

From the Department of Clinical Neuroscience  
Karolinska Institutet, Stockholm, Sweden

and

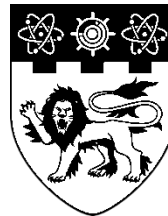
Lee Kong Chien School of Medicine  
Nanyang Technological University, Singapore

# **TRACKING SELECTIVE ATTENTION IN A MUSICAL COCKTAIL**

Cassia Low Manting



**Karolinska  
Institutet**



**NANYANG  
TECHNOLOGICAL  
UNIVERSITY**  
**SINGAPORE**

Stockholm 2021

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.

Printed by Universitetsservice US-AB

© Cassia Low Manting, 2021

ISBN 978-91-8016-298-2

Cover illustration by Zkorejo

Tracking Selective Attention in a Musical Cocktail

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

**Cassia Low Manting**

The thesis will be defended in public at Karolinska Institutet, Stockholm, Sweden, 10<sup>th</sup> of September 2021

*Principal Supervisor:*

Daniel Lundqvist  
Associate Professor  
Karolinska Institutet  
Department of Clinical Neuroscience  
The National Facility for  
Magnetoencephalography (NatMEG)

*Co-supervisor(s):*

Professor Balázs Zoltán Gulyás  
President's Chair Professor  
Nanyang Technological University  
Lee Kong Chian School of Medicine  
Centre for Neuroimaging Research at NTU

Fredrik Ullén  
Professor  
Karolinska Institutet  
Department of Neuroscience

Lau Møller Andersen  
PhD  
Aarhus University  
Aarhus Institute of Advanced Studies

*Opponent:*

Lauri Parkkonen  
Professor  
Aalto University  
Department of Neuroscience and Biomedical  
Engineering

*Examination Board:*

Elvira Brattico  
Professor  
Aarhus University  
Department of Clinical Medicine  
Center for Music In the Brain

Eugene I. Dairianathan  
Associate Professor  
Nanyang Technological University  
Department of Visual & Performing Arts

Tomas Furmark  
Professor  
Uppsala Universitet  
Department of Psychology  
Division of Emotion Psychology



For my teachers

## ACKNOWLEDGEMENTS

The journey here has been nothing short of incredible – challenging and stressful at times, but nonetheless incredibly fun, satisfying and rewarding. This journey would not have been possible without the countless support that I have most gratefully received from the wonderful people who have so kindly walked beside me. I would like to thank, first and foremost, my beloved friends and colleagues, family, and mentors, for your patience, company and support.

To many of my friends and family faraway in sunny Singapore (and also **Rachel** in Australia): I miss every one of you, and thank you for always offering a listening ear to whine to despite the distance and time difference. I look forward to reuniting with you soon! To my **brother** and **parents**, thank you for always having my back and showering me with care from across the world. I hope I have made you proud! To **Charlene**, **KK**, **Timo**, **Samuel**, **Jedi**, **Xuan**, I am very sorry for missing all the weddings and parties that I should have been there for, especially when covid restrictions came into play last year. I will be sure to make up for it whenever I can! To my university roommate/partner-in-crime **Hanny**, I look forward to meeting your little one, Rae! I also look forward to meeting **Charlene**'s little ninja, and **Samuel**'s otah. To my absolutely best childhood friend for more than a quarter century who is closer than a sister to me, **Li Rong** (a.k.a **Lelong**), congratulations on your engagement! I am absolutely elated for you and I will definitely be there for you wedding next year! For the rest of my dear friends whom I have missed a large chunk of your lives, **Hui Dan**, **Fernice**, **Adeline Liew**, **Adeline Lee**, **Jiankai**, my cousins **Ying** and **Tian Hui**, I am very happy and blessed to have you in my life and let's really catch up when I am back!

To my tight circle of amazing friends here in Stockholm (and some in Germany, Spain and America) who have coloured my life in beautiful ways, thank you for the fikas, dinners, tea parties, dancing, book clubs, maybe-dates, computer games, cooking experiments, dog walking and light-hearted, heartwarming as well as downright hilarious conversations. **Marcus**, **Kristin**, **Lulu**, **Leona**, **Jessica**, **Dilana**, **Andrea**, **Lisette**, **Laurent**, **Zach**, **Nicholas**, **Georg**, **Patrik**, **Johanna** and especially **John**, thank you for motivating me and keeping me company, I could not have made it this far without all of you. This last month of thesis writing had been the most stressful and intense period, and I cannot think of a better way to have written it other than with your support and companionship. I am truly blessed to have you in my life. To those of you that have left Stockholm in pursuit of further dreams, I hope the best for you and we must stay in touch often (and yes, I will visit but you must lend me your couch)!

To my supervisors, colleagues and ex-colleagues, **Daniel**, **Fredrik**, **Balazs**, **Lau**, **Mikkel**, **Allison**, my heartfelt thanks for generously sharing your wise input and knowledge, encouragement, and for experiencing this journey first-hand with me. I have learnt a tremendous amount from all of you, both professionally and personally, which I believed has made me into a much better scientist today. Also, thank you for letting me pick very dull Spanish learning podcasts during helium refills. To **Daniel**, a huge thank you for helping so

much with organizing the dissertation, I really appreciate it! To **Karin**, thank you for so graciously agreeing to chair my upcoming dissertation, I am really honoured! Also, thank you to my thesis advisory committee for their useful input throughout the years.

To both my universities **NTU** and **Karolinska Institutet**, thank you for granting me the opportunity to embark on this eye-opening joint doctoral programme. I hope that I can continue working in collaborative efforts in the future, which I truly believe are great for the advancement of scientific research. Thank you **Lars** for taking care of us, students from the joint programme, always checking to make sure that we are alive, happy and that our studies were running smoothly. I would also like to thank **NTU** for providing me with so many opportunities and resources throughout my undergraduate years which have helped to build my experiences in the research field that proved fruitful later during my PhD. To **May Chong** from the SBS office, and **Johanna Ackemar** from KI, thank you for always assisting with all the administration throughout these years. It would have been much more chaotic and stressful for me if I had not received your help in navigating through the necessary logistics and admin of this joint programme.

Last but not least, to all my teachers and mentors who have catalyzed my unquenchable thirst for knowledge and inspired my passion for science, thank you for being such amazing role models and always believing in me. This thesis is primarily dedicated to all of you who have taught and mentored me from primary school to university. I still remember my dear primary one form teacher, **Mrs Catherine Yew**, who was always caring and encouraging towards me. She was the first who kickstarted my love for learning that eventually grew into the foundations that motivate my scientific endeavours today, and I will be forever grateful to her. Teaching is indeed a noble profession and teachers do change lives.

While this beautiful chapter of mine is soon coming to a close, I am thrilled to welcome what fascination the future has in store for me. Whether from near or afar, I hope that each and every one of you will continue to share my exciting journeys that lie ahead.

Cassia Low Manting  
2 August 2021

# ABSTRACT

The cocktail party effect refers to mankind's ability to focus attention on a single sound within noisy or complex auditory environments, in order to extract the most behaviourally-relevant information present. To investigate this cognitive phenomenon in a precise manner, we used frequency-tagging to separate and identify neural auditory steady-state responses (ASSRs), which are specific to the driving frequency of their respective auditory sources, from a mixture elicited by multiple simultaneous stimuli. Participants directed attention to one out of a few musical streams as we examined how top-down selective attention to the pitch and timing of frequency-tagged musical tones influences the corresponding ASSR. In addition, bottom-up attention towards musical tones was also manipulated with salient changes in pitch. Using magnetoencephalography, we showed that the attentional enhancement of the ASSR can already be observed at general sensor level. Distributed source analyses revealed multiple ASSR sources distributed across the frontal, parietal and temporal cortices, with each of these areas modulated to different extents by selective attention. Notably, we uncovered the existence of ASSR attentional modulation in the frontal, parietal and insular lobes, which has not been previously reported in literature. Our results indicated that the ASSR enhancements from top-down and bottom-up attention were strongest at the frontal and temporal lobes respectively. The ASSR and its attentional modulation also displayed sensitivity towards individual musicality, demonstrating positive correlations with musical sophistication.

Surprisingly, we discovered stark differences in cortical representation as well as character between the ASSR during and after the stimulation period. The ASSR during, but not after, stimulation was affected by the stimulus properties, selective attention, and participants' level of musical sophistication. Moreover, the ASSR during stimulation was generated primarily from temporal sources whereas the post-stimulus ASSR originated mainly from the frontal cortex. We also assessed how the complexity of the stimulus cocktail affects the ASSR by comparing ASSRs elicited from simultaneous musical streams with ASSRs generated by streams that were completely separated in time. With two simultaneous streams, suppression of the ASSR power was observed, which furthermore varied across the cortical space. This caused a shift in the ASSR source distribution from temporal-dominance (separated streams) to proportionally stronger activity in the frontal and centro-parietal cortices (simultaneous streams). With the accumulation of evidence highlighting the differences between ASSR sources in the frontal and temporal regions, our results collectively advocate that these sources are characteristically unique, functionally distinct and largely independent from one another.

Taken together, this thesis revealed new aspects of the ASSR and ways to effectively extract the effect of selective attention and its interaction with individual auditory experiences such as musical training. Importantly, this work advocates a novel 'beyond the temporal cortex' perspective on ASSR modulation and advances the study of human cognition towards more complex and naturalistic soundscapes using frequency-tagging.



## LIST OF SCIENTIFIC PAPERS

- I. **Manting, C. L.**; Andersen, L. M.; Gulyas, B.; Ullén, F.; Lundqvist, D., Attentional modulation of the auditory steady-state response across the cortex. *NeuroImage* **2020**, *217*, 116930.
- II. **Manting, C. L.**; Gulyas, B.; Ullén, F.; Lundqvist, D., Auditory steady-state responses during and after a stimulus: Cortical sources, and the influence of attention and musicality. *NeuroImage* **2021**, *233*, 117962.
- III. **Manting, C. L.**; Gulyas, B.; Ullén, F.; Lundqvist, D., Selective Attention to Frequency-tagged Concurrent Melodies: The Interplay between Suppression, Top-down and Bottom-up Attention. *Manuscript in Preparation*

### *Scientific papers not included in the thesis*

**Vinding, M. C.**; Eriksson, A.; Manting, C. L.; Waldthaler, J.; Ferreira, D.; Ingvar, M.; Svenningsson, P.; Lundqvist, D., Different features of the cortical sensorimotor rhythms are uniquely linked to the severity of specific symptoms in Parkinson's disease. *medRxiv* **2021**, 2021.06.27.21259592. *pre-print*



# TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	THE COCKTAIL PARTY EFFECT & SELECTIVE AUDITORY ATTENTION	1
1.1.1	Brief History of the Cocktail Party Effect	1
1.1.2	Neuroimaging Approaches to the Cocktail Party Problem	1
1.1.3	Mechanisms Underlying Selective Attention & the Hierarchy of Auditory Processing	4
1.2	MAGNETOENCEPHALOGRAPHY	6
1.2.1	Basic MEG Theory & Instrumentation	6
1.2.2	Analysis of MEG Signals	7
1.3	FREQUENCY TAGGING & THE AUDITORY STEADY-STATE RESPONSE	10
1.3.1	The ASSR & its Origins	10
1.3.2	Technical Considerations for Using Frequency-Tagging	13
1.3.3	Frequency-Tagging Multiple Stimuli	15
1.3.4	ASSR in Attentional research	17
1.4	UTILIZING MUSIC FOR NEUROSCIENCE	19
1.4.1	Anatomical & Functional Neural Changes Associated with Musical Training	19
1.4.2	Transferability of Musical Abilities to General Cognitive Skills	23
<b>2</b>	<b>RESEARCH AIMS &amp; THESIS FRAMEWORK</b>	<b>25</b>
<b>3</b>	<b>MATERIALS &amp; METHODS</b>	<b>27</b>
3.1	PARTICIPANTS	27
3.2	THE GOLDSMITH MUSICAL SOPHISTICATION INDEX	27
3.3	EXPERIMENTAL TASKS & TRAINING	27
3.4	STIMULI	29
3.5	DATA ACQUISITION	30
3.6	DATA PROCESSING	30
3.6.1	Pre-processing	30
3.6.2	Sensor-space analysis	30
3.6.3	Source-space analysis	31
3.7	ETHICAL CONSIDERATIONS	31
<b>4</b>	<b>RESULTS &amp; DISCUSSION</b>	<b>33</b>
4.1	BEHAVIORAL RESULTS	33
4.2	ASSR SOURCES AND CHARACTERISTICS	33
4.3	ATTENTIONAL MODULATION OF THE ASSR	37
4.4	ASSR AND MUSICAL EXPERIENCE	40
<b>5</b>	<b>LIMITATIONS &amp; RECOMMENDATIONS FOR FUTURE WORK</b>	<b>43</b>
5.1	Generalizability of findings	43
5.2	Towards naturalistic Musical stimuli	43
5.3	Training Selective Attention & Other Cognitive Abilities	45

5.4	ASSR Across Space and Time .....	45
<b>6</b>	<b>CONCLUSIONS</b> .....	<b>47</b>
<b>7</b>	<b>REFERENCES</b> .....	<b>49</b>

## LIST OF ABBREVIATIONS

AM	Amplitude modulation
ASSR	Auditory Steady-State Response
BOLD	Blood Oxygenation Level-Dependent
CPE	Cocktail Party Effect
EEG	Electroencephalography
ERF/ERP	Event-Related Field/Event-Related Potential
$f_c$	Carrier Frequency
$f_m$	Modulation Frequency
fMRI	Functional Magnetic Resonance Imaging
FM	Frequency Modulation
ISI	Inter-Stimulus Interval
m	Modulation Depth
MDT	Music Development Tracking
MEG	Magnetoencephalography
MASTER	Multiple Auditory Steady-State Responses
MRI	Magnetic Resonance Imaging
MSI	Musical Sophistication Index
PSD	Power Spectral Density
ROI	Region of Interest
SQUID	Superconducting Quantum Interference Devices



# 1 INTRODUCTION

## 1.1 THE COCKTAIL PARTY EFFECT & SELECTIVE AUDITORY ATTENTION

“Everyone knows what attention is. It is taking possession of the mind, in clear and vivid form, of one out of what seems several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies a withdrawal from some things in order to deal effectively with others.” — William James (1890)

### 1.1.1 Brief History of the Cocktail Party Effect

In light of the brain’s limited capacity to fully process all incoming sensory information at the same time, mankind’s ability to exclusively attend to a particular sound within a complex and dynamic environment is highly essential as it allows us to extract and process the most important information amidst a noisy cocktail of sounds. This perceptual effect of selective attention was first coined the “Cocktail party effect (CPE)” by Cherry in 1953 from a study inspired by our inherent capacity to follow and understand a single conversation while concurrently filtering out irrelevant sounds in the background, such as the loud clinks of glasses, bursts of laughter, and other simultaneous conversations, that are competing for our attention<sup>1</sup>. On a related note, directional shifts of attention can be triggered voluntarily by top-down factors, or involuntarily via bottom-up mechanisms<sup>2</sup>. For example, the brain employs top-down mechanisms to focus on the teacher in a chaotic classroom, or a single instrument during a band performance. However, when a phone rings, a baby cries or an alarm sounds, attention is automatically directed to the salient sound in a bottom-up manner.

Following Cherry’s defining work, numerous related studies rapidly emerged and today, the field of CPE has built up quite a substantial library of literature seeking to explain the underlying cognitive mechanisms<sup>3-4</sup>. However, like most research studying human cognition, selective auditory attention has its ancestral roots stemming from psychophysical experiments. Hence in earlier times, the vast majority of this research belongs to the realm of psychophysics and it was not until later years that neuroimaging was introduced into the playing field. The advent of cutting-edge neuroimaging technology invited a surge of CPE-related neuroimaging literature, bringing along eye-opening insights into the neural correlates of the CPE. For the first time in history, scientists have the means to uncover where, when and how selective attention occurs in the brain.

### 1.1.2 Neuroimaging Approaches to the Cocktail Party Problem

Functional magnetic resonance imaging (fMRI)

With the emergence of functional magnetic resonance imaging (fMRI), scientists have managed to localize, within centimetre range, the anatomical brain areas that are responsible for selective auditory attention. fMRI is a non-invasive technique which measures the blood oxygenation level-dependent (BOLD) contrast across conditions at different regions of the

brain. To put simply, a larger BOLD signal reflects relatively higher oxygen consumption locally which strongly correlates with the underlying electric field potential, thereby indicating more activity<sup>5-6</sup>. However, it is important to take note that the BOLD measure is not a direct measure of neural electrical activity but rather one that makes use of its strong correlation with the local field potential, and reflects the total pre- and post-synaptic activities summed across mass neuronal populations<sup>5-6</sup>. While fMRI studies constitute a large fraction of neuroimaging work in human cognitive research at present<sup>7-8</sup>, the sluggishness of the BOLD signal limits its temporal resolution to the order of seconds. Hence, in order to track neural mechanisms underlying many cognitive processes (including attention), which often take place within the millisecond range, scientists have to resort to other more direct measurement techniques such as magneto- and electro-encephalography (MEG and EEG respectively).

#### Magnetoencephalography (MEG) & Electroencephalography (EEG)

Between these two electrophysiological techniques, EEG dominates the scene with an early head start (with the first EEG studies conducted in the 1920s and 1930s<sup>9</sup>), followed by MEG gaining popularity in later years. Both EEG and MEG are direct measures of electric<sup>10</sup> and magnetic fields<sup>11</sup> respectively on the brain surface, generated by the movement of ions due to biochemical processes occurring in neurons. The architecture and parallel arrangement of pyramidal cortical neurons is such that synchronous activity of these neurons produces a measurable signal, which is proportional to the sum of the momentary post-synaptic potentials generated<sup>12</sup>. In the study of brain function, the methods of MEG and EEG hold a unique position of being the only non-invasive techniques capable of directly measuring neuronal activity with a millisecond time resolution, and are thus optimally suited for elucidating the spatiotemporal sequences of brain activity that underlie different brain functions. Since the scalp, skull and cerebrospinal fluid have varying electrical conductivities but relatively constant magnetic permeability, compared to MEG, EEG signals experience significantly more distortion as the electric fields travel across these brain regions before reaching the surface where they are recorded by sensors<sup>12</sup>. From the two methods, MEG and EEG, MEG therefore has a unique capacity in terms of a spatio-temporally fine-resolved measure of ongoing brain activity. EEG and MEG are typically used non-invasively, although electrodes can also be implanted intracranially, usually in the cases of epileptic patients, for higher signal-to-noise ratio and more accurate source localization<sup>13-14</sup> (more details on MEG provided in Chapter 1.2). Because of their respective advantages and disadvantages, especially in terms of the spatial-temporal resolution trade-off, fMRI is well-suited for localizing precisely where an effect occurs whereas MEG and EEG are more useful for understanding the dynamics of neural activities overtime. It may be helpful to note that combinations of multiple neuroimaging techniques can sometimes be used together to study neural cognitive processes like the CPE<sup>15</sup>.

The bulk of magneto-electroencephalographic approaches to cognitive research, including selective auditory attention, can be broadly classified into one of two very well-established



approaches, namely, studies using (i) event-related potentials (or their magnetic equivalent event-related fields) or (ii) following responses<sup>16</sup>.

#### Event-related potentials

Typically, one can record event-related potentials (ERPs) that are time-locked to a stimulus event such as an onset or offset<sup>17-18</sup>. Due to the weak signals produced by the brain, ERPs can only be obtained via signal averaging across multiple trials to dilute the background noise in order to achieve a higher signal-to-noise ratio. This approach allows inferences to be made about the location and more exclusively, the precise timing, of sensory and cognitive processes<sup>19</sup>. The ERP method is compatible with a large variety of stimulus, ranging from the simplest tones<sup>20</sup> and speech stimuli<sup>21-22</sup> to more sophisticated musical chords<sup>23</sup>. Many components of the ERP can be studied individually, or in relation to other components. Prominent examples in auditory research demonstrated that the transient ERP components, including the P20-50<sup>24</sup>, N100<sup>25</sup>, sustained potential<sup>13</sup>, as well as their magnetic equivalents<sup>26</sup>, are enhanced by selective attention.

A considerable amount of research on auditory selective attention, in particularly bottom-up attention, has made use of a change-specific ERP component known as the mismatch negativity (MMN) response, that is elicited by sound deviants from regularity, for example, a sudden change in loudness, pitch, timbre, duration or spatial location of a repeating tone (see Näätänen 2007<sup>27</sup> for a comprehensive review). Moreover, the MMN can also be evoked by naturalistic stimuli such as deviant speech tokens<sup>28</sup> or footsteps within a noisy traffic environment<sup>29</sup>. The MMN is said reflect the automatic switching of attention to potentially important events occurring in the background acoustic environment, hindering performance on the main task<sup>30</sup>. Although the MMN is generated via bottom-up mechanisms, it can also be modulated by top-down attention<sup>31-32</sup>.

Even though ERPs can contribute valuable insights into the principles behinds the CPE, discretizing auditory stimuli into individual events may not be the best way to reflect natural processing of continuous speech in a real-world context. Importantly, it is difficult and sometimes even impossible to separate simultaneous ERPs originating from different sources, which is often the case for natural auditory stimuli. Instead, other researchers sought for answers using continuous measurements of oscillatory neural responses that phase-lock to incoming auditory stimuli, otherwise known as following responses.

#### Oscillatory following responses

Following responses generated in the brain are oscillatory in nature and can follow the amplitude envelope or the frequency of a sound. Should the sound envelope or frequency vary periodically in time, the corresponding following response becomes periodic as well to give an auditory steady-state response (ASSR)<sup>16</sup>. A thorough description of the ASSR is provided in Chapter 1.3.

### 1.1.3 Mechanisms Underlying Selective Attention & the Hierarchy of Auditory Processing

In neuroscience literature, selective attention has been observed to modulate the neural representation of a stimulus in several different ways. For instance, the effect of attention can manifest as an enhancement of neural activity<sup>24-26, 33</sup>, increased neural connectivity<sup>34-35</sup> and synchronization<sup>36</sup>, or more robust neural encoding<sup>37-38</sup> of the attended stimulus relative to the unattended stimulus. These effects have been observed in several areas including the auditory cortex, auditory association regions, as well as frontal and parietal regions<sup>24-26, 33-38</sup>. To explain some of these observations, many researchers proposed a gain and suppression mechanism by which selective attention operates, wherein attention enhances the neural representation of the attended stimulus while suppressing that of the unattended distractors<sup>25-26, 39-42</sup>. This interpretation is supported by studies utilizing single-unit neurophysiology<sup>43</sup> as well as broad-scale neuroimaging methodologies<sup>25-26, 39, 41-42</sup>, is able to accommodate a wide range of stimuli (even within each neuroimaging and sensory modality), and appears to be applicable across sensory modalities including vision, audition, and touch<sup>39, 41-42, 44-49</sup>.

Attentional modulation of neural activity has been found in several regions across both cortical<sup>33, 50-52</sup> and sub-cortical<sup>53</sup> structures of the brain. Typically, the effect of selective attention manifests in the auditory cortex (AC)<sup>26, 54</sup>, although its extent and precise location depend on specific experimental and stimulus parameters. For instance, in dichotic selective listening experiments, the Heschl's gyrus and planum polare were activated more strongly on the contralateral hemisphere with respect to the attended ear compared to the ipsilateral hemisphere<sup>53, 55-56</sup>. Moreover, attention selection based on spatial ("where") versus non-spatial ("what") features of auditory objects also modulate different parts of the brain<sup>57</sup>. Ahveninen et al.<sup>15</sup> demonstrated that spatial attention towards speech tokens exerted an effect on the AC anteriorly whereas attention towards phonetic features modulated posterior regions, approximately 30 ms later. Beyond the AC, studies contrasting the regions involved in spatial attention versus pitch-based attention found higher activity in the frontal eye fields, premotor, as well as parietal areas that was associated with spatial attention, and higher activity in the superior temporal gyrus during attention to pitch<sup>58-59</sup>.

More recently, neuroimaging evidence suggests that the attentional modulation of neural activity seems to, at least partly, operate on the level of processing required to discriminate the attended stimulus from other sounds during attentional selection, and consequently the effect of attention manifests in regions where neurons are optimized to carry out such level of processing<sup>60</sup>. This view is consistent with the hierarchical structure of auditory processing in the primate cortex, in which increasingly complex information is processed by different anatomical structures of the brain at progressive latencies from subcortical areas to the auditory core, and subsequently towards the belt, parabelt, secondary auditory regions and beyond<sup>61-68</sup>. Through this pathway, earlier processes resolve simpler acoustic features, such as tonotopy, in lower-level regions of the hierarchy (i.e. primary auditory regions), whereas for more complex or

abstract perceptual features (e.g. semantic and content-based) that require further processing to discriminate between stimuli, the process occurs later in higher-level regions ascending the auditory pathway.

