From Behavior to Brain

Citation for published version (APA):

Heynckes, M. (2022). *From Behavior to Brain: The contribution of rhythm to auditory perception*. [Doctoral Thesis, Maastricht University]. Maastricht University. https://doi.org/10.26481/dis.20220623mh

Document status and date:
Published: 01/01/2022

DOI:
10.26481/dis.20220623mh

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 25 Apr. 2024

From Behavior to Brain

The contribution of rhythm to auditory perception

Dissertation
To obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. Dr. Pamela Habibovic,
in accordance with the decision of the Board of Deans,
to be defended in public on
Thursday 23rd of June 2022, at 16:00 hours
by
Miriam Heynckes

Promotors

Prof. Dr. Elia Formisano

Prof. Dr. Peter De Weerd

Dr. Federico De Martino

Assessment Committee

Prof. Dr. Sonja Kotz (Chair)

Prof. Dr. Essa Yacoub
University of Minnesota

Prof. Dr. Christopher I. Petkov
Newcastle University

Dr. Isma Zulfiqar

The research presented in this thesis was supported by the Netherlands Organization for Scientific Research (NWO) Research Talent Grant 406.17.200 awarded to Miriam Heynckes, Elia Formisano and Peter De Weerd.

Contents

1

General Introduction

Understanding how the brain works is fascinating to scientists and non-scientists alike, as brain functioning is responsible for how we perceive the world and act upon it. One of the fundamental questions in (auditory) cognitive neuroscience asks how the brain makes sense of the acoustic information around us and how auditory perception and behavior relate to activity in the brain.

In recent years a paradigm-shift has occurred in our understanding of the mechanisms of auditory perception. Inspired by theories advanced over 150 years ago by Helmholtz (Helmholtz, 1896), generative models suggest that what we perceive does not reflect the sensory stimulus itself, but rather a combination of the stimulus and an internal (generative) model of its causes that predicts the external stimulus (Friston, 2010; Schroeder, Wilson, Radman, Scharfman & Lakatos, 2010). In audition, the temporal pattern of acoustic inputs is often highly predictable (think of the rhythm of sounds for instance), and the brain relies on this information by matching its neural response to presented rhythmic stimuli (Lakatos, Karmos, Mehta, Ulbert & Schroeder, 2008). This allows listeners to predict upcoming acoustic input to prepare for and focus on time points when useful information is likely to occur (Nobre & Van Ede, 2018). Such temporal expectations are often studied in the context of acoustic rhythms. The empirical questions of this thesis center around the contribution of acoustic rhythms to perception and where in the brain such perceptual processing occurs.

In search of a mechanism of auditory perception

Neural mechanisms

A general framework through which scientists try to explain the neural basis of perception is by investigating and characterizing *neural mechanisms*. A neural mechanism can be conceptualized as a *structure*, which is performing a certain *function*. Both structure and function of a mechanism can be further broken down into parts of the structure, each performing an organized functional set of operations (Bechtel & Abrahmsen, 2005). Cognitive neuroscientists aim to localize operations by assigning them to specific parts (Bechtel, 2008). They can achieve this using a variety of methods that measure the activity of neuronal ensembles in response to experimental manipulations (that aim to isolate operations). Some methods allow observing specifically in which regions of the brain this activity occurs. By doing so, researchers can investigate specific mechanisms in their structure - function relation. Box 1 highlights the neuroscientific methods employed in this thesis and their relation to other common methods.

The orchestrated functioning of the mechanism (and its parts) is responsible for one or more (behavioral) phenomena (Bechtel, 2006; Bogen & Woodward, 1988). Therefore, neural mechanisms are defined in terms of the phenomena for which they are responsible (Glennan, 1996; Kauffman, 1971). This point is crucial as it highlights the need to link (brain) parts and operations to behavior. Much of modern neuroscientific research is necessarily exploratory, studies focus on describing parts and operations by using a multitude of techniques and only consider behavior post hoc (Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017). These studies thereby run the risk of remaining at an observational level of describing parts and operations, instead of explaining neural mechanisms by explicitly linking them to behavioral phenomena. David Marr referred to the link between computations (behavior) and implementation (parts and operations) as the algorithm. Considering all three components (computation, algorithm and implementation) allows understanding a neural mechanism (Marr, 1982).

In the present thesis, acoustic rhythm perception is investigated behaviorally using psychophysics. The results of this research are then used to motivate hypotheses for neuroimaging investigations in which we attempt to explain behavior through measures of brain activity.

Box 1 - Neuroscientific methods

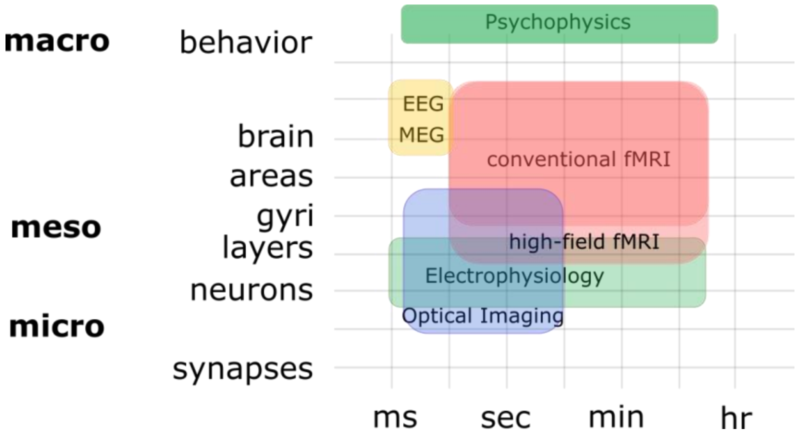


Figure 1 Neuroimaging methods can be organized based on the spatial scale they probe from micro to macroscopic, and with respect to the time at which these methods can capture observations of the brain from milliseconds to hours (adapted from Vizioli et al. 2021).

Psychophysics

Psychophysics aims at explaining perception by relating human sensation and perception to underlying neuronal mechanisms, via mathematical models or linking hypotheses (Morgan et al., 2013). Operations within a mechanism are quantified without directly probing neural responses but by observing behavior (Fechner, 1860). Typically, this involves the use of controlled stimuli, that recapitulate hypothetical operations and thus allow to infer the processes involved in their perception, thereby creating a vocabulary that describes the inner workings of the mind (Read, 2015). These experiments typically measure reaction times, accuracy scores and other so-called behavioral measures, through which inferences on the underlying mechanisms can be drawn. Psychoacoustics as a discipline concerns itself with human sensation and perception related to audition.

Functional magnetic resonance imaging

Functional magnetic resonance imaging (fMRI) is an indirect and noninvasive method of measuring brain activity in humans and animals. The most common fMRI signal exploits the differences in magnetic susceptibility of oxygenated and deoxygenated blood and has therefore been termed the blood oxygenation level dependent (BOLD) signal (Ogawa, 1990). Because the metabolic demand of neural activity results in blood being delivered to a specific brain area and in a relative increase of oxygenated blood (compared to de-oxygenated blood), a localized change in neural activity leads to a localized vascular response that can be measured with fMRI.

The spatial and temporal specificity of fMRI depend on how the vascular response relates to the triggering neural activity in space and time. The spatial specificity of the BOLD signal depends on several parameters (e.g., the magnetic field strength at which data are acquired or the specific acquisition strategy that is used) and is thought to be in the order of 1 mm for Gradient Echo (GE) BOLD (Chaimow, Yacoub, Ugurbil, Shmuel, 2018). Compared to conventional field strengths (3 Tesla and below), high magnetic fields (7 Tesla and above) have significant advantages, as the BOLD contrast sensitivity increases, resulting in higher signal to noise ratio (Vaughan et al., 2001), which can be traded to increase resolution (< 1 mm) allowing to partition the cortical ribbon in at least three compartments (the thickness of e.g. the auditory cortex is approximatively 3 mm - Zoellner et al., 2018). This allows, in combination with advanced data analysis strategies for high-field fMRI (Polimeni et al., 2017; De Martino et al., 2018), mapping brain activation at the mesoscopic scale (i.e., the one of cortical layers and columns). When considering depth dependent fMRI responses acquired with conventional GE-BOLD, the spatial specificity of the acquired signal varies between 0.8 - 1.8 mm, increasing from deep to superficial depth levels within the cortex (Fracasso et al., 2021).

From the temporal point of view, the vascular response exhibits a typical delay between the neural activity and the peak of the BOLD response of about 6 seconds. This delayed response creates a challenge that needs to be incorporated into the design of experiments. Yet auditory fMRI studies can exploit the “sluggishness’’ of the BOLD response to their advantage to tackle another problem: the loudness of the MRI scanner (in the regime of 100 decibels (Di Salle, 2003). The noise mostly comes from the fast switching of the gradients inside the scanner that are used to generate the spatial encoding of the signal. Because of the delayed peak of the BOLD response though, auditory neuroscientists can choose to present a sound in “silence” and, only after the sound has played, turn on the machine to measure the brain response to the sound (Moerel, Yacoub, Gulban, Lage-Castellanos & De Martino, 2021).

Avoiding a reductionist bias

As different neuroimaging approaches allow observing neural mechanisms at increasing spatial and temporal scales (see box 1), a relevant question emerges. Which scale is the appropriate one to identify neural mechanisms? Observing increasingly smaller parts and operations may not necessarily allow us to better understand brain function. The risk of incurring in a reductionist bias has been eloquently formulated by Marr: ‘‘trying to understand perception by understanding neurons is like trying to understand a bird’s flight by studying only its feathers. It just cannot be done’’ (Marr, 1982). What is required is to map the concept *flight* to the parts (*feathers*) and their operations (*the flapping of wings*) (Krakauer et al., 2017). Thus, keeping in mind the behavior of the organism is imperative to avoid a reductionist bias (Read, 2015; Krakauer et al. 2017). Yet, it is also important to identify a proper level at which we choose to inspect a mechanism. Will it be the feathers or even smaller parts? For brain mechanisms, a proper level is not defined by the quantity of neurons we can record (or if we can record smaller or larger ensembles of them), but by the quality of the mapping between the neural mechanism (parts and operations) and behavior (Fetsch, 2016; Krakauer et al., 2017).

The approach brought forward in this thesis links human auditory behavior to neural data at the mesoscale of cortical layers. We believe this scale to be appropriate to achieve a proper mapping between the neural signals we measure and the behavior we wish to explain. In what follows we elaborate why the mesoscale of cortical layers (and columns) may be the proper level at which to achieve this mapping.

The following sections provide the reader with an overview of these three components: (1) psychoacoustic factors of behavior, (2) structural brain parts and (3) functional operations relevant to the perception of acoustic rhythms.

Psychoacoustic factors

The neural basis of behavioral and perceptual phenomena cannot be properly characterized without first allowing for the independent detailed study of the behavior itself (Krakauer et al, 2015). With psychoacoustic experiments we can probe the behavior of humans change in response to manipulation of (low-level) acoustic features. This enables building of hypotheses on how the human auditory system processes these features (Moore, 2003). It is worth noting that often multiple experiments are required to carefully delineate the perceptual phenomenon for which a mechanism is responsible. One reason for this is that it is not always clear which variables influence the mechanism and thus are crucial for the occurrence of the phenomenon. Thus, building a model of a neural mechanism requires to treat some variables as sources of error variance that must be controlled (Waskom et al., 2019). Mechanistic explanation is thus achieved through a (sometimes iterative) process of (re) defining the phenomenon and probing parts of the structure or function and thereby the mechanism.

Before diving into the psychoacoustic factors that influence rhythm perception, it helps to understand what the most relevant physical features of sounds are. Sounds are air pressure waves. The simplest possible waves (pure tones) take a sinusoidal shape and induce in the listener a distinct pitch percept depending on the frequency of the sinusoidal wave, also termed carrier frequency. The acoustics of more complex sounds (e.g. speech) can be represented as the combination of constituent sinusoidal waves at various frequencies. Already at the periphery (i.e. the cochlea) the auditory system represents sounds by topographically encoding their frequency content. When a sound enters the ear, the frequencies composing that sound will be processed at distinct locations on the basilar membrane due to its biomechanical structure. This spatial representation of sound is referred to as a tonotopic map (Goldstein, 2009). Psychoacoustic studies have shown, for example, that the sensitivity to detect a difference between two tones is smallest at low frequencies and increases monotonically with increasing frequency (Sek and Moore, 1995).

In addition to spectrally encoding the content of a sound, the auditory system is sensitive to its temporal information. Natural sounds are characterized by temporal structure, which may be rhythmic as in the sounds produced by a metronome (creating the simplest possible rhythm, and termed isochronous), or (quasi-) rhythmic and more complex as in speech or music. The brain relies on such rhythmicity to predict upcoming acoustic input and to prepare for relevant events (Nobre & Van Ede, 2018). Rhythmic regularity of a stimulus decreases reaction times and improves accuracies of responses to supra-threshold stimuli when targets occur at an anticipated moment in time, compared to aperiodic unpredictable stimuli (Ellis et al.,2010; Rohenkohl et al., 2012), or targets occurring out of phase with a preceding rhythm (Barnes & Jones, 2000; Jones et al., 2002; Large & Jones, 1999) Rhythmic regularity also improved detection of sub-threshold stimuli (ten Oever et al., 2017).

The processing of the temporal structure and frequency content of a sound are not independent processes in the brain, as previous psychoacoustic studies on amplitude modulation detection have shown that rhythm and carrier frequency interact (Dau, Kohlmeier & Kohlrasuch 1997; Moore & Glasberg, 2001; Moore & Sek 1995; Simpson & McAlpine, 2015). In **Chapter 2** we conduct a psychoacoustic study, in which we examine these low-level sound features and their effect on the behavior of participants in a temporal detection task. In particular, we focus on the effects of specific temporal rates (i.e., changes in rhythm), the carrier frequency of the sounds and their interaction.

Temporal expectations can be formed based on different types of temporal structure of sounds. For example, temporal associations, such as the ones formed when learning the regularity between a cue and a target sound (like thunder following lightning) similarly enable temporal predictions. Therefore, understanding to what extent rhythm allows forming temporal predictions requires comparing the effects of rhythm to the effects of other forms of temporal structure. Early studies on temporal associations, also termed *foreperiods,* searched for optimal time intervals for behavioral preparation between the cue and target stimulus. They focused on supra-second foreperiods and revealed that the shortest reaction times to detect a target occurred between 2 and 4 seconds, suggesting this interval to be optimal for full attention (Woodrow,1914). Subsequent psychoacoustic studies that manipulated the variability of the temporal association interval have shown that response times become faster when the interval is more predictable (i.e., less variable over a block of trials; Bertelson & Boons, 1960) and temporal prediction efficiency is optimal at shorter time intervals and diminishes with longer intervals between warning signal and cue, as observed by increased reaction times (Näätänen & Niemi, 1981) and lower accuracy (Rolke & Hoffman, 2007). A large body of literature has since then been accumulated on the effects of foreperiods (Los et al., 2001). Whether temporal cueing and rhythmic processing are sub-served by the same neural (prediction) mechanism and thereby affect behavior in a similar way has remained an open question in the field (Nobre et al., 2007; Rimmele et al., 2018; Breska & Ivry, 2018; Bouwer et al., 2020). In **Chapter 3** we report the results of a psychoacoustic study investigating both types of temporal structure, periodic rhythms and temporal associations, by comparing them to aperiodic rhythms and assessing detection sensitivity and reaction times. Moreover, we manipulated the awareness of predictability of targets embedded within rhythms, by instructing participants and investigated its effect on our behavioral measures.

Structure of the auditory cortex

To identify parts of the brain relevant to auditory perception Richard L. Heschl dissected more than a thousand brains, identifying and meticulously describing the structure of a particular brain area (Heschl, 1878) - a gyrus (a convolution on the surface of the brain) that nowadays, due to his efforts, is known as Heschl's gyrus. At the time of his investigations, Heschl had no opportunity to assess the function of this gyrus, so he focused on describing its occurrence and macro-anatomical variability. To most, Heschl's contribution might have seemed quite futile at the time, yet quite remarkable from today's perspective, Heschl's gyrus is in fact the location of the primary auditory cortex (PAC), the first cortical area where acoustic information arriving from the ears is processed (Schreiner & Winer, 2011). The localization of PAC on Heschl's gyrus has been demonstrated by both cytoarchitectonic (Morosan et al., 2001; Rademacher et al., 2001) and functional neuroimaging studies (Formisano et al., 2003; Da Costa et al.,2011; Moerel et al., 2014; Saenz & Langers, 2014; De Martino et al., 2015). The asynchronous investigation of structure and function of Heschl's gyrus emphasizes the relevance of studies on parts and operations across spatial scales of description of a neural mechanism. The detailed macro-anatomical description of a part of the brain has been of invaluable use to scientists that later focused on its function.[1]

About two decades after Heschl's death in 1881, methods to stain neurons became available. The neuron doctrine arising from these methodological advances had identified the neuron as the relevant structural and functional unit of the brain (Golgi, 1906). Large parts of the cerebral cortex in mammals consist of six layers of cells (Figure 2A). These layers differ in thickness and cell type across distinct parts of the brain. The layered organization of cell bodies has been used to map the brain into different areas, the most famous example being Brodmann's cytoarchitectonic maps (Brodmann, 1910). Figure 2B shows a subdivision of the superior temporal gyrus (STG), into primary area 41, which is the approximate location of Heschl's gyrus and secondary auditory areas, labelled 42 and 22. Even today, proposed localizations of cognitive operations still refer to Brodmann's areas (Bechtel, 2008), while in later anatomical studies it also became clear that the cytoarchitectonic parcellation is more complex than the division suggested by Brodmann (Zilles & Amunts, 2010).

---

[1] We translated the original manuscript from German to English so that a larger audience can appreciate the detail of his meticulous work, see appendix.

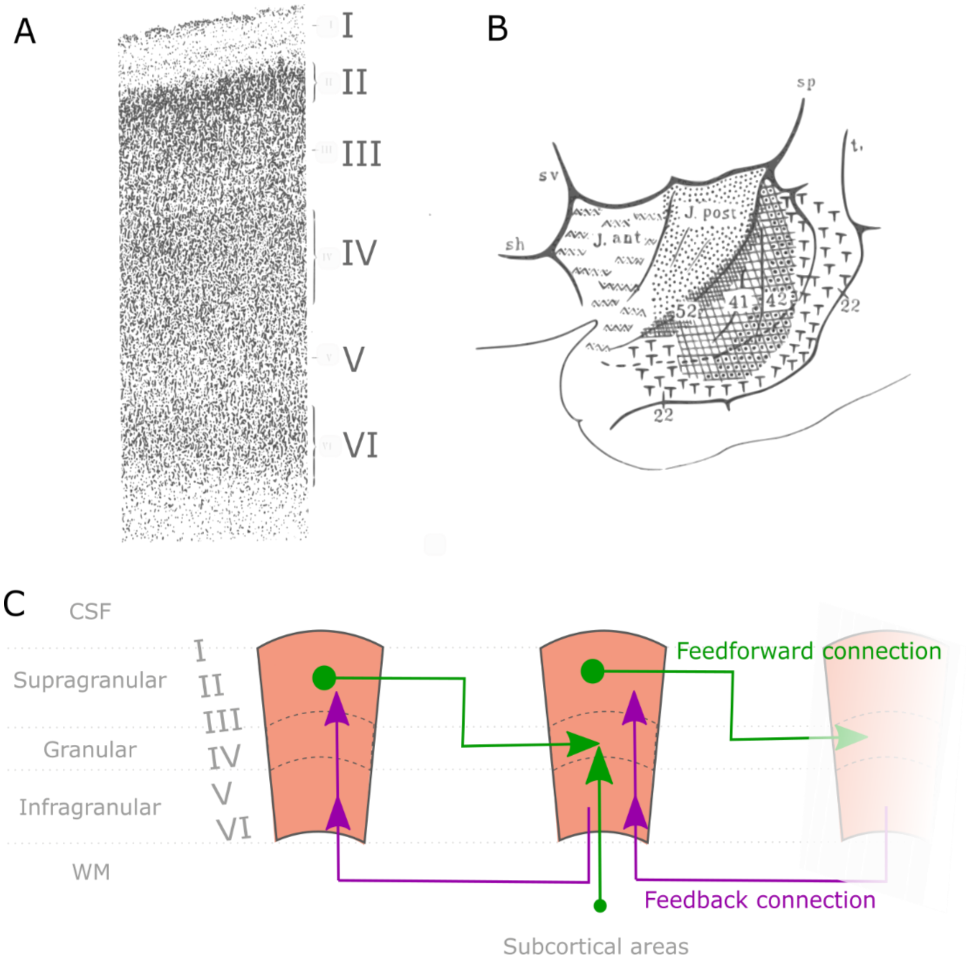


Figure 2 A Cyto-architectural six-layered structure of neocortex (Brodmann, 1910). **B** Cytoarchitectonic map by Brodmann (1910)**.** Auditory areas 41, 42 and 22 are shown. **C** Schematic summary of feedforward and feedback connections within the cortical hierarchy, drawn after Dumoulin (2017), based on Felleman & Van Essen (1991).

Functional operations of the auditory system

The information of sound frequency mapping seen in the peripheral auditory system (the cochlea) is preserved along the hierarchy of areas in the auditory pathway. Electrophysiological studies on the microscopic level have shown that neighboring neurons cluster based on their frequency preference, forming tonotopic maps (Merzenich & Brugge, 1973; Merzenich, Knight, & Roth, 1973). Following the seminal work relying on the use of invasive methodologies in animals, (f)MRI has allowed investigating the mechanisms underlying sound perception in humans non-invasively. With voxel sizes between 1 and 3 mm^3, conventional fMRI operates on a spatial scale larger than that of Brodmann's cortical layers, which combined span a cortical thickness of ~ 3 mm in auditory areas (Zoellner et al., 2018). Nevertheless, even using these relatively large voxel sizes, fMRI has allowed to non-invasively identify tonotopic maps on Heschl's gyrus in humans (Formisano, 2003; Da Costa et al., 2011, Saenz & Langers, 2014; De Martino et al., 2015). Moreover, with fMRI it has been possible to map the tonotopic organization outside Heschl's gyrus (Moerel et al., 2012). The tonotopic organization revealed by fMRI allows to non-invasively localize PAC although, based on fMRI alone, there is some debate on its exact borders and orientation (Moerel et al., 2014). Beyond the processing of sound frequencies, fMRI has allowed the topographical identification of spectro-temporal modulations (Santoro et al., 2017). While PAC performs a fine-grained specific analysis of low-level acoustic features (such as frequency and spectro-temporal modulations), it also plays a key role in the transformation from acoustic input to more complex representation and acoustic percepts (Moerel et al., 2019). After decoding of the acoustic input, the information is forwarded to non-primary auditory areas, such as the planum polare anterior to Heschl's gyrus, the planum temporale posterior to Heschl's gyrus, as well as the superior temporal gyrus (STG) spanning the lateral side of the temporal lobe. In both PP and PT stronger response to spectrally more complex band pass noise than tones have been reported (Petkov et al., 2006). Moreover, neuroimaging studies have shown that spectro-temporal modulation maps, along the STG, may relate to observed preference for more complex features, such as preference for slow modulations as seen in speech and music stimuli (Zatorre, Belin & Penhune, 2002).

While Chapter 2 and Chapter 3 describe psychoacoustic investigations of the contribution of rhythm to perception, in **Chapter 4** we focus on the parts and operations underlying the behavioral benefits induced by rhythms. In particular, we focus on rhythm induced perceptual detection. Human neuroimaging studies have identified the auditory cortex as a critical region for auditory perceptual detection (Giani et al., 2015; Gutschalk, Micheyl & Oxenham, 2008), and PAC specifically, (Wiegand & Gutschalk, 2012). Previous research suggests that feedforward information of stimuli and feedback processes (like attention) interact in early sensory areas. However, it has until recently not been possible to investigate in humans how feedforward and (modulating) feedback processes are embedded in the cortical circuitry of the auditory cortex and what their relevance for perception is. We can gain insights into these mechanisms by probing the mesoscopic scale at which feedforward, and feedback signals are spatially distinct (Douglas, Martin & Whiterridge, 1989; Douglas & Martin, 2004).

The mesoscale of cortical layers

The mesoscopic organizational level represents an intermediate level between single neurons (microscopic scale) and the organization of entire cortical areas (macroscopic scale). All neurons in the cortical circuitry, whether they are efferent, afferent, or intrinsic organize themselves with respect to cortical layers (Douglas & Martin, 2004). Thereby, the anatomical layered structure of the cortex supports the directional flow of information between cortical areas, which can be probed at the spatial mesoscale (Figure 2C; Shipp, 2005; Kaas & Hackett, 2000; Markov et al., 2014). Most information processing is local and contained within a cortical column, a small patch of cortex conserving certain response parameters across its depth (Linden & Schreiner, 2003), like frequency sensitivity (Abeles & Goldstein, 1970; De Martino et al., 2015). Yet the response of neurons within a column can be strongly influenced by lateral and feedback connections from different areas (Binzegger, Douglas & Martin 2004; Markov et al., 2013). The granular layer (IV) is considered the input layer, where feedforward sensory information arrives (Figure 2C, green arrows; Mitani et al.,1985), and layer IV is therefore most prominent in primary sensory cortices (Hackett et al., 1998). The primary auditory cortex receives massive feedforward input from the medial geniculate body of the thalamus (Cruikshank et al. 2002; Huang & Winer 2000). Feedback connections (Figure 2C, purple arrows) originate from deep pyramidal cells and terminate in supragranular and infragranular (i.e., non-granular) layers (Rockland & Pandya, 1979; Felleman & Van Essen,1991; Hackett et al., 2014).

Investigating the mesoscopic scale does not only allow bridging observations of microscopic brain structures, like neurons, with macroscopic brain areas, but it has the potential to allow linking the wealth of empirical observations at the microscopic and macroscopic scale with their respective theories of cognitive computations that require the interaction of both feedforward and feedback processes (Marquardt, 2019). Thus, by investigating the mesoscale of cortical layers, high-resolution fMRI research in conjunction with psychophysics, might decrease the gap between empirical observations and the required theory as marked by Krakauer et al (2017).

Non-invasively probing the spatial mesoscale using high-field fMRI entails further benefits. The larger field of view (here meant as how much of the brain we can inspect at once) compared to electrophysiological recordings, allows inspecting multiple cortical areas at the same time. In addition, fMRI allows observing the activity of areas typically difficult to reach using invasive recordings such as the anteromedial regions of STG, deep within the Sylvian Fissure (Schreiner & Winer, 2010).

Sampling the human mesoscopic cortical architecture with layer-dependent fMRI

To date, several studies have investigated the cortical architecture in humans using layer-dependent BOLD fMRI, but the majority has focused on the visual modality. This is in part due to the larger history and accumulated knowledge of neuroscience in vision (Hutmacher, 2019) and in part due to methodological considerations, such as the calcarine sulcus, the location of the primary visual cortex being (in some cases) a relatively flat piece of cortex, compared to Heschl’s gyrus where PAC is located. This feature allowed initial approaches to sample cortical layers and cortical columns (Menon & Goodyear, 1999; Yacoub, Shmuel, Logothetis & Ugurbil, 2007; Yacoub, Harel & Ugurbil, 2008). While early studies focused on mapping features through cortical depth (i.e. mapping columnar organization), subsequent studies observed depth-specific (i.e. layer) contributions in visual cortex (Polimeni, Fischl, Greve & Wald, 2010). Following this pioneering work, studies have shown a feedforward response to visual stimuli localized in middle cortical depths (Koopmans, Barth &

Norris, 2010; Olman et al., 2012), and increased feedback in deep (Kok, Bains, van Mourik, Norris, & de Lange, 2016) and superficial depths (Muckli et al., 2015; Marquardt et al., 2020). To disentangle feedforward and feedback, some studies have used illusory or predicted stimuli, while controlling for bottom-up (feedforward) stimulation (Kok et al., 2016; Muckli et al., 2015; Marquardt et al., 2020), which highlights the feasibility to investigate perceptual content using layer-fMRI.

In auditory cortex, the first layer-specific fMRI studies have focused on frequency sensitivity (and stability) across the cortical depth (De Martino et al., 2015; Ahveninen et al., 2016) as well as layer-specific feedback contributions targeting superficial depths (De Martino et al., 2015). Subsequent studies have probed the contributions of different depths to auditory (Moerel et al., 2019) and multi-sensory stimuli (Gau et al., 2020; Chai et al., 2021), highlighting the feasibility to investigate feedforward and feedback processing in the auditory system non-invasively in humans. In **Chapter 4**, we take the next step in auditory fMRI studies at high spatial resolution and attempt to link human perception and behavior to the mesoscale of cortical layers, thereby creating a mapping between the neural implementation and the behavioral level.

Outline of this thesis

Chapter 1 outlined the scientific rationale and context through which to look at the research brought forward here. We elaborated the idea of parts and operations forming a mechanism, which in turn explains a perceptual phenomenon and how to investigate these mechanisms. We briefly summarized a selection of relevant psychoacoustic and neuroimaging studies in anticipation of the behavioral findings in chapter 2 and chapter 3 and the neuroimaging findings in chapter 4.

Chapter 2 introduces the behavioral paradigm, probing the behavioral constraints of the task, such as sound- frequency (carrier) effects and rhythm effects.

Chapter 3 builds on the experiments conducted in chapter 2 and provides insights on the effect of temporal structure of stimuli as well as awareness of predictability affecting the perception of temporally shifted targets.

Chapter 4 takes a different step towards understanding the mechanism of temporal target detection, by localizing the neuronal computations within the brain. Specifically, we examined the layer-dependent fMRI activity in the auditory cortex in response to detected and undetected targets in rhythmic sounds

Chapter 5 forms a general discussion of the presented studies.

References

Abeles M and Goldstein MH, Jr. (1970) Functional architecture in cat primary auditory cortex: columnar organization and organization according

Ahveninen, J., Chang, W., Huang, S., Keil, B., Kopco, N., Rossi, S., Bonmassar, G., Witzel, T. & Polimeni, J.(2016). Intracortical depth analyses of frequency-sensitive regions of human auditory cortex using 7T fMRI." Neuroimage 143: 116-127.

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. Cognitive Psychology, 41(3), 254–311.

Bechtel, W. (2006). Discovering cell mechanisms: The creation of modern cell biology. Cam- bridge: Cambridge University Press.

Bechtel, W. (2008). Mechanisms in cognitive psychology: What are the operations? Philosophy of Science, 75(5), 983-994.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36(2), 421-441.

Bertelson, P., & Boons, J. P. (1960). Time uncertainty and choice reaction time. Nature, 187(4736), 531-532.

Binzegger, T., Douglas, R. J., & Martin, K. A. (2004). A quantitative map of the circuit of cat primary visual cortex. Journal of Neuroscience, 24(39), 8441-8453.

Bogen, J., & Woodward, J. (1988). Saving the phenomena. The philosophical review, 97(3), 303-352.

Bouwer, F. L., Honing, H., & Slagter, H. A. (2020). Beat-based and memory-based temporal expectations in rhythm: similar perceptual effects, different underlying mechanisms. Journal of Cognitive Neuroscience, 32(7), 1221-1241.

Brodmann, K. (1910). Feinere Anatomie des Großhirns. Handbuch der Neurologie (pp. 206-307). Springer, Berlin, Heidelberg.

Chai, Y., Liu, T. T., Marrett, S., Li, L., Khojandi, A., Handwerker, D. A., ... & Bandettini, P. A. (2021). Topographical and laminar distribution of audiovisual processing within human planum temporale. Progress in Neurobiology, 205, 102121.

Chaimow, D., Yacoub, E., Uğurbil, K., & Shmuel, A. (2018). Spatial specificity of the functional MRI blood oxygenation response relative to neuronal activity. NeuroImage, 164, 32–47.

Correa, A., Lupianez, J., & Tudela, P. (2006). The attentional mechanism of temporal orienting: Determinants and attributes. Experimental Brain Research, 169(1), 58-68.

Crick, F., & Koch, C. (2003). A framework for consciousness. Nature neuroscience, 6(2), 119-126.

Cruikshank SJ, Rose HJ, and Metherate R (2002) Auditory thalamo- cortical synaptic transmission in vitro. Journal of Neurophysiology 87:361–384

Da Costa, S., van der Zwaag, W., Marques, J. P., Frackowiak, R. S., Clarke, S., & Saenz, M. (2011). Human primary auditory cortex follows the shape of Heschl's gyrus. Journal of Neuroscience, 31(40), 14067-14075.

Dau, T., Kollmeier, B., & Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. The Journal of the Acoustical Society of America, 102(5), 2892–2905.

De Martino, F., Moerel, M., Ugurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2015). Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proceedings of the National Academy of Sciences, 112(52), 16036–16041.

De Martino, F., Yacoub, E., Kemper, V., Moerel, M., Uludag, K., De Weerd, P., Ugurbil, K., Goebel, R., & Formisano, E. (2018). The impact of ultra-high field MRI on cognitive and computational neuroimaging. NeuroImage, 168,366–382.

Di Salle, F., Esposito, F., Scarabino, T., Formisano, E., Marciano, E., Saulino, C., ... & Seifritz, E. (2003). fMRI of the auditory system: understanding the neural basis of auditory gestalt. Magnetic resonance imaging, 21(10), 1213-1224.

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. Neuroscience and Biobehavioral Reviews, 81, 181–187.

Douglas, R. J., & Martin, K. A. C. (2004). Neuronal circuits of the neocortex. Annual Review of Neuroscience, 27, 419–451.

Douglas, R. J., Martin, K. A. C., & Whitteridge, D. (1989). Microcircuit for Neocortex. Neural Computation, 1, 480–488.

Dumoulin, S. O., Fracasso, A., van der Zwaag, W., Siero, J. C. W., & Petridou, N. (2017). Ultra-high field MRI: Advancing systems neuroscience towards mesoscopic human brain function. NeuroImage, 1–13.

Fechner, G. T. (1860). Elemente der Psychophysik (Vol. 2). Breitkopf u. Härtel.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cerebral cortex (New York, NY: 1991), 1(1), 1-47.

Fetsch, C. R. (2016). The importance of task design and behavioral control for understanding the neural basis of cognitive functions. Current Opinion in Neurobiology, 37, 16–22.

Formisano, E. (2019). The “speech ready” auditory cortex. In P. Hagoort (Ed.), Human Language: From Genes and Brains to Behavior (pp. 481-499). MIT Press.

Formisano, E., Kim, D.-S., Salle, F. Di, Van De Moortele, P.-F., Ugurbil, K., & Goebel, R. (2003). Mirror-Symmetric Tonotopic Maps in Human Primary Auditory Cortex. Neuron, 40, 859–869.

Fracasso, A., Dumoulin, S. O., & Petridou, N. (2021). Point-spread function of the BOLD response across columns and cortical depth in human extra-striate cortex. Progress in Neurobiology, 202(February 2020), 102034.

Friston, K. (2010). The free-energy principle: a unified brain theory? Nature reviews neuroscience, 11(2), 127-138.

Gau, R., Bazin, P. L., Trampel, R., Turner, R., & Noppeney, U. (2020). Resolving multisensory and attentional influences across cortical depth in sensory cortices. Elife, 9, e46856.

Glennan, S. S. (1996). Mechanisms and the nature of causation. Erkenntnis, 44(1), 49-71.

Goldstein, E. B. (Ed.). (2009). Encyclopedia of perception. Sage.

Golgi, C. (1906). The neuron doctrine: theory and facts. Nobel lecture, 1921, 190-217.

Hackett, T. A., de la Mothe, L. A., Camalier, C. R., Falchier, A., Lakatos, P., Kajikawa, Y., & Schroeder, C. E. (2014). Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. https://doi.org/10.3389/fnins.2014.00072Frontiers in neuroscience, 8, 72.

Hackett, T. A., Stepniewska, I., and Kaas, J. H. (1998). Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J. Comp. Neurol. 394, 475–495.

Helmholtz, H.v., Handbuch der physiologischen Optik. 1896, Hamburg: Voss.

Heschl, R. L. (1878). - Über die vordere quere Schläfenwindung des menschlichen Großhirns. URL:

Huang CL and Winer JA (2000) Auditory thalamocortical projections in the cat: laminar and areal patterns of input. Journal of Comparative Neurology 427:302–331.

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. Proceedings of the National Academy of Sciences of the United States of America, 97(22), 11793–11799.

Kauffman, S. (1971). Articulation of parts explanations in biology and the rational search for them. In R.C. Bluck & R.S. Cohen (Eds.), PSA 1970 (pp. 257–272). Dordrecht: Reidel.

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective activation of the deep layers of the human primary visual cortex by top-down feedback. Current Biology, 26(3), 371-376.

Koopmans, P. J., Barth, M., & Norris, D. G. (2010). Layer-specific BOLD activation in human V1. Human brain mapping, 31(9), 1297-1304.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. Neuron, 93(3), 480-490.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. science, 320(5872), 110-113.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. Trends in Cognitive Sciences, 10(11), 494–501.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. Psychological Review, 106(1), 119–159.

Linden JF & Schreiner CE (2003) Columnar transformations in audi- tory cortex? A comparison to visual and somatosensory cortices. Cerebral Cortex 13:83–89

Los, S. A., Knol, D. L., & Boers, R. M. (2001). The foreperiod effect revisited: Conditioning as a basis for nonspecific preparation. Acta Psychologica, 106(1-2), 121-145.

Markov, N. T., Ercsey-Ravasz, M., Lamy, C., Gomes, A. R. R., Magrou, L., Misery, P., ... & Kennedy, H. (2013). The role of long-range connections on the specificity of the macaque interareal cortical network. Proceedings of the National Academy of Sciences, 110(13), 5187-5192.

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., ... & Kennedy, H. (2014). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. Journal of Comparative Neurology, 522(1), 225-259.

Marquardt, I. (2019). Investigations on bottom-up and top-down processing in early visual cortex with high-resolution fMRI. ProefschriftMaken Maastricht.

Marquardt, I., De Weerd, P., Schneider, M., Gulban, O. F., Ivanov, D., Wang, Y., & Uludağ, K. (2020). Feedback contribution to surface motion perception in the human early visual cortex. Elife, 9, e50933.

Marquardt, I., De Weerd, P., Schneider, M., Gulban, O. F., Ivanov, D., Wang, Y., & Uludağ, K. (2020). Feedback contribution to surface motion perception in the human early visual cortex. Elife, 9, e50933.

Marr, D. (1982). Vision: A Computational Approach (MIT Press)

Menon, R. S., & Goodyear, B. G. (1999). Submillimeter functional localization in human striate cortex using BOLD contrast at 4 Tesla: implications for the vascular point-spread function. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 41(2), 230-235.

Merzenich, M. M., & Brugge, J. F. (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain Research, 50(2), 275–296.

Merzenich, M. M., Knight, P. L., & Roth, G. L. (1973). Cochleotopic organization of primary auditory cortex in the cat. Brain Research, 63(C), 343–346.

Mitani, A., Shimokouchi, M., Itoh, K., Nomura, S., Kudo, M., & Mizuno, N. (1985). Morphology and laminar organization of electrophysiologically identified neurons in the primary auditory cortex in the cat. Journal of Comparative Neurology, 235(4), 430-447.

Mitra, P. P. (2014). The circuit architecture of whole brains at the mesoscopic scale. Neuron, 83(6), 1273–1283.

Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. Frontiers in Neuroscience, 8, 1–14.

Moerel, M., De Martino, F., Uğurbil, K., Yacoub, E., & Formisano, E. (2019). Processing complexity increases in superficial layers of human primary auditory cortex. Scientific reports, 9(1), 1-9.

Moerel, M., Martino, F. De, & Formisano, E. (2012). Processing of Natural Sounds in Human Auditory Cortex: Tonotopy, Spectral Tuning, and Relation to Voice Sensitivity. 32(41), 14205–14216.

Moerel, M., Yacoub, E., Gulban, O. F., Lage-Castellanos, A., & De Martino, F. (2021). Using high spatial resolution fMRI to understand representation in the auditory network. Progress in Neurobiology, 207,01887.

Moore, B. C. J. (2003). An Introduction to the Psychology of Hearing. Boston Academic Press, 3, 413.

Moore, B. C., & Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America, 110(2), 1067–1073.

Moore, B. C., & Sek, a. (1995). Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). The Journal of the Acoustical Society of America, 97(4), 2468–2478.

