) intersubject registration aligning sulcal patterns via spherical mapping of the surfaces (Fischl et al., 1999), (4) visualization as an inflated surface (Fischl et al., 1999), or a cut and flattened cortical patch (Van Essen et al., 2001). In Figure II-15, a typical example of a reconstructed cortical surface of the right hemisphere is given.
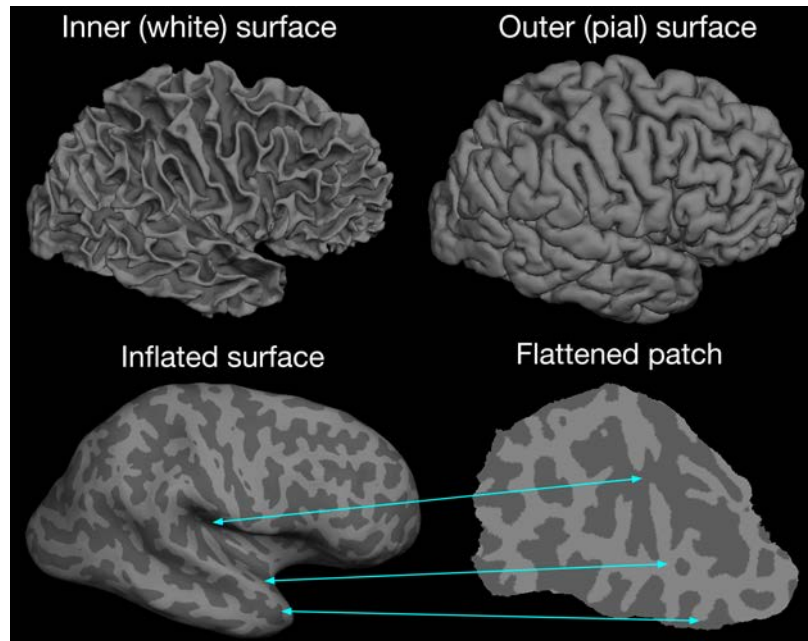
Figure II-15. Reconstructed cortical surfaces generated from a T1-weighted image of a healthy subject. Local curvature (K) on the surface is in a binary code (K<0, convex in lighter gray; K>0, concave in darker gray). For the flattened patch of the right supratemporal plane, insula, a part of parietal lobe is shown. The landmarks are marked for correspondence between surfaces.

II-3.3.2. Parametrization of a Convoluted Structure

One of the advantages of surface-based analysis is that mathematical parameterization can be used, such as the spherical harmonic descriptions (SPHARM) (Gerig et al., 2001; Styner et al., 2006), or intrinsic coordinate systems (i.e. specific to a certain structure as opposed to the Cartesian coordinate system) using Laplace-Beltrami (LB) operator eigenfunctions (Qiu et al., 2008; Reuter et al., 2009). This technique can be used for convoluted cortical surfaces, such as the entire cortex (Chung et al., 2007), a part of cortex such as planum temporale (Qiu et al., 2008), or subcortical structures such as amygdala (Chung et al., 2010) or hippocampus (Kim et al., 2012), in order to analyze the morphology of the structures.

Besides the shape of the structure, a certain biological measure (e.g., BOLD signal or cortical myelination) mapped on that brain structure can be also parameterized. In our own work (Kim et al., 2014), for instance, 1-D parameterization was done using the second LB eigenfunctions to Heschl's gyrus (HG) in order to investigate cortical myelination along the cortical columns. As in Figure II-16 (A), we estimated cortical columns (or 'profiles') across the cortical layers using an image processing software specifically designed for high-resolution image (i.e., 7-T) called CBS Tools (Bazin et al., 2014; Waehnert et al., 2014a). Manually segmented HGs were parameterized using the second LB eigenfunctions as in Figure II-16 (B). Based on the 1-D parameterization along the posteromedial-to-anteriolateral direction, 2-D myelin maps (the dimensions of the 2nd LB eigenfunction by the cortical depth) were created as shown in Figure II-16 (C), which enables group-wise statistical inference of the myelin of a convoluted cortical structure.
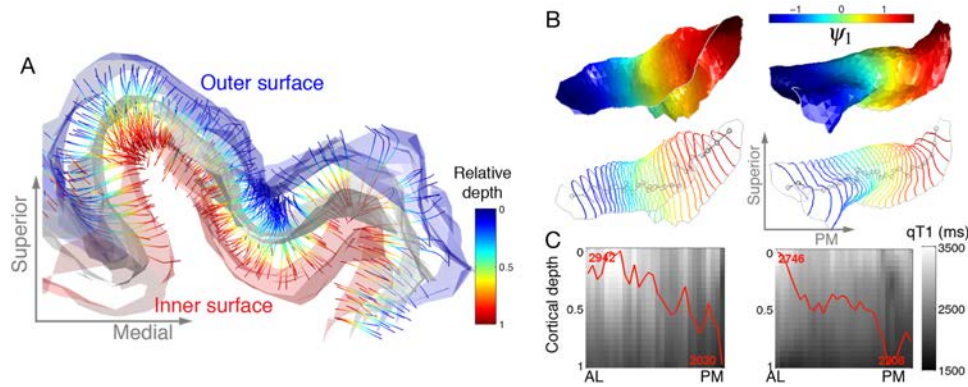
Figure II-16. Two-dimensional parameterization of the cortical myelin of Heschl's gyrus using the second Laplace-Beltrami eigenfunction and realistic cortical layering. The Figures are reproduced from (Kim et al., 2014), copyright reserved by IEEE.

II-3.3.3. Skull-Stripping of MP2RAGE Image

In ultra-high (> 7T) static magnetic fields, T1w imaging is highly affected by inhomogeneities in the transmit field. To minimize such biases, the T1w image in the MP2RAGE sequence is created by combining two inversion images (Marques et al., 2010). This results in uniform tissue contrast over the whole brain, but it also amplifies noise outside the head and in areas adjacent to the brain, which somewhat hampers the use of common image processing tools that are optimized for conventional T1w images at lower field strength (< 3T) MR images.

Although there have been previous attempts to do the task in various ways (Bazin et al., 2014; Fujimoto et al., 2014; O'Brien et al., 2014), unfortunately the proposed methods did not yield satisfactory outcomes for the current data. Thus, we performed brain extraction in two steps: (1) Brain Extraction Tool (BET) in FSL was applied to the second inversion image, which has a similar contrast as proton density images, with a liberal fractional intensity threshold of 0.1, in order to

create a rough brain mask that discards background noise outside the head. However, the noise in the dura mater still remains after BET. (2) We fed the brain mask to the Voxel-Based Morphometry Toolbox 8 (VBM8) on SPM to obtain tissue segmentations of gray matter, white matter, and cerebrospinal fluid. Finally, the skull-stripped T1w image was created by masking the unbiased T1w image with the union of the three segmentations from VBM8. Extracted brain images were visually inspected and manually corrected if necessary.

II-3.3.4. Correlation between Local Curvature and Cortical Myelin in MP2RAGE Images

Spurious correlations between cortical myelin density, cortical thickness, and local curvature were reported from histological samples (Annese et al., 2004) and a large-scale in-vivo MRI dataset (Shafee et al., 2015). Those relationships are known to be negative (greater myelination in a thinner or more convex region of the cortex) (Shafee et al., 2015). To construct a grid system free of such a bias, Waehnert et al. (2014b) implemented a novel layering method that preserves local volume so that the modeled layers mimic the physical bending process of cortical sheets. Although such layer modeling is highly attractive and worked better in ex-vivo images at 0.14-mm isotropic resolution, as shown in Waehnert et al. (2014b), the result of the isovolumetric model was very similar to that of other methods, such as an isopotential model using the Laplacian or an equidistant model, suggesting that the benefit of the novel layering method would be limited in our dataset at 0.7-mm isotropic resolution. Thus, we used an equidistant model (i.e., equal thickness of 'layers' at each point) by fixed proportion of cortical thickness.

In a real dataset that will be presented in Chapter III, negative correlations of qR1 values with curvature and thickness, respectively, were also found as previously reported in the literature (Dick et al., 2012; Lutti et al., 2014; Sereno et al., 2013). However, the effect sizes were very small and negligible (correlation between qR1 and curvature was -0.0058 ± 0.0026; correlation between qR1 and thickness was -0.0757 ± 0.0083). The significance level was very high ($p < 10^{-16}$), but this was mainly due to the very large number of measures (more than 140,000 vertices for each hemisphere). This also explains why the 'de-curved/de-thickened' qR1 values were very similar to the demeaned qR1 values in those studies (Dick et al., 2012; Sereno et al., 2013). Consequently, I did not use the geometrical covariates in statistical assessment presented in this thesis.

II-4. Statistical inference

II-4.1. General Linear Model

Conventionally, in the field of neuroimaging, a linear regression with multiple variables, or a general linear model (GLM), has been widely used in statistical inferences based on the Fisher-Neyman null-hypothesis-testing (NHT) framework (Fisher et al., 1960; Neyman and Pearson, 1992). In the NHT, the main goal is to compute the probability of a false positive under the null

hypothesis that states that the effect (or contrast) of interest equals zero. In other words, NHT finds the probability of observing a value that is equal to or greater than the actual measurement when there is no effect but only noise (called 'p-value') as $p = \Pr(y \geq y_{obs} \mid H_0)$ where y is the variable of interest, y_{obs} is the actual observation, and the null hypothesis is $H_0 : \mathrm{E}(y) = 0$.

Formally, a GLM can be written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \qquad\qquad (\text{II-4})$$

where $\mathbf{y}$ is a column vector of measurements, $\mathbf{X}$ is a design matrix, $\mathbf{b}$ is a column vector of unknown coefficients β corresponding to the variables coded in $\mathbf{X}$, and $\mathbf{e}$ is noise. The estimation of unknown coefficients can be done using least squared estimation (LSE) when the number of samples is greater than the number of coefficients as:

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \qquad\qquad (\text{II-5})$$

where $\mathbf{X}^T$ denotes matrix transposition and $\mathbf{X}^{-1}$ denotes matrix inversion. Then the null hypothesis can be posed on a specific coefficient (e.g., $H_0 : \beta_1 = 0$) or a specific linear combination of the coefficients, called 'contrast' (e.g., $H_0 : 2\beta_1 - \beta_2 - \beta_3 = 0$).

Computing a p-value depends on the assumptions on the noise. When the noise is assumed to be similar to that of theoretical distributions, p-values can be easily computed from the probability density function (PDF) of the theoretical distributions such as F or T distributions. This frequentist, analytic approach has the advantage of fast computation, but the weakness of being inaccurate when the distribution of the noise is far from the assumed theoretical distribution. Alternatively, the p-values can be computed without any assumption on the noise by randomization or permutation of the variable of interest (non-parametric testing). This approach provides more accurate p-value ('exact p-value' when all possible permutations were used; often 10,000 ~ 50,000 random samples of possible permutations make a good approximate) at the cost of greater computational load (a couple of days may be needed for high dimensional MRI data).

It should be stressed that a small p-value or a high value of statistics alone does not prove, but only support, that the targeted scientifically meaningful hypothesis is true. Under the NHT framework, the only possible outcome is either rejecting the null hypothesis or failing to reject it (while the absence of evidence is not the evidence of absence). In other words, NHT only can test one of the necessary conditions for our scientific hypothesis: for our hypothesis to be true, the null hypothesis should be rejected as well as other hypotheses that are different from our hypothesis. Moreover, the effect size in NHT is also an important issue. When the sample size is sufficiently large (which is getting increasingly feasible in the neuroimaging field), any non-zero effect can be declared to be 'significant' even if the effect is not physiologically meaningful. As an alternative to

the NHT framework, Bayesian inference recently has drawn the attention of scientists in various fields (Box and Tiao, 2011), which allows, for a particular model, computing the posterior probability distribution of certain parameters given the prior probability of those parameters and the data at hand.

II-4.2. Multiple Comparisons Correction

II-4.2.1. Omnibus p-value: Family-Wise Error Rate

'Massive univariate GLM', which is solving a GLM in every spatial element (e.g., voxel, vertex, electrode, sensor, and so on), have been extensively used for analyzing neuroimaging data for the sake of their simplicity in estimation. However, testing such many GLMs inflates the probability to find seemingly significant effects from any of that many spatial elements. This is a well-known problem called 'multiple comparisons' in the field of genetics, psychometrics, and many quantitative fields dealing with massive data. For instance, consider a questionnaire has 100 independent questions and the p-value for any group difference in any question is 0.05; then the probability to find any significant group difference in one or more questions is $1-(1-0.05)^{100} = 0.9940$, indicating almost always one would find seemingly significant results from a random question solely by chance. This collective probability of any false positives is called 'family-wise error rate' (FWER).

Many people tend to think that the FWER indicates the ratio of false positives under the omnibus null hypothesis (i.e., a set of null hypotheses for all voxels). For example, from 100 voxels, FWER of 5 % is often misunderstood as 5 voxels show false positives[14]. However, this notion is incorrect. Remind that p-value is a probability to observe a measure that is equal to or greater than a current observation over a sufficiently large number of realizations under the null hypothesis as:

$$p = \Pr(y \geq y_{obs} \mid \mathrm{H}_0). \tag{II-6}$$

That is, the p-value of 5 % indicates that, if one draws values from a Gaussian distribution for 100,000 times, it is expected that the one will obtain observations that exceed the current observation for 5,000 times.

Likewise, FWER indicates the probability of observing at least one false positive from any of k tests that exceeds a given threshold h as:

$$\mathrm{FWER} = \Pr(\sup_{x \in M} y(x) \geq h \mid \mathrm{H}_0), \tag{II-7}$$

---

[14] This interpretation (i.e., the number of false positives over the total number of tests) is even different from the false discovery rate, which is the number of false positives over the total number of positives.

where x is a particular test (or a voxel) included in a set M. Under the omnibus null hypothesis (i.e., all measures for all tests are from a null distribution), the FWER of 5 % indicates that, if one draws values from Gaussian distribution for k tests and n subjects for N (a sufficiently large number) times, the one would find any significant results from any of the p tests for 0.05*N times (not 0.05*k tests).

Then the question is how we should find the threshold h to control FWER as intended. In the remainder of this section, I will briefly summarize the different approaches for the multiple comparisons correction in the field of neuroimaging.

II-4.2.2. Parametric Approaches: Bonferroni Correction and False Discovery Rate

If all tests are independent, a simple solution to control FWER is to use a discounted (i.e., divided by the number of tests k) alpha-level. This procedure is called Bonferroni correction. For instance, for 100 tests and intended FWER of 0.05, the corrected alpha-level would be 0.05/100. If we compute the FWER under the null omnibus hypothesis, FWER = 1-(1-0.05/100)^{100} = 0.0488 resulting in a slight overcorrection. The other way around, one can 'correct' the nominal p-value p_n by multiplying them by the number of tests k. Then the corrected p-value p_c is given as:

$$p_c = \begin{cases} p_n k & \text{if } p_n k \leq 1 \\ 1 & \text{if } p_n k > 1 \end{cases}. \tag{II-8}$$

However, Bonferroni correction assumes independence between the measures, which is not true in many real-world datasets. Another simple and widely used approach that assumes positive dependency between the tests is called false discovery rate (FDR) (Benjamini and Hochberg, 1995). FDR is defined by the number of false positives F divided by the total number of positives (i.e., F and true positives T) as:

$$FDR = \begin{cases} \text{E}\left(\dfrac{F}{T+F}\right) & \text{if } T+F \geq 1 \\ 0 & \text{if } T+F = 0 \end{cases}, \tag{II-9}$$

where $\text{E}(\cdot)$ is the expectancy.

Although the definition of FDR is different from FWER, when the null omnibus hypothesis is correct (i.e., T = 0) as:

$$FDR = \begin{cases} 1 & \text{if } F \geq 1 \\ 0 & \text{if } F = 0 \end{cases} = \Pr(F \geq 1 \mid H_0) = FWER. \tag{II-10}$$

If there is any true positive, FWER will be greater than FDR.

The Benjamini-Hochberg procedure is quite simple. For k tests, to control FDR $\leq q$:

(1)   Sort *p*-values as $p_1 < p_2 < \cdots < p_k$.

(2)   Find the largest *i* that satisfies $p_i \leq \dfrac{i}{k} q$.

(3)   Declare $p_1, p_2, \cdots, p_i$ to be significant.

This procedure assumes independence between the tests (Benjamini and Hochberg, 1995), but a modification (Benjamini-Hochberg-Yekutieli procedure) for dependency also exists (Benjamini and Yekutieli, 2005). The step (2) is replaced by finding the largest i that satisfies $p_i \leq \dfrac{i}{k \cdot c(k)} q$ where $c(k)$ is given as 1 when the tests are independent or positively correlated; that is identical to the Benjamini-Hochberg procedure. If the tests are in arbitrary dependency, $c(k) = \sum_{j=1}^{k} \dfrac{1}{j}$ slightly relaxing the threshold.

The advantage of the FDR procedure is its simplicity enabling wide applications in neuroimaging studies (Genovese et al., 2002). However, it does not appreciate actual dependence in data, which may lead to improper correction for the multiple comparisons. For an illustration, consider 1-D data from the standard normal distribution $y \sim N(0,1)$ for *n*=20 subjects and *k*=200 independent tests (or voxels) were generated for N=2,000 realizations. For each realization, 200 one-sample T-tests were separately performed. In order to introduce spatial dependency in measurements, the simulated variable was smoothed by 1-D Gaussian kernel with FWHM of 3 voxels. The simulation results are depicted in Figure II-17. One example of a realization (one draw of a set of variables for multiple tests), correlation matrices between voxels are given for independent (Figure II-17 (A) and (E)) and dependent (Figure II-17 (E) and (F)) variables, respectively. In this realization, more than one false positives (FP) are found in both of independent and dependent data without any correction for multiple comparisons (Figure II-17 (C) and (G)). As noted above, the number or proportion of FP (in a single realization) should not be confused with the FWER (over sufficiently many realizations). Over 2,000 realizations (or draws), when the variables were independent, Bonferroni and FDR methods actually slightly under-corrected FWER. That is, whereas the intended FWER was 0.05, the FWER after Bonferroni was 0.0605 and that of FDR was 0.0615 (Figure II-17 (D)). Conversely, when the variables are spatially dependent, Bonferroni (0.0295) and FDR (0.036) over-corrected FWER (0.0295 for Bonferroni correction; 0.036 for FDR correction; Figure II-17 (H)), which would reduce sensitivity of the inference.
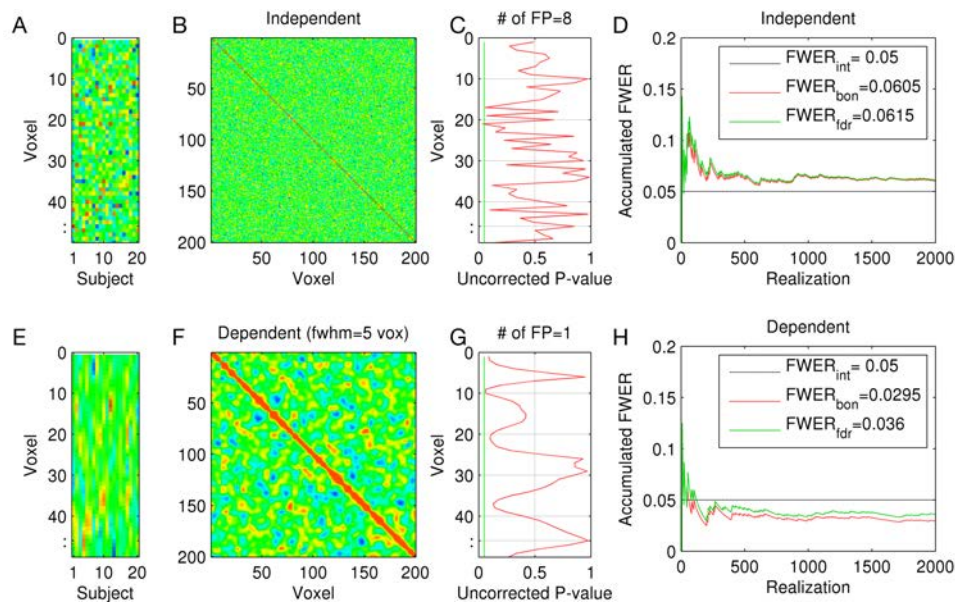
Figure II-17. Examples of parametric approaches to control family-wise error rates (FWERs) for independent and dependent data. The data was generated from zero-mean unit-variance Gaussian noise for 2,000 times: one realization of independent (A) and dependent (E) data. Their independency and dependency is depicted by voxel-by-voxel correlation. Unlike the independent case (B), smoothed noise show dependency between adjacent voxels (F). Uncorrected p-values for one realization (C) and (G) with the alpha level of 0.05 marked by a green vertical line. Accumulated FWER (the number of realization rejecting one or more null hypothesis over the number of realizations) are also given for independent (D) and dependent (H) cases. Abbreviations: FP, false positive; FWER$_{int}$, intended control of FWER; FWER$_{bon}$, FWER after Bonferroni correction; FWER$_{fdr}$, FWER after false discovery rate correction.

The problem of overcorrection by using Bonferroni and FDR is further depicted in Figure II-18. When the data is independent (FWHM = 0), Bonferroni (blue) and FDR (yellow) work well (i.e., after correction, FWER is more or less well controlled at intended level of 0.05). However, when the data is spatially dependent (FWHM > 0), both methods over-corrected FWER, more severely by Bonferroni correction. This illustration shows that, without assessment and corporation of dependency in the data, particularly neuroimaging data, the sensitivity and statistical power could be severely restricted. Because of that, sophisticated methods to properly control the FWER has been proposed since the early period of this newly arrived field of neuroimaging (Worsley et al., 1992). In the remainder of this section, I will discuss approaches that uses information on dependency from the data: random field theory and permutation tests.
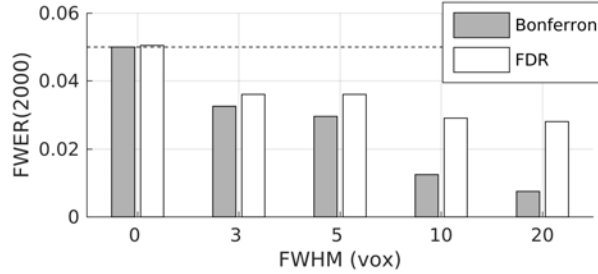
Figure II-18. Effect of smoothness (spatial dependency) on the family-wise error rates (FWERs) after multiple comparison correction with Bonferroni and false discovery rate (FDR) methods from 2,000 realizations of variables. Abbreviations: FWHM, full-width at half-maximum; vox, voxel.

II-4.2.3. Topological Approach: Random Field Theory

Suppose the noise in the neuroimaging data is a smooth and differentiable Gaussian random field $y(x)$ defined over Riemannian manifold M embedded in an arbitrary N-dimensional space (i.e., a 3-D volume or a 2-D surface space). As abovementioned in (II-7), FWER is given as a probability of s supremum of a smooth, differentiable Gaussian field to exceed a given threshold h under the null hypothesis. The analytic computation of the exact distribution of the supremum of the random field is difficult (Chung, 2012). However, there is an alternative way to compute the threshold h using the expected Euler Characteristic (EC) of the excursion set $A_h = \{x \in M : y(x) > h\}$, which has been introduced by Worsley (Worsley et al., 1992), inspired by Adler's random field theory (RFT) (Adler, 1981). The Euler Characteristic approach reformulates the geometric problem as a topological problem (Chung, 2012). Further details can be found in in literature (Taylor and Worsley, 2007).

When assuming a very high h (thus the excursion set is empty), the EC of the set is 0. We now decrease h until it touches the supremum of the random field. When $\sup_{x \in M} y(x) = h$, the EC of the excursion set is 1 (for one vertex). Between those two thresholds, the expected EC is bound between 0 and 1. For a sufficiently high h that bounds the EC of the excursion set between 0 and 1, the following is known from the RFT:

$$\Pr(\sup_{x \in M} y(x) > h) \approx E\left(\chi(A_h)\right) \qquad (\text{II-11})$$

where $E(\cdot)$ is the expectancy and $\chi(\cdot)$ is the EC of a given set (Chung, 2012). To compute the expected EC of the excursion set, the 'resel' (resolution element) count and EC density was introduced (Worsley et al., 1996). For instance, the FWER in the 3-D space is given by:

$$\Pr(\sup_{x \in M} y(x) > h) \approx \sum_{D=0}^{3} R_D(V)\rho_D(h) \ , \qquad \text{(II-12)}$$

where, for each dimension D, $R_D(V)$ is the number of 'resels' (resolution elements) in the search region V, which depends on the smoothness of the field and $\rho_D(h)$ is the EC density, which depends on the threshold (Worsley et al., 1996).

Follow-up studies on the use of RFT for proper control of FWER dealt with the accurate estimation of the smoothness of the field (Kiebel et al., 1999), with the correction for the non-stationary noise field (Worsley et al., 1999), and other problems. The RFT correction for a 3-D field has become the standard method to control FWER in SPM. For the 2-D random field, a software named SurfStat for surface-based analysis (Worsley et al., 2009a) was developed by Worsley and colleagues.

In addition to the height threshold, an alternative approach that regards height and the spatial extent of clusters was proposed (Friston et al., 1994). A non-parametric method for cluster-wise inference was later developed called threshold-free cluster enhancement (TFCE) (Smith and Nichols, 2009), which is in FSL. Very recently, concerns on the cluster-level (or extent-based) RFT correction, which has been very commonly used in many fMRI studies, drew a lot of attention in the neuroimaging field (Eklund et al., 2016). In the paper (Eklund et al., 2016), the authors claimed that FWER is poorly controlled with the extent-based than the height-based thresholding: from certain datasets, FWER with cluster-level threshold was greater than 15 % whereas the intended FWER level was 5 %.

However, the study (Eklund et al., 2016) was followed by a re-analysis demonstrating the poor performance of cluster-extent thresholding was due to up-sampling (Flandin and Friston, 2016). The reported inflation of FWER by the cluster-extent thresholding in Eklund et al. (2016) was only found when the functional images were resampled at the 2-mm-isotropic resolution, which is a default setting of a previous version of analysis software (i.e., SPM8). When the images were resampled at the 3-mm-isotropic resolution, which is a default setting of the current version (i.e., SPM12), the cluster-extent thresholding works as well as peak-height thresholding (Flandin and Friston, 2016). In the paper (Flandin and Friston, 2016), it was suggested that the only issue with the cluster-extent thresholding is the cluster-forming threshold, which is used to threshold the statistical image for defining the size of clusters. In fact, this point was already demonstrated earlier (Woo et al., 2014). In the study (Woo et al., 2014), authors showed that, from simulated and real datasets, when a liberal (>0.001, sometimes even 0.01) threshold was used to form clusters, the FWER was poorly controlled (i.e., FWER was greater than the intended upper bound), suggesting a strict (≤0.001) cluster-forming threshold in statistical inference of the functional neuroimaging data.

II-4.2.4. Non-parametrical Approach: Permutation Test

Permutation (also known as randomization) was earlier mentioned as a way to compute 'exact p-values' without any assumption on the noise distribution. Randomization can also be used to control FWER (Nichols and Holmes, 2002). The procedure is similar to that of a single test, except that instead of a single test statistics, the maximum of the multiple test statistics should be computed. Based on many permutations, a null distribution of the maximal test statistic will be obtained. Then, the FWER can be computed by the proportion of the permutations exceeding the actual observation over the total number of permutations. As already mentioned, this is a very powerful and straightforward method, but it is computationally very expensive and in some high dimensional cases, it could be impractical except some methodological studies validating parametric approaches (Eklund et al., 2016; Flandin and Friston, 2016; Woo et al., 2014).

III. Study #1: Intracortical myelination in musicians with absolute pitch — quantitative morphometry using 7-Tesla MRI

In this chapter, I will introduce a study on the intracortical myelination using 7-Tesla MRI with the magnetization-prepared two rapid gradient echo (MP2RAGE) imaging sequence. Important findings in this chapter were already published in a peer-reviewed journal (Kim and Knösche, 2016)[15]. As the main result, we found the cortical myelination in an anterior region of the right supratemporal plane (in the right planum polare; right PP) of musicians positively related to the degree to which these persons possess absolute pitch (AP). I will discuss the implications of this finding in terms of the electrical properties of intracortical connections, suppression of neuroplasticity after a 'critical period', and the ventral auditory pathway for processing object-related information.