## 1.2 MAGNETOENCEPHALOGRAPHY

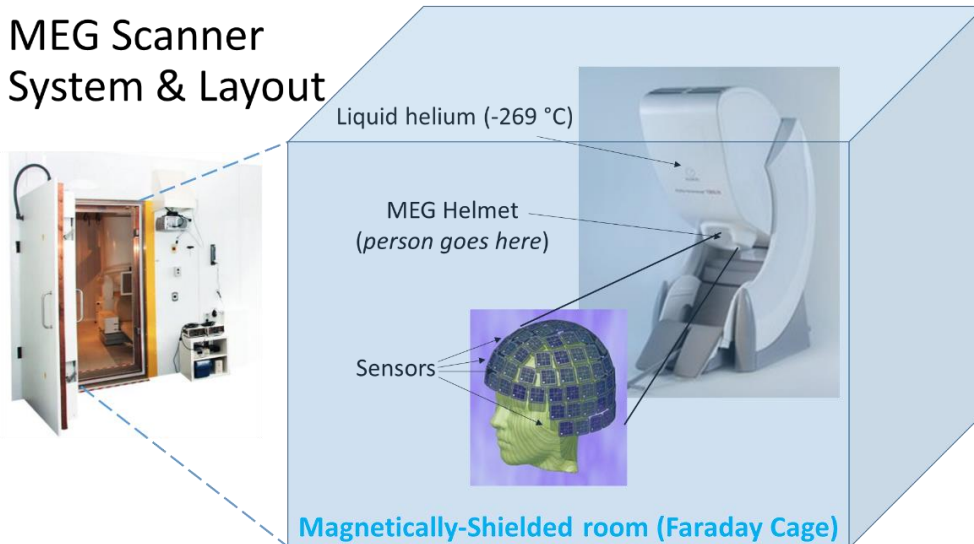
Magnetoencephalography (MEG) is a technique that uses an array of ultra-sensitive magnetic sensors to detect minute fluctuations in magnetic field strength arising from neural activity inside the brain. With regards to research in auditory cognition, MEG (& EEG) holds the unique advantage of being a mechanically silent (as opposed to fMRI scanner noise that can interfere with the sound stimulus during recordings), comfortable and non-invasive technique that provides a direct measure of neural activity with millisecond time resolution. Occupying only a small fraction of all neuroimaging research today, the field of MEG is far from being saturated, highlighting its immense potential for growth and innovative discoveries<sup>7-8</sup>.

### 1.2.1 Basic MEG Theory & Instrumentation

MEG measures the extracranial magnetic induction generated by the sum of post-synaptic potentials across 10 000 to 50 000 pyramidal cortical neurons that are synchronously activated and arranged in a parallel fashion<sup>8, 12</sup>. While MEG detects the primary intracellular currents in the apical dendrites of these pyramidal neurons, EEG measures the extracellular volume currents that compensate for the primary currents instead<sup>69</sup>. Since the neural magnetic flux is extremely weak in magnitude, on the scale of 10 to 100 fT<sup>70</sup> (that is about 10 to 100 million times smaller than Earth's magnetic field), ultra-sensitive sensors with superconducting property called superconducting quantum interference devices (SQUIDs) are used to pick up these signals<sup>11</sup>. Modern MEG systems contain more than 300 SQUIDs which are sampled at up to 30kHz simultaneously and arranged in helmet-shaped arrays to cover the whole cortex<sup>8</sup>. Furthermore, many systems use multiple types of SQUID sensors, namely magnetometers and gradiometers, within the same array. Planar gradiometers are maximally sensitive to magnetic fields directly under the sensors, and this sensitivity decreases steeply with distance. Magnetometers, on the other hand, have a circular sensitivity distribution and are sensitive to magnetic fields that originate from further distances<sup>71-73</sup>. The SQUIDs are contained in a thermally insulated tank, known as a dewar, and immersed in liquid helium at -269 °C – a condition required to maintain superconductivity which helps to minimize thermal noise and optimize data quality. The MEG system is housed in a magnetically-shielded room (i.e. a Faraday cage) with multi-layered walls that weigh approximately 7 tonnes.

As MEG signals experience decreased sensitivity with depth, superficial cortical sources are detected up to 100 times stronger than deeper subcortical sources with identical source strengths<sup>74-75</sup>. Nonetheless, modern advances in experimental design and signal extraction techniques have made it possible to study magnetic sources in deep brain regions such as the subcortical<sup>76-79</sup> areas, brainstem<sup>80</sup> and cerebellum<sup>81</sup>. With regards to the direction of neuronal current flow, MEG is more sensitive to tangential currents usually found in the gyri, than radial currents that typically flow in the sulci<sup>8, 82-83</sup>.

## MEG Scanner System & Layout



### 1.2.2 Analysis of MEG Signals

Conventionally, source level analysis of MEG data is used to estimate the neural origin of measured extracranial neural activity in brain space. The analytical steps required to transform raw MEG data into source level estimations can be broadly categorized into four processes, namely, (i) preprocessing, (ii) trial segmentation and averaging of event-related fields or power spectral density (PSD) estimations, (iii) forward modelling, and (iv) inverse modelling. In the (i) preprocessing step, noisy data is usually discarded or corrected. Sources of noise can arise from physiological causes, such as cardiac activity, muscle activity, eye-blinks and saccades, or from artefacts, like line noise, vibrations from nearby construction work, or movement of metal objects<sup>83-84</sup> (see Fig. 1A). Often, electrodes are used to simultaneously record ocular activity as well as cardiac activity alongside the MEG recording, to serve as references for the noise removal process later. Filters can also be applied in this step to exclude unwanted line noise frequency and to include only the frequencies of interest. Other types of noise that are more difficult to remove require more sophisticated statistical techniques known as source separation methods<sup>85-86</sup>. Often, (ii) trial-averaging is carried out on the preprocessed data to increase the signal-to-noise ratio (SNR) of the neurophysiological signal-of-interest relative to unwanted noise and unrelated brain signals. The two most common types of neural activity studied using MEG are Event-related fields (ERFs), which are responses that are time-locked to a stimulus event, and oscillatory responses (e.g. alpha/beta activity) that are continuous and may be related or unrelated to any stimulus (e.g. resting state data). ERFs are obtained by signal averaging the time-domain data across multiple trials, whereas oscillatory responses are interpreted in the frequency-domain after applying PSD estimations such as the Fourier transform. Using information regarding the physical parameters of the head (e.g. geometry, magnetic permeability) and MEG sensors, the (iii) forward model estimates the activity projected at the extracranial sensor locations from possible sources of activity inside the brain. The (iv) inverse model on the other hand, estimates the projection in the opposite direction from sensor to

source space. The biggest obstacle that MEG researchers face in estimating the inverse model is that there is an infinite number of solutions that can possibly explain the data (i.e. the solution is non-unique), and the solutions are noise sensitive<sup>83, 87</sup>. Anatomically and physiologically meaningful constraints are thus imposed to narrow down the number of possible solutions, while regularization helps to reduce the sensitivity to data noise. MEG source estimation approaches typically fall into one of the three categories: parametric source models<sup>83</sup> (e.g. dipole-fitting), distributed current estimates<sup>88-89</sup> (e.g. minimum-norm estimates), and scanning methods<sup>90-92</sup> (e.g. beamformer). Detailed explanation of each of these approaches is beyond the scope of this thesis but the reader is encouraged to refer to the cited references for further information. Additionally, anatomical information from magnetic resonance imaging (MRI) scans can be integrated into MEG analysis to provide a high-resolution brain template for better visualization and spatial accuracy<sup>8, 69, 83</sup>.

Over the years, MEG research has accumulated a growing library of contributions to the field of systems and behavioural neuroscience as well as clinical applications. This includes compelling discoveries in speech and language processing<sup>93-94</sup>, music perception<sup>95-96</sup>, consciousness<sup>97-98</sup>, attention<sup>99-100</sup>, memory<sup>101-102</sup>, epilepsy<sup>103</sup>, psychological diseases<sup>104-105</sup>, movement disorders<sup>106-107</sup> and more. As new advances in MEG recording and analytical techniques continue to develop and emerge, such as optically-pumped magnetometers<sup>108-110</sup> (OPM) systems and high-Tc SQUIDS<sup>111-112</sup>, there is much excitement to look forward to in the burgeoning field of MEG in the coming years.

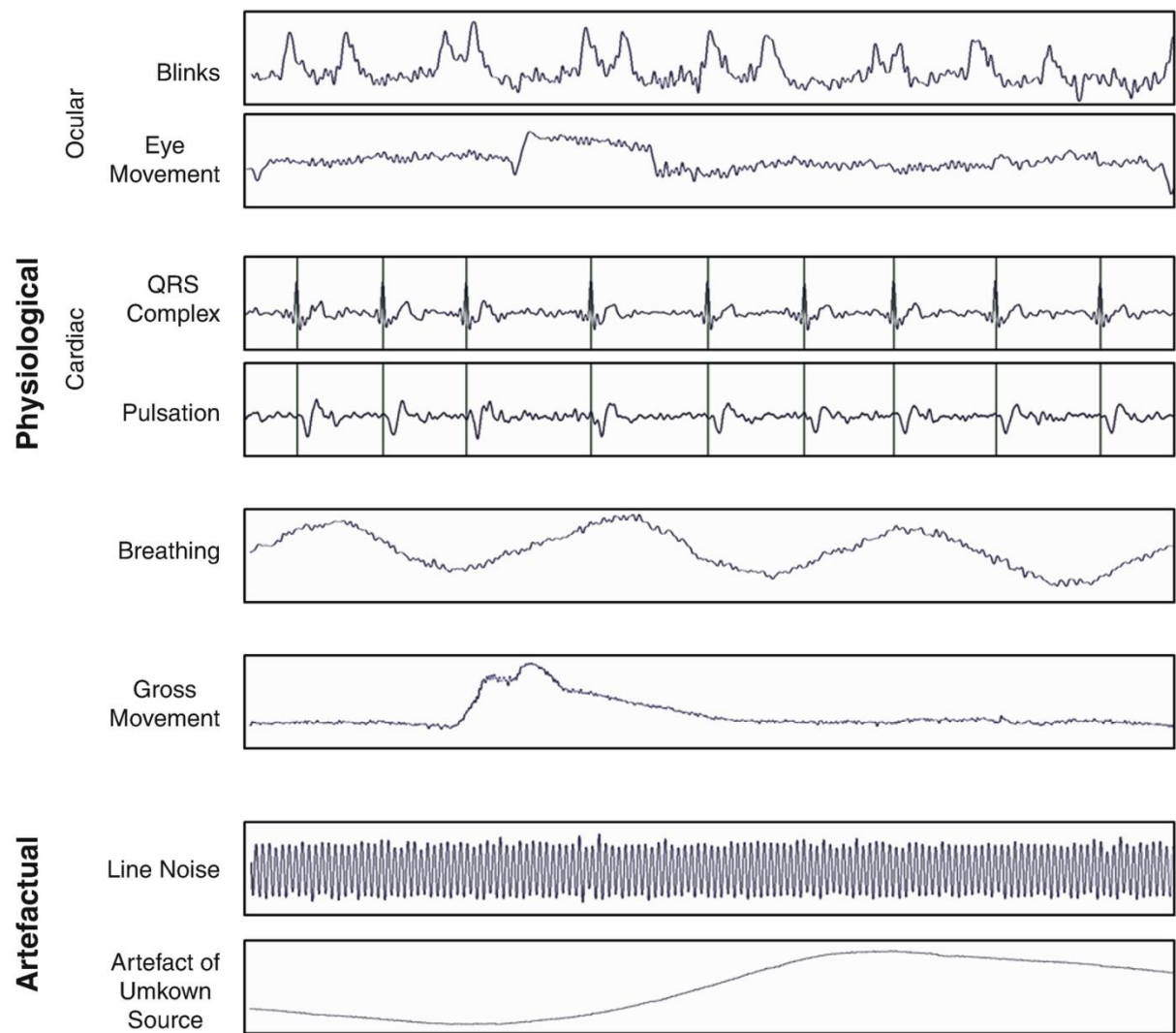


Figure 1A. Examples of common physiological (upper panels) and artefactual (lower panels) noise sources recorded from single MEG channels over a period of 10 s (image reproduced from Coffman 2020<sup>69</sup>).

### 1.3 FREQUENCY TAGGING & THE AUDITORY STEADY-STATE RESPONSE

Over the past decades, while neuroimaging technologies have unearthed important discoveries about the human brain, the neural mechanisms underlying many cognitive processes are still not completely understood. In the study of selective attention (among other processes), particularly when involving complex stimulation environments, this gap in knowledge is partially attributed to the difficulties associated with isolating specific brain activity that stems from one out of many simultaneous stimuli. Although challenging, the separation and identification of mixed neural signals is necessary as it allows researchers to track and study each individual stimulus-specific activity, such as activity generated by an individual instrument within an orchestra or a single speaker in a noisy cocktail party setting. Offering a simple, efficient and precise solution to this problem, frequency-tagging allows the generation of neural activity that is uniquely labelled according to its driving stimulus, known as the auditory steady-state response (ASSR), thereby allowing individual labelled neural responses within mixtures to be extracted and identified back to their respective sources.

#### 1.3.1 The ASSR & its Origins

The ASSR can be described as an oscillatory neural response that continuously phase-locks to the intrinsic fundamental frequency (and higher harmonics in some cases) of the driving stimulus over the time period of stimulus presentation<sup>113-114</sup>. It can be recorded using electro- and magnetoencephalography<sup>113-114</sup> and exhibits a maximum cortical power response at approximately 40 Hz in humans<sup>114-115</sup>. The ASSR stabilizes at around 200 ms from when the stimulus begins, and continues to oscillate at a constant phase throughout the duration of the stimulus<sup>116</sup>. The constituent discrete frequency components of the ASSR can be retrieved from recorded MEG or EEG data in a straightforward manner using power spectral density (PSD) estimation techniques such as Fourier analysis, and the resultant power spectrum would exhibit a peak specifically at the stimulus driving frequency. This driving frequency can be manipulated by adjusting the stimulus presentation rate<sup>117</sup>, or the modulation frequency ( $f_m$ ) of an amplitude- or frequency-modulated sound<sup>115</sup>. The focus of this thesis – amplitude modulation (AM) frequency-tagging, is a handy way to adjust the stimulus driving frequency, and consequently the ASSR frequency, while retaining much of the stimulus properties such as pitch, duration, and timbre (partially, depending on frequency-tagging parameters). AM frequency-tagging is done by increasing and decreasing the amplitude of the sound envelope (i.e. volume) at a precise rate defined by the modulation frequency (refer to Fig. 1B). Frequency-modulated (FM) frequency-tagging on the other hand, systematically modulates the carrier frequency of the stimulus according to  $f_m$  (FM frequency-tagging will not be covered in this thesis but interested readers may refer to Picton 2003<sup>115</sup> for more information). Frequency-tagging offers a simple, efficient and precise method to disentangle the processing of multiple sound streams presented simultaneously, since the neural ASSR activity corresponding to each stream can be distinguished by its unique  $f_m$  during analysis<sup>115, 118</sup> (see section 1.3.3 for further



details on multiple ASSRs). Today, the ASSR has found itself useful in many applications ranging from hearing assessments<sup>119-121</sup> to attention research<sup>13, 51, 99, 122</sup> and schizophrenia<sup>123-124</sup>.

## AM Frequency-Tagging

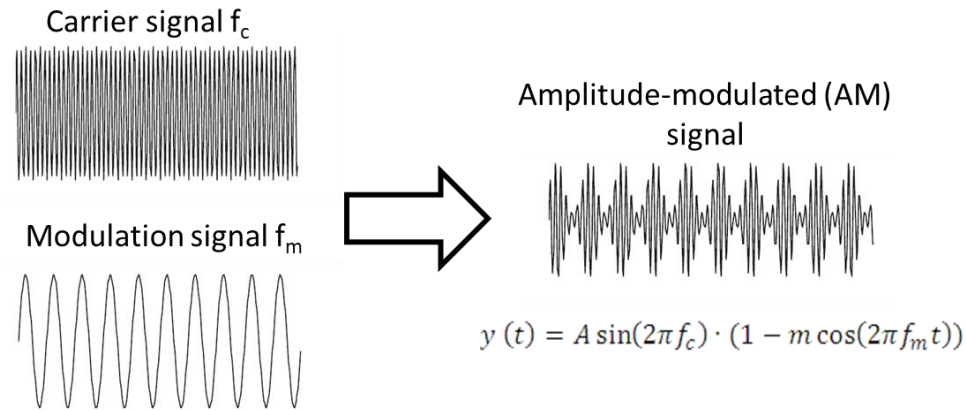


Figure 1B. AM frequency-tagging of a signal. The amplitude envelope of the carrier signal is modulated systematically according to the modulation frequency  $f_m$ . The resultant AM signal can be described by the equation displayed above. Image edited from Luo 2006<sup>125</sup>.

The origin of the ASSR is a topic that has sparked much debate between two main opposing camps. The first argument states that the ASSR is a linear superposition of consecutive transient ERPs, more specifically the mid-latency response or the transient gamma band response, and reflects the same neural processes<sup>126</sup> (Fig.1C). Proponents of this view use deconvolution techniques to show that the enhanced amplitude near 40 Hz is due to a better-fit overlap of the gamma band components<sup>127-128</sup>.

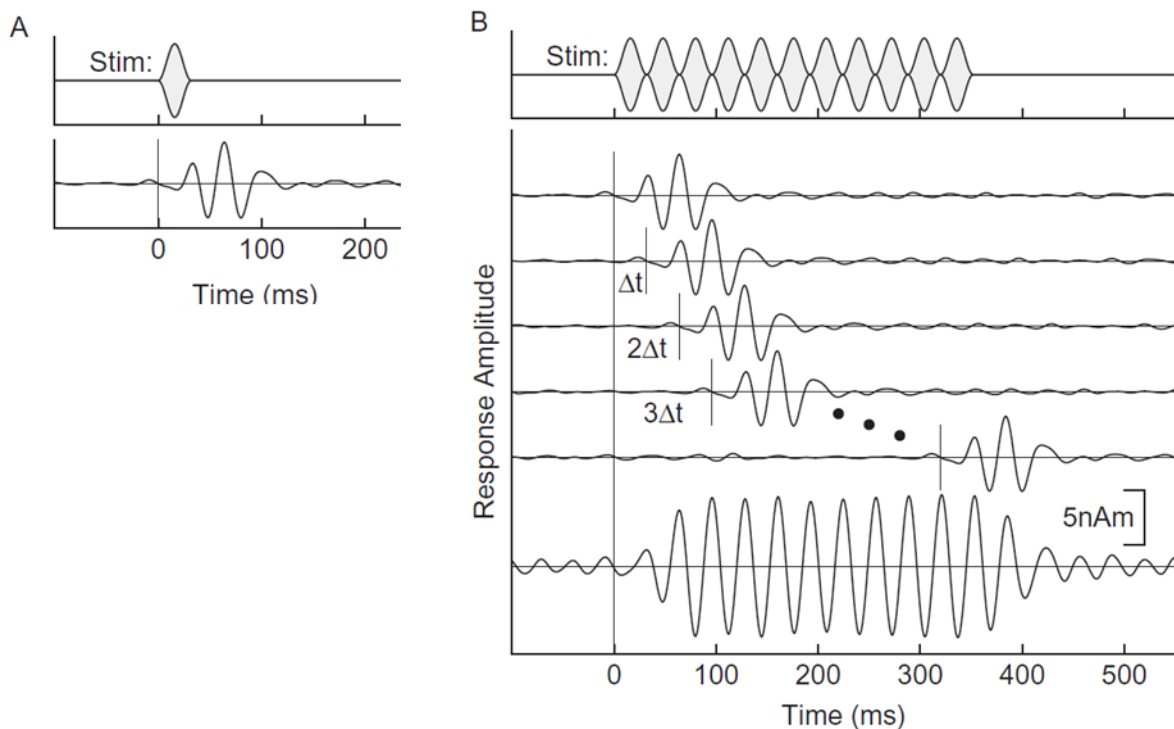


Figure.1C. Illustration of the superposition theory: (A) ERP to a single auditory stimulus; (B) ERPs to several consecutive stimuli shown in separate rows, and their summation is shown on the bottom row which resembles the ASSR<sup>129</sup>.

The conflicting camp however, contends that the ASSR is separate from the event-related potentials, and is generated when intrinsic neuronal oscillations synchronize with the stimulus temporal envelope, otherwise dubbed as the ‘Entrainment Hypothesis’<sup>130</sup>. One of the basis of this hypothesis stems from the fact that prominent neuronal oscillations, including beta and gamma bands 20 – 80 Hz, occur at frequencies that match that of the characteristic features in natural communication sounds such as speech<sup>131</sup>. Hence, these intrinsic neuronal oscillations are ideal for responding dynamically to fluctuations in the stimulus temporal envelope<sup>132</sup>. Selective attention entrains neuronal oscillations in a top-down manner, such that they phase-lock to the temporal envelope of the attended stream<sup>133</sup>. As a result, the neurons’ high excitability phases will coincide with major events occurring in that stream, thereby encoding the input information with maximal integrity and ensuring that the corresponding neural response is generated at high amplitude<sup>130</sup>. This happens at the expense of unattended streams which arrive out-of-phase with the neuron firing, resulting in neuronal suppression<sup>134</sup>. This process of neuronal tracking is most clearly illustrated in studies showing that distinct neural representations physically mirror the attended speech stream<sup>14, 37, 133</sup>. Intriguingly, the temporal envelope<sup>14, 133</sup> as well as spectrogram<sup>37</sup> of the attended stream can be reliably reconstructed from the recorded neural signals (Fig. 1D)<sup>133</sup>. Moreover, the neural representation of speech is dynamic and refines itself overtime, suggesting that its formation is an accumulative process that relies on spectral and temporal regularities<sup>14</sup>.

### Neural Encoding of Each Speech Stream

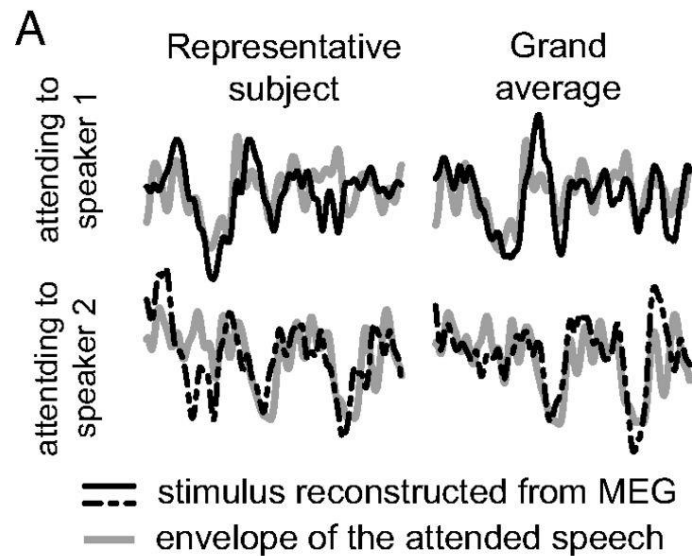


Figure 1D. Reconstruction of each attended speech stream from MEG recordings of neural activity (**black**), superimposed onto the corresponding stimuli’s original temporal envelope (**gray**). When subjects were presented with the same stimuli mixture containing two different speech streams, but instructed to attend to only one, different envelopes matching the attended streams can be decoded from neural responses (adapted from Ding 2012<sup>133</sup>).

### 1.3.2 Technical Considerations for Using Frequency-Tagging

The frequency-tagged ASSR can be elicited by presenting a subject with an auditory stimulus whose envelope and/or frequency changes periodically in time. A wide range of auditory stimuli has been used for this purpose, including clicks<sup>126, 135</sup>, tone-bursts<sup>126</sup>, binaural beats<sup>136</sup>, and via modulation using AM and FM of continuous tones<sup>137</sup>. The time needed to record a significant ASSR differs according to stimulus type, or more specifically, the broadness of their spectral bandwidths. A tone burst (first row in Fig. 1E) for example, has an acoustic spectrum that spans over a wide range of frequencies, thereby activating a large area of the basilar membrane. The summation of the resultant responses will thus be larger than that elicited from a sound with a narrower spectral bandwidth (e.g. pure sinusoidal amplitude-modulated or SAM tone), effectively decreasing the recording time needed<sup>115, 138</sup>. Figure 1E displays a comparison diagram showing the different types of stimuli used in most ASSR-related paradigms alongside their respective spectral compositions.

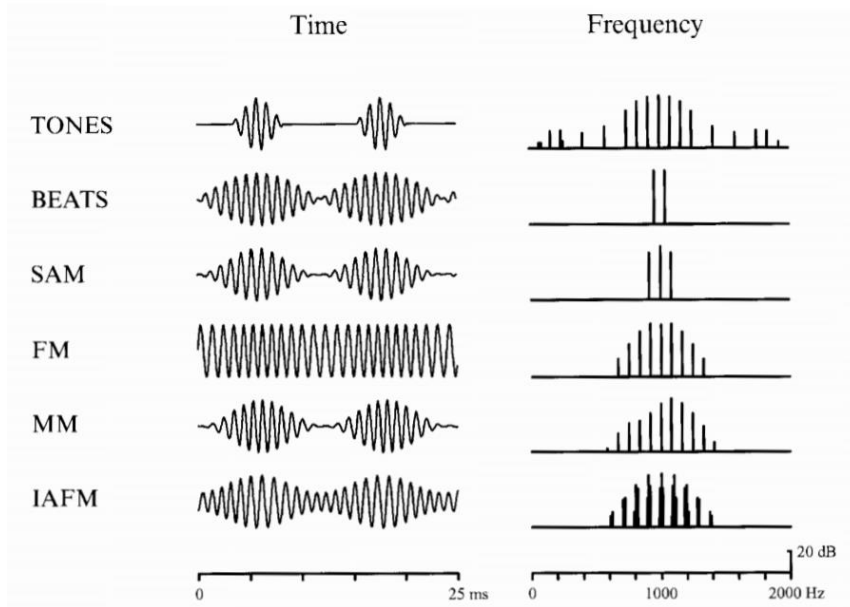


Figure 1E. Different types of stimuli used to elicit an ASSR. The waveform of each stimulus is displayed on a time axis with its spectral decomposition in the frequency domain shown on the right. For a sinusoidal amplitude-modulated (SAM) tone, three peaks centred at  $f_c$  are present and the difference between each peak equals to  $f_m$  (adapted from Picton 2003)<sup>115</sup>.