Morgan, M. J., Melmoth, D., & Solomon, J. A. (2013). Linking hypotheses underlying Class A and Class B methods. Visual Neuroscience, 30(5–6), 197–206.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage, 13(4), 684-701.

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., ... & Yacoub, E. (2015). Contextual feedback to superficial layers of V1. Current Biology, 25(20), 2690-2695. https://doi.org/10.1016/j.cub.2015.08.057

Nobre, A. C., & Van Ede, F. (2018). Anticipated moments: temporal structure in attention. Nature Reviews Neuroscience, 19(1), 34-48.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. proceedings of the National Academy of Sciences, 87(24), 9868-9872.

Olman, C. A., Harel, N., Feinberg, D. A., He, S., Zhang, P., Ugurbil, K., & Yacoub, E. (2012). Layer-specific fMRI reflects different neuronal computations at different depths in human V1. PloS one, 7(3), e32536.

Petkov, C. I., Kayser, C., Augath, M., & Logothetis, N. K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. PLoS biology, 4(7), e215.

Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010). Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. Neuroimage, 52(4), 1334-1346.

Polimeni, J. R., Renvall, V., Zaretskaya, N., & Fischl, B. (2017). Analysis strategies for high-resolution UHF-fMRI data. NeuroImage, April, 1–25.

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H. J., & Zilles, K. (2001). Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage, 13(4), 669-683.

Read, J. C. A. (2015). The place of human psychophysics in modern neuroscience. Neuroscience, 296, 116–129.

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. Trends in Cognitive Sciences, 22(10), 870–882.

Rockland, K. S., and Pandya, D. N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. Brain Res. 179, 3–20. doi: 10.1016/0006-8993(79)90485-2

Russel, W. M. S. and Burch, R. L. (n.d.). Russell 1959.pdf.

Saenz, M., & Langers, D. R. M. (2014). Tonotopic mapping of human auditory cortex. Hearing Research, 307, 42–52.

Santoro, R., Moerel, M., De Martino, F., Valente, G., Ugurbil, K., Yacoub, E., & Formisano, E. (2017). Reconstructing the spectrotemporal modulations of real-life sounds from fMRI response patterns. Proceedings of the National Academy of Sciences, 114(18), 4799–4804.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. Current opinion in neurobiology, 20(2), 172-176.

Shipp, S. (2016). Neural elements for predictive coding. Frontiers in Psychology, 7(NOV), 1–21.

Simpson, A. J. R. R., Reiss, J. D., & McAlpine, D. (2013). Tuning of Human Modulation Filters Is Carrier-Frequency Dependent. PLoS ONE, 8(8), 1–9.

ten Oever S, Schroeder CE, Poeppel D, van Atteveldt N, Mehta AD, Mégevand P, et al. Low-Frequency Cortical Oscillations Entrain to Subthreshold Rhythmic Auditory Stimuli. J Neurosci. 2017;37(19):4903–12.

Vaughan, J. T., Garwood, M., Collins, C. M., Liu, W., DelaBarre, L., Adriany, G., ... & Ugurbil, K. (2001). 7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 46(1), 24-30.

Vizioli, L., Yacoub, E., & Lewis, L. D. (2021). How pushing the spatiotemporal resolution of fMRI can advance neuroscience. Progress in Neurobiology, 207.

Waskom, M. L., Okazawa, G., & Kiani, R. (2019). Designing and interpreting psychophysical investigations of cognition. Neuron, 104(1), 100-112.

Wiegand, K., & Gutschalk, A. (2012). Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. NeuroImage, 61(1), 62–69.

Winer, J. A., & Schreiner, C. E. (Eds.). (2010). The auditory cortex. Springer Science & Business Media.

Woodrow H (1914) The measurement of attention. Psychol Mo- nogr 17:1–158

Yacoub, E., Harel, N., & Uğurbil, K. (2008). High-field fMRI unveils orientation columns in humans. Proceedings of the National Academy of Sciences, 105(30), 10607-10612.

Yacoub, E., Shmuel, A., Logothetis, N., & Uğurbil, K. (2007). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. Neuroimage, 37(4), 1161-1177.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. Trends in cognitive sciences, 6(1), 37-46.

Zilles, K., & Amunts, K. (2010). Centenary of Brodmann's map—conception and fate. Nature Reviews Neuroscience, 11(2), 139-145.

Zoellner, S., Benner, J., Zeidler, B., Seither-Preisler, A., Christiner, M., Seitz, A., Goebel, R., Heinecke, A., Wengenroth, M., Blatow, M., & Schneider, P. (2019). Reduced cortical thickness in Heschl’s gyrus as an in vivo marker for human primary auditory cortex. Human Brain Mapping, 40(4), 1139–1154. https://doi.org/10.1002/hbm.24434

2

Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences

Corresponding Manuscript:

Heynckes, M., De Weerd, P., Valente, G., Formisano, E., & De Martino, F. (2020). Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences. *PLOS One*, 15(6), e0234251.

Abstract

Regularity of acoustic rhythms allows predicting a target embedded within a stream, thereby improving detection performance and reaction times in spectral detection tasks. In two experiments we examine whether temporal regularity enhances perceptual sensitivity and reduces reaction times using a temporal shift detection task. Participants detected temporal shifts embedded at different positions within a sequence of quintet–sounds. Narrowband quintets were centered around carrier frequencies of 200 Hz, 1100 Hz, or 3100 Hz and presented at presentation rates between 1-8 Hz. We compared rhythmic sequences to control conditions where periodicity was reduced or absent and tested whether perceptual benefits depend on the presentation rate, the spectral content of the sounds, and task difficulty.

We found that (1) the slowest rate (1 Hz) led to the largest behavioral effect on sensitivity. (2) This sensitivity improvement is carrier-dependent, such that the largest improvement is observed for low-frequency (200 Hz) carriers compared to 1100 Hz and 3100 Hz carriers. (3) Moreover, we show that the predictive value of a temporal cue and that of a temporal rhythm similarly affect perceptual sensitivity. That is, both the cue and the rhythm induce confident temporal expectancies in contrast to an aperiodic rhythm, and thereby allow to effectively prepare and allocate attentional resources in time. (4) Lastly, periodic stimulation reduces reaction times compared to aperiodic stimulation, both at perceptual threshold as well as above threshold. Similarly, a temporal cue allowed participants to optimally prepare and thereby respond fastest.

Overall, our results are consistent with the hypothesis that periodicity leads to optimized predictions and processing of forthcoming input and thus to behavioral benefits. Predictable temporally cued sounds provide a similar perceptual benefit to periodic rhythms, despite an additional uncertainty of target position within periodic sequences. Several neural mechanisms may underlie our findings, including the entrainment of oscillatory activity of neural populations.

Introduction

Natural sounds are often characterized by a rhythm that enables the listener to predict and prepare for relevant events (Nobre & Van Ede, 2018), a finding confirmed by many studies probing the effects of periodicity and predictability on sound perception (Andreou, Kashino, & Chait, 2011; Ellis & Jones, 2010; Jones, Moynihan, Mackenzie, & Puente, 2002a; Lawrance, Harper, Cooke, & Schnupp, 2015; Morillon, Schroeder, Wyart, & Arnal, 2016; ten Oever, Schroeder, Poeppel, van Atteveldt, & Zion-Golumbic, 2014) . Recent results have shown that periodicity and predictability of rhythms have dissociable effects on reaction time and sensitivity. Periodic (isochronous) stimuli decrease the reaction time (Ellis & Jones, 2010; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Lange, 2009), even when their rhythm is not predictive of a target (Breska & Deouell, 2014; Sanabria, Capizzi, & Correa, 2011). On the other hand, predictability can increase perceptual sensitivity (Chang, Bosnyak, & Trainor, 2019; Henry & Herrmann, 2014; Lawrance et al., 2015; ten Oever et al., 2014) even when stimulus presentation is predictable but aperiodic (Herbst & Obleser, 2019; Morillon et al., 2016) .

Most previous studies have focused on broadband or a few selected frequencies and have often tested presentation rates around 1.5 Hz (Breska & Deouell, 2017; Hickok, Farahbod, & Saberi, 2015; Jones et al., 2002a; Lakatos et al., 2008, 2013; Morillon, Schroeder, & Wyart, 2014; ten Oever et al., 2017). This last choice is at odds with studies suggesting that the brain preferentially differentiates acoustic information covering a range of timescales through theta- (and gamma) frequency-band information (Teng & Poeppel, 2019), a result that highlights the relevance of the temporal modulations inherent in speech occurring at the syllabic-level (Ding et al., 2017; Poeppel & Assaneo, 2020). The importance of testing multiple carrier frequencies and presentation rates is stressed by psychoacoustic studies on amplitude modulation suggesting that rate and carrier frequency interact (Moore & Glasberg, 2001). To account for the interdependency between spectral and temporal processing, it has been suggested that the modulation filters modelling the auditory system (Dau, Kollmeier, & Kohlrausch, 1997) systematically change along the frequency axis. In particular, the auditory system may be optimized to track rapid modulations at high frequencies and slower modulations at lower carrier frequencies (Simpson, Reiss, & McAlpine, 2013).

Electrophysiological recordings in macaque auditory cortex suggest that the mechanism engaged by rhythmic sound processing is tonotopic and thus cannot be investigated when using broadband noise bursts (O’Connell et al., 2015), but requires the use of narrow-band stimuli or tones (Lakatos et al., 2013; O’Connell, Barczak, Schroeder, & Lakatos, 2014). These studies propose that the optimization of behavioral performance for periodic stimuli is achieved through the entrainment of neuronal oscillations, which temporally modulate the excitability of task-relevant neuronal populations (Henry & Obleser, 2012; Lakatos et al., 2008; Schroeder & Lakatos, 2009; ten Oever et al., 2017). Rhythmic external stimuli can entrain oscillations, during which neural delta (Lakatos et al., 2008) and theta oscillations (Ng, Schroeder, & Kayser, 2012) become aligned to the externally imposed rhythm. The entrained oscillations may form the basis of temporal predictions that can be beneficial for stimuli presented at the entrainment rhythm (Arnal & Giraud, 2012; Hickok et al., 2015).

Predictions can also be formed without periodicity, by using temporally cued associations. The behavioral benefits of temporal cueing have been shown in studies on foreperiod effects (Los, Knol, & Boers, 2001; Näätänen & Niemi, 1981). At the behavioral level, a predictable rhythmic sequence should lead to better perceptual discrimination at predicted moments in time than a predictable single interval. This has been tested in foreperiod-paradigms where the duration of an interval is judged, either in isolated pairs or with a preceding rhythm, and results indicate that discrimination thresholds improve (Henry & Herrmann 2014). Studies focusing on reaction times have compared effects of temporal cueing and periodic stimulus presentation both when the predictive information was provided symbolically, showing a benefit of periodicity over symbolic cueing (Ren et al., 2019), and temporally cued not showing a difference (Breska & Deouell, 2017; Breska & Ivry,

2018) and suggest a cumulative benefit of combining a temporal cue and a periodic rhythm (Ellis & Jones, 2010). Similarly, studies using pitch discrimination tasks (Herbst & Obleser, 2017; Jones, Moynihan, Mackenzie, & Puente, 2002b) have shown effects on accuracy and reaction times. Moreover, Herbst & Obleser (2019) have shown a benefit in accuracy derived from implicit predictability in temporally cued intervals. In line with a recent study by Chang et al. (2019) suggesting that delta-phase entrainment may relate differently to spectral-based and temporal sensitivity, it would be of interest to compare perceptual sensitivity and reaction times due to predictability derived from a periodic sequence and from a single acoustic temporal cue when using a temporal task. This would allow evaluating if a single mechanism underlies predictive effects induced by both a temporal cue and a periodic sequence.

In two behavioral studies, we systematically address these open points by investigating how rhythm and carrier frequency of periodically presented sounds contribute to the behavioral benefits of rhythmic sound perception (Experiment 1). In a pilot we compare a subset of these conditions to aperiodic sound sequences (see supplementary material). In experiment 2 we compare periodic and aperiodic sound sequences and additionally test whether the predictive advantage afforded by periodicity exceeds the advantage afforded by a temporal cue (Experiment 2). From previous findings (Chang et al., 2019; Morillon, Schroeder, Wyart, & Arnal, 2016; ten Oever et al., 2014) we expect a perceptual benefit in sensitivity and reaction times of periodic over aperiodic rhythms. In addition, we hypothesize that a periodic rhythm should result in a benefit over a single temporal cue, especially when targets are presented close to perceptual threshold. With regard to the difference between various periodic rhythms, we expect behavioral improvement to occur around the peaks of the external rhythm, as opposed to a general improvement due to periodic stimulation (Lakatos et al., 2013). In particular, we hypothesize that an external rhythm in the theta frequency range (~ 4 Hz), would result in improved behavioral performance compared to other presentation rates.

Materials & Methods

Participants

Prior to testing, participants were screened for normal hearing (≤ 20 dB) at audiometric test frequencies ranging from 0.25-8 kHz. In experiment 1a and 1b, two sets of 20 participants each took part in the experiments. Two participants were excluded, one due to failure to comprehend the task and one due to discomfort caused by the loudness of the stimuli, resulting in the abortion of the experiment. As a result, we analyzed data collected from n = 19 participants for experiment 1a (15 females, 4 males) and 1b (12 females, 7 males) each, and n = 20 participants for experiment 2 (13 females, 7 males). The size of n was determined based on studies reporting similar effect sizes (Morillon et al., 2016). The Ethics Review Committee of the Faculty of Psychology and Neuroscience (ERCPN) at Maastricht University granted approval for all studies and all participants gave informed consent.

Stimuli & Design

All stimulus presentation scripts were written in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA), using the Psychophysics toolbox (Brainard, 1997). Sounds were created at a 44.1 kHz sampling rate, 16-bit resolution and delivered through Sennheiser HD650 headphones. Data and analysis scripts are publicly available (https://doi.org/10.5281/zenodo.3695583). Exemplary stimuli can be found here (https://doi.org/10.5281/zenodo.3549376). Participants were seated in a sound-attenuated chamber. Instructions were presented on a computer monitor and responses collected with a standard keyboard. Participants were asked to detect a target; a temporal shift (TS) of a narrow-band sound embedded in a sequence of quintets. Narrowband sounds were centered around carrier frequencies of 200 Hz, 1100 Hz, or 3100 Hz. The passbands around the carriers were constructed using equivalent rectangular bandwidths (ERBS = 4; Moore, 2003). Each passband consisted of a summation of 21 sinusoids with amplitude normalized to 1 and a random onset phase. A quintet consisted of five 10ms narrowband sounds, each separated by 10ms (see inset 1 Fig 1C). Targets were constructed by shifting in time the third sound in a quintet (see inset 2 Fig 1C). Depending on the experiment, this shift was either fixed at 6ms (Experiment 1) or ranged between 1.5-7ms (Experiment 2; see Table 1). In both experiments, during a trial and up to 1 second after a quintet sequence finished, participants could press a button upon detecting a TS or another button at the end of a sequence indicating they did not perceive a TS. Quintet sequences had either a periodic or aperiodic repetition of quintets or occurred in the form of a temporal cue, where a single quintet cued the following quintet that possibly contained a TS. Table 1 reports the design parameters for each experiment. The distribution of TS was non-uniform such that TS occurred in 75% of the trials. Targets appeared at one of 3 possible quintet positions within the periodic or aperiodic sequences, or in the cued quintet. In experiment 1a targets TS were presented in quintet 9, 10 or 11 (e.g. occurring after 9, 10 or 11s for the 1 Hz rhythm). With increasing rhythm, targets thereby occurred earlier in time within a sequence (the 9th quintet at 2 Hz occurs at 4.5 s). Experiment 1b controls for the earlier occurrence of targets with increasing rhythm, by keeping the time until a TS constant across rhythm (a TS occurring in the 9th quintet at 1 Hz would occur in the 18th quintet at 2 Hz). Similar to experiment 1b, TS occurred at a fixed time in both the pilot experiment preceding experiment 2 (see supplementary material), as well as in experiment 2 itself.

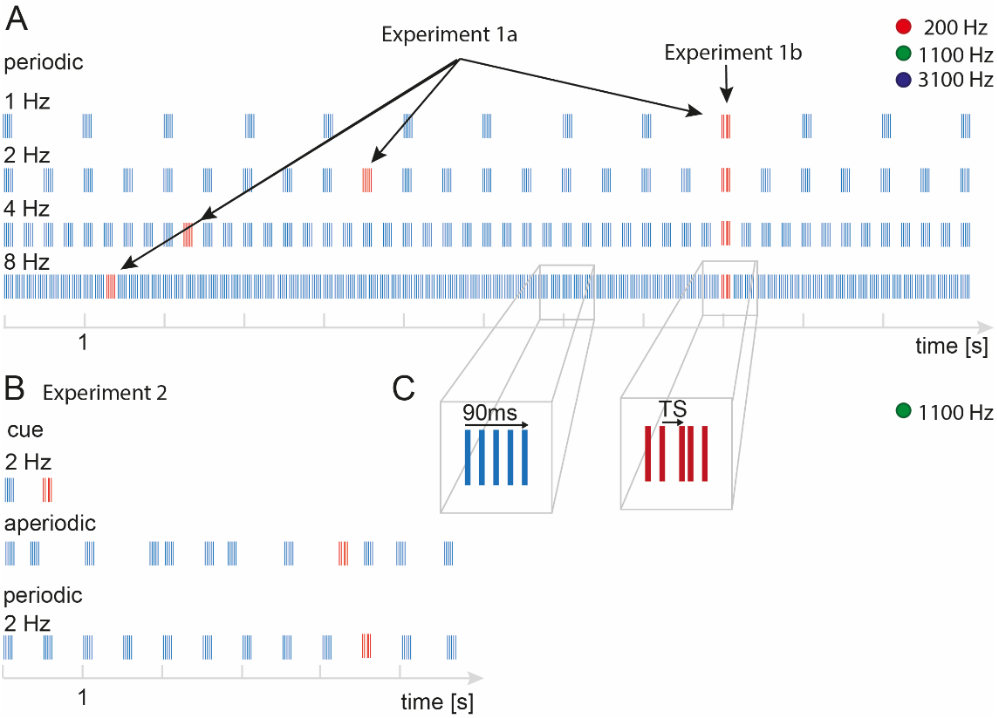


Figure 1. Stimuli. A. Experiment 1. Narrowband quintets were presented at 4 different presentation rates (1,2,4,8 Hz) and 3 different carrier frequencies (200, 1100, 3100 Hz). Target position (exemplars in red) was varied between experiment 1a and 1b. In experiment 1a targets occurred in either the 9th, 10th or 11th quintet within a sequence, while in experiment 1b targets appeared at the same time across rhythms (at 9,10 or 11 s). **B.** Stimuli in experiment 2 were 6s aperiodic and periodic sequences at 2 Hz, correspondingly the cue condition had an ISI of 500ms. The carrier frequency was 1100 Hz. **C.** Quintet structure. Narrowband sounds of 10ms length centered around the respective carrier frequencies were organized in a hierarchical rhythmic structure. Five sounds repeating at 50 Hz (10ms ISI) create a quintet (inset 1), while these quintets are repeated at a slow (a)periodic rhythm. Target stimuli (TS, see inset 2) had a different temporal structure: the third tone in a quintet was temporally shifted at 6ms in experiment 1a, 1b and 2 and between 1.5-7ms in experiment 3, as determined by the participant’s 70% detection threshold.

Table 1. Design and overview of stimuli conditions.

	Temporal structure	rhythm	Trials/Blocks	Carrier	Task difficulty	Motivation
Exp 1a	Periodic	1,2,4,8 Hz	48 trials (12 per rhythm) *10 blocks	200 Hz 1100 Hz 3100 Hz	6 ms TS	Test preferred rhythm and carrier
Exp 1b	Periodic	1,2,4,8 Hz	48 trials (12 per rhythm) *10 blocks	200 Hz 1100 Hz 3100 Hz	6 ms TS	Rhythm effect in Exp 1a due to target position?
Exp 2	Periodic Aperiodic Cue	2 Hz	72 trials (24 per condition) * 6 blocks	1100 Hz	1.5-7 ms TS (70% threshold)	Periodic benefit when task more difficult?

In experiment 1a (n =19) & 1b (n = 19) we investigated the effect of periodic rhythms (1, 2, 4, 8 Hz) and the carrier frequency (200, 1100 and 3100Hz) on target detection sensitivity (measured in d’) and reaction time. Every trial consisted of 12s periodic sequences of quintets. We conducted two versions of this experiment, in which we varied the position of the targets within the sequences. In experiment 1a we kept the number of preceding quintets until the TS constant across rhythms. This systematically reduced the time until a target appeared with increasing rhythm frequency. In experiment 1b, we kept the time at which targets appeared constant, thereby presenting an increasing number of preceding quintets with increasing rhythms prior to presenting a target quintet. The rationale to assess and compare these two experiment versions is that the strength of entrainment might increase with additional repetitions, and the effect of rhythm may be confounded by the systematic effect of target position in a sequence. Trials were counter-balanced and randomized with respect to target position and carrier frequency, and presented in (counterbalanced) blocks per rhythm, consisting of 12 trials. After four blocks (one per rhythm) subjects had a break. In total participants received 10*4 blocks, summing to a total of 480 trials. Subjects underwent a brief training session (8-12 trials) using a subset of the stimuli. During training, visual feedback on the performance was provided. Prior to the training and main experiment, participants adjusted the intensity of the sounds to equalize their perceived loudness. When comparing the three carrier frequencies at equal intensities, 200 Hz was generally perceived as softer and 3100 Hz perceived as louder relative to the reference frequency of 1100 Hz and were adjusted accordingly. These observations are in line with equal loudness contours (Moore, 2003).

Unless explicitly stated, the stimuli in the pilot experiment preceding experiment 2 (see supplementary material) and stimuli in experiment 2 were identical to those used in experiment 1b. We decided to limit the number of conditions and chose the carrier frequency and rhythm based on our findings in experiment 1 showing that with a 2 Hz rhythm and a 1100 Hz carrier frequency behavioral performance was intermediate, and thus we could expect the behavior to be modulated by the manipulations in experiment 2. In the pilot (n=20; 11 females, 9 males) we examined the effect of (average) rhythm on target detection, by comparing periodic predictable and aperiodic unpredictable sequences of quintets. The TS target remained fixed at 6ms. To create an aperiodic sequence, the main constraint was to present the same number of quintets in the same amount of time (compared to the periodic conditions) at aperiodic inter-stimulus intervals (ISIs). For instance, the aperiodic sequence corresponding to the 1 Hz rhythm had to be comprised of 12 quintets in 12 seconds. As a result, ISIs had to be sampled within two intervals, shorter or longer compared to the corresponding periodic condition. Periodic sequences at 1Hz have a 1s ISI, the corresponding aperiodic condition sampled ISI from two distributions with mean equal to 250ms and 1500ms. Similarly, periodic sequences at 2 Hz have a 500ms ISI, the corresponding aperiodic condition sampled ISI from two distributions with mean equal to 100ms and 733ms. In each aperiodic condition, the average over both sampled distributions approximates the periodic condition (1.07 Hz and 2.2 Hz respectively). Due to the nested temporal structure of the stimuli it was not possible to create aperiodic sequences at average rhythms of 4 Hz and 8 Hz, as the interval between quintets was too short. The participants performed 80 trials of the target detection task on aperiodic stimuli. Stimuli were presented in blocks of 16 trials, in which the average rhythm was constant.

D' scores of the pilot experiment approached ceiling for many of the participants. Therefore, in experiment 2 (n=20; 13 females, 7 males) we adjusted the difficulty range of the task to capture a modulation of behavior by periodicity. In addition, the stimuli were shortened to 6s, the (average) rhythm was fixed at 2 Hz and the carrier frequency at 1100 Hz, to limit the number of conditions and experiment duration. Aperiodic stimuli in experiment 2 were created similar to aperiodic stimuli in the pilot. Rhythmic stimuli (periodic predictable, aperiodic unpredictable) were compared to a temporal cue condition to test whether the predictive benefits derived from periodicity are larger than those derived from a single temporally predictive cue. Trials were presented in blocks (grouped by condition periodic, aperiodic, cue); within each block, trials were randomized, and the block order was counterbalanced across participants. Each participant completed a total of 432 trials. Task difficulty was determined through a staircase procedure preceding every block of trials, in which TS were set to achieve a behavioral performance at 70 % detection threshold, as determined by means of a 2 down 1 up procedure. The termination criterion was after 200 trials or 15 reversals. The TS varied on a fixed step-size of 10 logarithmically spaced steps between 7ms and 1.5 ms (see Figure S3 for average TS size per staircase preceding a block).

Statistical analysis

All analyses were conducted in MATLAB 2017a (The MATHWORKS Inc., Natick, MA, 234 USA). For each participant, the d' sensitivity index of signal detection theory and mean log-reaction times (logRT) of correct trials were calculated. Reaction time was calculated relative to target onset.

Statistical analysis of reaction times and sensitivity were carried out using a Generalized Linear Mixed Model (GLMM) with Matlab's *fitGLME* function. We assessed the model fits using likelihood ratio tests (using the function *compare* for GLMM). Contrasts were carried out performing an F-test on the specified fixed effects of the GLMM (using the function *coefTest*). Reaction times were fitted using the default identity link function, unlike the fitting of the sensitivity data for which a probit link function was used as described in more detail below.

Traditionally, d' is estimated by counting the frequency of an observer reporting 'yes' conditional on the presence and absence of a signal (i.e. the hit and false alarm rates) and taking the difference of these values on a z-transformed scale (Green & Swets, 1974). In the present work, statistical group analyses on d' are carried out using a GLMM, (Knoblauch & Maloney, 2012). We estimate both model parameters and d' simultaneously to determine the effect of carrier and target position, periodicity and temporal cueing on the population, rather than estimating d' on each condition separately and feeding the estimated values to a second level analysis. This statistical framework extends multiple linear regression to non-normal data such as count data and binary outcomes and it is more suited to handle extreme cases (100% hits or 0% false alarms). Within this framework, d' can be estimated by linearly modeling the behavioral outcomes (i.e. 'yes' or 'no') with a predictor X coding for the presence or absence of the target (see right side of eq. (1); and "Target" predictor in table 2). To fit an equal-variance Gaussian signal detection model an inverse Gaussian (probit) link function is used, where g is the link function and X represents the presence or absence of the signal.

$$g(E[\Pr(Resp =' Yes')]) = \beta_0 + \beta_1 X \tag{1}$$

When the signal is absent (i.e. $X = 0$), β_0 provides an estimate of the normal quantile of the false alarm rate. When the signal is present (i.e. $X = 1$), β_1 reflects the difference between hit and false alarm rate on the probit scale (hence, the difference between z-scaled hit and false alarm rates), or d'. The different experimental conditions are then added as predictors, and the estimated d' for each of these conditions (hence our effect of interest) is described by the interaction term between X ('target') and the respective condition predictor (see Knoblauch & Maloney, 2012 chapter 3.3.5). Unstandardized effect sizes (betas) are reported in units of the dependent variable (d ' or logRT), allowing for a meaningful comparison, in line with general recommendations on how to report effect sizes in psychological research (Pek & Flora, 2018).

For our visualization, we estimated d' as it is traditionally computed. Standard errors of d' group effects displayed in Fig 2 and 4 were obtained by non-parametric bootstrap sampling of estimated d' values, carried out at the subject level (N = 1000). The mean was used as a measure of central tendency around which 95% confidence intervals were created. All planned contrasts were corrected for multiple comparisons, using Bonferroni correction.

Table 2. Wilkinson notation of final model in each experiment.

Exp 1	d' ~ Criterion + Target*Rhythm*Carrier*Experiment + (1|Subject)
	logRT ~ 1+ Rhythm*Carrier*Experiment + (1|Subject)
Exp 2	d' ~ Criterion+ Target*temporalCondition + (1|Subject)
	logRT ~ 1+ temporalCondition

Note. Criterion is an additional predictor reflecting the intercept (normally notated as 1, here re-parameterized to -1 to reduce correlation between fixed effects (see p 262, Knoblauch & Maloney, 2012))

Results

Comparing periodic sequences at different carrier frequencies (experiment 1)

We investigated the benefits in perceptual sensitivity and reaction times associated with rhythmic sound presentation as a function of the rhythm and the carrier frequency in two experiment variants in which we varied the target position within a sequence (Experiment 1a and 1b). The data from both experiments were fitted using a single large GLMM (see table 2 for final notation of model).

Slow entraining rhythms improve target detection

The analysis showed a significant interaction of experiment and rhythm on d’ (F (3,864) = 4.528, $p < 0.01$), (see table 3). Therefore, we analyzed the effect of rhythm separately for experiment 1a and 1b, revealing an effect of rhythm on sensitivity in both, experiment 1a (Fig 2A), F (3,432) = 9.3704.19, $p < 0.001$) as well as experiment 1b (Fig 2B), (F (3,432) = 26.083, $p < 0.001$). Follow-up tests showed, for both experiment variants, a parametric effect of rhythm on sensitivity (see table 4). Counter to our hypothesis of an inverted U-shape where 4 Hz would perform best, we observed a parametric effect of rhythm. The slowest rhythms (1 Hz & 2 Hz) led to significantly higher sensitivity compared to the fastest (8 Hz) in both experiments. In addition, in experiment 1b the slowest rhythms (1 Hz & 2 Hz) led to significantly better sensitivity than 4 Hz as well. (see table 4 for specific contrasts). We did not observe this parametric effect of rhythm on reaction times (Fig 3) (F (3,431) = 2.498 $p > 0.05$).

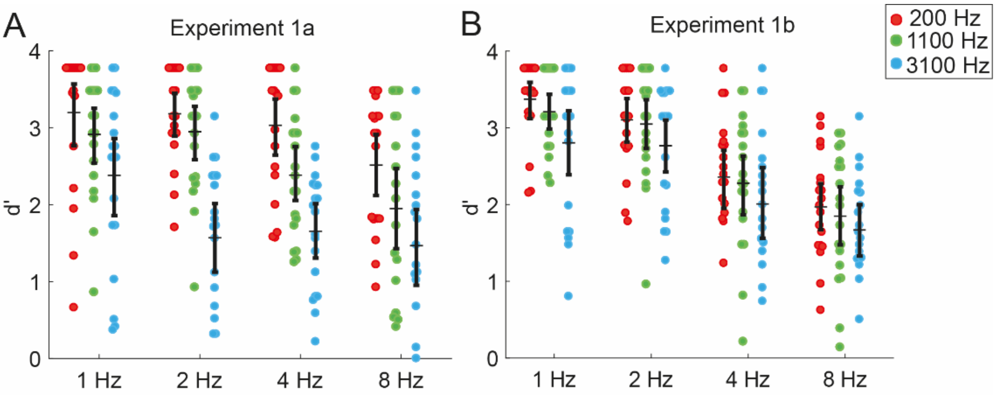


Figure 2. Both, the rhythm and carrier frequency parametrically affect sensitivity. Average d' of each participant (colored dots). Horizontal line indicates group mean, and errorbars depict bootstrapped 95% confidence intervals (at subject-level). **A** Experiment 1a; target occurred after a constant number of quintets per rhythm. **B** Experiment 1b; target occurred after constant time across rhythms.

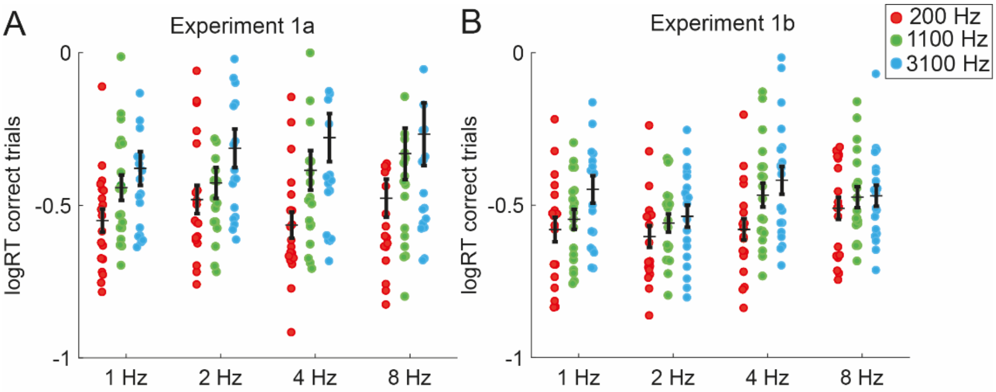


Figure 3. Carrier frequency, but not rhythm, affects reaction times in both experiments Mean logRT of each participant. Horizontal line indicates group mean, and errorbars depict SEM. **A** Experiment 1a; target occurred after a constant number of quintets per rhythm. **B** Experiment 1b; target occurred after constant time across rhythms.

Table 3. Planned contrasts sensitivity.

	β	**F**	**p**
(2 Hz-1 Hz):Experiment	-0. 8656	F(1, 864)=4.66	p>0.05
(4 Hz-1 Hz):Experiment	**-1.3884**	**F(1, 864)=12.82**	**p<0.01**
(8 Hz-1Hz):Experiment	**-1.1338**	**F(1, 864)=8.67**	**p <0.05**
(4 Hz-2 Hz):Experiment	-0.5228	F(1, 864)=2.63	p>0.05
(8 Hz-2 Hz):Experiment	-0.2682	F(1, 864)=0.71	p>0.05
(8 Hz-4 Hz):Experiment	0.2546	F(1, 864)=0.71	p>0.05
200 Hz- 1100 Hz	0.0983	F(1, 864)=0.18	p>0.05
200 Hz-3100 Hz	**0.9364**	**F(1, 864)=19.42**	**p <0.001**
1100 Hz-3100 Hz	**-0.8381**	**F(1, 864)=14.92**	**p <0.001**

Notes. Estimates are in d’. 1 Hz and 1100 Hz are reference categories for dummy coding scheme. Bold values indicate statistically significant results. $p < 0.05$, Bonferroni corrected. Each row refers to a contrast that interacts with the target predictor.

Table 4. Planned contrasts of rhythm per experiment.

	contrasts	**β**	**F**	**p**
Experiment 1a	(2 Hz-1 Hz)	0.0750	F(1, 432)=0.10	p>0.05
	(4 Hz-1 Hz)	-0.4874	F(1, 432)=4.13	p>0.05
	(8 Hz-1 Hz)	**-1.0152**	**F(1, 432)=19.29**	**p<0.001**
	(4 Hz-2 Hz)	0.5624	F(1, 432)=5.24	p>0.05
	(8 Hz-2 Hz)	**1.0903**	**F(1, 432)=21.13**	**p<0.001**
	(8 Hz-4 Hz)	0. 5279	F(1, 432)=5.07	p>0.05
Experiment 1b	(2 Hz-1 Hz)	-0.7902	F(1, 432)=6.11	p>0.05
	(4 Hz-1 Hz)	**-1.8744**	**F(1, 432)=37.83**	**p<0.001**
	(8 Hz-1 Hz)	**-2.1465**	**F(1, 432)=48.59**	**p<0.001**
	(4 Hz-2 Hz)	**1.0842**	**F(1, 432)=26.97**	**p<0.001**
	(8 Hz-2 Hz)	**1.3563**	**F(1, 432)=40.41**	**p<0.001**
	(8 Hz-4 Hz)	0.2721	F(1, 432)=2.05	p>0.05

Notes. Estimates are in d'. 1 Hz and 1100 Hz are reference categories for dummy coding scheme. Bold values indicate statistically significant results. $p < 0.05$, Bonferroni corrected. Each row refers to a contrast that interacts with the target predictor.

Carrier frequency of the stimulus affects target detection and reaction time

We observed a parametric main effect of carrier frequency on sensitivity F (2,864) =12.2, p<0.001). Participants were more sensitive in detecting a target when listening to a sequence with a 200 Hz carrier compared to the high carrier frequency at 3100 Hz (beta = -0.9364; F (1, 864) =19.42, p=<0.001), but also when comparing the 1100 Hz carrier against 3100 Hz carrier (beta = -0.8381; F (1, 864) =14.92, p<0.001). The difference in sensitivity between the 200 Hz and 1100 Hz was not significant (beta = 0.0983; F (1, 864) =0.18). In addition, we observe that the carrier frequency had a significant effect on logRT (F (2,431) =7.539, p< 0.001). Comparisons showed that the lowest carrier frequency led to significantly faster responses compared to the middle carrier (beta = -0.1078; F (1, 431) =5.87, p<0.05), and the highest carrier frequency (beta= -0.1709; F (1, 431) =14.74, p<0.001) (Fig 3).

Carrier Frequency and rhythm do not interact in their effects on sensitivity and reaction time

The interaction of rhythm by carrier was not significant in reaction times (F (6,431) = 1.125, p > 0.05) or sensitivity (F (6,864) = 2.0454, p > 0.05).

Effect of (a)periodic sequences and a temporal cue at perceptual threshold (Experiment 2)

Experiment 2 compared the effect of periodic predictable and aperiodic unpredictable sequences of 6 second length to a temporal cueing condition with a cueing interval matching the ISI of the sequences (500ms). We fitted two GLMMs for the reaction time data and d’ data respectively. Each model consisted of the fixed-effect within-subject factor *temporal structure* (3 levels; periodic, aperiodic, cue).

Temporally predictable stimulation (through a periodic rhythm or a cue) improves sensitivity

We compared target detection sensitivity in three temporal context conditions: predictable periodic, unpredictable aperiodic, (predictable) temporal cue. We found d’ to vary significantly as a function of temporal context (Fig 4A, F (2, 114) = 52.663, p< 0.001). Comparisons between the conditions revealed significant differences in d’ between the predictable periodic and unpredictable aperiodic sequences (beta = 0.5903; t (1, 114) =7.95, p< 0.001), as well as between the aperiodic sequence and the temporal cue (beta = 0.7154; t (1, 114) =9.3612, p<0.001). The difference in d' between predictable periodic sequences and the temporal cue was not significant (beta = 0.1251; F (1, 114) =2.49, p>0.05).

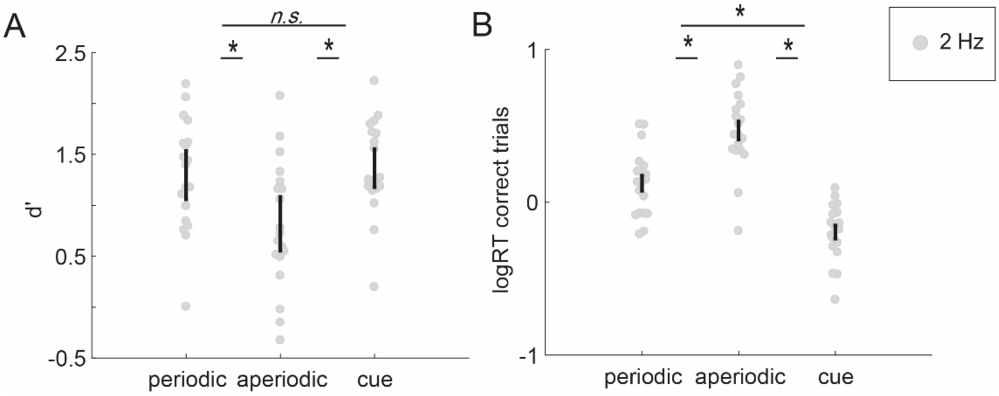


Figure 4. Experiment 2. When controlling for task difficulty, periodicity and a temporal cue improve hit reaction time and d’ compared to the aperiodic condition. TS size between 1.5-7ms (70% detection threshold). See Figure S3 for average TS size. A d’ per participant. Errorbars depict bootstrapped confidence intervals (at subject-level). B Mean logRT of each participant. Errorbars depict SEM.

Predictable temporal cue and a predictable periodic sequences improve reaction time compared to an aperiodic unpredictable sequence

The analysis of reaction times yielded a significant, yet different pattern between conditions (Fig 4B, F (2, 57) =48.1, p< 0.01). Comparisons revealed faster correct responses for periodic predictable than aperiodic unpredictable sequences (beta = -0.3317; t (1, 57) =-4.894, p< 0.001). Moreover, participants responded faster to temporally cued targets than periodic rhythms, (beta = -0.3331; F (1, 57) =24.15, p<0.001) as well as aperiodic (beta = -0.6647; t (1,57) =9.808, p<0.001) rhythms.

Thus, temporal predictability (whether through a cue or a periodic rhythm) led to an improvement of auditory sensitivity, while participants additionally benefit in their response times from the periodicity of stimulation compared to aperiodic stimulation, and a temporally predictable cue leading to the fastest response.