III-1. Motivation and Hypotheses

As extensively reviewed in the Chapter I, many neuroimaging studies indicated important candidates for the neural substrate of AP by describing macroscopic morphology. Briefly summarized, the reported findings include smaller right planum temporale area (PT; thus increasing leftward asymmetry index) (Keenan et al., 2001; Schlaug et al., 1995), greater volume of the right Heschl's gyrus (HG) (Wengenroth et al., 2014), greater cortical thickness in bilateral superior temporal gyri (STGs) and inferior frontal gyri (IFGs) and greater fractional anisotropy (FA) of diffusion tensors in the right occipitofrontal fasciculus (Dohn et al., 2015), greater FA of diffusion tensors in the left superior longitudinal fasciculus (Oechslin et al., 2009) in musicians with than without AP.

Despite of the known structural alterations localized in fronto-temporal networks, a mechanic account on how pitch chroma is extracted from sound and recognized in an absolute sense is still far from reached. In order to advance towards a computational model of AP, further detailed structural information is essential. Recently, in-vivo imaging of myeloarchitecture of the human cortex has drawn much attention in the neuroimaging field (Blackmon et al., 2011; De Martino et al., 2014; Dick et al., 2012; Geyer et al., 2011; Glasser and Van Essen, 2011; Hashim et al., 2015; Lutti et al., 2014; Shafee et al., 2015; Sigalovsky et al., 2006) due to its significant implication in the working principles of the human cerebral cortex (Nieuwenhuys, 2013b). Unlike conventional morphometric methods that only use arbitrary image intensity of T1-weighted images, such as

---

[15] http://doi.wiley.com/10.1002/hbm.23254

voxel-based morphometry (VBM) (Ashburner and Friston, 2000) or cortical thickness analysis (Fischl and Dale, 2000), this novel technique can infer the myelin content in a voxel based on the inverse relationship between myelin content and quantitative longitudinal relaxation time (qT1) (Geyer et al., 2011; Marques et al., 2010). Experiments that demonstrated the relevance of intracortical myelin to brain function and behavior already exist (Grydeland et al., 2013; Grydeland et al., 2015). Moreover, applications of in-vivo myelin mapping demonstrated that the separation of the core and belt regions of the human auditory cortex is feasible by in-vivo imaging (De Martino et al., 2014; Dick et al., 2012) based on the same myeloarchitectonic principles that are used in histological studies, namely, the core is defined by denser myelination (Wallace et al., 2002).

Intracortical myelination may be beneficial for neural circuits that enable AP, because cortical myelin provides electric insulation that reduces ephaptic crosstalk between nearby axons and thereby contributes to the specificity of transmission. That is, cortical myelin may support higher precision in categorization of pitch chroma. Moreover, cortical myelination development is known to suppress neuroplasticity after the "critical period" (McGee et al., 2005). This fixation effect of cortical myelination is particularly interesting given the known importance of musical experience during the early stage of life (i.e., between the ages of 4 and 7 years) in acquiring AP (Baharloo et al., 1998; Miyazaki, 2004a; Miyazaki et al., 2012). Thus, greater cortical myelination may indicate that more information is preserved in neural circuits, which lasts throughout life.

Hence, the degree of axonal myelination could be an important structural property of neural tissue with respect to AP ability, which goes beyond the mere size (area or volume) of macroscopically delineated sections of the cortex. This hypothesis will be tested in the present study by comparing cortical myelination between musicians with and without AP.

If the above hypothesis is confirmed and myelination is discriminative between AP and non-AP brains, we will also be able to shed new light on the location of the AP processor. Importantly, such a result is expected to directly point to the locations where increased or altered computational procedures associated to AP processing are carried out. Because of the speed and accuracy of AP processing indicated by behavioral investigations (Miyazaki, 1990) and the short latency of the AP-specific ERP component (Itoh et al., 2005), it is highly likely that early auditory processing is involved in chromatic categorization. Thus it is reasonable to expect greater myelination in regions of the auditory cortex (either primary or non-primary).

Finally, in light of cortical myelination, we will elucidate the question whether chromatic processing in AP musicians is spatially distinct from mere frequency discrimination, as suggested by previous behavioral evidence. In an early study (Siegel, 1974), AP listeners showed better pitch discrimination performance than control listeners only when the reference tone was tuned in to standard pitch (i.e., A4 = 440 Hz). This advantage of AP musicians was abolished when the tones were tuned in a non-standard scale (i.e., A4 is not exactly 440 Hz), implicating that pitch

discrimination ability is independent of AP acuity, as extensively discussed in the literature (Deutsch and Henthorn, 2004; Miyazaki, 1988; Miyazaki, 2004a; Takeuchi and Hulse, 1993). Accordingly, we hypothesize that myeloarchitectonic features related to the frequency discrimination threshold (FDT) (Micheyl et al., 2006) would be found in auditory cortical regions that are spatially distinct from the myeloarchitectonic correlates of AP. More specifically, because pitch-selective neurons were found near the anterolateral border of primary auditory cortex in marmoset monkeys (Bendor and Wang, 2005) and at the anterolateral end of HG in humans (Penagos et al., 2004), we hypothesize that pitch discrimination precision may be related to myeloarchitecture in the lateral HG extending to superior temporal gyrus (STG).

To achieve these aims, we map quantitative longitudinal relaxation rates ($qR_1 = 1/qT_1$) in the cortex using ultra-high-field (7 Tesla) MRI, which allows for sub-millimeter resolution, and we investigate whether and where there is a relationship between the concentration of myelin in the cortex on the one hand, and the AP and FDT abilities on the other.

III-2. Materials and Methods

III-2.1. Participants

Eight AP musicians (five women) and nine non-AP musicians (five women) participated in MRI and behavioral experiments at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany. Participants were recruited via the Institute's participant database and flyers posted in the University of Music and Theatre "Felix Mendelssohn Bartholdy" Leipzig and the University of Leipzig. Inclusion conditions comprised: age between 18 and 40 years, musical training of more than 10 years, right-handedness (i.e., laterality coefficients of handedness (LQ) ≥ 70 (Oldfield, 1971)), absence of any contraindication for MRI scanning, and absence of any history of neurological disorders, psychiatric diseases, use of psychiatric medications, head trauma, hearing loss, and tinnitus. An abbreviated audiometry, which was used to equalize stimuli presentation loudness across participants, confirmed intact hearing of all musicians. Prior to recruitment of musicians who identified themselves to have AP, a web-based AP test was used to confirm it (≥ 80 % correct answers). There was one case of self-reported absence of AP, which was re-categorized into the AP group based on that test. The local ethics committee approved the experimental protocol, and all participants submitted written informed consents prior to experiments.

Table III-1. Mean and standard deviation of demographic information

Variables	Non-AP (n= 9)		AP (n= 8)		t/z-statistic	p-value
	Mean	(Std.)	Mean	(Std.)		
Sex ratio (Women/all)	0.56	-	0.62	-	0.22	0.823

Age (year)	25.78	(5.19)	26.88	(2.95)	0.53	0.607
Handedness (LQ)	90.56	(11.54)	95	(9.26)	0.87	0.399
MET-melody hit rate (%)	82.26	(11.12)	91.11	(10.02)	1.71	0.107
FDT*	2.41	(0.23)	2.73	(0.68)	-1.34	0.199
Ethnicity ratio (Asian/all)	0.00	-	0.38	-	2.91	0.004
Musical training onset (year)	8.00	(2.96)	5.00	(1.51)	-2.88	0.012
Musical training duration (year)	16.9	(7.83)	23.00	(3.34)	2.04	0.059

* FDT is transformed in negative logarithm base to 10. Abbreviations: LQ, laterality coefficient, from -100 (exclusively left-handed) to 100 (exclusively right-handed) (Oldfield, 1971); MET, Musical Ear Test (Wallentin et al. 2010); FDT, frequency discrimination threshold (Micheyl et al., 2006); Std., standard deviation.

Demographic information of AP and non-AP musicians is listed in Table III-1. The musical aptitudes were estimated by the "Musical Ear Test" (Wallentin et al. 2010). Here, only the scores of the melodic tasks are reported, while rhythmic tasks are left out. Frequency discrimination threshold (FDT = F_1/F_0; F_1 and F_0 are the frequencies of target and reference, respectively) was measured using an abbreviated version of a protocol presented in Micheyl et al. (2006). Sex, age, handedness, musical aptitude, and logarithmic FDT were matched between the groups (minimum $p = 0.107$). The first musical instrument was for the AP musicians (n = 9): piano (6), violin (1), guitar (1), and clarinet (1); for the non-AP musicians (n = 8): piano (5), violin (2), and guitar (1).

There were mismatches between the AP and non-AP groups in terms of ethnicity. This imbalance is rooted in the much higher incidence of AP in Asian as compared to European musicians (Miyazaki et al., 2012), which makes it difficult to recruit balanced cohorts. Nonetheless, using a simulation, we demonstrated that the reliability in estimating the effect size of AP under the current design (relative error = -0.09 ± 13.16 %) is still comparable to a perfectly balanced design (relative error = 0.05 ± 9.41 %) and much better than the worst case (relative error = -0.39 ± 21.39 %). In an additional analysis, only with European participants (see the Section III-3.5.1), we replicated the main finding even from that small subset, which means that the result cannot be due to confounding influence from ethnicity and therefore this speaks for the robustness of the findings.

III-2.2. Behavioral Test

The AP performance of all musicians was measured by a behavioral test. Details of the AP test are described in Section II-1.1. In the test, the participants heard a random sequence of pure and piano tones tuned to the 12-key equal-tempered scale with A4 as 440.0 Hz and then pressed the corresponding keys on a muted digital piano. The target tones spanned 3 octaves: C3 (130.81 Hz) to B5 (987.77 Hz) for sine waves and E3 (164.81 Hz) to D#6 (1,244.50 Hz) for piano timbre.

The absolute pitch performance was analyzed in terms of (1) absolute error (AE), (2) absolute octave-corrected error (ACE), (3) difference between mean AE and mean ACE, as a measure of

octave error, (4) AP score (APS), (5) hit rate, and (6) reaction time (RT) for pure and piano tones. "Octave-error" is defined by an answer with correct pitch chroma but incorrect pitch height, which is known to be frequently observable in AP musicians (Miyazaki, 1988; Miyazaki, 2004a). APS is defined as 1-(octave-corrected error)/6 between 0 and 1 with 0.5 as a change level. 6 semitones is possible largest octave-corrected error.

III-2.3. Image Acquisition and Processing

Magnetization-prepared two rapid gradient echo (MP2RAGE) (Marques et al., 2010) images at 0.7-mm isotropic resolution were acquired using a 7-T whole-body MR system (Siemens, Erlangen, Germany) with an 8-channel head coil system (RAPID MR International, Ohio, USA). From the two inversion images, a T1-weighted (T1w) image and a qT1 image were derived (Marques and Gruetter, 2013). Further details can be found in Section II-2.3.

Cortical surfaces were reconstructed using FreeSurfer. A skull-stripped T1w image processed using VBM8 and manual adjustment (as explained in Section II-3.4) was down-sampled to 1-mm isotropic resolution and fed into the FreeSurfer standard script except the skull stripping. The qT1 images were mapped onto cortical surfaces generated by FreeSurfer (Fischl, 2012a) at various depths (i.e., 25, 50, and 75% of cortical thicknesses, measured from the interface between gray and white matter). Subsequently, the qR1 (= 1 / qT1) values were computed from the surface-mapped qT1 values and bounded within a range between 0.25 and 10 s^{-1}, which corresponds to the qT1 range between 100 and 4,000 msec. Following this, they were registered onto the template meshes.

III-2.4. Statistical Inference

Statistical tests were carried out using a GLM for each layer and each hemisphere with the nuisance covariates age, sex, and ethnicity. Multiple comparison correction based on the RFT was applied after surface-based smoothing with FWHM of 8 mm (i.e., 10 iterations on the mesh with a mean edge length of 0.77 mm) using the SurfStat Toolbox (Worsley et al., 2009b). Because the signal drop in the original 7-T image depends on the position in the head coil, systematic bias due to gross morphology such as the inferior-to-superior length of a head could alter qR1 values and statistical inference. Thus we limited the search region to accurately estimate the number of resells (resolution elements) in RFT (Worsley et al., 1992) by excluding ventral regions (i.e., temporal poles, fusiform gyri, entorhinal cortices, parahippocampal gyri, and inferior temporal lobes) as well as the medial wall (a cut section of the corpus callosum and subcortical structures to separate hemispheres) using an automatic parcellation based on the 'Desikan-Killiany-Atlas' (Desikan et al. 2006) from the FreeSurfer. In the current study, the FWER was controlled to be below 0.05 at the cluster level with a cluster-forming threshold of 0.001.

III-3. Results

III-3.1. Behavioral Results

In our experiment we allowed responses only within 4 seconds after the stimulus onset, causing some of the trials to be left unresponded. Additionally, in some cases, multiple keys were pressed within the time window. If the first response was given later than 100 msec after the stimulus onset, it was taken as a valid response; otherwise it was regarded as a late response to the previous trial. The overall percentage of valid responses was 93.7 ± 8.1 % (non-AP) and 94.4 ± 1.0 % (AP).
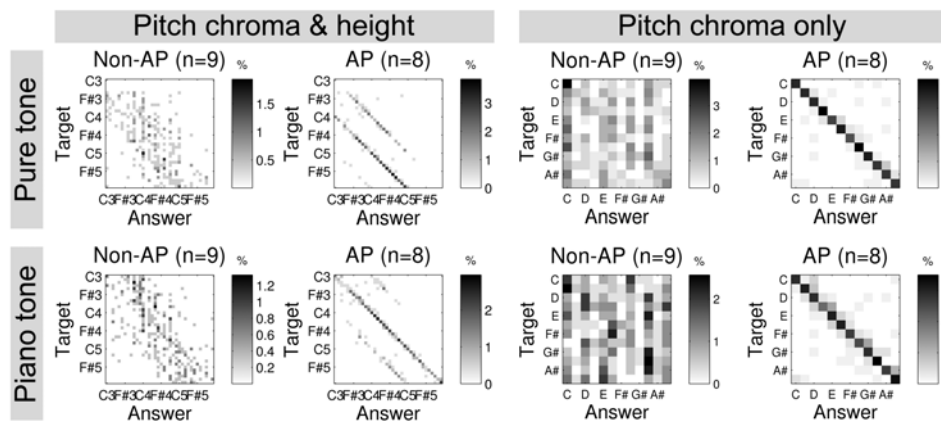
Figure III-1. Confusion matrices from the absolute pitch (AP) test for non-AP and AP groups. For each matrix, rows correspond to the target notes and columns correspond to the answers of the participants. For pure tones (upper row) and piano tones (lower row), the four matrices on the left side show raw responses ("Pitch chroma & height") and octave error corrected responses ("Pitch chroma only"). Gray scale codes the proportion of the trial numbers for a specific combination of target and answer notes over the total number of trials.

The group differences in the AP test are illustrated using confusion matrices in Figure III-1. The row of each matrix corresponds to the pitch of a presented tone. The column corresponds to the pitch of the answer: An element at the i-th row (e.g., C4) and j-th column (e.g., C3) represents how often (on average) subjects reported the j-th note (e.g., C3) when the target was actually the i-th note (e.g., C4). AP musicians showed distinctively sharp diagonal elements (meaning correct answers) whereas non-AP musicians did not show such a pattern, but a very broad scatter over nearly one octave. As commonly reported in the literature (Miyazaki, 1988; Takeuchi and Hulse, 1993), it was notable that AP musicians also made 'octave errors', which are defined as answers with correct pitch chroma but incorrect pitch height. In accordance with a previous study (Miyazaki, 1989), AP musicians made more octave errors with pure tones (59.6 ± 19.3% of total

valid answers) than with piano tones (14.6 ± 12.3 % of total valid answers; p < 0.0001), which could be explained by the relative unfamiliarity of the timbre of sine waves to musicians.

Table III-2. Behavioral measures for absolute pitch performance.

Measure	Timbre	Non-AP (n= 9)		AP (n= 8)		t-statistic	p-value
		Mean	(Std.)	Mean	(Std.)		
AE (semitones)	Pure tone	11.44	(6.90)	10.81	(4.52)	0.22	0.831
	Piano tone	8.59	(5.05)	3.69	(2.14)	2.54	0.023
ACE (semitones)	Pure tone	2.89	(0.26)	0.36	(0.51)	13.18	$<10^{-8}$
	Piano tone	2.99	(0.43)	0.53	(0.57)	10.11	$<10^{-6}$
APS	Pure tone	0.52	0.04	0.94	0.08	-13.18	$<10^{-8}$
	Piano tone	0.50	0.07	0.91	0.10	-10.12	$<10^{-7}$
Hit rate*	Pure tone	0.27	(0.07)	0.87	(0.21)	8.31	$<10^{-5}$
	Piano tone	0.29	(0.09)	0.89	(0.17)	9.02	$<10^{-5}$
RT (sec)	Pure tone	2.02	(0.64)	1.59	(0.65)	1.36	0.194
	Piano tone	1.92	(0.48)	1.69	(0.74)	0.79	0.441

Abbreviation: AE, absolute error; ACE, absolute corrected error; APS, absolute pitch score; RT, reaction time; Std., standard deviation. * An error with a semitone was considered as a correct response as in some of previous literature (Keenan et al., 2001; Schulze et al., 2013).

Table III-2 summarizes the statistics of AP performance including 'hit rates' (regarding errors up to one semitone as 'hit') in order to enable comparisons to previous studies (Keenan et al., 2001; Schulze et al., 2013). The mean values of AE were smaller in AP only for piano tones (p = 0.028) but not for pure tones (p = 0.777). In contrast, the mean values of ACE were smaller in AP both for pure and piano tones (p < 10^{-6}). The indifference in AE for pure tones is possibly due to the persistent octave errors in some musicians with AP (e.g., in an extreme case, octave errors were found in 83% of valid responses). The hit rates were also significantly different between APs and non-APs (p < 10^{-5}), consistently with the literature (Keenan et al., 2001; Schulze et al., 2013). The average hit rates were slightly lower than in previous studies (Keenan et al., 2001; Schulze et al., 2013) because those studies applied a more stringent threshold (> 90 %) (Miyazaki, 1988). In our case, hit rates from two musicians were below 90%, thus those participants could be classified as 'quasi-APs' (Wilson et al., 2009). Because of this, we also did correlation analysis with APS, which was measured for all musicians.

Unexpectedly, no significant difference in the RT was found between non-AP and AP musicians (min p = 0.179), although the participants were instructed to respond as quickly as possible while maintaining the accuracy of responses within the 4-sec time window. The mean RTs in the current study (1.97 sec in non-AP; 1.64 sec in AP) were much smaller than in a previous

study that allowed self-paced timing (7.6 sec in non-AP; 3.3 sec in AP) (Bermudez and Zatorre, 2009a). Because of the fixed response window, the non-AP participants could not take longer than AP participants as in the previous study (Bermudez and Zatorre, 2009a), which presumably is the reason for the observed lack of difference in RTs.

III-3.2. In-vivo Intracortical Myelination Mapping

Examples of surface-mapped qR1 values of a single subject and the overall mean (n= 17) are shown in Figure III-2. In high accordance with previous in-vivo myelination studies (De Martino et al., 2014; Geyer et al., 2011; Glasser and Van Essen, 2011; Shafee et al., 2015), high qR1 values (indicating high myelination) were observed in primary motor/somatosensory cortices (the pre/post-central gyri), primary auditory cortex (the medial region of Heschl's gyrus), and primary visual cortex (the occipital poles and the precuneus).
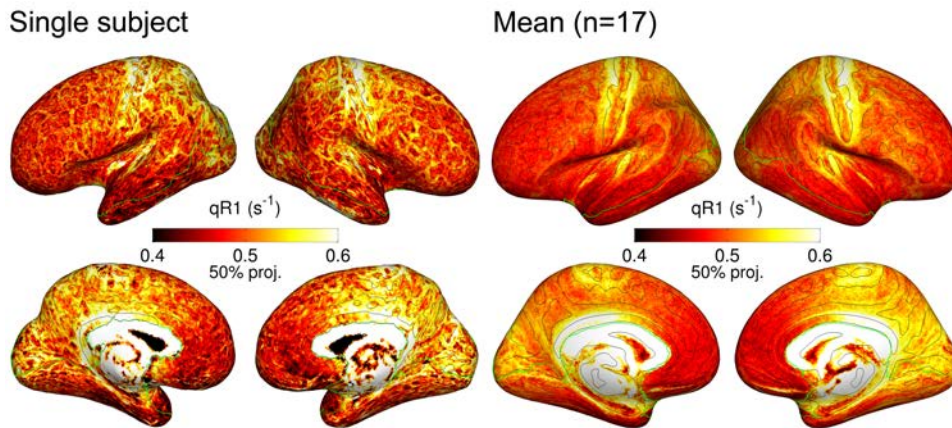
Figure III-2. Surface-mapped quantitative longitudinal relaxation rate (qR1) values of a single subject (left) and the average over all subjects (right, n=17). qR1 values were sampled at the 50% projection level of cortical thickness from the white matter surface. Masks excluding ventral regions are indicated by green contours (see the Section III-2.4). Note that the values of the single subject were minimally smoothed for visualization (FWHM= 1 mm) whereas the group mean values were not smoothed. Isocurvature contours (black) at zero level are overlaid to aid localization of landmarks of the inflated cortical surfaces.

III-3.3. Effects of Demographic Variables in Cortical Myelin

In assessing neuronal structures, nuisance demographic variables are commonly covaried for their influence in global and local morphometry. For instance, age (Goncalves et al., 2001; Thambisetty et al., 2010), sex (Good et al., 2001; Gur et al., 1999), and ethnicity (Chee et al., 2011) are known to correlate with macroscopic morphology such as local gray matter volume and cortical thickness. Also for the intracortical myelination, significant age effects in T1w/T2w ratio images were

reported (Grydeland et al. 2013; Shafee et al. 2015). Therefore it would be important to know whether the demographic variables significantly affected qR1 values in our current dataset as well. Thus we tested GLMs for the effects of age, sex, and ethnicity as:

$$qR_1 = \beta_0 + D\beta_1 + \varepsilon , \tag{III-1}$$

where D is a demographic variable, which was either age, sex, or ethnicity. As shown in Figure III-3 and Table III-3, we found significant effects over extensive areas including right inferior frontal gyrus, right anterior cingulate gyrus, and right superior temporal gyrus (age effect; Figure III-3, upper row), right superior temporal sulcus (sex effect; Figure III-3, middle row), and left planum temporale and supramarginal gyrus (ethnicity effect; Figure III-3, lower row). Given these significant effects of demographic variables, we incorporated them as covariate terms in order to control for possible confounding.

Figure III-3. The effect of demographic variables (age, sex, and ethnicity) on quantitative longitudinal relaxation rate (qR1).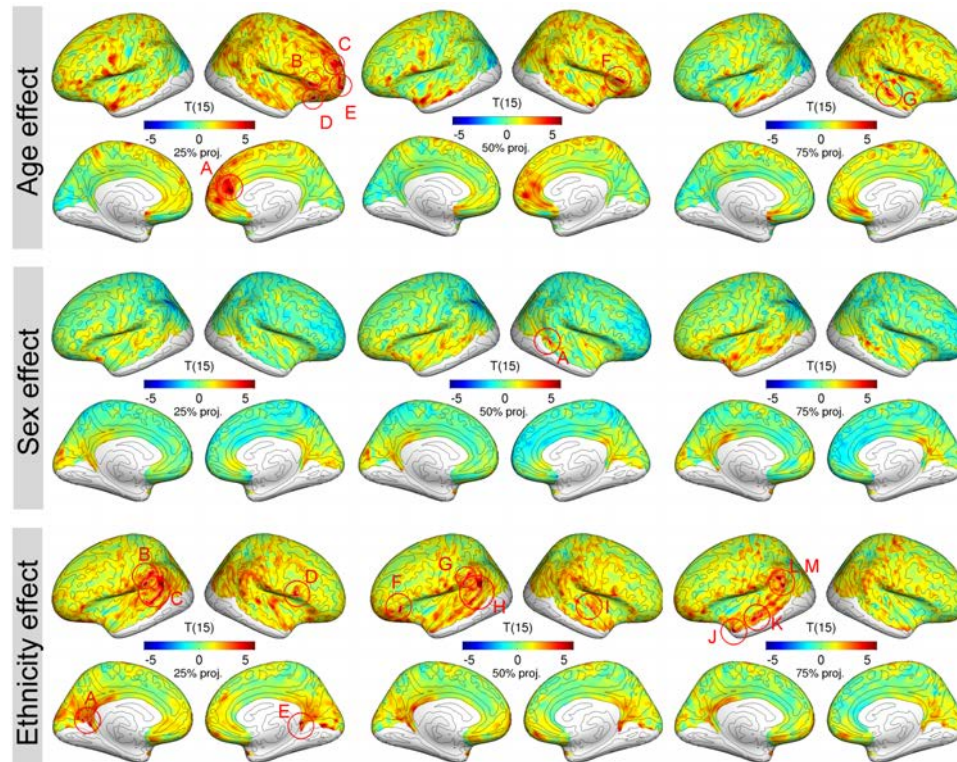 T-statistic maps are given for each projection level. Red circles mark the labeled local peaks. See Table III-3 for details of significant clusters.

Table III-3. Significant clusters for the effect of demographic variable

Label	Side	Level*	Structure name	T (15)	p-value+	Area (mm²)	MNI-305 (x, y, z mm)		
Age effect									
A	Right	25%	Right anterior cingulate cortex	6.15	0.0037	308.8	1.6	42.6	9.2
B	Right	25%	Right triangular part of the interior frontal gyrus	5.18	0.0111	108.2	51.1	26.3	-0.5
C	Right	25%	Right middle frontal sulcus	4.99	0.0188	314.0	23.7	44.2	19.6
D	Right	25%	Right orbital gyri	6.24	0.0184	72.0	37.2	25.7	-22.4
E	Right	25%	Right fronto-marginal gyrus/sulcus	4.99	0.0055	241.9	24.1	50.7	2.0
F	Right	50%	Right orbital part of the interior frontal gyrus	5.93	0.0092	81.4	53.0	30.3	-2.6
G	Right	75%	Right superior temporal sulcus	6.26	0.0047	109.1	65.3	-11.1	-2.5
Sex effect									
A	Right	50%	Right superior temporal sulcus	4.71	0.0434	46.4	44.3	-38.8	-0.9
Ethnicity effect									
A	Left	25%	Left mid-post cingulate cortex	5.26	0.0064	123.4	-20.8	-53.2	-1.4
B	Left	25%	Left supramarginal gyrus	5.50	0.0069	185.7	-64.9	-34.1	36.2
C	Left	25%	Left planum temporale	6.06	0.0001	426.9	-64.6	-52.5	18.3
D	Right	25%	Right opercular part of the interior frontal gyrus	4.78	0.0356	85.8	55.1	14.2	13.9
E	Right	25%	Right parahippocampal gyrus	5.48	0.0088	86.6	6.1	-45.0	1.0
F	Left	50%	Left orbital part of the interior frontal gyrus	4.95	0.0249	46.6	-50.5	32.3	-11.3
G	Left	50%	Left supramarginal gyrus	4.94	0.0436	103.8	-64.9	-34.5	35.6
H	Left	50%	Left planum temporale	10.14	0.0001	308.4	-64.6	-52.5	18.3
I	Right	50%	Right planum polare	4.68	0.0460	57.1	46.5	-14.0	3.7
J	Left	75%	Left temporal pole	7.80	0.0375	68.2	-45.8	18.8	-30.5
K	Left	75%	Left superior temporal sulcus	5.29	0.0296	145.3	-53.9	-13.3	-16.7
L	Left	75%	Left planum temporale	6.46	0.0069	83.8	-60.5	-43.8	28.1
M	Left	75%	Left planum temporale	7.17	0.0072	48.1	-63.7	-53.8	15.5

* Projection level indicates its sampling point from gray matter and white matter interface along an outward normal vector by 25 % (deep), 50 % (middle), or 75 % (superficial) of cortical thickness. + P-values are corrected for multiple comparison correction at cluster level.