Typically, sinusoidal AM (SAM) tones are defined by the equation:

$$y(t) = A \sin(2\pi f_c t) \cdot (1 - m \cos(2\pi f_m t))$$

where  $f_c$  is the carrier frequency,  $f_m$  is the modulation frequency,  $A$  is the amplitude and the modulation depth,  $m$ , is calculated as follows:

$$m = \frac{A_{max} - A_{min}}{A_{max} + A_{min}}$$

$A_{max}$  and  $A_{min}$  denote the maximum and minimum amplitudes of the resultant AM waveform<sup>114</sup>,

The resultant SAM tone will contain three spectral peaks at  $f_c - f_m$ ,  $f_c$  and  $f_c + f_m$  in the frequency domain (Fig. 1E). As seen in Fig. 1E, FM frequency-tagging produces a slightly different spectral composition with more peaks centred around  $f_c$  that have equal spacing of  $f_m$  between one another. Furthermore, AM and FM can be combined at the same (mixed modulation) or different  $f_m$  on one carrier stimulus (Independent amplitude and frequency modulation) to achieve a larger total ASSR response<sup>125, 139</sup>.

For ASSR recorded at the cortex, the amplitude varies with  $f_m$  or stimulus rate with the highest amplitude occurring at around 40 Hz (Fig. 1F)<sup>129</sup>. Additionally, the source waveforms that constitute the ASSR become more sinusoidal in shape with increasing  $f_m$  values (Fig. 1G(a)). ASSR peaks can also occur at multiple values of the  $f_m$  (Fig. 1G(b))<sup>114</sup>.

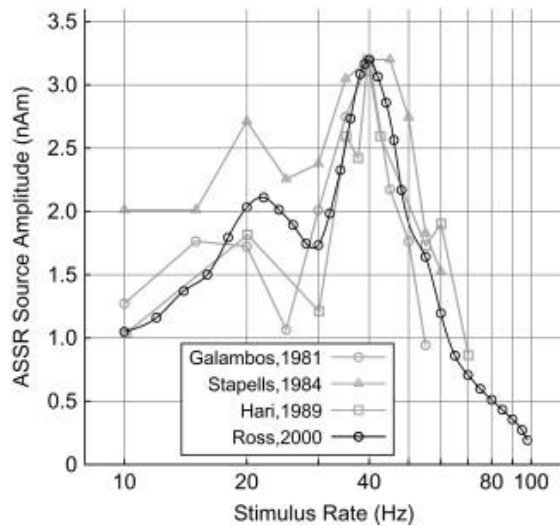


Figure 1F. Variability in ASSR amplitude, localized in the auditory cortex, with respect to stimulus rate or modulation frequency across four independent studies covering a range of stimuli used. A clear peak is observed at around 40 Hz, followed by a steep decline thereafter. Another smaller peak is also identified at around 20 Hz (adapted from Ross 2013<sup>129</sup>).

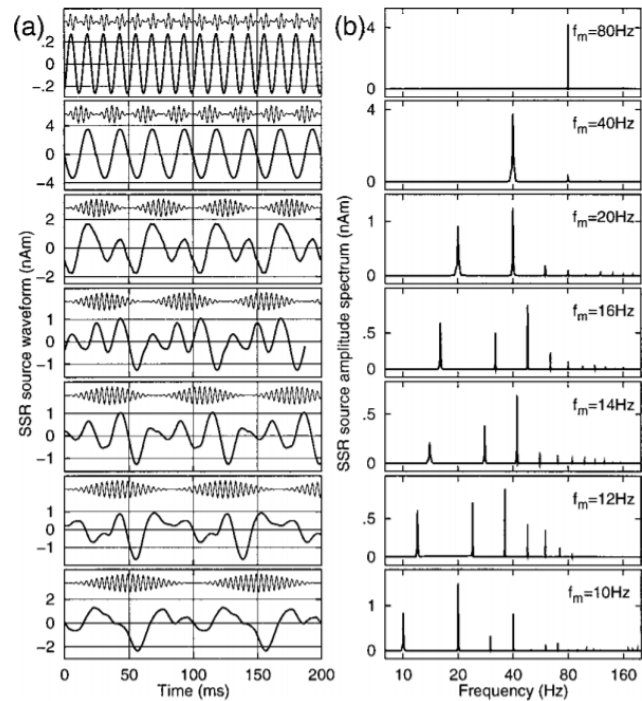


Figure 1G. ASSR data presented in time (left) and frequency domain (right). (a) For each  $f_m$ , the corresponding stimulus sound envelope is shown above the ASSR data along the time domain. The ASSR waveform tends towards a sinusoidal shape as  $f_m$  increases. (b) Fourier-transformed ASSR data showing peaks at  $f_m$  and subsequently at multiples of (adapted from Ross 2002<sup>114</sup>).

When designing an ASSR experiment, it is hence important to adjust the variable parameters in order to achieve a reasonable signal-to-noise ratio within a given time frame. Generally, the ASSR amplitude increases with modulation depth and stimulus intensity, but decreases with carrier frequency<sup>114-115</sup> (see Fig. 1H).

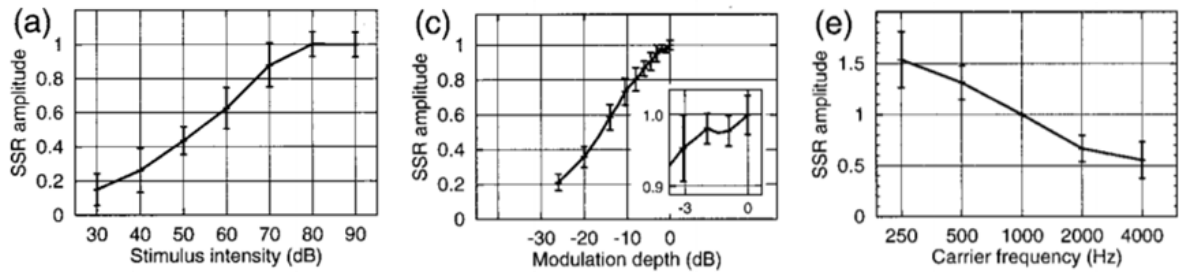


Figure 1H. Average effects of (a) stimulus intensity, (c) modulation depth and (e) carrier frequency on ASSR amplitude computed from eight subjects (adapted from Ross 2000<sup>114</sup>).

### 1.3.3 Frequency-Tagging Multiple Stimuli

It is also possible to record ASSRs from multiple AM frequency-tagged stimuli presented simultaneously (Multiple auditory steady-state responses or MASTER) with reasonably sufficient amplitudes, although careful adjustments of the necessary parameters must be considered to minimize between-stimuli interactions that can produce ASSR artefacts from unwanted distortion products<sup>138</sup>. Such distortion products arise due to the non-linear responsiveness and unidirectional nature of cochlear hair cells that produce the resultant ASSR. These cochlear properties can be modelled as equivalent to a compressive rectifier and they form the basis of sound envelope recognition<sup>138</sup>. As such, The ASSR at  $f_m$  is itself a distortion product that would not have been possible without this cochlear property. Beats are another type of distortion product that occur when two sound waves of similar frequency are presented simultaneously to a subject, giving rise to the perception of audible ‘beats’ at the frequency difference ( $< 30$  Hz) of the two waves<sup>136, 140-141</sup>. Distortion products can also occur at the various frequency differences between all possible combinations of the frequencies ( $f_c - f_m$ ,  $f_c$ ,  $f_c + f_m$  as well as their harmonics in more complex stimuli) present in the mixture (see Fig. 1I for an illustrative example). These artefacts give rise to a messy spectrum but the primary problem occurs when the ASSR at  $f_m$  coincides with the ASSR at the frequency of distortion products, causing ambiguities in determining the actual amplitude of the ASSR due to modulation<sup>138</sup>. In addition, previous experiments have showed that the amplitudes of 40 Hz ASSRs were significantly reduced when multiple AM tones were presented to the subject simultaneously<sup>137-138, 142</sup> (Fig. 1J).

Despite its limitations, with strategic planning, the MASTER technique is nonetheless very useful for studying auditory processing in complex auditory source mixtures and reducing recording time, for example in objective hearing assessments<sup>143</sup>.

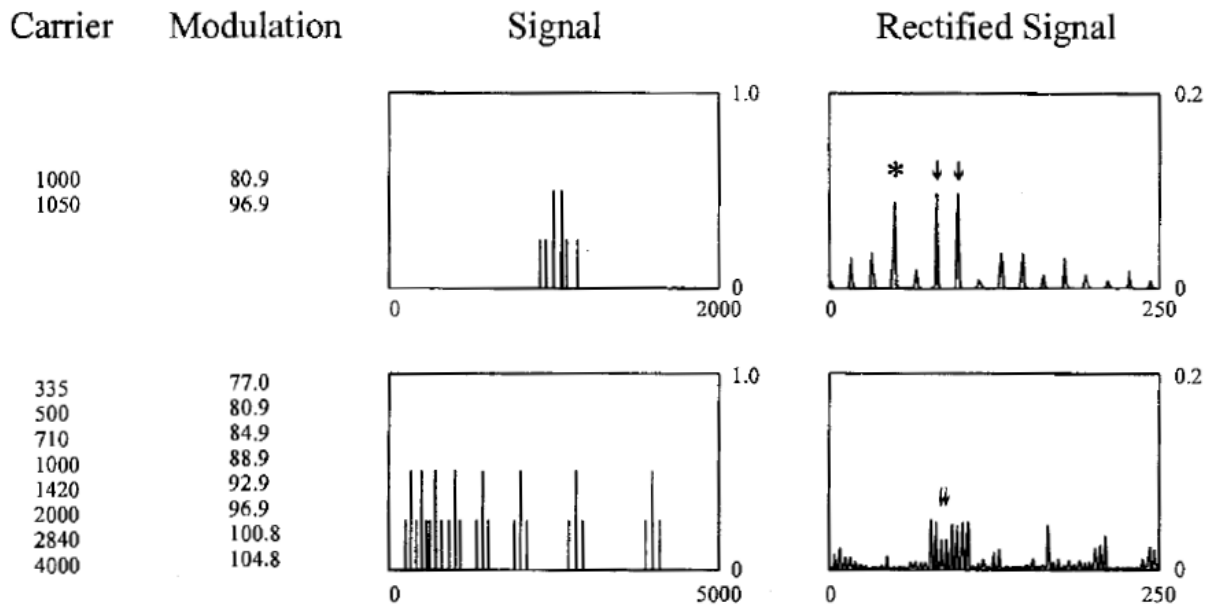


Figure 1I. Distortion products produced when two (top row) and eight (bottom row) stimuli with unique modulation frequencies were presented simultaneously. The distortion products occur at  $f_m$  and other frequencies equal to the differences between the various frequencies present. The two left-most columns display the carrier frequencies present in the mixture as well as their corresponding modulation frequencies. The centre 'Signal' column gives the resultant frequency spectra of the stimulus mixture. The rightmost column indicates the possible distortion products present at various frequencies: one of the beat frequencies is indicated with an asterisk (\*) and some of the modulation frequencies are marked with arrows (↓), while the rest are artefacts. Note that the example is for  $f_m \approx 80$  Hz but the same principle applies for  $f_m \approx 40$  Hz (adapted from John 1998<sup>138</sup>).

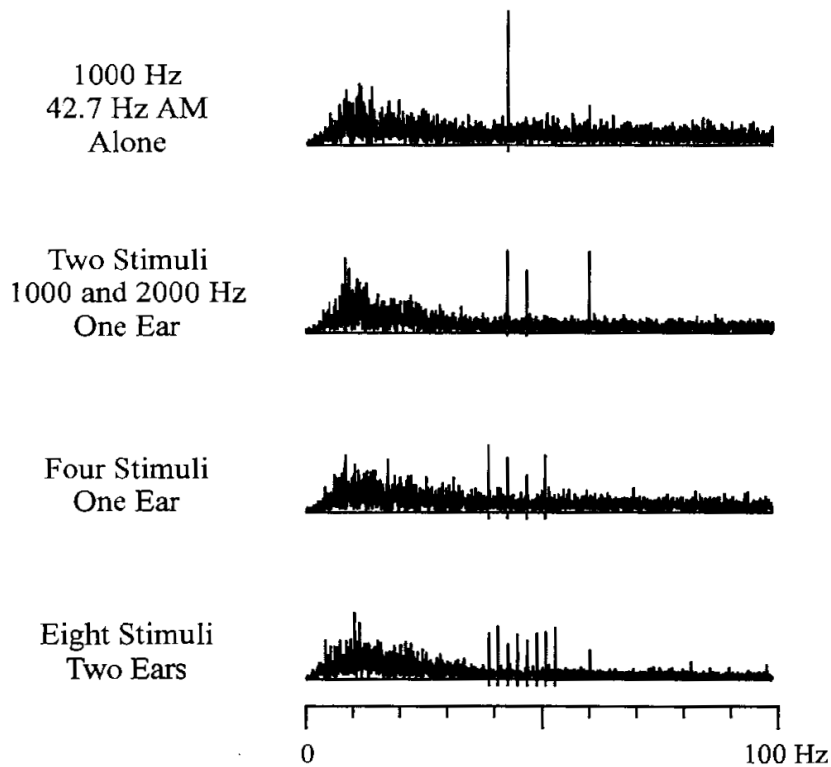


Figure 1J. Multiple ASSRs elicited in a typical subject when one, two, four and eight stimuli with different modulation frequencies were presented simultaneously. As observed, there is significant attenuation of the ASSR with simultaneous stimuli and this attenuation increases with the number of stimuli (adapted from John 1998<sup>138</sup>).

### 1.3.4 ASSR in Attentional research

A key application that highlights the usefulness of the ASSR for studying human cognitive functions is in the field of selective auditory attention. The ASSR at approximately 20 - 40 Hz is sensitive to attentional enhancement and has therefore been used in attention-related ASSR paradigms (see Fig.1K for illustrative example) where subjects were instructed to attend to a frequency-tagged auditory stream, presented alone or within a stimulus mixture, and the ASSR with and without attention were compared<sup>13, 99, 122, 144</sup>. Several intermodal studies have demonstrated that the cortically generated ASSR is enhanced when attention is directed towards an auditory stimulus from a competing visual stimulus<sup>145-147</sup>. Within the auditory domain (i.e. intramodal studies), the ASSR attentional enhancement was typically observed in simple cases containing a single auditory stimulus<sup>47, 148</sup> or where two different auditory stimuli were presented dichotically, wherein participants shift attention between the left and right ears<sup>122, 144, 149</sup>. Although relevant, one may argue that these paradigms were not representative enough of the complexity of the cocktail party problem in reality, as our natural auditory environment often contains several simultaneous sounds and is more stimulating and chaotic. In more complex scenarios, the experimental conclusions regarding the ASSR attentional modulation are mixed<sup>145, 150</sup>. The inconsistency in findings suggests that whether or not attention is found to affect the ASSR depends on several experimental design factors pertaining to the stimuli, task and analytical approach.

At source level, most studies have shown that the attentional enhancement of ASSRs is manifested in the auditory cortices<sup>13, 99, 122, 144</sup>. More recently, using complex musical stimuli, our group has demonstrated that the neural activity enhancement from intramodal selective auditory attention occurs also in regions beyond the temporal cortex, such as the parietal and frontal cortices, with regions in the pre-frontal cortex experiencing the largest degree of ASSR enhancement<sup>51-52</sup>. In addition, we also discovered that both the ASSR power itself and its degree of attentional modulation correlate with subjects' musicality<sup>52</sup>. Together, these findings reinforce the “gain” theory of selective attention<sup>25, 37, 41, 133, 151</sup>, wherein attention enhances the neural representation of a stimulus, and also show that selective attention recruits processes widespread across neural regions, some of which are influenced by long-term training.

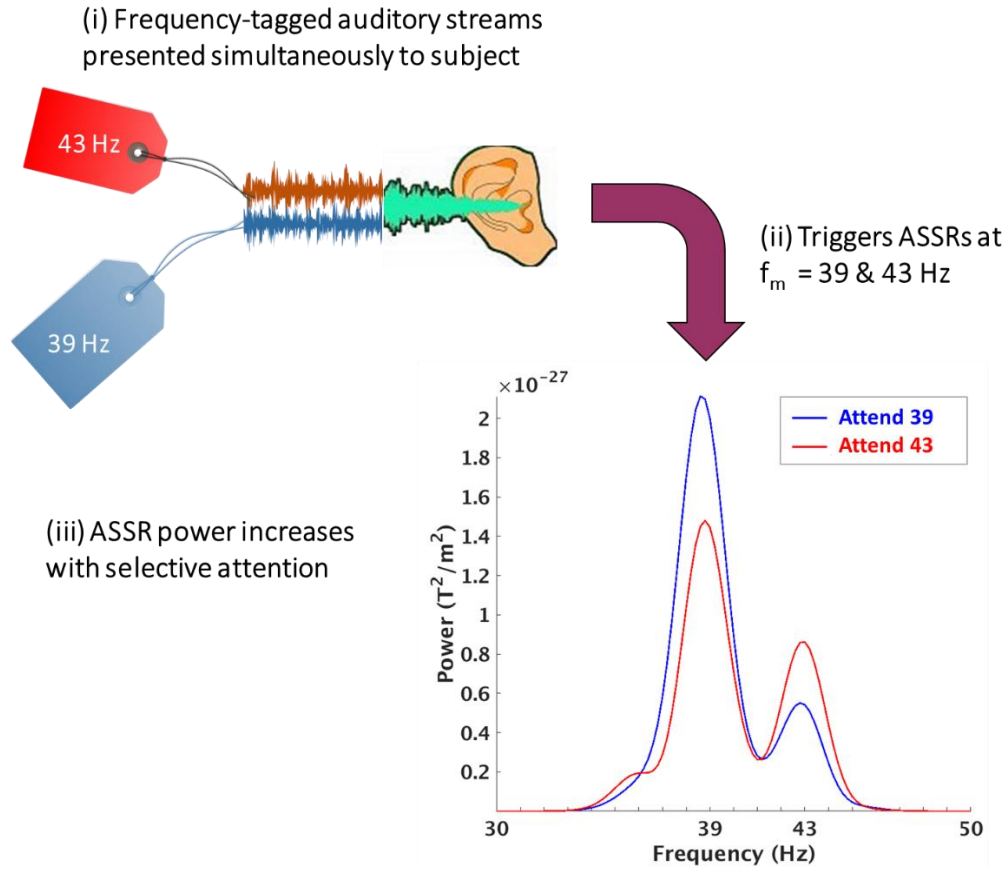


Figure 1K. Example of a frequency-tagging paradigm used to capture the effect of selective attention on ASSRs. (i) In this experiment, two auditory streams were frequency-tagged at 39 and 43 Hz respectively, and presented simultaneously to the subject. (ii) As the brain activity phase-locks to the incoming auditory stimuli, listening to these streams triggers ASSR power spectra peaks at 39 Hz and 43 Hz. (iii) When the subject attended selectively to the 39 Hz (red) stream, the corresponding ASSR power at 39 Hz was enhanced. Similarly, the 43 Hz ASSR peak increased in power when the subject focused on the other stream (power spectra obtained from unpublished data by our group).



## 1.4 UTILIZING MUSIC FOR NEUROSCIENCE

Engaging in music involves a multitude of sensory (e.g. motor coordination, auditory processing), cognitive (e.g. selective attention, working memory), as well as affective (e.g. emotional expression) processes, thus providing a platform to study how these processes operate, develop and are improved through training. Additionally, the complexity of music allows us to study neural processes in more naturalistic environments which typically involve complex auditory scenes, while remaining experimentally flexible and controllable to a large extent. For example, polyphonic music, which is composed of distinguishable melodies in different musical voices, can be useful in the study of selective auditory attention as the ability to shift attention between voices depend on several variables (e.g. interval or amplitude difference, melodic or rhythmic complexity, tonality etc.) that we can manipulate experimentally. Furthermore, ensemble and band playing present the opportunity to explore attention shifts between instruments, as well as between self-produced music and accompaniments generated by third-parties (e.g. fellow band members). Here, the experimentalist is granted the autonomy to customize the perceptual features (i.e. pitch, timbre) of the auditory stimuli according to the research question. As such, music offers a complex yet controllable platform to study the CPE. Thirdly, many researchers have found the musician's brain a useful model for studying neural plasticity and cognitive development, especially training-induced auditory expertise<sup>152-153</sup>. Neuroplasticity refers to changes in the brain's anatomy or activity in response to external stimuli, and is usually associated with the development of skills and behavioural adaptation. In the following sections, we will describe the neural changes brought about by musical training and how they relate to superior behavioural performances in music-related as well as non-music-related tasks.

### 1.4.1 Anatomical & Functional Neural Changes Associated with Musical Training

Studies across various neuroimaging modalities have shown both structural and functional differences between the brain of musicians and non-musicians<sup>154-160</sup>. As musical activities stimulate an extensive distribution of neural networks and regions, there exists no single area that is exclusively dedicated to music processing<sup>161-162</sup>. Hence, the music-induced changes manifest in varying degrees across different brain structures, depending on factors such as the type of musical training<sup>163-165</sup> (e.g. type of instrument, musical style), genetic predispositions<sup>166-167</sup>, training duration<sup>168</sup> and intensity<sup>156-157, 159</sup>, as well as age of training onset<sup>169-170</sup>. Other works in the field suggest that the neural changes and benefits accompanying musical training during childhood persist until adulthood even if the training did not continue<sup>171</sup>.

Structural changes in grey matter morphology or cortical thickness are assessed using MRI<sup>157-160</sup>, while alterations in white-matter integrity, microstructure and connections are evaluated through diffusion tensor imaging<sup>154-156</sup> (DTI). Compared to non-musicians, the musician's

brain contains increased grey matter volume and cortical thickness in the auditory, motor and visuo-spatial areas<sup>157, 168</sup>, frontal lobe<sup>160, 168</sup>, hippocampus<sup>157, 160, 168</sup>, and lingual gyrus<sup>160, 172</sup>, with enlargements in each region linked to distinct cognitive and sensory processes (see Olszewska et al. 2021<sup>173</sup> for a comprehensive review). DTI studies also revealed different white-matter architecture between musicians and non-musicians in several regions across the brain, especially in areas associated with fine motor control and sensory processing<sup>154-156</sup>.

Functional changes in brain activity can manifest as stronger neural responses<sup>158, 174-176</sup>, a migration of activated regions<sup>159, 176</sup>, or increased synchronization between brain regions<sup>177-178</sup>, and can be assessed using fMRI, EEG or MEG, typically with the involvement of a behavioural task such as listening to a tone or speech. These changes are observed with various stimulus types and paradigms, and have occurred in the auditory cortex<sup>158-159, 174, 179</sup>, subcortical regions<sup>175, 179</sup>, frontal lobe<sup>179-180</sup>, primary and supplementary motor areas<sup>181-182</sup>, as well as language-related regions<sup>181-182</sup> like Broca's area. Again, the modifications in brain activity are linked to enhancements in an array of sensory and cognitive abilities such as language processing, motor execution, selective attention, and emotional processing (see Olszewska et al. 2021<sup>173</sup> for more details). Often, the structural and functional changes brought about by musical training correlated with musical aptitude<sup>158, 176</sup> and cognitive performance<sup>179</sup>. Furthermore, these changes are strongly correlated with the intensity<sup>156-157, 159</sup> and duration<sup>168</sup> of musical training, as well as the age of training onset<sup>169-170, 176</sup>. Figure 1L on the following page provides an illustrative overview of the neural changes associated with musical training revealed by cross-sectional studies as summarized by Olszewska et al. 2021<sup>173</sup>.

However, cross-sectional studies that seek to link neural changes to musical training are faced with the classic “Nature versus Nurture” problem – scientists are still unsure to what extent these differences are the result of musical training per se or are confounded by certain genetic predispositions<sup>166-167</sup> (that may bias an individual to pursue a music education or persist longer in one). While studies demonstrating that the magnitude of the brain changes correlate with the amount of musical training<sup>156-157, 159, 168</sup> provide partial evidence supporting that these are indeed training-induced effects, natural predispositions between individuals can still bias their choices related to training intensity and duration. To address this conundrum, research on monozygotic twins, who should in principle carry identical genes and hence genetic predispositions, revealed that musical twins have increased cortical thickness and higher white-matter integrity in auditory-motor regions, higher white-matter integrity in the corpus callosum, as well as larger grey matter volume in the cerebellum, compared to their non-musical twin<sup>183</sup>. As the twins share an early environment and natural genetic predispositions, these brain differences are likely brought about by musical training.

Importantly, longitudinal studies also offer revealing insights into the temporal dynamics of neural changes during the musical training process. In such study designs, pre-existing

individual differences can be measured prior to training and accounted for, so that the training effects can be attributed to training-induced neuroplasticity. Longitudinal neuroimaging studies observed changes in brain anatomy<sup>184-185</sup>, activation level<sup>186-190</sup>, as well as functional connectivity<sup>177-178</sup> in either direction (i.e. increase or decrease) depending on brain region, which generally overlaps with the areas previously mentioned in other types of studies (e.g. frontal/temporal lobes, motor areas, cerebellum...etc.), and phase of training<sup>186-187</sup>. While these studies encompass a large range of training periods from several minutes to several months, most involve a relatively short time frame as they are expensive and difficult to implement over long periods.