Discussion

Using a temporal shift detection task we asked how the rate and the carrier frequency of a predictable periodic rhythm influence both reaction time and sensitivity of perceptual decisions. Moreover, we asked whether the predictive advantage derived from periodic stimulation is larger than that afforded by aperiodic stimulation or a single temporal cue. The data show that (1) the largest sensitivity improvement is observed for the slowest rhythm (1 Hz); (2) sensitivity improvement is larger for low-frequency (200 Hz) carriers compared to 1100 Hz and 3100 Hz carriers; (3) periodic stimulation significantly reduces reaction times compared to aperiodic stimulation (speeded responses were observed both at perceptual threshold as well as above threshold during the pilot experiment); (4) a response to a temporal cue is faster than a response to periodic stimulation (5) periodic stimulation and cueing significantly increase sensitivity compared to aperiodic stimulation.

Experiment 1. Largest sensitivity improvements occur at slowest rhythm

Psychophysical findings show that sensitivity towards amplitude modulation detection of noise is highest for humans in the (speech) range of 2-4 Hz, while highest for macaques in the range of 30-60 Hz (O’Connor, 2011). These and other findings have led to the notion that the human auditory cortex is considered to be speech-ready, therefore, we expected a peak in perceptual sensitivity with periodic sound presentation around 4 Hz. However, our results show that listeners’ performance was highest at a slow rhythm and decreased with increasing rate of rhythm. A similar pattern of preference for slow rhythms (e.g. 2 Hz) as opposed to faster rhythms (8 Hz) has also been shown for cortical synchronization (entrainment) to speech in noise. Phase-locking of neural activity to speech embedded in noise decreased from low (2 Hz) to high (8 Hz) frequencies, correlating with speech intelligibility (Ding and Simon, 2013). This pattern resembles the linear decrease across rhythms observed here and supports the predominant role of delta band frequencies in auditory processing.

Moreover, we speculate that our findings may relate to the nature of the task and stimuli we used. The acoustic features of the isochronous stimuli may be closer to music and its temporal modulations than to the modulations inherent in speech. Temporal modulations in western music peak between 0.5 and 3 Hz, depending on the instrument and may contribute to a preference for slower modulations. As to why the peak of the modulation spectrum in music may be lower than that of spoken speech, it has been suggested that music like language is limited by the dynamic rate of movement of the effector (i.e. the frequency range where movement is most efficient; usually hands and arms in the case of music and articulators in the case of language) (Ding et al., 2017). Slow rhythms (1 Hz and 2 Hz) approximate the rate of spontaneous, hence most efficient, motor tempo (around 1.5 Hz for adults) as measured by spontaneous tapping-tasks (McAuley, Jones, Holub, Johnston, & Miller, 2006). A recent review by Morillon et al. (2019) substantiates the link between auditory processing and the motor system, suggesting a downward propagation of temporal predictions from the motor system involving delta-oscillations that shape auditory perception by imposing temporal constraints.

Lastly, we show that reaction times were not modulated by different rhythms. Preparatory response processes are typically studied in foreperiod (FP) - reaction time experiments, in which it is a classical finding that both the duration of the FP (usually in the range of seconds) as well as the variability of FP across trials within a block have a considerable effect on reaction times (Näätänen & Niemi, 1981). In the present study, ISI within a block were constant (i.e. low variability across trials), allowing the participant to prepare a motor response equally probable across conditions. Moreover, the absence of a difference in reaction times suggests that the time-range tested here allowed for non-specific (motor) preparations across all rates. Despite no difference in reaction times across rates, a difference in sensitivity across rates was observed, highlighting the perceptual benefit of rhythmic sound presentation, especially for slower rhythms.

Experiment 1. Carrier-dependent improvement of sensitivity and reaction times

Surprisingly, we found that the sensitivity decreased with increasing the carrier frequency. Based on the literature on temporal modulation processing in humans, we would expect higher sensitivity for higher carrier frequencies, as sounds are encoded by auditory spectral filters (tonotopic mechanism). These spectral filters are narrower at lower frequencies and wider at higher frequencies and limit the temporal resolution of the auditory system. Therefore detection performance of a temporal shift should decrease at lower carrier frequencies where the bandwidth is narrower (Moore, Peters, & Glasberg, 1990). Indeed, Moore et al. (1993) and Viemeister (1979) observed increasing modulation detection thresholds for decreasing center frequencies. We controlled for this effect, by adjusting the stimuli to have equivalent rectangular bandwidth (ERB, Moore 1990) and equivalent perceived intensity. Despite this equalization we observe an effect of carrier frequency. We suggest that the perceptual benefit at the low carrier frequency may be a product of temporal coding mechanisms. Phase locking up to 250 Hz has been observed in human intracortical recordings using click trains (Nourski et al., 2013). This temporal encoding seems to provide a perceptual benefit when making judgments about the presence of a temporally shifted target thereby improving sensitivity. Additionally, this benefit in sensitivity for the lowest carrier frequency was accompanied by an increase in response speed. Simpson, Reiss and McAlpine (2013) estimated sensitivity to a range of amplitude modulation frequencies (0.5 Hz to 33 Hz) across a large number of frequency carriers including and beyond the range of carriers tested here. Their results suggest that at low carrier frequencies cortical modulation filters are most sensitive to slow modulation rates, similar to the rates used here (1-8 Hz). They speculate that such a frequency-dependent modulation tuning is related to the neural processing of acoustic properties of speech (Giraud & Poeppel, 2012). Lastly, these findings suggest that behavioral effects of entrainment, may depend on the type of spectral stimulation used for entraining and probing, which may be especially relevant in the context of spectral tasks (O'Connell et al., 2015).

Experiment 1. Effect of rhythm not confounded by target position

The number of preceding quintets was varied between experiment 1a and 1b. In experiment 1a we kept the number of preceding quintets until the TS constant across rhythms. This systematically reduced the time until a target appeared with increasing rhythm. In experiment 1b, we kept the time at which targets appeared constant across rhythms, thereby presenting an increasing number of preceding quintets with increasing rhythms prior to presenting a target quintet. Our rationale to assess and compare these two experiment versions being, that the strength of entrainment might increase with additional repetitions. The effect of rhythm would then be confounded by the systematic effect of target position in a sequence. Indeed, we show in experiment 1a as well as 1b an effect of rhythm. By having controlled the position of the target across the different presentation rates we therefore conclude that there is a difference between rhythms and said effect was not confounded by the systematic effect of target position in a sequence. In experiment 2 targets embedded in aperiodic and periodic sequences occurred late within a sequence (similar to Experiment 1b). It would be interesting to see what the effect of aperiodic and periodic sequences is when targets are presented earlier. We may speculate that this would further enhance the detection difficulty of the task enhancing the benefit of periodicity.

Experiment 2. Effects of periodicity and cueing diverge in logRT and sensitivity

The results of experiment 2 show that both perceptual sensitivity and reaction times are improved when stimuli are presented in periodic rhythms compared to aperiodic rhythms, when using a temporal detection task at perceptual threshold. Note that these effects are only apparent when controlling for task difficulty. In a pilot study, where this was not done, we did not observe an effect of periodicity on sensitivity (see supplementary material). The results of experiment 2 support earlier findings reporting a benefit of predictability in periodic over aperiodic stimulation(Jones et al., 2002b; Lawrance et al., 2015; Morillon et al., 2016; Rimmele, Jolsvai, &

Sussman, 2011; Rohenkohl, Cravo, Wyart, & Nobre, 2012; Schmidt-Kassow, Schubotz, & Kotz, 2009; ten Oever et al., 2014; Wollman & Morillon, 2018). This is in line with the idea of oscillatory entrainment and dynamic attending theory (Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010) and highlights the relevance of using a task with sufficient difficulty, in contrast to the pilot. However, this finding points to a more general question of the benefit of entrainment in everyday life as most stimuli we encounter are seldom at perceptual threshold.

In addition, we were interested in contrasting the benefit in reaction times and sensitivity of a predictable periodic rhythm to a predictable (but not periodic) temporal cue. We expected a benefit in reaction times similar to (albeit different) Morillon et al. (2015) showing a benefit of a periodic predictable sequence over an aperiodic, predictable condition of increasing tempo. Moreover, we expected a benefit in sensitivity for the periodic rhythm compared to the temporal cue. We show in experiment 2, that a temporal cue enables a participant to respond significantly faster than a periodic rhythm, while there is no difference in perceptual sensitivity between these two forms of temporal structure. The cue predicts target occurrence with a 75% validity (25% of trials were catch-trials without a target). Similarly, the periodic rhythm allowed participants to predict when a subsequent target may occur with a similar validity for the trial. Yet participants were faced with the additional uncertainty as to which quintet within a sequence may contain the target. This uncertainty may be the reason for the observed slowing of reaction times in the periodic condition compared to the cue. The greater effectiveness of the temporal cue compared to the rhythm in terms of reaction time suggests that the temporal cue induced a more confident temporal expectation, in line with the finding that target-occurrence uncertainty impairs reorienting, thereby lengthening of reaction times (Correa, Lupiáñez, & Tudela, 2006). It would be of interest to compare these two conditions under similar uncertainty of target occurrence.

Interestingly though, despite said larger uncertainty the periodic condition is not significantly worse than the temporal cue in terms of sensitivity, therefore we suggest that a benefit of a periodic rhythm to some extent countered the increased uncertainty. Ten Oever et al. (2014) have shown that entrainment of low-frequency oscillations in the delta - range serves a mechanistic role in enhancing perceptual sensitivity of subthreshold periodic, predictable sound sequences compared to aperiodic sequences. Under the hypothesis that oscillations align more efficiently to a rhythmic structure as compared to a single interval, it is surprising that the periodic sequences here did not result in a sensitivity benefit over the temporal cue. See for instance Barnes & Jones (2000) or Drake & Botte (1993) showing an accuracy benefit of periodic sequence over cue in duration estimation. We speculate that the additional uncertainty of when within a sequence a target may occur, may have countered a benefit of the entraining rhythmicity of the sequence. Again, it would be of interest to compare instances of a predictable periodic rhythm and a predictable temporal cue conveying the same uncertainty. We would then predict sensitivity of the periodic predictable condition to be higher than the temporal cue condition.

At the neurophysiological level, we speculate that such a mechanism might be implemented by a more flexible phase reset model of neuronal oscillations (Breska & Deouell, 2017; J. M.; ten Oever, van Atteveldt, & Sack, 2015; Wilsch, Henry, Herrmann, Maess, & Obleser, 2015), (see Rimmele, Morillon, Poeppel, & Arnal, 2018, for a recent review), in which the motor system tracks temporal regularities (Morillon et al., 2019). Further research will be necessary to elucidate the mechanism and nature of top-down predictions and how these affect auditory perception.

Conclusion

We show that overall temporal modulations in the range of 1-8 Hz are better processed at lower carrier frequencies, as measured by reaction times and sensitivity (experiments 1a and 1b). Additionally, the same results point to the perceptual benefit of slow rhythms (1 and 2 Hz) over faster ones (4 and 8 Hz). The regularity of rhythms enables the use of prediction to make more precise inferences about when we should expect to find a target embedded within the stream and, as a result, improve detection performance. Indeed, we show in experiment 2 a perceptual benefit of periodic predictable sequences over aperiodic unpredictable sequences in terms of reaction times and sensitivity (the latter only present when using a sufficiently difficult task). Crucially, in experiment 2 we show that the predictive value of a cue and that of a temporal rhythm do not differ in terms of the sensitivity in detecting a target, albeit it has to be noted that the periodic condition contained a larger uncertainty where the target would appear. These findings encourage us to reflect on what the perceptual benefits of periodicity and predictability respectively are, as these effects may diverge when teased apart using different tasks thereby allowing to make assumptions about the underlying mechanisms involved. Here we showed that both the cue and the rhythm induce confident temporal expectancies about the future occurrence of targets to effectively prepare and allocate attentional resources. Taken together, we may speculate that multiple processes may co-occur that facilitate the processing of rhythmic and predictable stimuli, in which oscillations form an intrinsic temporal constraint, controlled by temporal predictions. Potentially, cueing effects occur due to a single phase-reset of ongoing oscillations and similarly a rhythmic benefit occurs due to either a stimulus driven entrainment of oscillations or repeated top-down phase resets.

Supplementary Material

Pilot Experiment 2

Materials & Methods

Participants

20 subjects (11 females, 9 males) participated in the pilot experiment. The Ethics Review Committee of the Faculty of Psychology and Neuroscience (ERCPN) at Maastricht University granted approval for all studies and all participants gave informed consent.

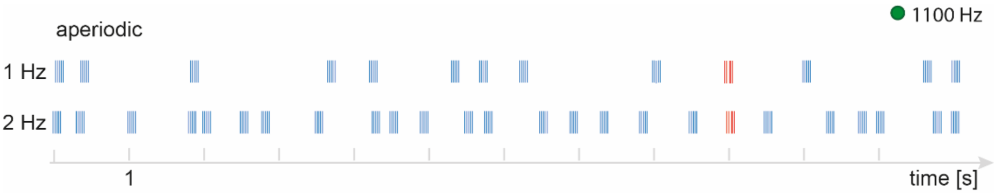


Figure S1. Stimuli pilot. Stimuli were presented with a carrier frequency of 1100 Hz. Two aperiodic conditions were presented with sequences of 12s length that had a similar number of quintets as periodic sequences at 1 Hz and 2 Hz of experiment 1.

Stimuli & Design

Similar to experiment 1, participants were asked to detect targets embedded in sound sequences of 12s length. The participants performed 80 trials of the target detection task on aperiodic stimuli. See methods experiment 2 of the main text on details of aperiodic stimuli. Stimuli were presented in blocks of 16 trials, in which the average rhythm was constant. The size of the temporal shift (TS) was fixed at 6ms occurring after a fixed time (similar to experiment 1b). During a trial and up to 1 second after a quintet sequence finished, participants could press a button upon detecting a TS or another button at the end of a sequence indicating they did not perceive a TS.

Statistical analysis

For details on statistical analysis see main text.

Table S1. Wilkinson notation of final model in pilot experiment

Pilot d’ ~ Criterion + Target*Rhythm*Experiment + (1 + Target:Rhythm|Subject)
logRT ~ 1 + Rhythm*Experiment + (1 |Subject)

Note. Criterion is an additional predictor reflecting the intercept (normally notated as 1, here re-parameterized to -1 to reduce correlation between fixed effects (see p 262, Knoblauch & Maloney, 2012))

Results

We examined the effect of (average) rhythm on target detection, by comparing periodic predictable sequences (matching conditions from experiment 1b) to aperiodic unpredictable sequences of quintets presented in the pilot.

Comparing aperiodic (pilot) and periodic sequences (experiment 1b) at 1 and 2 Hz above perceptual threshold

Two GLMMs with *rhythm* as within-subject factor (2 levels; 1 Hz, 2 Hz) and *experiment* as between-subjects factor were created to analyze log-reaction times and sensitivity data, comparing aperiodic and periodic presentation (Fig S1).

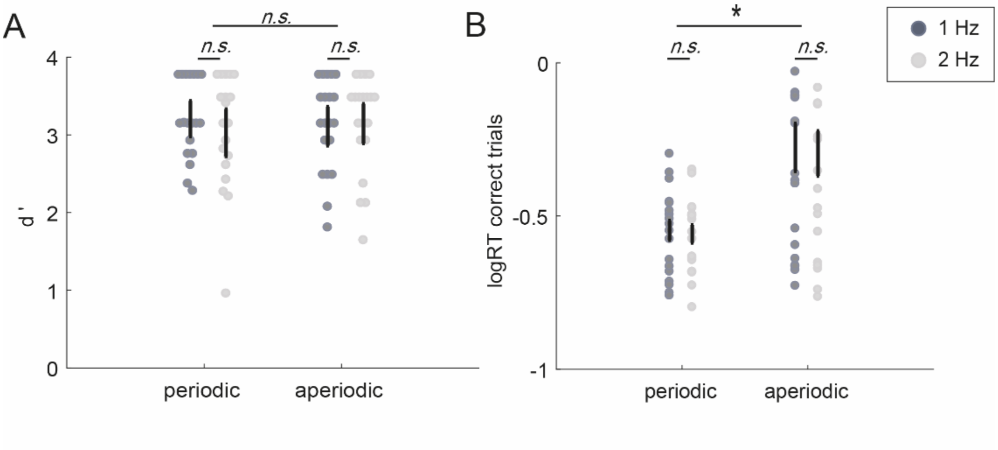


Figure S2. Periodicity improves hit reaction time but not d’ when TS fixed at 6 ms. Comparing matching conditions from experiment 1b (periodic) to experiment 2 (aperiodic). No difference between rhythms (1 & 2 Hz). **A** d’ per participant.

Errorbars depict bootstrapped confidence intervals (at subject-level). **B** Mean logRT of each participant. Errorbars depict SEM.

Periodicity improves reaction times but not sensitivity

Periodic stimulation improved reaction times when compared to the aperiodic stimulation (Fig S1 A, beta = 0.2708; (t (1, 74) = 3.336, $p < 0.01$). We did not find an effect of periodicity on perceptual sensitivity (Fig S1 B, beta = -0.0211; (t (1,148) = -0.051 $p>0.05$), when a detection task using a fixed target of 6ms was used (Fig 4; Experiments 1 and 2). We did not observe significant differences between the 1 and 2 Hz rhythm in either experiment in their effect on reaction times (beta = -0.012; t (1, 74) =-0.509, $p>0.05$) or on d' (beta= -0.4886; t (1, 148) =-1.160 $p>0.05$).

Discussion

The pilot results are discussed in the main body of the paper.

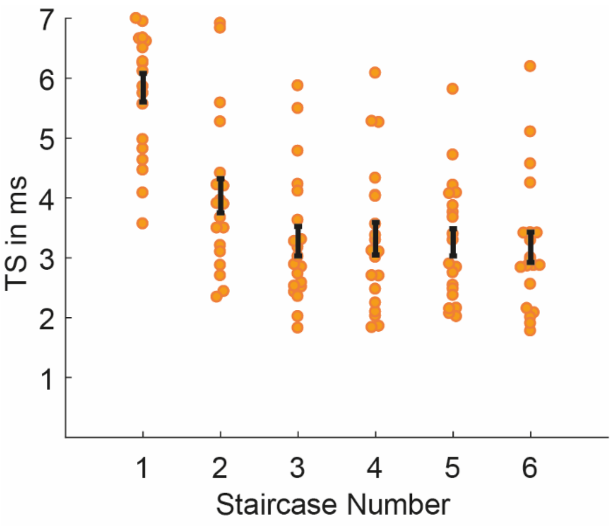


Figure S3. Decrease of TS size over course of experiment 2.TS size per participant of in total six staircase blocks. Each staircase preceded a trial block of experiment 2, ensuring sufficient task difficulty in the latter. Group average of to be detected TS decreases over the course of the experiment. During the staircase the TS varied on a fixed step-size of 10 logarithmically spaced steps between 7ms and 1.5 ms. The termination criterion was after 200 trials or 15 reversals. Errorbars depict SEM centered on group mean.

Acknowledgments

The authors thank Jacek Chudy and Louisa Zielke for help with data collection. We thank Sanne ten Oever for feedback on the manuscript and the reviewers for their valuable comments that helped improve this work. This work was funded by The Netherlands Organization for Scientific Research (NWO) Research Talent Grant 406.17.200 awarded to MH, EF and PdW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Andreou, L. V., Kashino, M., & Chait, M. (2011). The role of temporal regularity in auditory segregation. *Hearing Research*, *280*(1–2), 228–235.

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398.

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. *Cognitive Psychology*, *41*(3), 254–311. https://doi.org/10.1006/cogp.2000.0738

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. *Journal of Cognitive Neuroscience*, *26*(7), 1555–1571.

Breska, A., & Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biology*, *15*(2), 1–30.

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson’s disease. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), 12283–12288.

Chang, A., Bosnyak, D. J., & Trainor, L. J. (2019). Rhythmicity facilitates pitch discrimination: Differential roles of low and high frequency neural oscillations. *NeuroImage*, *198*, 31–43.

Correa, Á., Lupiáñez, J., & Tudela, P. (2006). The attentional mechanism of temporal orienting: Determinants and attributes. *Experimental Brain Research*, *169*(1), 58–68.

Dau, T., Kollmeier, B., & Kohlrausch, A. (1997). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *The Journal of the Acoustical Society of America*, *102*(5), 2892–2905.

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews*, *81*, 181–187.

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, *54*(3), 277–286.

Ellis, R. J., & Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. *Attention, Perception & Psychophysics*, *72*(8), 2274–2288.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100.

Henry, Molly J., & Herrmann, B. (2014). Low-Frequency Neural Oscillations Support Dynamic Attending in Temporal Context. *Timing and Time Perception*, *2*(1), 62–86.

Herbst, S. K., & Obleser, J. (2017). Implicit variations of temporal predictability:

Shaping the neural oscillatory and behavioural response. *Neuropsychologia*, *101*, 141–152.

Herbst, S. K., & Obleser, J. (2019). Implicit temporal predictability enhances pitch discrimination sensitivity and biases the phase of delta oscillations in auditory cortex. *NeuroImage*, *203*, 116198.

Hickok, G., Farahbod, H., & Saberi, K. (2015). The Rhythm of Perception: Entrainment to Acoustic Rhythms Induces Subsequent Perceptual Oscillation. *Psychological Science*, *26*(7), 1006–1013.

Jones, M. R., Moynihan, H., Mackenzie, N., & Puente, J. (2002a). Temporal Aspects of Stimulus-Driven Attending in Dynamic Arrays. *American Psychological Society*, *13*(4), 313–319.

Jones, M. R., Moynihan, H., Mackenzie, N., & Puente, J. (2002b). Temporal aspects of stimulus-driven attending in dynamoc arrays. *Psychological Science*, *13*(4), 313–319.

Knoblauch, K., & Maloney, L. T. (2012). *Modeling Psychophysical Data in R*. New York, NY: Springer New York.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science (New York, N.Y.)*, *320*(5872), 110–113.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The Spectrotemporal Filter Mechanism of Auditory Selective Attention. *Neuron*, *77*(4), 750–761.

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, *69*(1), 127–137.

Lawrance, E. L. A., Harper, N. S., Cooke, J. E., & Schnupp, J. W. H. (2015). Temporal predictability enhances auditory detection, *135*(6), 1–10.

Los, S. A., Knol, D. L., & Boers, R. M. (2001). The foreperiod effect revisited: Conditioning as a basis for nonspecific preparation. *Acta Psychologica*, *106*(1–2), 121–145.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367.

Moore, B. C., & Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *110*(2), 1067–1073.

Moore, B. C. J. (2003). An Introduction to the Psychology of Hearing. *Boston Academic Press*, *3*, 413.

Moore, B. C. J., Peters, R. W., & Glasberg, B. R. (1993). Detection of temporal gaps in sinusoids: Effects of frequency and level. *The Journal of the Acoustical Society of America*, *93*(3), 1563–1570.

Moore, B. C. J., Peters, R. W., & Glasberg, B. R. (2005). Auditory filter shapes at low center frequencies. *The Journal of the Acoustical Society of America*, *88*(1), 132–140.

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal Prediction in lieu of Periodic Stimulation. *Journal of Neuroscience*, *36*(8), 2342–2347.

Morillon, B, Arnal, L. H., Schroeder, C. E., & Keitel, A. (2019). Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. *Neuroscience & Biobehavioral Reviews*, *107*, 136–142.

Morillon, B, Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, *5*, 1–9.

Näätänen, R., & Niemi, P. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162.

Ng, B. S. W., Schroeder, T., & Kayser, C. (2012). A Precluding But Not Ensuring Role of Entrained Low-Frequency Oscillations for Auditory Perception. *The Journal of Neuroscience*, *32*(35), 12268–12276.

Nobre, A. C., & Van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48. https://doi.org/10.1038/nrn.2017.141

Nourski, K. V., Brugge, J. F., Reale, R. A., Kovach, C. K., Oya, H., Kawasaki, H., … Howard, M. A. (2013). Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. *Journal of Neurophysiology*, *109*(5), 1283–1295.

O’Connell, M. N., Barczak, A., Ross, D., McGinnis, T., Schroeder, C. E., & Lakatos, P. (2015). Multi-Scale Entrainment of Coupled Neuronal Oscillations in Primary Auditory Cortex. *Frontiers in Human Neuroscience*, *9*, 655.

O’Connell, M. N., Barczak, A., Schroeder, C. E., & Lakatos, P. (2014). Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(49), 16496–16508.

Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychological Methods*, *23*(2), 208–225.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*.

Ren, Y., Xu, Z., Wu, F., Ejima, Y., Yang, J., Takahashi, S., … Wu, J. (2019). Does Temporal Expectation Driven by Rhythmic Cues Differ From That Driven by Symbolic Cues Across the Millisecond and Second Range? *Perception*.

Rimmele, J., Jolsvai, H., & Sussman, E. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. *Journal of Cognitive Neuroscience*, *23*(5), 1136–1147. https://doi.org/10.1162/jocn.2010.21437

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882.

Rohenkohl, G., Cravo, a. M., Wyart, V., & Nobre, a. C. (2012). Temporal Expectation Improves the Quality of Sensory Information. *Journal of Neuroscience*, *32*(24), 8424–8428.

Sanabria, D., Capizzi, M., & Correa, Á. (2011). Rhythms That Speed You Up. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 236–244.

Schmidt-Kassow, M., Schubotz, R. I., & Kotz, S. A. (2009). Attention and entrainment: P3b varies as a function of temporal predictability. *NeuroReport*, *20*(1), 31–36.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176.

Simpson, A. J. R., Reiss, J. D., & McAlpine, D. (2013). Tuning of Human Modulation Filters Is Carrier-Frequency Dependent. *PLoS ONE*, *8*(8), 1–9.

ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., Mehta, A. D., Mégevand, P., … Zion-Golumbic, E. (2017). Low-Frequency Cortical Oscillations Entrain to Subthreshold Rhythmic Auditory Stimuli. *The Journal of Neuroscience*, *37*(19), 4903–4912.

ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., & Zion-Golumbic, E. (2014). Rhythmicity and cross-modal temporal cues facilitate detection. *Neuropsychologia*, *63*(1), 43–50.

ten Oever, S., van Atteveldt, N. M., & Sack, A. T. (2015). Increased Stimulus Expectancy Triggers Low-Frequency Phase reset during restricted vigilance. *J Neurosci*.

Teng, X., & Poeppel, D. (2019). Theta and Gamma Bands Encode Acoustic Dynamics over Wide-ranging Timescales. *BioRxiv*, (February), 547125.

Viemeister, N. F. (2005). Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America*, *66*(5), 1364–1380.

Wilsch, A., Henry, M. J., Herrmann, B., Maess, B., & Obleser, J. (2015). Slow-delta phase concentration marks improved temporal expectations based on the passage of time. *Psychophysiology*, *52*(7), 910–918.

Wollman, I., & Morillon, B. (2018). Organizational principles of multidimensional predictions in human auditory attention. *Scientific Reports*, *8*(1), 13466.

3

Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues

Corresponding manuscript:

Heynckes, M., Hoffmann, K., Formisano, E. & De Martino, F., De Weerd, P. (*submitted*). Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues.

Abstract

Sounds following a cue or embedded in a periodic rhythm are processed more effectively than sounds that are part of an aperiodic rhythm. One might expect also that a sound embedded in a periodic rhythm is processed more effectively than a sound following a single temporal cue. Such finding would follow from the idea that the entrainment of neural rhythmic activity by periodic stimuli renders the prediction of upcoming stimuli more efficient. We conducted two experiments in which we tested this idea. In a first experiment, targets in periodic and aperiodic rhythms, if they occurred, always appeared at the same moment in time, and thus were fully predictable, although participants remained unaware of this. In a second experiment, explicit instructions on the temporal location of the targets embedded in rhythms were provided. We assessed sensitivity and reaction times to the target stimuli in a difficult temporal detection task, and contrasted performance in this task to that obtained for targets temporally cued by a single preceding cue. Irrespective of explicit information about target predictability, target detection performance was always better in the periodic and temporal cue conditions, compared to the aperiodic condition. However, we found that the mere predictability of an acoustic target within a periodic rhythm did not allow participants to detect the target any better than in a condition where the target's timing was predicted by a single temporal cue. Only when participants were made aware of the specific moment in the periodic rhythm where the target could occur, did sensitivity increase. This finding suggests that neural entrainment by a periodic rhythm is not always sufficient to provide perceptual benefits, and that in some conditions these benefits may only occur in interaction with other factors such as explicit instruction and directed attention.

Introduction

Acoustic stimuli often have temporal structure which can be leveraged by the brain. Temporal structure may be rhythmic as in the sounds produced by a metronome (termed isochronous), or (quasi-) rhythmic as in speech or music. Temporal structure can also take the form of a temporal association between a cue and a stimulus presented after a given time interval, as in thunder following lightning. It is generally accepted that the efficiency with which stimuli are processed benefits from them being embedded in temporal structures. The present paper investigates to which extent instructions that inform participants on specific features of the temporal structure, which represent a manipulation of top-down attention, contribute to these processing benefits. In the context of this manipulation, benefits for auditory processing will be compared between target stimuli temporally associated with a single cue, and targets embedded in periodic or aperiodic rhythms.

Temporal associations consist of a warning cue followed by a potential target and have typically been investigated using temporal foreperiod tasks (Los et al. 2001; Näätänen & Niemi, 1981) and temporal cueing tasks (Nobre, Correa & Coull, 2007; Nobre & van Ede, 2018). In constant foreperiod tasks, a warning signal and target stimulus are presented, between which the time interval remains constant within a block of trials and is varied between blocks. These tasks have shown that temporal prediction efficiency is optimal at shorter time intervals and diminishes with longer intervals between warning signal and cue, as evidenced by increased reaction times (Näätänen & Niemi, 1981) and lower accuracy (Rolke & Hoffman, 2007). In temporal interval cueing tasks that use two cueing intervals, a valid temporally predictive cue (compared to the invalid cue) considerably improves reaction times and accuracy for targets at the shorter of two possible intervals (Griffin, Miniussi & Nobre, 2002; Stefanics et al 2010; for a review see Nobre & van Ede, 2018). If two cueing intervals are used, participants are explicitly instructed on the cue-target relationships, and it can therefore be assumed that participants in these studies are aware of the time interval value following each cue, predicting the occurrence of the target stimulus. Similarly, we assume that if one constant cue interval is used within a block, participants are aware of its explicit timing for execution of the task. This may lead to top–down attention focused to a specific moment in time, thereby enhancing processing of stimuli presented around the expected time. The temporal predictability of targets afforded by a temporal cue preceding a target may also be implicit, which occurs when the temporal interval between cue and target on any given trial is drawn from a distribution. Under these conditions, it has been shown that the less predictable the cue-target relation becomes (i.e. the wider the distribution), the more listeners' reaction times are slowed (Herbst and Obleser, 2017; Herbst, Fiedler & Obleser 2018).

Rhythmicity is another form of temporal structure that can be leveraged by the brain. The properties of stimuli are processed more efficiently by virtue of them being preceded by a periodic rhythm. In line with this idea, seminal work by Jones and colleagues (Barnes & Jones, 2000; Jones et al., 2002; Large & Jones, 1999) has shown accuracy benefits for targets occurring in-phase compared to out-of-phase with a preceding rhythm. Perceptual benefits of periodic rhythms can be attributed to the entrainment of ongoing neural oscillations (Schroeder & Lakatos,2009; Henry & Herrmann 2014; ten Oever et al, 2017), where the high-excitability phase of neural populations become aligned to the timing of relevant events. Although the present study reports only behavioral data, entrainment theory provides a useful framework within which to consider benefits of periodic over aperiodic rhythms. When entrainment is present, stimuli presented periodically should entail perceptual benefits for targets occurring in phase with the stimulation, even after the stimulation has ended, which has been demonstrated by a large number of studies (Barnes & Jones, 2000; Sanabria, 2011; Sanabria & Correa 2013; Breska & Deouell 2014; Breska & Ivry, 2018; Ren et al., 2019; Trivino et al., 2011; Ellis et al, 2010). A processing advantage of periodicity has been shown even when the periodic structure was task irrelevant (Rimmele, Jolsvai & Sussman, 2011) or when the task required focusing on a different (predictive) stimulus feature (e.g. stimulus color) (Rohenkohl, Coull, Nobre, 2011; Breska & Deouell, 2014). These suggested automatic benefits of rhythms have proven advantageous in speeded response tasks for frontal lobe patients in whom temporal associative processing of a cue and a stimulus was impaired (Trivino et al., 2011). While these findings suggest that rhythmic structure is processed in a primarily bottom-up stimulus driven fashion, an alternative explanation for the perceptual benefits of periodicity is that it guides top-down attention (Jones et al., 2002). According to this hypothesis, (rhythmic) attention permits processing targets that are in phase with the stimulation more efficiently, explaining the advantage of periodic versus aperiodic stimulation. Thus, whether auditory neural activity automatically entrains to rhythmic input or whether its entrainment is under active top-down control remains a topic of debate (Haegens & Zion-Golumbic, 2018; Bouwer, 2022). Previous studies comparing a rhythm to a temporal cue have shown that the expectancy profile to a preceding rhythm is sharper than that produced by a single cue (Experiment 4, Barnes & Jones, 2000) and that target perception benefits more from preceding rhythms than from preceding single cues (Ren et al., 2019). Combining a rhythm and a temporal cue has been shown to have an additive effect in reducing reaction times (Ellis et al., 2010), suggesting perhaps a similarity in (part of) the underlying mechanisms (Rimmele et al., 2018).

Here, we aim to gain insight in factors influencing the benefits induced by periodicity, by manipulating awareness (through task instruction) of the position in the rhythm where targets may be presented. We opted for having a constant location within a rhythm for potential target presentation, and we ran one study in which participants were explicitly informed about the location in the sequence where the target might occur, and another study in which they were not informed. To the best of our knowledge, targets embedded within rhythms, and thus being fully predictable while listeners remain unaware of this, have not been investigated and compared to targets with more explicit form of predictability due to task instruction. We compared the perceptual benefits of periodic rhythms (with and without instruction) to benefits of aperiodic rhythms and those of temporal interval cueing. Using this experimental design, we aimed to answer two principal questions. (1) Does a periodic rhythm lead to improved behavioral performance relative to a temporal cue and an aperiodic rhythm? (2) Does instructing in which position of the periodic and aperiodic rhythm to expect the possible presentation of a target affect the behavioral benefit of a rhythm compared to a temporal cueing interval? Our main finding was that the expected superiority of periodic rhythms over single cueing only emerged when participants were explicitly informed about the constant location of the targets within the periodic stimulus sequence.

Materials &Methods

Participants

In experiment 1, we collected data from 24 participants of which 20 participants (8 male, 12 female) were included in the final analysis. Two participants were excluded due to failure to comprehend and/or execute the task and two participants due to an abnormal audiogram assessed prior to testing. In experiment 2, we collected data from 17 participants, 2 were excluded due to failure to comprehend and/or execute the task, resulting in a total of 15 participants (1 male, 14 female). Participants were screened for normal hearing at audiometric test frequencies from 0.25 - 8 kHz. The sample size of each experiment is comparable to our previous study (Heynckes et al., 2020). The Ethics Review Committee of the Faculty of Psychology and Neuroscience (ERCPN) at Maastricht University granted approval for all studies and all participants gave written informed consent.

Stimuli

Acoustic stimuli were created in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA) at a 44.1kHz sampling rate, 16-bit resolution.

Macro-temporal structure. Three types of acoustic stimuli with different temporal structure were created: periodic and aperiodic rhythms of 5.5 s length, and a temporal cue condition of 700 ms length (see Figure 1). All stimuli had an (average) interstimulus interval (ISI) of 500ms, corresponding to a 2 Hz rate at which sound quintets were repeated. To create the aperiodic rhythm with a total stimulus duration of 5.5 seconds, on each trial we randomly sampled ISI (in ms) from a vector (x = [110 130 210 230 280 445 740 700 965 1190], with a mean of 500 ms).

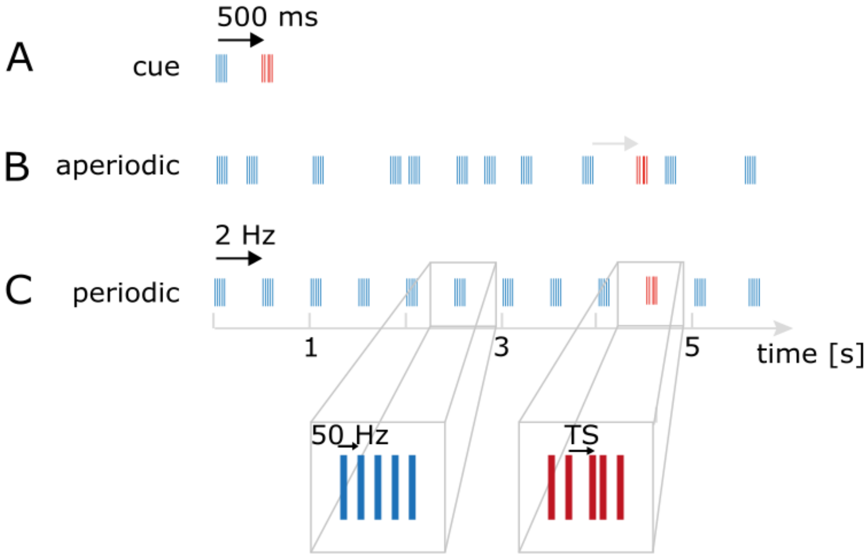


Figure 1 Temporal structure of stimuli. A. a temporal cue precedes the target (TS). **B**. Aperiodic rhythm with TS. Quintet TS-1 precedes at 2 Hz ISI. **C.** Periodic rhythm at 2 Hz with TS. Five sounds repeating at 50 Hz (10ms ISI) create a quintet (*inset 1*). Target stimuli (*inset 2*) had a different temporal structure: the third sound in a quintet was temporally shifted between 1.5 -7 ms, depending on a subject's perceptual threshold.

Quintets (Micro-temporal structure). The quintets consisted of five narrowband sounds of 10ms duration that were presented at a frequency of 50 Hz (i.e. an ISI of 10 ms). The narrow passband was centered on a carrier frequency of 1100 Hz (ERBS = 4; Moore, 2003), consisting of a summation of 21 sinusoids, each with a random phase onset.

Targets in the quintets. In a subset of quintets, the 3rd of five narrowband sounds was shifted in time. Participants had to report the presence of the temporal shift (TS) whenever it occurred in a quintet. Per participant the size of this shift was determined by a staircase procedure, converging on a 70% detection threshold (2 down 1 up) and ranged between 1.5 and 7ms in ten logarithmically spaced steps (Levitt, 1971). The termination criterion was after 200 trials or 15 reversals and the perceptual threshold was computed as the mean over the last 12 reversals. Supplementary figure S2 shows the staircase performance over the course of the experiments across participants.

The TS distribution over the quintets was non-uniform. In the temporal cueing condition, a TS, if present, occurred at the second (cued) quintet. The TS in the periodic and aperiodic conditions (if present) occurred at the 10th quintet (i.e. after ~4.5 s). Additionally, in the aperiodic condition the interval between the target quintet and the preceding quintet was kept constant at 500ms.

The TS occurred in 75% of the trials. Detecting a TS was a difficult perceptual task, to the extent that in the first experiment most participants remained unaware (as assessed by a post-experiment questionnaire) of the constant position of the target within periodic and aperiodic stimuli. In particular, 8 participants reported noticing targets occurring more frequently towards the end of a stimulus, while 2 participants reported noticing that the target was always at the same position, yet were unable to count and name the specific quintet position. In the second experiment, participants received an instruction when to expect a target within the periodic and aperiodic rhythm. They were verbally informed to expect targets more frequently at the 10th position and that counting may be beneficial in executing the task.

Design

All stimulus presentation scripts were written in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA), using the Psychophysics toolbox (Brainard, 1997) and are publicly available at (Heynckes, 2022a). Participants sat in front of a computer placed within a sound-attenuated chamber while listening to the stimuli delivered through Sennheiser HD650 headphones. Instructions were presented on the computer monitor and responses collected with a standard keyboard. Participants were asked to detect a target; a temporal shift (TS) of a narrow-band sound. During a trial, and up to 1s after a sound finished, participants could indicate whether they perceived a target. After a variable interval the next trial was initiated. Participants performed a total of 432 trials, divided over 6 blocks. Each block contained three mini-blocks in which the temporal structure was constant, with a mini-block consisting of 24 trials of periodic stimuli, followed by a mini-block of 24 trials aperiodic stimuli and a mini-block of 24 cue stimuli. The order of mini-blocks within each block was fully counterbalanced within each participant and randomized across participants.