III-3.4. Effects of Absolute Pitch in Cortical Myelin

We tested the effect of AP while controlling for confounding effects from the nuisance demographic variables, as discussed above, using a GLM as:

$$qR_1 = \beta_0 + age \cdot \beta_1 + sex \cdot \beta_2 + ethnicity \cdot \beta_3 + AP \cdot \beta_4 + \varepsilon, \quad \text{(III-2)}$$

where qR_1 is a column vector with one qR1 value per vertex; *age*, *sex*, *ethnicity*, and *AP* are column vectors coding age (years), sex (male or female), ethnicity (Asian or European) of musicians, and AP level (non-AP or AP); β_i is the *i*-th unknown coefficient to estimate; ε is zero-mean Gaussian noise.
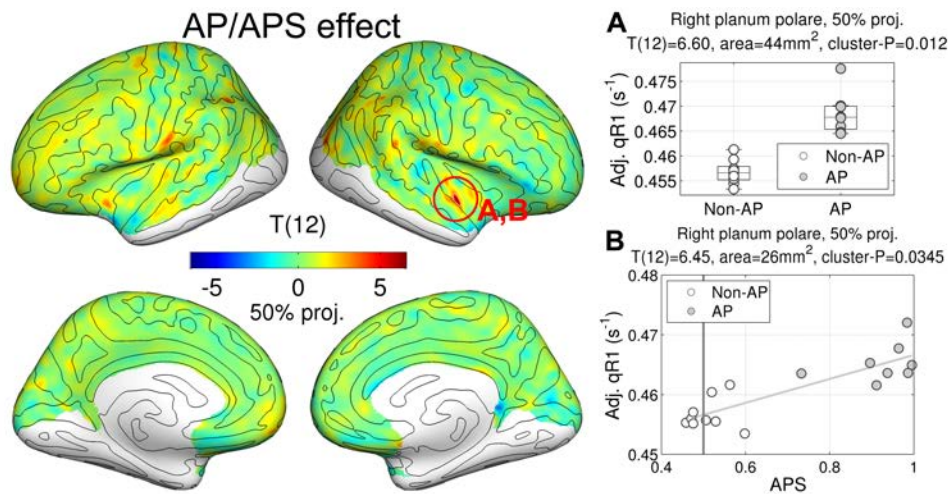
Figure III-4. Effects of absolute pitch (AP) and AP score (APS) in quantitative longitudinal relaxation rate (qR1). T-statistic maps for group differences are shown on inflated cortical surfaces. T-statistic maps for APS effect are not shown because they are very similar to those of AP effect. The degree of freedom is noted in parentheses above the color bar. Regression plots for AP effect (A) and APS effect (B) are given. The expected APS by chance (0.5) is marked by a gray vertical line in the plot (B). Note that the qR1 values in scatterplots are adjusted for the nuisance variables in the GLM using (III-2).

We found greater qR1 values in the AP group compared to the non-AP group in the right planum polare (PP) (max $T(12) = 6.60$, $p = 0.012$, area = 40 mm^2 at 50% projection level, peak MNI-305 coordinate = (46, -8, -12) mm) as shown in Figure III-4 (A). Moreover, we analyzed the correlation between qR1 values and APS. Because APSs for pure tones and piano tones were very similar ($r = 0.96$, $p < 10^{-8}$), the APSs for the two timbres were averaged for the correlational analysis. We used the same GLM (III-2) replacing the binary variable of AP group membership with a continuous behavioral variable of APS. We found a significant positive correlation in the right PP in the same position as for the AP (max $T(12) = 6.45$, $p = 0.035$, area = 23 mm^2 at 50 % projection level, peak MNI-305 coordinate = (46, -5, -12) mm) as shown in Figure III-4 (B).

In order to appreciate the result in the original anatomical context, the significant clusters for the AP contrast in Figure III-4 were transformed back into the native volumes of participants as shown in Figure III-5. The cluster voxels are all aligned along the medial part of the PP in the right hemisphere.
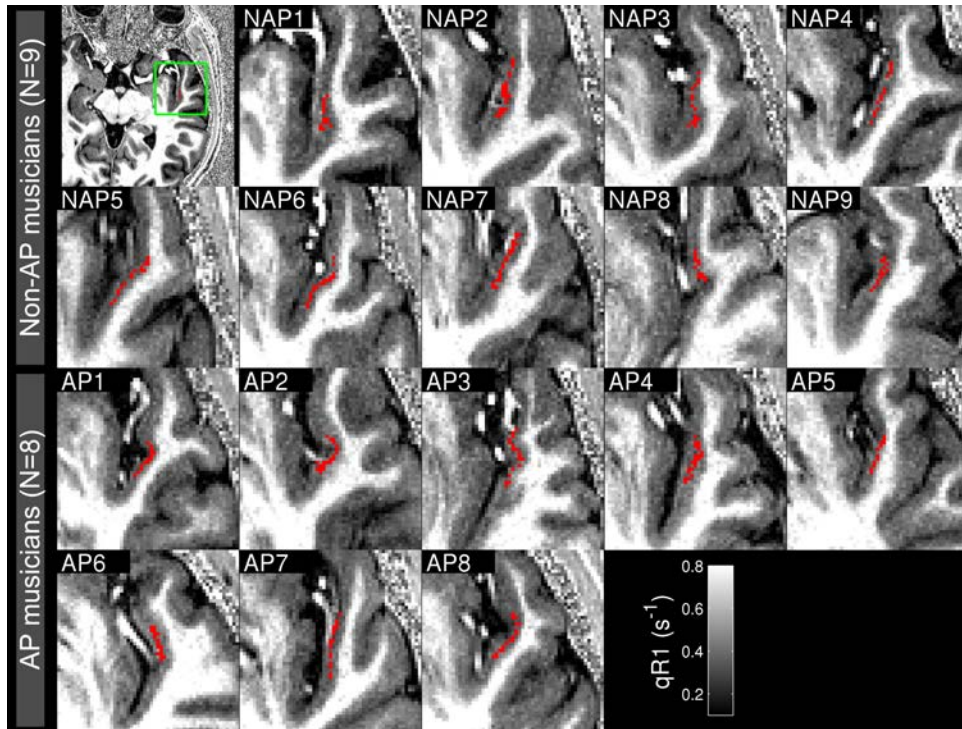
Figure III-5. Axial slices showing significant cluster voxels (red) in native volumes of individuals. Abbreviation: NAP, non-absolute pitch musician; AP, absolute pitch musician; qR1, quantitative longitudinal relaxation rate.

III-3.5. Ruling out Confounds from Ethnicity

III-3.5.1. Replication of AP Effect only with Europeans Subjects

Although ethnicity was included in the previous analysis as a nuisance variable, one may still be concerned about the risk of confound. In order to demonstrate that the AP effect in the right PP cannot be due to the mismatch in ethnicity, we replicated the same analysis with only the European musicians (i.e., 3 AP Europeans vs. 9 non-AP Europeans). Because there was no Asian musician included in this subset, the ethnicity term was discarded as:

$$qR_1 = \beta_0 + age \cdot \beta_1 + sex \cdot \beta_2 + AP \cdot \beta_3 + \varepsilon , \qquad (\text{III-3})$$

Please note that, due to the small number of samples in one group (i.e., 3 APs), the sensitivity of the analysis is reduced. However, if it does yield a significant result, it cannot be confounded by ethnicity. The result is shown in Figure III-6. At the same level in the same location, we reproduced the AP effect (max T(8) = 10.68, p = 0.005, area = 37 mm² at the 50 % projection level, peak MNI-305 coordinate= [45, -9, -11] mm). The effect size of AP at the peak in the right PP was 0.0171 s⁻¹,

which is very close to the effect size of AP estimated with the whole dataset (0.0169 s^{-1}; relative error= 1.18 %). If the current finding of greater myelin in the right PP had been due to confounding effect from the mismatch of ethnicity in experiment design, there should not be such a cluster without the Asian musicians. Therefore we can rule out the possibility of an ethnicity confound in the results presented in Figure III-3.
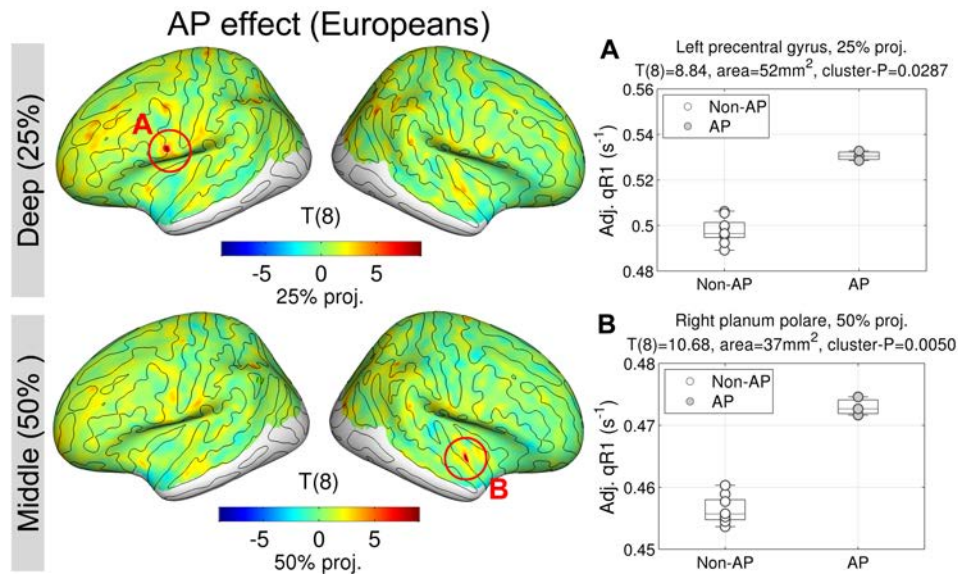
Figure III-6. Effect of absolute pitch (AP) within only European musicians in quantitative longitudinal relaxation rate (qR1). T-statistic maps are shown at the deep (25% of thickness) and the middle (50% of thickness) projection levels from the white matter surfaces. Medial views are omitted because no significant differences were found there. Note that the qR1 values in scatterplots are adjusted for nuisance variables.

An additional cluster in the left precentral gyrus was found (max T(8) = 8.84, p= 0.029, area = 52 mm² at the 25 % projection level, peak MNI-305 coordinate = [-56, 6, 15] mm). However, the estimated effect size (0.0344 s^{-1}) was much larger than that of the whole dataset (0.0118 s^{-1}; relative error= 192.3 %). Thus, the AP effect in the left precentral gyrus might be because of certain characteristics that are specific to the 3 European AP musicians, which were cancelled out by the Asian musicians when the GLM (III-2) was fitted to the full dataset. Hence, any further interpretation of this finding based on the current data is difficult. A dedicated study explicitly targeting the ethnicity effect would be needed here.

III-3.5.2. Simulation for the Estimate Reliability of the Current Between-subject Design

Although the effect of AP was still found only with the European musicians, one might still wonder whether it is tractable to separate the effect of ethnicity from that of AP. The question is, in other

words, how collinearity between the ethnicity and AP affects in estimation of the effect of AP. To compare the current design with an ideal one, a numerical simulation was carried out. The group effect was estimated using the model (III-2).

From the peak vertex at the middle depth (50 % of cortical thickness) in the right planum polare, the estimated coefficients were 0.4437 (interceptor), 0.0004 (age), 0.0040 (sex), -0.0093 (ethnicity), and 0.0169 s^{-1} (AP). The standard deviation of the residuals was 0.0033 and the mean was 10^{-14} s^{-1}. The Kolmogorov-Smirnov test did not reject the null hypothesis that the distribution of residuals is normal (p = 0.845). Therefore, we used the normal distribution to generate noise with the observed mean and standard deviation. Assuming the GLM (III-2) as well as effect size and noise level estimated from the current dataset as 'ground truth', simulated datasets (k = 10,000) were generated for different design matrices varying the number of European AP musicians. Because the number of European AP musicians controls the separation between the ethnicity and AP variables, its change varies the collinearity between ethnicity and AP. The procedure of simulation is as follows:

(1) Change the original design matrix X to make it more or less balanced by changing the ethnicity vector x'_e while leaving other vectors x_{-e} constant:

$$\mathbf{X} \leftarrow \begin{bmatrix} \mathbf{x}_{-e} & \mathbf{x}'_e \end{bmatrix}$$

(2) Generate data with assumed 'true' parameters of the effect size $\mathbf{b} = [\beta_0, \beta_1, \cdots, \beta_n]$, mean μ, and standard deviation σ of a normal distribution N:

$$y \leftarrow \mathbf{X}\mathbf{b} + N(\mu, \sigma^2)$$

(3) Least-square-estimation (LSE) fitting where superscripted T means transposition and + means pseudoinverse:

$$\hat{\mathbf{b}} \leftarrow (\mathbf{X}^T\mathbf{X})^+ + \mathbf{X}^T y$$

(4) Compute and store the relative error of estimation for the i-th regressor of interest:

$$e_{rel} \leftarrow (\hat{\beta}_i - \beta_i)/\beta_i$$

(5) Repeat steps from (1) to (3) for 10,000 times.

(6) Repeat step (1) for all possible design matrices.

Boxplots of relative errors of estimated effect sizes of AP over the correlation between ethnicity and AP are given in Figure III-7. Because the generated noise was unbiased, the median of estimated $\hat{\beta}_i$ was very close to the true β_i in all cases. However, as the design becomes more

imbalanced (i.e., greater correlation), the reliability of estimation decreases (i.e., more distant range between outliers and quantiles), and vice versa.

The result indicates that the reliability of estimating the effect size of AP under the current design (relative error= -0.09 ± 13.16 %; the blue squared boxplot in Figure III-7) is still comparable to a perfectly balanced design (relative error= 0.05 ± 9.41 %, the leftmost boxplot in Figure III-7) and much better than the worst case (relative error= -0.39 ± 21.39 %; the rightmost boxplot in Figure III-7). Additionally, in terms of the minima and maxima of relative errors from 10,000 randomizations, the current design shows reasonable bounds (between -46.16 and 49.06 %) with respect to the ideal one (between -38.01 and 35.51 %) and certainly narrower ones compared to the worst one (between -79.38 and 91.16 %). Taken together with the additional GLM only with European musicians (Section III-3.5.1), we conclude that there is no strong evidence to believe that our result (i.e., the effect of AP) was driven by the ethnicity.
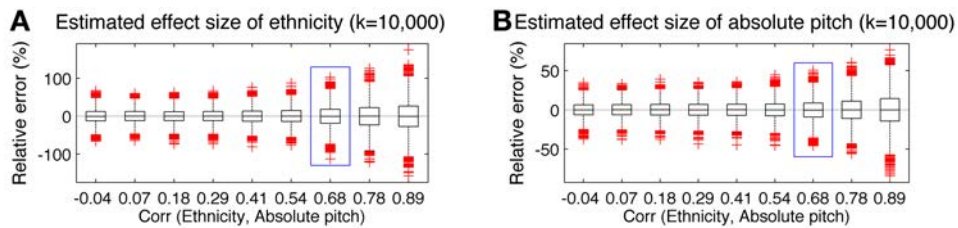
Figure III-7. Relative errors in estimating effect size of absolute pitch over correlation between ethnicity and AP. The blue rectangle indicates the correlation between the two variables under the current design.

III-3.6. Intracortical Myelination and Frequency Discrimination Threshold

To differentiate the observed neuronal correlate of AP from the sensitivity to mere pitch discrimination power (i.e., the relative pitch ability), we tested a GLM as:

$$qR_1 = \beta_0 + age \cdot \beta_1 + sex \cdot \beta_2 + ethnicity \cdot \beta_3 + AP \cdot \beta_4 + (-\log_{10} FDT) \cdot \beta_5 + \varepsilon \qquad \text{(III-4)}$$

where FDT=F1/F0 is the frequency discrimination threshold (see the Section III-2.1), which was not significantly different between AP and non-AP groups.

The results are given in Figure III-8 and Table III-4. An effect of frequency discrimination threshold (FDT) in the lateral region of the right Heschl's gyrus (HG) extending to the lateral superior temporal gyrus (STG) was found across all layers, as shown in Figure III-8 (A), (B), (C). Differences in other regions were found only in the superficial layer (75% of thickness from the white matter surface): i.e., in the opercular part of the right inferior frontal gyrus (Figure III-8 (D)) and in the right lateral superior transverse gyrus of the parietal lobe (Figure III-8 (E)).

Figure III-8. The effect of frequency discrimination threshold (FDT) on quantitative longitudinal relaxation rate (qR1). T-statistic maps are given for each projection level.

Table III-4. Significant clusters for the effect of frequency discrimination threshold

Label	Side	Level*	Structure name	T (11)	p-value+	Area (mm²)	MNI-305 (x, y, z mm)		
A	Right	25 %	Right transverse gyrus of Heschl	6.51	0.0331	46.8	59.9	-2.6	-0.8
B	Right	50 %	Right transverse gyrus of Heschl	6.26	0.0208	61.8	60.0	-1.4	-1.7
C	Right	75 %	Right lateral superior transverse gyrus	5.48	0.0240	37.3	60.1	2.6	-7.9
D	Right	75 %	Right intra-/trans-parietal sulci	6.01	0.0480	37.6	20.6	-59.1	51.2
E	Right	75 %	Right opercular part of the inferior frontal gyrus	5.78	0.0088	70.7	54.5	-5.5	6.4

* Projection level indicates its sampling point from gray matter and white matter interface along an outward normal vector by 25 % (deep), 50 % (middle), or 75 % (superficial) of cortical thickness. + P-values are corrected for multiple comparison correction at cluster level.

In Figure III-9, a closer view of the one remarkable cluster is shown with a scatter plot. Strikingly, the positive correlation of cortical myelin with FDT in the right lateral HG extending into lateral STG (across all layers) was localized in an area of the right STP (Figure III-9, inset, red) that is close to but distinct from that of the effect of AP (Figure III-9, inset, green). This clearly demonstrates the distinctiveness of the AP-related temporal myeloarchitecture (for pitch chroma recognition) and the FDT-related one (for relative comparison of two successive tones).
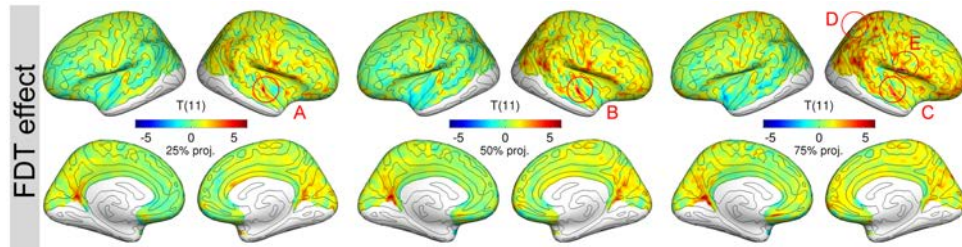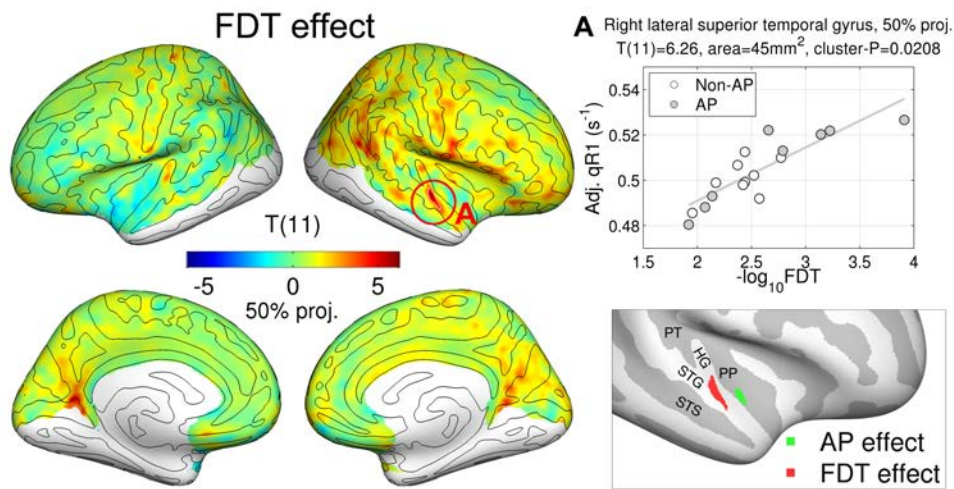
Figure III-9. Effect of frequency discrimination threshold (FDT) in quantitative longitudinal relaxation rate (qR1). T-statistic maps only for the middle level (50% of thickness) are shown. A regression plot at the peak vertex and an inset showing locations of the significant clusters (p< 0.05) for AP effect (green) and FDT effect (red) are shown in the right column. For the FDT effect, the cluster is a union of all clusters across layers.

III-3.7. Subcortical White Matter Myelination

To further examine the found effect of AP, we further probed whether we can localize white matter (WM) myelin difference using qR1 mapping. Because the effect of AP in the cortical myelination in the right PP could either be due to long-range connections from/to the right PP or due to intracortical connections within that region, myelin alteration beneath the right PP would indicate that the altered cortical myelin is actually reflection of long-range WM connections. However, qR1 mapping is not an optimal choice to investigate white matter myelination compared to magnetization transfer imaging or diffusion-weighted imaging. Here, we only explored such a possibility from qR1 data, but it should be further examined using white matter myelin imaging in a future study.

We sampled qR1 values from subcortical white matter at 25, 50, and 75 % of cortical thickness along the inward normal vectors pointing to the white matter. All sampling points were visually inspected (see Figure III-10 for an example) and it was computationally confirmed that coordinates of all sampling points are confined in the white matter mask in a 3-D volume.
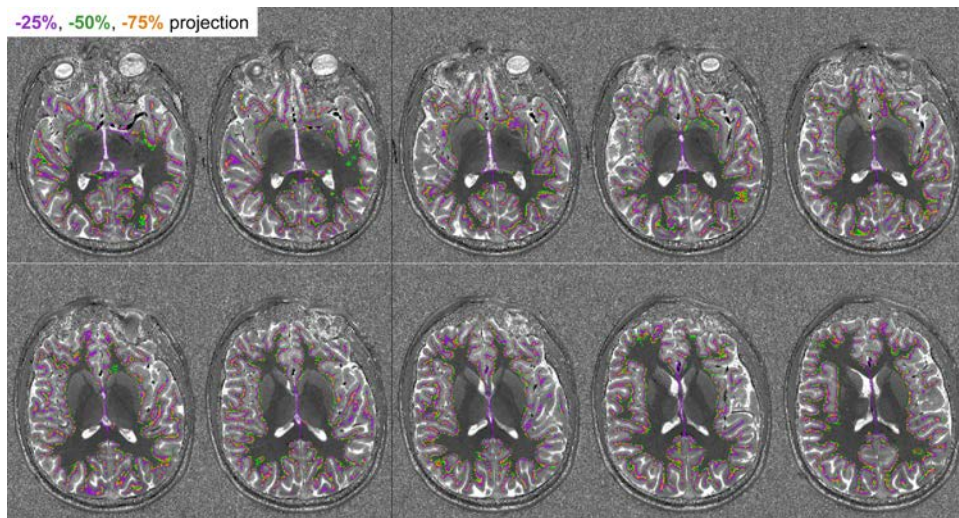
Figure III-10. Subcortical white matter sampling points overlaid on the qT1 (= 1/qR1) image of a single subject. Purple, green, and orange pixels indicate sampling position in white matter which are shifted inward by the 25 %, 50 %, and 75 % of cortical thickness from the gray and white matter interface, respectively. The sampling points are positioned within the white matter, which appears in dark in the qT1 image.

Using the same model (III-3) that was used to test for AP effects in intracortical myelination, we found an AP effect in the right angular gyrus (AG) (max T(12) = 6.29, p = 0.018, area =25 mm² at the -75 % projection level, peak MNI-305 coordinate = [35, -75, 42] mm) as seen in Figure III-11 ('AP effect'). However, we did not find any differences in subcortical WM near the right PP, suggesting that the greater cortical myelination in the AP musicians might reflect horizontal, local connections within the cortex rather than vertical, long-range connections.

In addition, we investigated the FDT effect in qR1 values in the subcortical WM using a GLM (III-4), because the intersubject variability in FDT might be related to variability in myelination of the acoustic radiation (i.e., thalamic projection into primary auditory cortex (PAC). However, we only found a positive correlation between FDT and subcortical myelin in the left orbital sulci (max T(12) = 9.40, p = 0.0002, area = 254 mm² at the -25 % projection; max T(12) = 8.41, p = 0.0158, area = 48 mm² at the -50% projection) as shown in Figure III-11 (lower row), but no effect of FDT in qR1 values beneath the PAC. Significant clusters for the effects of AP and FDT are listed in Table III-5. However, as mentioned earlier, more evidence from other imaging techniques that are better at localizing alteration of WM myelination is needed.
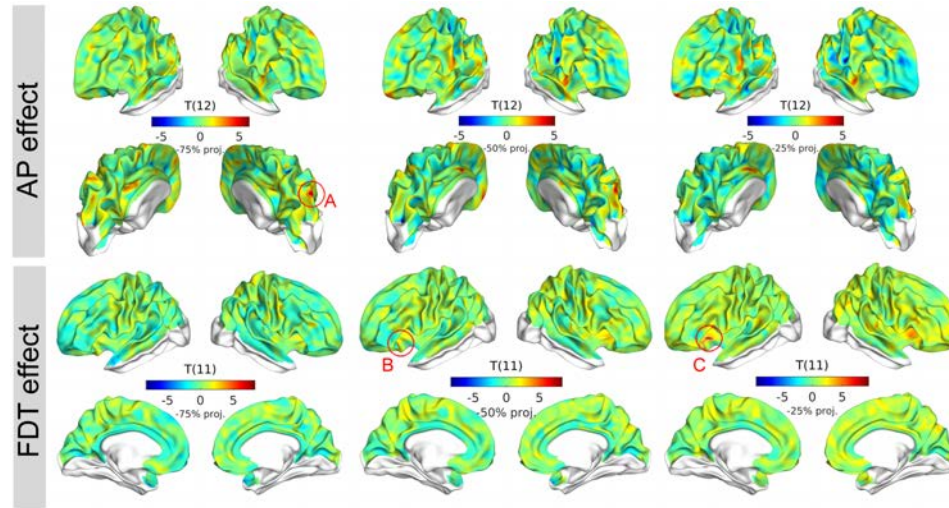
Figure III-11. The effect of absolute pitch (AP) and frequency discrimination threshold (FDT) on quantitative longitudinal relaxation rate (qR1) in subcortical white matter. T-statistic maps are given for each projection level. See Table III-5 for details of significant clusters.

Table III-5. Significant clusters for the effect of absolute pitch (AP) and frequency discrimination threshold (FDT) in subcortical white matter qR1 values.

Label	Side	Level*	Structure name	T (11)	p-value+	Area (mm2)	MNI-305 (x, y, z mm)		
AP effect									
A	Right	-75 %	Right angular gyrus	6.29	0.0323	29.55	34.6	-75.4	42.4
FDT effect									
B	Left	-50 %	Left orbital sulci	8.41	0.0158	47.5	-40.5	31.2	-16.4
C	Left	-25 %	Left orbital sulci	6.01	0.0480	37.6	20.6	-59.1	51.2

* Projection level indicates its sampling point within the subcortical white matter at an inward projection by -75 % (deep), -50 % (middle), or -25 % (superficial) of cortical thickness. + P-values are corrected for multiple comparison correction at cluster level.

III-4. Discussion

Searching for the myeloarchitectonic correlate of AP, we found greater myelination in AP musicians compared to non-AP musicians in the right PP. As hypothesized, the group effect of AP and the correlation effect of APS were located in the STP. In addition, we found a positive correlation with the frequency discrimination threshold in the lateral Heschl's gyrus demonstrating spatially distinct neural processes of absolute recognition of pitch chroma and relative frequency resolution.

While our findings associate the AP ability with the myeloarchitecture of the right anterior STP that is part of the ventral auditory pathway, some previous studies have found heightened leftward specialization in the posterior STP (i.e., the area, volume, GMV as well as the functional activation of the PT) in AP musicians compared to non-AP musicians (Keenan et al., 2001; Luders et al., 2004; Schlaug et al., 1995). In addition, a recent diffusion tensor imaging study reported greater fractional anisotropy (FA) in the left superior longitudinal fascicle (i.e., the dorsal auditory pathway) (Oechslin et al., 2009). Although this might seem surprising at first glance, it is not necessarily a contradiction, because we analyzed a different aspect of anatomy, namely cortical myelination as opposed to cortical thickness (Dohn et al., 2015), area size (Keenan et al., 2001; Schlaug et al., 1995), or diffusion anisotropy in the white matter (Oechslin et al., 2009) in the previous studies. These aspects do not need to be correlated. For example, an increase or decrease of GMV at the border of the cortex would affect the image intensity gradient, thus changing the apparent cortical thickness. However, in the present study, we found an increase in myelination in the middle depth of the cortex, which is less likely to affect the cortical surface reconstruction (Dale et al., 1999). As previously discussed in the Introduction, the notion of leftward asymmetry of AP processing (Keenan et al., 2001; Schlaug et al., 1995) is not uncontended, because recent studies reported AP effects also in the right STP (Bermudez et al., 2009; Dohn et al., 2015; Wengenroth et al., 2014). In particular, a positive correlation between the right HG volume and AP proficiency suggests that AP processing is not strictly left hemispheric (possibly bilaterally in parallel) (Wengenroth et al., 2014). Moreover, because AP mainly involves processing spectral aspects of the auditory input, the idea of hemispheric differences in temporal integration time windows (Poeppel, 2003) and the hemispheric specialization theory (Zatorre and Zarate, 2012) support a prominent role of the right hemisphere. Here, we discuss the biological implications of intracortical myelination, particularly in the context of the role of the anterior STP in pitch chroma processing.