# HOW MUSICAL TRAINING SHAPES THE BRAIN

INCREMENTS IN STRUCTURE AND FUNCTION  
MUSICIANS COMPARED TO NON-MUSICIANS

## FRONTAL LOBE

INCREASED GREY MATTER VOLUME  
superior, medial, inferior frontal gyrus  
– executive functions  
INCREASED fMRI ACTIVATION  
primary & supplementary motor areas  
– motor preparation and execution  
Broca's area – speech production

## TEMPORAL LOBE

INCREASED GREY MATTER VOLUME &  
INCREASED fMRI ACTIVATION  
Heschl's gyrus – primary auditory cortex,  
pitch perception  
superior temporal gyrus – auditory  
processing  
INCREASED MISMATCH NEGATIVITY (EEG)

## BASAL GANGLIA & LIMBIC SYSTEM

INCREASED GREY MATTER VOLUME  
hippocampus – memory formation & retrieval  
INCREASED EEG RESPONSE  
temporal-limbic areas – emotions and memory

## PARIETAL LOBE

INCREASED GREY MATTER VOLUME  
primary somatosensory cortex – touch perception  
INCREASED fMRI ACTIVATION  
supramarginal gyrus – syntax processing  
and attention

## WHITE MATTER TRACTS

INCREASED WHITE-MATTER INTEGRITY  
corpus callosum – connects brain hemispheres  
corticospinal tracts – connect motor areas with  
spinal cord  
short fibres – connect sensory and motor areas  
striatum – motor planning and reward perception

## OCCIPITAL LOBE

INCREASED GREY MATTER VOLUME  
lingual gyrus – score reading

## CEREBELLUM

INCREASED GREY & WHITE MATTER VOLUME  
movement coordination and motor learning

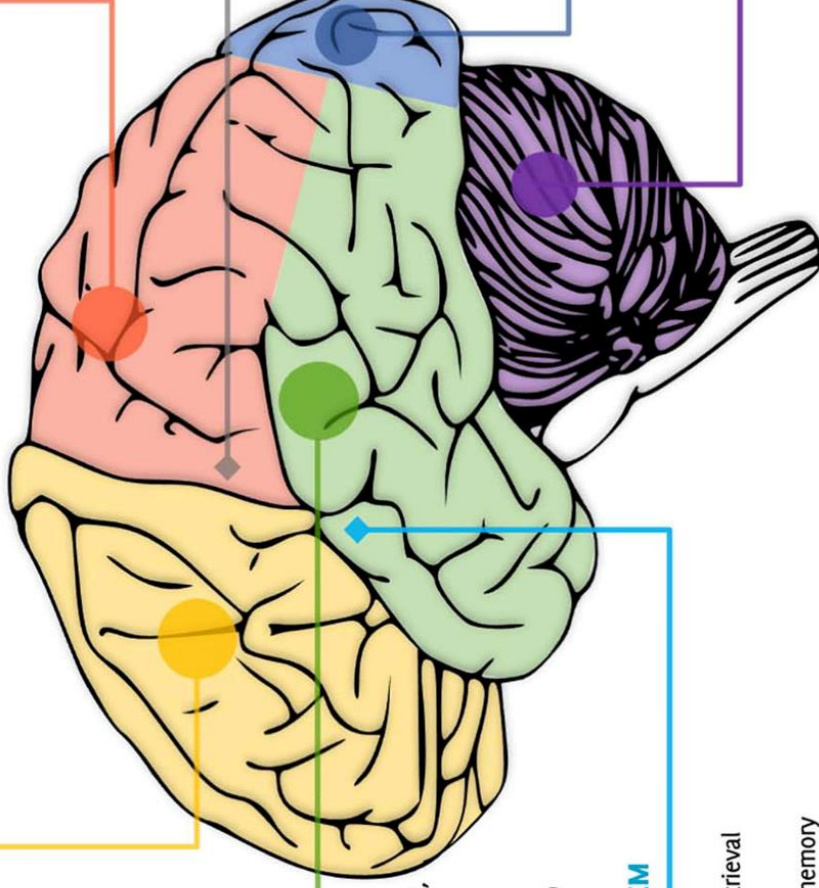


Figure 1L. Overview of neural changes associated with musical training taken from Olszewska et al. 2021<sup>173</sup>

### 1.4.2 Transferability of Musical Abilities to General Cognitive Skills

Championed by a library of literature, it is known that musical training brings about a wide array of benefits to the learner, ranging from auditory perception, linguistic ability, non-verbal reasoning, executive function such as attention and working memory, emotional processing and perhaps even spatial and mathematical ability<sup>191-201</sup> (see Miendlarzewska 2014<sup>202</sup> for a review). Out of these, enhancements to auditory processing and linguistic abilities are two of the most well-studied aspects.

Substantial evidence in the literature indicates that the neuroplastic changes associated with musical training is not just specific to the processing of musical sounds, but can be carried over to general auditory skills, although which aspects of auditory processing is enhanced depend on the nature of training<sup>164-165, 192</sup>. For example, musically-trained individuals demonstrated better selective listening in noisy environments<sup>203-205</sup>, as well as superior memory traces of and higher sensitivity to acoustic features<sup>206-208</sup> (e.g. pitch discrimination ability). Furthermore, musical training also promotes faster maturity<sup>209</sup> and less age-related degradation<sup>210-211</sup> of auditory processing mechanisms.

Moreover, the perception of music and that of speech share many similar neural processes, leading to the belief that musical practice can enhance linguistic abilities<sup>212-216</sup>. A myriad of studies support this idea, showing that musical training benefits syntax processing, reading skills, verbal memory, prosody perception, pitch and consonant discrimination in speech, language acquisition, as well as speech-in-noise processing<sup>203, 205, 217-229</sup>. Many of these behavioral advantages coincided with stronger response to, or more robust neural encoding of speech stimuli and components of speech stimuli<sup>200, 208, 228-233</sup>. An advantage in pitch discrimination can help speech perception in various aspects, for instance, in the judgment of the speaker's emotion, intention (e.g. question or statement) or identity, and meaning in tonal languages<sup>192</sup>. Moreover, musical training teaches individuals to discriminate relevant sounds from a complex auditory cocktail, which allows them to better track regularities in the sound environment<sup>192, 234</sup>. This can play a role in improving speech-in-noise ability in both healthy individuals and populations with related deficiencies<sup>204, 235-236</sup>. In addition, the enhancement in speech-in-noise ability correlates with the number of years of musical experience<sup>237</sup>.



## 2 RESEARCH AIMS & THESIS FRAMEWORK

### General research aims and motivation

A crucial and challenging problem faced by researchers investigating the phenomenon of selective auditory attention, particularly when the research question involves sophisticated and naturalistic scenarios, stems from the complexity of the neural activity mix that is elicited when multiple auditory sources are simultaneously present. Although difficult, the separation and identification of mixed neural signals is necessary to allow researchers to study each individual stimulus-specific activity, such as activity generated by an individual instrument within an orchestra or a single speaker in a noisy cocktail party setting. As a potential solution to this challenge, frequency-tagging offers a simple, efficient and precise method of labelling neural activity uniquely according to its driving stimulus, allowing individual labelled neural responses within mixtures to be extracted and identified back to their respective source. Thus, the primary objective of this thesis is to develop and optimize methods utilizing frequency-tagging to track neural activity in complex auditory environments, and extract the effect of selective attention on the corresponding ASSR signals, for the purpose of investigating the neural underpinnings governing selective auditory attention. Since frequency-tagging benefits greatly from measurement technologies with a high degree of spatio-temporal precision, MEG's superior spatial (complemented by anatomical MRI templates) and temporal resolution makes it an unparalleled choice for recording ASSRs.

An important step to learning how to effectively utilize the ASSR for cognitive research is to acquire a more comprehensive understanding of the characteristics of the ASSR in order to adapt its use in different situations and study designs. The present thesis aimed to accomplish this by examining what brain regions generate the ASSR and what their functional implications are, how the ASSR develops across time, and how it is influenced by auditory training (e.g. musical expertise).

Even though the ASSR has been used in previous research for studying human cognition, one may argue that the experimental paradigms used were too simplistic and thus not sufficiently representative of the complexity of the cocktail party problem in reality, as our natural auditory environment often contains several simultaneous sounds and is more stimulating and chaotic. Thus, a key component of this thesis is to progress towards applying the frequency-tagging method to more complex and naturalistic contexts such as multi-voiced music, potentially steering research away from conventional simplistic and artificial experimental set-ups (e.g., clicks and beeps) towards more realistic soundscapes.

Additionally, by inspecting the association between ASSR and musical experience, the present thesis intended to explore the effects of long-term auditory training on auditory

processing, sound separation and selective attention. In the long run, we would like to employ the ASSR to devise more effective strategies for the enhancement of these skills.

From a more general perspective beyond selective attention, we hope that the frequency-tagging technique developed by this thesis can be useful for understanding how sound information is processed in the brain, as well as how cognitive abilities are acquired, developed and executed.

#### Study-specific aims

In accordance with the aforementioned general aims, we carried out three studies utilizing variations of a melody tracking task with different degrees of complexity, during which participants were instructed to exclusively focus on a single melody stream in a mixture of two to three melody streams that were presented diotically (i.e., identically to both ears). The streams were AM frequency-tagged to elicit unique ASSRs between 39 – 43 Hz, and were differentiable only by their respective onset timing and pitch, so that attention was based on these sound features (rather than spatial location).

In the first study, **Study I**, the main goal was to assess if selective attention directed towards a specific melody stream modulates the corresponding ASSR power, and if this can be already observed in sensor level MEG analyses. A secondary goal of **Study I** was also to gain insights into the structural distribution of cortical ASSR sources and their attentional modulation.

Although exploratory in nature, the secondary results from **Study I** served as a spatial template for **Study II** to more precisely pin-point *a priori* which cortical sites were more likely to be susceptible to ASSR attentional modulation and perform confirmatory analyses via statistical testing at regions across the bihemispheric frontal, parietal, temporal and insular lobes. In addition, **Study II** also investigated the relationship between ASSR attentional modulation and musicality at each of these locations. A parallel aim of **Study II** was to compare the ASSR's characteristics during and after auditory stimulation, in terms of their sensitivity to the (i) stimulus physical properties such as volume and carrier frequency, (ii) selective auditory attention, (iii) musicality, as well as their respective (iv) cortical source distributions.

Finally, **Study III** sought to extend the application of ASSRs to more complex auditory mixtures containing simultaneous musical streams that overlapped in time (in Study I and II the streams were separated in time). To this aim, we investigated how the suppression effect from simultaneous ASSRs affected the ASSR activity and its cortical source distribution, as well as whether the suppression would affect the separability of mixed ASSR signals, or their ability to assess attentional modulation. For this final study, both top-down and bottom-up selective attention were engaged to manipulate the ASSR as we compared and contrasted how each of them modulate ASSR sources across the cortical space.



### **3 MATERIALS & METHODS**

In this section, we will summarize the common methodologies adopted for each of the three studies included in the present thesis. For details on additional methods specific to each study, please refer to the appended articles in the Appendix.

#### **3.1 PARTICIPANTS**

For each study, twenty-eight to twenty-nine participants with normal hearing volunteered to take part in the experiment. Participants with less-than-chance performance in the behavioural task (see section 1.3) were excluded for all MEG analyses.

#### **3.2 THE GOLDSMITH MUSICAL SOPHISTICATION INDEX**

A subset of the Goldsmith Musical Sophistication Index (Gold-MSI) self-report questionnaire (v1.0)<sup>238-239</sup> containing 22 questions was used to estimate each participant's level of musical sophistication. The MSI quantifies a participant's level of musical skills, engagement and behaviour in multiple facets, and is ideal for testing amongst a general population that includes both musicians and non-musicians. For relevance to the selective auditory attention task, we focused on questions from the perceptual ability, musical training and singing abilities subscales. These include all questions from the musical training and singing abilities subscales, selected questions from the perceptual abilities and active engagement subscales, but no questions from the emotions subscale. The emotions subscale was omitted because our melodic stimulus is non-emotional and therefore this subscale was deemed task-irrelevant. Several questions from the active engagement subcategory were also omitted as we wanted to focus on musical aptitude and ability, rather than exposure (e.g. questions such as "Music is kind of an addiction for me – I couldn't live without it" that does not directly affect musical ability, were omitted). This specific combination of questions has demonstrated strong correlations between the resultant MSI and performance in all of our three selective auditory attention tasks. Across all participants, we obtained MSI scores ranging from 40 - 132, out of a maximum score of 154. A copy of the questionnaire used for this study can be found in the appendix (Supplementary Information S1 of Paper 2).

#### **3.3 EXPERIMENTAL TASKS & TRAINING**

Each of the three studies used a slightly different variation of the Melody Development Tracking (MDT) task. All of them require the participant to selectively attend to one out of two or three melody streams with different pitch (i.e. carrier frequency range) that are termed 'voices'. Study I and II used three voices and are identical except that the pitch difference between voices was further apart in Study II. In the first two studies, a new tone from each voice was presented every 1 s, with a 250 ms silent inter-stimulus interval (ISI) between tones. Using only two voices, Study III adopted the same pitch difference between the lowest and highest pitched voice as Study II, but contained voices that were completely overlapping

in time with no silent gaps between them (whereas the voices were completely separated in time in Study I and II). The following Figure 3A visually illustrates the differences in experiment design between the non-overlapping and overlapping variations of the MDT task. For a summary of the technical details outlining each study, please refer to Table 3B.

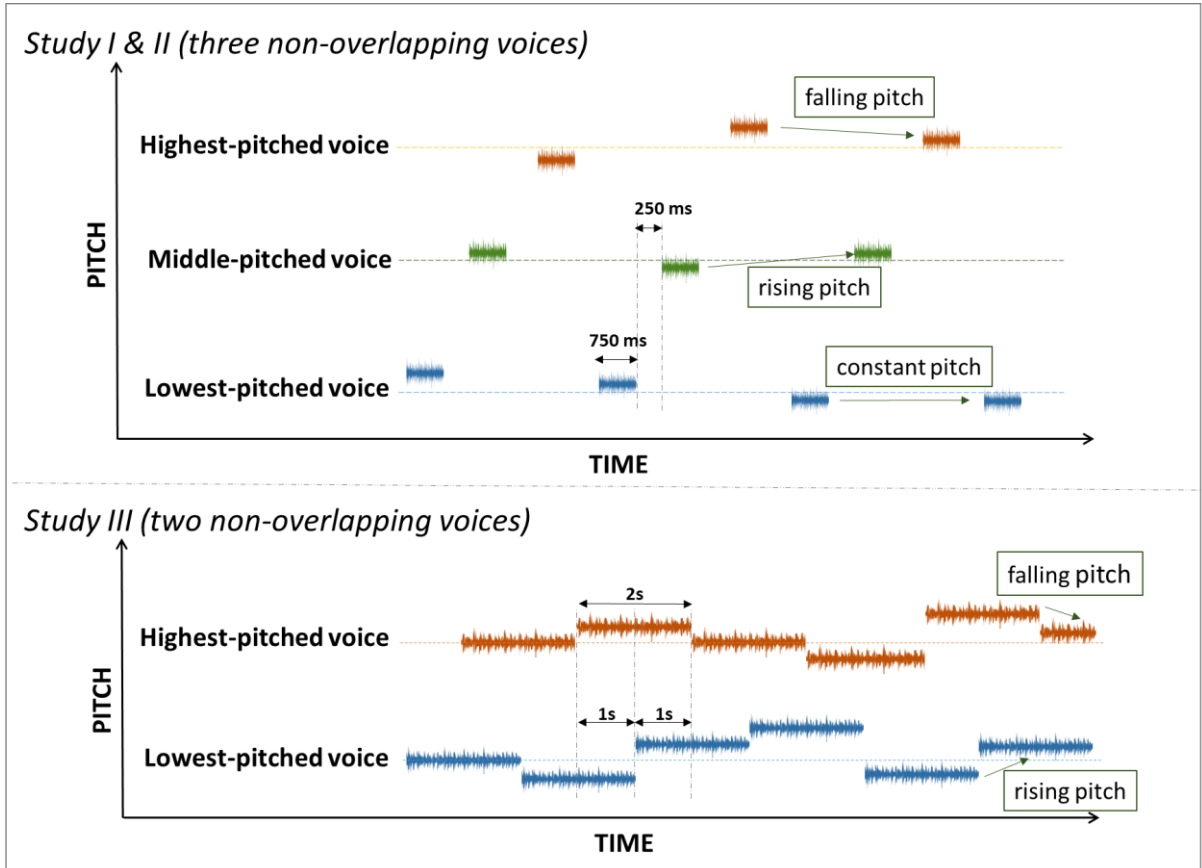


Figure 3A. Illustrative description of the non-overlapping (Study I and II) and overlapping (Study III) experimental tasks (not drawn to scale).

Before each set of melody begins, the participants were cued to direct attention exclusively to the lowest-pitched voice or highest-pitched voice. When the melody stopped at a random time point, participants were required to report the last direction of pitch change for the attended voice (either *falling*, *rising* or *constant* pitch) with a button press. In total, 28 of these behavioural responses per study were collected for each participant. To assess response accuracy in the MDT task, mean task performance scores for each participant were calculated as the percentage of correct answers out of all 28 responses.

Prior to the actual MEG recording, participants received 10 to 15 min of training to familiarize themselves with the task. Participants were deemed ready to commence with the actual experiment once they managed to report the correct answers for at least five consecutive trials. As the task was designed to require continuous selective attention to the cued melody stream, it was imperative to maintain alertness and alleviate fatigue. We therefore introduced a brief break in the task every ~5 min, during which the general attentiveness of the participant was also assessed using the Karolinska sleepiness scale<sup>240</sup>. To

minimize movement artefacts, participants were asked not to move when listening to each melody segment, which was at most 30 s long.

### 3.4 STIMULI

Each voice was constructed using a stream of 750 ms (Study I and II) or 2 s (Study III) long sinusoidal tones of  $f_c$  between 131 – 523 Hz (see Table 3B for exact  $f_c$  of each voice), generated using the Ableton Live 9 software (Berlin, Germany). For Study I and II, each tone is followed by 250 ms of silence before the next tone was played, whereas no silent gap exists between tones in Study III. The tones were amplitude-modulated sinusoidally in Ableton Live 9 at  $f_m$  of 39 (Bottom or Low voice), 41 (Middle voice) and 43 (Top or High voice) Hz, and a modulation depth of 100% to achieve maximum ASSR power<sup>8</sup>. For simplicity, only tones in the C major harmonic scale were used. The duration of melody presentation was randomized to be between 9 – 30 seconds long to reduce predictability of the stop point and thereby maintain high attention throughout the melody. The relative volume of each voice was adjusted to account for differences in subjective loudness for different frequency ranges<sup>241</sup>. The respective settings for the lowest-pitched, middle-pitched and highest-pitched voices were 0 dB, -6 dB and -10 dB, resulting in their raw volume decreasing in the same order. In order to focus on feature-based sound separation (in this case, pitch and timing) rather than location-based (i.e. left-ear versus right-ear) source separation, the mixture of streams was presented identically to both ears via foam inserts, ears with the volume calibrated to approximately 75 dB SPL per ear using a soundmeter (Type 2235, Brüel & Kjær, Nærum, Denmark), subjected to individual comfort level.

	<i>Number of voices</i>	<i><math>f_c</math> range per voice</i>	<i>Tone duration + ISI</i>
<i>STUDY I</i>	3 non-overlapping	Low: 131 – 220 Hz Middle: 147 – 294 Hz High: 196 – 329 Hz	750 ms (tone) + 250 ms (silence)
<i>STUDY II</i>	3 non-overlapping	Low: 131 – 220 Hz Middle: 175 – 349 Hz High: 329 – 523 Hz	750 ms (tone) + 250 ms (silence)
<i>STUDY III</i>	2 overlapping	Low: 131 – 220 Hz High: 329 – 523 Hz	2 s (tone) + no silent gap

Table 3B. Summary of Study I, II and III with respect to the number of voices, their respective carrier frequency ( $f_c$ ) ranges, tone duration and inter-stimulus interval (ISI).

### 3.5 DATA ACQUISITION

MEG measurements were carried out using a 306-channel whole-scalp neuromagnetometer system (Elekta TRIUX™, Elekta Neuromag Oy, Helsinki, Finland). Data was recorded at a 1 kHz sampling rate, on-line bandpass filtered between 0.1 - 330 Hz and stored for off-line analysis. Horizontal eye-movements and eye-blinks were monitored using horizontal and vertical bipolar electrooculography (EOG) electrodes. Cardiac activity was monitored with bipolar electrocardiography (ECG) electrodes attached below the left and right clavicle. Internal active shielding was active during MEG recordings to suppress electromagnetic artefacts from the surrounding environment. In preparation for the MEG-measurement, each participant's head shape was digitized using a Polhemus FASTRAK. The participant's head position and head movement were monitored during MEG recordings using head-position indicator coils. Anatomical MRIs were acquired using hi-res Sagittal T1 weighted 3D IR-SPGR (inversion recovery spoiled gradient echo) images by a GE MR750 3 Tesla scanner with the following pulse sequence parameters: 1 mm isotropic resolution, FoV  $240 \times 240$  mm, acquisition matrix:  $240 \times 240$ , 180 slices 1 mm thick, bandwidth per pixel=347 Hz/pixel, Flip Angle=12 degrees, TI=400 ms, TE=2.4 ms, TR=5.5 ms resulting in a TR per slice of 1390 ms.

### 3.6 DATA PROCESSING

The acquired MEG data was pre-processed using MaxFilter (-v2.2)<sup>85-86</sup>, and subsequently analysed and processed using the Fieldtrip toolbox<sup>242</sup> in MATLAB (Version 2016a, Mathworks Inc., Natick, MA), as well as the MNE-Python software<sup>243</sup>. Cortical reconstruction and volumetric segmentation of all participants' MRI was performed with the Freesurfer image analysis suite<sup>244</sup>.

#### 3.6.1 Pre-processing

MEG data was MaxFiltered by applying temporal signal space separation (tSSS) to suppress artefacts from outside the MEG helmet and to compensate for head movement during recordings<sup>85-86</sup>, before being transformed to a default head position. The tSSS had a buffer length of 10 s and a cut-off correlation coefficient of 0.98. The continuous MEG data was divided into 1 s-long epochs from stimulus onset (i.e. onset of each tone). Epochs were then visually inspected for artefacts and outliers with high variance were rejected using *ft\_rejectvisual*<sup>242</sup>. After cleaning, approximately 70 % of all epochs remained and were kept for further analyses.

#### 3.6.2 Sensor-space analysis

Sensor-space analysis was carried out on cleaned MEG epochs obtained after the pre-processing steps above, to extract the ASSR power for each condition. The minimum number of cleaned epochs averaged was ~100 per condition. To obtain the *timelocked* ASSR data

across time, a 30 – 50 Hz bandpass filter was applied to the epochs which were then averaged per condition. Power spectral density (PSD) estimation methods were then applied on a truncated time window of the *timelocked* ASSR, zero-padded to 1 s to acquire the ASSR power spectra with a frequency resolution of 1 Hz. The ASSR power at  $f_m$ , defined as 39, 41, and 43 Hz for the lowest-pitched, middle-pitched and highest-pitched voices respectively, was extracted to give the mean ASSR power per condition, and then further averaged across MEG sensors. In Study I, ERF sustained fields (300 – 800 ms post-stimulus<sup>13, 18</sup>) were also extracted, using a 20 Hz low-pass filter, to check for the manipulation of attention by the task, since it has already been well-documented in literature that attention enhances the ERF<sup>4-7</sup>. Typically, calculated values, including ASSR power and computed ratios, were converted to the base 10 logarithmic scale to achieve more normal data distributions across participants for parametric statistical analysis (e.g. t-tests, ANOVAs). Pearson's correlation tests were used to investigate the relationships between variables, for example between ASSR power and MSI.

### 3.6.3 Source-space analysis

We used a minimum-norm estimate<sup>243</sup> (MNE) distributed source model containing 20484 dipolar sources on the cortical surface to produce individual-specific anatomical layouts of the ASSR sources. These models were generated by entering sensor-space *timelocked* ASSR data into the MNE computation, before applying a Welch PSD estimation (hanning windowed, frequency resolution = 1 Hz), with zero-padding to 1 s. Subsequently, the individual MNE solutions were morphed to a common fsaverage template. The Brainnetome Atlas<sup>245</sup> and PALS-B12 Atlas<sup>246</sup> (based on Brodmann areas) were used to demarcate the cortical space for analyses comparing region-of-interests.

## 3.7 ETHICAL CONSIDERATIONS

All three studies were approved by the Regional Ethics Review Board in Stockholm (Dnr: 2017/998-31/2). Both written and oral informed consents were obtained from all participants prior to the experiment. All participants received a monetary compensation of SEK 600 (~EUR 60). The risks associated with this project are generally small and can be identified as the following:

- i) Participant experiences discomfort during MRI scanning due to claustrophobia. If such instances occur, although rarely, the scan is immediately cancelled. All personnel running the MRI scan are required to undergo safety training and follow standardized safety routines prior to the scanning. While there are risks associated with the presence of metal in the body, experimenters are trained to screen participants for the presence of metal before entry into the MR room.
- ii) All MRI scans will be sent to a qualified medical physician/radiographer for screening of any possible health indication. Procedures are put in place to ensure that any medically-relevant findings will be followed-up as in regular medical care. Findings of

non-biological significance are not disclosed to the participant in accordance with standard practice in clinical assessment of MRI.

- iii) MEG (including EOG and ECG) is considered completely risk free.
- iv) The risks associated with auditory stimulation are considered to be very small. The auditory stimulation equipment has undergone safety assessment, and its output volume is calibrated to be 70 – 75 decibels according to each participant's comfort level. The output volume is controlled via an amplifier in a closed cabinet in the MEG lab to prevent unintentional adjustments of volume.
- v) Written information related to all associated risks is provided to the participant before inclusion in the experiment. Participants are informed before the experiment that their participation is strictly voluntary and that they reserve full right to withdraw from participation without providing any reason at any point of time. The experimenter is also obligated to answer any supplementary questions regarding the experiment should the participant enquires.