Statistical analysis

All analyses were conducted in MATLAB 2017a (The MATHWORKS Inc., Natick, MA, 234 USA). The data and analysis code can be found at (Heynckes, 2022b). For each participant, we computed sensitivity (d prime) as well as mean log- reaction time (logRT) for correct trials. Reaction time was calculated relative to TS target onset. Statistical group analyses were carried out using generalized linear mixed models (GLMM), which were assessed using likelihood ratio tests. The full notation of the models is found in the supplementary material. Post hoc tests were performed by specifying contrasts and corrected for multiple comparisons (where applicable) using Bonferroni correction. D prime scores and model parameters were estimated in one step, by linearly modeling the behavioral outcomes of the detection task (yes & no) with a predictor target coding for the presence or absence of the target and employing an inverse Gaussian link (probit) function, as previously described in more detail in Heynckes et al. (2020). The different experimental conditions were then added as categorical predictors, and the estimated d’ for each of these conditions was described by the interaction term between target and the respective condition predictor (see Knoblauch & Maloney 2012, chapter 3.3.5). Unstandardized effect sizes (betas) are reported in units of the dependent variable (d ‘or logRT), allowing for a meaningful comparison, in line with general recommendations on how to report effect sizes in psychological research (Pek et al, 2018). For our visualization, we estimated d’ as it is traditionally computed. Standard errors of d’ group effects displayed in figure 2 were obtained by non-parametric bootstrap sampling of estimated d’ values, carried out at the subject level (N = 1000). The mean was used as a measure of central tendency around which 95% confidence intervals were created.

Results

Instruction improves sensitivity of auditory target perception in periodic rhythm compared to a temporal cue.

We fitted a GLMM to the d' sensitivity data for target detection. This analysis showed a significant three-way interaction of target, instruction and temporal structure on d' (F (2,198) = 48,55, p < 0.001). Notably, the d' advantage afforded by the temporal cue and by the periodic rhythm over the aperiodic rhythm was larger in the experiment with instruction (Fig. 2A) than in the experiment without instruction (Fig. 2B) as assessed in a set of post-hoc contrasts. The difference between periodic and aperiodic rhythm (given the target predictor) differed between the experiments (t (1, 198) = 49.84, p< 0.001). Likewise, the difference between the temporal cue and aperiodic rhythm (given the target predictor) differed between the experiments (t (1, 198) = 38.43, p< 0.0001), as did the difference between the temporal cue and periodic rhythm (given the target predictor) (t (1, 198) = 7.01, p< 0.01). Because of the significant three-way interaction of target, instruction and temporal structure, we analyzed the interaction of target and temporal structure separately for the two experiments. In both the instructed and the uninstructed experiment, d' varied significantly as a function of the temporal structure (Fig2A, *no Instruction*; F (2, 114) = 25.469, p< 0.001; Fig2B, *with Instruction*; F (2, 84) = 170.66, p< 0.001). Follow-up comparisons of the temporal conditions in the experiment without instruction showed a d' benefit for periodic rhythms over aperiodic rhythms (Fig2A, beta = -0.3663; t (1, 114) = 20.98, p< 0.001). Similarly, a d' benefit was present of periodic over aperiodic rhythms in the experiment with instruction (Fig2B, beta = -1.55; t (1, 84 = 241.14, p< 0.001). A d' benefit was also observed for the temporal cue over the aperiodic rhythm both without instruction (Fig2A, *beta* = -0.5646; t (1, 114) = 48.87, p< 0.001) and with instruction (Fig2B, *beta* = -1.42; t (1, 84) = 228.05, p< 0.001). These findings replicate our previously reported results (Heynckes et al., 2020).

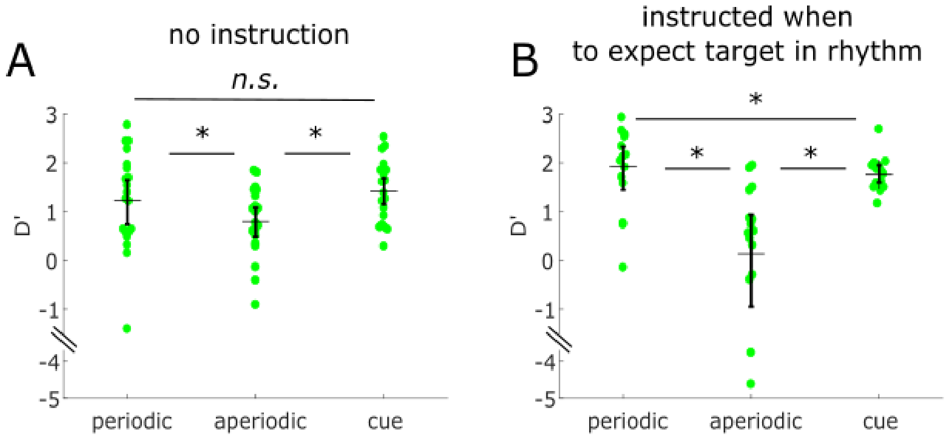


Figure 2 Instruction when to expect a target improves detection in periodic rhythm compared to a temporal cue. In experiment 1 **(A)**, participants received no instruction on possible target positions within the temporal sequences and remained unaware that targets occurred always at the same position. In this case sensitivity does not differ between a periodic rhythm and a temporal cue. **(B)** When being instructed on the temporal contingency of target occurrence in periodic and aperiodic stimuli (participants were instructed in which position to expect the possible presentation of a target), a higher sensitivity for the periodic rhythm compared to temporal cue (as well as aperiodic) occurs. Filled circles depict d' per participant. Black horizontal line shows the group mean. Errorbars depict bootstrapped confidence intervals (at subject-level). Note the discontinued y-axis for visualization purposes.

Crucially, whether or not participants received an instruction on the position of target occurrence within periodic rhythms compared to a temporal cue had a significant effect on their detectability. When the temporal location of potential target presentations within a periodic rhythm was not made explicit by instruction, sensitivity due to the periodic rhythm did not differ compared to that afforded by a single temporal cue (Fig2A, *no Instruction*; beta = 0.217; t (1, 114) = 1.37; p>0.5). In contrast, when target predictability was made explicit by instructing participants, sensitivity in the periodic condition was significantly higher compared to a temporal cue (Fig2B, *with Instruction*; beta = 2.7464; t (1, 84) = 173.78; p<0.001).

*Limited effect of instruction on processing speed of targets in periodic and aperiodic r*hythms.

We fitted a GLMM to the response times obtained in the two experiments. The analysis of reaction times showed a non-significant interaction of instruction and temporal condition (F (2,99) = 0.99; p>0.05). Providing an instruction, did not have a significant effect on reaction times, compared to when no instruction was provided (F (1,99) = 2.72; p=0.1). The temporal condition had a significant effect on reaction times across the two experiments (F (2,99) = 24.2; p<0.001). Follow-up tests showed that participants responded faster to targets following a single temporal cue, then to targets in periodic rhythms, (F (1, 99) = 15.44, p<0.001) as well as compared to aperiodic rhythms, (F (1, 99) = 81.56, p<0.001). Moreover, comparisons revealed faster correct responses for periodic than aperiodic rhythms F (1, 99) = 43.83, p<0.001).

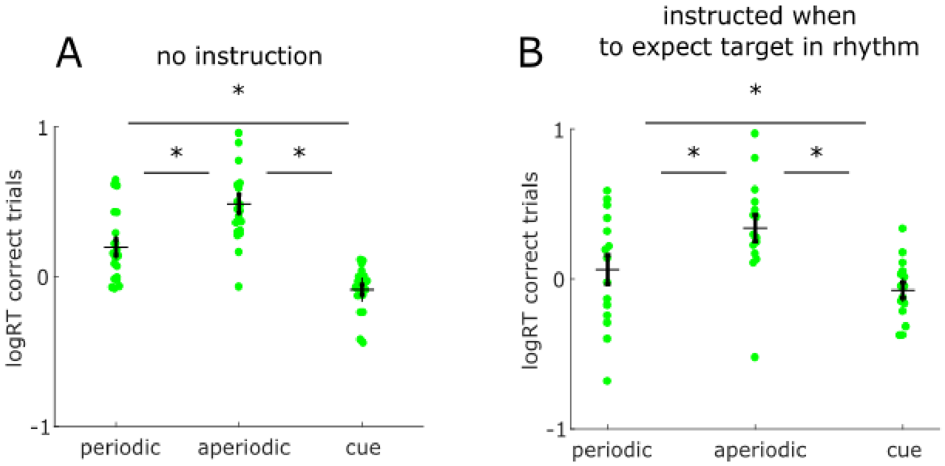


Figure 3 Log reaction times, are fastest following the temporal cue, independent of instruction. (**A**) Participants had received no instruction on target position, but an implicit temporal prediction of the target position within the temporal rhythms and remained unaware that targets occurred at the same position. (**B**) In experiment 2 participants were instructed in which position to expect the possible presentation of a target in the periodic and aperiodic rhythm. Whether or not an instruction was provided did not significantly interact with effects of temporal structure on logRT. Main effect of temporal structure was therefore analyzed across experiments. A temporal cue lead to fastest logRT, followed by the periodic and aperiodic conditions. Filled circles show single participant mean logRT on correct trials. Black horizontal line shows the group mean. Errorbars depict SEM.

Discussion

Temporal expectations facilitate sensory processing and perception, and they play an important role in perceiving the outside world (Nobre & van Ede, 2018). Currently, little is known about whether shared or separate mechanisms contribute to temporal expectations based on different temporal structures of sounds. Specifically, it is unclear if periodic rhythms and temporal interval cues affect behavioral performance in similar ways, relative to aperiodic rhythms. Moreover, to what extent these effects rely on task instruction and attention is also still unclear.

In the present study, we conducted two experiments in which a difficult temporal detection task was used to study the effect of three temporal structures (periodic, aperiodic and temporal cueing) on reaction times and sensitivity (d’). In the experiments, we either gave no instruction about when to expect a possible target in any of the three conditions (experiment 1) or provided an instruction on when to expect a target within a periodic or an aperiodic rhythm (experiment 2). In both experiments and in all conditions, targets (if they occurred) always were presented at the same temporal position within periodic and aperiodic rhythms, or after the same time interval if following a single temporal cue. With this design we set out to investigate (1) whether a periodic rhythm led to improved sensitivity relative to a temporal cue and an aperiodic rhythm and (2) whether instructing in which position of the periodic and aperiodic rhythm to expect the possible presentation of a target affected the behavioral benefit of a periodic rhythm compared to a temporal cueing interval and an aperiodic rhythm.

With respect to perceptual sensitivity, we showed in both experiments that periodic rhythms led to improved sensitivity compared to aperiodic rhythms, and that the size of this effect was increased when instructions informing participants on specific features of the temporal structure (representing a manipulation of top-down attention) were provided. This finding is in line with the frequently reported behavioral benefit of periodic stimulation. When comparing the periodic rhythm and temporal cue, we showed that both a rhythm and a temporal cue led to a similar sensitivity when participants were not receiving any explicit instruction of when to expect a target (experiment 1), replicating our previous findings (Heynckes et al., 2020). However, providing an instruction led to improved behavioral detection performance in the periodic rhythm compared to the temporal cue (experiment 2).

The entrainment of ongoing oscillations can be seen as a framework for bottom up benefits of rhythmicity (Haegens & Zion-Golumbic ,2018). Here we show that attention can boost expectations elicited by periodic stimulation leading to enhanced sensitivity in target detection. In the absence of an instruction, the rhythm provided no behavioral advantage in sensitivity compared to a temporal cue. This argues against the idea that there are always automatic benefits of rhythmic stimulus presentations. Therefore, our results may suggest that, in the absence of instruction, entrainment of oscillations alone would not provide a vehicle for the generation of predictions that are more efficient than the predictions provided by a single cue of the temporal interval after which a target follows. An additional possible explanation of the findings in the periodic stimulation condition in the absence of instruction could be that the longer stimulus duration of a periodic rhythm (compared to a temporal cue) may have had consequences for maintaining top-down attention. This may have interacted negatively with the bottom-up benefits afforded by the periodic stimulation. Specifically, participants may have evaluated each stimulus in the rhythm, thereby dividing (or exhausting) attentional resources. When we instructed participants about the location within periodic and aperiodic rhythms where targets could occur, this led to significantly higher sensitivity to targets in the periodic condition than in the temporal cue condition. We suggest that explicit instruction allowed participants to intentionally use the rhythm, guiding attention to a specific moment in time (Junker, Park, Shin & Cho, 2020). The focusing of attention to a precise point in time in the periodic sequence may have been aided by the accumulated evidence of the precise length of inter-stimulus intervals, providing an advantage over both aperiodic and interval cue conditions. Thus, the temporal narrowing of the window of attention may have facilitated behavioral sensitivity in addition to the benefit of in-phase tone presentation in the periodic condition. This suggestion is corroborated by previous findings showing behavioral beneficial effects on reaction times to targets through the intentional use of a rhythm (Breska and Deouell, 2014), and by the previously reported link between attention and changes in sensory gain associated with entrainment (Lakatos et al., 2008; Henry & Herrmann 2014). In addition, proposals that temporal predictions (instantiated through a rhythm or a temporal cue) modulate ongoing oscillations through a top-down reset (Rimmele et al., 2018) are compatible with this view. In the same vein, by inserting cues within ongoing rhythms, advantages of attentional allocation for the detection of targets embedded in a rhythms have also been demonstrated, crucially with a concomitant delta-phase reset, in line with entrainment theories (Stefanics et al.; 2010; Herbst et al., 2022).

In contrast, reaction times in the present study were not affected by the presence or absence of instruction about the temporal position of the target, and hence appeared unaffected by attention. In both experiments, reaction times were fastest for the temporal cue condition followed by the periodic and aperiodic conditions, similar to previous research (Bouwer et al., 2019). Although participants were instructed to be fast and accurate, the difficulty of the perceptual task may have led participants to prioritize accuracy, so that explicit information about temporal target position in the experiment with instruction may have provided benefits exclusively for accuracy and not for processing time.

A number of studies have behaviorally compared the benefits of single cues, periodic and aperiodic rhythms for the accuracy of auditory target perception. Several studies have assessed the benefits of periodicity for the sensitivity of target perception by comparing performance for a periodic rhythm relative to an aperiodic rhythm (Jones et al, 2002; Jones et al., 2006; Rohenkohl et al.; 2011; Lawrance et al., 2014; ten Oever et al., 2014; ten Oever et al. 2017; Morillon et al., 2016; Schmidt-Kassow et al., 2009). These studies showed facilitated target perception for the periodic relative to the aperiodic conditions, in line with our data.

A number of other studies (Lin et al. 2021; Breska & Deouell, 2014; Ren et al., 2019; in control participants: Breska & Ivry, 2018; Breska & Ivry, 2021) have compared the perceptual benefits of temporal interval cues with periodic rhythm conditions for targets placed after the termination of the periodic rhythm. Overall, these studies have shown only limited evidence that periodicity may lead to a better behavioral performance than interval cueing. Among those five studies, only one study (Ren et al., 2019) reported a benefit in reaction times following the rhythmic cue compared to a temporal interval, whereas the other studies did not report any benefit of rhythms over temporal interval cues, neither in sensitivity nor in reaction times. Although the five above-cited studies in one aspect of their design resemble our non-instructed experiment (in the sense that no explicit instructions were given in these studies as to when to expect the target), their strategy to present the target stimuli after the termination of a periodic rhythm limits the extent to which their results can be compared with ours. Therefore, it is not certain that applying an instruction to the design of the above-referred studies to direct attention towards the temporal location of the target, would have increased the benefits of periodic over single temporal cue conditions for target perception. Instead, we speculate that the difficulty of the target detection task in our experiments, partially due to inserting the target within the rhythm (as done in Bouwer et al. 2019; Heynckes et al., 2020; Stefanics et al.; 2010; Herbst et al., 2022), may have in the present study created the conditions in which an explicit instruction indicating the temporal location where targets occurred in the periodic rhythm led to sensitivity advantages over single temporal cueing.

To conclude, our results shed light on the possibility that target stimuli embedded in a periodic rhythm benefit from automatically generated predictions providing enhanced target processing. We evaluated this possibility by comparing target processing efficiency in periodic rhythms to that afforded by temporal cues and aperiodic rhythms. In our experimental design, we found no evidence for automatic benefits of periodic rhythms over single temporal cues for target processing. Instead, better target processing in periodic rhythms than after temporal cues only emerged after an explicit task instruction pointing to the precise moment in the rhythm when a target was to be expected. Hence, here, the benefits of periodic rhythms for perception only occurred in interaction with attention. We suggest that our data increase insight into the conditions in which periodic rhythms can provide perceptual advantages to targets embedded in these rhythms. The hypothesized explanatory mechanisms, in which we suggest interactions between oscillatory entrainment and top-down attentional effects should be tested by neurophysiological or electrophysiological experiments (ten Oever et al., 2022).

Supplementary materials

Table S1. Wilkinson notation of final models

Large Model (both experiments) Sensitivity	d’ ~ Criterion + Target*Temporal Structure*Experiment + (1 |Subject) -1
Large Model (both experiments) Reaction times	logRT ~ Temporal Structure*Experiment + (1 |Subject)
No Instruction Exp Sensitivity	d’ ~ Criterion + Target*Temporal Structure + (1 |Subject) -1
With Instruction Exp Sensitivity	d’ ~ Criterion + Target*Temporal Structure + (1 |Subject) -1

Note. Criterion is an additional predictor reflecting the intercept (normally notated as 1, here re-parameterized to -1 to reduce correlation between fixed effects (see p 262, Knoblauch & Maloney, 2012))

Acknowledgements

We thank Anna Yi and Lauren Huizenga for help with the data collection. This work was funded by The Netherlands Organization for Scientific Research (NWO) Research Talent Grant 406.17.200 awarded to MH, EF and PDW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. Cognitive Psychology, 41(3), 254–311. https://doi.org/10.1006/cogp.2000.0738

Bouwer FL, Honing H, Slagter HA. Beat-based and memory-based temporal expectations in rhythm: Similar perceptual effects, different underlying mechanisms. J Cogn Neurosci. 2019;32(7):1221–41. https://doi.org/10.1162/jocn_a_01529

Bouwer, F. (preprint retrieved 2022, January 25). Neural entrainment to auditory rhythms: automatic or top-down driven? https://doi.org/10.31234/osf.io/p9uza

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433–436. https://doi.org/10.1163/156856897X00357

Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. Journal of Cognitive Neuroscience, 26(7), 1555–1571.https://doi.org/10.1162/jocn_a_00564

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson’s disease. Proceedings of the National Academy of Sciences of the United States of America, 115(48), 12283–12288. https://doi.org/10.1073/pnas.1810596115

Breska, A., & Ivry, R. B. (2021). The human cerebellum is essential for modulating perceptual sensitivity based on temporal expectations. Elife, 10, e66743. https://doi.org/10.7554/eLife.66743

Ellis, R. J., & Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. Attention, Perception, & Psychophysics, 72(8), 2274-2288.

Griffin, I. C., Miniussi, C., & Nobre, A. C. (2001). Orienting attention in time. Frontiers in Bioscience. https://doi.org/10.2741/griffin

Haegens, S., & Golumbic, E. Z. (2018). Rhythmic facilitation of sensory processing: A critical review. Neuroscience & Biobehavioral Reviews, 86, 150-165.

Henry, M. J., & Herrmann, B. (2014). Low-Frequency Neural Oscillations Support Dynamic Attending in Temporal Context. Timing and Time Perception, 2(1), 62–86. https://doi.org/10.1163/22134468-00002011

Herbst, S. K., & Obleser, J. (2019). Implicit temporal predictability enhances pitch discrimination sensitivity and biases the phase of delta oscillations in auditory cortex. NeuroImage, 203(August), 1–33. https://doi.org/10.1016/j.neuroimage.2019.116198

Herbst, S. K., Fiedler, L., & Obleser, J. (2018). Tracking temporal hazard in the human electroencephalogram using a forward encoding model. eneuro, 5(2).

Herbst, S. K., Stefanics, G., & Obleser, J. (2022). Endogenous modulation of delta phase by expectation–A replication of Stefanics et al., 2010. Cortex, 149, 226-245.

Heynckes, Miriam. (2022a). Presentation scripts of "Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues" (Version 1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6472448

Heynckes, M. (2022b). Data and analysis scripts of "Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues" (Version 1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.6473030

Heynckes, M., De Weerd, P., Valente, G., Formisano, E., & De Martino, F. (2020). Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences. PLoS ONE, 15(6), 1–17. https://doi.org/10.1371/journal.pone.0234251

Jones, M. R., Moynihan, H., Mackenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. Psychological Science, 13(4), 313–319. https://doi.org/10.1111/1467-9280.00458

Junker, M. S., Park, B. Y., Shin, J. C., & Cho, Y. S. (2020). Adaptive Changes in the Dynamics of Visual Attention With Extended Practice. Frontiers in Psychology, 2633.

Knoblauch, K., & Maloney, L. T. (2012). Modeling Psychophysical Data in R. Springer New York. https://doi.org/10.1007/978-0-387-78171-6

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. science, 320(5872), 110-113.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. Psychological Review, 106(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Lawrance, E. L. A., Harper, N. S., Cooke, J. E., & Schnupp, J. W. H. (2014). Temporal predictability enhances auditory detection. The Journal of the Acoustical Society of America, 135(6), EL357–EL363. https://doi.org/10.1121/1.4879667

Lawrance, E. L., Harper, N. S., Cooke, J. E., & Schnupp, J. W. (2014). Temporal predictability enhances auditory detection. The Journal of the Acoustical Society of America, 135(6), EL357-EL363.

Levitt, H. C. C. H. (1971). Transformed up-down methods in psychoacoustics. The Journal of the Acoustical Society of America, 49(2B), 467-477.

Lin WM, Oetringer DA, Bakker-Marshall I, Emmerzaal J, Wilsch A, ElShafei HA, et al. No behavioural evidence for rhythmic facilitation of perceptual discrimination. Eur J Neurosci. 2021;(March):1–13. https://doi.org/10.1111/ejn.15208

Los, S. A., Knol, D. L., & Boers, R. M. (2001). The foreperiod effect revisited: Conditioning as a basis for nonspecific preparation. Acta Psychologica, 106(1–2), 121–145. https://doi.org/10.1016/S0001-6918(00)00029-9

Moore, B. C. (2003). An introduction to the psychology of hearing. Brill.

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal Prediction in lieu of Periodic Stimulation. Journal of Neuroscience, 36(8), 2342–2347. https://doi.org/10.1523/JNEUROSCI.0836-15.2016

Näätänen, R., & Niemi, P. (1981). Foreperiod and simple reaction time. Psychological Bulletin, 89(1), 133–162.

Nobre, A. C., & Van Ede, F. (2018). Anticipated moments: Temporal structure in attention. Nature Reviews Neuroscience, 19(1), 34–48. https://doi.org/10.1038/nrn.2017.141

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. Current opinion in neurobiology, 17(4), 465-470. https://doi.org/10.1016/j.conb.2007.07.006

Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A discussion and tutorial. Psychological Methods, 23(2), 208–225. https://doi.org/10.1037/met0000126

Ren, Y., Xu, Z., Wu, F., Ejima, Y., Yang, J., Takahashi, S., ... & Wu, J. (2019). Does temporal expectation driven by rhythmic cues differ from that driven by symbolic cues across the millisecond and second range?. Perception, 48(6), 515-529.; https://doi.org/10.1177%2F0301006619847579

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. Trends in Cognitive Sciences, 22(10), 870–882. https://doi.org/10.1016/j.tics.2018.08.003

Rimmele, J., Jolsvai, H., & Sussman, E. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. Journal of cognitive neuroscience, 23(5), 1136-1147.

Rohenkohl G, Nobre AC. Alpha oscillations related to anticipatory attention follow temporal expectations. J Neurosci. 2011;31(40):14076–84. https://doi.org/10.1523/JNEUROSCI.3387-11.2011

Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. PLoS One, 6(1), e14620.

Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. Psychonomic Bulletin and Review, 14(3), 522–526. https://doi.org/10.3758/BF03194101

Sanabria, D., & Correa, Á. (2013). Electrophysiological evidence of temporal preparation driven by rhythms in audition. Biological Psychology, 92(2), 98–105. https://doi.org/10.1016/j.biopsycho.2012.11.012

Sanabria, D., Capizzi, M., & Correa, Á. (2011). Rhythms that speed you up. Journal of Experimental Psychology: Human Perception and Performance, 37(1), 236.

Schmidt-Kassow, M., Schubotz, R. I., & Kotz, S. A. (2009). Attention and entrainment: P3b varies as a function of temporal predictability. Neuroreport, 20(1), 31-36.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends in Neurosciences, 32(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Stefanics, G., Hangya, B., Hernadi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase Entrainment of Human Delta Oscillations Can Mediate the Effects of Expectation on Reaction Speed. Journal of Neuroscience, 30(41), 13578–13585. https://doi.org/10.1523/JNEUROSCI.0703-10.2010

ten Oever S, Schroeder CE, Poeppel D, van Atteveldt N, Mehta AD, Mégevand P, et al. Low-Frequency Cortical Oscillations Entrain to Subthreshold Rhythmic Auditory Stimuli. J Neurosci. 2017;37(19):4903–12. http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3658-16.2017 ten Oever S, Schroeder CE, Poeppel D, van Atteveldt N, Zion-Golumbic E. Rhythmicity and cross-modal temporal cues facilitate detection. Neuropsychologia. 2014;63(1):43–50. http://dx.doi.org/10.1016/j.neuropsychologia.2014.08.008

ten Oever, S., van der Werf, O. J., Schuhmann, T., & Sack, A. T. (2022). Absence of behavioral rhythms: noise or unexplained neuronal mechanisms? (response to Fiebelkorn, 2021). European Journal of Neuroscience.

Trivino, M., Arnedo, M., Lupiánez, J., Chirivella, J., & Correa, Á. (2011). Rhythms can overcome temporal orienting deficit after right frontal damage. Neuropsychologia, 49(14), 3917-3930. https://doi.org/10.1016/j.neuropsychologia.2011.10.009

4

Layer-specific correlates of detected and undetected auditory targets during attention

Corresponding Manuscript:

Heynckes, M., De Weerd, P., Formisano, E. & De Martino, F. (*in preparation*). Layer-specific correlates of detected and undetected auditory targets during attention

Abstract

In everyday life, the processing of acoustic information allows us to react to subtle changes in the soundscape. Yet even when closely attending to sounds in the context of a task, we occasionally miss task-relevant features. Here, we investigated the neural representations of attended target stimuli that were either correctly detected or missed. The computations that underlie our ability to detect behavioral relevant sound changes are thought to be embedded in both feedforward and feedback processes within the auditory hierarchy. In the present study, we exploited the high spatial resolution of high-field (7T) functional magnetic resonance imaging (fMRI) in humans to assess the role of feedforward and feedback contributions in primary and non-primary auditory areas during behavioral detection of target sounds. We demonstrated that the successful detection of subtle temporal shifts in target sounds leads to a selective increase of activation in superficial layers of primary auditory cortex (PAC) and qualitatively stronger responses in middle layers of planum polare (PP) compared to undetected targets. These results are suggestive of feedback signals reaching as far back as PAC and possibly propagating from PAC to secondary areas.

Introduction

When thinking of someone cracking a safe, the image of trying to open a safe by carefully listening to the mechanisms of the slot comes to mind, as it is depicted in some movies. The underlying principle being that some safe mechanisms are based on wheels, whose notches need to line up for a locking bar to slide and open them. When closely paying attention, mechanical imperfections of the wheels touching the bar within the locking mechanism can be heard and used to open the lock without knowing the combination. The auditory system possesses a remarkable ability to process information, often subtle, on which we base our decisions. We know that focusing one's attention on specific aspects of a sound can help in better detecting subtle changes in the soundscape. Yet, despite an overall benefit when attending, we experience lapses in our detection. Given the known segregation of feedforward and feedback processes within the laminar organization of the cortex (Douglas, Martin & Whitteridge, 1989; Douglas & Martin, 2004), we hypothesize that feedback mediated attentional processing is involved in the successful detection of sounds.

Electrophysiological research in animals has investigated the neural correlates of attention and highlighted changes in cortical oscillations in superficial layers (Lakatos, Musacchia, O'Connell, Falchier, Javitt, & Schroeder, 2013; O'Connell, Barczak, Schroeder & Lakatos, 2014). In humans, the modulation of cortical layers by attention in both vision and audition has been probed non-invasively using high-field functional magnetic resonance imaging (fMRI) (Lawrence, Norris, & De Lange, 2019; De Martino, Moerel, Ugurbil, Goebel, Yacoub, Formisano, 2015; Gau, Bazin, Trampel, Turner & Noppeney, 2016; Klein, Fracasso, van Dijk, Paffen, Te Pas & Dumoulin, 2018; Liu et al., 2021). In these studies, attentional modulation was probed by either drawing attention towards or away from the relevant stimulus (or stimulus feature). In particular, in the auditory domain, attending to an auditory stimulus (compared to a concurrently presented visual stimulus) has highlighted changes in frequency tuning (i.e. tuning width) (De Martino et al., 2015) and an increase in activation in superficial layers of (primary) auditory cortex (Gau et al., 2016). In the visual domain, within-modality attentional manipulations (spatial or feature based attention) have been used in layer-specific studies, which demonstrated activity modulations in superficial layers (De Lange et al., 2019; Liu et al. 2021) as well as changes in population receptive fields in deep layers (Klein et al., 2018) of primary visual cortex (V1). Altogether, these data indicate that the presence or absence of attention to stimuli modulates activity in superficial layers. Here we asked where in the auditory cortex and in which cortical layers, neural activity variations would be present that could explain why identical auditory stimuli presented under identical attentional instructions would be detected in some trials, and not in others. In line with literature ascribing a role of superficial layers in receiving attentional feedback, we hypothesized that small variations of activity in superficial layers of auditory cortex may be related to variations in perception of physically identical stimuli.

Apart from the segregation of feedforward and feedback signals across cortical depths, the auditory cortex has a tonotopic organization (Merzenich & Brugge, 1973; Merzenich, Knight, & Roth, 1973; Formisano, Kim, Di Salle, Van de Moortele, Ugurbil, & Goebel, 2003). Attention to frequency specific targets gain-modulates frequency-specific (tonotopic) regions in a layer dependent manner (O’Connell et al., 2014). Task-dependent changes in the receptive fields have been shown using invasive electrophysiology in animals in superficial layers of the auditory cortex (Francis, Elgueda, Englitz, Fritz, & Shamma ,2018) and have been suggested as neural correlates of selective attention. Similarly, frequency-specific effects of attention have been shown non-invasively in humans at a macroscopic level (Riecke, Peters, Valente, Poser, Kemper, Formisano & Sorger, 2018; Da Costa, van der Zwaag, Miller, Clarke & Saenz, 2013; De Martino et al., 2015). We presented narrowband stimuli at two distinct frequencies (high and low), to understand whether the topographic organization of human auditory cortical areas interacts with laminar processing when detecting a relevant sound.

In particular, we asked human listeners to perform an auditory temporal detection task while concurrently acquiring layer-specific fMRI data. By comparing responses to (acoustically identical) perceptually detected and undetected targets, we localized responses related to the detection of sounds under constant, demanding attentional conditions and hypothesized that behavioral relevance of attention is reflected in the change of population level activity in superficial cortical layers of the primary auditory cortex (Fig. 1C) and further hypothesized that this effect may be tonotopic.

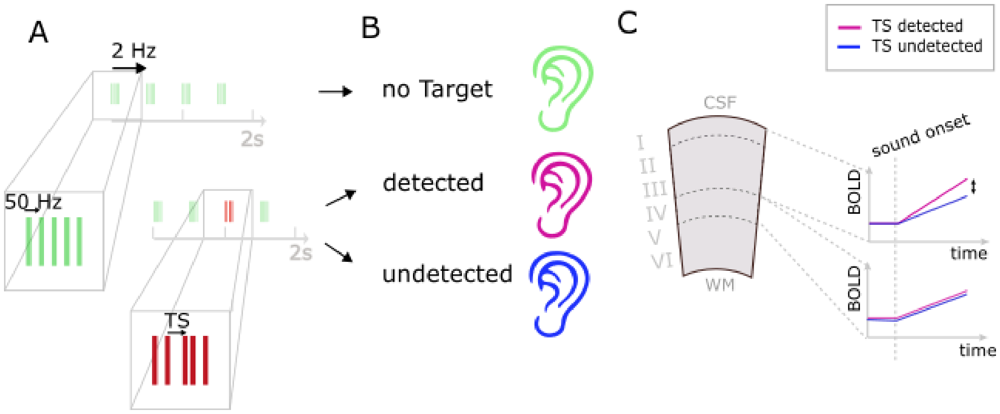


Figure 1. Task and Hypotheses. (A) Stimuli were periodic sequences of narrowband quintets repeating at 2 Hz. Two narrowband frequency ranges around 200 Hz and 1100 Hz were used to create low and high pitch sounds. Five sounds repeating at 50 Hz (10ms ISI) formed a quintet (inset 1). 75% of the stimuli contained a target. Target sounds (TS; inset 2) had a different temporal structure: the third sound in a quintet was temporally shifted between 1.5 -7 ms, depending on subject’s perceptual threshold. Figure S1.1 shows behavioral detection rates per subject. **(B)** Target trials were sorted based on percept. **(C)** Expected laminar response profile. Both detected and undetected TS would entail a feedforward increase in middle layers, but a detected TS (magenta line) additionally increases the BOLD response in superficial layers compared to undetected TS (blue line).

Results

We examined the laminar response profile of human auditory cortex (AC), using 2-D gradient echo (GE) blood oxygen level dependent (BOLD) fMRI at 7T, with sub-millimeter resolution, during perceptual detection of temporally shifted target sounds (TS) embedded in rhythmic sound sequences (Fig. 1A). Specifically, we contrasted different percepts of acoustically identical sound sequences containing a target (Fig. 1B).

Based on subjects' responses, we labeled trials as detected (i.e., target present and detected), undetected (target present and not detected) and no Target (target not present). For our first-level analysis, we fitted the fMRI responses using a general linear model (GLM) with separate predictors for every trial and condition. Predictors were convolved with a standard hemodynamic response function (HRF). Our analysis focused on multiple regions of interest (ROIs), consisting of primary and non-primary areas of human auditory cortex (Fig. 2B-C) which we delineated following macro-anatomical landmarks (Kim et al., 2000). Laminar profiles were obtained by sampling the response estimates (beta weights) of each trial onto 11 equivolume depth surfaces (Waehnert et al., 2014) (Fig. 2D). Responses were then averaged across trials of the same perceptual condition per depth level (Fig. 2G). Multiple inclusion criteria guided the selection of vertices that were used to obtain average layer profiles (per condition and ROI - see Fig. S3.1). First, vertices had to be within a particular ROI (primary auditory cortex [PAC], Heschl's Gyrus [HG], planum polare [PP], planum temporale [PT]). Second, the linear model had to explain a certain amount of variance in the measured signal (F-test) in an independent tonotopic localizer (F>2) and in the in the main experiment exceeded (F>0.1), thereby ensuring that all vertices with a positive (average across conditions) BOLD response to sounds in the main experiment were included, independent of depth. Figure S3.2 shows the laminar profile per subject ROI and condition.

Second-level group statistics (n=10) were carried out on the differences of the mean (across subjects) betas between the perceptual conditions across cortical depths (Fig. 3D-F). Specifically, we statistically tested whether detected and undetected sounds differentially modulated the response compared to the no Target condition (i.e. we compared the difference of differences, [detected vs. no Target] vs. [undetected vs. no Target]). By subtracting the response to the no target condition from both the responses to detected and undetected sounds we control for the layer dependent signal increase towards the superficial gray matter elicited by the draining vasculature (Turner, 2002; Polimeni et al., 2010). To assess the frequency-specificity of detection effect we fitted a generalized linear mixed effects (GLME) model with four predictors (*depth* [linear], *condition* [detected minus no Target; undetected minus no Target], *BFandSound* [highSoundHighBF, highSoundLowBF, lowSoundHighBF, lowSoundlowBF] and their *interactions*). Upon determining that detection was non-frequency-specific we collapsed data across frequency-preferring voxel populations and sounds. We then used three predictors (depth [linear], condition [detected minus no Target; undetected minus no Target] and their interaction) in a separate generalized linear mixed effects (GLME) model per ROI.

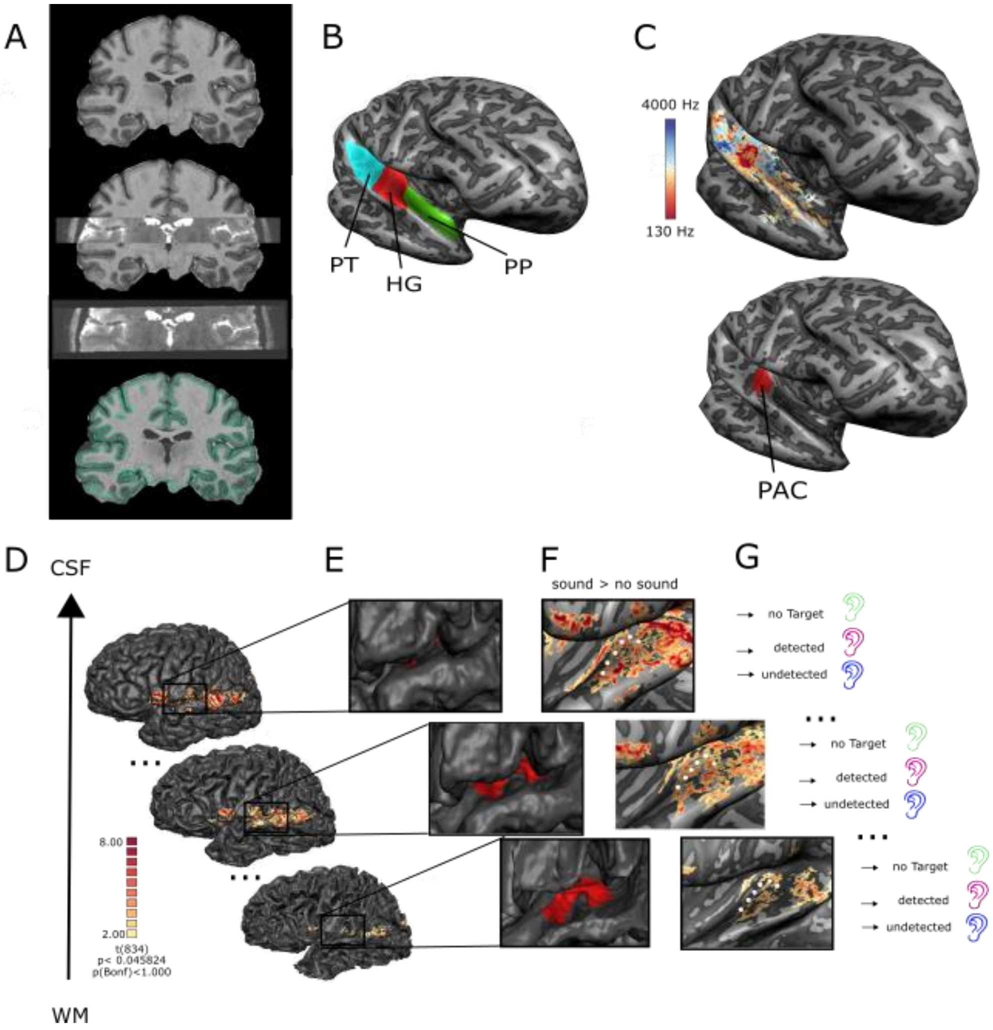

A
B
PT
HG
PP
C
4000 Hz
130 Hz
PAC
D
CSF
WM
8.00
2.00
t(834)
p< 0.045824
p(Bonf)<1.000
E
F
sound > no sound
G
no Target
detected
undetected
no Target
detected
undetected
no Target
detected
undetected

Figure 2. Analysis approach (previous page) (A) Interleaved anatomical image and functional volume, highlighting correspondence between datasets, anatomical images are segmented (and manually corrected around the ROI) to identify white and gray matter. See Figure S.5 for enlarged view. **(B)** The segmentation is used to reconstruct cortical surfaces (inflated view, with cortical curvature; light gray, gyrus; dark gray, sulcus). Anatomical ROIs (planum temporale (PT), Heschl's gyrus (HG) and planum polare (PP) are defined based on major anatomical landmarks (Kim et al, 2000) in every participant. **(C)** Right hemisphere showing tonotopic map and primary auditory cortex (PAC) in a single participant. PAC is functionally defined using the tonotopic localizer (Moerel et al, 2014). **(D)** We sampled eleven equivolume surfaces along the depth of cortical gray matter from white matter (WM) to cerebrospinal fluid (CSF). The most superficial, middle and deepest surface are shown. Overlaid activation map depicting the overall response to sounds (Ftest) in the main experiment. **(E)** Zoomed view onto temporal lobe with HG shown in red. Note the varying curvature across depth. (**F**) Inflated surfaces at each depth, functional map same as in D. The typical GE-BOLD increase of activation towards the cortical surface is visible. White dotted line demarks HG. **G)** Last step in the analysis is to inspect the response per perceptual condition at each cortical depth level.