III-4.1. Implication of the Intracortical Myelination for AP

Myelination in the cortex, and alterations thereof, can in principle be due to myelination of intracortical fibers or due to long-range axonal connections that start from or end in the cortical region. In our study, myelination effects were found in the right PP at middle depth (i.e., 50% of cortical thickness) and not in deeper or more superficial cortical layers. Moreover, we did not find a group difference in qR_1 values in the subcortical WM underneath the right PP. Although the subcortical white matter needs to be further investigated using imaging techniques that are more sensitive to white matter microstructures, such as magnetization transfer imaging or diffusion-weighted imaging, the absence of strong differences in the subcortical white matter, along with the restriction of the cortical findings to the middle layer, suggests that the myelination effect in the anterior STP is indeed mainly related to local connections.

Ephaptic crosstalk between two adjacent axons depends on the electric resistance between a node of Ranvier in one axon and the nearest, myelinated part of the other axon (Binczak et al.,

2001). By means of electrical insulation, myelination of local neural circuitries could reduce ephaptic coupling between nearby axons, therefore, increasing specificity of information transfer. Extraction of pitch chroma based on the long-term templates of 12 tones in the Western scale would require precise processing of pitch information, which is already a cortical representation based on acoustic inputs (Griffiths and Hall, 2012). If such neural operations occur in the non-primary auditory cortex, greater local myelination might support processing with heightened specificity. It is known from *post mortem* investigations that non-primary auditory cortex in the human brain has more widespread tangential intracortical connections compared to primary auditory cortex (Tardif and Clarke, 2001). The current myeloarchitectonic findings might be interpreted as an alteration of the tangential connections in the non-primary auditory cortex, which contribute to the computation of pitch chroma.

Moreover, intracortical myelination can restrict experience-driven neuroplasticity. In a mouse model (McGee et al., 2005), an abrupt increase of gray matter myelination was found in normal mice after the critical period, which precluded further reorganization of ocular dominance. The importance of cortical myelin in suppression of neuroplasticity in adult mice was also demonstrated by the observation that focal demyelination enabled recovery from amblyopia (Bavelier et al., 2010). A recent intriguing human study found that administration of volproic acid, which is known to cause hypomyelination of the corpus callosum in developing mice (Shen et al., 2005), can enable human adults to acquire an AP-like long-term memory for a limited number of tones (Gervain et al., 2013). Given that the acquisition of AP is associated with an early onset of musical training/experience, typically between age 4 and 7 years (Baharloo et al., 1998; Miyazaki and Ogawa, 2006), the greater myelination in the AP musicians may imply that AP requires a prohibitory mechanism of neuroplasticity, which preserves the spectro-temporal templates obtained during the critical period in children with predispositions for AP (Baharloo et al., 2000).

III-4.2. Functional Role of the Anterior STP in AP

We found heavier cortical myelination in the anterior STP in the AP musicians. Specifically, the myelin difference was localized on the medial part of the PP, which corresponds to anterior auditory (AA) region as defined based on cyto-/receptor architecture (Rivier and Clarke, 1997). The AA region, which is a part of the 'belt' region, was characterized by strong (but weaker than primary auditory area; AI) densities of soma and fibers of pyramidal cells in layer III (Rivier and Clarke, 1997). A tracing study found that the intrinsic connectivity of area TG (in PP) is mainly of longer range (greater than 4 mm and up to 7 mm) compared to that of HG (less than 2.5 mm) and anisotropic: the connections spread more posteriorly towards HG (Tardif and Clarke, 2001). Based on the intermediate cytoarchitectonic properties (i.e., cell density greater than lateral STG but smaller than AI) AA has been considered as an early non-primary auditory cortex (Rivier and Clarke, 1997) and also a part of the ventral pathway based on neuroimaging studies that

demonstrated selectivity of AA to sound recognition rather than sound localization (Clarke and Morosan, 2012).

The functional and structural dissociation of the spatial and non-spatial information, being separately processed in parallel in non-human primates (Kusmierek and Rauschecker, 2009; Rauschecker, 2015; Tian et al., 2001) and humans (Arnott et al., 2004b; Warren and Griffiths, 2003a), has been theorized as 'dual auditory pathways'. From a number of human neuroimaging studies (Altmann et al., 2007; Arnott et al., 2004b; Barrett and Hall, 2006; Hart et al., 2004), the spatial information was found to be processed in the dorsal auditory pathway whereas the non-spatial information was found to be in the ventral auditory pathway. Particularly, the ventral auditory pathway, including the right anterior region of STP and superior temporal sulcus (STS), has been believed to play a distinctive role in identifying auditory object properties (Griffiths and Warren, 2004), such as voice identity (Belin et al., 2000) and acoustic properties of vocal tracts (von Kriegstein et al., 2006). This further demonstrated that the right anterior STS is sensitive to the change of the sound-source identity using environmental sounds, of which distinctiveness was parametrically manipulated.

The differential processing of pitch chroma and pitch height has often been suggested based on the occurrence of octave errors in AP musicians (Deutsch, 2013; Deutsch and Henthorn, 2004; Miyazaki, 1988; Takeuchi and Hulse, 1993). The dissociated recognition ability of AP musicians on pitch chroma and pitch height implies that the AP musicians are especially sensitive to the pitch chroma, not to the pitch height. A neuroimaging study that independently manipulated pitch chroma and pitch height revealed that the anterior STP is more sensitive to change in pitch chroma than the posterior STP (Warren et al. 2003), supporting the notion that recurring representation of pitch chroma is handled in a pathway that is distinct from the one that processes the pitch height. Thus, we speculate that the AP-related myelination increase in the anterior STP found in our data might be related to the parallel processing of pitch chroma and pitch height. More specifically, for AP musicians, pitch chroma may be an identifiable property of an auditory object, while pitch height is processed as a relative position on a linear continuum, similarly to the non-AP population. As discussed in the previous section, local myelination may increase specificity of the tone representations and hinder neuroplasticity (Bavelier et al., 2010), thus enabling precise matching between the spectro-temporal images (Griffiths and Warren, 2004) of online inputs and of the templates in long-term memory.

Localization of the long-term memory traces of pitch chroma in AP musicians remains unclear. Dohn et al. (2015) suggested an involvement of the right hippocampus based on their finding of a correlation between the FA of the white matter in the ventral pathway and cortical thickness in the parahippocampal gyrus. However, it was shown that the hippocampus is selectively involved in the retrieval of recollection-like context-based episodic memory, but not in familiarity-based recognition (Eldridge et al., 2000; Fortin et al., 2004). Strikingly, two very rare cases of epileptic

patients with AP (Suriadi et al., 2015; Zatorre, 1989) demonstrated that the musical abilities and AP recognition remained intact after temporal lobectomy. In one case (Zatorre, 1989), the patient underwent an anterior temporal lobectomy in the left hemisphere that removed the amygdala and a sizable portion of hippocampal structures. In another case (Suriadi et al., 2015), the patient underwent a selective amygdalohippocampectomy in the right hemisphere to minimize damages in other temporal structures. The unimpaired AP listening after the operation implies that the hippocampus might not be necessary in retaining AP.

We suggest that the right anterior STP plays a key role in the pitch chroma extraction process, which corresponds to the early perceptual component of AP and should be a unique process only in the AP musicians but not in the non-AP musicians. For further clarification, functional measurements (e.g., through fMRI) might be useful. If there exists a separate cortical module in the right PP that extracts pitch chroma, the region would be selectively sensitive to pitch chroma but not to pitch height. In other words, when an AP musician hears two tones an octave apart (e.g., C4 and C5), in the primary auditory cortex the activation peaks for the two tones should be distant along the tonotopic gradient (Moerel et al., 2012), whereas in the right PP the two peaks should be at the same position because of the identical pitch chromas of the two tones. Alternatively, instead of a geometrical space, one could measure a high dimensional distance in a functional space as commonly utilized in multivariate pattern analysis methods such as a searchlight method (Haxby, 2012). Under the same assumption, the functional distance amongst the local multivariate patterns in the right PP responding to the two tones with the same pitch chroma is expected to be closer than that in the primary auditory cortex. Such analyses will help to localize the AP template in the future.

III-4.3. Intracortical Myelination and FDT

Further to some recent studies reporting in-vivo measurements of myelin concentration and tonotopic organization in human auditory cortex (De Martino et al., 2014; Dick et al., 2012), we found a correlation between in-vivo mapping of intracortical myelination in the right anterolateral HG extending to the lateral STG, which corresponds to Te3 in non-primary auditory cortex (Morosan et al., 2001), and the frequency resolution in auditory perception of musicians for the first time. In the following paragraphs we briefly discuss how this finding relates to previous research.

The importance of the right lateral HG in pitch perception has been established in a number of functional and lesion studies in human and non-human primate subjects (Bendor and Wang, 2005; Norman-Haignere et al., 2013; Patterson et al., 2002; Penagos et al., 2004). In marmoset monkeys, pitch-selective neurons were found near the anterolateral border of the primary auditory cortex (Bendor and Wang, 2005). In humans, pitch-sensitive regions were also localized in the anterolateral part of the primary auditory cortex (Norman-Haignere et al., 2013; Patterson et al.,

2002). Another human fMRI study, in which pitch salience was manipulated while controlling for temporal regularity of auditory stimuli, found a correlate of pitch salience at the anterolateral end of HG (Penagos et al., 2004). Moreover, microelectrode recordings from bilateral HG in epileptic patients showed that responses of single neurons are selectively tuned with a very fine frequency resolution (frequency ratio of about 3%), which corresponds to the just-noticeable difference in the untrained normal population (Bitterman et al., 2008).

In addition, greater involvement of the right hemisphere, compared to the left hemisphere, in pitch perception (Zatorre, 1998) has been further demonstrated in human studies using magnetoencephalography (MEG) (Patel and Balaban, 2001; Schneider et al., 2002), positron emission tomography (PET) (Zatorre, 1998), and fMRI (Zatorre and Belin, 2001). Even more strikingly, human lesion studies showed that resection of the right HG greatly worsened the pitch discrimination ability, whereas it remained intact when only the left HG was removed (Johnsrude et al., 2000; Russell and Golfinos, 2003). More recently, Coffey and colleagues demonstrated that the frequency-locked auditory response (estimated from MEG source reconstruction) in the right auditory cortex was inversely correlated with a frequency discrimination threshold across individuals whereas such a relationship was not found in the left auditory cortex (Coffey et al., 2016a). There are also some morphological findings supporting the role of the right HG in pitch discrimination. Schneider and colleagues reported increased grey matter volume related to pitch discrimination performance (Schneider et al., 2002), albeit at a much coarser frequency resolution (>1 semitone) compared to the present study. In this context, our finding of a correlation between the cortical myelination in the right lateral HG and FDT supports the crucial role of the right lateral HG in pitch extraction and may provide an insight into the neural implementation of this process.

The separate locations of pitch chroma identification and fine pitch frequency discrimination in the right auditory cortex, which is specialized to spectral as opposed to temporal processing according to some theories (Poeppel, 2003), in terms of local cortical myelination (which might reflect local processing capabilities) highlights the distinctness of the underlying mechanisms of both faculties.

III-4.4. Limitation

In the current dataset, there was a mismatch between the AP and non-AP groups in terms of ethnicity (more Asian musicians than European musicians in the AP group), which was rooted in recruiting difficulties. However, we did demonstrate that the current findings with respect to AP could not be due to confounding influence of ethnicity.

The question remains as to whether there are some specific differences in cortical myelination that could be attributed to the ethnic background of the subjects. Although up to now no study has reported ethnic differences in cortical myelination, differences in cortical thickness and GM local

volume were found in various areas of the cortex; frontal, parietal, median temporal, and polymodal association cortices (Chee et al., 2011). However, as discussed earlier, cortical thickness and myelination within the cortex may not be necessarily related. Although we did find effect of ethnicity in our current data, a large-scale study is needed to elucidate the effects of ethnicity and any possible interaction with AP.

III-4.5. Conclusion

We found greater intracortical myelination in an area within the right PP in AP musicians compared to non-AP musicians, which was spatially distinct from another area in the anterolateral HG that correlated with FDT. We argue that heavier myelination of local wiring may be beneficial to the preservation of precise representations of pitch chroma after the acquisition of AP. Because it is known that the identity of an auditory object is processed along the ventral auditory pathway including the anterior STP, we suggest that the pitch chroma may be processed as an identifiable object property in AP musicians. Our findings indicate that the extraction of pitch chroma might occur in the anterior STP; however, further studies based on structural and functional brain imaging are needed to clarify whether and to what extent that region is merely a part of a wider network. Moreover, the localization of distinct cortical myelination related to pitch chroma identification and frequency discrimination in the right hemisphere is in favor of theories that postulate specialization of the left and right hemispheres in temporal and spectral processing, respectively.

IV. Study #2: Intrinsic Functional Connectivity of the Right Planum Polare and Absolute Pitch

In this chapter, I will present a functional connectivity study on the resting-state fMRI. The main findings of this chapter are now in the review process for publication in a peer-reviewed journal. In Chapter III, we found a positive effect of the level of absolute pitch (AP) in cortical myelination in the right planum polare (PP). Based on this finding, we further revealed greater functional connectivity between the right PP and bilateral auditory cortices, medial frontal structures, and the left ventrolateral prefrontal cortex. I will discuss the implication of the findings of functional connectivity in the context of the 'dual auditory pathway' hypothesis.

IV-1. Motivation and Hypotheses

In the Chapter III, the implication of the greater cortical myelin in the right PP in musicians with AP was discussed (Kim and Knösche, 2016). Based on the dual-pathway framework (as earlier mentioned in the Section I-5.1), the anterior part of the supratemporal plane (STP) is known to be sensitive to non-spatial auditory information such as frequency modulation (Hart et al., 2004) and pitch change (Altmann et al., 2007; Barrett and Hall, 2006). In particular, Warren and colleagues suggested that the ventral pathway is involved in invariant properties of an auditory object (i.e., irrelevant to the spatial location of the object), whereas the dorsal pathway initiates auditory object/scene segregation (Warren et al., 2003a).

As discussed in the Section I-4.1, the pitch chroma constitutes discrete (i.e., categorizable) non-spatial information of an auditory object. Non-AP listeners perceive pitch chroma only based on tonal context (thus tonal function of the pitch) or an external reference (similarity perception of octave-spaced tones). Thus in the Chapter III, we postulated that the involvement of the right ventral pathway, which encompasses the right PP, in AP perception could be crucial (Kim and Knösche, 2016).

From a line of behavioral evidence (see Section I-4.1), it has been suggested that AP perception is fairly spontaneous and difficult to inhibit. For instance, AP listeners unintentionally recognize the pitch chroma of glasses clinking (Miyazaki, 2004a) and show a Stroop-like effect from incongruent solfeggio (Itoh et al., 2005; Schulze et al., 2013). This suggests that the AP process is dominantly active regardless of efforts and intentions, implying an AP-related functional network, if any, might be persistent even during rest (resting state network, RSN). Considering this, we further tested our conjecture on the importance of the right PP in AP perception using resting-state functional MRI (rs-fMRI), which has been extensively used to investigate the relationship amongst neural activities in human brains over the decades (Smith et al., 2013).

We hypothesized that if the increased local connectivity in the right PP (as indexed by increased cortical myelination) is relevant for AP, then this region should exhibit greater RSFC with (1) adjacent auditory cortices, (2) homologous auditory regions in the left hemisphere, and (3) the left hemisphere frontotemporal network. The rationale of our hypotheses is as follows.

(1) Given that cortical myelin would provide additional electric insulation and decrease local transmission time, which may increase efficiency and specificity of information transfer, neural communication via intracortical connections would benefit from the myelination of local circuits. Therefore, we expect greater short-range RSFC of the right PP with neighboring auditory cortex (AC) areas in individuals with more accurate AP perception. A similar link between local morphology (local gray matter volume) and RSFC was recently found in the left STG in musicians compared to non-musicians (Fauvel et al., 2014).

(2) Studies using independent component analysis (ICA) based methods have consistently shown that the auditory RSN bears high modularity within bilateral STPs, including HG, HS, PP, and PT, both in general populations (Damoiseaux et al., 2006; Shirer et al., 2012; Smith et al., 2013; Smith et al., 2009; Smith et al., 2012) and in musicians (Fauvel et al., 2014; Luo et al., 2012). Additionally, in a histological study with rhesus monkeys (Pandya et al., 1969), the PAC and the posterior part of the STP were found to be structurally connected to their contralateral homologues via the corpus callosum, whereas the anterior parts of the STPs were connected via the anterior commissure. This line of evidence suggests high functional connectivity (FC) between the homologous auditory regions. As mentioned above, the ventral auditory pathway is likely to process pitch chroma (i.e., non-spatial information) processing (Warren et al., 2003a). Therefore, it can be expected that the left PP would also be involved in non-spatial information processing that the right PP undertakes in an individual with proficient AP. This could result in increased FC between two homologous auditory cortices, in addition to the common communication between bilateral auditory cortices in non-AP musicians.

(3) As largely accepted, if the perceptual and associative components of AP perception are localized in the temporal and frontal structures, respectively, integration of the two parts is crucial for the successful realization of AP. In particular, the ventral auditory pathway from PAC, via PP, to the ventrolateral prefrontal cortex (VLPFC) is expected to show greater RSFC in relation to the acuity of AP. From a previous study (Wengenroth et al., 2014), the left inferior frontal gyrus (known as Broca's area) showed greater BOLD activation in passive response to musical tones. Considering the dominancy of the left hemisphere in language functions, the association of the pitch chroma representation with the verbal label is likely to be processed in the left hemisphere.

Importantly, a previous rs-fMRI study reported greater overall degree centrality in AP than in non-AP musicians (Loui et al., 2012), suggesting a global reorganization of brain networks. In contrast, in this study, we tested specific hypotheses on the role of particular anatomical regions

that are expected to be involved in the AP specific RSN, based on the results describe in the Section III-3.4 (Kim and Knösche, 2016).

IV-2. Materials and Methods

IV-2.1. Participants and Behavioral Tests

We analyzed resting-state fMRI data from 17 musicians (8 AP musicians and 9 non-AP musicians) who also participated in the previous study (Kim and Knösche, 2016) described in the Section III-2.1. The musicians were matched for musical aptitude (Wallentin et al., 2010), age, gender, and handedness. Absolute pitch performance of all musicians were behaviorally assessed and summarized in a scaled index (between 0 and 1 with a chance level of 0.5).

IV-2.2. Image Acquisition

Echo-planar imaging (EPI), magnetization-prepared pulses and rapid gradient-echo (MPRAGE) were acquired using a 3-T MR system Magnetom Prisma (Siemens, Erlangen, Germany) at 2.3-mm isotropic resolution. While the participant lay still with eyes open and was instructed not to fall asleep for 10 minutes, 420 volumes of EPIs were obtained. T1-weighted images using MPRAGE were also taken at 1-mm isotropic resolution during the same session. Further details about parameters of MR sequences can be found in the Section II-2.3.

IV-2.3. Image Processing

Preprocessing of EPI images was done using SPM12[16]. Images were first corrected for the multiband slice timing, unwarped and realigned, and resampled at 2.3-mm isotropic resolution using a 4th degree B-spline. In order to minimize spurious correlation due to head movements (Power et al., 2012), we adopted the 'anatomical CompCor' approach. The main idea of the approach is that the signal fluctuation from the white matter (WM) and cerebrospinal fluid (CSF) voxels would not be due to neuronal activities (Behzadi et al., 2007). Further details about CompCor approach can be found in the Section II-3.2.2. Finally, we used 6 rigid-motion parameters (3 translations and 3 rotations with respect to the first volume) and frame-wise displacement to regress out motion artifacts. Then the residual time series were temporally filtered by a band-pass-filter between 9 and 80 mHz (Satterthwaite et al., 2013) for the FC measure of cross-correlation.

Cortical surfaces were created using FreeSurfer from T1w image and the residual timeseries was registered using a boundary-based registration (BBR) approach (Greve and Fischl, 2009) at the

---

[16] http://www.fil.ion.ucl.ac.uk/spm/software/spm12/

middle cortical depth for high signal-to-noise ratio (Yeo et al., 2011). Details in spatial processing can be found in Section III-3.3. The surface-projected residual time series were non-linearly registered onto a low-resolution template mesh called 'fsaverage5' in FreeSurfer and spatially smoothed using a 2-D Gaussian kernel with the full-width at half-maximum (FWHM) of 10 mm. For the given seed mask in the right PP (which is the significant cluster for the AP effect in cortical myelin from the previous study (Chapter III), the first principal component (PC1) was extracted. FC measures of cross-correlation and cross-coherence were calculated from the PC1 time series against the whole cortex.

IV-2.4. Functional Connectivity Analysis

We used two different FC measures for analysis: cross-correlation and cross-coherence. Cross-correlation of two time series i and j at time lag h is given by:

$$\rho_{ij}(h) = \frac{\text{cov}_{ij}(t,\ t+h)}{\sqrt{\text{var}_i(t)\ \text{var}_j(t+h)}},$$

(IV-1)

where $\rho_{ij}(h) = \rho_{ji}(-h)$, which is restricted to the interval [-1, 1]. If the time lag h equals to zero, $\rho_{ij}(h)$ corresponds to the zero-lag cross-correlation, which equals Pearson's coefficient of correlation.

Cross-coherence of two time series i and j at frequency λ is given by:

$$\text{coh}_{ij}(\lambda) = |R_{ij}(\lambda)|^2 = \frac{|f_{ij}(\lambda)|^2}{f_{ii}(\lambda)\ f_{jj}(\lambda)},$$

(IV-2)

where $R_{ij}(\lambda)$ is the complex valued coherency, $f_{ii}(\lambda)$ is the spectral density function at frequency λ, and $f_{ij}(\lambda)$ is the cross-spectral density function. As squared, coherence is a positive (bounded in [0, 1]) and symmetric function (i.e., ($\text{coh}_{ij}(\lambda) = \text{coh}_{ij}(-\lambda)$).

Zero-lag cross-correlation quantifies the linear relationship between two time series without time lag. Due to its simplicity and clarity, most of the resting state fMRI (rs-fMRI) literature used this measure for FC. Conversely, coherence can capture the frequency-dependent FC, regardless of the time lag between BOLD time series because we took the magnitude of coherency.

Frequency dependency of BOLD fluctuation at rest has been reliably demonstrated across many studies (Achard and Bullmore, 2007; Achard et al., 2006; Cordes et al., 2001; De Luca et al., 2006; Qian et al., 2015; Salvador et al., 2005a; Salvador et al., 2005b; Sasai et al., 2014; Sun et al., 2004). It should be noted that the temporal scale in rs-fMRI data is much slower (2–3 orders of magnitude) compared to that in electrophysiological data. The normal EEG frequency range is between 1 and

100 Hz, while a common filtering band range for rs-fMRI is between 0.01 and 0.1 Hz. The difference in temporal scale is due to the hemodynamic nature of the BOLD signal. Nonetheless, the slow RSFC in rs-fMRI data is in agreement with RSFC in electrophysiology data. Using simultaneous recording of EEG and fMRI (Britz et al., 2010), EEG-microstates (quasi-stationary for every 100 ms), convoluted by hemodynamic response function (HRF), were spatially correlated with the independent components (ICs) from rs-fMRI data (dominant for about 10 s). Britz and colleagues suggested that the neural activities underlying the BOLD fluctuation could be faster (as EEG-microstates) and scale invariant (Britz et al., 2010). More directly, simultaneous recording of local field potential (LFP) and multi-unit activity (MUA) in the primary auditory cortex of monkeys showed high correlation with BOLD fluctuation with a peak in a very low frequency range (0.033–0.1 Hz) (Shmuel and Leopold, 2008).

Studies using graph theory demonstrated that BOLD signals from human brains at rest exhibit meaningful topological properties such as small-worldness or rich-club coefficients at very slow frequencies: e.g., 0–15 mHz (Qian et al., 2015), 10–50 mHz (De Luca et al., 2006), 10–60 mHz (Achard et al., 2006), and 60–120 mHz. A recent study combining fMRI and fNIRS (i.e., functional near infrared spectroscopy) also reported topological features that are commonly found in human brains in low frequency bands (10–30 mHz and 70–90 mHz) (Sasai et al., 2014). Moreover, partial coherence between distant regions, especially interhemispheric homologous regions was found to be substantial in slow frequency bands of 0.4–151.8 mHz (Salvador et al., 2005b) and 0–150 mHz (Sun et al., 2004). Taken together, the slow (< 100 mHz) BOLD fluctuation during rest (without external stimuli) should be considered as being due to complex neurovascular coupling and reflecting underlying neuronal activities (Logothetis et al., 2001).

Despite the general behavioral relevance of infra-slow frequency BOLD activities in humans (Britz et al., 2010) and monkeys (Shmuel and Leopold, 2008), specific relevance of particular frequency bands to cognitive functions and behaviors is yet to be established. Therefore, we segmented the frequency of interest (0–100 mHz) into small frequency bins (bandwidth of 2 mHz; 50 bins).

IV-2.5. Statistical Inference

We tested the effect of APS as:

$$FC = \beta_0 + motion\beta_1 + ethnicity\beta_2 + APS\beta_3 + \varepsilon ,$$ (IV-3)

where *FC* is either cross-correlation or cross-coherence, *motion* is maximal frame-wise displacement of the head (Power et al., 2015), ethnicity is a binary variable for being either Asian or European, APS is a behavioral measure of the acuity of AP, and ε is the Gaussian error variable. Note that, instead of testing for the AP group index (denoting whether a participant belongs to the

AP or the non-AP group), we used the APS index, which is a more direct representation of the behavioral information (essentially, AP group affiliation is only a binarization of the APS).

Statistical inference on the correlation maps with multiple comparisons correction based on the random field theory (RFT) was done using SurfStat (Worsley et al., 2009a) in the MATLAB environment (version 8.2; Mathworks, Inc., Natick, MA, USA). As additional Bonferroni correction for zero-lag cross-correlation, RFT-corrected cluster-level p-values in each hemisphere were further multiplied by 4 for the number of tests (2 contrasts by 2 hemispheres). For cross-coherence, coherence maps in 2 mHz bins were highly similar between adjacent bins (correlation over the spatial domain between 0.62 and 0.98 with a mean of 0.89 ± 0.05). Such dependency in the frequency domain suggests that using Bonferroni correction for the number of bins (i.e., 50) would be an overly stringent correction. For M/EEG time-frequency analysis, the use of RFT-based correction for 4-D M/EEG data (2D-sensor-grid by frequency by time) was previously suggested (Kilner and Friston, 2010). However RFT correction for triangular-mesh-mapped high dimensional data such as time-frequency representation of fMRI time series over multiple cortical layers (i.e., 5-D) is yet to be implemented. Instead, we used singular value decomposition (SVD) to approximate the number of independent tests in the frequency domain in analogy to the resel, which is the number of independent tests in the spatial domain. In order to explain more than 99% of the variance for all individuals, 19 independent components were required. Thus, we applied Bonferroni correction with the factor of 76 (2 contrasts by 2 hemispheres by 19 independent components).

IV-3. Results

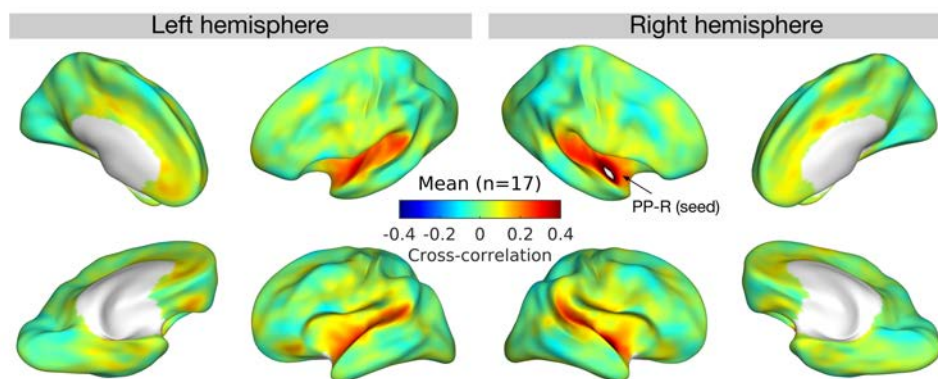### IV-3.1. Seed-based Functional Connectivity Analysis

Figure IV-1. Averaged zero-lag cross-correlation. Zero-lag cross-correlations between the right planum polare (PP-R; the white patch marked by the arrow) and the rest of the

cortex are averaged across all individuals (n=17). Correlations are projected onto minimally inflated cortical surfaces, which show major landmarks such as central sulcus, temporal/parietal convexity, frontal operculum, and middle frontal gyrus. Four oblique views, i.e., (medial/lateral) × (top/bottom), are shown for the left and right hemispheres.