All data collected, including personal information, are stored confidentially and only used for research purposes. Reusing of MEG or MRI data is only allowed upon the consent of the participant.

## 4 RESULTS & DISCUSSION

The key findings from Study I – III can be summarized into the following four categories: 1.1) Behavioral Results, 1.2) ASSR Sources and Characteristics, 1.3) Attentional Modulation of the ASSR, and 1.4) ASSR and MSI. This section will discuss each of the four categories listed above by consolidating results from across all studies. For technical details regarding each result (e.g. p-values), please refer to the original articles in the Appendix.

### 4.1 BEHAVIORAL RESULTS

Overall, participants performed significantly above chance in the MDT tasks for all studies. There was no significant difference in performance between directing attention to the lowest-pitched or highest-pitched voice, and neither was there any difference in performance between the three variations of the MDT task. This suggests that the MDT task difficulty was unaffected by the widths of pitch difference between voices (Study I vs Study II), or whether the voices were separated or overlapping in time (Study II vs Study III). In all three studies, performance correlated significantly with MSI (Fig. 4A).

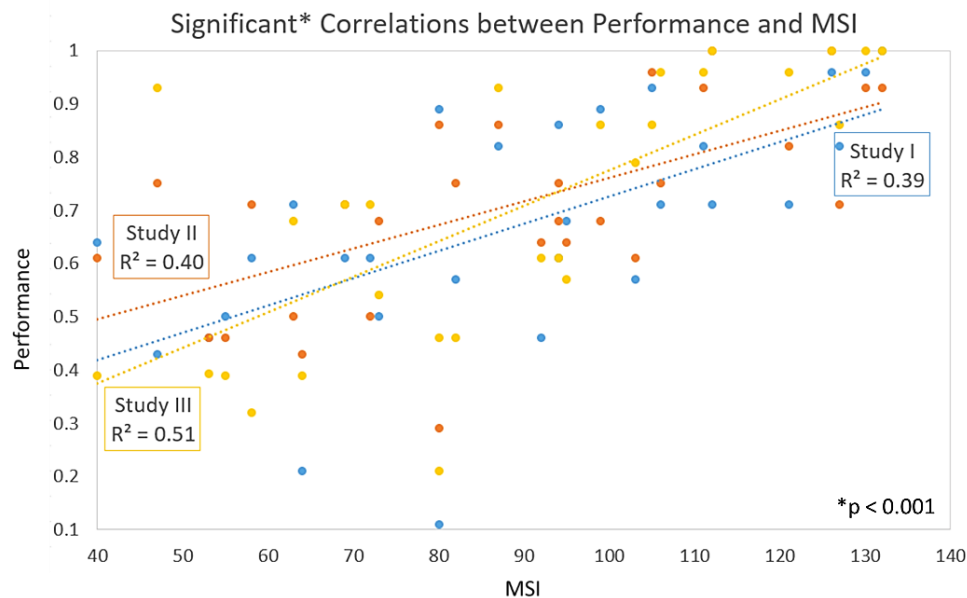


Figure 4A. Behavioural performance in the MDT tasks from Study I to III correlated significantly with MSI with  $p < 0.001$ . Note that the  $R^2$  value for Study II varies slightly from the value reported in the published article (Study II) as the above calculation included all participants whereas some participants were excluded in the publication.

### 4.2 ASSR SOURCES AND CHARACTERISTICS

In Study I, MNE distributed source models identified multiple ASSR sources from a variety of frontal, temporal and parietal regions. Unsurprisingly, sources with the strongest power were found in the primary auditory cortical regions, followed by parietal and frontal sources. To address the question if these sources were truly active and independent or caused by field spread from strong sources within the primary auditory cortex, we used a point spread

function for a simulated 41 Hz sine wave at the left and right auditory cortices, which showed large and systematic differences between the point spread model and our observed ASSR sources (see Fig. 4B). Specifically speaking, the modelled point spread function displayed a much lower maximum power and less extensive coverage of the cortex compared to the MNE solution, indicating that the observed frontal and parietal activity cannot be explained solely by the signal spread of the primary auditory source. Although several activated regions in the parietal and secondary auditory cortices overlap with the point spread function, the much larger power generated by the MNE solution indicates that additional sources outside the primary auditory cortex must be present to contribute to the additional signal power. As further support for this interpretation, previous EEG<sup>42</sup> and positron emission tomography (PET)<sup>40-41</sup> studies have also found multiple sources generating the 40 Hz ASSR, including many regions outside the auditory pathway. These regions, especially the frontal areas, are commonly overlooked in ASSR-attention studies, which typically place exclusive focus on stronger sources within the primary auditory cortex<sup>12-17</sup>, especially when the ASSR sources are modelled as dipoles.

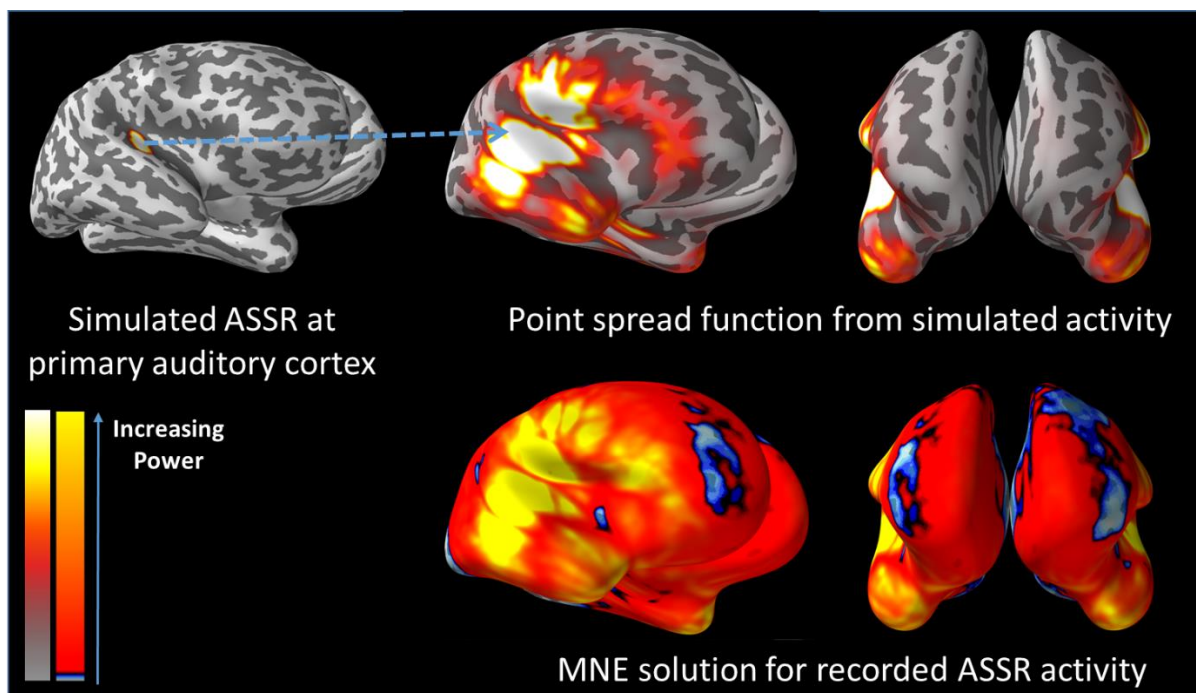
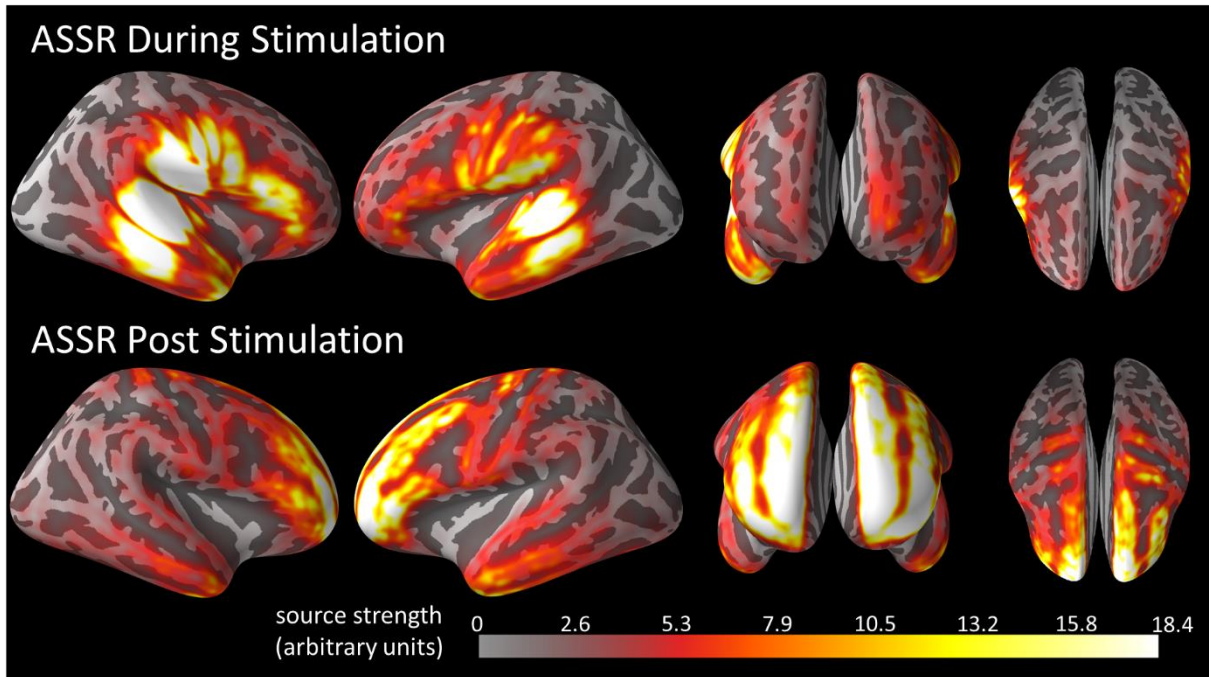


Figure 4B. (Top row) Point spread function from stimulated 41 Hz ASSR activity at the primary auditory cortex in both hemispheres. (Bottom row) Computed MNE solution of recorded 41 Hz ASSR activity during actual experiment. The colorbars on the left are for reference only to indicate the direction of increasing power and not drawn to scale as the simulated and recorded activities are plotted with different scales. The recorded activity was approximately  $10^4$  times stronger than the point spread function. Only the right lateral and frontal views are shown in the above figure. Please refer to Figures 4A and 4B of Study I for the other views and the exact range of the respective scales used.

Following these results, Study II revealed that the activation patterns of the frontal and temporal ASSR sources vary across time (Fig. 4C). In particular, temporal sources were more prominent during stimulus playback whereas frontal sources dominated in the post-stimulus time segment. This discovery was surprising and unprecedented as frontal ~40 Hz ASSR



activity is typically much weaker than temporal ASSR activity<sup>240, 247-251</sup>, and hence this is the first time any study has reported stronger frontal than temporal  $\sim 40$  Hz ASSR sources, albeit only during post-stimulation.

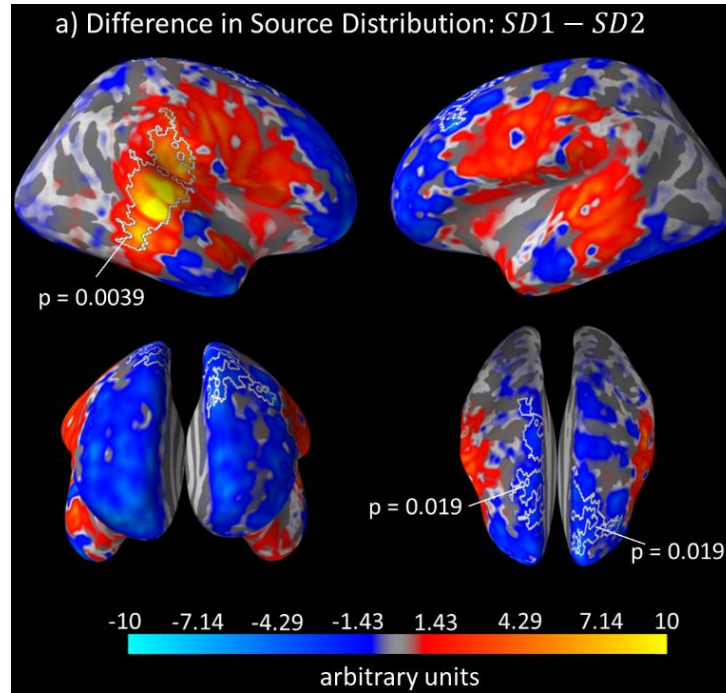


**Figure 4C.** Comparison of the 41 Hz ASSR cortical source distributions during (top row) and post (bottom row) stimulation. As shown above, the strongest ASSR sources during stimulation were found in the temporal primary and secondary auditory cortices and parts of the parietal cortex. In contrast, the strongest post-stimulus ASSR sources resided mainly in the frontal cortices. Orientation views from left to right: right lateral, left lateral, frontal, top.

Moreover, the results also showed that the ASSR during and after stimulation exhibited different sensitivities to experimental factors and stimulus properties. On one hand, the ASSR during stimulation was modulated by several factors: (i) It clearly decreased in power across voices, reflecting the physical differences in volume and  $f_c$  between these voices, (ii) its power correlated positively with individual MSI (see section 4.4), and (iii) increased with selective attention. On the other hand, the post-stimulus ASSR remained relatively unaffected by the same factors. Firstly, the pattern of decreasing power corresponding to the volume decrement from the lowest to highest pitched voice was not observed in the post-stimulus ASSR. Since no external driving input was present after the auditory stimulation, we regard the post-stimulus ASSR as an endogenously generated oscillatory neural activity. This could explain why the ASSR during stimulation, which we believe to be primarily generated by bottom-up driven neural components, is modulated by the acoustic properties (i.e. loudness) of the auditory stimulus. These bottom-up driven components are absent in the post-stimulus ASSR, thereby rendering it unaffected by the acoustic differences across voices. By similar arguments, the effect of individual musical sophistication on the ASSR may also occur on a bottom-up driven auditory processing level, possibly explaining why the post-stimulus ASSR seemed unaffected by the MSI. In addition, the results also show that the power modulating effects of selective attention, at least in the context of the ASSR, occur during the active

process of competition between auditory objects, but do not persist beyond the stimulation period. While the ASSR itself reflects a neural representation of the periodic envelope of the stimulus, the results suggest that the brain progresses from a sensory processing and discrimination state (i.e. identifying and paying attention to a tone) driving the ASSR at the auditory cortices during stimulation, to a different frontal-dominated state during which the periodic neural representation is still maintained (i.e. the post-stimulus ASSR), but is not sensitive to external factors such as the stimulus' properties or to the perception of and attention to it. Collectively, these findings indicate that the ASSR during and after stimulation reflect very different cortical processes, and that the post-stimulus ASSR is not simply a ringing extension generated by the same sources underlying the ASSR during stimulation. While evidence from Study II supports the existence of the post-stimulus ASSR, its functional relevance remains largely speculative and thus a question to be addressed by future studies. To the best of our knowledge, this is the first time any study has characterized the post-stimulus ASSR, thereby adding a new dimension to the existing literature and potential uses of the ASSR for research, especially in the neuroscientific understanding of higher executive processing such as selective attention and music perception.

While Study II investigated the development of the ASSR overtime, Study III focused on how the ASSR power and source distribution changes as the number of concurrent overlapping voices increases. Generally across sensors, the ASSR decreases to approximately a third of its power when two voices were presented simultaneously (dual-voice), as compared to when the voices were presented one at a time (single-voice). Source analysis showed that the extent of ASSR power suppression due to simultaneous sources was significant across hemispheres, lobes and voices, although voice effects were likely to be confounded by the adjusted volume differences and thus not discussed further. To elaborate, the temporal lobe experienced the largest suppression (reduced to 43%) from simultaneous voices followed by the parietal (reduced to 50%) then frontal (reduced to 68%) lobe. Additionally, during dual-voice listening, the fraction of activated ASSR sources decreased in the temporal-parietal regions but increased in the frontal regions compared to single-voice listening (Fig. 4D), explaining why the frontal ASSR power experienced the least degree of suppression from simultaneous ASSR sources (because part of this suppression was compensated by the increase in fractional resources allocated to the frontal region). We believe that this could be because identifying, separating, selectively directing and maintaining attention to the target voice was more cognitively demanding in the dual-voice scenario, leading to more resources being allocated to the frontal region, an area known to be central in the execution and maintenance of selective attention<sup>252-254</sup>. As before, this is the first time any study has attempted to spatially characterize the effect of an additional simultaneous source on the ASSR power and its source distribution, thus advancing our understanding of the ASSR and its potential applications.



**Figure 4D.** Differences in the distribution of ASSR sources between single-voice and dual-voice processing across the frontal, temporal and parietal cortices. Each source (vertex) was normalized by division over the sum of the ASSR power across all 20484 vertices at a single-subject level, thereby expressing the power of each source as a fraction of the total ASSR power. These values were also averaged across the lowest-pitched and highest-pitched voices. All clusters with  $p < 0.05$  obtained from a cluster-based permutation test of the single-voice minus dual-voice difference ( $SD1 - SD2$ ) are demarcated in white with labelled corresponding p-values. Orientation views (clockwise starting from top-left): right lateral, left lateral, top, frontal.

Taken together, the findings from all three studies collectively support the interpretation that ASSR sources in the frontal and temporal lobes are independent, characteristically different, as well as functionally distinct. Firstly, the frontal and temporal ASSR sources exhibited different sensitivities towards top-down selective attention, between top-down and bottom-up selective attention, as well as the physical parameters of the driving stimulus. Moreover, their respective activation patterns fluctuated uniquely across time and the sources were also suppressed to different degrees when a competing simultaneous voice was added.

### 4.3 ATTENTIONAL MODULATION OF THE ASSR

At a general sensor level, results from Study I and II indicated that top-down selective attention enhanced the ASSR by 14 % and 11% respectively. Furthermore, in Study I, no effect of attentional enhancement was observed on the middle-pitched reference voice which participants were never instructed to attend to, providing further evidence that the enhancement stemmed from top-down selective attention towards the attended voices (i.e. lowest-pitched and highest-pitched voices). In addition, ERF results mirrored these findings, demonstrating significant attentional enhancement specifically for the lowest-pitched and highest-pitched voices but not the middle-pitched voice. Since the attentional enhancement effect on ERFs is already a well-established phenomenon in literature<sup>13, 21-22, 26</sup>, the ERF results indicated that selective attention was successfully manipulated (i.e. directly exclusively to the instructed voice) as intended by the MDT task. With two simultaneous

voices in Study III however, no significant enhancement by top-down attention was observed at sensor level, although there was a significant 15% average enhancement by bottom-up attention. We believe that the suppression effect from simultaneous voices led to weaker signal-to-noise ratios and thus reduced sensitivity at picking up the modulation effect from top-down selective attention. This problem was mitigated with subsequent source level analyses, which managed to extract the attentional effect on the ASSRs eventually. The modulation in ASSR power due to selective attention supports the notion of a top-down regulated gain control mechanism of attention, proposed by many authors in the past<sup>25-26, 39, 41-42</sup>. For all three studies, there was no significant difference between the attentional modulation across voices.

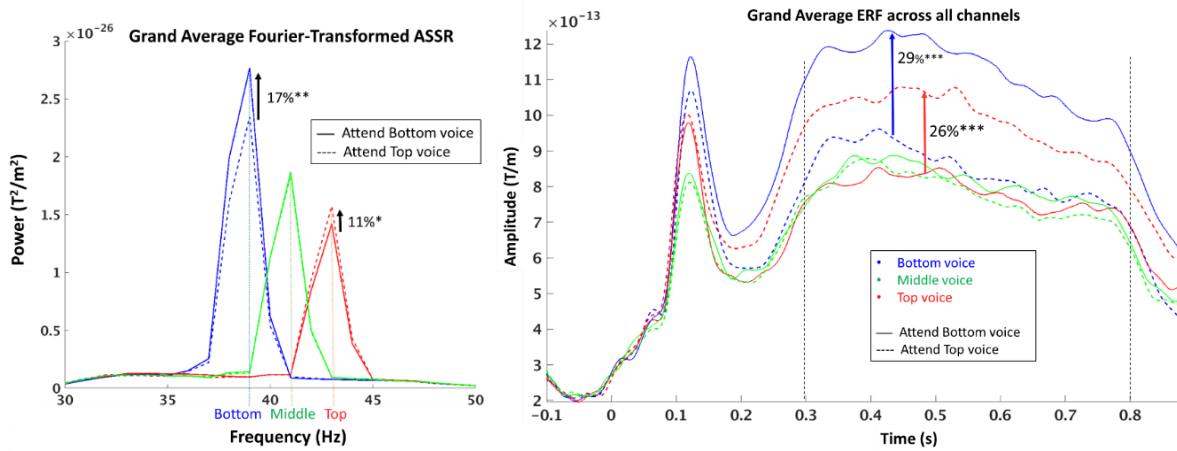


Figure 4E. ASSR (left) and ERF (right) results correspond with one another, showing that attentional enhancement occurred exclusively on the attentionally-manipulated lowest-pitched (**Bottom**) and highest-pitched (**Top**) voices, but not the middle-pitched (**Middle**) voice that participants were never instructed to attend to.

The majority of studies that investigated the effect of attention on the ASSR adopted intermodal designs<sup>47, 145, 255</sup>, or an intramodal dichotic listening task<sup>13, 99, 122, 144</sup>, wherein participants shift attention between the left and right ears. Importantly, our results provide the first clear evidences that selective attention enhances the ASSR when attention was directed to perceptual features (i.e. pitch and timing), rather than spatial separation between sounds as in previous dichotic listening designs. This is important and relevant as feature-based sound separation and identification is a significant component of selective attention in natural cocktail party-like settings.

Source level analysis was carried out primarily to compare the degree of attentional modulation across different cortical regions and identify areas that are most optimal for assessing selective attention. Cluster-based permutation results from Study I suggested that apart from the temporal cortical regions, sources in the frontal and parietal regions could also be modulated by selective attention. These regions cover the middle frontal gyrus, inferior frontal gyrus, orbital gyrus, inferior parietal lobule, precentral and postcentral gyrus, insular gyrus, superior temporal gyrus, middle temporal gyrus and posterior superior temporal sulcus (see Fig. 4F). Confirmatory statistical evidence from Study II later agreed with these findings,

showing significant ASSR attentional enhancement in each of the bi-hemispheric frontal, temporal, parietal and insular lobes. Both studies also showed that the frontal cortex experienced the largest attentional effect compared to the temporal and parietal cortices. This is not a surprising finding as the frontal regions has long been regarded as the established centre for attentional control in neuroscience literature, including auditory attention<sup>254, 256-257</sup> as well as attention in other sensory modalities<sup>252-253</sup>. With respect to the temporal regions, attentional enhancement of the ASSR in the auditory cortex has been reported by several studies, although limited to spatial<sup>13, 99, 122, 144</sup> and intermodal<sup>47, 145, 255</sup> attention as mentioned earlier. Evidence of auditory attentional modulation in the parietal cortex has also been reported in previous studies<sup>58, 258-261</sup>, although not within the ASSR domain, owing perhaps to the lack of documentation on ASSR sources outside the auditory cortex. In relation to this, the motor cortex, housed by the parts of the frontal and parietal lobes, is known to exhibit a robust entrainment to sensory stimulation rhythms that is also enhanced from attention<sup>134, 260, 262-263</sup>. Since the ASSR may be conceptualized as an entrainment (to the stimulus) itself, it is reasonable that ASSR activity and its attentional modulation was found in the motor cortex.

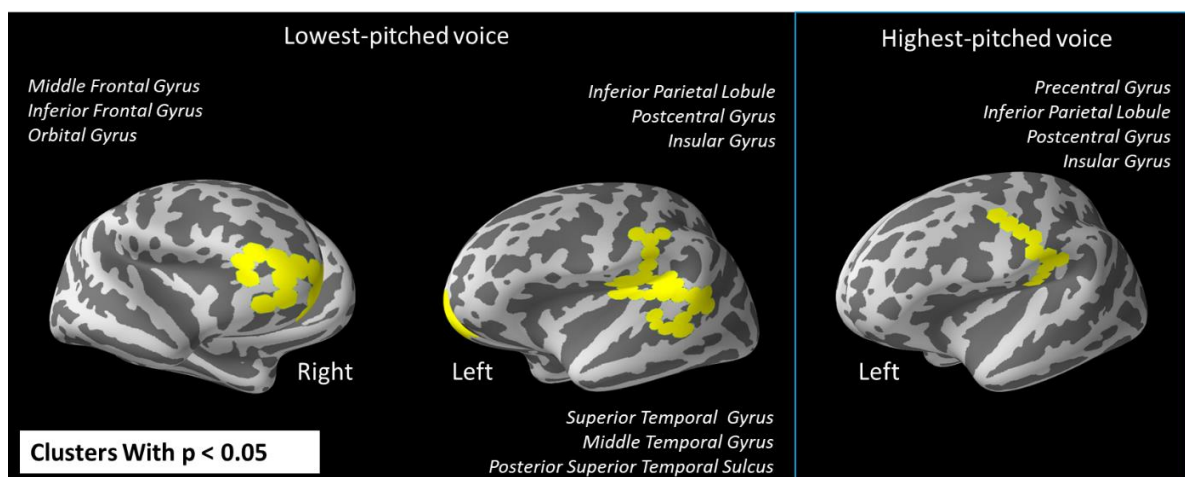


Figure 4F. Clusters with  $p < 0.05$  obtained from cluster-based permutation tests of the Attend-Unattend contrast for the lowest-pitched (left and centre) and highest-pitched (right) voice. ROIs containing vertices that belong to the cluster are described in *italics* above and below the brain. Please refer to Figure 5 of Study I for exact p-values corresponding to each clusters.