Layer-specific detection modulation is not frequency-specific

In a first step, we compared if low and high frequency preferring voxel populations determined in the tonotopic localizer retain their frequency preference in the main experiment by visualizing the degree of frequency preference in PAC between the localizer and the main experiment, for every participant (Fig. S4.1). We confirmed that on average for each participant, frequency preference is retained as determined by relatively larger z-scored responses to the preferred frequency compared to the non-preferred carrier frequency of the population. The subsequent analyses were carried out on non- z-scored responses of the main experiment. To assess whether the layer-specific detection effect in superficial areas of PAC is tonotopic, we first sub-divided trials based on the presentation of a high or a low narrowband sound stimulus, and averaged the responses of each perceptual condition (detected, undetected, no Target) separately for low and high sounds. We identified high and low preferring voxel populations in PAC using the tonotopic localizer and inspected laminar response profiles of each perceptual condition separately for low and high sounds within each frequency-preferring population. This permitted assessing whether detection of a temporal shift embedded within a frequency-specific sound sequence elicited a frequency-specific response in the respective voxel population. At the cortical depth dependent level we did not observe a significant 3-way interaction (Depth:Condition:BFandSound (F (1,864) = 0.735, p >0.05), indicating that the detection effect was not significantly different for high and low preferring targets in high and low preferring sub-regions of PAC. We therefore collapsed the data across these conditions for the main analysis of this publication.

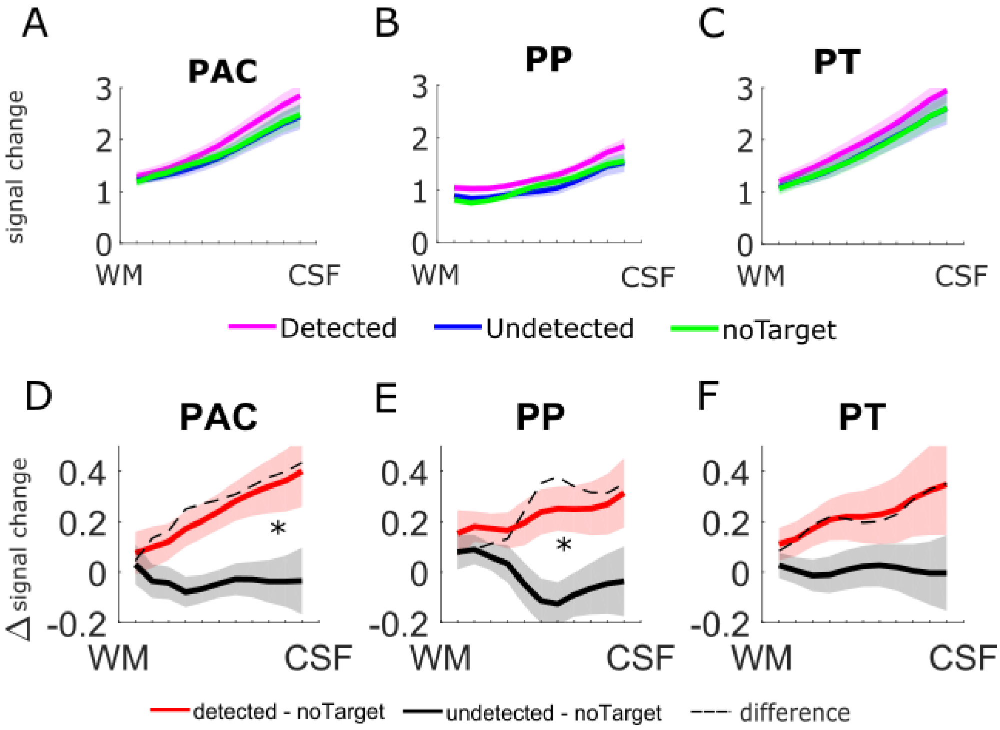


Figure 3. Layer-specific BOLD response for the different perceptual conditions per ROI. A. BOLD response to detected (magenta), undetected (blue) and no Target (green) sounds in the different layers of PAC, averaged over trials and participants. **B.** Same as A but in planum polare (PP). **C.** Same as A but in planum temporale (PT). Laminar profiles in all ROIs show an increase towards the cortical surface (closer to CSF). **D.** Difference in BOLD response between detected and no Target sounds (red; detected - no Target) and, undetected and no Target sounds (black; undetected - no Target) show a modulation of the BOLD response towards superficial layers of PAC driven by detection. Dashed line depicts difference between red and black line. **E.** Same as in D but for PP. A detection modulation in middle and superficial layers within PP occurs. **F.** Same as in D but for PT. No significant differences between BOLD response to detected and undetected targets are observed in area PT. Shading indicates the standard error of the mean across participants. Figures S3.1-S3.6 show single subject plots and the results for the HG ROI (not depicted here).

Detection of a target selectively increases activation in superficial layers of PAC

In PAC all three perceptual conditions show an increase in response from deep to superficial layers (Fig. 3 A). When subtracting the no Target condition, the additional modulation induced by detection of a target is apparent as an increase from deep to superficial layers (Fig. 3D - red line), while undetected targets do not result in a significant change in response compared to no Target trials (Fig. 3D - black line). This led to a significant interaction between depth and condition (F (1,216) = 22.48, $p< 0.001$). All three perceptual conditions show an increase in response from deep to superficial layers (Fig 3 A). When subtracting the no Target condition, the additional modulation induced by detection of a target is apparent as an increase from deep to superficial layers (Fig. 3D - red line), while undetected targets do not result in a change in response compared to no Target trials (Fig 3D - black line). This result is indicative of feedback related signals affecting the processing of superficial layers of PAC in relation to the processing of detected targets. A similar effect was present when considering the whole of Heschl's gyrus as region of interest (F (1,216) = 6.06, $p < 0.05$; Fig. S3.6).

Detection effects in non-primary areas

Middle layers of planum polare (PP) show a differential modulation of their response to detected and undetected sounds, when compared to no Target trials (Fig. 3B-E – significant interaction (F (1,216) = 6.85, $p< 0.01$). Qualitative inspection of this interaction suggests that this appears to be driven by a signal decrease in the undetected trials. This effect may be indicative of a signal being forwarded to the middle layers of PP from the superficial layers of PAC when targets are detected (increase in red line). We did not observe any modulation of the responses with detection in the planum temporale (Fig. 3 C-F – non-significant interaction (F (1,216) = 2.75, $p >0.05$) while all conditions elicited a strong increase in the response towards superficial layers (as expected due to draining and captured in a significant main effect of depth (F (1,216) = 5.25, $p< 0.05$).

Discussion

Layer-specific effects of detection in PAC

Previous fMRI research has highlighted the modulation of superficial cortical layers of (primary) auditory cortex when attending and responding to auditory stimuli (and ignoring visual ones) (Gau et al., 2016; De Martino et al.,2015). Here, we aimed to understand whether feedback mechanisms can explain why physically identical stimuli presented under identical attentional instructions are detected in some trials and not in others. To do so, we measured laminar fMRI responses from human auditory cortex, while participants performed a temporal target detection task at perceptual threshold. This allowed us to contrast the response to detected and undetected targets, while the bottom-up acoustic information remained identical. We showed that detected targets elicited a comparatively stronger response in superficial layers of the primary auditory cortex, indicating the relevance of feedback processing.

We have reported our results after subtracting the response to no Target trials from the responses to detected and undetected sounds. By doing so, we were able to control for offset effects induced by local vascular contributions to the BOLD signal, which should be consistent across the experimental conditions. This permitted highlighting the modulation induced by the detection of a target, despite the overall increase of the GE -EPI signal towards the pial surface (Uludag & Blinder, 2018). Acquisition techniques such as 3D-GRASE (Oshio & Feinberg, 1991) and VASO (Huber et al., 2017) which are not (or less) affected by vascular draining exist, GE-EPI offers increased sensitivity (compared to both 3D-GRASE and VASO), coverage (compared to 3D GRASE) and temporal efficiency (compared to VASO) all of which were essential to our study (Moerel, Yacoub, Gulban, Lage-Castellanos & De Martino, 2021).

It is conceivable that the increase in response we observed in superficial layers of PAC could have been the result of a fluctuation of attentional sampling, which is known to modulate long-latency sensory responses (Snyder et al., 2012). Multiple recent laminar fMRI studies located top-down effects of attention in superficial layers of human visual and auditory cortex either by attending to different modalities (auditory and visual) or by studying feature-based attention within a modality (Lawrence et al., 2019, De Martino et al., 2015; Gau et al.,2016 Liu et al., 2021; but see van Mourik et al.,2021). Invasive electrophysiological studies have also related changes in sensory gain of superficial layers to fluctuations of attention (Henry & Herrmann, 2014; Lakatos et al., 2013; O'Connell et al., 2014). Our results are thus consistent with the idea that attention modulates superficial layers of (auditory) cortical areas and provide first evidence that these small fluctuations can make the difference between detecting or not detecting an otherwise identical stimulus.

Our results identify PAC as a target of such feedback signals. Previous MEG research has also suggested that feedback to PAC (a unique long-latency response ranging between 50ms - 300ms) may be relevant to the detection of target sounds (Giani, Belardinelli, Ortiz, Kleiner, & Noppeney, 2015; Gutschalk, Micheyl & Oxenham, 2008). At the macroscopic level, increased fMRI BOLD responses in PAC in response to detected targets have been suggested to be the result of feedback signals (Wiegand & Gutschalk, 2012), potentially originating in parietal areas (Giani et al., 2015; Cusack, 2005). The plausible involvement of feedback to primary (auditory) cortical areas in determining the detectability of a stimulus is also corroborated by studies on bistable perception, or auditory streaming. These studies related variations in (primary) auditory cortex responses to changes in percept evoked by the identical physical stimulus (Micheyl, Tian, Carlyon & Rauschecker, 2005). Using fMRI, for example, responses in regions adjoining PAC have been associated with the perceptual interpretation of acoustically identical sounds (Kilian-Hütten, Valente, Vroomen & Formisano 2011) as well as to perceptual streaming (Hill, Bishop, Yadav & Miller, 2011).

Responses to detected targets are not modulated by frequency of the sounds

Contrary to previous invasive electrophysiology studies and non-invasive human studies (Lakatos et al. 2013; O'Connell et al. 2014; Riecke et al., 2018), we did not find the detection effect to be specific to cortical regions whose preference was maximal for the carrier frequency of the sounds (e.g. high vs. low frequency). While the absence of evidence is not evidence of the absence, a possible explanation for such inconsistency may stem from the nature of the task we employed. In previous research reporting frequency specific effects in auditory cortical regions, the task entailed focusing attention to the spectral content of the sounds (Lakatos et al., 2013; O’Connell et al., 2015; Riecke et al., 2018). In our task, the carrier frequency of the sounds was not the target of attention as participants were instructed to detect temporal shifts embedded in the stream of sounds. This line of reasoning, and our results are in line with previous investigations showing an attentional enhancement in layer 2/3, independent of the preferred frequency of the recording site when sound frequency was not task-relevant (Francis et al., 2018).

Layer-specific effect of detection in non-primary areas

Beyond the effect of detection in primary auditory cortex, we observed a layer-specific effect of detection in non-primary areas, namely of planum polare (PP), rostral to Heschl’s gyrus. Qualitative inspection of the responses showed a decrease in activation for the undetected trials compared to the no Target condition, resulting in a modulation of middle layers for the difference between detected and undetected trials (dashed line in Fig. 3E). Planum polare is considered secondary auditory cortex (Kim et al., 2000). We consider this region as part of the belt, which has been shown to have a stronger response to band pass noise than to tones (Petkov et al., 2006). Belt regions provide a clear step in auditory processing, and are situated at an anatomical intermediary position between primary core and parabelt regions (Kaas & Hackett, 2000; Wallace, Johnston & Palmer, 2002; Winer & Schreiner, 2010). It is thus conceivable that the modulation of middle layers of PP in response to detected targets may reflect feedforward signaling from superficial layers of PAC. While the posterior planum temporale (PT) is also part of the auditory belt and contains anatomical projections from PAC (Kaas & Hackett, 2000), we did not find a selective modulation of middle layers of PT. This indicates that if the middle layer modulation of PP is indeed related to feedforward signaling from PAC, such signals are sent only rostral from PAC in response to detected sounds. While our tasks required volunteers to detect temporal shifts, these shifts were small (less than 7 ms) and perceptually may have resulted in quintets with a shift being recognized as a different object compared to quintets without a shift. Our results of a modulation of middle layers of PP are thus in line with the known what/when division of the ventral and dorsal auditory pathway (Rauschecker & Tian, 2000; Jasmin, Lima & Scott, 2019).

In conclusion, the current study shows that when detecting a temporally shifted target, the response of neural populations in superficial layers of primary auditory cortex increases (in a non-frequency specific manner). This modulation is compatible with feedback signals targeting the primary auditory cortex. Planum polare also showed a layer depended effect, and the qualitative inspection of the responses revealed a modulation of middle layers, suggestive of the information related with target detection being forwarded from primary to secondary cortical areas. Future studies may be directed at identifying the source of the feedback signal we identified here.

Materials and Methods

Participants

Ten healthy participants (4 females, 6 males) were recruited for a total of 14 sessions. All participants were students or employees of Maastricht University. The study was approved by the research ethics committee of the Faculty of Psychology and Neuroscience at Maastricht University. For every subject we acquired 1 run of the tonotopic localizer, between 3 - 9 runs of the target detection experiment and a high-resolution anatomical scan. Most participants had previous experience with high-resolution fMRI studies.

Experimental design and stimuli

All stimulus presentation scripts were written in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA), using the Psychophysics toolbox (Brainard, 1997) and custom-code. Participants underwent a training session (~ 20 minutes) followed by a scanning session (~ 2 ½ hours). Participants 01, 02, 03 and 08 underwent two scan sessions, to acquire additional functional runs. Prior to each scan session the sound intensity of stimuli was adjusted individually to (perceptually) equalize the loudness of the stimuli presented within the localizer as well as in the training session and main experiment.

Target detection experiment. Participants were asked to detect a target; a temporal shift (TS) of a narrow-band sound embedded in a sequence of quintets repeating at 2 Hz. Narrowband sounds were centered around carrier frequencies of 200 Hz or 1100 Hz. The passbands around the carriers were constructed using equivalent rectangular bandwidths (ERBS = 4; Moore, 2003). Each passband consisted of a summation of 21 sinusoids with amplitude normalized to 1 and a random onset phase. A quintet consisted of five 10ms narrowband sounds, each separated by 10ms (see inset 1 Fig. 1A). Targets were constructed by shifting in time the third sound in a quintet (see inset 2 Fig. 1A). During a training session participants' 70% detection threshold was determined by means of a 2 down 1 up staircase, in which the size of a temporal shift (TS; ranging between 1 – 9 ms) determined the difficulty of a detection task. A 70% detection threshold outside the scanner was used as a starting threshold during the scanning session. The more challenging (i.e. louder scan environment) led to a (desirable) detection accuracy ~ 50% during scanning. Maintaining task difficulty to achieve a detection of 50% required adjusting the TS individually after every run by the experimenters, to ensure an approximately equal number of detected and undetected trials per participant, to be contrasted later on in the analysis. Supplementary Figure S1 displays the behavioral detection rates per participant for high and low sounds separately. All sounds were presented in intervals between acquisitions. After the sound finished, participants were cued by a green fixation cross to respond whether they had detected a target or not and instructed to press 1 or 2 on the button box. The cue for a button press was randomly jittered on each trial between 0 and 200ms after the sound had finished. Each run consisted of a total of 30 trials, 15 per carrier frequency, of which 3 trials per carrier were without a target and 12 containing a target. In addition, 3 silent trials per run were randomly interspersed functioning as baseline for sound vs silence contrasting.

Tonotopic localizer. To map tonotopic organization in the AC, a frequency localizer was performed (Formisano, 2003). We presented 7 center frequencies (130 Hz, 200 Hz, 306 Hz, 721 Hz, 1100 Hz, 1700 Hz and 4000 Hz) in blocks. Each block consisted of three amplitude modulated tones centered on one of the center frequencies (center frequencies +/- 0.1 octaves). Five center frequencies were log-spaced between 130 Hz and 4000 Hz, and two additional center frequencies were inserted (200 and 1100 Hz, the carrier frequencies employed in the target detection experiment). Tones were amplitude modulated (8 Hz, modulation depth of 1) and presented for 800 ms. During the localizer, participants were asked to fixate and passively listen to the sounds. The duration of the localizer was 7 ½ min.

MRI acquisition

Data acquisition was performed on a whole-body Magnetom scanner (nominal field strength 7 Tesla (T) (Siemens Medical Systems, Erlangen, German) at the Maastricht Brain Imaging Center, The Netherlands. All images were acquired using a 32-channel head coil (Nova Medical Inc. Wilmington, MA, USA).

Target detection experiment. For the sub-millimeter measurements, we used an event-related (sparse) design, with a 2D-GE-EPI sequence (TE/TR = 25 /3500ms, TA = 1400, silent gap = 2100; in-plane FoV 1120 x 1120mm; matrix size 200 X 200; slices = 42; GRAPPA factor = 3; partial Fourier = 6/8; phase-encoding direction anterior - posterior, with multiband factor = 2, and ascending slice order, yielding a nominal resolution of 0.8mm isotropic - see Fig S2.1 for design and coverage). Before acquisition of the first functional run, we acquired 10 volumes for distortion correction (5 volumes with opposite phase-encoding directions AP and PA).

Tonotopic localizer. We acquired the tonotopic localizer using a block design with a 2D-GE-EPI sequence (TE/TR = 21.2 /2600 ms; TA = 1200 ms; silent gap= 1400ms; in-plane FoV 1140x 1140 mm; matrix size 136 X 136; slices = 46, GRAPPA factor = 2, multiband factor= 2, partial Fourier = 6/8, phase-encoding direction anterior-posterior - yielding a voxel resolution of 1.2 mm isotropic). Preceding the localizer, 10 volumes in opposite phase-encoding direction (5 volumes AP and PA each) were acquired for distortion correction.

Anatomical scans. For visualization of the functional results and to obtain high-quality segmentations of the gray and white matter we obtained anatomical scans at a nominal voxel resolution of 0.65 mm isotropic. For this we used a MP2RAGE (Marques et al, 2010) sequence (TR = 5000 ms; TE = 2.5 ms; TI1 = 900 ms; TI2 =2700 ms; FoV 207x207 mm; matrix size 320 X 320; FA1 = 5 degrees, FA2 = 3 degrees; GRAPPA factor = 3 with an overall TA = 10:55 min). For four subjects a second scan session was performed, in which a lower resolution, hence faster, MPRAGE sequence at 1 mm isotropic, was used to acquire T1-weighted images for in-session alignment of functional data (TR = 2370 ms; TE = 2.3 ms; TI = 1500 ms; FoV 256x256 mm; matrix size 256x256; FA = 5 degrees; GRAPPA factor = 3 with an overall TA = 05:03 min).

Behavioral data analysis

Behavioral data were analyzed in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA). For every subject we determined the number of detected and undetected trials for the low and high carrier sounds separately (see Fig. S1.1 for behavioral performance per subject). Reaction times have not been analyzed as participants were cued to respond.

Anatomical data processing

Preprocessing. Anatomical images were processed using the advanced segmentation tools in BrainVoyager 21.4 (Brain Innovation, Maastricht, Netherlands), SPM's bias correction (Ashburner & Friston, 2005), ITK SNAP (Yushkkevich et al., 2006) and FSL BET (Jenkinson et al., 2012). If not otherwise indicated, default parameters were used.

The second inversion image of the MP2RAGE was subjected to the automated segmentation in SPM to obtain tissue-probability maps. The non-brain tissue-probability maps (C3, C4, C5) were manually thresholded and combined with a brain mask, obtained from the second inversion image using FSL BET. By combining these, we obtain a brain-mask that allows removing non-brain tissue and large veins (for a stepwise procedure see Kashyap, Ivanov, Havlicek, Poser & Uludag, 2019). This anatomical pre-processing workflow was developed particularly to work well for MP2RAGE data (https://github.com/srikash/presurfer). The resulting mask was inspected, had the cerebellum manually removed and was further manually polished using ITK SNAP in combination with a graphics tablet (Intuos Art; Wacom Co.). The resulting mask was applied to the T1w image (UNI) of the MP2RAGE. We then used BrainVoyager's intensity inhomogeneity correction and up-sampled the image to a resolution of 0.4 mm isotropic, using the spatial transformation option in BrainVoyager's 3D Volume tools. Lastly, the image was transformed from native space into a space in which the anterior and posterior commissure were in the same plane (ACPC space). We refer to this space as the voxel space.

Segmentation. The resulting image was input to BrainVoyager's advanced segmentation routine to obtain a white matter (WM) mask. This initial WM mask was inspected and manually polished in ITK SNAP, where emphasis was placed on corrections in the region of interest (bilateral auditory cortex [AC]). The polished WM mask was input to the subsequent step of the advanced segmentation routine in BrainVoyager to obtain a GM mask. This GM mask tended to be too inclusive, containing blood vessels, posing a challenge especially around the strongly vascularized AC. Therefore, we manually polished the GM definition and GM/CSF boundary in ITK SNAP. As a last step the obtained GM/WM segmentation was manually split into two hemispheres.

Cortical depth sampling. Using the GM/WM segmentation at 0.4 mm isotropic resolution, we measure the cortical thickness of individual segmented cortical hemispheres in volume space. Based on the cortical thickness we can perform whole-mesh cortical depth sampling, where we create surface meshes at equivolume cortical depth levels between the WM/GM boundary and the GM/CSF boundary (Waehnert et al. 2014). The created set of meshes at different cortical depth were then used to sample the functional data using trilinear interpolation. Surface visualizations are always based on the mid GM surface reconstruction.

Anatomical ROI selection. Based on macro-anatomical landmarks (sulci and gyri) and following the definition reported in Kim et al. (2000), the temporal lobe of each subject was divided into three anatomical ROIs in each hemisphere: Heschl's gyrus (HG), planum temporale (PT), planum polare (PP), drawn onto the inflated hemispheric surfaces, see Fig. 2B.

Functional data processing - Tonotopic localizer

*Preprocessing. P*reprocessing of the localizer data was performed in BrainVoyager 21.4, the NeuroElf toolbox in Matlab, as well as custom code in Matlab R2017a (The MathWorks Inc). Where not specified otherwise, default settings were used. Slice-scan-time correction, 3D motion correction (with sinc interpolation), linear trend removal and high-pass filtering (7 cycles) was performed. BrainVoyager's COPE plugin was used to correct EPI geometric distortions using a pair of opposite-phase encoded data.

Statistical analysis. All statistical computations were performed on the single subject level, by fitting a general linear model with a predictor for each center frequency to the data of the tonotopic localizer, obtaining a beta (response-strength) for every predictor and computing a statistical activation maps (FMap) of all predictors combined (contrast: sound> no sound). Fig. S3 shows the overall response to sounds in the localizer at statistical significance threshold qFDR > 0.05 for every participant.

Tonotopic maps. Tonotopic maps were derived following the standard procedure of z-scoring the response of voxels on the temporal lobe per frequency predictor, thereby removing a response bias towards low frequencies, and then color coding each voxel according to the frequency to which it best responded (i.e. its preferred frequency, indicated by the beta value – Formisano et al., 2003).

Functional ROI definition. In addition to dividing the human auditory cortex in terms of its major anatomical landmarks, we define primary auditory cortex functionally using the main tonotopic gradient obtain in the localizer (Moerel et al., 2014), as the auditory cortex in humans displays large macro-anatomical variability (Heschl, 1825; see appendix of this thesis).

The statistical activation map (FMap) in response to sounds, and the tonotopic map derived from the localizer were up-sampled from their native resolution at 1.2 mm isotropic by linearly interpolating to 0.8 mm isotropic to match the high-resolution functional data of the main experiment. The obtained up-sampled tonotopic map was then projected on the individual's reconstruction of the inflated mid-GM surface for each hemisphere, which allowed locating the main tonotopic gradient. The most likely position of the primary auditory cortex was localized using the tonotopic gradient of high frequency (posteromedial HG) to low frequency (medial portion HG) and back to high frequency. Supplementary figure S4 displays the tonotopic maps of all ten participants.

Functional data processing - Target detection experiment

Functional data were processed using BrainVoyager 21.4, the NeuroElf toolbox in Matlab, as well as custom code in Matlab R2017a (The MathWorks Inc). Where not specified otherwise, default settings were used.

Preprocessing. Preprocessing for all high-resolution functional data was performed in the default order in BrainVoyager (slice-scan time correction, 3D motion correction [with sinc interpolation and across runs] and linear trend removal and high-pass filtering (7 cycles). We corrected all functional images for EPI geometric distortions using BrainVoyager's COPE plugin based on the AP/PA images.

Co-registration of functional to anatomical images. The functional data of the first run were registered to the pre-processed anatomical data in native space using BrainVoyager's FMR-VMR co-registration. The positional information provided in the header is used for an initial alignment followed by fine-tuning co-registration using boundary-based registration. The result for the first run was visually inspected by overlaying the functional and anatomical images acquired in the same session and manually improved where necessary. The obtained initial alignment and fine-tuning alignment transformation files were used for the remaining runs within a session in combination with an ACPC transformation file to create a volume timecourse per run in the voxel space, using sinc interpolation. When a second session was acquired, co-registration of functional images was performed to in-session MPRAGE anatomical data. In a second step, between session anatomical data were then aligned using BrainVoyager's vmr-vmr co-registration and the resulting transformation matrix applied when creating volume time courses.

Functional data – statistical analysis

Statistical analysis per ROI. We computed a GLM with a separate predictor for every trial, classified as either being low detected, low undetected, low no Target or high detected, high undetected or high no Target, where high and low refers to the carrier frequency of the sound. Fig. 2A shows the overall response to sounds compared to baseline silence, corrected at qFDR < 0.05 for an exemplary subject. (See Fig. S2.2 for all subjects).

In a second step we sampled these single trial beta maps on 11 reconstructed depth dependent surfaces and averaged across trials of the same perceptual condition (Fig. 2D-G). To obtain laminar profiles multiple inclusion criteria guided the selection of vertices for sampling the mean beta surface maps (see Fig. S3.1). Vertices had to be within a particular ROI (PAC, HG, PP, PT). Their statistical F-value in response to sounds in the localizer needed to exceed F>2 and statistical F-value in response to sounds in the main experiment exceeded F>0.1, thereby ensuring that voxels with an (average) positive BOLD response to sounds in the main experiment were included, independent of depth. In each subject we extracted the mean (beta) across these vertices, per perceptual condition per depth. The perceptual conditions depended on the behavior of the subject and could lead to unequal condition size. Therefore, we bootstrapped a 95% confidence interval of the mean of trials per perceptual condition per depth (n=100).

Per ROI, second-level group statistics (n=10) were carried out on the mean differences of the bootstrapped betas between the perceptual conditions extracted from each subject (detected minus no Target and undetected minus no Target. Fig S3.4 and 3.5). The no Target condition served as control condition, that is we subtracted its effect from the effect of detection or non-detection of a target. We used three predictors (depth [linear], condition [detected minus no target; undetected minus no target] and their interaction) in a separate generalized linear mixed effects (GLME) model per ROI. Model fits were compared using likelihood ratio tests.

Checking frequency selectivity of detection effect in PAC. We expected voxels to retain their frequency preference (high vs low frequency) across the localizer and main experiment. To test this, we selected voxels whose time courses were modulated in response to sounds, exceeding a statistical threshold of F>2 in the tonotopic localizer. In the localizer data we divided the voxels as preferring low or high frequency by performing tonotopic mapping with only two predictors (200Hz and

1100 Hz). This allowed us to directly compare the preference from the tonotopy to the main experiment. For these groups of voxels (i.e., labelled as preferring low or high frequency in the localizer), we then plotted the response (after z-scoring as customary in tonotopic mapping) to the high and low preferring sounds (separately) in the main experiment (see Fig. S4.1)

Tonotopic analysis of main experiment. For the tonotopic analysis of the data we selected vertices in PAC as outlined in the previous section. Meaning that, we extracted the mean (beta) across vertices per perceptual condition (detected, undetected, no Target), per depth (11 levels), in low- and high-preferring groups of voxels within PAC, for low and high presented sounds in the main experiment.

Second-level group statistics (n=10) were carried out on the differences between perceptual conditions extracted from each subject (detected minus no Target and undetected minus no Target, Fig S4.2). For the tonotopic analysis of PAC we used four predictors (*depth* [linear], *condition* [detected minus no target; undetected minus no target], *BFandSound* [highSoundHighBF, highSoundLowBF, lowSoundHighBF, lowSoundlowBF] and their *interactions*) in a separate generalized linear mixed effects (GLME).

Supplementary materials

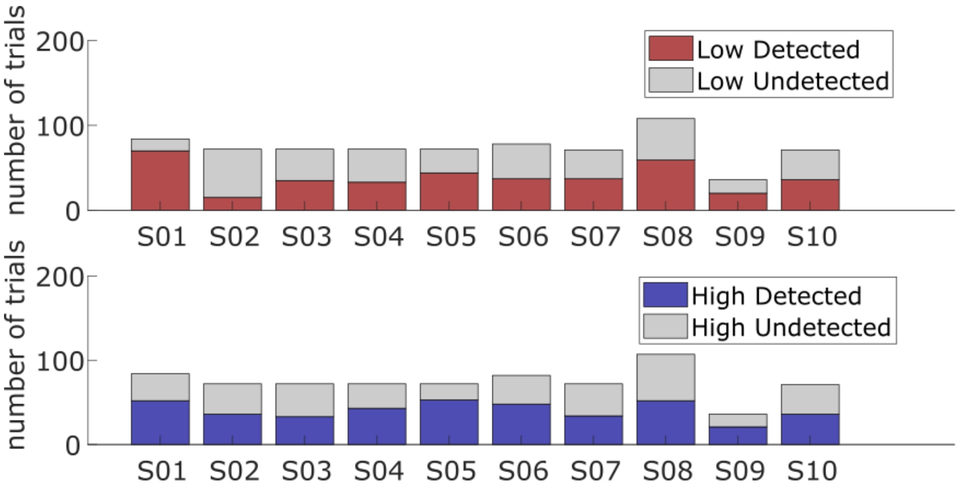


Figure S1.1. Behavioral detection rates per participant for low (top, red) and high (bottom, blue) frequency sounds. For most participants the behavioral task led to similar amounts of detected and undetected trials. S01 and S02 show unequal detection rates for the low carrier sounds.

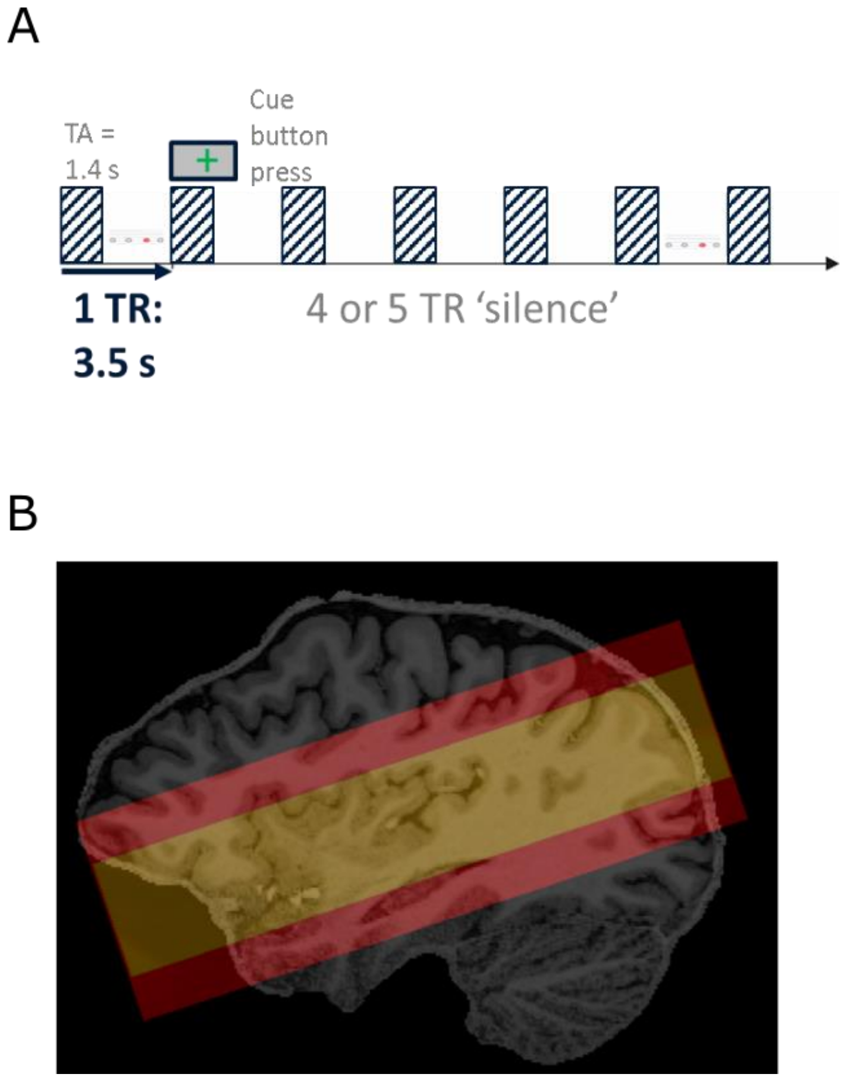


Figure S2.1. Experimental design and coverage (A) To acquire a trial of high-resolution BOLD data during our auditory task we used an event related design with a TR of 3.5s and an inter trial interval of 4/5 TRs. The acquisition of a functional volume lasted 1.4s after which we introduced a silent gap of 2s in which we presented the sounds. After every presented sound the participant was cued to respond whether they detected a target within the sound. **(B)** Yellow colored slab depicts the coverage (including the superior temporal plane and gyrus) of the acquired functional data (0.8 mm isotropic voxels) in the main experiment. We also acquired a tonotopic localizer centered on the same location (red colored slab) and whole-brain anatomical data using MP2RAGE.

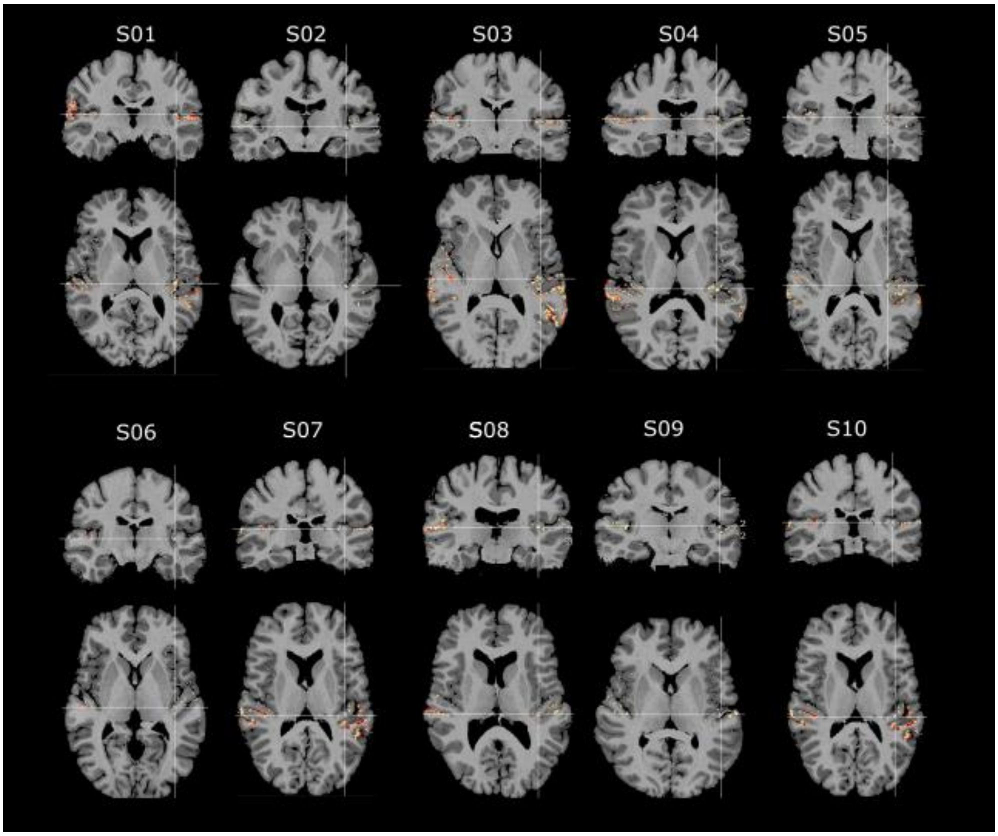


Figure S2.2. Single-subject results: The overall response to sounds in the main experiment. (F-Map; FDR corrected $q < 0.05$), except for S02, where $F>2$ has been used for visualization, projected on anatomical UNI image (up-sampled to 0.4 mm isotropic) and limited to GM voxels of the temporal lobe included in subsequent analyses.

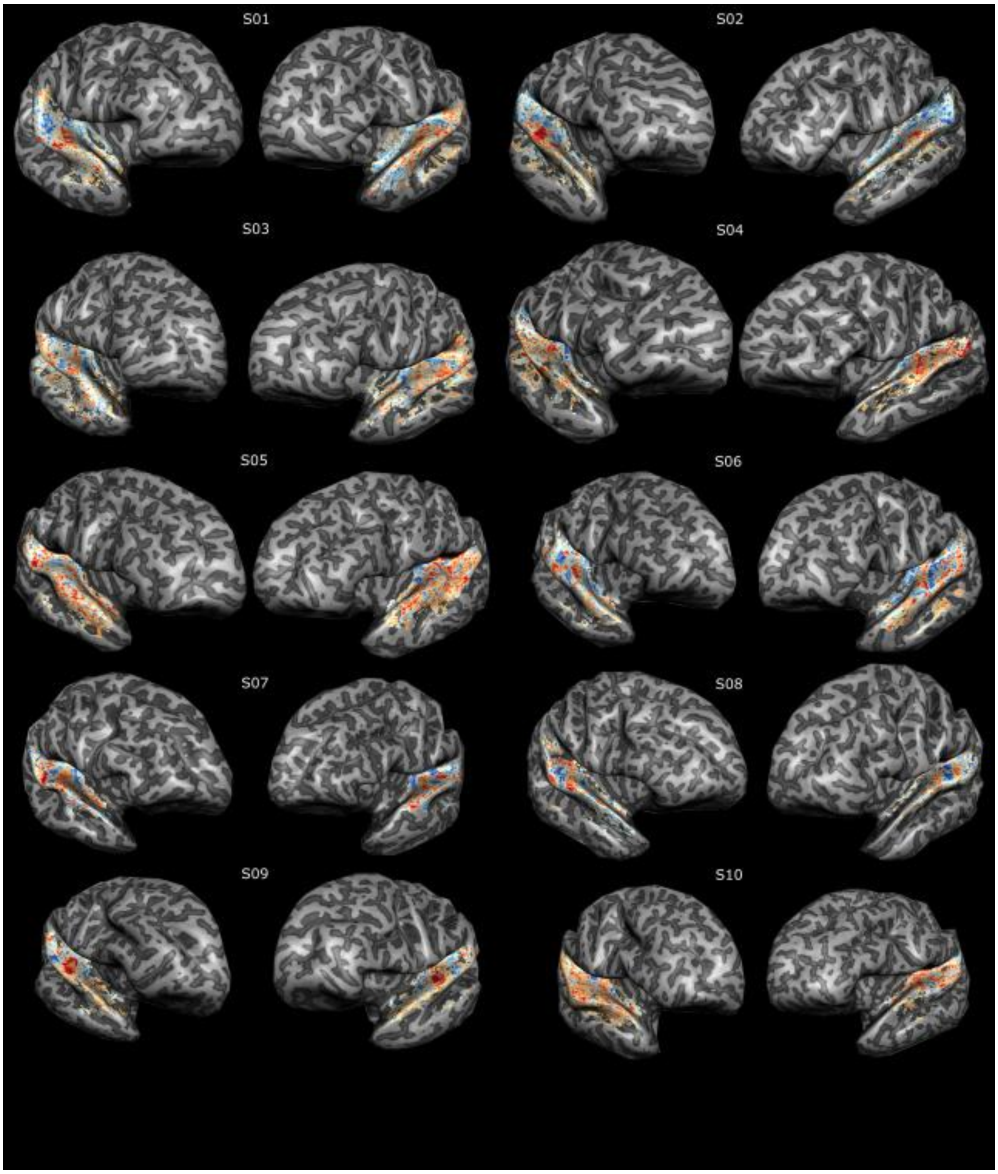


Figure S2.3. Tonotopic maps of independent localizer projected on inflated hemispheres of each participant (n=10). Maps are constructed using seven localizer frequencies. Each voxel was color-coded according to its best frequency response (i.e., the highest voxel activation in response to each of the frequencies). We used a red-yellow-blue color scale, where voxel tuning to low frequencies is indicated in red colors and tuning to high frequencies in blue. The maps shown here show consistent similarity across participants, with a typical tonotopic gradient in auditory cortical areas.