From the right PP, strong correlation with the bilateral STPs was observed across all musicians (n = 17) as shown in Figure IV-1. We found significant effects of APS (Figure IV-2) in cortical regions that are adjacent to the right PP: the right lateral STG (LSTG-R) and the right inferior segment of the circular sulcus of the insula (ICSI-R). Moreover, we also found APS effects in frontal regions: the left triangular gyrus (TIFG-L) and the right medial orbitofrontal cortex (MOFC-R). Detailed statistics are given in Table IV-1.
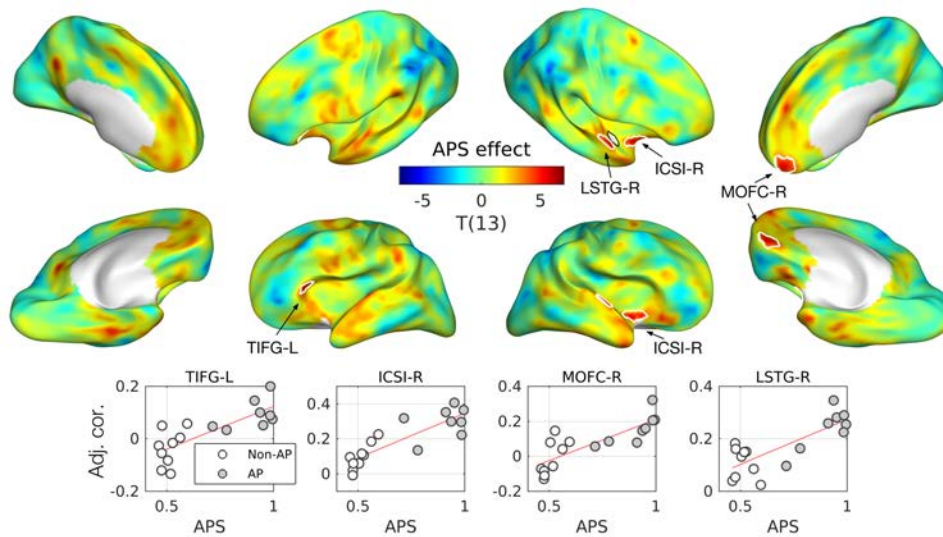

Figure IV-2. Effects of the absolute pitch score (APS) in zero-lag cross-correlation. T-statistic maps are projected onto semi-inflated cortical surfaces. Significant clusters are indicated by white contours. Below the cortical surface maps, scatterplots with regression lines are given for the peak of each cluster with an open circle for non-AP and a gray circle for AP musicians. Abbreviation: TIFG, triangular part of the inferior frontal gyrus; ICSI, inferior segment of the circular sulcus of the insula; MOFC, medial orbitofrontal cortex; LSTG, lateral superior temporal gyrus.

Table IV-1. Effects of absolute pitch score (APS) in zero-lag correlation.

Label	Effect size	Max T(13)	P-value*	Area (mm²)	MNI305 (mm)	Full name of the structure[†]
TIFG-L	0.66	5.38	0.0364	80.8	-42, 29, -2	Left triangular part of the inferior frontal gyrus
ICSI-R	0.88	5.83	0.0080	90.6	41, 3, -15	Right inferior segment of the circular sulcus of the insula
MOFC-R	0.91	5.60	0.0171	136.3	5, 54, -12	Right medial orbitofrontal cortex
LSTG-R	0.84	6.51	0.0241	39.0	62, 4, -5	Right lateral aspect of the superior temporal gyrus

*P-values are corrected for multiple comparisons. † Identification of anatomical nomenclature is generally based on Destrieux Atlas (a2009s) and Desikan-Killiany-Tourville Atlas (DKTatlas40) in FreeSurfer.

For cross-coherence, averaged coherence maps across all musicians are given in Figure IV-3. In contrast to the zero-lag cross-correlation, the coherence maps were bilaterally distributed, in particular between the STPs. The spatial range of the coherence was more extensive in the lowest frequency bin (0-2 mHz), especially between anterior regions (including the left PP in Figure IV-3 (A)) of the STPs, compared to higher frequency bins (24–26, 48–50 mHz).
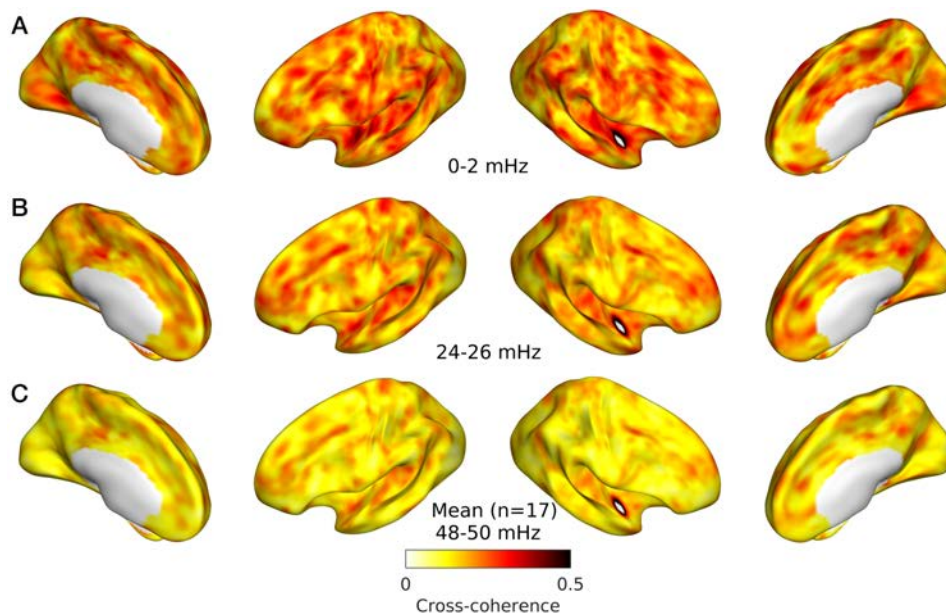
Figure IV-3. Averaged cross-coherence. Cross-coherence between the right planum polare (PP; marked by a white patch) and the rest of the cortex is averaged across all individuals (n=17) for three example frequency bins: (A) 0-2 mHz, (B) 24-26 mHz, and (C) 48-50 mHz.

We found significant effects of APS in cross-coherence in the left PP extending towards LSTG (PP/LSTG-L) as well as in medial structures such as the left anterior cingulate cortex (ACC-L) and the medial aspect of the right superior frontal gyrus (SFG-R) (Figure IV-4). Detailed statistics are listed in Table IV-2.
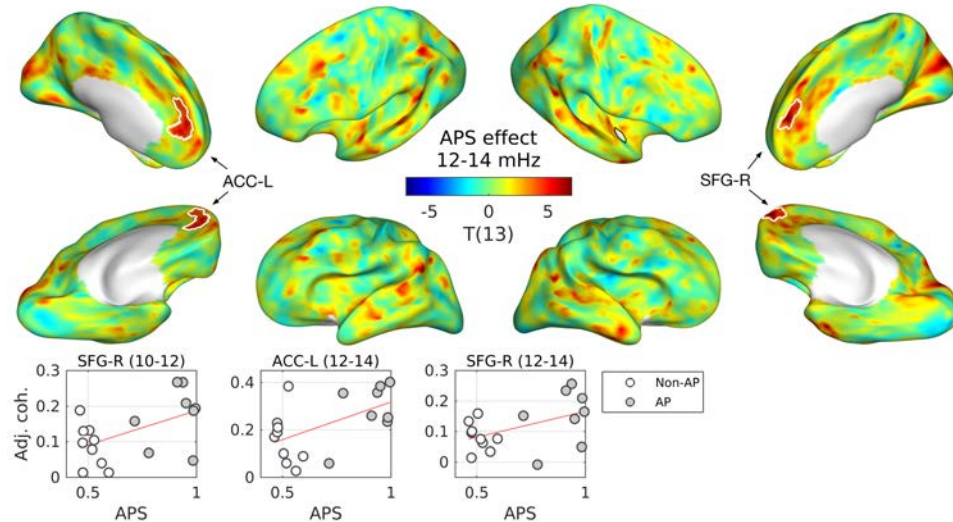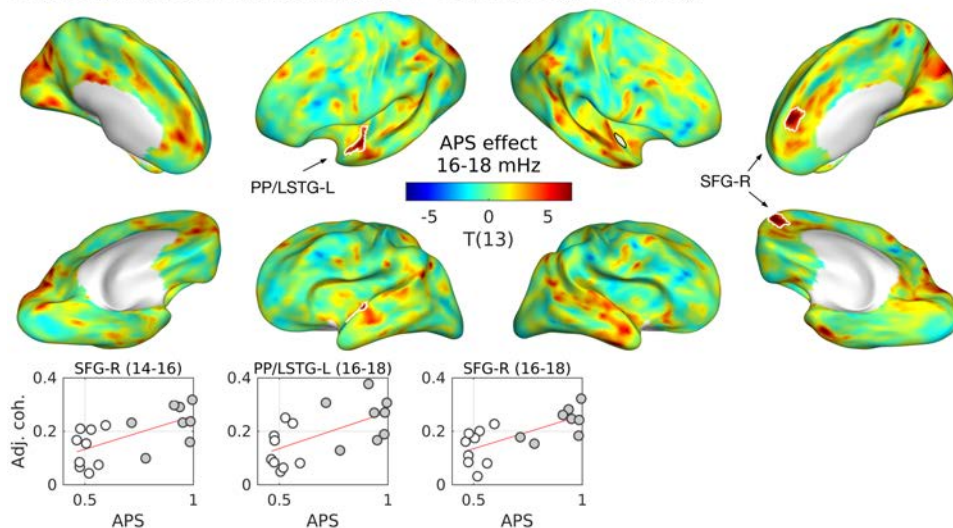
Figure IV-4. Effects of the absolute pitch score (APS) in cross-coherence. Significant effects were found for medium low frequencies (10–18 mHz) only. T-statistic maps for two frequency bands are given: (A) 12–14 (B) 16–18 mHz, while 10–12 and 14–16 mHz are omitted for their similarity to the adjacent frequency bands. Significant clusters are indicated by white contours. Scatterplots with regression lines are given for the peak of each cluster with an open circle for non-AP and a gray circle for AP musicians. The coherence coefficients in the scatterplots were adjusted for the nuisance variables (i.e., Adj. coh., adjusted cross-coherence). Abbreviation: SFG, superior frontal gyrus; ACC, anterior cingulate cortex; PP, planum polare; LSTG, lateral aspect of the superior temporal gyrus.

Table IV-2. Effects of absolute pitch score (APS) in cross-coherence.

Label	Freq. (mHz)	Effect size	Max T(13)	P-value*	Area (mm²)	MNI-305 (mm)	Full name of structure[†]
SFG-R	10-12	1.28	8.24	0.0412	163.2	5, 53, 19	Right superior frontal gyrus
ACC-L	12-14	1.38	5.93	0.0279	285.8	-7, 48, 11	Left anterior part of the cingulate gyrus and sulcus
SFG-R	12-14	1.25	8.41	0.0430	166.6	6, 50, 20	Right superior frontal gyrus
SFG-R	14-16	1.36	9.22	0.0389	167.2	7, 52, 22	Right superior frontal gyrus
PP/LSTG-L	16-18	1.36	7.79	0.0131	147.4	-54, 3, -4	Left lateral aspect of the superior temporal gyrus
SFG-R	16-18	1.26	10.36	0.0333	163.0	7, 52, 22	Right superior frontal gyrus

*P-values are corrected for multiple comparisons. † Identification of anatomical nomenclature is generally based on the Destrieux Atlas (a2009s) and the Desikan-Killiany-Tourville Atlas (DKTatlas40) in FreeSurfer.

IV-3.2. Lagged Cross-correlation in Regions-of-interest

We further studied the nature of RSFC between the ROIs where we found APS effects: TIFG-L, ICSI-R, OFC-R, and LSTG-R for zero-lag correlation and SFG-R, ACC-L, and PP/LSTG-L for coherence. The cross-correlations were computed with the filtered data (i.e., band-pass filtering between 9 and 80 mHz). Group averaged lagged cross-correlations are depicted in Figure IV-5. An interesting difference between the auditory ROIs and non-auditory ROIs was found. As one can clearly see from the group-averaged correlograms, the auditory regions near the PAC (i.e., PP/LSTG-L, ISCI-R, LSTG-R; Figure IV-5, upper row) show very distinct peaks at zero-lag. In contrast, the non-auditory cortical regions (i.e., TIFG-L, ACC-L, MOFC-R, SFG-R; Figure IV-5, lower row) did not show such a pronounced peak at any time lag. This contrast between the auditory vs. non-auditory cortices seems to corroborate the high modularity of the auditory RSN from already mentioned rs-fMRI studies (Damoiseaux et al., 2006; Fauvel et al., 2014; Luo et al., 2012; Shirer et al., 2012; Smith et al., 2013; Smith et al., 2009; Smith et al., 2012). It suggests that the right PP is very tightly coupled with a small time lag to other auditory cortices (even in the contralateral hemisphere).
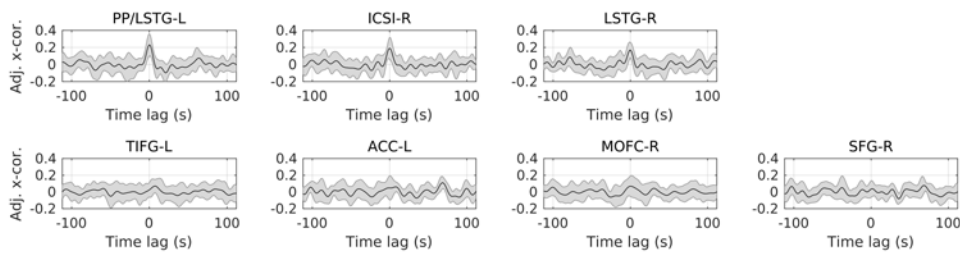

Figure IV-5. Group averaged (n = 17) correlograms are plotted with standard deviation shaded in gray for adjacent auditory cortical regions (upper row) and distant non-auditory cortical regions (lower row). Cross-correlation coefficients were adjusted for the effect of nuisance covariates.

IV-3.3. Correlation between Cortical Myelin and Resting-state Functional Connectivity

We hypothesized a positive relationship between the intracortical myelination of the right PP and its RSFC to adjacent areas, based on the assumption that intracortical myelination enhances short range connectivity. Theoretically, it is plausible that cortical myelination of local neural circuits would increase information transfer efficiency and synchrony by reducing variance in information transfer time lag. However, empirical evidence for such facilitation in the human cerebral cortex remains to be found.

In the previous work (Chapter III; Kim and Knösche (2016)), we found a positive correlation between the APS and the cortical myelin content in the right PP. In the current study, we also found a positive relationship between the APS and RSFC of the right PP to a number of neighboring and distant cortical areas. Although we found that both intracortical myelination and RSFC of the right PP were correlated with APS, it does not necessarily indicate a correlation between myelination and RSFC. As proven by Langford and colleagues, positive correlation is not *transitive* unless the correlation coefficient is sufficiently close to 1 (Langford et al., 2001). That is, a positive correlation between X and Y and one between Y and Z does not necessarily imply a positive correlation between X and Z.

Therefore, we directly tested the relationship between intracortical myelination (indexed by the longitudinal relaxation rate qR1) in the right PP and RSFC between the right PP and the rest of the cortex using a simple GLM without covariates as:

$$FC = \beta_0 + qR_1\beta_1 + \varepsilon ,$$ \hfill (IV-4)

At a slightly liberal alpha level of 0.051, we found a positive effect of qR1 in the zero-lag cross-correlation between the right PP and the right first transverse sulcus (FTS) and LSTG as shown in Figure IV-6: T(15) = 5.53, beta = 15.8, p = 0.051, area = 70.8 mm², MNI-305 = (63, 0, -1) mm. This suggests that the myeloarchitecture of a local circuit has a significant impact on its FC with adjacent regions. Further investigation with the general population (i.e., including non-musicians) may find a more generalizable relationship between cortical myelination and FC.
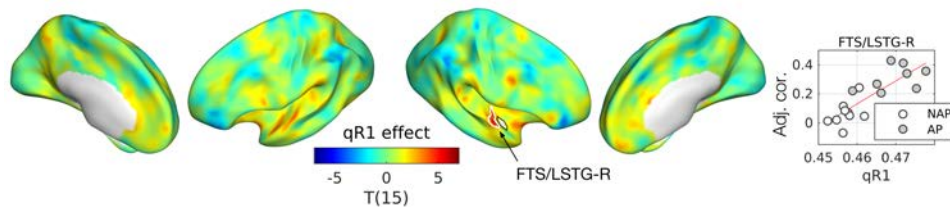
Figure IV-6. Effect of cortical myelin (qR1) on cross-correlation. A significant cluster in the right first transverse sulcus and lateral superior temporal gyrus (FTS/LSTG-R) is marked with a scatterplot at the peak vertex. Abbreviation: qR1, quantitative longitudinal relaxation rate; Adj. cor., adjusted zero-lag cross-correlation.

IV-3.4. Replication of Main Findings only with Europeans

Since there is a mismatch between the musicians with and without AP in terms of ethnicity in the present data, one might be concerned whether the found effect of APS in RSFC could be confounded by such an imbalance. To rule out such possibilities, we reanalyzed the effects of APS in zero-lag cross-correlation and cross-coherence against the right PP using only European musicians (3 APs vs. 9 non-APs) as shown in Figure IV-7.

We replicated the main findings in principle. For zero-lag correlation, we found APS effects in (1) TIFG-L, (2) ICSI-R. For cross-coherence, we found (3) PP/LSTG-L, (4) FTS-R, and (5) ACC-R. There were other regions that showed significant effect of APS possibly due to the small sample size in the AP group (N=3). Nonetheless, it must be noted that the main findings that overlap with this result with only European musicians cannot be due to ethnicity confounding. The full list of results can be found in Table IV-3
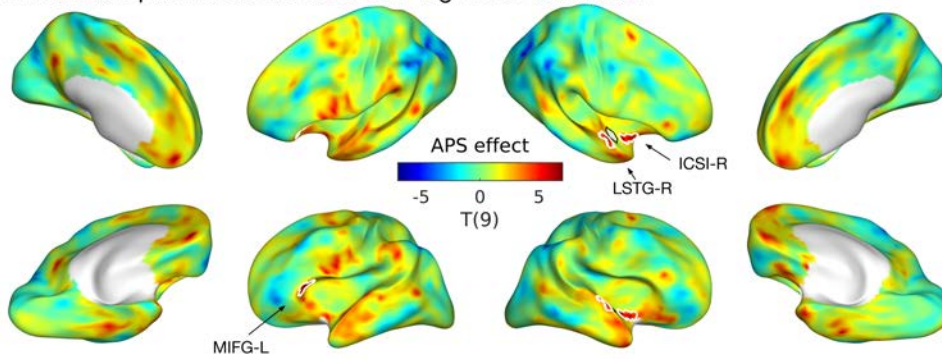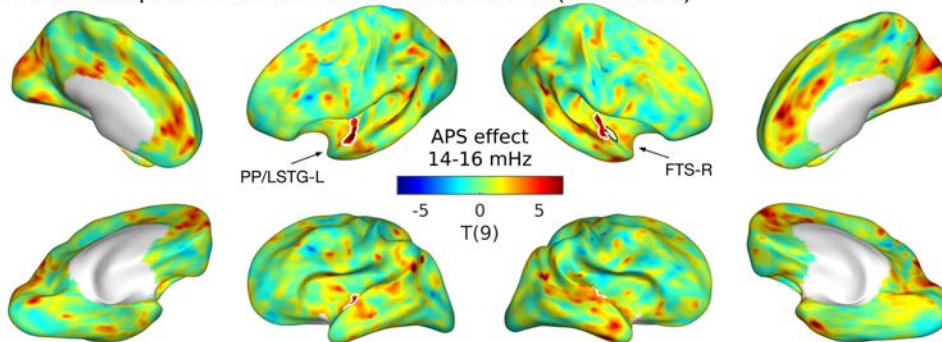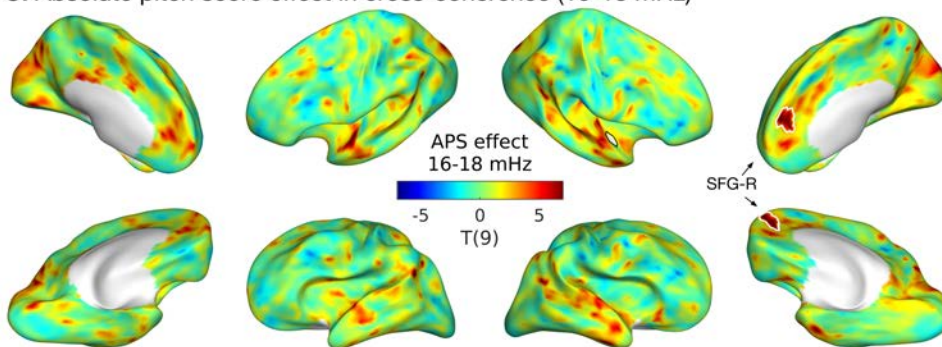
Figure IV-7. Effects of the absolute pitch score (APS) in (A) zero-order cross-correlation and (B, C) cross-coherence only with European musicians. T-statistic maps for the effect of the acuity of AP in correlation are projected onto semi-inflated cortical surfaces. Significant clusters are indicated by white contours. See Table IV-3 for details of the clusters.

Table IV-3. Effects of absolute pitch score in cross-coherence and cross-coherence.

| RSFC Label		Freq. (mHz)	Beta	Max T(9)	P-value*	Area (mm²)	MNI-305 (mm)	Full name of structure†
Cor.	TIFG-L		0.73	7.44	0.0123	109.1	-39, 26, 3	Left triangular part of the inferior frontal gyrus
	ICSI-R		0.79	7.45	0.0096	65.8	42, -5, -16	Right inferior segment of the circular sulcus of the insula
Coh.	PP/STG-L	14-16	1.16	14.95	0.0189	152.2	-58, 6, -5	Left planum polare and lateral superior temporal gyrus
	FTS-R	14-16	0.96	7.45	0.04	129.9	60, 0, -1	Right first transverse sulcus
	ACC-R	16-18	1.21	10.53	0.0226	189.3	6, 47, 19	Right anterior part of cingulate and sulcus

*P-values are corrected for multiple comparison correction. † Identification of anatomical nomenclature is generally based on automatic parcellation by FreeSurfer (Destrieux et al., 2010). Abbreviation: Cor., correlation; Coh., coherence.

IV-4. Discussion

In this study, we investigated the effect of acuity of AP perception on the RSFC of the right PP, which was highly myelinated in the Section III-3.4 (Kim and Knösche, 2016). We found functional connectomes in favor of the involvement of the right ventral auditory pathway in AP perception and further propose a potential relationship between local myelination within a cortical region and its intrinsic and extrinsic FC.

As hypothesized, we found greater RSFC of the right PP with the adjacent ipsilateral auditory cortex (i.e., the right LSTG) as well as the contralateral auditory cortices (i.e., the left PP/LSTG). Furthermore, we found additional evidence for the fronto-temporal network integration and a close relationship between the auditory RSN and the DMN in AP listeners. In the following sections, we will discuss the physiological plausibility and behavioral relevance of the current findings.

IV-4.1. Resting-state Functional Connectivity and Neural Interaction

It is commonly accepted that BOLD signals are closely related to neuronal activity, as, for example, supported by simultaneous measurement of local field potentials (LFPs) and BOLD in optogenetic rodents (Lee et al., 2010). Consequently, RSFC should be reflective of interactions between neural populations, mediated by neurovascular coupling. Although the dynamics of such coupling remains unclear, a line of studies suggests that the connectivity amongst neural activities is indeed reflected in the FC based on BOLD signals (Achard and Bullmore, 2007; Achard et al., 2006; Britz et al., 2010; Cordes et al., 2001; De Luca et al., 2006; Ko et al., 2011; Qian et al., 2015; Salvador et al., 2005a; Salvador et al., 2005b; Sasai et al., 2014; Shmuel and Leopold, 2008; Sun et al., 2004).

One of the points of discussion with respect to this relationship is spectral scale. As mentioned above, the dynamics of EEG and fMRI signals are different by about 2–3 orders of magnitude (typically, EEG, 1–100 Hz; fMRI, 0.001–1 Hz). Therefore, the spectral relationship between the

BOLD and electrophysiological signals has been tested in various studies, especially in terms of spontaneous neural activities. One possibility is to look at very slow components of electrophysiological signals that match the frequency range of BOLD (i.e., < 1 Hz). For instance, ICA-based blind source separation on simultaneous recordings of EEG and fMRI during rest extracted RSNs based on BOLD signal and EEG (Hiltunen et al., 2014). EEG data were low-pass filtered at 55 mHz and down-sampled to match the fMRI data. Several BOLD-RSNs and EEG-RSNs including the DMN and the *salience network* (SN) were found spatially overlapped, suggesting common neural sources that are reflected in both scalp potentials and BOLD signals. Also, in an animal model, strong cross-correlation between LFP and BOLD signals, which were both low-pass-filtered at 100 mHz, was found at the time lag of 4 sec in a group of anesthetized rats (Pan et al., 2013). However, direct relationships between BOLD and high frequency electrophysiological signals have also been reported. For example, in a study using human intracortical EEG (iEEG), high coherence between DMN nodes (i.e., the ACC and precuneous) was found from slow (15 mHz) BOLD signal and high gamma (65–110 Hz) activities (Ko et al., 2011).

Another point is the relationship between the frequency of BOLD signals and the range of FC. In one of the early seminal studies on resting-state fMRI (Salvador et al., 2005a), partial correlation and coherence in a relatively low fMRI frequency range (0.4–152 mHz) generally decayed over the increasing Euclidian distance between the regions. However, the connectivity between homologous areas in different hemispheres remained strong despite their relatively long distances (Salvador et al., 2005a). This effect was absent in patients with acute brainstem ischemia (Salvador et al., 2005a), suggesting that the strong interhemispheric connectivity between homologous regions in the lower BOLD frequency range may reflect common input from the brainstem rather than transcallosal communication. The specific involvement of lower frequencies in long-range connections was further supported by a follow-up study (Salvador et al., 2005b): in a relatively low frequency range (i.e., 4–152 mHz), mutual information between the homologous regions remained high over distance between ROIs, whereas such an effect was not found in a higher frequency range (i.e., 300–455 mHz).

In this study, we observed interhemispheric connectivity between the right PP and homologous or near-homologous cortical regions in the left hemisphere. As shown in Figure IV-3, the lowest frequency bin (0–2 mHz) showed greater coherence over the whole cortex, and especially between the homologous regions, in comparison to the relatively high frequencies (24–26, 48–50 mHz). This reconfirms and refines the picture of long-range interhemispheric connections being prominent in lower BOLD frequencies. Given that the strong low-frequency coherence between the homologous pairs was destroyed by brainstem lesion (Salvador et al., 2005a), the current observation of low-frequency coherence could imply that AP-related functional/structural alteration might exist even from the brainstem level, and that inputs from there give rise to FC between distant areas, especially in both auditory cortices. This also agrees with the finding that

cortical myelination seems to be related to short-range connections within the right auditory cortex, but not to long-range connections.

IV-4.2. Ventral and Dorsal Auditory Pathways

In the Chapter III, a postulation was discussed that the right PP is important in AP perception in the context of the 'dual auditory stream' hypothesis (i.e., dorsal and ventral pathways) (Kaas and Hackett, 1999; Rauschecker, 2012; Rauschecker and Tian, 2000), based on the enhanced cortical myelination found in the region for AP musicians as in the Section III-3.4 (Kim and Knösche, 2016). As discussed earlier in III-4.2, the anterior part of the STP has been consistently found to be involved in processing non-spatial aspects of auditory information, such as pitch, timbre, and phoneme identity (Altmann et al., 2007; Barrett et al., 2005; Barrett and Hall, 2006; Hart et al., 2004). In particular, as mentioned earlier, the anterior part of the STP including PP showed greater sensitivity compared to the posterior part (Warren et al., 2003b). The recognition of pitch chroma is what AP listeners are good at, while they may not be better than non-AP musicians at perceiving pitch height (Deutsch and Henthorn, 2004; Miyazaki, 1988; Miyazaki, 2004a; Takeuchi and Hulse, 1993).