As for Study III (i.e with overlapping voices), the ASSR enhancement by top-down selective attention was found significant only in the left cortical frontal lobe, whereas at all other areas (i.e. right frontal lobe, bi-hemispheric temporal and parietal lobes) it failed to reach significance after Bonferroni correction over six tests. On the other hand, the effect of bottom-up attention was most significant in the temporal lobe, where the lower level sensory auditory cortices are situated. One possible explanation to these findings could be that involuntary bottom-up attention is associated with automatic stimulus processing mechanisms that predominate in the sensory cortices<sup>264</sup>, in contrast to how top-down factors such as greater utilization of voluntary attention and working memory modulate higher-level executive regions located frontally<sup>252-254, 264</sup>. This explanation aligns with studies demonstrating that top-down and bottom-up mechanisms are mediated by specialized neural



networks, albeit with partially-overlapping regions<sup>264-265</sup>. Moreover, additional analysis into interaction effects between bottom-up and top-down attention revealed a synergistic relationship between the two, in that the enhancing effect of top-down attention was more observable when bottom-up attention was also directed towards the same tone and vice versa. We postulate that bottom-up driven attention based on stimulus saliency on tone onset may help the listener to 'find' the tone more quickly and consciously direct resources towards it via top-down attentional mechanisms. However, top-down direction of attention toward a target tone can also be counteracted by bottom-up attention towards a competing simultaneous tone<sup>266</sup>, which may explain why the effect of top-down attention on the cued voice was not observable when the other competing voice experienced a salient change in pitch (and thus drew attention away from the cued voice through bottom-up mechanisms). As studies that directly address the effect of bottom-up attention on the ASSR are rare, even more so for how the interplay between top-down and bottom-up attention affects the ASSR, this work offers refreshing and novel insights into the field.

To summarize the findings from our three studies with regards to the attentional modulation of the ASSR, we have demonstrated the feasibility of using the ASSR to extract top-down attentional effects, first at sensor level, and then improved the sensitivity of the approach via source level analysis. Finally, we extended the application from voices separated in time (Study I and II) to overlapping voices, with the added component of extracting bottom-up selective attention (Study III).

#### **4.4 ASSR AND MUSICAL EXPERIENCE**

Study II showed that that musical sophistication influences both the ASSR power and its degree of modulation by top-down selective attention. Notably, the behavioural results showed a significant positive correlation between the MSI of participants and their performance scores in the MDT task ( $r = 0.67$ ), demonstrating the musical sensitivity of the task. More importantly, our results show that the ASSR power correlates strongly with MSI scores ( $r = 0.40$ ). These results are in agreement with previous studies<sup>174, 200, 267</sup> showing that neural processing of auditory stimuli are enhanced by musical abilities and experience, and may reflect an improvement in auditory skills owing to musical training. Additionally, we found a strong positive correlation between MSI and the degree of attentional enhancement in the left ( $r = 0.39$ ) and right ( $r = 0.44$ ) parietal cortices with no such relationships to attentional enhancement in temporal and frontal cortices. The following Figure 4G. depicts the correlations between ASSR power and attentional modulation against MSI.

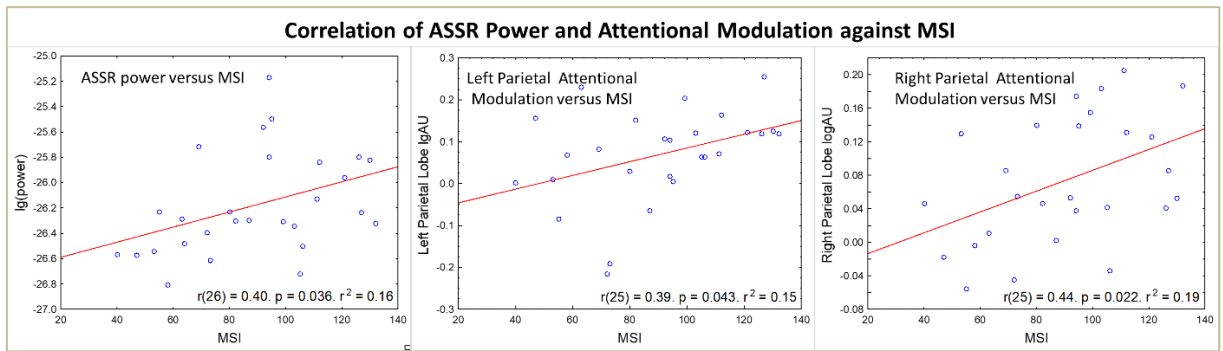


Figure 4G. Correlations between ASSR power (left), and attentional modulation at the parietal cortices (centre and right), against MSI. Pearson's  $r$ ,  $R^2$  and  $p$ -values are indicated directly below each graph.

In earlier studies<sup>268-270</sup>, parietal regions have been shown to exhibit sensitivity to musical experience and ability. This has been attributed to the fact that these regions play important roles for musical skill learning and performance. These evidences further strengthen the belief that musical experience is related to selective listening ability, a phenomenon which can be reflected in recorded neural signals like the ASSR. Speculatively, even better correlations between the ASSR and musical experience may be achieved by using more naturally musical stimulus or musically challenging experimental tasks with greater specificity and sensitivity to individual musical experiences and listening ability in future studies.





## 5 LIMITATIONS & RECOMMENDATIONS FOR FUTURE WORK

While we believe that our present results make novel contributions to the existing literature on ASSR methodology as well as to the neuroscientific understanding of selective auditory attention, the studies bear several limitations and calls for further work to clarify the current findings.

### 5.1 GENERALIZABILITY OF FINDINGS

The behaviour of ASSRs is notorious for being rather sensitive towards variations in experimental design, stimulus properties and analytical approaches, making cross-study comparisons tricky to interpret. For instance, the ASSR waveform changes with type of stimulus<sup>115</sup>, and its power and activity distribution changes with the number of simultaneous sources<sup>138</sup>, carrier frequency, modulation frequency and modulation depth<sup>114-115, 249, 271</sup>. Many of these cross-experimental variations affect the signal-to-noise ratios for ASSRs and increase the minimum recording time needed to separate individual ASSRs or capture ASSR modulations across conditions, which may make it problematic or even infeasible to directly apply the use of the ASSR from one experimental setting to another, as in the case of comparing single-source to multi-source designs. These problems are further complicated by the fact that the ASSR is generated by an extensive collection of sources across the neural cortex<sup>51-52</sup>, which differ in strength as well as how they are modulated by behavioural and individual factors like selective attention and musicality respectively. Related to this point, our results build on ASSR sources generated by AM frequencies close to 40 Hz and may not be generalizable across ASSRs at other frequencies which have the tendency to display different source distribution patterns<sup>249</sup>. Furthermore, the ASSR modulation, by attention for example, can also be influenced by task complexity (e.g. number of sources, ease of separating sources), stimulus properties and the chosen analysis route, as observed in previous works with mixed conclusions about the attention effect<sup>99, 122, 144-145, 149</sup>, and also in our three studies where the computed amount of attentional modulation varied substantially across studies. Hence, it is important to examine how the different experimental and individual factors can influence the manifestation of the ASSR, in order to adapt the use of ASSRs in different situations and study designs, as well as make cross-experimental comparisons more valid and easier to interpret.

### 5.2 TOWARDS NATURALISTIC MUSICAL STIMULI

To date, many different types of stimuli, ranging from clicks to tone bursts, have been used to evoke the ASSR<sup>128, 136, 272</sup>. These stimuli however, carry a clearly discernable artificial sound and do not usually exist in our everyday auditory environment. This presents researchers the problem of associating findings from using these artificial stimuli with real-life listening situations. While the use of sine tones may not be an accurate representation of natural

auditory mixtures such as a large choir or a symphony orchestra, the complexity of the stimulus cocktail in our work is the first of its kind and serves as a stepping stone for future studies on selective attention in more natural and sophisticated environments. Potentially, a deeper understanding of multi-source ASSR could pave the way towards applying frequency-tagging to more complex and naturalistic soundscapes for translating experimental findings into real-life applications.

For future studies, we recommend the progression towards using more natural auditory stimuli which better reflect real-life listening situations. Although tone sequences are not new to ASSR research, only pure sine tones were used before<sup>272-274</sup>. No current study, to the best of our knowledge, has attempted to use natural-sounding instrumental tones. The selection of the type of instrument to use is restricted by the availability of instruments with relatively constant-amplitude sound waveform that is suitable for frequency-tagging. For example, a flute sound's stable and smooth natural acoustic envelope would be more suitable for frequency-tagging than that of a piano tone which has steep initial amplitude decay. Moreover, there is a clear trade-off between ASSR power and naturalness of the stimulus – increasing the modulation depth (m) elicits a stronger ASSR but also makes the stimulus sound rougher and less natural. Thus, careful optimization is needed to find a balance between keeping the stimuli sounding natural whilst maintaining sufficient ASSR power to capture behavioral effects like selective attention. Using natural instrumental stimuli in experimental setups allows for better association with the implications of musical expertise, especially instrument-specific expertise. By recruiting expert musician participants, namely instrumentalists corresponding to the type of stimulus to be used (e.g. flute players for flute tone stimulus), one can explore the effect of instrument-specific expertise on ASSR modulation. For example, while it may be intuitive that musicians should outperform non-musicians in auditory stream separation tasks, little is known about whether musicians are more easily distracted by tones that are familiar due to their training (i.e. is it harder for a violinist to ignore a violin playing in the background?) and how this affects selective attention. In the arena of music, many have conjectured theoretical models to describe attention in multi-voiced music<sup>275,276</sup>. Yet, none have succeeded in substantiating their models with conclusive empirical data. Our methodological approach for tracking attention can potentially achieve this challenging goal, as the combination of MEG and frequency-tagging makes it possible to directly assess attentional modulation with a unique spatiotemporal resolution at the neural level, even in real-time theoretically speaking. By delving deeper into the neural underpinnings governing auditory selective attention and musical training, such studies can have profound implications for learning, especially in aiding individuals with attention-deficits or other learning disabilities.

From another perspective, while the above suggested studies provide useful insights on how attention is maneuvered in the auditory system through understanding how it is processed in music, the converse is also true – that is, exploring how attention is modulated in music may

potentially unearth novel discoveries pertaining to the realm of musical improvisation and creativity, which until today, remains an elusive black box. Taking further advantage of our unique frequency-tagging technique, it is possible to capture spontaneous attentional modulations between and within instruments during live group performance and improvisation (i.e. in an ensemble/band). This approach can probe into how attention is distributed among a performer's own instrument's output and that produced by his group mates' instruments.

### **5.3 TRAINING SELECTIVE ATTENTION & OTHER COGNITIVE ABILITIES**

One of the long-term goals of our research is to investigate the potential of enhancing auditory processing and cognition, such as selective listening, via training. Through studying the acquisition of musical expertise, we can understand how training facilitates better auditory streaming and selective attention, along with what neural mechanisms are involved. To this aim, our experimental design is sufficiently flexible and controllable for investigating the fundamental components of acoustic perception, namely, pitch, timing and timbre<sup>277</sup>, separately or in their various combinations, providing a comprehensive yet malleable avenue for research into the workings of auditory processing, as well as the neuroplastic changes induced by training auditory cognitive abilities.

In addition, with the knowledge that musical training can induce benefits that transfer to other related cognitive abilities<sup>191-201, 203, 231, 237</sup> (e.g. linguistic ability, working memory, motor coordination), our music-based approach offers the potential of using a more intriguing and relatable domain, that is music, to enhance various skills that are otherwise relatively mundane and general, such as selective attention or working memory.

### **5.4 ASSR ACROSS SPACE AND TIME**

When discussing the individual signal contribution from different cortical areas, a critical point to consider is the degree of independence of these sources as field spread is an unavoidable aspect of the MEG signal. We addressed this issue in depth with point spread simulations in Study I and characteristic differences in Study II and III (see section 4.2 under Results & Discussion), demonstrating a clear independence between frontal and temporal sources. However, the ASSR sources in the temporal, parietal and insular lobes often appear relatively close to one another, and in some cases even connected. It is then important to consider whether some of these sources arise from independent sources overlapping in space that cannot be differentiated by our source localization, or whether they are artefacts of signal spreading.

Although our sensor-space and source-space analyses provide statistical evidence affirming the attentional modulation of the ASSR per se, and our method could also demarcate regional contributions to the ASSR modulation relatively broadly (i.e. attentional effect likely in the frontal-parietal cortices), the current approach cannot pin-point more precisely which

locations, especially at a vertex level, are more susceptible to ASSR attentional modulation in a confirmatory manner. One major limitation that precludes our present analysis from demarcating smaller regions-of-interests was the presence of spatial variation between individuals in their respective attentional modulations, which necessitated the use of moderately large predefined regions-of-interests to cover most of the relevant areas when using a common atlas across all subjects. This approach resulted in the inclusion of numerous vertices with low signal power or weak attentional modulation in the calculation, consequently diluting and reducing the average computed extent of effect. Therefore, alternative methods to characterize the area of effect while accounting for such individual variation are needed for more precise mapping and effective capture of the ASSR attentional modulation.

Another obvious limitation of our analytical approach is that the adopted PSD estimation methods collapse data across-time and thus lose information about the latency and duration of effect (i.e. attentional modulation), as well as the development of the effect over time. This is a practical problem heavily dependent on the acquired signal power (longer time windows needed for weaker signals), and can be mitigated with better signal-to-noise ratios. A more concentrated localization approach of each individual's ASSR activity can help to achieve higher signal-to-noise, thus allowing for shorter time windows in the PSD estimation and better time resolution with time-frequency analysis. Using time-frequency analyses, we recommend that future research aim to investigate and characterize the pattern of migration of ASSR sources and their attentional modulation across time and space, as well as how they may vary across individuals, which may be an important step towards a deeper understanding of the neural mechanisms involved in selective auditory attention and auditory perception.

## 6 CONCLUSIONS

In pursuit of a method to separate mixed neural signals generated by multiple simultaneous sounds into the individual constituent activities, we turned to frequency-tagging as a means to label, identify and isolate neural ASSRs uniquely, granting us the opportunity to study stimulus-specific neural responses even amidst the presence of other ongoing brain activity. While this approach may appear straightforward in theory, it has proven to be a tricky task to put into practice, owing to the sensitivity of the ASSR to a myriad of factors such as the signal-to-noise ratio, type of stimulus, modulation depth, complexity of task, as well as number of simultaneous auditory sources<sup>114-115, 138, 278</sup>. To better understand the nature of the ASSR for optimizing its usage in cognitive research, our work sought to characterize the ASSR across space and time, as well as examine its feasibility as a tool for extracting the effect of selective attention. To these aims, we employed a cognitively demanding “Melody Development Tracking” task that entailed participants to selectively focus attention to a designated target melody presented within a mixture of competing melodies. This paradigm has allowed us to test the feasibility of using the ASSR to study selective auditory attention in complex multi-source auditory scenarios.

Our results demonstrated that selective attention enhances the ASSR, and that this effect can be robustly observed at sensor as well as source level analysis of MEG data. Furthermore, the ASSR attention modulation varied considerably across the cortex, and was strongest in the frontal regions, which is well-aligned with current literature marking the pre-frontal cortex as the centre for attentional control<sup>252-254</sup>. We also showed that the ASSR exhibits significant attentional enhancement in the bilateral frontal, temporal, parietal and insular lobes, with the degree of attentional enhancement correlating positively with individual musical sophistication scores specifically at the left and right parietal cortices – areas that are commonly associated with musical training<sup>268-270</sup>. In addition, the ASSR was primarily driven by frontal sources during the post-stimulation period, whereas temporal sources account for most of the ASSR activity during auditory stimulation, wherein the ASSR was also sensitive to the stimulus physical features, selective attention and participant’s musicality. In the case of simultaneously overlapping voices, suppression of the ASSR power was observed to be inhomogeneous across the neural cortex, with more suppression in the temporal-parietal regions than in the frontal regions. This was partly attributed to the reallocation of total neural resources from temporal-parietal to frontal areas during the processing of simultaneous voices, which can be explained by the notion that the higher complexity in the simultaneous auditory mix engaged more strongly higher-level cognitive processing mechanisms housed in the frontal regions<sup>252-254, 264</sup>. Moreover, we learnt that while the suppression rendered the ASSR less sensitive at picking up the effect of top-down attention at sensor level, the attentional modulation could still be detected using a more localized source level approach. On a related note, the effect of ASSR enhancement due to bottom-up auditory attention could be observed both at sensor and source level. While the

effects of top-down and bottom-up attention seemed to complement one another in terms of power enhancement, they were concentrated at different regions, with top-down attention strongest at the frontal lobe and bottom-up attention centred at the lower-level auditory sensory areas.

In conclusion, our results present clear evidence that selective auditory attention to the perceptual features of musical streams increases the ASSR power of the attended stream in the frontal (strongest attention effect), temporal and parietal cortices. The many novel contributions of the present work advanced our understanding of the ASSR and demonstrated its suitability as an effective tool for investigating the neural correlates of cognitive processes such as music perception and selective attention. Notably, our work highlights the importance of including non-auditory areas in ASSR application studies and advocates a novel ‘beyond the temporal cortex’ perspective on ASSR modulation. These findings are a significant leap towards separating and extracting neural signals in complex soundscapes, from the conventional simplistic experimental set-ups. Eventually, this would pave the way towards using frequency-tagging for precise tracking of individual auditory sources in more naturalistic soundscapes with greater degree of complexity, that can better translate experimental findings into explanations of real-life cognitive phenomena.

## 7 REFERENCES

1. Cherry, E. C., Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America* 1953, 25 (5), 975-979.
2. Corbetta, M.; Shulman, G. L., Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* **2002**, 3 (3), 201-215.
3. Broadbent, D. E., *Perception and communication*. Pergamon Press: 1958.
4. Treisman, A. M., Selective Attention In Man. *British Medical Bulletin* **1964**, 20 (1), 12-16.
5. Logothetis, N. K., What we can do and what we cannot do with fMRI. *Nature* **2008**, 453 (7197), 869-878.
6. Ekstrom, A., How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. *Brain research reviews* **2010**, 62 (2), 233-244.
7. Hämäläinen, M. S.; Lundqvist, D., MEG as an Enabling Tool in Neuroscience: Transcending Boundaries with New Analysis Methods and Devices. In *Magnetoencephalography: From Signals to Dynamic Cortical Networks*, Supek, S.; Aine, C. J., Eds. Springer International Publishing: Cham, 2019; pp 3-39.
8. Baillet, S., Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience* **2017**, 20 (3), 327-339.
9. Tudor, M.; Tudor, L.; Tudor, K. I., [Hans Berger (1873-1941)--the history of electroencephalography]. *Acta medica Croatica : casopis Hrvatske akademije medicinskih znanosti* **2005**, 59 (4), 307-13.
10. Berger, H., Uber das Elektrenkephalogramm des Menschen. *Arch. Psychiatr. Nervenkr.* **1929**, 87, 527.
11. Cohen, D., Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer. *Science* **1972**, 175 (4022), 664-6.
12. Lopes da Silva, F., EEG and MEG: Relevance to Neuroscience. *Neuron* **2013**, 80 (5), 1112-1128.
13. Bidet-Caulet, A.; Fischer, C.; Besle, J.; Aguera, P.-E.; Giard, M.-H.; Bertrand, O., Effects of Selective Attention on the Electrophysiological Representation of Concurrent Sounds in the Human Auditory Cortex. *The Journal of Neuroscience* **2007**, 27 (35), 9252-9261.
14. Zion Golumbic, Elana M.; Ding, N.; Bickel, S.; Lakatos, P.; Schevon, Catherine A.; McKhann, Guy M.; Goodman, Robert R.; Emerson, R.; Mehta, Ashesh D.; Simon, Jonathan Z.; Poeppel, D.; Schroeder, Charles E., Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a “Cocktail Party”. *Neuron* **2013**, 77 (5), 980-991.

15. Ahveninen, J.; Jääskeläinen, I. P.; Raij, T.; Bonmassar, G.; Devore, S.; Hämäläinen, M.; Levänen, S.; Lin, F.-H.; Sams, M.; Shinn-Cunningham, B. G.; Witzel, T.; Belliveau, J. W., Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences* **2006**, *103* (39), 14608-14613.
16. Picton, T., Hearing in time: evoked potential studies of temporal processing. *Ear Hear* **2013**, *34* (4), 385-401.
17. Hari, R.; Aittoniemi, K.; Järvinen, M.-L.; Katila, T.; Varpula, T., Auditory evoked transient and sustained magnetic fields of the human brain localization of neural generators. *Experimental Brain Research* **1980**, *40* (2), 237-240.
18. Picton, T. W.; Woods, D. L.; Proulx, G. B., Human auditory sustained potentials. I. The nature of the response. *Electroencephalography and clinical neurophysiology* **1978**, *45* (2), 186-197.
19. Niedermeyer, E., *Electroencephalography : Basic Principles, Clinical Applications, and Related Fields*. LWW: Philadelphia, US, 2004.
20. Hansen, J. C.; Hillyard, S. A., Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and clinical neurophysiology* **1980**, *49* (3-4), 277-90.
21. Hansen, J. C.; Dickstein, P. W.; Berka, C.; Hillyard, S. A., Event-related potentials during selective attention to speech sounds. *Biological Psychology* **1983**, *16* (3-4), 211-224.
22. Woods, D. L.; Hillyard, S. A.; Hansen, J. C., Event-related brain potentials reveal similar attentional mechanisms during selective listening and shadowing. *Journal of Experimental Psychology: Human Perception and Performance* **1984**, *10* (6), 761-777.
23. Brattico, E.; Pallesen, K. J.; Varyagina, O.; Bailey, C.; Anourova, I.; Jarvenpaa, M.; Eerola, T.; Tervaniemi, M., Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study. *J Cogn Neurosci* **2009**, *21* (11), 2230-44.
24. Woldorff, M. G.; Hillyard, S. A., Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and clinical neurophysiology* **1991**, *79* (3), 170-191.
25. Hillyard, S. A.; Hink, R. F.; Schwent, V. L.; Picton, T. W., Electrical Signs of Selective Attention in the Human Brain. *Science* **1973**, *182* (4108), 177-180.
26. Woldorff, M. G.; Gallen, C. C.; Hampson, S. A.; Hillyard, S. A.; Pantev, C.; Sobel, D.; Bloom, F. E., Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90* (18), 8722-8726.
27. Näätänen, R.; Paavilainen, P.; Rinne, T.; Alho, K., The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology* **2007**, *118* (12), 2544-2590.



28. Pulvermüller, F.; Shtyrov, Y.; Hasting, A. S.; Carlyon, R. P., Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language* **2008**, *104* (3), 244-253.
29. Winkler, I.; Teder-Salejarvi, W. A.; Horvath, J.; Naatanen, R.; Sussman, E., Human auditory cortex tracks task-irrelevant sound sources. *Neuroreport* **2003**, *14* (16), 2053-6.
30. Schröger, E.; Paavilainen, P.; Näätänen, R., Mismatch negativity to changes in a continuous tone with regularly varying frequencies. *Electroencephalography and clinical neurophysiology* **1994**, *92* (2), 140-7.
31. Woldorff, M. G.; Hackley, S. A.; Hillyard, S. A., The Effects of Channel-Selective Attention on the Mismatch Negativity Wave Elicited by Deviant Tones. *Psychophysiology* **1991**, *28* (1), 30-42.
32. Woldorff, M. G.; Hillyard, S. A.; Gallen, C. C.; Hampson, S. R.; Bloom, F. E., Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology* **1998**, *35* (3), 283-92.
33. Jäncke, L.; Mirzazade, S.; Joni Shah, N., Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neuroscience Letters* **1999**, *266* (2), 125-128.
34. Tóth, B.; Farkas, D.; Urbán, G.; Szalárdy, O.; Orosz, G.; Hunyadi, L.; Hajdu, B.; Kovács, A.; Szabó, B. T.; Shestopalova, L. B.; Winkler, I., Attention and speech-processing related functional brain networks activated in a multi-speaker environment. *PLOS ONE* **2019**, *14* (2), e0212754.
35. Lipschutz, B.; Kolinsky, R.; Damhaut, P.; Wikler, D.; Goldman, S., Attention-Dependent Changes of Activation and Connectivity in Dichotic Listening. *NeuroImage* **2002**, *17* (2), 643-656.
36. Doesburg, S. M.; Green, J. J.; McDonald, J. J.; Ward, L. M., Theta modulation of inter-regional gamma synchronization during auditory attention control. *Brain Research* **2012**, *1431*, 77-85.
37. Mesgarani, N.; Chang, E. F., Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **2012**, *485* (7397), 233-236.
38. O'Sullivan, J. A.; Power, A. J.; Mesgarani, N.; Rajaram, S.; Foxe, J. J.; Shinn-Cunningham, B. G.; Slaney, M.; Shamma, S. A.; Lalor, E. C., Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex* **2014**, *25* (7), 1697-1706.
39. Lee, A. K. C.; Larson, E.; Maddox, R. K.; Shinn-Cunningham, B. G., Using neuroimaging to understand the cortical mechanisms of auditory selective attention. *Hearing Research* **2014**, *307*, 111-120.
40. Kauramäki, J.; Jääskeläinen, I. P.; Sams, M., Selective Attention Increases Both Gain and Feature Selectivity of the Human Auditory Cortex. *PLoS ONE* **2007**, *2* (9), e909.