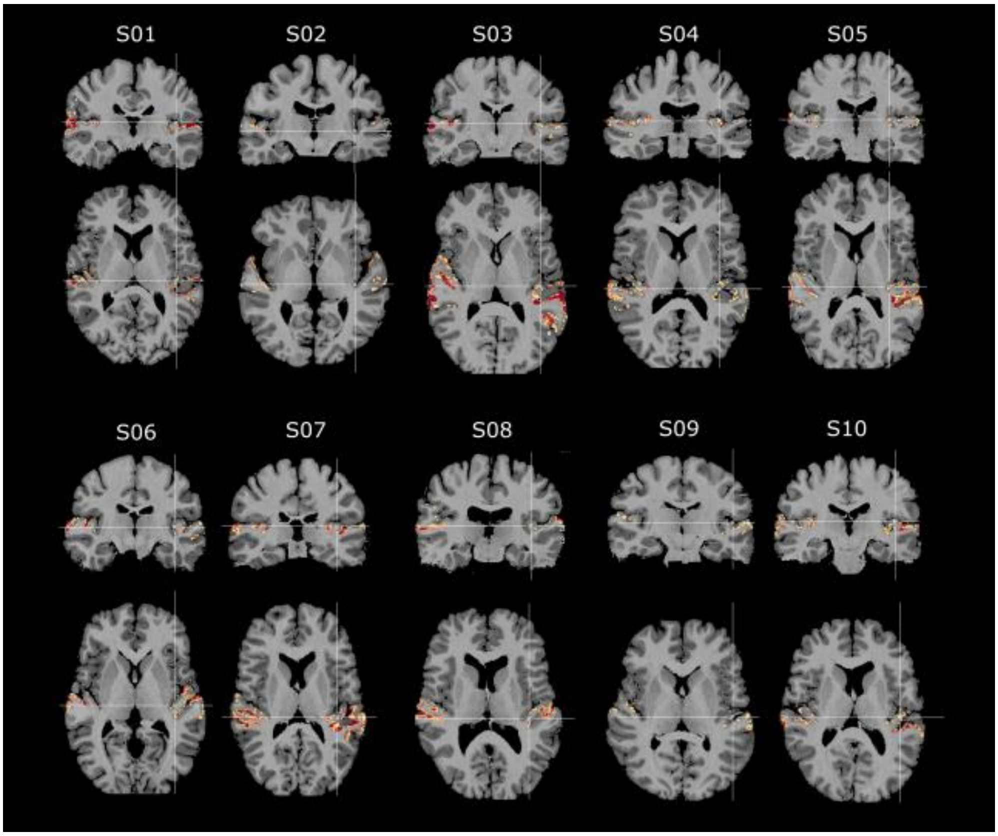


Figure S2.4. Single-subject results: The overall response to sounds in the localizer. (F-Map; FDR corrected $q < 0.05$), projected on anatomical UNI image (up-sampled to 0.4 mm isotropic) and limited to GM voxels of the temporal lobe included in subsequent analyses.

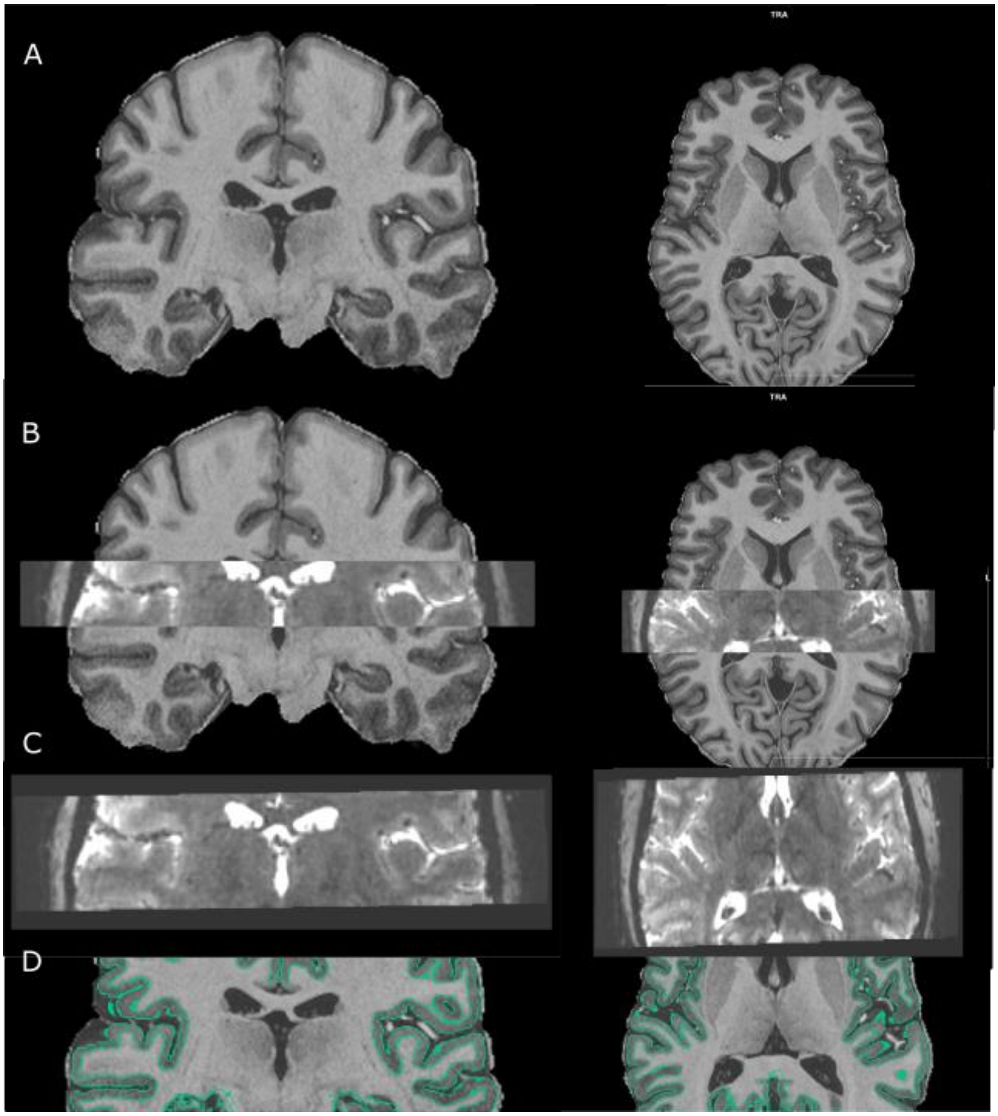


Figure S2.5 Individual coregistration and segmentation. Images show coronal slices (left) and transverse slices (right) in radiological convention. (A) Anatomical UNI image. (B) Anatomical image and functional volume are interleaved to highlight co-registration between datasets. (C) Functional volume of GE-EPI data showing lower intensity values in WM and higher intensity values in GM. (D) anatomical image with overlaid GM/WM and GM/CSF boundaries in green showing segmentation quality. Manual corrections were done around regions of interest.

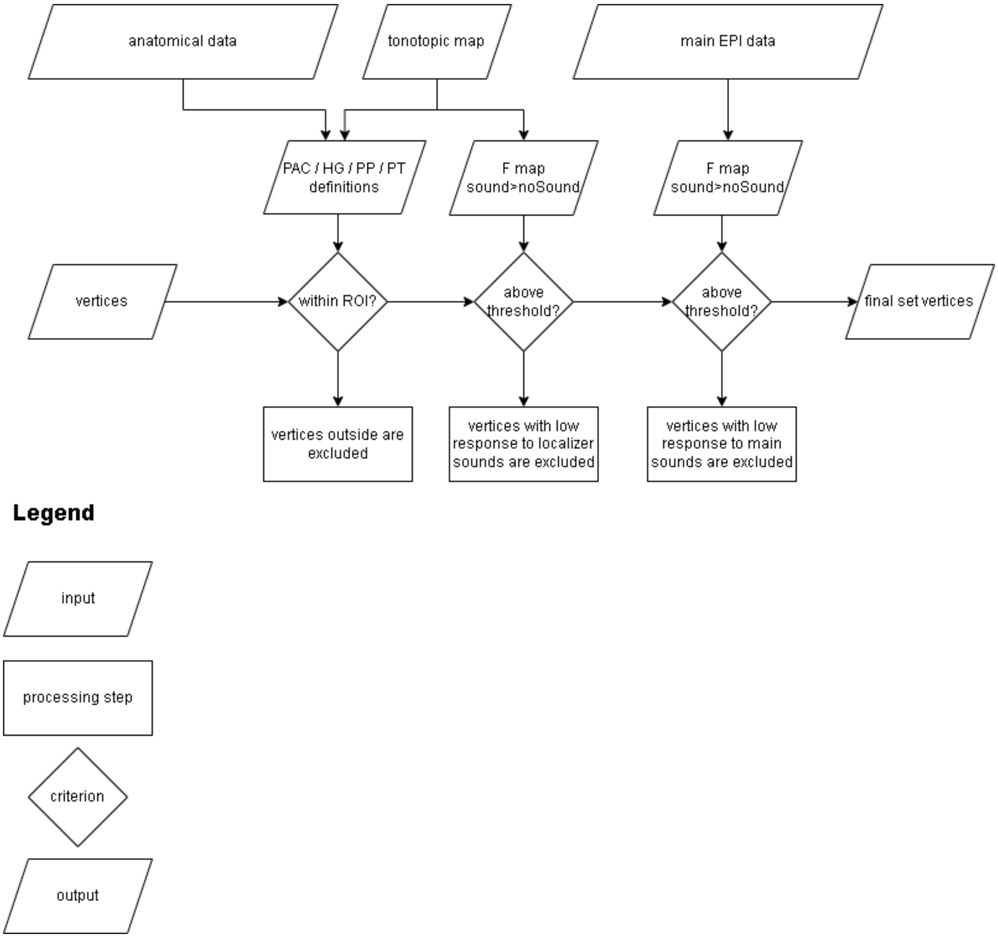


Figure S3.1. Overview of selection steps. Vertices were selected based on multiple criteria. Vertices had to be within a particular ROI (PAC, HG, PP, PT). Their statistical F>value in response to sounds in the localizer needed to exceed F >2. Then the set size was further reduced by only including vertices whose statistical F-value in response to sounds in the main experiment exceeded F>.0.1, thereby ensuring that all voxels with a positive BOLD response to sounds in the main experiment were included.

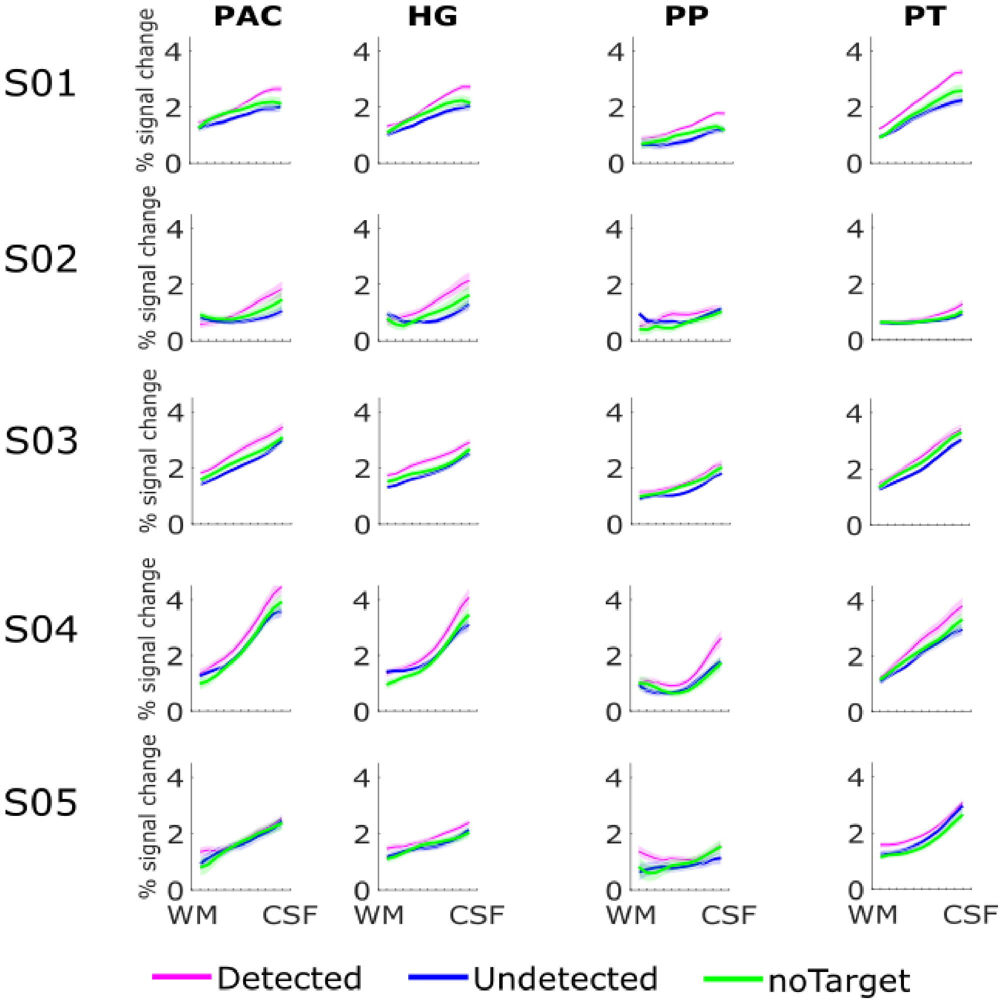


Figure S3.2. Single-subject laminar response profiles per ROI for perceptual conditions for subject S01-S05. BOLD response to detected (magenta), undetected (blue) and no Target (green) sounds in the different layers per ROI. Laminar profiles in all ROIs show an increase towards the cortical surface (closer to CSF). WM = white matter; CSF = cerebrospinal fluid; PAC = primary auditory cortex; HG = Heschl's gyrus; PP = planum polare; PT = planum temporale.

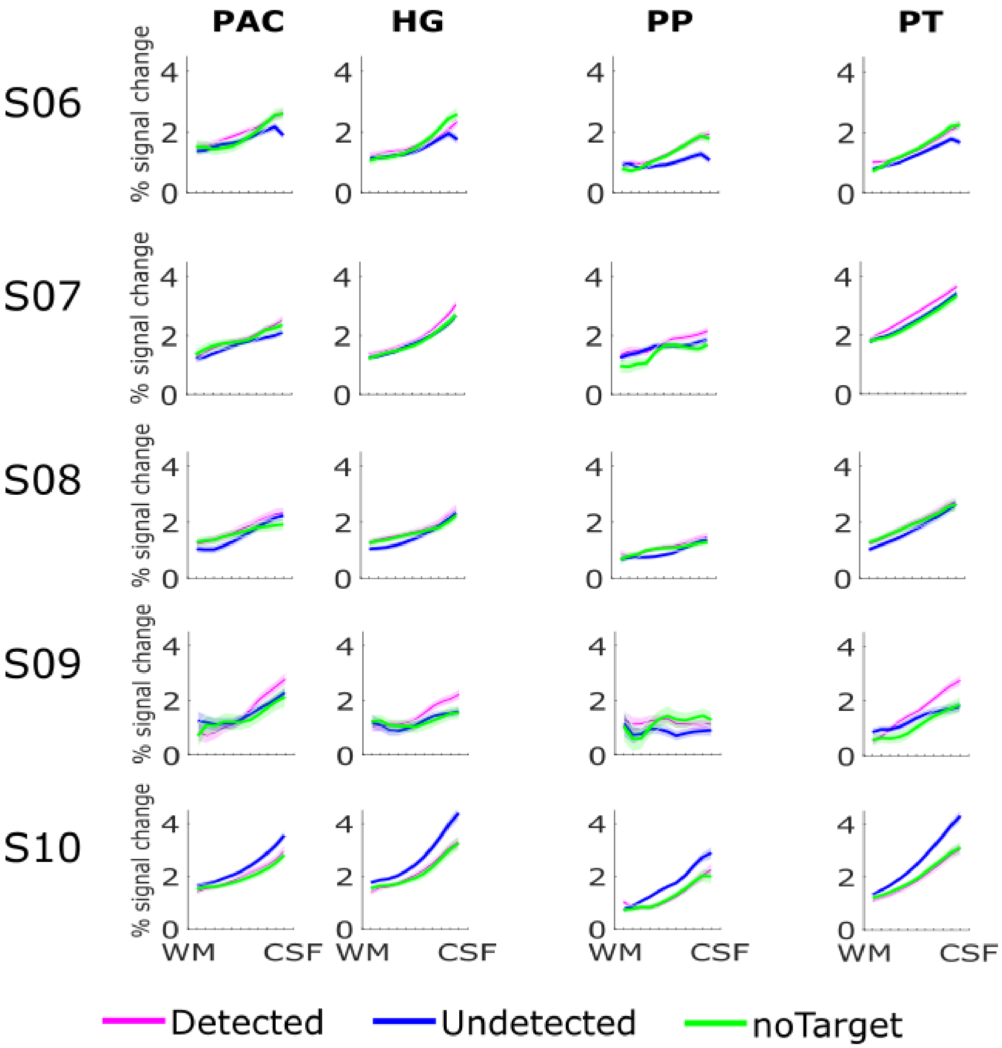


Figure S3.3. Single subject laminar response profiles per ROI for perceptual conditions for subject S06-S10. BOLD response to detected (magenta), undetected (blue) and no Target (green) sounds in the different layers per ROI. Laminar profiles in all ROIs show an increase towards the cortical surface (closer to CSF). WM = white matter; CSF = cerebrospinal fluid; PAC = primary auditory cortex; HG = Heschl's gyrus; PP = planum polare; PT = planum temporale.

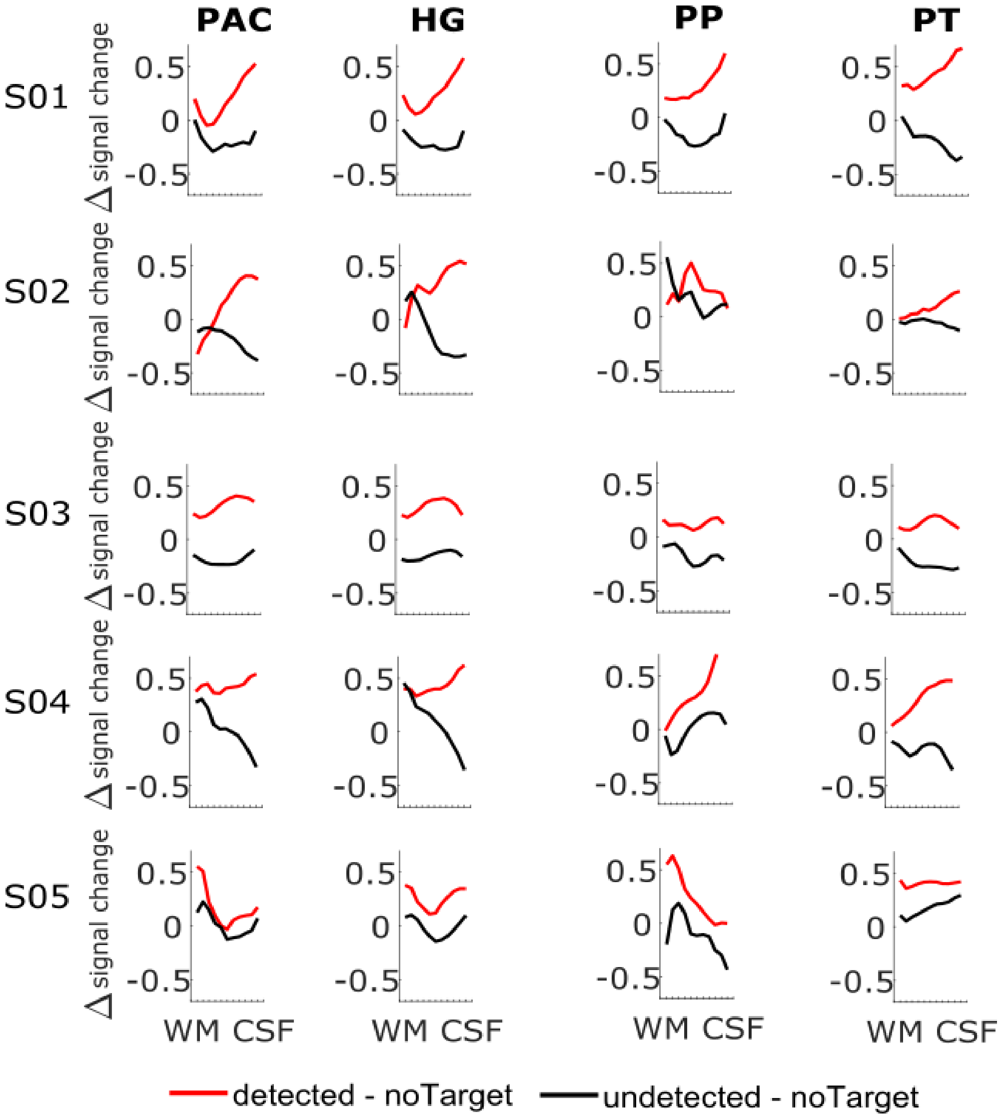


Figure S3.4. Single subject laminar response profiles per ROI for difference between perceptual conditions for subject S01-S05. Difference in BOLD response between detected and no Target sounds (red; detected – no Target) and, undetected and no Target sounds (black; undetected – no Target). WM = white matter; CSF = cerebrospinal fluid; PAC = primary auditory cortex; HG = Heschl's gyrus; PP = planum polare; PT = planum temporale.

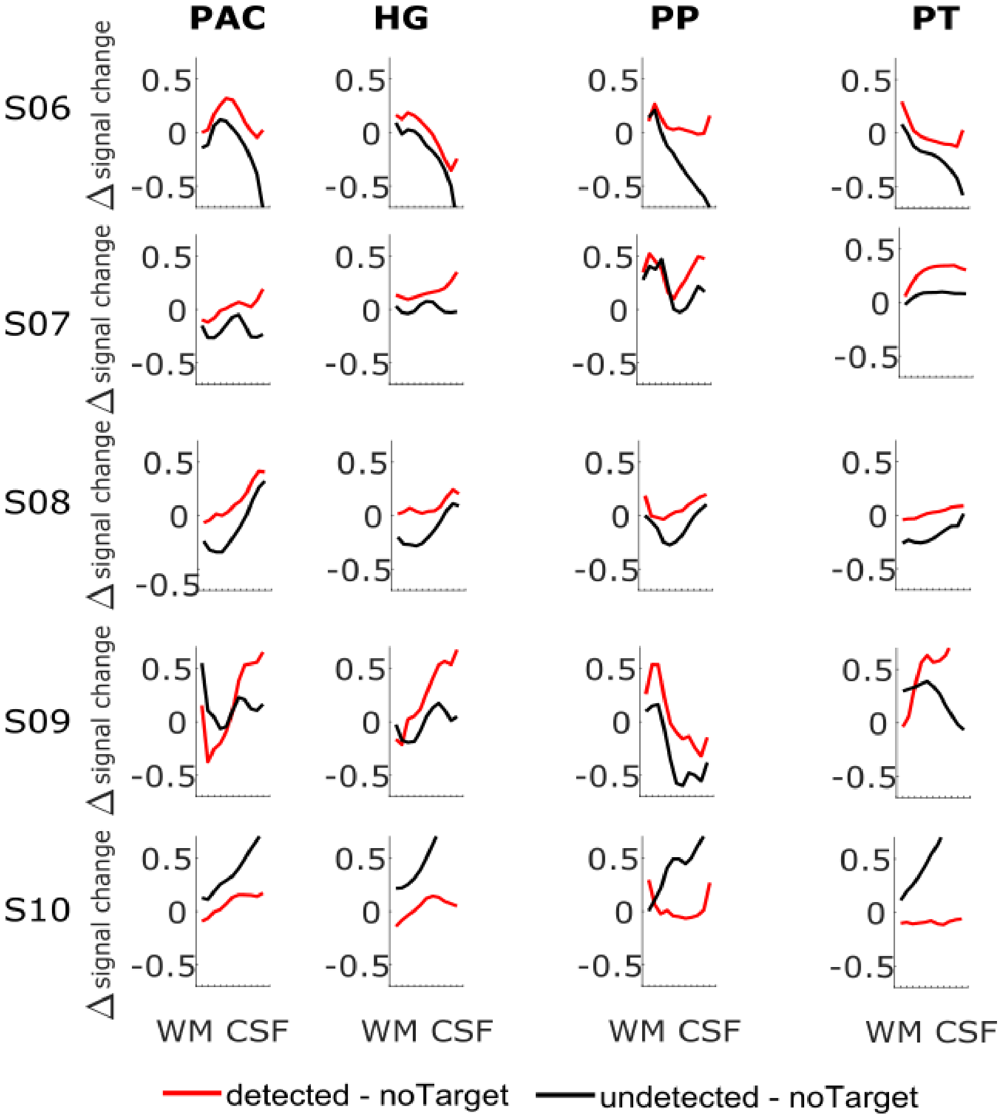


Figure S3.5. Single subject laminar response profiles per ROI for difference between perceptual conditions for subject S06-S10. Difference in BOLD response between detected and no Target sounds (red; detected – no Target) and, undetected and no Target sounds (black; undetected – no Target). WM = white matter; CSF = cerebrospinal fluid; PAC = primary auditory cortex; HG = Heschl's gyrus; PP = planum polare; PT = planum temporale.

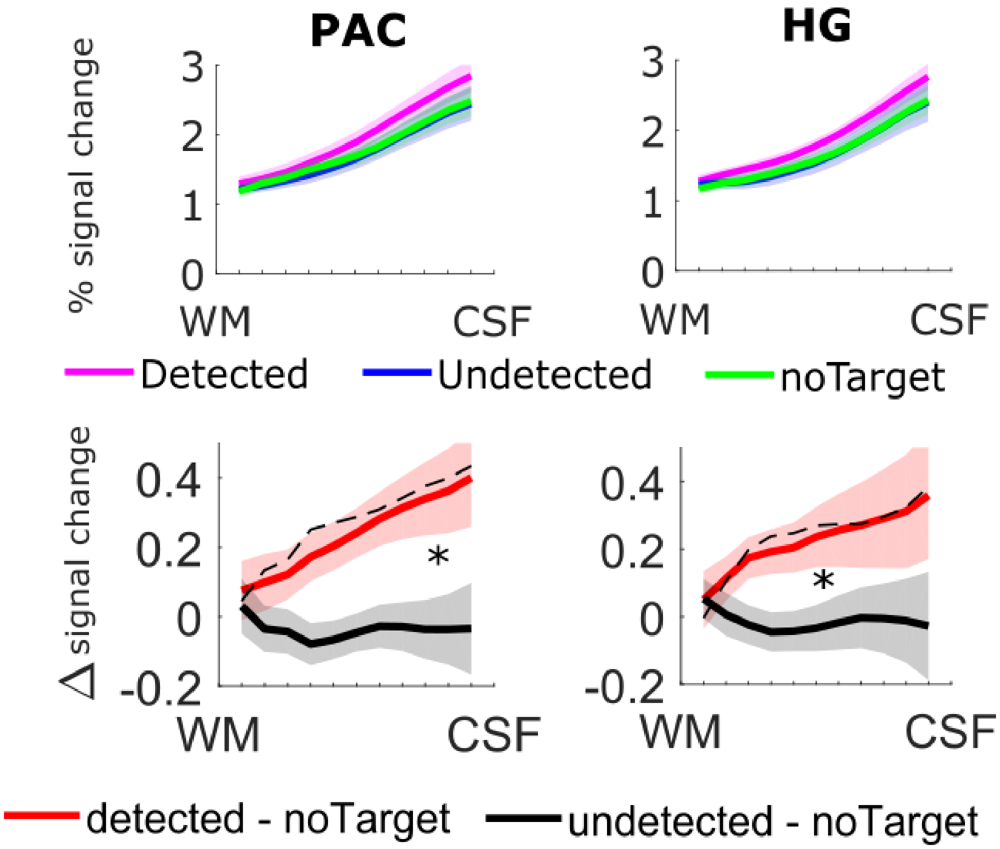


Figure S3.6. Laminar response profile in primary auditory cortex (PAC) and Heschl's gyrus (HG) are similar. Top row: BOLD response to detected (magenta), undetected (blue) and no Target (green) sounds in the different layers of PAC and HG, averaged over trials and participants. Bottom row: Difference in BOLD response between detected and no Target sounds (red; detected – no Target) and, undetected and no Target sounds (black; undetected – no Target) show a modulation of the BOLD response towards superficial layers of PAC as well as in HG driven by detection. Dashed line depicts difference between red and black line. PAC = primary auditory cortex; HG = Heschl's gyrus; WM = white matter; CSF = cerebrospinal fluid; Shading of indicates the standard error of the mean across participants.

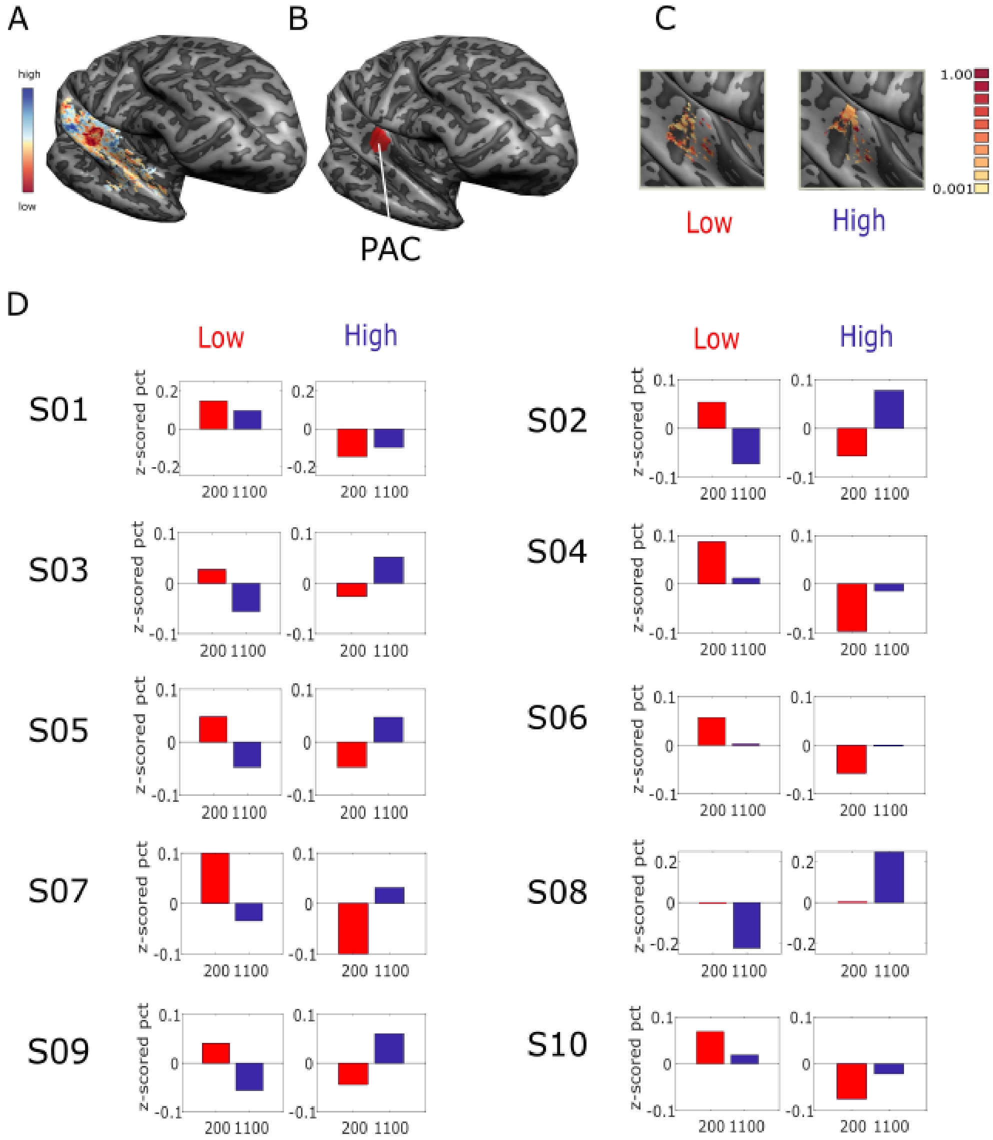

A
high
low
B
PAC
C
1.00
0.001
Low
High
D
Low
High
Low
High
S01
S02
S03
S04
S05
S06
S07
S08
S09
S10
z-scored pct
0.2
0
-0.2
0.1
-0.1
200 1100

Figure S4.1 (previous page). Analysis approach tonotopic detection showing responses in main experiment are tonotopic. (A) A tonotopic localizer map on 3D hemispheric surface reconstruction (inflated view, with cortical curvature; light gray, gyrus; dark gray, sulcus). **(B)** Reconstructed hemisphere showing primary auditory cortex (PAC) in a single participant. PAC is functionally defined using the frequency-gradient from the tonotopic localizer (Moerel et al, 2014). **(C)** Within PAC we determine BF- populations of voxels preferring either high or low sounds in localizer. Selectivity for BF is computed as relative contribution and ranges between 0 and 1, where 1 is high selectivity. **(D)** Responses of main experiment to low (200Hz) and high (1100Hz) sounds, sampled per BF populations (columns) as identified in localizer (average of both hemispheres). On average, voxels retain their frequency preference between localizer and main experiment. To assess tonotopic quality of the main experiment we z-scored responses to low and high sounds of the main experiment as done in tonotopic mapping. Each population of voxels as identified in the localizer, also displays a frequency preference (relative higher z-scored beta response to preferred than non-preferred frequency) in the main experiment, confirming that localizer can be used to identify best-frequency populations that maintain their selectivity in the main experiment. Note that the relative difference between high and low for each ROI.

Acknowledgements

We thank Agustin Lage-Castellanos for help with the statistical analysis, Sriranga Kashyap for help with the anatomical preprocessing, and our participants for their endurance during long scan sessions.

References

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. Neuroimage, 26(3), 839-851.

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433–436. https://doi.org/10.1163/156856897X00357

Cusack, R. (2005). The intraparietal sulcus and perceptual organization. Journal of cognitive neuroscience, 17(4), 641-651. https://doi.org/10.1162/0898929053467541

Da Costa, S., van der Zwaag, W., Miller, L. M., Clarke, S., & Saenz, M. (2013). Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. Journal of Neuroscience, 33(5), 1858-1863. https://doi.org/10.1523/JNEUROSCI.4405-12.2013

De Martino, F., Moerel, M., Ugurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2015). Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proceedings of the National Academy of Sciences, 112(52), 16036–16041. https://doi.org/10.1073/pnas.1507552112

Douglas, R. J., & Martin, K. A. (2004). Neuronal circuits of the neocortex. Annu. Rev. Neurosci., 27, 419-451. https://doi.org/10.1146/annurev.neuro.27.070203.144152

Douglas, R. J., Martin, K. A., & Whitteridge, D. (1989). A canonical microcircuit for neocortex. Neural computation, 1(4), 480-488. https://doi.org/10.1162/neco.1989.1.4.480

Formisano, E., Kim, D. S., Di Salle, F., Van de Moortele, P. F., Ugurbil, K., & Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron, 40(4), 859-869. https://doi.org/10.1016/S0896-6273(03)00669-X

Francis, N. A., Elgueda, D., Englitz, B., Fritz, J. B., & Shamma, S. A. (2018). Laminar profile of task-related plasticity in ferret primary auditory cortex. Scientific Reports, 8(1), 1-10.

Gau, R., Bazin, P. L., Trampel, R., Turner, R., & Noppeney, U. (2020). Resolving multisensory and attentional influences across cortical depth in sensory cortices. Elife, 9, e46856.

Giani, A. S., Belardinelli, P., Ortiz, E., Kleiner, M., & Noppeney, U. (2015). Detecting tones in complex auditory scenes. NeuroImage, 122, 203–213. https://doi.org/10.1016/j.neuroimage.2015.07.001

Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. PLoS biology, 6(6), e138.

Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. Timing & Time Perception, 2(1), 62-86.

Heschl, R. L. (1878). - Ueber die vordere quere Schläfenwindung des menschlichen Grosshirns.

Hill, K. T., Bishop, C. W., Yadav, D., & Miller, L. M. (2011). Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. BMC neuroscience, 12(1), 1-8.

Huber, L., Handwerker, D. A., Jangraw, D. C., Chen, G., Hall, A., Stüber, C., ... & Bandettini, P. A. (2017). High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. Neuron, 96(6), 1253-1263.

Jasmin, K., Lima, C. F., & Scott, S. K. (2019). Understanding rostral–caudal auditory cortex contributions to auditory perception. Nature Reviews Neuroscience, 20(7), 425-434.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. Neuroimage, 62(2), 782-790.

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. Proceedings of the National Academy of Sciences of the United States of America, 97(22), 11793–11799. https://doi.org/10.1073/pnas.97.22.11793

Kashyap, S., Ivanov, D., Havlicek, M., Poser, B., & Uludag, K. (2019). Laminar CBF and BOLD fMRI in the human visual cortex using arterial spin labelling at 7T. In Proceedings of the 27th Scientific Meeting of ISMRM (p. 609).

Kilian-Hütten, N., Valente, G., Vroomen, J., & Formisano, E. (2011). Auditory cortex encodes the perceptual interpretation of ambiguous sound. Journal of Neuroscience, 31(5), 1715-1720. https://doi.org/10.1523/JNEUROSCI.4572-10.2011

Kim, J. J., Crespo-Facorro, B., Andreasen, N. C., O'Leary, D. S., Zhang, B., Harris, G., & Magnotta, V. A. (2000). An MRI-based parcellation method for the temporal lobe. Neuroimage, 11(4), 271-288.

Klein, B. P., Fracasso, A., van Dijk, J. A., Paffen, C. L., Te Pas, S. F., & Dumoulin, S. O. (2018). Cortical depth dependent population receptive field attraction by spatial attention in human V1. NeuroImage, 176, 301-312.

Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron [Internet]. 2017;93(3):480–90.

Lakatos, P., Musacchia, G., O’Connell, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. Neuron, 77(4), 750-761. https://doi.org/10.1016/j.neuron.2012.11.034

Lawrence, S. J., Norris, D. G., & De Lange, F. P. (2019). Dissociable laminar profiles of concurrent bottom-up and top-down modulation in the human visual cortex. Elife, 8, e44422. https://doi.org/10.7554/eLife.44422

Liu, C., Guo, F., Qian, C., Zhang, Z., Sun, K., Wang, D. J., ... & Zhang, P. (2021). Layer-dependent multiplicative effects of spatial attention on contrast responses in human early visual cortex. Progress in Neurobiology, 207, 101897.