We discovered a positive relationship between the APS and RSFC between the right PP and ipsilateral (right LSTG) as well as contralateral (left PP/LSTG) auditory areas. This finding supports our hypothesis on the involvement of the ventral auditory pathway in AP perception. Moreover, the significant relationship between intracortical myelination in the right PP and its RSFC to neighboring non-primary auditory cortices in the right ventral pathway indicates that heightened myelination in the right PP is functionally meaningful in auditory processing, supposedly preserving specificity of transferred information due to dense cortical myelination.

Conversely, previous studies have consistently implicated the dorsal auditory pathway with an emphasis on structural and functional laterality in favor of the left hemisphere (Keenan et al., 2001; Ohnishi et al., 2001; Schlaug et al., 1995). Moreover, the involvement of the left fronto-temporal network was demonstrated by increased phase synchrony between the frontal and temporal source activities based on EEG data (Elmer et al., 2015) and greater FA of the superior longitudinal fasciculus (Oechslin et al., 2009) in AP listeners. However, recent studies (Bermudez et al., 2009; Dohn et al., 2015; Wengenroth et al., 2014) suggested that the AP-related processes might not be exclusively processed in one hemisphere. Given the rapid processing of AP (Itoh et al., 2005; Miyazaki, 1988), chromatic categorization of a given tone and linking that to a lexical label and/or a motor program should be done very efficiently. Such efficiency would be maximized only when specialized subprocesses are properly integrated. For instance, it could be possible that the perceptual process of AP is implemented in the right STP whereas linking the transferred pitch chroma representation with verbal or non-verbal labeling is processed in the left VLPFC, as previously proposed by Wengenroth and colleagues (Wengenroth et al., 2014). One piece of

evidence for such integration could be our current finding of the high degree of RSFC between the right PP and the left IFG in AP musicians.

IV-4.3. Link to the Default-mode Network

The DMN was initially suggested from the decrease of blood flow during language-related tasks compared to resting during PET experiments (Shulman et al., 1997), and has been further established by subsequent fMRI studies (Greicius et al., 2003; Smith et al., 2009). Given that the notion of 'default-mode' came from the deactivation, many studies related the DMN to internally generated processes, such as, for example, self-referential episodic memory, prospection, and theory of mind (Greicius and Menon, 2004; Spreng et al., 2009). The auditory RSN usually shows anticorrelation with the DMN and positive correlation with the SN (Menon, 2011), which directs attention to the sensory inputs. This anticorrelation between the DMN and SN seems to reflect 'deactivation' during rest (i.e., negative contrast from the subtraction of 'rest' condition from 'stimulus' condition) in block-designed fMRI experiments (Greicius and Menon, 2004).

In this study, interestingly, the auditory RSN based on the right PP showed APS-related neural coupling to the medial frontal structures, that is, ACC, medial SFG, and medial OFG, which are known to be crucial nodes of the DMN (Greicius et al., 2003; Greicius et al., 2009; Smith et al., 2009; Zald et al., 2014). The frequency band of the coherence (10–18 mHz) in the current finding was also in accordance with one of the few multimodal BOLD imaging studies using simultaneous fMRI and fNIRS (Sasai et al., 2014) that reported a high degree centrality of the ACC in a low frequency band (10–30 mHz). Increased RSFC between the right PP and the DMN network might be related to spontaneous AP-related processing that is active even during rest, which could result in unintentional AP listening (Miyazaki, 2004a).

IV-4.4. Auditory insular cortex

We found a positive APS effect for the zero-lag correlation between the right PP and the anterior part of the inferior segment of the circular sulcus of the insula (i.e., ICSI). Auditory regions of the insula in human brains have been implicated from cytoarchitecture studies (Rivier and Clarke, 1997), microelectrode recording in patients (Remedios et al., 2009) and fMRI studies (Bamiou et al., 2003). It has been previously shown that the anterior insula is involved in auditory object identification (Binder et al., 2004). Moreover, dissection of the anterior insula in epileptic patients impaired verbal naming of familiar faces with intact conceptual knowledge (Papagno et al., 2010). Taken together, higher RSFC between the right PP and the right ICSI may indicate closer interaction for sound source identification in individuals with greater acuity of AP.

IV-4.5. Conclusion

The right PP, which had been identified as more heavily myelinated in AP musicians, was found to be part of an AP related network of RSFC. Importantly, this network comprised adjacent ipsilateral auditory areas, thus corroborating the prominent role of the right ventral auditory stream in AP categorization, and a left hemispheric fronto-temporal network, possibly indexing the integration of the categorization and (verbal) labeling/association phases of AP processing. The fact that a direct relationship with cortical myelination of the PP was only found for RSFC with neighboring ipsilateral areas suggests that this structural change is indeed relevant for the improved categorization abilities of AP musicians, but is probably not the only structural alteration giving rise to the observed functional changes, in particular with respect to the association phase.

V. General Discussion

The present thesis was strongly motivated by the current lack of mechanistic explanation on how the absolute pitch (AP) is realized in some humans. Previous neuroimaging studies only localized alteration of morphology, function, and connectivity without any suggestion of neural mechanism behind the perception of AP. For instance, if AP listeners have bigger right Heschl's gyrus and greater MEG and BOLD activation (Wengenroth et al., 2014), how does it facilitate AP? But it would be only fair to mention that it is not a unique problem in studies on AP but somewhat ubiquitous in the neuroimaging field. This thesis also only adds localization of a physical quantity that is believed to be related to myeloarchitecture of human cortex, and altered functional coupling related to the found cortical area. However, given that more primary and fundamental neural processing of audition and pitch perception still remains to be explained in terms of the working mechanisms of engaged neural systems, computational modeling of AP certainly requires a lot more studies to gain enough knowledge in how human auditory systems work.

However, the other way around, maybe we could learn more about the general principle of human auditory systems by studying human subjects with distinguished auditory perception such as AP, which is believed to be an outcome of interaction between genetic and environmental factors (Zatorre, 2003). In other words, the studies on AP are not just for understanding the unusual auditory perception per se but also could provide insights into the underlying mechanism of higher-order auditory processing in general humans.

I believe the current thesis also contributes to the knowledge on the pitch chroma perception in a general sense. For instance, it has been well known in musicology (Deutsch, 1982) and acoustics (Shepard, 1964) that people perceive the pitch of sound, which conveys one-dimensional information (frequency for pure tones or periodicity for complex tones), as a two-dimensional psychological representation (i.e., pitch chroma and pitch height). However, there are only three neuroimaging studies (Briley et al., 2013; Moerel et al., 2015; Warren et al., 2003b) that investigated the neural underpinning of pitch chroma perception. In the fMRI studies (Moerel et al., 2015; Warren et al., 2003b), the anterior part of the superior temporal plane (STP) was found to be sensitive to pitch chroma information. Also from the equivalent current dipole (ECD) fitting based on the event-related potential (ERP) responding to pitch chroma (Briley et al., 2013), the two-dipole model localized bilateral regions on the superior temporal planes (STPs) that were anterior and lateral with respect to primary auditory cortex (PAC). In the first study of the current thesis, we found myeloarchitectonic correlates of the AP, which entails absolute recognition of pitch chroma but not of pitch height. Thus our finding suggests that the processor of pitch chroma information lies in the anterior STP and converges with the previous functional studies from a different perspective.

In this final chapter, I will summarize key findings of the individual studies, contributions of the current thesis, and possible implications to the pitch chroma perception in general population. In addition, I will discuss limitations of the current thesis and ideas for future studies that could eventually contribute to the computational modeling of the absolute recognition of pitch chroma.

V-1. Main findings

Our research questions were (1) whether the microstructure of the auditory cortex is related to AP, (2) if any, to what extent the alteration of myelination affects functional connectivity, and (3) its relevance to the ventral auditory pathway. We discovered novel findings as follows:

(1)   There is a positive effect of AP (i.e., group difference) and AP score (APS) in the quantitative relaxation rate (qR1), reflecting cortical myelin density, in the right planum polare (PP).

(2)   The AP-related myelination was differentiated from another association between the relative pitch discrimination ability and qR1 in the lateral Heschl's gyrus (HG).

(3)   There are positive effects of APS in the resting-state functional connectivity (RSFC; i.e., zero-lag cross-correlation and cross-coherence) based on the right PP (which was more myelinated in individuals with higher APS) in:

   a.   adjacent auditory cortex (i.e., the right lateral superior temporal gyrus (LSTG)) and auditory insular cortex (i.e., the anterior region of the right inferior circular sulcus of the insula (ICSI)),

   b.   contralateral homologues (i.e., the left PP and the left LSTG),

   c.   anteromedial structures (i.e., the medial orbitofrontal cortex (OFC), the left anterior cingulate cortex (ACC), the right medial superior frontal gyrus (SFG)),

   d.   ventrolateral frontal cortex (i.e., the left inferior frontal cortex (IFG)).

First, the myeloarchitecture of auditory cortex, in particular the right PP, was found to correlate with the APS. Because the APS effect was found only at the middle depth (i.e. at 50 % of cortical thickness from the inner boundary toward the outer boundary of cortex), we speculated that the increase of myelin content would be related to lateral connections between adjacent groups of neurons in the PP. From histological works, it has been known that a subfield of the PP called the anterior area (AA) has a dense distribution of pyramidal cell bodies in layer III (Rivier and Clarke, 1997). It has been also known that an anterior subfield of the PP called *Temporalgebiet* (TG; temporal area; (von Economo and Koskinas, 1925)) has widely spread (i.e., about 5–6 mm; compared to 1–2 mm of PAC) connections towards Heschl's gyrus and non-primary auditory cortices (Tardif and Clarke, 2001). Given such a prominent myeloarchitecture of the PP implying

its integrative function, the cortical myelin in the region is highly likely to be engaged in higher level auditory processing at the cortex level, which could be essential for the perceptual stage of AP.

Second, we found a close relationship between the cortical myelin in the right LSTG and frequency discrimination thresholds (FDT). The pitch discrimination test for FDT did not require verbal coding or absolute pitch memory as the two sine tones (i.e., reference and target) were presented without a lag in-between. The measure showed clear distinction between musicians and non-musicians ($T(32) = 2.6$, $p = 0.013$; not previously reported data, comparing 19 musicians vs. age/gender matched 15 non-musicians) whereas no significant difference found between musicians with and without AP ($T(17) = 0.56$, $p = 0.57$). Thus the behavioral measure is irrelevant to AP but related to the spectral resolution of auditory systems, presumably at the pitch extraction level. We found neuronal correlates of the FDT in the right LSTG consistently with previous studies that found higher selectivity in the frequency domain of the right auditory cortex, particularly the lateral borders of HG and LSTG (Bendor and Wang, 2005; Norman-Haignere et al., 2013; Patterson et al., 2002; Penagos et al., 2004). It is very intriguing that the myeloarchitectonic correlate of AP was in close proximity, but clearly distinguished from that of FDT, which reflects the behavioral dissociation of the AP precision and the relative spectral resolution that has been constantly found throughout literature (Deutsch and Henthorn, 2004; Miyazaki, 1988; Miyazaki, 2004a; Takeuchi and Hulse, 1993). This dissociation of myeloarchitectures suggests that precise extraction of pitch occurs in the right LSTG (or a sub-network centering at LSTG) while chromatic categorization of the extracted pitch takes place in the right PP (or its sub-network).

Third, the highly myelinated subregion of the right PP showed greater functional connectivity with various cortical regions during rest in individuals with higher APS supporting the importance of the right PP in the AP-related resting-state networks. We hypothesized that the RSFC between the right PP and the other auditory regions along the ventral pathway would correlate with APS, if the right PP bears a central position in the AP-specific ventral pathway. As expected, we found that the resting-state fluctuations of the BOLD signal in the right PP were more tightly coupled with the adjacent auditory cortex (i.e., the right LSTG) and the contralateral homologues (i.e., the left PP/LSTG) in musicians with better APS. In addition, we also found increased RSFC between the right PP and the right ICSI in musicians with greater APS, of which the implication in auditory object identification has been known from an fMRI study (Binder et al., 2004), an MEG study (Ahveninen et al., 2016), and a lesion study (Papagno et al., 2010). This finding suggests that the right PP in fact has a pivotal role that is relevant to AP in the ventral auditory pathway.

Moreover, we found greater RSFC between the right PP and anterior medial structures (i.e., the left ACC and the right medial SFG) that are known as important nodes of the *default-mode network* (DMN) (Greicius et al., 2003; Greicius et al., 2009; Smith et al., 2009; Zald et al., 2014). Commonly auditory cortex is encapsulated in the *salience network*, which negatively correlates

with DMN. This alteration may reflect the spontaneity of AP even during rest, which enables unintended perception of pitch chroma of ambient noise.

Finally, we found a positive effect of APS in the RSFC between the right PP and the left ventrolateral prefrontal cortex (i.e., the left IFG, loosely corresponding to Broca's area). This increased connectivity between the right PP and the left IFG could be an indication of communication between the perceptual stage and the (verbal) associative stage of the AP.

V-2. Contributions of the Current Thesis

As previously mentioned, I suggest that the current thesis contributes to the knowledge on the pitch chroma perception in AP listeners as well as in the general population. In this section, I will discuss how the present thesis expands our understanding of the neural mechanism of AP and how it could be related to general auditory perception in people without AP.

First of all, the current thesis stresses the importance of the perceptual part of the AP in contrast to the view that most people have 'implicit AP' and the only difference between AP and non-AP listeners is just the speed of the association between the 'absolute memory' and verbal labels (Levitin, 1994; Smith and Schmuckler, 2008). It is well known that many people have relatively good memory of the pitch in the absolute sense of ones' favorite songs (Levitin, 1994) or frequently exposed tones such as landline ring tones (Smith and Schmuckler, 2008) or the censoring tone at 1K Hz used for masking inappropriate words in movies or TV shows (Van Hedger et al., 2016). In the trials where a participant was required to sing or hum one's familiar songs (Levitin, 1994), the average absolute deviations corrected for octave-error were 1.84 and 2.25 semitones for the first and second songs, respectively (the data was not numerically given in the text, so computed from the histograms in the paper). Given that the chance level is 3 semitones for octave-corrected deviations, the performance is not as impressive as the authors suggested, particularly for the second trial. Moreover, the proportion of the people who sang in exact pitch was about 20 % out of 43 subjects. Especially, the sample included subjects with and without musical background. Thus, that study (Levitin, 1994) only weakly supports the argument that the 'implicit AP' is ubiquitous in general population.

Previously, Itoh and colleagues observed the 'AP negativity' (Itoh et al., 2005), which is a negative EEG component peaking around 150 ms after auditory stimulus onset. Because this early EEG component was found only in musicians with AP, it was suggested that early auditory process is crucial in AP perception. Conversely, another EEG study reported absence of the AP negativity in musicians (Elmer et al., 2013) arguing that there is no AP-specific activity in the early stage of auditory processing. However, Itoh and colleagues observed the AP negativity at a left posterior-temporal electrode (i.e., T5 in the conventional 10-20 EEG lattice) (Itoh et al., 2005) whereas electrodes in posterior-temporal regions were simply discarded in the analysis of Elmer and

colleagues (Elmer et al., 2013). Thus the negative result in the study (Elmer et al., 2013) did not disprove alteration in the perceptual stage of AP. Instead, there is a line of evidence that function and structures of the temporal lobes of musicians with AP are different from that of musicians without AP from fMRI (Dohn et al., 2015; Itoh et al., 2005; Keenan et al., 2001; Luders et al., 2004; Ohnishi et al., 2001; Schlaug et al., 1995; Wengenroth et al., 2014; Zatorre et al., 1998). Given the latency of the P2 (i.e., 200-300 ms) that is known as to be generated from anterior planum temporale (PT) from musicians with and without AP (Wengenroth et al., 2014) and the latency of the AP negativity (i.e., 150 ms), the M/EEG responses are likely to reflect functional alternation in the non-primary auditory cortex, presumably rendering distinctive pitch chroma extraction. Taken together, by focusing on the auditory cortices of AP listeners, the current thesis draws attention to the most important and intriguing part of AP.

Second, more importantly, the current thesis suggests a link between the absolute recognition of pitch chroma and the ventral auditory pathway based on the dual pathway hypothesis. According to the hypothesis and several related pieces of evidence, the ventral pathway processes object identity related features. In many fMRI studies sensitivity to location-invariant features that are related to the identity of a sound source such as timbre, phoneme, frequency modulation, melody contour, and pitch chroma was found in constituents of the ventral pathway (Altmann et al., 2007; Barrett et al., 2005; Barrett and Hall, 2006; DeWitt and Rauschecker, 2012; Hart et al., 2004; Warren et al., 2003b) whereas sensitivity to spatial information and somatosensory/motor related information was found throughout the dorsal pathway including parietal, and sensory/motor cortices (Arnott and Alain, 2011; Rauschecker, 2011; Warren et al., 2005).

Given this distinction, the chromatic categorization is more likely to be realized in the ventral pathway than the dorsal pathway, because pitch chroma is a discrete (i.e., categorizable) feature of a tone to AP listeners unlike non-AP listeners. Therefore, we suggested that the heavy myelination in the right PP has a close relationship with the chromatic categorization in AP listeners as mentioned in the Chapter III (Kim and Knösche, 2016), which has not been considered as a framework to explain the perceptual stage of the AP recognition before our publication. In fact, the alteration of fractional anisotropy in the right ventral pathway, in particular inferior longitudinal fasciculus (ILF), had already been reported (Dohn et al., 2015), but the dual pathway hypothesis for an underlying mechanism of AP had not been introduced.

It should be noted that the suggested importance of the ventral pathway in AP is not mutually exclusive with the previous literature that consistently reported alteration in the left dorsal pathway (i.e., posterior planum temporale (PT), superior longitudinal fasciculus (SLF), dorsolateral prefrontal cortex (DLPFC) in the left hemisphere) in musicians with AP (Itoh et al., 2005; Keenan et al., 2001; Luders et al., 2004; Oechslin et al., 2009; Ohnishi et al., 2001; Schlaug et al., 1995; Wengenroth et al., 2014; Zatorre et al., 1998). Although we did not find alteration in the left dorsal pathway, it is plausible that the left hemisphere could actively contribute to the other processes of

AP recognition. Other than the relationship with AP, the left PT was also implicated from auditory processing of rapidly changing phonetic cues in musicians compared to non-musicians (Elmer et al., 2012), which could be understood in the context of hemispheric-specialization in terms of processing window duration (i.e., the left PT mainly processes at the low-spectral and high-temporal resolution) (Zatorre and Belin, 2001). Furthermore, a close relationship between the posterior STP and motor/somatosensory cortices was found while well-trained musicians were watching a video clip of playing the piano (Hasegawa et al., 2004). Rauschecker extended the role of the dorsal pathway (from the auditory cortex towards premotor cortex via inferior parietal lobules) as including somatosensory and premotor cortices with respect to vocalization motor/feedback systems (Rauschecker, 2011). All things considered, the role of the left dorsal pathway could be association between the chromatic categorization and the verbal labels (i.e., solfeggio) and sensory/motor programs (i.e., playing their own musical instruments) followed by pitch chroma extraction, which is more likely to be processed along the right ventral auditory pathway as given in the Chapter III (Kim and Knösche, 2016).

The auditory pathways that are possibly involved in AP are depicted in Figure V-1. As already discussed above, only the left dorsal pathway (a blue arrow in Figure V-1) has been emphasized in previous studies concerning functional and structural alteration of temporal cortex (Itoh et al., 2005; Keenan et al., 2001; Luders et al., 2004; Oechslin et al., 2009; Ohnishi et al., 2001; Schlaug et al., 1995). However, in our current studies presented in the Chapter III and Chapter IV (Kim and Knösche, 2016; Kim and Knösche, in preparation), the importance of the ventral pathway in AP was suggested based on greater cortical myelin in the right PP and greater functional connectivity between the right PP (in green) and the left PP and anterior LSTG (in red). Additionally, the right HG (including lateral aspects) (Wengenroth et al., 2014) was also found greater in AP musicians. We have not found significant alterations in the right dorsal pathway (in gray), which could be related to pitch height perception given the importance of octaves in auditory scene segregation (Warren et al., 2003a). Moreover, it remains unclear to what extent subcortical structures are also involved in AP processing. Given that the coherence between the right and left PPs in a low frequency bin was stronger in musicians with better AP and that the interhemispheric coherence across long distance was affected by lesions in brainstem (Salvador et al., 2005a), the involvement of subcortical structures in AP might be possible although we did not find significant effect in the current data.
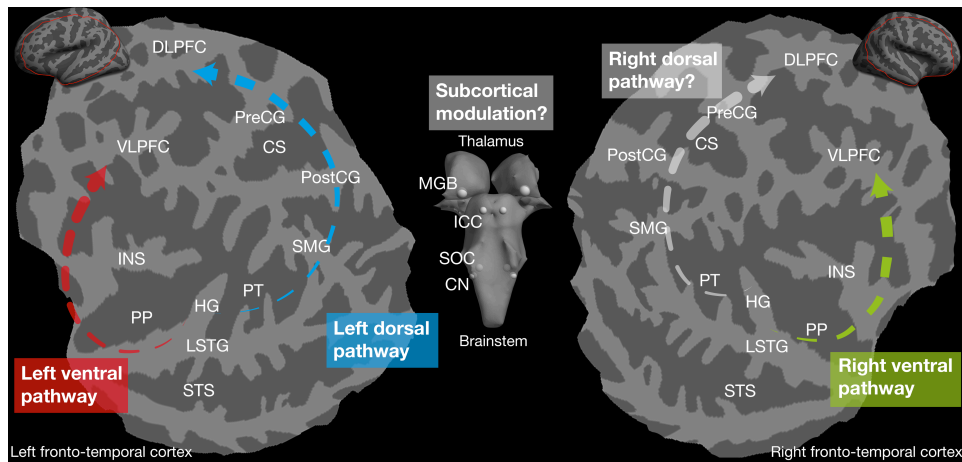
Figure V-1. Overview of fronto-temporal networks and subcortical networks. Abbreviations: CN, cochlear nucleus; SOC, superior olive complex; ICC, inferior colliculus; MGB, medial geniculate body; HG, Heschl's gyrus; PP, planum polare; PT, planum temporale; LSTG, lateral superior temporal gyrus; STS, superior temporal sulcus; SMG, supramarginal gyrus; INS, insula; CS, central sulcus; Pre-/PostCG, pre/post-central gyrus; VLPFC, ventrolateral prefrontal cortex; DLPFC, dorsolateral prefrontal cortex.

Finally, as mentioned earlier, it has been known that the sensitivity to pitch chroma was localized in the bilateral anterior STPs in non-musicians without AP from the fMRI study (Warren et al., 2003b). The abovementioned EEG study also localized the pitch onset response by two dipoles on the bilateral STPs that were more anterior and lateral compared to that of the sound onset response (Briley et al., 2013). In the current thesis, the right PP showed heavier cortical myelination and its functional connectivity between the cortical regions that are included in the ventral pathway. As behaviorally demonstrated throughout the literature (Deutsch and Henthorn, 2004; Miyazaki, 1988; Miyazaki, 2004a; Takeuchi and Hulse, 1993), some AP listeners make errors in pitch height without errors in pitch chroma, especially for unfamiliar timbre or pitch range. Thus, the cortical myelination in the right PP may have implication even for non-AP listeners. Since the acuity of AP perception was positively correlated with the cortical myelin in the PP over musicians with various degree of AP (i.e., non-AP, quasi-AP, genuine-AP, and so on), it is likely that the pitch chroma extraction process, possibly matching the pitch of the current input to long-term (i.e., AP-specific) or short-term (i.e., temporary tonal context or an explicit reference) templates, takes place in the right PP. Musicians with AP can be thought as a human model with certain genetic/environmental interactions that render internalization of pitch chroma references as previously suggested (Zatorre, 2003). Therefore, studies on AP provide valuable insights into how the human auditory system encodes tonal context and matches pitch chroma with references.

V-3. Remaining Questions and Ideas for Future Studies

While the current thesis adds knowledge on the myeloarchitecture and intrinsic functional connectivity of the auditory cortex in musicians with AP, there are abundant questions that remain to be revealed to answer how AP is realized in the human auditory systems. In the following subsections, I will discuss some of these questions and ideas for further investigations.

V-1.1. Separation of Pitch Chroma and Pitch Height Processing

Inspired by the study by Warren and colleagues (Warren et al., 2003b), the distribution of neurons that are sensitive to pitch chroma and pitch height, respectively, over the STPs should be examined in AP listeners. Already our structural findings and their functional relevance with respect to the ventral auditory pathway implicate that pitch chroma information would be processed in the anterior STP. However, there has been no direct functional evidence that this separation is enhanced in AP listeners. Functional mapping of sensitivity of chroma and height in AP listeners, if found, is expected to consolidate our hypothesis that the absolute pitch chroma processing takes place in the right PP and adjacent auditory cortices along the ventral pathway.

There is a report that pure tones (sine waves) failed to attenuate cortical response to pitch-onset in non-AP non-musicians (Briley et al., 2013). Moreover, previous pitch chroma studies used complex or musical tones (Moerel et al., 2015; Warren et al., 2003b). It would be even more interesting whether such an effect of timbre interacts with the AP perception by using pure tones and complex tones for experiment stimuli.

V-1.2. Pitch Exaction and Chromatic Categorization

We have assumed the pitch chroma would be categorized after the extraction of pitch (presumably in the lateral HG and anterior LSTG) in AP listeners. Because we found distinctive myeloarchitectonic correlates for the AP score (APS) and frequency discrimination threshold (FDT) in the right PP and the right LSTG, respectively as in the Chapter III (Kim and Knösche, 2016), it seems are adjacent but different groups of neurons work on fine-resolution pitch extraction (i.e., < 0.1 semitone) and chromatic categorization (i.e., > 1 semitone). However, there have been no direct functional studies so far. Thus it is still unclear whether the two supposed extraction processes really exist and whether they are consequential or simultaneous.

For pitch extraction, the pitch onset response (POR, sensitive to 'pitchness', or prominent periodicity that can be detected by autocorrelation) is known to be distinctive from the sound onset response (SOR, sensitive to energy change). In an early study (Krumbholz et al., 2003) to segregate POR from SOR, an adaptation paradigm (see Figure V-2 for an example) was used: White noise was first presented followed by iterated ripple noise (IRN) (Yost, 1996) with the same energy but evoking pitchness to attenuate SOR and elicit only POR (Krumbholz et al., 2003). In a

following study with human deep electrode recordings (Schönwiesner and Zatorre, 2008), the medial HG showed only SOR (latency of 100-130 ms) but no POR whereas the lateral HG showed only strong POR (latency of 220 ms) with no SOR. This double dissociation serves as strong evidence that the pitch extraction process is likely to engage the lateral HG but not the medial region. Moreover, it is still possible that other non-primary auditory cortices, presumably LSTG and/or PT, would engage in pitch extraction (Barker et al., 2012; Griffiths and Warren, 2002; Hall and Plack, 2009).

AP-related early processing in the temporal lobes has been suggested from ERP/ERF studies in forms of 'AP negativity' (Itoh et al., 2005) and N100 (Hirose et al., 2004; Hirose et al., 2005; Itoh et al., 2005) from temporal electrodes. Such early responses were not observed in the frontal/central electrodes (Elmer et al., 2013; Rogenmoser et al., 2015). Considering the short latency (i.e., 100-150 ms) of AP-related ERPs and the relatively long latency of POR (i.e., later than 200 ms), it may be possible that the pitch extraction process is altered in AP listeners, possibly integrating the extraction of pitch and chroma. Alternatively, it could be also possible that the POR is much earlier in AP listeners than in the general population.

To differentiate the pitch and chroma extraction processes, independent parameterization of periodicity and pitch chroma can be easily achieved using IRN. Using the adaptation paradigm, but instead of white noise using IRN with various degree of periodicity and pitch chroma (Briley et al., 2013), it may possible to disentangle pitch extraction and chromatic categorization.
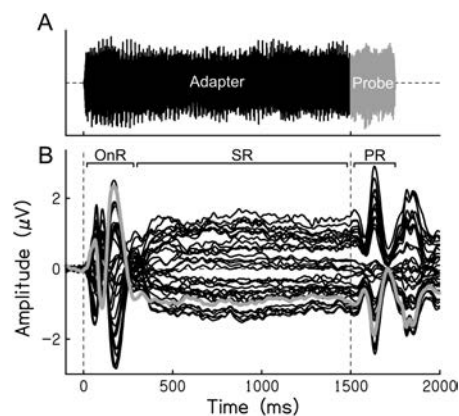
Figure V-2. An example of adaptation paradigm. After presenting an adapter for 1.5 s, a probe stimulus was briefly given for 0.25 s. If the same group of neurons would process both of the adaptor and the probe, the probe response should be similar to the sustained, attenuated response. Conversely, if different groups of neurons would process the adaptor and probe respectively, the probe response should be as strong as the adapter onset response Abbreviations: OnR, onset response; SR, sustained response; PR, probe response.