41. Hillyard, S. A.; Vogel, E. K.; Luck, S. J., Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **1998**, 353 (1373), 1257-70.
42. Kerlin, J. R.; Shahin, A. J.; Miller, L. M., Attentional Gain Control of Ongoing Cortical Speech Representations in a “Cocktail Party”. *The Journal of Neuroscience* **2010**, 30 (2), 620-628.
43. Luck, S. J.; Chelazzi, L.; Hillyard, S. A.; Desimone, R., Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *Journal of neurophysiology* **1997**, 77 (1), 24-42.
44. Corbetta, M.; Miezin, F. M.; Dobmeyer, S.; Shulman, G. L.; Petersen, S. E., Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J Neurosci* **1991**, 11 (8), 2383-402.
45. Johansen-Berg, H.; Christensen, V.; Woolrich, M.; Matthews, P. M., Attention to touch modulates activity in both primary and secondary somatosensory areas. *Neuroreport* **2000**, 11 (6), 1237-1241.
46. Mima, T.; Nagamine, T.; Nakamura, K.; Shibasaki, H., Attention modulates both primary and second somatosensory cortical activities in humans: a magnetoencephalographic study. *Journal of neurophysiology* **1998**, 80 (4), 2215-2221.
47. Saupe, K.; Widmann, A.; Bendixen, A.; Müller, M. M.; Schröger, E., Effects of intermodal attention on the auditory steady-state response and the event-related potential. *Psychophysiology* **2009**, 46 (2), 321-327.
48. Karns, C. M.; Knight, R. T., Intermodal Auditory, Visual, and Tactile Attention Modulates Early Stages of Neural Processing. *Journal of Cognitive Neuroscience* **2009**, 21 (4), 669-683.
49. Hötting, K.; Rösler, F.; Röder, B., Crossmodal and intermodal attention modulate event-related brain potentials to tactile and auditory stimuli. *Experimental Brain Research* **2003**, 148 (1), 26-37.
50. Grady, C. L.; Van Meter, J. W.; Maisog, J. M.; Pietrini, P.; Krasuski, J.; Rauschecker, J. P., Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport* **1997**, 8 (11), 2511-6.
51. Manting, C. L.; Andersen, L. M.; Gulyas, B.; Ullén, F.; Lundqvist, D., Attentional modulation of the auditory steady-state response across the cortex. *NeuroImage* **2020**, 217, 116930.
52. Manting, C. L.; Gulyas, B.; Ullén, F.; Lundqvist, D., Auditory steady-state responses during and after a stimulus: Cortical sources, and the influence of attention and musicality. *NeuroImage* **2021**, 233, 117962.

53. Rinne, T.; Balk, M. H.; Koistinen, S.; Autti, T.; Alho, K.; Sams, M., Auditory selective attention modulates activation of human inferior colliculus. *Journal of neurophysiology* **2008**, *100* (6), 3323-7.
54. Xiang, J.; Simon, J.; Elhilali, M., Competing Streams at the Cocktail Party: Exploring the Mechanisms of Attention and Temporal Integration. *The Journal of Neuroscience* **2010**, *30* (36), 12084-12093.
55. Jäncke, L.; Shah, N. J., Does dichotic listening probe temporal lobe functions? *Neurology* **2002**, *58* (5), 736-43.
56. Rinne, T., Activations of human auditory cortex during visual and auditory selective attention tasks with varying difficulty. *The open neuroimaging journal* **2010**, *4*, 187-93.
57. Warren, J. D.; Griffiths, T. D., Distinct Mechanisms for Processing Spatial Sequences and Pitch Sequences in the Human Auditory Brain. *The Journal of Neuroscience* **2003**, *23* (13), 5799-5804.
58. Hill, K. T.; Miller, L. M., Auditory Attentional Control and Selection during Cocktail Party Listening. *Cerebral Cortex (New York, NY)* **2010**, *20* (3), 583-590.
59. Lee, A.; Rajaram, S.; Xia, J.; Bharadwaj, H.; Larson, E.; Hämäläinen, M.; Shinn-Cunningham, B., Auditory Selective Attention Reveals Preparatory Activity in Different Cortical Regions for Selection Based on Source Location and Source Pitch. *Frontiers in Neuroscience* **2013**, *6* (190).
60. O'Sullivan, J.; Herrero, J.; Smith, E.; Schevon, C.; McKhann, G. M.; Sheth, S. A.; Mehta, A. D.; Mesgarani, N., Hierarchical Encoding of Attended Auditory Objects in Multi-talker Speech Perception. *Neuron* **2019**, *104* (6), 1195-1209.e3.
61. Inui, K.; Okamoto, H.; Miki, K.; Gunji, A.; Kakigi, R., Serial and Parallel Processing in the Human Auditory Cortex: A Magnetoencephalographic Study. *Cerebral Cortex* **2005**, *16* (1), 18-30.
62. Nourski, K. V.; Steinschneider, M.; McMurray, B.; Kovach, C. K.; Oya, H.; Kawasaki, H.; Howard, M. A., 3rd, Functional organization of human auditory cortex: investigation of response latencies through direct recordings. *Neuroimage* **2014**, *101*, 598-609.
63. Davis, M. H.; Johnsrude, I. S., Hierarchical Processing in Spoken Language Comprehension. *The Journal of Neuroscience* **2003**, *23* (8), 3423-3431.
64. Puvvada, K. C.; Simon, J. Z., Cortical Representations of Speech in a Multitalker Auditory Scene. *The Journal of Neuroscience* **2017**, *37* (38), 9189-9196.
65. Kaas, J. H.; Hackett, T. A., Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences* **2000**, *97* (22), 11793-11799.
66. Obleser, J.; Zimmermann, J.; Van Meter, J.; Rauschecker, J. P., Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb Cortex* **2007**, *17* (10), 2251-7.

67. Steinschneider, M.; Nourski, K. V.; Rhone, A. E.; Kawasaki, H.; Oya, H.; Howard, M. A., 3rd, Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings. *Front Neurosci* **2014**, *8*, 240.
68. de Heer, W. A.; Huth, A. G.; Griffiths, T. L.; Gallant, J. L.; Theunissen, F. E., The Hierarchical Cortical Organization of Human Speech Processing. *The Journal of Neuroscience* **2017**, *37* (27), 6539-6557.
69. Coffman, B. A.; Salisbury, D. F., MEG Methods: A Primer of Basic MEG Analysis. In *Neuroimaging in Schizophrenia*, Kubicki, M.; Shenton, M. E., Eds. Springer International Publishing: Cham, 2020; pp 191-210.
70. Cohen, D., Magnetoencephalography: Evidence of Magnetic Fields Produced by Alpha-Rhythm Currents. *Science* **1968**, *161* (3843), 784-786.
71. Garcés, P.; López-Sanz, D.; Maestú, F.; Pereda, E., Choice of Magnetometers and Gradiometers after Signal Space Separation. *Sensors* **2017**, *17* (12), 2926.
72. Malmivuo, J.; Plonsey, R., Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields. Oxford University Press, USA: 1995.
73. Vrba, J.; Robinson, S., SQUID sensor array configurations for magnetoencephalography applications. *Superconductor Science and Technology* **2002**, *15* (9), R51.
74. Hillebrand, A.; Barnes, G. R., A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage* **2002**, *16* (3 Pt 1), 638-50.
75. Attal, Y.; Bhattacharjee, M.; Yelnik, J.; Cottureau, B.; Lefèvre, J.; Okada, Y.; Bardinet, E.; Chupin, M.; Baillet, S., Modeling and detecting deep brain activity with MEG & EEG. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference* **2007**, 2007, 4937-40.
76. Samuelsson, J. G.; Khan, S.; Sundaram, P.; Peled, N.; Hämäläinen, M. S., Cortical Signal Suppression (CSS) for Detection of Subcortical Activity Using MEG and EEG. *Brain Topography* **2019**, *32* (2), 215-228.
77. Roux, F.; Wibrals, M.; Singer, W.; Aru, J.; Uhlhaas, P. J., The Phase of Thalamic Alpha Activity Modulates Cortical Gamma-Band Activity: Evidence from Resting-State MEG Recordings. *The Journal of Neuroscience* **2013**, *33* (45), 17827-17835.
78. Cornwell, B. R.; Arkin, N.; Overstreet, C.; Carver, F. W.; Grillon, C., Distinct contributions of human hippocampal theta to spatial cognition and anxiety. *Hippocampus* **2012**, *22* (9), 1848-59.
79. Cornwell, B. R.; Carver, F. W.; Coppola, R.; Johnson, L.; Alvarez, R.; Grillon, C., Evoked amygdala responses to negative faces revealed by adaptive MEG beamformers. *Brain Res* **2008**, *1244*, 103-12.

80. Parkkonen, L.; Fujiki, N.; Mäkelä, J. P., Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. *Human brain mapping* **2009**, *30* (6), 1772-82.
81. Samuelsson, J. G.; Sundaram, P.; Khan, S.; Sereno, M. I.; Hämäläinen, M. S., Detectability of cerebellar activity with magnetoencephalography and electroencephalography. *Human brain mapping* **2020**, *41* (9), 2357-2372.
82. Ahlfors, S. P.; Han, J.; Belliveau, J. W.; Hamalainen, M. S., Sensitivity of MEG and EEG to source orientation. *Brain Topogr* **2010**, *23* (3), 227-32.
83. Hämäläinen, M.; Hari, R.; Ilmoniemi, R. J.; Knuutila, J.; Lounasmaa, O. V., Magnetoencephalography---theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics* **1993**, *65* (2), 413-497.
84. Jousmäki, V.; Hari, R., Cardiac Artifacts in Magnetoencephalogram. *Journal of Clinical Neurophysiology* **1996**, *13* (2), 172-176.
85. Taulu, S.; Simola, J., Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Physics in medicine and biology* **2006**, *51* (7), 1759-68.
86. Taulu, S.; Kajola, M.; Simola, J., Suppression of Interference and Artifacts by the Signal Space Separation Method. 2004; Vol. 16, p 269-75.
87. Baillet, S., Forward and Inverse Problems of MEG/EEG. In *Encyclopedia of Computational Neuroscience*, Jaeger, D.; Jung, R., Eds. Springer New York: New York, NY, 2013; pp 1-8.
88. Hämäläinen, M. S.; Ilmoniemi, R. J., Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & Biological Engineering & Computing* **1994**, *32* (1), 35-42.
89. Gramfort, A.; Kowalski, M.; Hämäläinen, M., Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. *Physics in medicine and biology* **2012**, *57* (7), 1937-61.
90. Sekihara, K.; Nagarajan, S. S., *Adaptive spatial filters for electromagnetic brain imaging*. Springer Science & Business Media: 2008.
91. Veen, B. D. V.; Buckley, K. M., Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Magazine* **1988**, *5* (2), 4-24.
92. Mosher, J. C.; Leahy, R. M., Recursive MUSIC: a framework for EEG and MEG source localization. *IEEE transactions on bio-medical engineering* **1998**, *45* (11), 1342-54.
93. Salmelin, R.; Kujala, J.; Liljeström, M., Magnetoencephalography and the cortical dynamics of language processing. *Handbook of Neurolinguistics* **2019**.
94. Ding, N.; Melloni, L.; Zhang, H.; Tian, X.; Poeppel, D., Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci* **2016**, *19* (1), 158-164.

95. Doelling, K. B.; Poeppel, D., Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences* **2015**, *112* (45), E6233-E6242.
96. Lappe, C.; Steinstrater, O.; Pantev, C., Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection. *Front Hum Neurosci* **2013**, *7*, 260.
97. Tallon-Baudry, C., On the neural mechanisms subserving consciousness and attention. *Frontiers in psychology* **2012**, *2*, 397.
98. Dehaene, S.; Changeux, J.-P., Experimental and theoretical approaches to conscious processing. *Neuron* **2011**, *70* (2), 200-227.
99. Bharadwaj, H.; Lee, A. K. C.; Shinn-Cunningham, B., Measuring auditory selective attention using frequency tagging. *Frontiers in Integrative Neuroscience* **2014**, *8* (6).
100. Degerman, A.; Rinne, T.; Sarkka, A. K.; Salmi, J.; Alho, K., Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields. *The European journal of neuroscience* **2008**, *27* (12), 3329-41.
101. Alho, K.; Huotilainen, M.; Tiitinen, H.; Ilmoniemi, R. J.; Knuutila, J.; Näätänen, R., Memory-related processing of complex sound patterns in human auditory cortex: A MEG study. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience* **1993**, *4* (4), 391-394.
102. Martin, T.; McDaniel, M. A.; Guynn, M. J.; Houck, J. M.; Woodruff, C. C.; Bish, J. P.; Moses, S. N.; Kičić, D.; Tesche, C. D., Brain regions and their dynamics in prospective memory retrieval: A MEG study. *International Journal of Psychophysiology* **2007**, *64* (3), 247-258.
103. Kharkar, S.; Knowlton, R., Magnetoencephalography in the presurgical evaluation of epilepsy. *Epilepsy & Behavior* **2015**, *46*, 19-26.
104. Wilson, T. W.; Rojas, D. C.; Reite, M. L.; Teale, P. D.; Rogers, S. J., Children and Adolescents with Autism Exhibit Reduced MEG Steady-State Gamma Responses. *Biological Psychiatry* **2007**, *62* (3), 192-197.
105. Port, R. G.; Anwar, A. R.; Ku, M.; Carlson, G. C.; Siegel, S. J.; Roberts, T. P., Prospective MEG biomarkers in ASD: pre-clinical evidence and clinical promise of electrophysiological signatures. *The Yale journal of biology and medicine* **2015**, *88* (1), 25.
106. Schnitzler, A.; Timmermann, L.; Gross, J., Physiological and pathological oscillatory networks in the human motor system. *Journal of Physiology-Paris* **2006**, *99* (1), 3-7.
107. Vinding, M. C.; Tsitsi, P.; Waldthaler, J.; Oostenveld, R.; Ingvar, M.; Svenningsson, P.; Lundqvist, D., Reduction of spontaneous cortical beta bursts in Parkinson's disease is linked to symptom severity. *Brain Communications* **2020**, *2* (1).
108. Boto, E.; Meyer, S. S.; Shah, V.; Alem, O.; Knappe, S.; Kruger, P.; Fromhold, T. M.; Lim, M.; Glover, P. M.; Morris, P. G.; Bowtell, R.; Barnes, G. R.; Brookes, M. J., A new

generation of magnetoencephalography: Room temperature measurements using optically-pumped magnetometers. *NeuroImage* **2017**, *149*, 404-414.

109.Boto, E.; Holmes, N.; Leggett, J.; Roberts, G.; Shah, V.; Meyer, S. S.; Muñoz, L. D.; Mullinger, K. J.; Tierney, T. M.; Bestmann, S.; Barnes, G. R.; Bowtell, R.; Brookes, M. J., Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* **2018**, *555* (7698), 657-661.

110.Hill, R. M.; Boto, E.; Rea, M.; Holmes, N.; Leggett, J.; Coles, L. A.; Papastavrou, M.; Everton, S. K.; Hunt, B. A. E.; Sims, D.; Osborne, J.; Shah, V.; Bowtell, R.; Brookes, M. J., Multi-channel whole-head OPM-MEG: Helmet design and a comparison with a conventional system. *NeuroImage* **2020**, *219*, 116995.

111.Andersen, L. M.; Oostenveld, R.; Pfeiffer, C.; Ruffieux, S.; Jousmäki, V.; Hämäläinen, M.; Schneiderman, J. F.; Lundqvist, D., Similarities and differences between on-scalp and conventional in-helmet magnetoencephalography recordings. *PLoS One* **2017**, *12* (7), e0178602.

112.Faley, M. I.; Poppe, U.; Borkowski, R. E. D.; Schiek, M.; Boers, F.; Chocholacs, H.; Dammers, J.; Eich, E.; Shah, N. J.; Ermakov, A. B.; Slobodchikov, V. Y.; Maslennikov, Y. V.; Koshelets, V. P., Magnetoencephalography using a Multilayer hightc DC SQUID Magnetometer. *Physics Procedia* **2012**, *36*, 66-71.

113.Regan, D., Human brain electrophysiology : evoked potentials and evoked magnetic fields in science and medicine. Elsevier: New York, 1989.

114.Ross, B.; Borgmann, C.; Draganova, R.; Roberts, L. E.; Pantev, C., A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *J Acoust Soc Am* **2000**, *108* (2), 679-91.

115.Picton, T. W.; John, M. S.; Dimitrijevic, A.; Purcell, D., Human auditory steady-state responses. *Int J Audiol* **2003**, *42* (4), 177-219.

116.Ross, B.; Picton, T. W.; Pantev, C., Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hearing Research* **2002**, *165* (1), 68-84.

117.Azzena, G. B.; Conti, G.; Santarelli, R.; Ottaviani, F.; Paludetti, G.; Maurizi, M., Generation of human auditory steady-state responses (SSRs). I: Stimulus rate effects. *Hear Res* **1995**, *83* (1-2), 1-8.

118.Lins, O. G.; Picton, T. W., Auditory steady-state responses to multiple simultaneous stimuli. *Electroencephalography and clinical neurophysiology* **1995**, *96* (5), 420-32.

119.Aimoni, C.; Crema, L.; Savini, S.; Negossi, L.; Rosignoli, M.; Sacchetto, L.; Bianchini, C.; Ciorba, A., Hearing threshold estimation by auditory steady state responses (ASSR) in children. *Acta Otorhinolaryngol Ital* **2018**, *38* (4), 361-368.

120.Luts, H.; Desloovere, C.; Kumar, A.; Vandermeersch, E.; Wouters, J., Objective assessment of frequency-specific hearing thresholds in babies. *International journal of pediatric otorhinolaryngology* **2004**, *68* (7), 915-26.

- 121.Luts, H.; Wouters, J., Hearing assessment by recording multiple auditory steady-state responses: the influence of test duration. *Int J Audiol* **2004**, 43 (8), 471-8.
- 122.Müller, N.; Schlee, W.; Hartmann, T.; Lorenz, I.; Weisz, N., Top-Down Modulation of the Auditory Steady-State Response in a Task-Switch Paradigm. *Frontiers in Human Neuroscience* **2009**, 3, 1.
- 123.Kwon, J. S.; O'Donnell, B. F.; Wallenstein, G. V.; Greene, R. W.; Hirayasu, Y.; Nestor, P. G.; Hasselmo, M. E.; Potts, G. F.; Shenton, M. E.; McCarley, R. W., Gamma Frequency–Range Abnormalities to Auditory Stimulation in Schizophrenia. *Archives of General Psychiatry* **1999**, 56 (11), 1001-1005.
- 124.O'Donnell, B. F.; Vohs, J. L.; Krishnan, G. P.; Rass, O.; Hetrick, W. P.; Morzorati, S. L., Chapter 6 - The auditory steady-state response (ASSR): a translational biomarker for schizophrenia. In *Supplements to Clinical Neurophysiology*, Başar, E.; Başar-Eroğlu, C.; Özerdem, A.; Rossini, P. M.; Yener, G. G., Eds. Elsevier: 2013; Vol. 62, pp 101-112.
- 125.Luo, H.; Wang, Y.; Poeppel, D.; Simon, J. Z., Concurrent Encoding of Frequency and Amplitude Modulation in Human Auditory Cortex: MEG Evidence. *Journal of neurophysiology* **2006**, 96 (5), 2712-2723.
- 126.Galambos, R.; Makeig, S.; Talmachoff, P. J., A 40-Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A* **1981**, 78 (4), 2643-7.
- 127.Gutschalk, A.; Mase, R.; Roth, R.; Ille, N.; Rupp, A.; Hähnel, S.; Picton, T. W.; Scherg, M., Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. *Clinical Neurophysiology* **1999**, 110 (5), 856-868.
- 128.Presacco, A.; Bohórquez, J.; Yavuz, E.; Özdamar, Ö., Auditory steady-state responses to 40-Hz click trains: Relationship to middle latency, gamma band and beta band responses studied with deconvolution. *Clinical Neurophysiology* **2010**, 121 (9), 1540-1550.
- 129.Ross, B., Chapter 8 - Steady-state auditory evoked responses. In *Handbook of Clinical Neurophysiology*, Gastone, G. C., Ed. Elsevier: 2013; Vol. Volume 10, pp 155-176.
- 130.Schroeder, C. E.; Lakatos, P., Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences* **2009**, 32 (1), 9-18.
- 131.Giraud, A.-L.; Poeppel, D., Speech Perception from a Neurophysiological Perspective. In *The Human Auditory Cortex*, Poeppel, D.; Overath, T.; Popper, A. N.; Fay, R. R., Eds. Springer New York: New York, NY, 2012; pp 225-260.
- 132.Zion Golumbic, E. M.; Poeppel, D.; Schroeder, C. E., Temporal context in speech processing and attentional stream selection: A behavioral and neural perspective. *Brain and Language* **2012**, 122 (3), 151-161.
- 133.Ding, N.; Simon, J. Z., Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences* **2012**, 109 (29), 11854-11859.



- 134.Horton, C.; D'Zmura, M.; Srinivasan, R., Suppression of competing speech through entrainment of cortical oscillations. *Journal of neurophysiology* **2013**, *109* (12), 3082-3093.
- 135.Hari, R.; Hamalainen, M.; Joutsiniemi, S. L., Neuromagnetic steady-state responses to auditory stimuli. *J Acoust Soc Am* **1989**, *86* (3), 1033-9.
- 136.Draganova, R.; Ross B Fau - Wollbrink, A.; Wollbrink A Fau - Pantev, C.; Pantev, C., Cortical steady-state responses to central and peripheral auditory beats. **2008**, (1460-2199 (Electronic)).
- 137.Picton, T. W.; Skinner, C. R.; Champagne, S. C.; Kellett, A. J. C.; Maiste, A. C., Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *The Journal of the Acoustical Society of America* **1987**, *82* (1), 165-178.
- 138.John, M. S.; Lins, O. G.; Boucher, B. L.; Picton, T. W., Multiple auditory steady-state responses (MASTER): stimulus and recording parameters. *Audiology : official organ of the International Society of Audiology* **1998**, *37* (2), 59-82.
- 139.Ding, N.; Simon, J. Z., Neural representations of complex temporal modulations in the human auditory cortex. *Journal of neurophysiology* **2009**, *102* (5), 2731-43.
- 140.Oster, G., Auditory beats in the brain. *Scientific American* **1973**, *229* (4), 94-102.
- 141.Schwarz, D. W.; Taylor, P., Human auditory steady state responses to binaural and monaural beats. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **2005**, *116* (3), 658-68.
- 142.Lins, O. G.; Picton, T. W.; Boucher, B. L.; Durieux-Smith, A.; Champagne, S. C.; Moran, L. M.; Perez-Abalo, M. C.; Martin, V.; Savio, G., Frequency-Specific Audiometry Using Steady-State Responses. *Ear and Hearing* **1996**, *17* (2), 81-96.
- 143.Dimitrijevic, A.; John, M. S.; Van Roon, P.; Purcell, D. W.; Adamonis, J.; Ostroff, J.; Nedzelski, J. M.; Picton, T. W., Estimating the audiogram using multiple auditory steady-state responses. *Journal of the American Academy of Audiology* **2002**, *13* (4), 205-24.
- 144.Lazzouni, L.; Ross, B.; Voss, P.; Lepore, F., Neuromagnetic auditory steady-state responses to amplitude modulated sounds following dichotic or monaural presentation. *Clinical Neurophysiology* **2010**, *121* (2), 200-207.
- 145.Gander, P. E.; Bosnyak, D. J.; Roberts, L. E., Evidence for modality-specific but not frequency-specific modulation of human primary auditory cortex by attention. *Hearing Research* **2010**, *268* (1), 213-226.
- 146.Saupe, K.; Schröger, E.; Andersen, S. K.; Müller, M. M., Neural Mechanisms of Intermodal Sustained Selective Attention with Concurrently Presented Auditory and Visual Stimuli. *Frontiers in Human Neuroscience* **2009**, *3*, 58.
- 147.Keitel, C.; Schröger, E.; Saupe, K.; Müller, M. M., Sustained selective intermodal attention modulates processing of language-like stimuli. *Experimental Brain Research* **2011**, *213* (2), 321-327.