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P. F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage, 49(2), 1271-1281.

Merzenich, M. M., & Brugge, J. F. (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain research, 50(2), 275-296. https://doi.org/10.1016/0006-8993(73)90731-2

Merzenich, M. M., Knight, P. L., & Roth, G. L. (1973). Cochleotopic organization of primary auditory cortex in the cat. Brain research, 63, 343-346. https://doi.org/10.1016/0006-8993(73)90101-7

Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron, 48(1), 139-148. https://doi.org/10.1016/j.neuron.2005.08.039

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

Moerel, M., Yacoub, E., Gulban, O. F., Lage-Castellanos, A., & De Martino, F. (2021). Using high spatial resolution fMRI to understand representation in the auditory network. Progress in Neurobiology, 207(January 2020), 101887. https://doi.org/10.1016/j.pneurobio.2020.101887

Moore, B. C. J. (2003). An Introduction to the Psychology of Hearing. Boston Academic Press, 3, 413. https://doi.org/10.1016/j.tins.2007.05.005

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., ... & Yacoub, E. (2015). Contextual feedback to superficial layers of V1. Current Biology, 25(20), 2690-2695. https://doi.org/10.1016/j.cub.2015.08.057

O'Connell, M. N., Barczak, A., Schroeder, C. E., & Lakatos, P. (2014). Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex. Journal of Neuroscience, 34(49), 16496-16508. https://doi.org/10.1523/JNEUROSCI.2055-14.2014

Oshio, K., & Feinberg, D. A. (1991). GRASE (gradient-and spin-echo) imaging: a novel fast MRI technique. Magnetic resonance in medicine, 20(2), 344-349.

Petkov, C. I., Kayser, C., Augath, M., & Logothetis, N. K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. PLoS biology, 4(7), e215.

Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010). Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. Neuroimage, 52(4), 1334-1346.

Riecke, L., Peters, J. C., Valente, G., Poser, B. A., Kemper, V. G., Formisano, E., & Sorger, B. (2018). Frequency-specific attentional modulation in human primary auditory cortex and midbrain. NeuroImage, 174, 274-287. https://doi.org/10.1016/j.neuroimage.2018.03.038

Snyder JS, Gregg MK, Weintraub DM, Alain C. Attention, awareness, and the perception of auditory scenes. Front Psychol. 2012;3(FEB):1–17.

Turner, R. (2002). How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. Neuroimage, 16(4), 1062-1067.

Uludağ, K., & Blinder, P. (2018). Linking brain vascular physiology to hemodynamic response in ultra-high field MRI. Neuroimage, 168, 279-295.

van Mourik, T., Koopmans, P. J., Bains, L. J., Norris, D. G., & Jehee, J. F. (2021). Investigation of layer specific BOLD in the human visual cortex during visual attention. bioRxiv.

Waehnert, M. D., Dinse, J., Weiss, M., Streicher, M. N., Waehnert, P., Geyer, S., ... & Bazin, P. L. (2014). Anatomically motivated modeling of cortical laminae. Neuroimage, 93, 210-220.

Wallace, M. N., Johnston, P. W., & Palmer, A. R. (2002). Histochemical identification of cortical areas in the auditory region of the human brain. Experimental Brain Research, 143(4), 499-508.

Wiegand, K., & Gutschalk, A. (2012). Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. Neuroimage, 61(1), 62-69. https://doi.org/10.1016/j.neuroimage.2012.02.067

Winer, J. A., & Schreiner, C. E. (Eds.). (2010). The auditory cortex. Springer Science & Business Media.

Yushkevich, P. (2006). ITK-SNaP integration, NLM insight. Webpage: http://www. itk. org/index. html

5

General Discussion

Summary

In three empirical studies, we investigated the behavioral sensitivity and reaction times to rhythms and other forms of temporal structure, as well as the neural correlates underlying the detection of target sounds embedded in rhythmic stimuli. In **Chapter 2** we first developed a behavioral paradigm, in which we examined the contributions of temporal rates (i.e. changes in rhythm) and carrier frequencies of sounds, as well as their interaction to the detection of temporally shifted targets. This paradigm was then applied in **Chapter 3** to study the effect of temporal structure and predictability awareness on behavioral sensitivity and response times. Finally, in **Chapter 4** we explored the neural correlates underlying the detection of temporally shifted targets, when attending to rhythmic sounds. Specifically, we examined layer-specific responses in auditory areas using 7T fMRI. While the specific results of conducted studies are discussed in the separate chapters, the following section aims at summarizing and integrating the findings across chapters. In particular, speculations on the role of top-down attention contributions in rhythmic perception are suggested.

The brain uses regularities in the environment to predict upcoming sensory input. In audition, one form of such a regularity is rhythm. Current theories of neural processing propose that the brain formulates predictions of upcoming sensory events to facilitate perception (Friston, 2011). Such predictions are based on external input from the environment, and can be formed both about the content as well as the timing (Arnal & Giraud, 2012). Prediction mechanisms have been shown to affect behavior in humans. In audition, the ability to predict an upcoming sound facilitates sensitivity and response times when the sound is task relevant (van Ede & Nobre, 2018). The underlying neural mechanisms may be an alignment of ongoing oscillations in auditory areas when attending to rhythmic stimuli (shown in non-human primates; Lakatos et al.,2014), supporting a top-down centered view of rhythmic facilitation. On the other hand, behavioral research has also shown that presenting sounds in a predictable periodic stream (vs an aperiodic one) entails perceptual benefits even when this is not task-relevant (Breska & Deouell, 2014). It has therefore been suggested, that the underlying neural mechanism might be a (partially) automatic response to rhythmic sensory input (Haegens & Zion-Golumbic, 2018).

In **Chapter 2** we aimed to investigate whether some rhythms might be better suited than others in engaging an underlying neural mechanism. We assessed this by investigating the behavioral sensitivity and reaction time to targets embedded within these rhythms. Moreover, we controlled for carrier frequency and repetition i.e. position of target within sequences. Our results allowed us to conclude that given the specific task and stimuli, slow rhythmic streams of low frequency sounds (200 Hz) lead to largest behavioral benefits in reaction times and sensitivity. The perceptual benefit of slow rates is in line with propositions that a temporal prediction mechanism may be instantiated by the motor system, such that auditory perception is influenced by motor activity, particularly in the delta band, imposing temporal constraints on the sampling of sensory information (Morillon et al., 2019).

In **Chapter 3** the temporal shift detection task was employed again, to probe effect of different temporal structures while keeping predictability of targets constant. Moreover, we investigated whether being aware of the predictability of targets affects behavioral detection and reaction times. We showed that while rhythmic structure allowed to proactively enhance perceptual sensitivity, temporal cueing conferred an equal behavioral advantage as periodic rhythms, when no specific instructions on the relevance of the temporal structure were provided. In contrast, in a second experiment we showed that explicit instruction further improve behavioral benefits of rhythms compared to a temporal cue, suggesting that attention can boost expectations elicited by periodic stimulation leading to enhanced sensitivity in target detection. These results highlight that attention can be directed in time (Nobre & Coull, 2010) and that it can be leveraged to enhance rhythmic effects.

What becomes clear from these investigations is that attention interacts with the bottom-up rhythmic stimulus, such that attending to specific moments in periodic rhythms increases perceptual sensitivity. Yet despite constant efforts by participants to attend and detect the relevant target stimuli, in some trials the targets were not detected. In **Chapter 4** we aimed to investigate the cortical circuitry underlying our ability to detect behaviorally relevant sound changes. In particular, we examined the feedforward and feedback processes within the auditory hierarchy during detection (by comparing detected and undetected targets). While previous human neuroimaging studies had identified primary auditory cortex (PAC) as contributing to target detection, the exact role of feedforward and feedback processing remained speculative (Giani et al., 2015; Gutschalk, Micheyl & Oxenham, 2008; Wiegand & Gutschalk, 2012). We sampled the spatial mesoscale noninvasively and observed an increased response in superficial layers of PAC; suggestive of a feedback signature of detection, in line with the known hierarchical organization of the auditory cortex (Hackett et al., 2014).

Taken together, the results in this thesis show that stimulus detection is proactively optimized by using temporal regularities in the environment, both afforded via rhythmicity and temporal cueing. The magnitude of a behavioral benefit of predictability by a rhythm and temporal cue compared to a less predictable aperiodic condition is similar. However, directing attention through explicit instruction on target predictability can have a modulating effect on the effect of temporal structure on detection (sensitivity), such that in periodic rhythms sensitivity is improved relative to a temporal cue. When such a stimulus is attended and a target is detected a feedback signature of this detection in superficial layers of the primary auditory cortex can be observed.

General outlook and limitations

As elaborated in the introduction of this thesis, the different levels of analysis described by Marr (computation, algorithm and implementation) mutually inform each other and ultimately allow understanding neural mechanisms. As researchers we are often confined to certain levels of investigation on spatial and temporal scales and the link between results from different observations is left to be made by the scientific community. In the present thesis we are, on the one hand employing one of the most established forms of investigation; namely psychophysical studies. This allowed us to be close to the relevant phenomenon, i.e. we used psychoacoustics to observe behavior in a controlled, quantitative manner. By manipulating low-level (acoustic) features we can gain an understanding of the system and which algorithms it may use. We followed this rationale in Chapters 2 and 3. Identifying the computation and probing how the underlying algorithm may work (by probing and redefining the behavioral phenomenon), yet this leaves open the question of how a mechanism is implemented biologically. We used 7T high-field fMRI to investigate the level of cortical layers, which is the finest scale at which functional operations can be investigated non-invasively in living humans. By combining these methods of investigation and bridging from the implementation to computational level, the present thesis shows that computations in cortical layers subserve the detection of behaviorally relevant events in rhythmic sounds.

Assessing responses to (often subtle) perceptual changes is a challenging task. It requires the observation of changes in neural responses co-occurring with changes in perception without concomitant stimulus-induced changes in neural activity. The latter form a common confound in studies attempting to bridge the gap between neural responses and perception (Logothetis and Schall, 1989). While these perceptual modulations have been successfully captured in human studies, the effects are usually small. They can for instance be seen by means of decoding of the responses over larger regions using multi-voxel pattern analysis (MVPA) (Kilian-Hütten et al., 2011; Riecke et al., 2018; Muckli et al., 2015). Especially, for rhythm perception, a phenomenon evolving over time, these methods may prove important for future avenues of investigation, as they are potentially sensitive to neural differences at finer spatial or temporal scales (Vizioli et al., 2018).

Investigating these subtle perceptual changes at high spatial resolution is even more challenging, as the signal to noise ratio (in fMRI) is known to decrease with increased spatial resolution. Therefore, methodological considerations guided our choices. First, to maximize sensitivity, we chose for a GE EPI BOLD acquisition, which we know leads to a contribution of large draining veins (Polimeni et al., 2010). In the literature several approaches have been proposed to tackle the draining vein problem of GE EPI BOLD. The most straightforward option considers the microvascular task-related changes as (linear) modulations on top of a constant macrovascular contribution. The latter can thereby be controlled through subtraction of closely matched task-conditions, similar to a linear de-trending (Fracasso, Luijten, Dumoulin & Petridou, 2018). Alternative acquisition techniques other than GE EPI are less sensitive to large draining veins, (e.g. 3D-GRASE (Kemper, De Martino, Yacoub & Goebel, 2016; De Martino et al., 2013) or VASO (Huber et al., 2017), yet these acquisitions suffer from a larger decrease in sensitivity relative to GE EPI and it is an open question, whether we would be able to capture the subtle perceptual effects shown in the present thesis. From a methodological perspective, studies employing VASO in auditory areas will be promising to probe the specificity of laminar responses in auditory areas, with the potential to assess whole-brain submillimeter functional resolution and layer-dependent functional connectivity (Huber et al., 2021). Second, after an initial piloting with a sparse design, we decided to compromise the ability to present sounds in long periods of silence (as we were trying to do with the sparse design) to instead favor statistical power by reducing the silent gap (and collecting more volumes for the statistical analysis). Finally, we had to adjust task difficulty online during scanning, such that an approximately equal numbers of presented targets would be detected or undetected by participants.

Conscious of these methodological limitations, we see a large potential in the use of behaviorally driven layer fMRI as a tool to empirically test the functional relevance of cortical layers (for perception) in humans. In the introduction of this thesis, we highlighted the importance of behavior for neuroscience. Here, we would like to conclude that the opposite holds true as well. Neuroimaging, and fMRI in particular has been criticized for being expensive relative to its potential to inform cognitive theories (i.e. refine algorithms) (Colthart, 2006). Yet, neuroscientific methods are of important use when the behavioral results do not allow two competing models to be distinguished. In chapter 2 and 3 both rhythmic acoustic processing and cueing interval based temporal processing affect detection sensitivity in similar magnitudes. While psychoacoustics enables narrowing down potential algorithms and computations involved in perception of rhythms and temporal intervals, the question whether a common mechanism or different mechanism is implemented in the brain, may in the future ultimately only be resolved using (behaviorally driven) functional neuroimaging (Grahn et al., 2012). Future studies will have to investigate possible different mechanisms underlying different types of temporal expectations such as cueing and rhythms. We speculate here that a network involving the cerebellum may underlie interval based processing while the basal ganglia and frontal network for rhythm processing (Breska & Ivry, 2018; Breska & Ivry, 2021; Kotz, Schwartze, Schmidt-Kassow, 2009). Determining the origin of the here observed feedback signal may be achieved by adjusted slab placement, including either frontal regions, motor regions or cerebellar regions. Specifically, interactions between auditory and motor regions and their role in rhythm processing (Morillon, Hackett, Kajikawa & Schroeder, 2015), as fMRI studies have shown increased coupling during rhythm processing (Chen, Zatorre, Penhune, 2006; Grahn & Rowe, 2009). In addition, future studies are needed to further probe the frequency-specificity of the observed detection effect. This could be achieved by employing a similar paradigm, but incorporating a spectral detection task instead of a temporal detection task. Changing the nature of the task, might increase the relevance of finer spectral analysis of sounds, possibly by engaging a spectrally-specific detection mechanism (Lakatos et al., 2013; O’Connell et al., 2014; De Martino et al., 2015; Riecke et al., 2018).

Concluding remarks

In sum, the present thesis combined investigations of behavior and investigations of the brain, attempting to link the two. The degree of insight we can get from an experiment is not limited by the number of neurons we can record from (the spatial resolution we can achieve), but by the quality of the mapping we can create between internal brain states and behavioral reports. This thesis makes a leap forward in meeting both of these conditions, as we not only show an effect at the resolution of the cortical laminar response profile, but directly link this to behavior.

Samenvatting

In drie empirische studies onderzochten wij de gedragsmatige gevoeligheid en reactietijden voor ritmes en andere vormen van temporele structuur alsook de neurale correlaten die ten grondslag liggen aan de detectie van target geluiden ingebed in ritmische stimuli. In **Hoofdstuk 2** ontwikkelde we een gedragsmatig paradigma waarin we de bijdrage van temporele snelheden en draaggolven van geluiden onderzochten evenals hun interactie met de detectie van temporeel verschoven targets. Dit paradigma werd toegepast in **Hoofdstuk 3** om de effecten van temporele structuur en de kennis van voorspelbaarheid op gedragsmatige gevoeligheid en reactietijden te onderzoeken. Ten slotte hebben we in **Hoofdstuk 4** de neurale correlaten onderzocht die ten grondslag liggen aan de detectie van temporeel verschoven targets terwijl de aandacht gericht is op ritmische geluiden. We hebben laag-specifieke responsen in auditieve gebieden onderzocht met 7T fMRI. Terwijl de specifieke resultaten van de studies worden besproken in de aparte hoofdstukken heeft deze sectie als doel om de resultaten samen te vatten en te integreren. Speculaties over de rol van gerichte aandacht in ritmische perceptie worden voorgesteld.

De hersenen gebruiken regelmatigheden in de omgeving om toekomstige sensorische input te voorspellen. In auditie is ritme een regelmatige structuur. Huidige theorieën van neurale processen stellen voor dat het brein voorspellingen van toekomstige sensorische gebeurtenissen formuleert om perceptie te faciliteren (Friston, 2011). Dergelijke voorspellingen zijn gebaseerd op externe input van de omgeving en kunnen zowel over de inhoud als over de timing worden gevormd (Arnal & Giraud, 2012). Het is aangetoond dat voorspellingsmechanismen gedrag van mensen kunnen beïnvloeden. In auditie kan het vermogen om toekomstige geluiden te voorspellen de gedragsmatige gevoeligheid en reactietijden verbeteren als het geluid taakrelevant is (van Ede & Nobre, 2018). Het in lijn brengen van aanwezige oscillaties in auditieve gebieden met ritmische stimuli waar de aandacht op gericht is zou een onderliggend neuraal mechanisme kunnen zijn (dit is aangetoond bij niet-menselijke primaten; Lakatos et al., 2014). Dit ondersteund een beeld van ritmische facilitering als gerichte aandacht. Aan de ander kant heeft gedragsonderzoek laten zien dat het presenteren van geluiden in een voorspelbare periodieke (versus een niet-periodieke) structuur perceptuele voordelen met zich meebrengt ook al zijn de geluiden niet taakrelevant (Breska & Deouell, 2014). Het werd daarom gesuggereerd dat het onderliggende neurale mechanisme waarschijnlijk een (deels) automatische response is ten aanzien van ritmische sensorische input (Haegens & Zion-Golumbic, 2018).

In **Hoofdstuk 2** onderzochten we of sommige ritmes beter geschikt zijn dan andere ritmes om onderliggende neurale mechanismes aan te drijven. We hebben dit gedaan door de gedragsmatige gevoeligheid en reactietijden te onderzoeken op targets die ingebed zijn in ritmes. Bovendien controleerden we voor de frequentie van de draaggolf en de positie van de target binnen de sequentie. Onze resultaten lieten zien dat onder specifieke taak en stimulus omstandigheden langzame ritmische structuren van laagfrequente draaggolven (200 Hz) tot de grootste gedragsmatige voordelen in reactietijden en gevoeligheid leiden. De perceptuele voordelen van langzame snelheden is in overeenstemming met het idee dat een temporeel voorspellingsmechanisme wordt geïnitieerd door het motor systeem. Auditieve perceptie wordt beïnvloed door motor activiteit, met name in de delta band, waardoor temporele beperkingen worden opgelegd aan het samplen van sensorische informatie (Morillon et al., 2019).

In **Hoofdstuk 3** werd de temporele verschuivingsdetectietaak weer toegepast om het effect van verschillende temporele structuren te onderzoeken terwijl we de voorspelbaarheid van de targets constant hielden. We onderzochten bovendien of detectie en reactietijden worden beïnvloed door kennis over de voorspelbaarheid van de target. We lieten zien dat alhoewel ritmische structuur ons toestaat om proactief perceptuele gevoeligheid te verbeteren, temporele hints net zoveel gedragsmatige voordelen gaven in het geval wanneer er geen specifieke instructies werden gegeven over de relevantie van de temporele structuur. In een tweede experiment toonden we daarentegen aan dat expliciete instructies gedragsmatige voordelen van ritmes verder versterkten in vergelijking met een temporele hint. Dit suggereert dat aandacht de verwachting van periodieke stimulatie kan vergroten wat leidt tot versterkte gevoeligheid in target detectie. Deze resultaten laten zien dat aandacht kan worden gestuurd in de tijd (Nobre & Coull, 2010) en dat het kan worden gebruikt om ritmische effecten te versterken.

Wat duidelijk wordt van deze experimenten is dat aandacht interacteert met de sensorische verwerking van de ritmische stimulus zodat de aandacht richten op specifieke momenten in periodieke ritmes de perceptuele gevoeligheid versterkt. Ondanks de constante inspanning van proefpersonen om de aandacht te richten op target stimuli en deze te detecteren werden in sommige trials de targets niet gedetecteerd. In **Hoofdstuk 4** onderzochten we de corticale circuits die ten grondslag liggen aan ons vermogen om relevante geluidsveranderingen waar te nemen. We onderzoeken specifiek de feedforward en feedback processen binnen de auditieve hiërarchie tijdens detectie (door gedetecteerde en niet gedetecteerde targets te vergelijken). Terwijl vorige menselijke neuroimaging-onderzoeken de primaire auditieve cortex (PAC) identificeerde als het gebied dat bijdraagt aan target detectie bleef de exacte rol van feedforward en feedback processen speculatief (Giani et al., 2015; Gutschalk, Micheyl & Oxenham, 2008; Wiegand & Gutschalk, 2012). Wij onderzochten de spatiele meso-schaal non-invasief en observeerde een verhoging in activatie van de oppervlakkige lagen van de PAC suggestief voor een feedback signaal gerelateerd aan detectie. Dit is in overeenstemming met de bekende hiërarchische structuur van de auditieve cortex (Hackett et al., 2014).

Alles bij elkaar genomen laten de resultaten in deze thesis zien dat stimulus detectie proactief wordt geoptimaliseerd door gebruik te maken van temporele regelmatigheden in de omgeving die te vinden zijn in ritmes en temporele hints. De grootte van de gedragsmatige voordelen van voorspelbaarheid in een ritme en temporele hints in verhouding tot niet-periodieke condities is vergelijkbaar. Het richten van de aandacht door middel van expliciete instructies over de voorspelbaarheid van een target kan echter een modulerend effect hebben op het effect van temporele structuur op detectie (gevoeligheid) zodat in periodieke ritmes gevoeligheid is verbeterd relatief tot een temporele hint. Wanneer aandacht wordt gericht op zo’n stimulus en de target wordt gedetecteerd kan er een feedback signaal van deze detectie worden geobserveerd in de oppervlakkige lagen van de primaire auditieve cortex.

References

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. Trends in cognitive sciences, 16(7), 390-398.

Breska, A., & Deouell, L. Y. (2014). Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. Journal of Cognitive Neuroscience, 26(7), 1555–1571.

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. Proceedings of the National Academy of Sciences of the United States of America, 115(48), 12283–12288.

Breska, A., & Ivry, R. B. (2021). The human cerebellum is essential for modulating perceptual sensitivity based on temporal expectations. Elife, 10, e66743.

Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. Neuroimage, 32(4), 1771-1781.

Coltheart, M. (2006). What has functional neuroimaging told us about the mind (so far). Cortex, 42, 323–331.

De Martino, F., Moerel, M., Ugurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2015). Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proceedings of the National Academy of Sciences, 112(52), 16036–16041.

De Martino, F., Zimmermann, J., Muckli, L., Ugurbil, K., Yacoub, E., & Goebel, R. (2013). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. PloS one, 8(3), e60514.

Fracasso, A., Luijten, P. R., Dumoulin, S. O., & Petridou, N. (2018). Laminar imaging of positive and negative BOLD in human visual cortex at 7 T. Neuroimage, 164, 100-111.

Friston, K. (2010). The free-energy principle: a unified brain theory? Nature reviews neuroscience, 11(2), 127-138.

Giani, A. S., Belardinelli, P., Ortiz, E., Kleiner, M., & Noppeney, U. (2015). Detecting tones in complex auditory scenes. NeuroImage, 122, 203–213.

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: current findings and future perspectives. Topics in cognitive science, 4(4), 585-606.

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. Journal of Neuroscience, 29(23), 7540-7548.

Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. PLoS biology, 6(6), e138.

Hackett, T. A., de la Mothe, L. A., Camalier, C. R., Falchier, A., Lakatos, P., Kajikawa, Y., & Schroeder, C. E. (2014). Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. https://doi.org/10.3389/fnins.2014.00072Frontiers in neuroscience, 8, 72.

Haegens, S., & Golumbic, E. Z. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews, 86*, 150-165.

Huber, L., Finn, E. S., Chai, Y., Goebel, R., Stirnberg, R., Stöcker, T., ... & Poser, B. A. (2021). Layer-dependent functional connectivity methods. Progress in Neurobiology, 207, 101835.

Huber, L., Handwerker, D. A., Jangraw, D. C., Chen, G., Hall, A., Stüber, C., ... & Bandettini, P. A. (2017). High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. Neuron, 96(6), 1253-1263.

Kemper, V. G., De Martino, F., Yacoub, E., & Goebel, R. (2016). Variable flip angle 3D-GRASE for high resolution fMRI at 7 tesla. Magnetic resonance in medicine, 76(3), 897-904.

Kilian-Hütten, N., Valente, G., Vroomen, J., & Formisano, E. (2011). Auditory cortex encodes the perceptual interpretation of ambiguous sound. Journal of Neuroscience, 31(5), 1715-1720.

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. Cortex, 45(8), 982-990.

Lakatos, P., Musacchia, G., O'Connell, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. Neuron, 77(4), 750-761.

Logothetis, N., and Schall, D. (1989). Neuronal correlates of subjective visual perception. Science 245, 761–763.

Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron, 48(1), 139-148.

Moerel, M., Yacoub, E., Gulban, O. F., Lage-Castellanos, A., & De Martino, F. (2021). Using high spatial resolution fMRI to understand representation in the auditory network. Progress in Neurobiology, 207(January 2020), 101887.

Morillon, B., Hackett, T. A., Kajikawa, Y., & Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. Current opinion in neurobiology, 31, 230-238.

Morillon B, Arnal LH, Schroeder CE, Keitel A, Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception, Neuroscience and Biobehavioral Reviews (2019), doi:

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., ... & Yacoub, E. (2015). Contextual feedback to superficial layers of V1. Current Biology, 25(20), 2690-2695.

Nobre, A. C., & Van Ede, F. (2018). Anticipated moments: temporal structure in attention. Nature Reviews Neuroscience, 19(1), 34-48.

Nobre, A. C., & Coull, J. T. (Eds.). (2010). Attention and time. Oxford University Press, USA.

O'Connell, M. N., Barczak, A., Schroeder, C. E., & Lakatos, P. (2014). Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex. Journal of Neuroscience, 34(49), 16496-16508.

Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010). Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. Neuroimage, 52(4), 1334-1346.

Riecke, L., Peters, J. C., Valente, G., Poser, B. A., Kemper, V. G., Formisano, E., & Sorger, B. (2018). Frequency-specific attentional modulation in human primary auditory cortex and midbrain. NeuroImage, 174, 274-287.

Vizioli, L., Bratch, A., Lao, J., Ugurbil, K., Muckli, L., & Yacoub, E. (2018). Temporal multivariate pattern analysis (tMVPA): A single trial approach exploring the temporal dynamics of the BOLD signal. Journal of neuroscience methods, 308, 74-87.

Wiegand, K., & Gutschalk, A. (2012). Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. NeuroImage, 61(1), 62–69.

Appendix

English translation of
“ Über Die Vordere Quere Schläfenwindung
Des Menschlichen Großhirns”
by R.L. Heschl

Corresponding Manuscript:

Heynckes, M.*, Gulban, O, F.*, & De Martino, F. (under review at *Brain Multiphysics*). English translation of “ Über Die Vordere Quere Schläfenwindung Des Menschlichen Großhirns” by R.L. Heschl.

* both authors contributed equally

Abstract

This perspective accompanies a translation of R.L. Heschl's "Über Die Vordere Quere Schläfenwindung Des Menschlichen Großhirns", which translates to "On the anterior transverse temporal gyrus of the human cerebrum". The original manuscript reports an anatomical description of the transverse temporal gyrus, Heschl's gyrus, in humans. Within this description, Heschl reports a detailed description of a particular morphological shape deviation, one in which the superior temporal gyrus is split. Here, we contextualize the significance of said diverging morphology for modern neuroscience, specifically for but not limited to auditory human neuroimaging, accompanied by the full English translation of the original manuscript.

Keywords

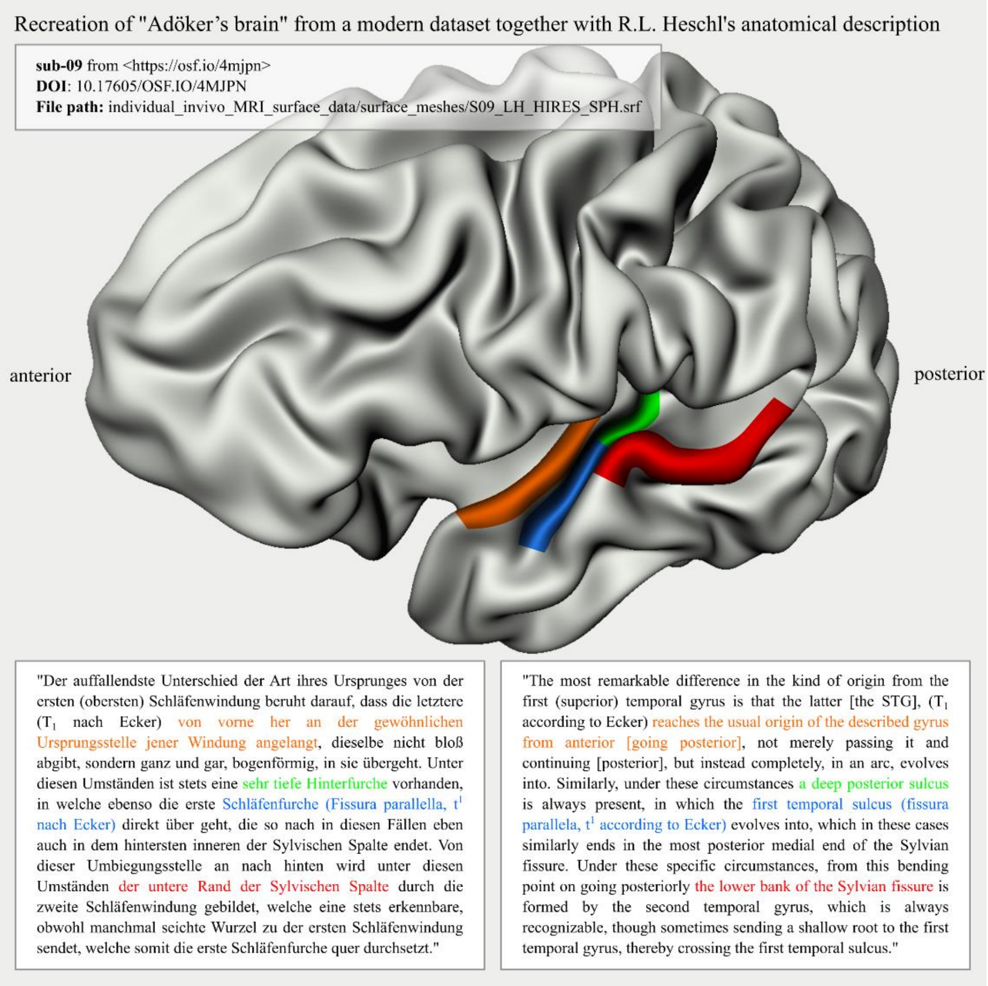
human, cortex, auditory, anatomy, superior temporal gyrus, Heschl's gyrus

Abbreviations

STG: Superior temporal gyrus: (referred to as T_1)
HG: Heschl's gyrus (referred to by Heschl as T^1)
MRI: magnetic resonance imaging

Introduction

Every neuroscientist has encountered the name “Heschl’s gyrus” (also known as “transverse temporal gyrus”) as one of the main human brain anatomical landmarks which also happens to host the primary auditory cortex [1]. However, how many neuroscientists know where the name “Heschl’s gyrus” originates from? Richard Ladislaus Heschl (1824-1881) was an Austrian anatomist who reported his observations on the human temporal gyrus in a manuscript published on the occasion of the 25-year anniversary celebration of the Vienna psychiatric hospital [2]. Within this manuscript, R.L. Heschl described the work he devoted the best years of his life to: observations over the anatomy of the transverse temporal gyrus he gathered from dissecting an astonishing number of human brains (632 male, 455 female). Interestingly, the work reports a particularly strong variation of shape in the human superior temporal cortex, one in which the superior temporal gyrus (STG) continues in the transverse temporal gyrus forming and arc and thus splitting the STG (in the anterior to posterior direction) with an intermediate sulcus. This variation occurs in approximately 10% of the population (see **Figure 1** and **Table 1**). Heschl pointed out that the contemporary anatomist Burdach [3], might even have mistaken this particular shape to be the most frequently occurring shape of the transverse temporal gyrus. With great foresight, R.L. Heschl himself noted that, in his era, the usefulness of his work might have been limited; yet, he noted, it would surely be appreciated in the future. Indeed, 144 years later, his observations still remain relevant. Here, we first briefly contextualize R.L. Heschl’s work within the modern neuroimaging framework, and then proceed to our English translation of his original manuscript written in German. In doing so, we hope to rekindle interest in the “strong variation of shape in the human superior temporal cortex” mentioned by R.L. Heschl.



Recreation of "Adöker’s brain" from a modern dataset together with R.L. Heschl's anatomical description

sub-09 from <https://osf.io/4mjpn>
DOI: 10.17605/OSF.IO/4MJPN
File path: individual_invivo_MRI_surface_data/surface_meshes/S09_LH_HIRES_SPH.srf

"Der auffallendste Unterschied der Art ihres Ursprunges von der ersten (obersten) Schläfenwindung beruht darauf, dass die letztere (T_1 nach Ecker) von vorne her an der gewöhnlichen Ursprungsstelle jener Windung angelangt, dieselbe nicht bloß abgibt, sondern ganz und gar, bogenförmig, in sie übergeht. Unter diesen Umständen ist stets eine sehr tiefe Hinterfurche vorhanden, in welche ebenso die erste Schläfenfurche (Fissura parallella, t^1 nach Ecker) direkt über geht, die so nach in diesen Fällen eben auch in dem hintersten inneren der Sylvischen Spalte endet. Von dieser Umbiegungsstelle an nach hinten wird unter diesen Umständen der untere Rand der Sylvischen Spalte durch die zweite Schläfenwindung gebildet, welche eine stets erkennbare, obwohl manchmal seichte Wurzel zu der ersten Schläfenwindung sendet, welche somit die erste Schläfenfurche quer durchsetzt."

"The most remarkable difference in the kind of origin from the first (superior) temporal gyrus is that the latter [the STG], (T_1 according to Ecker) reaches the usual origin of the described gyrus from anterior [going posterior], not merely passing it and continuing [posterior], but instead completely, in an arc, evolves into. Similarly, under these circumstances a deep posterior sulcus is always present, in which the first temporal sulcus (fissura parallela, t^1 according to Ecker) evolves into, which in these cases similarly ends in the most posterior medial end of the Sylvian fissure. Under these specific circumstances, from this bending point on going posteriorly the lower bank of the Sylvian fissure is formed by the second temporal gyrus, which is always recognizable, though sometimes sending a shallow root to the first temporal gyrus, thereby crossing the first temporal sulcus."

Figure 1. Recreation of “Adöker’s brain” referenced by R.L. Heschl from a publicly accessible dataset. The original human brain mold seems to be lost in history. Therefore, we have recreated a modern visual following the anatomical descriptions of R.L. Heschl. The matching cortical surface data is acquired from <https://osf.io/4mjpn/> [4] and visualized using Brainvoyager v22.2 [5].

Table 1. Occurrence of split STG by sex and hemisphere reported by R.L Heschl.

	Female (n = 455)		Male (n = 632)	
Bilateral	0	*(0%)*	3	*(0.47%)*
Only left hemisphere	19	*(4.2%)*	91	*(14.4%)*
Only right hemisphere	1	*(0.2%)*	2	*(0.3%)*

Modern contextualization

Standard neuroimaging practice using magnetic resonance imaging (MRI) often includes group averaging of individual brains. The common practice of group averaging involves non-linear alignments of whole brain volumes [6] or cortical surfaces (e.g. [7,8]), all of which are based on anatomical images. Studying brain function by anatomically aligning individual brains, to each other or to a common atlas, assumes that functional localization of different brain areas is correlated with - at least- macro anatomical landmarks [9]. Therefore, when a major divergence from the most common macro anatomical landmarks occurs, for that individual or hemisphere, we are running the risk of including a strong outlier in our group results. The relevance of these outliers will depend on the on the prevalence of such anatomical divergence and the sample size. For studies with large sample sizes, if the divergent macro anatomical shape is not prevalent, the effect of such outliers *might* be ignored. However, with lower sample sizes or higher prevalence, more attention needs to be paid to such outliers. Indeed, accounting for the macro anatomical variance between subjects improves correspondence between micro anatomically defined brain areas (defined post-mortem and in a small sample) [4]. This highlights that micro and macro characteristics are correlated to a certain extent [10] and that macro anatomical variations should not be ignored given the relation between micro anatomical characteristics and the functional properties of brain areas. This is particularly true for the human auditory cortex, having a large macro anatomical variability [1]. Therefore, the “strong variation of shape in the human superior temporal cortex”, reported in approximately 10% of the population, should be carefully considered in the (auditory) neuroimaging community. For example, the actual percentage of occurrence of this variation in large modern datasets [11,12], and how strong it impacts the group aligned functional (and anatomical) results reported within these studies is currently unknown.

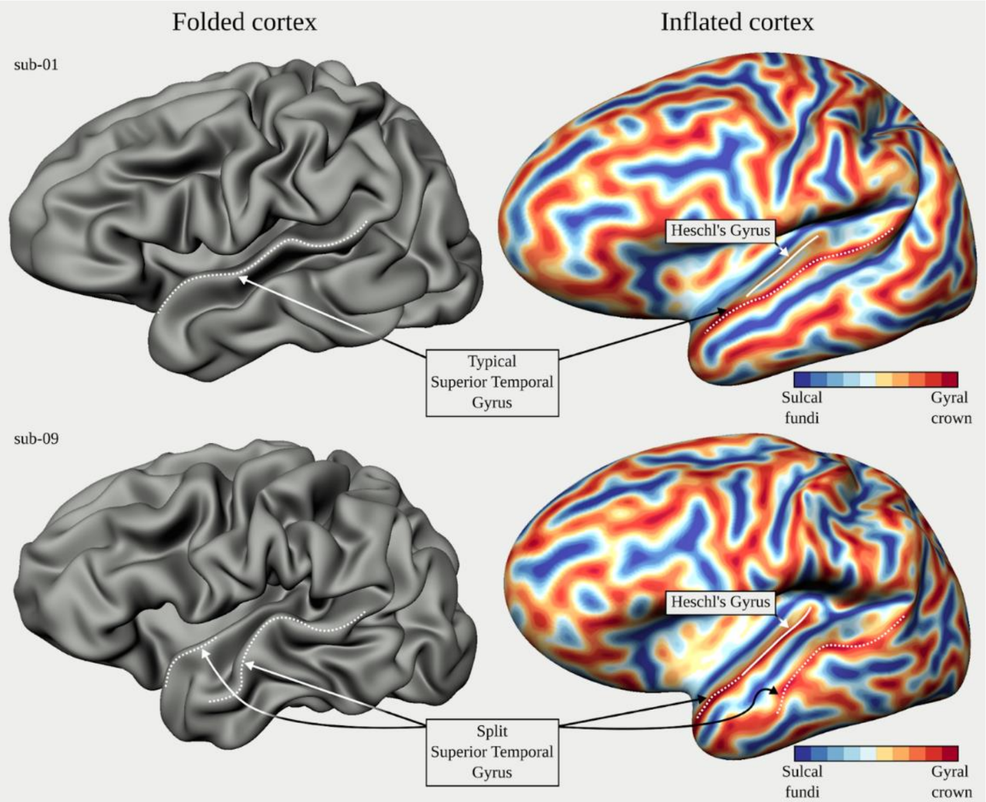


Figure 2: Comparison between the "typical superior temporal gyrus" shape to the "split superior temporal gyrus" described by R.L. Heschl (see **Figure 1**). Approximately 10% of humans have the "split superior temporal gyrus" morphology (see **Table 1**) which shows a deep sulcal pattern splitting the superior temporal gyrus towards its anterior end. The cortical surface data used in this figure is acquired from <https://osf.io/4mjpn/> [4] and visualized using Brainvoyager v22.2 [5].

Translation notes

We aimed to produce a translation that would give the reader a sense of the detail R. L. Heschl used in describing his work. German, especially around that time period, tends to be composed of very long sentences. Therefore, we have chosen to change this stylistic element in order to increase the fluidity of the text. Moreover, we employed a context-dependent translation, rather than a literal one, by reordering certain phrases to reduce illogical elements. In doing so, we aimed to not obscure or alter the intended meaning of the author to the best of our knowledge. Occasionally we included [Square brackets] to refer to the subject of a sentence and thereby improve readability.