V-1.3. Chromatic categorization in non-APs

The abovementioned question on the distinction between the pitch extraction and the pitch chroma extraction raises a related question: is the chromatic categorization in AP listeners a similar process to that in non-AP listeners with tonal context? It has been consistently shown that the general population (regardless of AP) perceives pitch as a 2-D helical representation (Briley et al., 2013; Shepard, 1964; Warren et al., 2003a; Warren et al., 2003b). That is, a tone with the frequency of f 'sounds closer' to that of $2f$ then g when $f < g < 2f$. However, the perception of pitch chroma in non-AP listeners could be different from that of AP listeners. As illustrated in Figure V-3, non-AP musicians possess only 'tonal pitch perception'[17]: they cannot directly recognize the pitch chroma of a tone but only the tonal function (e.g., dominant) of a tone in a given tonal context. Only when explicit information is given (e.g., the scale was G major), non-AP listeners can 'compute' its pitch chroma (e.g., the perfect 5th from G is D, thus that tone was D). On the other hand, AP listeners can directly perceive the chroma (e.g., D) and find the key from the tonal context (e.g., that scale is G). Based on that, AP musicians may 'compute' the tonal function (e.g., the interval between G and D is the perfect fifth, which is called dominant) or recognize the tonal function directly with sufficient training in tonal function processing similar to non-AP musicians. It should be noted that musicians in general, including ones with AP, are trained for processing tonal functions and tonal contexts (which allows musical tasks such as transposition or coordination with other instruments).
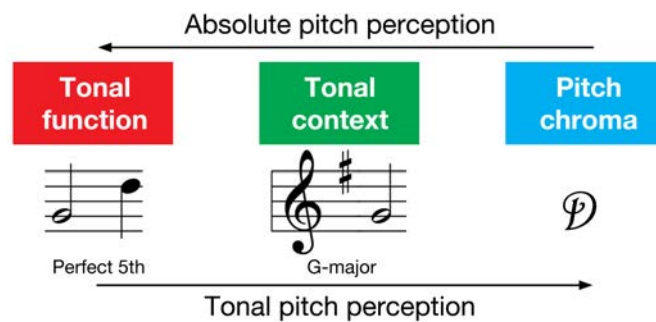
Figure V-3. Schematic view of tonal pitch based on tonal function and absolute pitch based on pitch chroma. Musicians are trained for tonal pitch, thus many AP musicians are also excellent in processing tonal function and context.

---

[17] The described perception is commonly called 'relative pitch' in music education. However, in order to avoid confusion with pitch discrimination, which is also relative pitch processing, here I name it 'tonal pitch'.

Given the opposite direction of pitch processing, the direct recognition of pitch chroma in APs may not be compatible to the relative recognition of tonal function. This could be tested by priming a task where non-AP musicians are required to report the pitch chroma of a given tone with various tonal contexts and in absence of it. Comparing this tonal function recognition in non-AP musicians with direct pitch chroma recognition would provide important evidence on the similarity and dissimilarity of the two processes.

V-1.4. Chromatic Categorization of Ambiguous Pitch

For the categorical perception (or perceptual decision making), ambiguous stimuli provide a means to separate the reflection of physical input and its psychological perception. For instance, Kilian-Hutten and colleagues, using decoding analysis, localized information that distinguishes different interpretations of physically identical sound (morphed phonemes) in the left STP (Kilian-Hutten et al., 2011). Inspired by that study, one could try to differentiate brain activities that are solely involved in chromatic categorization from other neural signals that are related to specific pitch.

In a preliminary study, musicians with AP responded to tones that are 'mistuned' by sub-semitone intervals form standard tuning (i.e., A4 = 440 Hz). The results show stochastic response patterns generally showing sigmoid curves as given in Figure V-4. However, the sigmoid curves would be also possible when AP listeners perceived deviant tones as clearly shifted tones (e.g., "right between the C and C#"), but just because the response device was a piano they pressed either one of two keys randomly. In this case, the assumption of different perceptions does not hold. Therefore, behavioral studies and structured interviews should be done before designing imaging experiments.
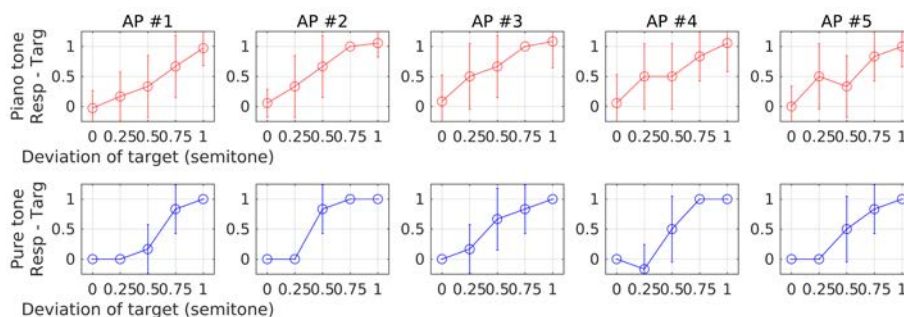

Figure V-4. Psychoacoustic responses to sub-semitone-deviated tones in the timbre of piano tone (upper row, red) and pure (sinusoid) tone (lower row, blue) from 5 musicians with absolute pitch.

VI. References

Achard S, Bullmore E. (2007): Efficiency and cost of economical brain functional networks. PLoS Comput. Biol. 3(2):174-183.

Achard S, Salvador R, Whitcher B, Suckling J, Bullmore E. (2006): A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J. Neurosci. 26(1):63-72.

Adler RJ. 1981. The geometry of random fields. Chichester, UK: John Wiley&Sons.

Ahveninen J, Huang S, Ahlfors SP, Hämäläinen M, Rossi S, Sams M, Jääskeläinen IP. (2016): Interacting parallel pathways associate sounds with visual identity in auditory cortices. Neuroimage 124:858-868.

Akiva-Kabiri L, Henik A. (2012): A Unique Asymmetrical Stroop Effect in Absolute Pitch Possessors. Exp. Psychol. 59(5):272-278.

Altmann CF, Bledowski C, Wibral M, Kaiser J. (2007): Processing of location and pattern changes of natural sounds in the human auditory cortex. Neuroimage 35(3):1192-1200.

Annese J, Pitiot A, Dinov ID, Toga AW. (2004): A myelo-architectonic method for the structural classification of cortical areas. Neuroimage 21(1):15-26.

ANSI. 1973. Psychoacoustical Terminology: American National Standards Institute, New York, NY.

Arnott SR, Alain C. (2011): The auditory dorsal pathway: orienting vision. Neurosci. Biobehav. Rev. 35(10):2162-2173.

Arnott SR, Binns MA, Grady CL, Alain C. (2004a): Assessing the auditory dual pathway model in humans. Neuroimage 22.

Arnott SR, Binns MA, Grady CL, Alain C. (2004b): Assessing the auditory dual-pathway model in humans. Neuroimage 22(1):401-408.

Ashburner J. (2012): SPM: a history. Neuroimage 62(2):791-800.

Ashburner J, Friston KJ. (2000): Voxel-based morphometry - the methods. Neuroimage 11(6):805-821.

Ashburner J, Hutton C, Frackowiak R, Johnsrude I, Price C, Friston K. (1998): Identifying global anatomical differences: deformation-based morphometry. Hum. Brain Mapp. 6(5-6):348-57.

Athos EA, Levinson B, Kistler A, Zemansky J, Bostrom A, Freimer N, Gitschier J. (2007): Dichotomy and perceptual distortions in absolute pitch ability. Proc. Natl. Acad. Sci. U. S. A. 104(37):14795-800.

Bachem A. (1955): Absolute pitch. The Journal of the Acoustical Society of America 27(6):1180-1185.

Baharloo S, Johnston PA, Service SK, Gitschier J, Freimer NB. (1998): Absolute pitch: an approach for identification of genetic and nongenetic components. Am. J. Hum. Genet. 62(2):224-31.

Baharloo S, Service SK, Risch N, Gitschier J, Freimer NB. (2000): Familial aggregation of absolute pitch. Am. J. Hum. Genet. 67(3):755-8.

Bamiou D-E, Musiek FE, Luxon LM. (2003): The insula (Island of Reil) and its role in auditory processing: Literature review. Brain Research Reviews 42(2):143-154.

Barker D, Plack CJ, Hall DA. (2012): Reexamining the Evidence for a Pitch-Sensitive Region: A Human fMRI Study Using Iterated Ripple Noise. Cereb. Cortex 22(4):745-753.

Barrett D, Hall D, Akeroyd M, Summerfield AQ. (2005): The role of Heschl's (HG) gyrus in the analysis of pitch and spatial compactness. J. Cogn. Neurosci.:180-180.

Barrett DJ, Hall DA. (2006): Response preferences for "what" and "where" in human non-primary auditory cortex. Neuroimage 32(2):968-977.

Bavelier D, Levi DM, Li RW, Dan Y, Hensch TK. (2010): Removing brakes on adult brain plasticity: from molecular to behavioral interventions. J. Neurosci. 30(45):14964-71.

Bazin P-L, Weiss M, Dinse J, Schäfer A, Trampel R, Turner R. (2014): A computational framework for ultra-high resolution cortical segmentation at 7Tesla. Neuroimage 93:201-209.

Behzadi Y, Restom K, Liau J, Liu TT. (2007): A component based noise correction

method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37(1):90-101.

Belin P, Fecteau S, Bedard C. (2004): Thinking the voice: neural correlates of voice perception. Trends Cogn. Sci. 8(3):129-135.

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. (2000): Voice-selective areas in human auditory cortex. Nature 403(6767):309-12.

Bendor D, Wang X. (2005): The neuronal representation of pitch in primate auditory cortex. Nature 436(7054):1161-5.

Benjamini Y, Hochberg Y. (1995): Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological):289-300.

Benjamini Y, Yekutieli D. (2005): False discovery rate–adjusted multiple confidence intervals for selected parameters. Journal of the American Statistical Association 100(469):71-81.

Bermudez P, Lerch JP, Evans AC, Zatorre RJ. (2009): Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. Cereb. Cortex 19(7):1583-96.

Bermudez P, Zatorre RJ. (2009a): The absolute pitch mind continues to reveal itself. J. Biol. 8(8):75.

Bermudez P, Zatorre RJ. (2009b): A Distribution of Absolute Pitch Ability as Revealed by Computerized Testing. Music Percept. 27(2):89-101.

Bianco ML. 2015. Understanding and dealing with the loss of absolute pitch as one ages. Oakland, California: Mills College. 41 p.

Binczak S, Eilbeck JC, Scott AC. (2001): Ephaptic coupling of myelinated nerve fibers. Physica D 148(1-2):159-174.

Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. (2004): Neural correlates of sensory and decision processes in auditory object identification. Nat. Neurosci. 7(3):295-301.

Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I. (2008): Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. Nature 451(7175):197-201.

Blackmon K, Halgren E, Barr WB, Carlson C, Devinsky O, DuBois J, Quinn BT, French J, Kuznciecky R, Thesen T. (2011): Individual differences in verbal abilities associated with regional blurring of the left gray and white matter boundary. J. Neurosci. 31(43):15257-63.

Bloch F. (1946): Nuclear induction. Physical review 70(7-8):460.

Box GE, Tiao GC. 2011. Bayesian inference in statistical analysis: John Wiley & Sons.

Briley PM, Breakey C, Krumbholz K. (2013): Evidence for pitch chroma mapping in human auditory cortex. Cereb. Cortex 23(11):2601-2610.

Britz J, Van De Ville D, Michel CM. (2010): BOLD correlates of EEG topography reveal rapid resting-state network dynamics. Neuroimage 52(4):1162-1170.

Brunetti M, Belardinelli P, Caulo M, Del Gratta C, Della Penna S, Ferretti A, Lucci G, Moretti A, Pizzella V, Tartaro A. (2005): Human brain activation during passive listening to sounds from different locations: an fMRI and MEG study. Hum. Brain Mapp. 26(4):251-261.

Bullmore E, Sporns O. (2009): Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10(3):186-198.

Buxton RB. 2009. Introduction to functional magnetic resonance imaging: principles and techniques: Cambridge university press.

Callaghan MF, Freund P, Draganski B, Anderson E, Cappelletti M, Chowdhury R, Diedrichsen J, FitzGerald TH, Smittenaar P, Helms G. (2014): Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging. Neurobiol. Aging 35(8):1862-1872.

Chai XJ, Castañón AN, Öngür D, Whitfield-Gabrieli S. (2012): Anticorrelations in resting state networks without global signal regression. Neuroimage 59(2):1420-1428.

Chee MWL, Zheng H, Goh JOS, Park D, Sutton BP. (2011): Brain Structure in Young and Old East Asians and Westerners: Comparisons of Structural Volume and Cortical Thickness. J. Cogn. Neurosci. 23(5):1065-1079.

Chin CS. (2003): The Development of Absolute Pitch: A Theory Concerning the Roles of Music Training at an Early Developmental Age and Individual Cognitive Style. Psychology of Music 31(2):155-171.

Chung MK. 2012. Computational neuroanatomy: The methods: World Scientific.

Chung MK, Dalton KM, Shen L, Evans AC, Davidson RJ. (2007): Weighted Fourier series representation and its application to quantifying the amount of gray matter. Medical Imaging, IEEE Transactions on 26(4):566-581.

Chung MK, Robbins SM, Dalton KM, Davidson RJ, Alexander AL, Evans AC. (2005): Cortical thickness analysis in autism with heat kernel smoothing. Neuroimage 25(4):1256-1265.

Chung MK, Worsley KJ, Nacewicz BM, Dalton KM, Davidson RJ. (2010): General multivariate linear modeling of surface shapes using SurfStat. Neuroimage 53(2):491-505.

Chung MK, Worsley KJ, Paus T, Cherif C, Collins DL, Giedd JN, Rapoport JL, Evans AC. (2001a): A unified statistical approach to deformation-based morphometry. Neuroimage 14(3):595-606.

Chung MK, Worsley KJ, Taylor J, Ramsay J, Robbins S, Evans AC. (2001b): Diffusion smoothing on the cortical surface. Neuroimage 13(6):S95-S95.

Clark GM. (2013): The multichannel cochlear implant for severe-to-profound hearing loss. Nat. Med. 19(10):1236-9.

Clarke S, Morosan P. 2012. Architecture, connectivity, and transmitter receptors of human auditory cortex. In: Poeppel D, Overath T, Popper AN, Fay RR, editors. The Human Auditory Cortex. New York, NY: Springer New York. p 11-38.

Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ. (2016a): Cortical contributions to the auditory frequency-following response revealed by MEG. Nature Communications 7.

Coffey EBJ, Colagrosso EMG, Lehmann A, Schˆnwiesner M, Zatorre RJ. (2016b): Individual Differences in the Frequency-Following Response: Relation to Pitch Perception. PLoS One 11(3):e0152374.

Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME. (2001): Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. AJNR Am. J. Neuroradiol. 22(7):1326-33.

Dale AM, Fischl B, Sereno MI. (1999): Cortical surface-based analysis - I. Segmentation and surface reconstruction. Neuroimage 9(2):179-194.

Damoiseaux JS, Rombouts SA, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF. (2006): Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. U. S. A. 103(37):13848-53.

De Cheveigne A. 2006. Pitch perception models. In: CJ P, AJ O, RR F, editors. Pitch: neural coding and perception: Springer. p 169-233.

De Luca M, Beckmann CF, De Stefano N, Matthews PM, Smith SM. (2006): fMRI resting state networks define distinct modes of long-distance interactions in the human brain. Neuroimage 29(4):1359-67.

De Martino F, Moerel M, Xu J, van de Moortele PF, Ugurbil K, Goebel R, Yacoub E, Formisano E. (2014): High-Resolution Mapping of Myeloarchitecture In Vivo: Localization of Auditory Areas in the Human Brain. Cereb. Cortex 25(10):3394-3405.

Destrieux C, Fischl B, Dale A, Halgren E. (2010): Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53(1):1-15.

Deutsch D. 1982. The Psychology of music. New York: Academic Press. xvii, 542 p. p.

Deutsch D. 2013. The psychology of music. Amsterdam: Academic Press.

Deutsch D, Henthorn T. (2004): Absolute pitch, speech, and tone language: Some experiments and a proposed framework. Music Percept. 21(3):339-356.

Deutsch D, Henthorn T, Marvin E, Xu HS. (2006): Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. J. Acoust. Soc. Am. 119(2):719-722.

DeWitt I, Rauschecker JP. (2012): Phoneme and word recognition in the auditory ventral stream. Proc. Natl. Acad. Sci. U. S. A. 109(8):E505-14.

Dick F, Tierney AT, Lutti A, Josephs O, Sereno MI, Weiskopf N. (2012): In vivo functional and myeloarchitectonic mapping of human primary auditory areas. J. Neurosci. 32(46):16095-105.

Dohn A, Garza-Villarreal EA, Chakravarty MM, Hansen M, Lerch JP, Vuust P. (2015): Gray- and white-matter anatomy of absolute pitch possessors. Cereb. Cortex 25(5):1379-88.

Dohn A, Garza-Villarreal EA, Heaton P, Vuust P. (2012): Do Musicians with Perfect Pitch Have More Autism Traits than Musicians without Perfect Pitch? An Empirical Study. PLoS One 7(5):e37961.

Dudel J, Menzel R, Schmidt RF. 2013. Neurowissenschaft: Vom Molekül zur Kognition: Springer-Verlag.

Eickhoff S, Walters NB, Schleicher A, Kril J, Egan GF, Zilles K, Watson JD, Amunts K. (2005): High‐resolution MRI reflects myeloarchitecture and cytoarchitecture of human cerebral cortex. Hum. Brain Mapp. 24(3):206-215.

Eklund A, Nichols TE, Knutsson H. (2016): Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the National Academy of Sciences:201602413.

Eldridge LL, Knowlton BJ, Furmanski CS, Bookheimer SY, Engel SA. (2000): Remembering episodes: a selective role for the hippocampus during retrieval. Nat. Neurosci. 3(11):1149-1152.

Elmer S, Meyer M, Jäncke L. (2012): Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. Cereb. Cortex 22(3):650-658.

Elmer S, Rogenmoser L, Kuhnis J, Jancke L. (2015): Bridging the Gap between Perceptual and Cognitive Perspectives on Absolute Pitch. J. Neurosci. 35(1):366-371.

Elmer S, Sollberger S, Meyer M, Jancke L. (2013): An Empirical Reevaluation of Absolute Pitch: Behavioral and Electrophysiological Measurements. J. Cogn. Neurosci. 25(10):1736-1753.

Fauvel B, Groussard M, Chételat G, Fouquet M, Landeau B, Eustache F, Desgranges B, Platel H. (2014): Morphological brain plasticity induced by musical expertise is accompanied by modulation of

functional connectivity at rest. Neuroimage 90:179-188.

Fischl B. (2012a): FreeSurfer. Neuroimage 62(2):774-81.

Fischl B. (2012b): FreeSurfer. Neuroimage 62(2):774-781.

Fischl B, Dale AM. (2000): Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97(20):11050-5.

Fischl B, Sereno MI, Dale AM. (1999): Cortical surface-based analysis - II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9(2):195-207.

Fisher SRA, Fisher RA, Genetiker S, Fisher RA, Genetician S, Fisher RA, Généticien S. 1960. The design of experiments: Oliver and Boyd Edinburgh.

Flandin G, Friston KJ. 2016. Analysis of family-wise error rates in statistical parametric mapping using random field theory. ArXiv e-prints.

Fletcher H. (1924): The Physical Criterion for Determining the Pitch of a Musical Tone. Physical Review 23(3):427-437.

Fortin NJ, Wright SP, Eichenbaum H. (2004): Recollection-like memory retrieval in rats is dependent on the hippocampus. Nature 431(7005):188-191.

Fox MD, Zhang D, Snyder AZ, Raichle ME. (2009): The Global Signal and Observed Anticorrelated Resting State Brain Networks. J. Neurophysiol. 101(6):3270-3283.

Friston KJ, Worsley KJ, Frackowiak R, Mazziotta JC, Evans AC. (1994): Assessing the significance of focal activations using their spatial extent. Hum. Brain Mapp. 1(3):210-220.

Fujimoto K, Polimeni JR, van der Kouwe AJW, Reuter M, Kober T, Benner T, Fischl B, Wald LL. (2014): Quantitative comparison of cortical surface reconstructions from MP2RAGE and multi-echo MPRAGE data at 3 and 7 T. Neuroimage 90:60-73.

Gaab N, Keenan JP, Schlaug G. (2003): The effects of gender on the neural substrates of pitch memory. J. Cogn. Neurosci. 15.

Gaab N, Schulze K, Ozdemir E, Schlaug G. (2006): Neural correlates of absolute pitch differ between blind and sighted musicians. Neuroreport 17(18):1853-1857.

Genovese CR, Lazar NA, Nichols T. (2002):
    Thresholding of statistical maps in
    functional neuroimaging using the false
    discovery rate. Neuroimage 15(4):870-
    878.

Gerig G, Styner M, Jones D, Weinberger D,
    Lieberman J. Shape analysis of brain
    ventricles using spharm; 2001. IEEE. p
    171-178.

Gervain J, Vines BW, Chen LM, Seo RJ, Hensch
    TK, Werker JF, Young AH. (2013):
    Valproate reopens critical-period
    learning of absolute pitch. Front. Syst.
    Neurosci. 7:102.

Geyer S, Weiss M, Reimann K, Lohmann G,
    Turner R. (2011): Microstructural
    Parcellation of the Human Cerebral
    Cortex - From Brodmann's Post-Mortem
    Map to in vivo Mapping with High-Field
    Magnetic Resonance Imaging. Front.
    Hum. Neurosci. 5:19.

Glasser MF, Van Essen DC. (2011): Mapping
    Human Cortical Areas In Vivo Based on
    Myelin Content as Revealed by T1- and
    T2-Weighted MRI. J. Neurosci.
    31(32):11597-11616.

Goebel R. (2012): BrainVoyager — Past, present,
    future. Neuroimage 62(2):748-756.

Goncalves MS, Hall DA, Johnsrude IS, Haggard
    MP. (2001): Can meaningful effective
    connectivities be obtained between
    auditory cortical regions? Neuroimage
    14(6):1353-60.

Good CD, Johnsrude I, Ashburner J, Henson
    RNA, Friston KJ, Frackowiak RSJ.
    (2001): Cerebral asymmetry and the
    effects of sex and handedness on brain
    structure: A voxel-based morphometric
    analysis of 465 normal adult human
    brains. Neuroimage 14(3):685-700.

Gregersen PK, Kowalsky E, Lee A, Baron-Cohen
    S, Fisher SE, Asher JE, Ballard D,
    Freudenberg J, Li W. (2013): Absolute
    pitch exhibits phenotypic and genetic
    overlap with synesthesia. Hum. Mol.
    Genet. 22(10):2097-104.

Greicius MD, Krasnow B, Reiss AL, Menon V.
    (2003): Functional connectivity in the
    resting brain: A network analysis of the
    default mode hypothesis. Proceedings of
    the National Academy of Sciences
    100(1):253-258.

Greicius MD, Menon V. (2004): Default-mode
    activity during a passive sensory task:
    uncoupled from deactivation but

impacting activation. J. Cogn. Neurosci.
    16(9):1484-1492.

Greicius MD, Supekar K, Menon V, Dougherty
    RF. (2009): Resting-State Functional
    Connectivity Reflects Structural
    Connectivity in the Default Mode
    Network. Cereb. Cortex 19(1):72-78.

Greve DN, Fischl B. (2009): Accurate and robust
    brain image alignment using boundary-
    based registration. Neuroimage 48(1):63-
    72.

Griffiths TD, Hall DA. (2012): Mapping Pitch
    Representation in Neural Ensembles
    with fMRI. J. Neurosci. 32(39):13343-
    13347.

Griffiths TD, Warren JD. (2002): The planum
    temporale as a computational hub.
    Trends Neurosci. 25(7):348-353.

Griffiths TD, Warren JD. (2004): What is an
    auditory object? Nat. Rev. Neurosci.
    5(11):887-892.

Grydeland H, Walhovd KB, Tamnes CK, Westlye
    LT, Fjell AM. (2013): Intracortical
    Myelin Links with Performance
    Variability across the Human Lifespan:
    Results from T1-and T2-Weighted MRI
    Myelin Mapping and Diffusion Tensor
    Imaging. J. Neurosci. 33(47):18618-
    18630.

Grydeland H, Westlye LT, Walhovd KB, Fjell
    AM. (2015): Intracortical Posterior
    Cingulate Myelin Content Relates to
    Error Processing: Results from T1- and
    T2-Weighted MRI Myelin Mapping and
    Electrophysiology in Healthy Adults.
    Cereb. Cortex.

Gur RC, Turetsky BI, Matsui M, Yan M, Bilker
    W, Hughett P, Gur RE. (1999): Sex
    differences in brain gray and white
    matter in healthy young adults:
    Correlations with cognitive performance.
    J. Neurosci. 19(10):4065-4072.

Hagmann P, Cammoun L, Gigandet X, Meuli R,
    Honey CJ, Wedeen VJ, Sporns O. (2008):
    Mapping the Structural Core of Human
    Cerebral Cortex. PLoS Biol. 6(7):e159.

Hagmann P, Kurant M, Gigandet X, Thiran P,
    Wedeen VJ, Meuli R, Thiran J-P. (2007):
    Mapping Human Whole-Brain
    Structural Networks with Diffusion MRI.
    PLoS One 2(7):e597.

Hall DA, Plack CJ. (2009): Pitch Processing Sites
    in the Human Auditory Brain. Cereb.
    Cortex 19(3):576-585.

Hallam S. 2015. The Oxford handbook of music psychology. Newyork, NY: Oxford University Press. pages cm p.

Han X, Pham DL, Tosun D, Rettmann ME, Xu C, Prince JL. (2004): CRUISE: cortical reconstruction using implicit surface evolution. Neuroimage 23(3):997-1012.

Hart HC, Palmer AR, Hall DA. (2004): Different areas of human non-primary auditory cortex are activated by sounds with spatial and nonspatial properties. Hum. Brain Mapp. 21(3):178-190.

Hasegawa T, Matsuki K-I, Ueno T, Maeda Y, Matsue Y, Konishi Y, Sadato N. (2004): Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. Cognitive Brain Research 20(3):510-518.

Hashim E, Rowley CD, Grad S, Bock NA. (2015): Patterns of myeloarchitecture in lower limb amputees: an MRI study. Front. Neurosci. 9:15.

Hawley ML, Melcher JR, Fullerton BC. (2005): Effects of sound bandwidth on fMRI activation in human auditory brainstem nuclei. Hear. Res. 204(1):101-110.

Haxby JV. (2012): Multivariate pattern analysis of fMRI: The early beginnings. Neuroimage 62(2):852-855.

He Y, Chen ZJ, Evans AC. (2007): Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. Cereb. Cortex 17(10):2407-19.

Hiltunen T, Kantola J, Abou Elseoud A, Lepola P, Suominen K, Starck T, Nikkinen J, Remes J, Tervonen O, Palva S and others. (2014): Infra-Slow EEG Fluctuations Are Correlated with Resting-State Network Dynamics in fMRI. The Journal of Neuroscience 34(2):356-362.

Hirose H, Kubota M, Kimura I, Yumoto M, Sakakihara Y. (2004): N100m in adults possessing absolute pitch. Neuroreport 15(9):1383-1386.

Hirose H, Kubota M, Kimura I, Yumoto M, Sakakihara Y. (2005): Increased right auditory cortex activity in absolute pitch possessors. Neuroreport 16(16):1775-1779.

Hopf A. (1954): Zur Frage der Konstanz und Abgrenzbarkeit myeloarchitektonischer Rindenfelder (Uniformity and definability of myeloarchitectonic cortical areas). Dtsch. Z. Nervenheilkd. 172(2):188-200.

Hua X, Leow AD, Parikshak N, Lee S, Chiang M-C, Toga AW, Jack CR, Weiner MW, Thompson PM, Initiative AsDN. (2008): Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. Neuroimage 43(3):458-469.

Itoh K, Suwazono S, Arao H, Miyazaki K, Nakada T. (2005): Electrophysiological correlates of absolute pitch and relative pitch. Cereb. Cortex 15(6):760-9.