148. Ross, B.; Picton, T. W.; Herdman, A. T.; Pantev, C. The effect of attention on the auditory steady-state response *Neurology & clinical neurophysiology : NCN* [Online], 2004, p. 22. PubMed. <http://europepmc.org/abstract/MED/16012602> (accessed 2004).
149. Mahajan, Y.; Davis, C.; Kim, J., Attentional Modulation of Auditory Steady-State Responses. *PLoS ONE* **2014**, 9 (10), e110902.
150. Riecke, L.; Scharke, W.; Valente, G.; Gutschalk, A., Sustained Selective Attention to Competing Amplitude-Modulations in Human Auditory Cortex. *PLoS ONE* **2014**, 9 (9), e108045.
151. Ding, N.; Simon, J. Z., Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of neurophysiology* **2012**, 107 (1), 78-89.
152. Habib, M.; Besson, M., What do music training and musical experience teach us about brain plasticity? *Music Perception* **2009**, 26 (3), 279-285.
153. Zatorre, R., Music, the food of neuroscience? *Nature* **2005**, 434 (7031), 312-315.
154. Schmithorst, V. J.; Wilke, M., Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neurosci Lett* **2002**, 321 (1-2), 57-60.
155. Abdul-Kareem, I. A.; Stancak, A.; Parkes, L. M.; Al-Ameen, M.; Alghamdi, J.; Aldhafeeri, F. M.; Embleton, K.; Morris, D.; Sluming, V., Plasticity of the superior and middle cerebellar peduncles in musicians revealed by quantitative analysis of volume and number of streamlines based on diffusion tensor tractography. *Cerebellum (London, England)* **2011**, 10 (3), 611-23.
156. Bengtsson, S. L.; Nagy, Z.; Skare, S.; Forsman, L.; Forssberg, H.; Ullen, F., Extensive piano practicing has regionally specific effects on white matter development. *Nat Neurosci* **2005**, 8 (9), 1148-1150.
157. Gaser, C.; Schlaug, G., Brain Structures Differ between Musicians and Non-Musicians. *The Journal of Neuroscience* **2003**, 23 (27), 9240-9245.
158. Schneider, P.; Scherg, M.; Dosch, H. G.; Specht, H. J.; Gutschalk, A.; Rupp, A., Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci* **2002**, 5 (7), 688-694.
159. Schneider, P.; Sluming, V.; Roberts, N.; Scherg, M.; Goebel, R.; Specht, H. J.; Dosch, H. G.; Bleeck, S.; Stippich, C.; Rupp, A., Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat Neurosci* **2005**, 8 (9), 1241-7.
160. Bermudez, P.; Lerch, J. P.; Evans, A. C.; Zatorre, R. J., Neuroanatomical Correlates of Musicianship as Revealed by Cortical Thickness and Voxel-Based Morphometry. *Cerebral Cortex* **2008**, 19 (7), 1583-1596.
161. Peretz, I.; Zatorre, R. J., Brain Organization for Music Processing. *Annual Review of Psychology* **2005**, 56 (1), 89-114.

162. Warren, J., How does the brain process music? *Clinical medicine (London, England)* **2008**, 8 (1), 32-6.
163. Strait, D. L.; Chan, K.; Ashley, R.; Kraus, N., Specialization among the specialized: auditory brainstem function is tuned in to timbre. *Cortex* **2012**, 48 (3), 360-2.
164. Seppänen, M.; Brattico, E.; Tervaniemi, M., Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiology of Learning and Memory* **2007**, 87 (2), 236-247.
165. Nager, W.; Kohlmetz, C.; Altenmüller, E.; Rodriguez-Fornells, A.; Münte, T. F., The fate of sounds in conductors' brains: an ERP study. *Cognitive Brain Research* **2003**, 17 (1), 83-93.
166. Zatorre, R. J., Predispositions and Plasticity in Music and Speech Learning: Neural Correlates and Implications. *Science* **2013**, 342 (6158), 585-589.
167. Zuk, J.; Gaab, N., Evaluating predisposition and training in shaping the musician's brain: the need for a developmental perspective. *Ann N Y Acad Sci* **2018**.
168. Groussard, M.; Viader, F.; Landeau, B.; Desgranges, B.; Eustache, F.; Platel, H., The effects of musical practice on structural plasticity: the dynamics of grey matter changes. *Brain Cogn* **2014**, 90, 174-80.
169. Schlaug, G.; Jäncke, L.; Huang, Y.; Staiger, J. F.; Steinmetz, H., Increased corpus callosum size in musicians. *Neuropsychologia* **1995**, 33 (8), 1047-1055.
170. Steele, C. J.; Bailey, J. A.; Zatorre, R. J.; Penhune, V. B., Early Musical Training and White-Matter Plasticity in the Corpus Callosum: Evidence for a Sensitive Period. *The Journal of Neuroscience* **2013**, 33 (3), 1282-1290.
171. Skoe, E.; Kraus, N., A Little Goes a Long Way: How the Adult Brain Is Shaped by Musical Training in Childhood. *The Journal of Neuroscience* **2012**, 32 (34), 11507-11510.
172. Sato, K.; Kirino, E.; Tanaka, S., A Voxel-Based Morphometry Study of the Brain of University Students Majoring in Music and Nonmusic Disciplines. *Behavioural neurology* **2015**, 2015, 274919.
173. Olszewska, A. M.; Gaca, M.; Herman, A. M.; Jednoróg, K.; Marchewka, A., How Musical Training Shapes the Adult Brain: Predispositions and Neuroplasticity. *Frontiers in Neuroscience* **2021**, 15 (204).
174. Fujioka, T.; Trainor, L. J.; Ross, B.; Kakigi, R.; Pantev, C., Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of cognitive neuroscience* **2004**, 16.
175. Lee, K. M.; Skoe, E.; Kraus, N.; Ashley, R., Selective Subcortical Enhancement of Musical Intervals in Musicians. *The Journal of Neuroscience* **2009**, 29 (18), 5832-5840.
176. Ohnishi, T.; Matsuda, H.; Asada, T.; Aruga, M.; Hirakata, M.; Nishikawa, M.; Katoh, A.; Imabayashi, E., Functional anatomy of musical perception in musicians. *Cereb Cortex* **2001**, 11 (8), 754-60.

177. Amad, A.; Seidman, J.; Draper, S. B.; Bruchhage, M. M. K.; Lowry, R. G.; Wheeler, J.; Robertson, A.; Williams, S. C. R.; Smith, M. S., Motor Learning Induces Plasticity in the Resting Brain—Drumming Up a Connection. *Cerebral Cortex* **2016**, *27* (3), 2010-2021.
178. Li, Q.; Wang, X.; Wang, S.; Xie, Y.; Li, X.; Xie, Y.; Li, S., Musical training induces functional and structural auditory - motor network plasticity in young adults. *Human brain mapping* **2018**, *39* (5), 2098-2110.
179. Bianchi, F.; Hjortkjær, J.; Santurette, S.; Zatorre, R. J.; Siebner, H. R.; Dau, T., Subcortical and cortical correlates of pitch discrimination: Evidence for two levels of neuroplasticity in musicians. *NeuroImage* **2017**, *163*, 398-412.
180. Habermeyer, B.; Herdener, M.; Esposito, F.; Hilti, C. C.; Klarhöfer, M.; di Salle, F.; Wetzel, S.; Scheffler, K.; Cattapan-Ludewig, K.; Seifritz, E., Neural correlates of pre-attentive processing of pattern deviance in professional musicians. *Human brain mapping* **2009**, *30* (11), 3736-47.
181. Bangert, M.; Peschel, T.; Schlaug, G.; Rotte, M.; Drescher, D.; Hinrichs, H.; Heinze, H. J.; Altenmüller, E., Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* **2006**, *30* (3), 917-26.
182. Limb, C. J.; Kemeny, S.; Ortigoza, E. B.; Rouhani, S.; Braun, A. R., Left hemispheric lateralization of brain activity during passive rhythm perception in musicians. *The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology* **2006**, *288* (4), 382-9.
183. de Manzano, Ö.; Ullén, F., Same Genes, Different Brains: Neuroanatomical Differences Between Monozygotic Twins Discordant for Musical Training. *Cerebral Cortex* **2017**, *28* (1), 387-394.
184. Tavor, I.; Botvinik-Nezer, R.; Bernstein-Eliav, M.; Tsarfaty, G.; Assaf, Y., Short-term plasticity following motor sequence learning revealed by diffusion magnetic resonance imaging. *Human brain mapping* **2020**, *41* (2), 442-452.
185. Li, Q.; Wang, X.; Wang, S.; Xie, Y.; Li, X.; Xie, Y.; Li, S., Dynamic reconfiguration of the functional brain network after musical training in young adults. *Brain Structure and Function* **2019**, *224* (5), 1781-1795.
186. Chen, J. L.; Rae, C.; Watkins, K. E., Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *Neuroimage* **2012**, *59* (2), 1200-8.
187. Brown, R. M.; Penhune, V. B., Efficacy of Auditory versus Motor Learning for Skilled and Novice Performers. *J Cogn Neurosci* **2018**, *30* (11), 1657-1682.
188. Herholz, S. C.; Boh, B.; Pantev, C., Musical training modulates encoding of higher-order regularities in the auditory cortex. *European Journal of Neuroscience* **2011**, *34* (3), 524-529.
189. Lappe, C.; Herholz, S. C.; Trainor, L. J.; Pantev, C., Cortical plasticity induced by short-term unimodal and multimodal musical training. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **2008**, *28* (39), 9632-9639.

190. Pantev, C.; Lappe, C.; Herholz, S. C.; Trainor, L., Auditory-somatosensory integration and cortical plasticity in musical training. *Ann N Y Acad Sci* **2009**, *1169*, 143-50.
191. Forgeard, M.; Winner, E.; Norton, A.; Schlaug, G., Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One* **2008**, *3* (10), e3566.
192. Kraus, N.; Chandrasekaran, B., Music training for the development of auditory skills. *Nat Rev Neurosci* **2010**, *11* (8), 599-605.
193. Trainor, L. J.; Shahin, A. J.; Roberts, L. E., Understanding the benefits of musical training: effects on oscillatory brain activity. *Ann N Y Acad Sci* **2009**, *1169*, 133-42.
194. Bialystok, E.; Depape, A. M., Musical expertise, bilingualism, and executive functioning. *Journal of experimental psychology. Human perception and performance* **2009**, *35* (2), 565-74.
195. Moreno, S.; Bialystok, E.; Barac, R.; Schellenberg, E. G.; Cepeda, N. J.; Chau, T., Short-term music training enhances verbal intelligence and executive function. *Psychol Sci* **2011**, *22* (11), 1425-1433.
196. Roden, I.; Grube, D.; Bongard, S.; Kreutz, G., Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music* **2014**, *42* (2), 284-298.
197. Strait, D. L.; Kraus, N.; Skoe, E.; Ashley, R., Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. *The European journal of neuroscience* **2009**, *29* (3), 661-8.
198. Lima, C. F.; Castro, S. L., Speaking to the trained ear: musical expertise enhances the recognition of emotions in speech prosody. *Emotion* **2011**, *11* (5), 1021-31.
199. Patel, A. D.; Iversen, J. R., The linguistic benefits of musical abilities. *Trends in cognitive sciences* **2007**, *11* (9), 369-372.
200. Musacchia, G.; Sams, M.; Skoe, E.; Kraus, N., Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences* **2007**, *104* (40), 15894-15898.
201. Schlaug, G.; Norton, A.; Overy, K.; Cronin, K.; Lee, D. J.; Winner, E. In *Effects of music training on children's brain and cognitive development*, Proceedings of the 8th International Conference on Music Perception & Cognition, 2005; pp 133-134.
202. Miendlarzewska, E.; Trost, W., How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in Neuroscience* **2014**, *7* (279).
203. Parbery-Clark, A.; Skoe, E.; Kraus, N., Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *The Journal of Neuroscience* **2009a**, *29* (45), 14100-14107.
204. Chandrasekaran, B.; Kraus, N., Music, Noise-Exclusion, and Learning. *Music Perception* **2010**, *27* (4), 297-306.

- 205.Parbery-Clark, A.; Skoe, E.; Lam, C.; Kraus, N., Musician enhancement for speech-in-noise. *Ear and hearing* **2009**, *30* (6), 653-661.
- 206.Nikjeh, D. A.; Lister, J. J.; Frisch, S. A., Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: influence of music training. *Ear Hear* **2009**, *30* (4), 432-46.
- 207.Moreno, S.; Marques, C.; Santos, A.; Santos, M.; Castro, S. L.; Besson, M., Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex* **2009**, *19* (3), 712-723.
- 208.Magne, C.; Schön, D.; Besson, M., Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J Cogn Neurosci* **2006**, *18* (2), 199-211.
- 209.Habibi, A.; Cahn, B. R.; Damasio, A.; Damasio, H., Neural correlates of accelerated auditory processing in children engaged in music training. *Developmental Cognitive Neuroscience* **2016**, *21*, 1-14.
- 210.Zendel, B. R.; Alain, C., Musicians experience less age-related decline in central auditory processing. *Psychol Aging* **2012**, *27* (2), 410-7.
- 211.Parbery-Clark, A.; Anderson, S.; Hittner, E.; Kraus, N., Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging* **2012**, *33* (7), 1483.e1-1483.e4.
- 212.Patel, A. D.; Gibson, E.; Ratner, J.; Besson, M.; Holcomb, P. J., Processing syntactic relations in language and music: an event-related potential study. *J Cogn Neurosci* **1998**, *10* (6), 717-33.
- 213.Patel, A. D., Language, music, syntax and the brain. *Nature Neuroscience* **2003**, *6* (7), 674-681.
- 214.Zatorre, R. J.; Belin, P.; Penhune, V. B., Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences* **2002**, *6* (1), 37-46.
- 215.Schön, D.; Gordon, R.; Campagne, A.; Magne, C.; Astésano, C.; Anton, J. L.; Besson, M., Similar cerebral networks in language, music and song perception. *Neuroimage* **2010**, *51* (1), 450-61.
- 216.Maess, B.; Koelsch, S.; Gunter, T. C.; Friederici, A. D., Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience* **2001**, *4* (5), 540-545.
- 217.Jentschke, S.; Koelsch, S., Musical training modulates the development of syntax processing in children. *NeuroImage* **2009**, *47* (2), 735-744.
- 218.Moreno, S.; Marques, C.; Santos, A.; Santos, M.; Castro, S. L.; Besson, M., Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb Cortex* **2009**, *19* (3), 712-23.
- 219.Besson, M.; Chobert, J.; Marie, C., Transfer of Training between Music and Speech: Common Processing, Attention, and Memory. *Frontiers in psychology* **2011**, *2*, 94-94.

- 220.Chan, A. S.; Ho, Y.-C.; Cheung, M.-C., Music training improves verbal memory. *Nature* **1998**, 396 (6707), 128-128.
- 221.Jakobson, L. S.; Cuddy, L. L.; Kilgour, A. R., Time tagging: A key to musicians' superior memory. *Music Perception* **2003**, 20 (3), 307-313.
- 222.Milovanov, R.; Huotilainen, M.; Välimäki, V.; Esquef, P. A. A.; Tervaniemi, M., Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research* **2008**, 1194, 81-89.
- 223.Coffey, E. B. J.; Chepesiuk, A. M. P.; Herholz, S. C.; Baillet, S.; Zatorre, R. J., Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. *Frontiers in Neuroscience* **2017**, 11 (479).
- 224.Schön, D.; Magne, C.; Besson, M., The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* **2004**, 41 (3), 341-9.
- 225.Chobert, J.; Marie, C.; François, C.; Schön, D.; Besson, M., Enhanced passive and active processing of syllables in musician children. *J Cogn Neurosci* **2011**, 23 (12), 3874-87.
- 226.Francois, C.; Schön, D., Musical expertise boosts implicit learning of both musical and linguistic structures. *Cereb Cortex* **2011**, 21 (10), 2357-65.
- 227.Chobert, J.; Besson, M., Musical expertise and second language learning. *Brain Sci* **2013**, 3 (2), 923-940.
- 228.Marie, C.; Delogu, F.; Lampis, G.; Belardinelli, M. O.; Besson, M., Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *J Cogn Neurosci* **2011**, 23 (10), 2701-15.
- 229.Dittinger, E.; Chobert, J.; Ziegler, J. C.; Besson, M., Fast Brain Plasticity during Word Learning in Musically-Trained Children. *Frontiers in Human Neuroscience* **2017**, 11 (233).
- 230.Martínez-Montes, E.; Hernández-Pérez, H.; Chobert, J.; Morgado-Rodríguez, L.; Suárez-Murias, C.; Valdés-Sosa, P.; Besson, M., Musical expertise and foreign speech perception. *Frontiers in Systems Neuroscience* **2013**, 7 (84).
- 231.Wong, P. C. M.; Skoe, E.; Russo, N. M.; Dees, T.; Kraus, N., Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* **2007**, 10 (4), 420-422.
- 232.Bidelman, G. M.; Weiss, M. W.; Moreno, S.; Alain, C., Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience* **2014**, 40 (4), 2662-2673.
- 233.Besson, M.; Schön, D.; Moreno, S.; Santos, A.; Magne, C., Influence of musical expertise and musical training on pitch processing in music and language. *Restorative neurology and neuroscience* **2007**, 25 (3-4), 399-410.

- 234.Winkler, I.; Denham, S. L.; Nelken, I., Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences* **2009**, *13* (12), 532-540.
- 235.Hornickel, J.; Skoe, E.; Nicol, T.; Zecker, S.; Kraus, N., Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci U S A* **2009**, *106* (31), 13022-7.
- 236.Chandrasekaran, B.; Hornickel, J.; Skoe, E.; Nicol, T.; Kraus, N., Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron* **2009**, *64* (3), 311-319.
- 237.Parbery-Clark, A.; Skoe, E.; Lam, C.; Kraus, N., Musician enhancement for speech-in-noise. *Ear and hearing* **2009b**, *30* (6), 653.
- 238.Müllensiefen, D.; Gingras, B.; Musil, J.; Stewart, L., Measuring the facets of musicality: The Goldsmiths Musical Sophistication Index (Gold-MSI). *Personality and Individual Differences* **2014**, *60* (Supplement), S35.
- 239.Müllensiefen, D.; Gingras, B.; Musil, J.; Stewart, L., The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE* **2014**, *9* (2), e89642.
- 240.Lamminmäki, S.; Parkkonen, L.; Hari, R., Human Neuromagnetic Steady-State Responses to Amplitude-Modulated Tones, Speech, and Music. *Ear and Hearing* **2014**, *35* (4), 461-467.
- 241.Robinson, D. W.; Dadson, R. S., A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics* **1956**, *7* (5), 166.
- 242.Oostenveld, R.; Fries, P.; Maris, E.; Schoffelen, J.-M., FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience* **2011**, *2011*, 9.
- 243.Gramfort, A.; Luessi, M.; Larson, E.; Engemann, D.; Strohmeier, D.; Brodbeck, C.; Goj, R.; Jas, M.; Brooks, T.; Parkkonen, L.; Hämäläinen, M., MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience* **2013**, *7* (267).
- 244.Fischl, B., FreeSurfer. *Neuroimage* **2012**, *62* (2), 774-81.
- 245.Fan, L.; Chu, C.; Li, H.; Chen, L.; Xie, S.; Zhang, Y.; Yang, Z.; Jiang, T.; Laird, A. R.; Wang, J.; Zhuo, J.; Yu, C.; Fox, P. T.; Eickhoff, S. B., The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex* **2016**, *26* (8), 3508-3526.
- 246.Van Essen, D. C., A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *NeuroImage* **2005**, *28* (3), 635-662.
- 247.Reyes, S. A.; Lockwood, A. H.; Salvi, R. J.; Coad, M. L.; Wack, D. S.; Burkard, R. F., Mapping the 40-Hz auditory steady-state response using current density reconstructions. *Hearing Research* **2005**, *204* (1), 1-15.



- 248.Reyes, S. A.; Salvi, R. J.; Burkard, R. F.; Coad, M. L.; Wack, D. S.; Galantowicz, P. J.; Lockwood, A. H., PET imaging of the 40 Hz auditory steady state response. *Hearing Research* **2004**, *194* (1), 73-80.
- 249.Farahani, E. D.; Goossens, T.; Wouters, J.; van Wieringen, A., Spatiotemporal reconstruction of auditory steady-state responses to acoustic amplitude modulations: Potential sources beyond the auditory pathway. *NeuroImage* **2017**, *148*, 240-253.
- 250.Schoonhoven, R.; Boden, C. J.; Verbunt, J. P.; de Munck, J. C., A whole head MEG study of the amplitude-modulation-following response: phase coherence, group delay and dipole source analysis. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **2003**, *114* (11), 2096-106.
- 251.Steinmann, I.; Gutschalk, A., Potential fMRI correlates of 40-Hz phase locking in primary auditory cortex, thalamus and midbrain. *NeuroImage* **2011**, *54* (1), 495-504.
- 252.Cohen, R. A., Attention and the Frontal Cortex. In *The Neuropsychology of Attention*, Springer US: Boston, MA, 2014; pp 335-379.
- 253.Foster, J. K.; Eskes, G. A.; Stuss, D. T., The cognitive neuropsychology of attention: A frontal lobe perspective. *Cognitive Neuropsychology* **1994**, *11* (2), 133-147.
- 254.Plakke, B.; Romanski, L. M., Auditory connections and functions of prefrontal cortex. *Frontiers in neuroscience* **2014**, *8*, 199-199.
- 255.Keitel, C.; Maess, B.; Schröger, E.; Müller, M. M., Steady-state responses track attentional shifts towards audition and vision. *Human Neuroscience Archive* **2011**.
- 256.Pugh, K. R.; offywitz, B. A.; Shaywitz, S. E.; Fulbright, R. K.; Byrd, D.; Skudlarski, P.; Shankweiler, D. P.; Katz, L.; Constable, R. T.; Fletcher, J.; Lacadie, C.; Marchione, K.; Gore, J. C., Auditory selective attention: an fMRI investigation. *Neuroimage* **1996**, *4* (3 Pt 1), 159-73.
- 257.Tzourio, N.; El Massioui, F.; Crivello, F.; Joliot, M.; Renault, B.; Mazoyer, B., Functional Anatomy of Human Auditory Attention Studied with PET. *NeuroImage* **1997**, *5* (1), 63-77.
- 258.Shomstein, S.; Yantis, S., Parietal Cortex Mediates Voluntary Control of Spatial and Nonspatial Auditory Attention. *The Journal of Neuroscience* **2006**, *26* (2), 435-439.
- 259.Wu, C. T.; Weissman, D. H.; Roberts, K. C.; Woldorff, M. G., The neural circuitry underlying the executive control of auditory spatial attention. *Brain Res* **2007**, *1134* (1), 187-98.
- 260.Besle, J.; Schevon, C. A.; Mehta, A. D.; Lakatos, P.; Goodman, R. R.; McKhann, G. M.; Emerson, R. G.; Schroeder, C. E., Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *The Journal of Neuroscience* **2011**, *31* (9), 3176-3185.
- 261.Alain, C.; Arnott, S. R.; Hevenor, S.; Graham, S.; Grady, C. L., "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences* **2001**, *98* (21), 12301-12306.

262. Fujioka, T.; Ross, B.; Trainor, L. J., Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. *The Journal of Neuroscience* **2015**, *35* (45), 15187-15198.
263. Sameiro-Barbosa, C. M.; Geiser, E., Sensory Entrainment Mechanisms in Auditory Perception: Neural Synchronization Cortico-Striatal Activation. *Frontiers in Neuroscience* **2016**, *10* (361).
264. Huang, S.; Belliveau, J. W.; Tengshe, C.; Ahveninen, J., Brain Networks of Novelty-Driven Involuntary and Cued Voluntary Auditory Attention Shifting. *PLOS ONE* **2012**, *7* (8), e44062.
265. Salmi, J.; Rinne, T.; Koistinen, S.; Salonen, O.; Alho, K., Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Research* **2009**, *1286*, 155-164.
266. Huang, N.; Elhilali, M., Push-pull competition between bottom-up and top-down auditory attention to natural soundscapes. *eLife* **2020**, *9*, e52984.
267. Parbery-Clark, A.; Anderson, S.; Hittner, E.; Kraus, N., Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Front Aging Neurosci* **2012**, *4*, 30-30.
268. Rosenkranz, K.; Williamon, A.; Rothwell, J. C., Motorcortical excitability and synaptic plasticity is enhanced in professional musicians. *J Neurosci* **2007**, *27* (19), 5200-6.
269. Gärtner, H.; Minnerop, M.; Pieperhoff, P.; Schleicher, A.; Zilles, K.; Altenmüller, E.; Amunts, K., Brain morphometry shows effects of long-term musical practice in middle-aged keyboard players. *Frontiers in Psychology* **2013**, *4* (636).
270. de Aquino, M. P. B.; Verdejo-Román, J.; Pérez-García, M.; Pérez-García, P., Different role of the supplementary motor area and the insula between musicians and non-musicians in a controlled musical creativity task. *Scientific reports* **2019**, *9* (1), 13006.
271. Ross, B., *Steady-state auditory evoked responses*. 2013; Vol. 10, p 155-176.
272. Kuriki, S.; Kobayashi, Y.; Kobayashi, T.; Tanaka, K.; Uchikawa, Y., Steady-state MEG responses elicited by a sequence of amplitude-modulated short tones of different carrier frequencies. *Hearing Research* **2013**, *296*, 25-35.
273. Patel, A. D.; Balaban, E., Temporal patterns of human cortical activity reflect tone sequence structure. *Nature* **2000**, *404* (6773), 80-84.
274. Patel, A. D.; Balaban, E., Human auditory cortical dynamics during perception of long acoustic sequences: phase tracking of carrier frequency by the auditory steady-state response. *Cereb Cortex* **2004**, *14* (1), 35-46.
275. Bigand, E.; McAdams, S.; Forêt, S., Divided attention in music. *International Journal of Psychology* **2000**, *35* (6), 270-278.

276. Demany, L.; Erviti, M.; Semal, C., Auditory attention is divisible: segregated tone streams can be tracked simultaneously. *Journal of experimental psychology. Human perception and performance* **2015**, *41* (2), 356-63.
277. Kraus, N.; Skoe, E.; Parbery-Clark, A.; Ashley, R., Experience-induced Malleability in Neural Encoding of Pitch, Timbre, and Timing: Implications for Language and Music. *Annals of the New York Academy of Sciences* **2009**, *1169*, 543-557.
278. Korczak, P.; Smart, J.; Delgado, R.; Strobel, T.; Bradford, C., Auditory Steady-State Responses. *Journal of the American Academy of Audiology* **2012**, *23*, 146-70.