Translation of: On the anterior transverse temporal gyrus of the human cerebrum

On the occasion of the 25-year anniversary celebration of the Vienna
Landesirrenanstalt (psychiatric hospital)
Published by
Dr. Richard L. Heschl.

One and a half years ago, I described under the name of the ‘anterior transverse temporal gyrus’ („der vorderen queren Schläfenwindung “) a gyrus progressing on the superior surface of the human temporal lobe, which had received little attention from the researchers who had previously been involved in the study of the gyri of the human brain. But, soon after I had inspected several fresh, unhardened brains regarding their gyri, this gyrus appeared to be a constant to me. I declared said gyrus as such and Bischoff and Rüdinger agreed.

In order to find this gyrus, the most convenient way is to dissect a brain, taken out of the skull as usual, first by means of a horizontal cut through the colliculi and cerebral peduncle into the cerebrum and cerebellum and then to dissect the former into its hemispheres by means of a sagittal cut through the corpus callosum and the center of the base. One takes a hemisphere, its convex surface facing down [note: the medial cut facing up], either on the inner surface of the left hand or on a previously wetted cup or similar board and commences by pulling off the meninges at the Gyrus Fornicatus [limbic lobe], continuing down to the edges of the upward-turned medial brain. Then the brain is turned onto the medial surface and one removes the meninges from all sides towards the Sylvian fissure. A tweezer is rarely needed to peel off [the meninges], sometimes a scissor is needed to cut off an artery running across a gyrus and sulcus. Usually it suffices to use the base of the thumb in combination with the tip of the index finger. Upon reaching the edge of the Sylvian fissure, one presses carefully with the edge of the thumb the operculum upwards and can then, in most cases, easily reach the parts of the pia meninx lying in the depth [of the Sylvian fissure]. Only towards the tip of the occipital lobe are tweezers required, as the meninges are most tender here. After removing the meninges, one becomes aware of the firmness of the brain mass provided by the meninges, since, upon removal of the meninges, even relatively solid hemispheres do not hold their shape, but form oval disks when laid on the table. After opening the Sylvian fissure and removing the meninges one can distinguish three surfaces forming the Sylvian fissure. A dorsal surface – the inferior surface of the front parietal part of the operculum with its transverse gyri; a medial surface – the lateral surface of the bulky protruding insula having 5-7 convolutions; and an inferior surface – the dorsal surface of the temporal lobe. The latter is usually of a triangular shape with curved sides, the lateral longest side is convex towards the outside, forming the dorsal bank of the superior temporal gyrus (inferior lip of the Sylvian fissure); the medial medium-length edge surrounds the insula in a short and sharp curve; and the most posterior and shortest edge connects the two posterior ends of the longer edges. The curve of its shape varies.

Upon uncovering said triangular surface, its sharpest angle pointing anteriorly, one notices that on the posterior half one or more gyri progress transversely across the surface from lateral, anterior to medial, posterior. There is always at least one anterior [transverse gyrus], very often a second and often a third [transverse gyrus]. Sometimes even a fourth or fifth [transverse gyrus] is present, either fanning out or running in parallel across the dorsal surface of the temporal lobe.

The first, or anterior, transverse temporal gyrus emerges gradually from the center of the lateral edge of the dorsal surface of the first temporal gyrus of the temporal lobe (T_1 according to Ecker[2]); being the longest of these gyri [the first transverse temporal gyrus] merging on occasion with the second transverse temporal gyrus and ending either by itself or in connection with the second in the most posterior corner of the Sylvian fissure, about 1cm from the entrance to the inferior horn [of the lateral ventricles] while the other end more laterally. This gyrus was first reported by

---

[2] Note: Ecker denoted gyri with uppercase letters and subscript numbers (T_1) and sulci with lower case letters and superscript numbering (t^1).

Burdach, Bau und Leben des Gehirns II Band S. 174 §. 208 in the following way: ‘The border (ora lobi inferioris) forming the edge of the tip of the inferior lobe is a … gyrus, progressing on its [the inferior lobe’s] lateral surface, below the Sylvian fissure upward, then progressing along the lateral edge of the dorsal surface and thereupon winds itself dorsally and posterior medial and finally joining with the posterior convolution of the operculum, being formed by the arcuate fasciculus.’ We will see shortly, that this somewhat unclear description does not correspond to the general case, the scheme, but is rather an exceptional case, as discussed later in more detail.

Otherwise, I only find the gyrus in question in Barkow’s, described as Gyrus magnus sinus operti, and the remaining brain anatomists ignore it, probably due to the investigation of hardened brains, in which the Sylvian fissure remains closed, and the gyrus invisible. I mentioned in my first publication that said gyrus is depicted in Henle, Reichert etc. but is otherwise neglected.

I termed this complex of gyri progressing along the dorsal surface of the temporal lobe, the (superior) transverse temporal gyri, and the most anterior of these 3-4 gyri as a constant, even if present in different forms. I refer to this gyrus with the letter T^1 [note: not Ecker’s T_1 referring to the STG].

I have sought it in no less than 1087 listed, and in hundreds of other not explicitly recorded, brains and have found it every time as a well-characterized gyrus of consistent shape. Its length varies with the depth of the Sylvian fissure: its height between 4 and 12 mm, its thickness is usually between 12-15mm. Occasionally the gyrus is anteriorly as well as posteriorly, and sometimes only posteriorly, delimited from the dorsal surface of the temporal lobe by a 6-8 mm deep sulcus, such that it rises with a 3-4mm wide foot – on the cross-section– sitting mushroom-caped shape on the lobe.

Concerning the histology of the discussed anterior transverse temporal gyrus, I can so far report the following: the direction of the fibers of the white matter is almost straight, or a slightly curved line drawn from the peduncles, from which piece by piece bundles consisting of few fibers branch off, suddenly making a sharp turn, rising up, entering into the cortical gray matter.

The latter has relatively small ganglion-cells located in the lower part, close to the white matter, such that the form of these cells ascribes this gyrus to the sensitive areas of the brain. A more detailed pursuit of its origin, progression and connection with the posteriorly located shorter temporal gyri results in extraordinary conditions, which bring about remarkable differences between male and female brains and between right and left hemispheres. The most remarkable difference in the origin of the anterior transverse temporal gyrus from the first (superior) temporal gyrus is that the latter [the STG], (T_1 according to Ecker) reaches the usual origin of the described gyrus from anterior [going posterior], not merely passing it and continuing [posterior], but instead completely, in an arc, evolves into it[3].

Similarly, under these circumstances a deep posterior sulcus is always present, in which the first temporal sulcus (fissura parallela, t^1 according to Ecker) evolves into, which in these cases similarly ends in the most posterior medial end of the Sylvian fissure. Under these specific circumstances, from this bending point going posteriorly, the lower bank of the Sylvian fissure is formed by the second temporal gyrus, which is always recognizable, though sometimes sending a shallow root to the first temporal gyrus, thereby crossing the first temporal sulcus. The described general pattern of the temporal gyri then becomes the arched transition of the first (lateral) [temporal gyrus] into the anterior transverse temporal gyrus, as was apparently meant by Burdach in his messages, as quoted above. One should accordingly assume that this description corresponds to the more common occurrence, however, this is not the case.

I mentioned earlier the number of investigated brains, whose results I noted down, to be 1087; this number is composed of 632 male and 455 female brains. To begin with, one would assume that the occurrence of that arched transition from T_1 [STG] in T^1 [HG] may be incidental. The comparison of these cases, however, reveals a regularity, differentiating between right and left as well male and female brain.

In the 632 male brains the described arched transition had occurred three times in both hemispheres, twice only in the right, whereas ninety-one times in the left

---

[3] See the mold of Adöker's brain (number 47 of the exhibition)

hemisphere. This occurrence therefore happens, in both hemispheres once in 210 cases (0.47%), in the right hemisphere once in 316 cases (0.31%), while in the left hemisphere already in (632:910 =) 6.9 cases once (14.4%). This cannot be a coincidence but must be considered as a peculiarity of the left hemisphere.

My tables further show that in 455 female brains the arched transition never occurred in both hemispheres, once in the right (0.2%), nineteen times in the left, that's once in (455:19=) 24 (actually 23.9) cases (4.2%).

Also this rather seldom occurrence of the arched transition in females is certainly no coincidence, even more so, as the ratios of the occurrences between right and left hemisphere in males being 1:17 (5:94) and in females being 1:19 are almost identical, meaning that: this peculiar formation occurs in males in absolute terms more frequently than in females, however, for the cases in which it does occur the ratio between right and left is almost the same for the brains of both sexes. These things will of course currently mostly have value as an anatomical fact rather than a practical value; the gyrus, however, is a solid constant as any other generally accepted constant.

Occasionally, in my first publication on this object, have I referred to the fact that said gyrus is also present early in intrauterine life. I repeat said comment here, because Herr Professor Schwalbe added a question mark in parenthesis in his paper in the yearly report for anatomy, without stating a reason.

I must object against such a method of criticism, as it is convenient but also cheap; I however, do not want to say a word about it anymore.

Therefore, I stated repeatedly, that the anterior superior temporal gyrus is easily visible as a smooth rise in the center of the lower bank of at that time fully opened Sylvian Fissure, already at a time when no other gyrus is visible, in the beginning of the 5th month. However, towards the end of the 5th month, at a time, when the three primary sulci on the lateral surface of the cerebrum are already present and the depth of the Sylvian fissure is increasing due to the substantial increase in brain substance, becomes said smooth rise in the center of the lower bank already considerable and continues in a smooth rising bulge, ending at the exact same place in the most posterior corner of the Sylvian fissure, as is the case later in the adult brain.

I have observed this behavior in total eight times in brains dating around the aforementioned period. It is obvious that if the earlier discussed arched transition from T_1 [STG] into T^1 [HG] ought to be present in the later brain, the lower bank of the Sylvian fissure must possess a crossing cleft; I have not seen this in early fetal brains. However, I did at 8 months old as well as older children observe the characteristic form.

I must speak a few words on Pansch's repeatedly urged view considering the sulci of the cerebrum. Pansch wants to put the emphasis concerning the conception of the various forms of cortex not on the gyri, but solely on the sulci, and especially their depth.

While this consideration [of the importance of sulci] certainly has its merits, it is in the by Pansch formulated exclusivity decidedly exaggerated. Certainly, two sulci correspond to every gyrus, as do two gyri to every sulcus; the deeper the sulcus, the taller the gyrus and so forth. Going back to the time of development, one finds that the first real and permanent sulci emerge from said undefined, in my experience, doubtfully shallow sulci. [This occurs] around the middle of the 5 month, so around the time, when the first noticeable movement of the child occurs, emerging even before the sulci, now generally referred to as primary sulci.

These sulci do not form because the surface of the brain sinks in, but rather, that it rises in other places, and the growing white matter pushes the young cortex in front of it. Thereby, the cerebrum divides itself in multiple larger areas; the formation of sulci results from a lagging growth of the white matter of strip-like areas in certain locations. A while later, the white matter again falls behind in growth in different places, and so this process repeats until the final form is reached. The investigation of newborn brains has taught me such, as the same shape can be observed as is the case later.

Therefore, I must, in contrast to Pansch, consider the formation of gyri as the active moment of the cortex: the various depths of the sulci merely correspond to the time in which the surrounding gyri in fetal life arose, given that the base of the sulci did not experience an unusual rise, falling outside the general laws of growth.

The typicality of a gyrus, therefore, has nothing to do with the formation of the surrounding sulci, as its height is solely dependent on its initial occurrence in fetal life. In other words, it is conceivable that a typical gyrus is low, and an atypical gyrus is high, as it is not required that the early emergence of a gyrus or sulcus is equal to its later presence, for instance around time of maturity. And in reality, the picture of gyri during development continuously changes. We will gain even better insights into these conditions, once the histogenetic circumstances of gyri are known more precisely. But already the aforementioned considerations should suffice, to restrict the claims by Pansch.

Simultaneously with the rise of the gyri and their gradual formation, commences the gradual considerable increase in white matter, of which branches continue into the substance of the gyri. Commonly, the formation of both matters occurs in parallel; there are also few cases, in which the white matter does not or only very slightly increase, namely that the compact mass of white matter of the centrum semiovale stays relatively small, while its protrusions entering the convolutions, while staying slim, rise and multiply, as is required for the later developing convolutions. This gives rise to a curious, little known phenomenon. A brain, in its entirety, stunted in volume, but with abundant, frequently winding, notably, narrow, but tightly packed convolutions, whose formation hardly allows one to identify the known pattern.

A dissection through such brains reveals that the sulci reach close to the lateral ventricles, as the semiovale center merely forms a very narrow white strip – a 2-3 mm thick disk – of which very long and narrow cut slices of white matter of the convolutions rise as elongated protrusions, which remind of the behavior of the cerebellum. I termed this condition microgyria, which in the cases I have observed has always occurred on the lateral and medial surfaces of the cerebrum but never on the ventral surface. Individuals with such brains die in early youth, the latest during their second year of life, are always mentally retarded, and suffer strongly from external hydrocephalus; such that the skull is not or only slightly smaller than a regular sized [skull]. Such cases may show, that the depth of sulci depend on the increase in mass of their white matter below; in my opinion it therefore only matters, whether a gyrus or sulcus is typical, that’s in certain location present in the developed brain or not, but not in primary or secondary instance how tall or shallow a convolution is, or from which time of development it originates.

Acknowledgements

We thank Karen Emmendorfer for her support in translating this document. A special thanks and acknowledgement of contribution go to Bruno Blüggel, from Greifswald University, Bruno Sperl from Vienna, and Eduard Winter from the Naturhistorisches Museum Vienna, who spent their time reaching out and exploring in various archives in the search of an imprint, termed “Adöker’s brain” mentioned in Heschl’s publication. Unfortunately, the imprint has been lost to time. However, we found the same split shape of the transverse temporal gyrus in our data and created a modern version of this imprint in Figure 1 to aid our readers in visualizing the original description of R.L. Heschl.

Funding

MH was funded by The Netherlands Organization for Scientific Research (NWO) under the Research Talent Grant 406.17.200. OFG is an employee of Brain Innovation. FDM has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. ERC - CoG 2020 - 101001270).

References

1. Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, *8*, 225. https://doi.org/10.3389/fnins.2014.00225
2. Heschl, R. L. (1878). *Ueber die vordere quere Schlafenwindung des menschlichen Grosshirns*. URL: <https://nbn-resolving.org/urn:nbn:de:gbv:9-g-4879281
3. Burdach, K. F. (1826). Vom Baue und Leben des Gehirns (Vol. 3). Dyk.
4. Gulban, O. F., Goebel, R., Moerel, M., Zachlod, D., Mohlberg, H., Amunts, K., & De Martino, F. (2020). Improving a probabilistic cytoarchitectonic atlas of auditory cortex using a novel method for inter-individual alignment. *ELife*, *9*, 1–29. https://doi.org/10.7554/eLife.56963
5. Goebel, R. (2012). BrainVoyager--past, present, future. *NeuroImage*, *62*(2), 748–756.
6. Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., & Brain Development Cooperative Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. Neuroimage, 54(1), 313-327. https://doi.org/10.1016/j.neuroimage.2010.07.033
7. Fischl, B., Sereno, M. I., Tootell, R. B. H., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*(4), 272–284. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4
8. Frost, M. A., & Goebel, R. (2012). Measuring structural–functional correspondence: Spatial variability of specialized brain regions after macro-anatomical alignment. *NeuroImage*, *59*(2), 1369–1381. https://doi.org/10.1016/j.neuroimage.2011.08.035
9. Rademacher, J., Rademacher, J., Caviness, V. S., Steinmetz, H., & Galaburda, A. M. (1993). Topographical variation of the human primary cortices: implications for neuroimaging, brain mapping, and neurobiology. *Cerebral Cortex*, *3*(4), 313–329. https://doi.org/10.1093/cercor/3.4.313
10. Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage, 13(4), 684-701.https://doi.org/10.1006/nimg.2000.0715
11. Marie, D., Jobard, G., Crivello, F., Perchey, G., Petit, L., Mellet, E., … Tzourio-Mazoyer, N. (2015). Descriptive anatomy of Heschl’s gyri in 430 healthy volunteers, including 198 left-handers. *Brain Structure & Function*, *220*(2), 729–743. https://doi.org/10.1007/s00429-013-0680-x
12. Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178. https://doi.org/10.1038/nature18933

Impact paragraph

The impact paragraph forms a reflection in layman's terms on the scientific impact of the results of the research described in the thesis and, if applicable, also the social impact already achieved or anticipated.

The empirical work presented in this thesis first and foremost was aimed at better understanding rhythm perception, (temporal) target detection and the nature of the underlying neural mechanisms. Therefore, the present findings are mostly of interest to the (auditory) neuroscience community, and the neuroimaging community developing methods we employed such as the small community centered around layer-fMRI. At time of writing approximately 180 papers have been published which employed this method[4]. In recent years this community has been growing (see Figure 1) and received increasing interest from the broader neuroscience community for the exciting possibility to use this approach to further our understanding of cortical processing in humans and address new (and old) research questions with greater detail. Of the studies published to date, only a quarter are neuroscience application studies, as opposed to methodological studies developing and benchmarking the technique.

---

[4] www.layerfmri.com/papers

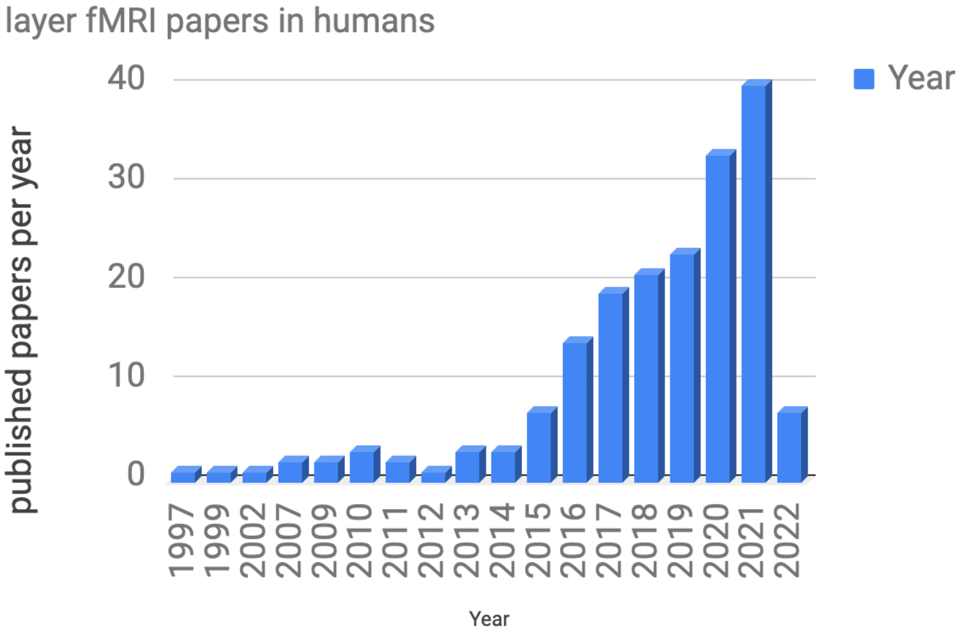


Figure 1. layer fMRI papers in humans (date: February 2022). The included publications fulfill the following criteria; they have a focus on functional imaging, cortical layers, human imaging and preprints are included. Data and visualization courtesy of Renzo Huber at www.layerfmri.com/papers.

Often it is assumed that fundamental research and application follow a one-way path. Fundamental research would lay the ground for applied questions, technologies, or treatments[5]. We would like to emphasize that this is usually not a one-directional path but bi-directional and involving trial and error. To illustrate this, as in the introduction of this thesis where we borrow once again David Marr's analogy of studying all the feathers of a bird to understand how flying works. The flight of airplanes preceded the understanding of the fundamental mechanisms - the laws of aerodynamics were in fact described only after people were already flying (Ridley, 2020). This is an example of why innovation in science is a two-way path. Indeed, often application is the parent of fundamental research. Techniques and processes that work are

---

[5] Even governments endorsed this reasoning. From 2013 to 2018 the impact chapter of this PhD thesis would have been referred to as valorization chapter, based on advice from the national commission of valorization, with the aim to outline the relevance of the present research for society with a focus on application (see UM promotiereglement version 3, July 2013).

developed first, but the full understanding comes later. As a result, science should not wait to make use of techniques until their full understanding is gained. On the contrary, to strengthen and improve techniques and further their fundamental understanding often field testing is needed.

Methods and analyses for fMRI in general, and layer-dependent fMRI specifically, partake in this iterative dance of (neuroscientific) research and innovation. While we may currently not yet precisely know what the (laminar) BOLD response signal reflects (Goense, Borhaus, & Logothetis 2016; Havlicek & Uludag, 2020), and many methodological challenges have been and are being faced (Polimeni et al., 2017); it should not keep fundamental research from asking neuroscientific questions and contributing insights while advancing method development (for fundamental examples see Huber et al., 2017 or De Martino et al., 2015; for a clinical example see for instance Stephan et al., 2017).

When considering the available human layer-specific fMRI publications, the brain modality most extensively studied by far has been vision (~70%), with a tenfold difference compared to audition (~7%)[6]. This fact mirrors a larger asymmetry in neuroscience (Hutmacher, 2019). Therefore, this thesis, particularly **Chapter 4** focusing on layer fMRI to probe basic mechanisms of the auditory system might inspire new research. In the future, the results of chapter 4 may contribute to computational modeling efforts in the auditory system. Such fundamental scientific results have an impact by contributing a puzzle piece of knowledge.

The importance of fundamental research is reinforced by the current zeitgeist at time of writing of this thesis. This chapter was written in February 2022, two years into the COVID-19 pandemic. Two scientific principles in particular, reflected in the presented thesis, are in common with some of the scientific values seen in response to the pandemic.

---

[6] www.layerfmri.com/papers

(1) Valuing basic and curiosity driven research for the sake of knowledge

The enormous impact of COVID-19 has resulted in significant efforts towards preventing and treating the novel coronavirus, particularly the development of a vaccine was of paramount importance. In the global fight against the pandemic mRNA vaccines represented a clear breakthrough. However, it would be naive to think that the severity of the global crisis alone fostered such an achievement. mRNA vaccines were preceded by decades of (publicly funded) research starting in the 1960s long before the current pandemic (Dolgin, 2021). Back then researchers were not thinking about using mRNA as a medical product and patenting the technology. Instead, they hoped to use it to interrogate basic molecular and genetic processes.

This shows that while there might not be any immediate scientific or social impact, scientists still need to do the work. Because we don’t know which knowledge will be helpful in the future and which not.

(2) Open science

Fundamental research and the scientific process is not completed by asking the necessary questions through research, a crucial element is the subsequent dissemination of the results. Open science aims to optimize the scientific process by facilitating access to various steps of the scientific process and their results. This can include access to analysis source code and data (*open source and open data*) via public repositories like Zenodo[7]. We tried to facilitate the dissemination of knowledge, by following several open science practices. The data, stimuli and analysis scripts underlying **Chapter 2** and **Chapter 3** are publicly available in a Zenodo data repository[8,910].

---

[7] https://zenodo.org/

[8] https://doi.org/10. 5281/zenodo.3695583

[9] https://doi.org/10.5281/zenodo.6473030

[10] https://doi.org/10.5281/zenodo.6472448

The most well-known form of open science and closest to more traditional forms of academic publishing is *open access*. The pandemic has acted as a catalyst for an enormous amount of COVID-related preprint publications being added to public archives (but see Besancon et al (2021) for a more critical evaluation). In general, if knowledge is accessible, impact is created. The scientific publications underlying this thesis are all or will be published open access. Similarly, this thesis will be made publicly available on the website of Maastricht University. In addition, to promote access to historic research, we translated the work ‘On the anterior transverse temporal gyrus of the human cerebrum’ by Richard Heschl, relevant to the content of this thesis and originally published in German in 1878.

To conclude we would like to describe the results of this thesis in everyday terms.

Chapter 2: Hearing something change in a rhythmic sound is easier than in a random sound. This detection is even better when the rhythm of the sound is slow and has a low tone frequency.

Chapter 3: However, rhythms by themselves don’t make you better in detecting a change than temporally cued events, unless somebody tells you.

Chapter 4: Sometimes we don’t hear changes in rhythms even though they are there. Some regions in the brain respond to these changes more when we actually heard them.

References

Besançon L, Peiffer-Smadja N, Segalas C, Jiang H, Masuzzo P, Smout C, et al. Open science saves lives: lessons from the COVID-19 pandemic. *BMC Med Res Methodol*. 2021;21(1):1–18.

Dolgin BE. the Tangled History. Nature. (2021);597:318–24. https://doi.org/10.1038/d41586-021-02483-w

De Martino, F., Moerel, M., Ugurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2015). Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proceedings of the National Academy of Sciences, 112(52), 16036–16041. https://doi.org/10.1073/pnas.1507552112

Goense J, Bohraus Y, Logothetis NK. (2016). fMRI at High Spatial Resolution: Implications for BOLD-Models. *Front Comput Neurosci.*;10:66. https://doi.org/10.3389/fncom.2016.00066

Havlicek M, Uludağ K (2019). A dynamical model of the laminar BOLD response. Neuroimage. 2020;204 https://doi.org/10.1016/j.neuroimage.2019.116209

Huber, L., Handwerker, D. A., Jangraw, D. C., Chen, G., Hall, A., Stüber, C., … Bandettini, P. A. (2017). High-Resolution CBV-fMRI Allows Mapping of Laminar Activity and Connectivity of Cortical Input and Output in Human M1. Neuron, 96(6), 1253-1263.e7. https://doi.org/10.1016/j.neuron.2017.11.005

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality?. *Frontiers in psychology*, 2246. https://doi.org/10.3389/fpsyg.2019.02246

Polimeni JR, Renvall V, Zaretskaya N, Fischl B. (2017). Analysis strategies for high-resolution UHF-fMRI data. *Neuroimage.*:1–25. http://dx.doi.org/10.1016/j.neuroimage.2017.04.053

Stephan, K.E., et al., Laminar fMRI and computational theories of brain function, NeuroImage (2017). https://doi.org/10.1016/j.neuroimage.2017.11.001

Ridley, M. (2020). How innovation works: And why it flourishes in freedom.

Acknowledgements

A PhD is always a team effort and with gratitude I look back to the support, guidance, inspiration, motivation, friendship and knowledge I found in this time. Without these people behind me, this dissertation would not exist.

Federico, we first met during the fMRI course. I'm thankful that you took a chance on me and suggested writing the NWO grant together, while starting from scratch with a behavioral paradigm during my master internship. I appreciate your deep and broad scientific insights. The right balance on how to progress on the data with precision, yet keeping an eye on the big picture. And something I got to appreciate especially in these last months: closing a project. I have learned a lot from you. Not only on scientific matters, but offering advice and helping me grow as a person throughout my PhD. Thank you for your support! I'm grateful we got to work together, as well as for the nice moments, like occasional dinners, outside of work.

Peter, you were my mentor during the research master program, and for me, unofficially remained my mentor throughout our PhD journey. I am thankful that I felt your door was always open to stop by for a coffee and chat, the occasional bike rides and the many social drinks we had in times when this was possible. These moments made all the difference. I'm also grateful we got to work so closely together in teaching the 'perception and attention' course. This time has taught me a lot and your rigorous approach to science and data, but also your view on education has left a mark on me. If everybody would do psychophysics experiments with an eye for detail as you … Thank you for believing in me and for leading by example what a rigorous scientist and academic looks like.

Elia, from writing the grant together with Peter and Federico up to our weekly zoom meetings marking the second half of my PhD, I could always count on you, coordinating many of the bureaucratic matters in the background, on questioning the data, sometimes sparking lively discussions in our meetings, and overall your optimistic nature that things are looking good when I felt more doubtful at times. All this has made a big difference to me. Thank you for supporting me and for the encouragement and helpful feedback.

Thank you all for having me experience an enjoyable PhD, where I learnt a lot and in which I always felt welcome with my ideas and input, where I had the opportunity to go in different directions (like the ECoG project, that for various reasons didn't commence), but most of all that I felt there was a space to make mistakes and learn from them. I don't take this for granted.

My **thesis committee**, thank you for taking the time to thoroughly review my thesis and discuss it with me during my defense.

Colleagues and friends

Alex, my paranymph, research master comrade, vegan master chef (I miss our weekly dinners!) and most important, great friend with whom I have grown together so much in the past years! Thank you for the many amazing years we shared, your relentless support in good and bad times. And last but not least, thank you for being my constant source of memes on being a PhD student and the many wines we drank together. **Sanne**, it's a full circle! Our scientific journey began with you being my supervisor when you were doing a PhD and I was just a young bachelor student, and it ends with me getting a PhD and you being my paranymph and great friend. I'll always remember our "conference vacation" to Tel Aviv and Marseille and the many times you and Mehrdad had me stay over when I needed a place to stay in Maastricht. I know it comes with baby-sitting duties from now on and I look forward ;-) **Agustin**, I'm so thankful for your friendship and your continuous support, much beyond Matlab, and even from the other side of the world. Distance and time don't matter and I'm so grateful for that! **Stefano**, thank you for our many bike adventures and our Italian/Cuban cooking nights and the many social drinks we joined (read: crashed) with Agustin at Oxfordlaan. **Alix**, thank you for our lockdown walks and book discussions, motivating me for 7AM workout classes, wonderful dinners together and for continuing them now in Amsterdam. I'm so excited for your future and I know you’ll be an amazing mom! **Milene**, thank you for the Italian dinners and our bike rides together. Limburg’s scenery will always be a cherished memory. **Kamil**, your presence in the department is missed. Thank you for the scientific, but also non-scientific discussions over the years. **Kirsten,** it’s time to admit it! You actually did a PhD in Maastricht, not Louvain. Thank you for the many koffietjes and wijntjes and creative crafty challenges. **Till,** I am thankful to have had your friendship and support, experiencing your grounding and calm advice over the years. **Felix,** thank you for the motivational (and let’s call them realistic) chats in the hallway. **Shanice**, since the bachelor we have walked the same path and in my perception you have always excelled at being a scientist. I wish you all the best for your future. **Sebastian**, after-hour scans, many lunches, post-work drinks and conference visits. They all become all the more enjoyable when you're around! How you approach doing science and ponder over data and analyses tells me that your future holds a lot for you and I look forward to seeing this happen! **Fabian**, thank you for multiple iterations of volunteering in the 7T as my pilot participant. Short notice and technical issues didn’t stop you from offering your help again and I much appreciate that! **Sven**, always a pleasure to run into you in the hallway and tell me enthusiastically about the latest device you built! Keep your inquisitive and creative mind! I would also like to thank my undergraduate teachers who taught me valuable lessons. In particular, **Arie** and **Stefan**, for being a source of guidance and opportunities, critical thinking skills and inspiration in my early (MARBLE) research attempts. I am still thankful that you made it possible for me to go abroad during the bachelor and do a Marble project upon returning, because you believed in me. **Minye**, **Giada**, **Linda**, **Yawen** thanks to all of you for being my officemates over the years. I always enjoyed sharing an office with you all, perfectly combining productivity and fun! Yawen, from sharing a physical office with Linda, to our virtual zoom office and co-writing sessions. Thank you for always having an open ear, for your support and more than once letting me explain my messy code to then realize while explaining where the problem was.

The current and past members of the **auditory group: Elia, Michelle, Giancarlo, Lars H., Isma, Agustin, Lars R., Federico, Martin, Fren, Valentin, Vittoria, Shruti, Martha, Niels, Michele, Lonike, Min, Xueying, Lidongsheng.** Thank you all for the many interesting discussions and creating an open, collaborative atmosphere. **Ben**, **Mathilde** and **Peter**, thank you for efforts to organize the MBIC graduate school.

And the **PALM group**
Hannah, Peter, **Vincent**, **Selma**, **Mark**, **Maryam**, **Gesa**, the many students Peter always managed to find, and some of whom became my research assistants. This group was always up for drinks and social get-togethers. I feel we entered the lockdown with lab drinks and met up as soon as it was possible again. **Mark** and Hannah, a special thanks for collaborating in our attempt of the ECoG project. Glimpses of working with patients have taught me a lot. Above all, how difficult obtaining patient data can be. **Hannah**, you're my reminder that while being fully immersed in a job in which you already write a lot, you can choose to spend even more time writing and that having a passion outside of science is a wonderful thing! Please always stop by for more bookstore visits in Amsterdam!

The **layer seminar**, what a special group of people! **Renzo**, **Faruk**, **Sebastian**, **Lonike**, **Alessandra**, **Till**, **Johannes**, **Yawen**, **Maite** as well as **Mahdi**, **Elaine**, **Kenshu** and **Jorie**. All from different sections within our department, yet united by the common interest in the methodology. Many coffee get-togethers, fruitful discussions, general support, data support and random things, feedback on presentations, open science, submissions for conferences and of course the many papers we discussed over the past years. I'm proud of what we co-created. A special thanks to Faruk; I ascribe the success of having a regularly occurring meeting in the early stages of the group to your tenacity.

To my **CN colleagues**, there are too many names to mention. I have always felt welcome and enjoyed working and socializing with you. I always looked forward to Christmas parties, department days and social drinks. A big shout out here to Christl for being the initiating force more than once using her great karaoke skills! **Christl,**

Riny, Jose and **Eva** thank you for your support and solving many of my administrative challenges over the years and always being kind and understanding. **Kwinten**, **Sean**, **Dienne** and everybody else having worked in the past years or still working at **Bandito's**. I was always proud to tell people in which building my office was located, as we had by far the best coffee and (vegan) lunch thanks to you! From **Scannexus**. **Liel**, thank you very much for your kind help! **Bianca** en **Esther**, **Rick**, **Job**, and **Chris**, for your technical support, and that I could always approach you for scanning related questions. **Kim, Lauren, Louisa, Anna, Jacek, Emil,** thank you for your help with collecting behavioral data, you have made a valuable contribution to the work presented here! A big thank you to the many amazing instructors of **UM sports!** For many years you have contributed to keeping me healthy and sane with great sports classes.

My life outside academia

Leonie, everyone needs a friend like you, bringing laughter, craziness, deep conversations and caring support into my life. Thank you for reminding me that it's normal to feel that a PhD is hard at times. **Jorrit**, thank you for bringing adventure into my life, inspiring me and always challenging me to do better. **Boy**, for your friendship, support and learning a lot from each other. **Dennis** and **Sandy,** our weekly dinners in your kitchen will always have a special place in my heart. **Niklas**, you are my reminder that you can study very different things, end up in very different places in the world and somehow still both know how to code in MATLAB. Yet, most of all I am grateful that despite the distance our friendship has maintained its essence since high school! I'm excited to see what the future will hold for you and Biljana. **Guido** for listening and always making me feel supported in my endeavors. You have taught me a lot. **Yvette** my early morning workout, co- and remote working buddy. But most of all, dear kind-hearted and supportive friend! How much I cherish that our paths crossed again after many years. Never will I forget spending a month together with you and **Mandy** secluded in beautiful France writing this book. **Lisa,** thank you for your friendship, support and exploring Amsterdam together! And for always watching over Nala and Oreo, especially when I was heading to Maastricht for a few days. You’re officially the best cat aunt in the world! **Tobi**, your tenacity to study new things, desire to learn and flow with life are admirable. Let’s go to the beach again soon. **Maastricht goes vegan** and **the Improv group.** Never would I have enjoyed my PhD time nearly as much, if it wasn’t for the wonderful people I got to meet here. What else does one need other than great company, great food and tons of laughter? Especially our MGV team, we have created and built on something great over the years, that I look back to with pride and a deep sense of joy.

An dieser Stelle einen nostalgischen Dank an **Hannah**, **Maya**, **Wenke** und **Becca**, oder besser gesagt *naknaknak*, und die glorreichen Zeiten in der Mockeborg, die diesem PhD vorausgingen und dazu beigetragen haben. Ich bin gespannt was die Zukunft für euch bereithält.

Helga und **Leonie**, ihr seid beide zu früh von dieser Welt gegangen und habt mein PhD Leben nicht mehr mitbekommen, umso mehr wart ihr ein wichtiger Teil meines Lebens davor. Dafür bin ich euch dankbar und denke oft an euch.

My deepest gratitude goes out to my family. My parents, **Ruth** and **Martin** and my brothers **Maik** and **Patrick.** Mama & Papa, ohne eure Unterstützung wäre es mir nie möglich gewesen so weit zu kommen. Ihr habt immer an mich geglaubt und mir so vieles ermöglicht. Diese Doktorarbeit ist mit euer Verdienst! Ich liebe euch sehr. **Oma Lene.** Vielen Dank für dein Vertrauen und deine Unterstützung, *‘Du machst das schon alles richtig Kind’*, bedeutet die Welt für mich. Danke an meine Familie; meine Tanten und Onkels, Cousins und Cousinen die mich motiviert und unterstützt haben, und mir immer das Gefühl geben, dass ich raus in die Welt gehen kann und gleichzeitig jederzeit zuhause willkommen bin.

There are many more people whom I could list here. Please know I am grateful, that we crossed paths during this special time in my life!

About the author

Miriam Heynckes was born on April 22nd 1993 in Nettetal, Germany. She completed her early education in Schwalmtal, at the Janusz-Korczak Realschule in 2009 and the Gymnasium St. Wolfhelm in 2012, interleaved by a semester at Long Prairie Grey Eagle High School in Minnesota, USA between 2009 and 2010. As first in her family she continued higher education at Maastricht University, where she obtained her Bachelor's Degree in Psychology and Neuroscience *cum laude* in 2015. In 2017, she completed her Master's degree Cognitive and Clinical Neuroscience *cum laude*, with a specialization in cognitive neuroscience. Her thesis focused on developing a behavioral paradigm for auditory perceptual processing of rhythmic stimuli under supervision of Federico De Martino and Elia Formisano. In 2017, Miriam together with Elia Formisano, Peter De Weerd and Federico De Martino successfully applied for a research talent grant from the Dutch Research Council (NWO) to continue her research during her PhD, which she finished in 2022. Miriam currently lives in Amsterdam.

List of publications

Peer-reviewed publications

Heynckes M, De Weerd P., Valente G., Formisano E., De Martino F., (2020) Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences. *PLOS One*;15(6):1–7

Heynckes, M. (2015). The predictive vs. the simulating brain: a literature review on the mechanisms behind mimicry. *Maastricht Student Journal of Psychology and Neuroscience,* 4.

Manuscripts submitted or in preparation

Heynckes, M., Gulban, O, F., & De Martino, F. (under review at *Brain Multiphysics*). English translation of “ Über Die Vordere Quere Schläfenwindung Des Menschlichen Großhirns” by R.L. Heschl

Heynckes, M., Hoffmann, K., Formisano, E. & De Martino, F., De Weerd, P. (*submitted*). Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues.

Heynckes, M., De Weerd, P., Formisano, E. & De Martino, F. (*in preparation*). Layer-specific correlates of detected and undetected auditory targets during attention

Public Datasets

Heynckes M., de Weerd P., Valente G., Formisano E., & De Martino, F (2020). Data & analysis scripts of " Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences" (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3695583

Koiso, K., Dresbach, S., Wiggins C. J., Gulban O. F., Mueller A. K., **Heynckes M.**, Chai, Y., Miyawaki, Y., Poser, B., & Huber, R. (2021). Whole brain layer-fMRI: An open dataset for methods benchmarking. https://doi.org/10.18112/openneuro.ds003216.v2.1.1

Heynckes, M. (2022). Presentation scripts of "Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues" (Version 1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.6472448

Heynckes, M. (2022). Data and analysis scripts of "Predictability awareness rather than mere predictability enhances the perceptual benefits for targets in auditory rhythms over targets following temporal cues" (Version 1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.6473030

Conference Contributions

Heynckes, M., Formisano, E., De Weerd, P., & De Martino, F. (2021). Temporal deviant detection in human auditory cortex using high-field fMRI. ISMRM (virtual). Poster talk.

Heynckes M, de Weerd P., Valente G., Formisano E., & De Martino, F. (2019). Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences. BryCoCo (Tel Aviv, Israel). Poster.

Heynckes M, de Weerd P., Valente G., Formisano E., & De Martino, F. (2019). Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences. The Predictive Brain Conference (Marseille, France). Poster.