Jäncke L, Langer N, Hänggi J. (2012): Diminished Whole-brain but Enhanced Peri-sylvian Connectivity in Absolute Pitch Musicians. J. Cogn. Neurosci. 24(6):1447-1461.

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. (2012): Fsl. Neuroimage 62(2):782-90.

Johnsrude IS, Penhune VB, Zatorre RJ. (2000): Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123 (Pt 1):155-63.

Jones DK, Knösche TR, Turner R. (2013): White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. Neuroimage 73:239-254.

Kaas JH, Hackett TA. (1999): 'What' and 'where' processing in auditory cortex. Nat. Neurosci. 2(12):1045-7.

Keenan JP, Thangaraj V, Halpern AR, Schlaug G. (2001): Absolute pitch and planum temporale. Neuroimage 14(6):1402-1408.

Kiebel SJ, Poline J-B, Friston KJ, Holmes AP, Worsley KJ. (1999): Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. Neuroimage 10(6):756-766.

Kilian-Hutten N, Valente G, Vroomen J, Formisano E. (2011): Auditory cortex encodes the perceptual interpretation of ambiguous sound. J. Neurosci. 31(5):1715-20.

Kilner JM, Friston KJ. (2010): Topological inference for EEG and MEG. The Annals of Applied Statistics:1272-1290.

Kim S-G, Chung MK, Schaefer SM, Van Reekum C, Davidson RJ. Sparse shape representation using the laplace-beltrami eigenfunctions and its application to

modeling subcortical structures; 2012. IEEE. p 25-32.

Kim S-G, Jung WH, Kim SN, Jang JH, Kwon JS. (2013): Disparity between dorsal and ventral networks in patients with obsessive-compulsive disorder: evidence revealed by graph theoretical analysis based on cortical thickness from MRI. Front. Hum. Neurosci. 7:302.

Kim S-G, Knösche TR. (2016): Intracortical myelination in musicians with absolute pitch: quantitative morphometry using 7-T MRI. Hum. Brain Mapp.

Kim S-G, Knösche TR. (in preparation): On the functional connectivity of absolute pitch.

Kim S-G, Stelzer J, Bazin P-L, Viehweger A, Knösche T. Group-wise analysis on myelination profiles of cerebral cortex using the second eigenvector of Laplace-Beltrami operator; 2014. IEEE. p 1007-1010.

Ko AL, Darvas F, Poliakov A, Ojemann J, Sorensen LB. (2011): Quasi-periodic Fluctuations in Default Mode Network Electrophysiology. The Journal of Neuroscience 31(32):11728-11732.

Koelsch S, Schulze K, Sammler D, Fritz T, Muller K, Gruber O. (2009): Functional architecture of verbal and tonal working memory: an FMRI study. Hum. Brain Mapp. 30.

Kries Jv. (1892): Über das absolute Gehör. Zeitschrift für Psychologie und Physiologie der Sinnesorgane 3:257-279.

Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhöner B. (2003): Neuromagnetic Evidence for a Pitch Processing Center in Heschl's Gyrus. Cereb. Cortex 13(7):765-772.

Krumhansl CL. (1995): Music psychology and music theory: Problems and prospects. Music Theory Spectrum 17(1):53-80.

Krumhansl CL, Shepard RN. (1979): Quantification of the hierarchy of tonal functions within a diatonic context. J. Exp. Psychol. Hum. Percept. Perform. 5(4):579.

Kusmierek P, Rauschecker JP. (2009): Functional specialization of medial auditory belt cortex in the alert rhesus monkey. J. Neurophysiol. 102(3):1606-22.

Langford E, Schwertman N, Owens M. (2001): Is the property of being positively correlated transitive? The American Statistician 55(4):322-325.

Langner G. (1997): Neural processing and representation of periodicity pitch. Acta Otolaryngol. 117(sup532):68-76.

Langner G, Schreiner CE. (1988): Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J. Neurophysiol. 60(6):1799-1822.

Lee JH, Durand R, Gradinaru V, Zhang F, Goshen I, Kim D-S, Fenno LE, Ramakrishnan C, Deisseroth K. (2010): Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. Nature 465(7299):788-792.

Lee K, Barber D, Paley M, Wilkinson I, Papadakis N, Griffiths P. (2002): Image‐based EPI ghost correction using an algorithm based on projection onto convex sets (POCS). Magn. Reson. Med. 47(4):812-817.

Levitin DJ. (1994): Absolute Memory for Musical Pitch - Evidence from the Production of Learned Melodies. Percept. Psychophys. 56(4):414-423.

Levitin DJ, Rogers SE. (2005): Absolute pitch: perception, coding, and controversies. Trends Cogn. Sci. 9(1):26-33.

Liu XZ, Yan D. (2007): Ageing and hearing loss. The Journal of Pathology 211(2):188-197.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. (2001): Neurophysiological investigation of the basis of the fMRI signal. Nature 412(6843):150-157.

Lopez HS, Liu F, Poole M, Crozier S. (2009): Equivalent Magnetization Current Method Applied to the Design of Gradient Coils for Magnetic Resonance Imaging. IEEE Transactions on Magnetics 45(2):767-775.

Loui P, Li HC, Hohmann A, Schlaug G. (2011): Enhanced Cortical Connectivity in Absolute Pitch Musicians: A Model for Local Hyperconnectivity. J. Cogn. Neurosci. 23(4):1015-1026.

Loui P, Zamm A, Schlaug G. (2012): Enhanced functional networks in absolute pitch. Neuroimage 63(2):632-640.

Luders E, Gaser C, Jancke L, Schlaug G. (2004): A voxel-based approach to gray matter asymmetries. Neuroimage 22(2):656-664.

Luo C, Guo Z-w, Lai Y-x, Liao W, Liu Q, Kendrick KM, Yao D-z, Li H. (2012): Musical Training Induces Functional Plasticity in Perceptual and Motor

Networks: Insights from Resting-State fMRI. PLoS One 7(5):e36568.

Lutti A, Dick F, Sereno MI, Weiskopf N. (2014): Using high-resolution quantitative mapping of R1 as an index of cortical myelination. Neuroimage 93:176-188.

MacDonald D, Kabani N, Avis D, Evans AC. (2000): Automated 3-D Extraction of Inner and Outer Surfaces of Cerebral Cortex from MRI. Neuroimage 12(3):340-356.

Mahling F. (1926): Das Problem der" Audition colorée": eine historisch-kritische Untersuchung. Arch. Gesamte Psychol.

Marques JP, Gruetter R. (2013): New developments and applications of the MP2RAGE sequence--focusing the contrast and high spatial resolution R1 mapping. PLoS One 8(7):e69294.

Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele PF, Gruetter R. (2010): MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage 49(2):1271-81.

McGee AW, Yang Y, Fischer QS, Daw NW, Strittmatter SM. (2005): Experience-Driven Plasticity of Visual Cortex Limited by Myelin and Nogo Receptor. Science 309(5744):2222-2226.

Menon V. (2011): Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn. Sci. 15(10):483-506.

Meyer M. (1899): Is the Memory of Absolute Pitch Capable of Development by Training? Psychol. Rev. 6(5):514-516.

Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. (2006): Influence of musical and psychoacoustical training on pitch discrimination. Hear. Res. 219(1–2):36--47.

Miyazaki K. (1988): Musical Pitch Identification by Absolute Pitch Possessors. Percept. Psychophys. 44(6):501-512.

Miyazaki K. (1990): The Speed of Musical Pitch Identification by Absolute-Pitch Possessors. Music Percept. 8(2):177-188.

Miyazaki K. (2004a): How well do we understand absolute pitch? Acoust. Sci. Technol. 25(6):426-432.

Miyazaki K. (2004b): Recognition of transposed melodies by absolute-pitch possessors. Jpn. Psychol. Res. 46(4):270-282.

Miyazaki K, Makomaska S, Rakowski A. (2012): Prevalence of absolute pitch: A comparison between Japanese and Polish music students. J. Acoust. Soc. Am. 132(5):3484-3493.

Miyazaki K, Ogawa Y. (2006): Learning absolute pitch by children: A cross-sectional study. Music Percept. 24(1):63-78.

Miyazaki Ki. (1989): Absolute Pitch Identification: Effects of Timbre and Pitch Region. Music Percept. 7(1):1-14.

Moerel M, De Martino F, Formisano E. (2012): Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity. J. Neurosci. 32(41):14205-16.

Moerel M, De Martino F, Santoro R, Ugurbil K, Goebel R, Yacoub E, Formisano E. (2013): Processing of Natural Sounds: Characterization of Multipeak Spectral Tuning in Human Auditory Cortex. J. Neurosci. 33(29):11888-11898.

Moerel M, De Martino F, Santoro R, Yacoub E, Formisano E. (2015): Representation of pitch chroma by multi-peak spectral tuning in human auditory cortex. Neuroimage 106:161-9.

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.

Morosan P, Schleicher A, Amunts K, Zilles K. (2005): Multimodal architectonic mapping of human superior temporal gyrus. Anat. Embryol. (Berl.) 210(5-6):401-406.

Mugler JP, Brookeman JR. (1990): Three‐dimensional magnetization‐prepared rapid gradient‐echo imaging (3D MP RAGE). Magn. Reson. Med. 15(1):152-157.

Neuweiler G, Schmidt S. (1993): Audition in echolocating bats. Curr. Opin. Neurobiol. 3(4):563-569.

Neyman J, Pearson ES. 1992. On the problem of the most efficient tests of statistical hypotheses: Springer.

Nichols TE, Holmes AP. (2002): Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15(1):1-25.

Nieuwenhuys R. 2013a. The myeloarchitectonic studies on the human cerebral cortex of

the Vogt-Vogt school, and their significance for the interpretation of functional neuroimaging data. Microstructural Parcellation of the Human Cerebral Cortex: Springer. p 55-125.

Nieuwenhuys R. (2013b): The myeloarchitectonic studies on the human cerebral cortex of the Vogt-Vogt school, and their significance for the interpretation of functional neuroimaging data. Brain Struct. Funct. 218(2):303-352.

Norman-Haignere S, Kanwisher N, McDermott JH. (2013): Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex. J. Neurosci. 33(50):19451-69.

O'Brien KR, Kober T, Hagmann P, Maeder P, Marques J, Lazeyras F, Krueger G, Roche A. (2014): Robust T1-weighted structural brain imaging and morphometry at 7T using MP2RAGE. PLoS One 9(6):e99676.

Oechslin MS, Imfeld A, Loenneker T, Meyer M, Jancke L. (2009): The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. Front. Hum. Neurosci. 3:76.

Ohnishi T, Matsuda H, Asada T, Aruga M, Hirakata M, Nishikawa M, Katoh A, Imabayashi E. (2001): Functional anatomy of musical perception in musicians. Cereb. Cortex 11(8):754-760.

Oldfield RC. (1971): The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9(1):97-113.

Owens B. (2016): Data sharing: Access all areas. Nature 533(7602):S71-S72.

Pan WJ, Thompson GJ, Magnuson ME, Jaeger D, Keilholz S. (2013): Infraslow LFP correlates to resting-state fMRI BOLD signals. Neuroimage 74:288-97.

Pandya DN, Hallett M, Mukherjee SK. (1969): Intra-and interhemispheric connections of the neocortical auditory system in the rhesus monkey. Brain Res. 14(1):49-65.

Papagno C, Miracapillo C, Casarotti A, Romero Lauro LJ, Castellano A, Falini A, Casaceli G, Fava E, Bello L. (2010): What is the role of the uncinate fasciculus? Surgical removal and proper name retrieval. Brain.

Patel AD, Balaban E. (2001): Human pitch perception is reflected in the timing of stimulus-related cortical activity. Nat. Neurosci. 4(8):839-844.

Patterson RD, Allerhand MH, Giguere C. (1995): Time‐domain modeling of peripheral auditory processing: A modular architecture and a software platform. The Journal of the Acoustical Society of America 98(4):1890-1894.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. (2002): The processing of temporal pitch and melody information in auditory cortex. Neuron 36(4):767-76.

Penagos H, Melcher JR, Oxenham AJ. (2004): A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J. Neurosci. 24(30):6810-5.

Plack CJ, Barker D, Hall DA. (2014): Pitch coding and pitch processing in the human brain. Hear. Res. 307:53-64.

Plack CJ, Oxenham AJ, Fay RR. 2006. Pitch: neural coding and perception: Springer Science & Business Media.

Poeppel D. (2003): The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Communication 41(1):245-255.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. (2013): Steps toward optimizing motion artifact removal in functional connectivity MRI; a reply to Carp. Neuroimage 76(1):439-441.

Power JD, Schlaggar BL, Petersen SE. (2015): Recent progress and outstanding issues in motion correction in resting state fMRI. Neuroimage 105:536-551.

Qian L, Zhang Y, Zheng L, Shang Y, Gao JH, Liu Y. (2015): Frequency dependent topological patterns of resting-state brain networks. PLoS One 10(4):e0124681.

Qiu A, Younes L, Miller MI. (2008): Intrinsic and extrinsic analysis in computational anatomy. Neuroimage 39(4):1803-1814.

Rajagopalan S, Mukherjee D, Mohler ER. 2005. Manual of vascular diseases: Lippincott Williams & Wilkins.

Rauschecker JP. (2011): An expanded role for the dorsal auditory pathway in sensorimotor control and integration. Hear. Res. 271(1):16-25.

Rauschecker JP. (2012): Ventral and dorsal streams in the evolution of speech and language. Front. Evol. Neurosci. 4:7.

Rauschecker JP. (2015): Auditory and visual cortex of primates: a comparison of two sensory systems. Eur. J. Neurosci. 41(5):579-85.

Rauschecker JP, Tian B. (2000): Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 97(22):11800-6.

Rauschecker JP, Tian B, Hauser M. (1995): Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268(5207):111-114.

Remedios R, Logothetis NK, Kayser C. (2009): An Auditory Region in the Primate Insular Cortex Responding Preferentially to Vocal Communication Sounds. The Journal of Neuroscience 29(4):1034-1045.

Reuter M, Wolter F-E, Shenton M, Niethammer M. (2009): Laplace–Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. Computer-Aided Design 41(10):739-755.

Rivier F, Clarke S. (1997): Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. Neuroimage 6(4):288-304.

Roche A. (2011): A Four-Dimensional Registration Algorithm With Application to Joint Correction of Motion and Slice Timing in fMRI. IEEE Trans. Med. Imaging 30(8):1546-1554.

Rogenmoser L, Elmer S, Jancke L. (2015): Absolute Pitch: Evidence for Early Cognitive Facilitation during Passive Listening as Revealed by Reduced P3a Amplitudes. J. Cogn. Neurosci. 27(3):623-637.

Russell SM, Golfinos JG. (2003): Amusia following resection of a Heschl gyrus glioma - Case report. J. Neurosurg. 98(5):1109-12.

Saad ZS, Gotts SJ, Murphy K, Chen G, Jo HJ, Martin A, Cox RW. (2012): Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. Brain connectivity 2(1):25-32.

Sakakibara A. (2014): A longitudinal study of the process of acquiring absolute pitch: A practical report of training with the 'chord identification method'. Psychology of Music 42(1):86-111.

Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, Bullmore E. (2005a): Neurophysiological architecture of functional magnetic resonance images of human brain. Cereb. Cortex 15(9):1332-42.

Salvador R, Suckling J, Schwarzbauer C, Bullmore E. (2005b): Undirected graphs of frequency-dependent functional connectivity in whole brain networks. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360(1457):937-46.

Sasai S, Homae F, Watanabe H, Sasaki AT, Tanabe HC, Sadato N, Taga G. (2014): Frequency-specific network topologies in the resting human brain. Front. Hum. Neurosci. 8:1022.

Schellenberg EG, Trehub SE. (2003): Good Pitch Memory Is Widespread. Psychol. Sci. 14(3):262-266.

Schlaug G, Jancke L, Huang Y, Steinmetz H. (1995): In vivo evidence of structural brain asymmetry in musicians. Science 267.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. (2002): Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat. Neurosci. 5(7):688-694.

Schneider P, Sluming V, Roberts N, Scherg M, Goebel R, Specht HJ, Dosch HG, Bleeck S, Stippich C, Rupp A. (2005): Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. Nat. Neurosci. 8(9):1241-7.

Schönwiesner M, Zatorre RJ. (2008): Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. Exp. Brain Res. 187(1):97-105.

Schulze K, Gaab N, Schlaug G. (2009): Perceiving pitch absolutely: Comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci. 10.

Schulze K, Mueller K, Koelsch S. (2013): Auditory stroop and absolute pitch: an fMRI study. Hum. Brain Mapp. 34(7):1579-90.

Seither-Preisler A, Parncutt R, Schneider P. (2014): Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. The Journal of Neuroscience 34(33):10937-10949.

Sereno MI, Lutti A, Weiskopf N, Dick F. (2013): Mapping the human cortical surface by combining quantitative T(1) with retinotopy. Cereb. Cortex 23(9):2261-8.

Shafee R, Buckner RL, Fischl B. (2015): Gray matter myelination of 1555 human brains using partial volume corrected MRI images. Neuroimage 105:473-485.

Shattuck DW, Leahy RM. (2002): BrainSuite: an automated cortical surface identification tool. Med. Image Anal. 6(2):129-142.

Shen S, Li J, Casaccia-Bonnefil P. (2005): Histone modifications affect timing of oligodendrocyte progenitor differentiation in the developing rat brain. J. Cell Biol. 169(4):577-89.

Shepard RN. (1964): Circularity in Judgments of Relative Pitch. The Journal of the Acoustical Society of America 36(12):2346-2353.

Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. (2012): Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb. Cortex 22(1):158-65.

Shmuel A, Leopold DA. (2008): Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: implications for functional connectivity at rest. Hum. Brain Mapp. 29(7):751-761.

Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE. (1997): Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. J. Cogn. Neurosci. 9(5):648-663.

Siegel JA. (1974): Sensory and verbal coding strategies in subjects with absolute pitch. J. Exp. Psychol. 103(1):37-44.

Siegel JA, Siegel W. (1977): Absolute identification of notes and intervals by musicians. Percept. Psychophys. 21.

Siegel W. (1972): Memory effects in the method of absolute judgment. J. Exp. Psychol. 94(2):121.

Sigalovsky IS, Fischl B, Melcher JR. (2006): Mapping an intrinsic MR property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences. Neuroimage 32(4):1524-37.

Sladky R, Friston KJ, Tröstl J, Cunnington R, Moser E, Windischberger C. (2011): Slice-timing effects and their correction in functional MRI. Neuroimage 58(2):588-594.

Smith AM, Lewis BK, Ruttimann UE, Frank QY, Sinnwell TM, Yang Y, Duyn JH, Frank JA. (1999): Investigation of low frequency drift in fMRI signal. Neuroimage 9(5):526-533.

Smith NA, Schmuckler MA. (2008): Dial A440 for absolute pitch: Absolute pitch memory by non-absolute pitch possessors. The Journal of the Acoustical Society of America 123(4):EL77-EL84.

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP and others. (2013): Resting-state fMRI in the Human Connectome Project. Neuroimage 80:144-68.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR and others. (2009): Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. U. S. A. 106(31):13040-13045.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM and others. (2006): Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. Neuroimage 31(4):1487-1505.

Smith SM, Miller KL, Moeller S, Xu J, Auerbach EJ, Woolrich MW, Beckmann CF, Jenkinson M, Andersson J, Glasser MF and others. (2012): Temporally-independent functional modes of spontaneous brain activity. Proc. Natl. Acad. Sci. U. S. A. 109(8):3131-6.

Smith SM, Nichols TE. (2009): Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44(1):83-98.

Spreng RN, Mar RA, Kim AS. (2009): The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-

analysis. J. Cogn. Neurosci. 21(3):489-510.

Stecker GC, Harrington IA, Middlebrooks JC. (2005): Location coding by opponent neural populations in the auditory cortex. PLoS Biol. 3(3):e78.

Stehling MK, Turner R, Mansfield P. (1991): Echo-planar imaging: magnetic resonance imaging in a fraction of a second. Science 254(5028):43-50.

Stuber C, Morawski M, Schafer A, Labadie C, Wahnert M, Leuze C, Streicher M, Barapatre N, Reimann K, Geyer S and others. (2014): Myelin and iron concentration in the human brain: A quantitative study of MRI contrast. Neuroimage 93:95-106.

Styner M, Oguz I, Xu S, Brechbühler C, Pantazis D, Levitt JJ, Shenton ME, Gerig G. (2006): Framework for the statistical shape analysis of brain structures using SPHARM-PDM. The insight journal(1071):242.

Sun FT, Miller LM, D'Esposito M. (2004): Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. Neuroimage 21(2):647-658.

Suriadi MM, Usui K, Tottori T, Terada K, Fujitani S, Umeoka S, Usui N, Baba K, Matsuda K, Inoue Y. (2015): Preservation of absolute pitch after right amygdalohippocampectomy for a pianist with TLE. Epilepsy Behav. 42:14-17.

Sylvester CM, Barch DM, Corbetta M, Power JD, Schlaggar BL, Luby JL. (2013): Resting State Functional Connectivity of the Ventral Attention Network in Children With a History of Depression or Anxiety. J. Am. Acad. Child Adolesc. Psychiatry 52(12):1326-1336.

Takeuchi AH, Hulse SH. (1993): Absolute Pitch. Psychol. Bull. 113(2):345-361.

Tardif E, Clarke S. (2001): Intrinsic connectivity of human auditory areas: a tracing study with DiI. Eur. J. Neurosci. 13(5):1045-1050.

Taylor JE, Worsley KJ. (2007): Detecting sparse signals in random fields, with an application to brain mapping. Journal of the American Statistical Association 102(479):913-928.

Thambisetty M, Wan J, Carass A, An Y, Prince JL, Resnick SM. (2010): Longitudinal changes in cortical thickness associated with normal aging. Neuroimage 52(4):1215-1223.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. (2001): Functional specialization in rhesus monkey auditory cortex. Science 292(5515):290-3.

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. (2001): An integrated software suite for surface-based analyses of cerebral cortex. J. Am. Med. Inform. Assoc. 8(5):443-459.

van Essen DC, Maunsell JHR. (1980): Two-dimensional maps of the cerebral cortex. The Journal of Comparative Neurology 191(2):255-281.

van Essen DC, Zeki SM. (1978): The topographic organization of rhesus monkey prestriate cortex. The Journal of Physiology 277:193-226.

Van Hedger SC, Heald SL, Koch R, Nusbaum HC. (2015): Auditory working memory predicts individual differences in absolute pitch learning. Cognition 140:95-110.

Van Hedger SC, Heald SL, Nusbaum HC. (2016): What the [bleep]? Enhanced absolute pitch memory for a 1000Hz sine tone. Cognition 154:139-150.

von Economo C, Horn L. (1930): Über Windungsrelief. Maße und Rindenarchitektonik der.

von Economo CF, Koskinas GN. 1925. Die cytoarchitektonik der hirnrinde des erwachsenen menschen: J. Springer.

von Kriegstein K, Warren JD, Ives DT, Patterson RD, Griffiths TD. (2006): Processing the acoustic effect of size in speech sounds. Neuroimage 32(1):368-375.

Waehnert M, Dinse J, Weiss M, Streicher M, Waehnert P, Geyer S, Turner R, Bazin P-L. (2014a): Anatomically motivated modeling of cortical laminae. Neuroimage 93:210-220.

Waehnert MD, Dinse J, Weiss M, Streicher MN, Waehnert P, Geyer S, Turner R, Bazin PL. (2014b): Anatomically motivated modeling of cortical laminae. Neuroimage 93:210-220.

Wakefield GH. Mathematical representation of joint time-chroma distributions; 1999. International Society for Optics and Photonics. p 637-645.

Wallace MN, Johnston PW, Palmer AR. (2002): Histochemical identification of cortical areas in the auditory region of the

human brain. Exp. Brain Res. 143(4):499-508.

Wallentin M, Nielsen AH, Friis-Olivarius M, Vuust C, Vuust P. (2010): The Musical Ear Test, a new reliable test for measuring musical competence (vol 20, pg 188, 2010). Learning and Individual Differences 20(6):705-705.

Ward J, Huckstep B, Tsakanikos E. (2006): Sound-Colour Synaesthesia: to What Extent Does it Use Cross-Modal Mechanisms Common to us All? Cortex 42(2):264-280.

Ward W, Burns E. 1982. Absolute Pitch. In: Deutsch D, editor. The Psychology of Music. New York: Academic Press.

Ward WD. (1985): Absolute tonality versus absolute piano. The Journal of the Acoustical Society of America 78(S1):S76-S76.

Warren JD, Griffiths TD. (2003a): Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. J. Neurosci. 23(13):5799-5804.

Warren JD, Griffiths TD. (2003b): Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. The journal of neuroscience 23(13):5799-5804.

Warren JD, Uppenkamp S, Patterson RD, Griffiths TD. (2003a): Analyzing pitch chroma and pitch height in the human brain. Annual N Y Acad Sci 999:212-4.

Warren JD, Uppenkamp S, Patterson RD, Griffiths TD. (2003b): Separating pitch chroma and pitch height in the human brain. Proc. Natl. Acad. Sci. U. S. A. 100(17):10038-42.

Warren JE, Wise RJ, Warren JD. (2005): Sounds do-able: auditory–motor transformations and the posterior temporal plane. Trends Neurosci. 28(12):636-643.

Wengenroth M, Blatow M, Bendszus M, Schneider P. (2010): Leftward lateralization of auditory cortex underlies holistic sound perception in Williams syndrome. PLoS One 5(8):e12326.

Wengenroth M, Blatow M, Heinecke A, Reinhardt J, Stippich C, Hofmann E, Schneider P. (2014): Increased volume and function of right auditory cortex as a marker for absolute pitch. Cereb. Cortex 24(5):1127-37.

Wiegrebe L, Meddis R. (2004): The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. The Journal of the Acoustical Society of America 115(3):1207-1218.

Wilson SJ, Lusher D, Wan CY, Dudgeon P, Reutens DC. (2009): The neurocognitive components of pitch processing: insights from absolute pitch. Cereb. Cortex 19(3):724-32.

Wong YK, Wong AC-N. (2014): Absolute pitch memory: Its prevalence among musicians and dependence on the testing context. Psychonomic bulletin & review 21(2):534-542.

Woo C-W, Krishnan A, Wager TD. (2014): Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. Neuroimage 91:412-419.

Worsley K, Andermann M, Koulis T, MacDonald D, Evans A. (1999): Detecting changes in nonisotropic images. Hum. Brain Mapp. 8(2-3):98-101.

Worsley K, Taylor JE, Carbonell F, Chung M, Duerden E, Bernhardt B, Lyttelton O, Boucher M, Evans A. (2009a): SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. Neuroimage 47:S102.

Worsley KJ, Evans AC, Marrett S, Neelin P. (1992): A three-dimensional statistical analysis for CBF activation studies in human brain. J. Cereb. Blood Flow Metab. 12(6):900-18.

Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. (1996): A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4(1):58-73.

Worsley KJ, Taylor JE, Carbonell F, Chung MK, Duerden E, Bernhardt B, Lyttelton O, Boucher M, Evans AC. (2009b): SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. Neuroimage 47, Supplement 1:S102.

Worsley KJ, Taylor JE, Tomaiuolo F, Lerch J. (2004): Unified univariate and

multivariate random field theory. Neuroimage 23 Suppl 1:S189-95.

Wundt WM. 1904. Principles of Physiological Psychology. Titchener EB, translator: Sonnenschein.

Yost WA. (1996): Pitch of iterated rippled noise. The Journal of the Acoustical Society of America 100(1):511-518.

Zald DH, McHugo M, Ray KL, Glahn DC, Eickhoff SB, Laird AR. (2014): Meta-analytic connectivity modeling reveals differential functional connectivity of the medial and lateral orbitofrontal cortex. Cereb. Cortex 24(1):232-248.

Zatorre RJ. (1989): Intact absolute pitch ability after left temporal lobectomy. Cortex 25(4):567-80.

Zatorre RJ. (1998): Functional specialization of human auditory cortex for musical processing. Brain 121:1817-1818.

Zatorre RJ. (2003): Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nat. Neurosci. 6(7):692-5.

Zatorre RJ, Belin P. (2001): Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11(10):946-53.

Zatorre RJ, Perry DW, Beckett CA, Westbury CF, Evans AC. (1998): Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. Proc. Natl. Acad. Sci. U. S. A. 95(6):3172-7.

Zatorre RJ, Zarate JM. 2012. Cortical Processing of Music. 261-294 p.

Zimmer U, Lewald J, Erb M, Karnath H-O. (2006): Processing of auditory spatial cues in human cortex: an fMRI study. Neuropsychologia 44(3):454-461.