



HAL
open science

Rhythmic neural and perceptual dynamics in the auditory system

Benedikt Zoefel

► **To cite this version:**

Benedikt Zoefel. Rhythmic neural and perceptual dynamics in the auditory system. Cognitive Sciences. Université Toulouse 3 - Paul Sabatier, 2023. tel-04757183

HAL Id: tel-04757183

<https://hal.science/tel-04757183v1>

Submitted on 28 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paul Sabatier – Toulouse III
Ecole Doctorale CLESCO

Mémoire présenté en vue de l'obtention de
Habilitation à Diriger des Recherches
Date de soutenance : 2 juin 2023

Rhythmic neural and perceptual dynamics
in the auditory system

Benedikt Zoefel

Composition du jury :

M. Pierre PAYOUX (Président du Jury, Université Toulouse III Paul Sabatier)

Mme Isabelle BERRY (Université Toulouse III Paul Sabatier)

Mme Emmanuelle COURTIOL (Université Lyon I)

M. Gregor THUT (University of Glasgow)

M. Nathan WEISZ (University of Salzburg)

M. Bruno GIORDANO (Université d'Aix-Marseille)

Mme Andrea MARTIN (Max Planck Institute Nijmegen)

Centre de Recherche Cerveau et Cognition
(CerCo) - CNRS UMR 5549

TABLE OF CONTENT

1	SUMMARY	3
2	CURRICULUM VITAE	4
3	INTRODUCTION	9
3.1	Auditory Oscillations – an attentional filter that operates in time.....	9
4	PREVIOUS RESEARCH	12
4.1	PhD work (2012-2015).....	12
4.1.1	Perceptual impact of information loss during rhythmic sampling of stimulus input.....	12
4.1.2	Neural entrainment to speech without slow changes in spectral energy.....	14
4.1.3	Neural oscillations structure acoustic information.....	21
4.1.4	Summary: PhD work.....	23
4.2	Post-doc work (2016-2019)	24
4.2.1	Causal role of neural entrainment for speech processing	24
4.2.2	Optimised methods to test for oscillatory phase effects	29
4.2.3	Intelligibility improves the perception of speech rhythm	31
4.2.4	Summary: Post-doc work	33
4.3.	First steps as principal investigator (2019 – 2022)	34
4.3.1	Entrainment echoes in the MEG	35
4.3.2	Entrainment echoes after tACS	37
4.3.3	Entrainment echoes in auditory perception	40
5	FUTURE RESEARCH	45
5.1	Rhythmic entrainment echoes.....	45
5.1.1	Entrainment echoes in the cerebellum	45
5.1.2	Entrainment echoes after tACS	48
5.2	Auditory Oscillations: Useful or harmful?	49
5.2.1	Effects of relevance and predictability on phasic modulation of auditory perception.....	50
5.2.2	Auditory phase-reset.....	52
5.3	Manipulating oscillations to control auditory perception and attention ..	54
5.3.1	Modulating speech perception with tACS.....	54
5.3.2	Neural entrainment vs alpha oscillations: Two competing “modes” of auditory stimulus processing?	59
5.3.3	Modulating speech perception with neurofeedback.....	60
5.3.4	Modulating speech perception with AM-kTMP	65
6	CONCLUSION: PERSPECTIVES IN FUNDAMENTAL, APPLIED, AND CLINICAL RESEARCH.....	70
7	REFERENCES.....	72

1 SUMMARY

The human brain is continuously bombarded with sensory information and is therefore in need of a “tool” to select and filter its input. The overarching goal of my research is to understand whether and how rhythmic brain activity, also termed *neural oscillations*, can be used for this purpose. These oscillations lead to a rhythmic alternation between moments that “open the gate” for important events needing attention, and moments that inhibit the processing of irrelevant, to-be-ignored information. My focus is on the auditory system, as a modality that is not only currently under-investigated with respect to oscillatory processes, but also seems to be different from others. During my PhD, I explored and confirmed the possibility that, due to its rapidly changing input, the auditory system is particularly vulnerable to information loss that can occur when the “gate is closed”. I also developed the hypothesis that the often-described synchronisation between neural oscillations and rhythmic stimuli (“*neural entrainment*”) avoids such information loss by aligning high-excitability moments (“open gates”) with relevant events in the stimulus. I showed that neural entrainment to human speech persists even in the absence of pronounced changes in spectral energy, suggesting the involvement of a higher-level, potentially linguistic, process. During my post-doc, I developed brain stimulation methods to manipulate neural entrainment. I showed that a change in entrainment leads to changes in neural and perceptual responses to speech, further corroborating the causal, specific role that neural entrainment plays for the perception of speech. Together with several students and researchers I supervise, I am currently applying a state-of-the-art combination of neuro-/electrophysiological and brain stimulation methods to further develop these lines of research. One of our aims is to test the hypothesis that neural oscillations in the auditory system can be “useful” or “harmful” and can therefore be “turned on and off”, depending on the circumstances the system is confronted with. This approach will address the question of how neural oscillations operate in a world that changes rapidly and is not always predictable. We also develop novel protocols to control auditory perception and attention, via a manipulation of neural oscillations. Such protocols are not only designed to advance fundamental research, but are also important to improve conditions associated with a malfunctioning of oscillatory processing. They might also inspire future clinical and technological applications for everyday life scenarios in which sustained attention is critical.

2 CURRICULUM VITAE

Candidate

ZOEFEL, Benedikt, born on 7th September 1985 in Tübingen (Germany)

Education

2015	PhD in Neuroscience Paul Sabatier University, Toulouse , France Advisor: Rufin VANRULLEN
2011	M.Sc. in Integrative Neuroscience Otto-von-Guericke University, Magdeburg , Germany Advisor: Peter HEIL
2009	B.Sc. in Psychology, Otto-von-Guericke University, Magdeburg , Germany Advisor: Christoph S. HERRMANN

Academic Positions

2020 –	CNRS Researcher (Chargé de Recherche) Centre de Recherche Cerveau et Cognition (CerCo), Toulouse, France
2016 – 2019	Postdoctoral Researcher MRC Cognition and Brain Sciences Unit, University of Cambridge, UK
2014 – 2015	Visiting Scholar Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, US

Grants, Awards & Scientific Honours

2023	Research Grant (€47,250) <i>Fondation Fyssen</i>
2021	Research Grant (€340,000) <i>Agence Nationale de la Recherche (ANR) - JCJC</i> Research Grant (€294,000) <i>Fondation pour l'Audition</i>
2021 2019	Selected for interview stage of ERC Starting Grant application Merit Award <i>Society for the Neurobiology of Language, US</i>
2018 – 2019	Marie Skłodowska-Curie Individual Fellowship (€195,455) <i>European Union</i>
2018	Leverhulme Small Grant (£9,980) <i>British Academy</i>
2017	Research Workshop Grant (£4,500) <i>Experimental Psychology Society, UK</i>
2016 - 2017	Postdoctoral Fellowship (€78,000) <i>German Academic Exchange Service (DAAD)</i>
2012 – 2015	PhD Scholarship (€58,000) <i>Studienstiftung des Deutschen Volkes, Germany</i>
2009 – 2011	M.Sc. Scholarship (€13,500) <i>Studienstiftung des Deutschen Volkes, Germany</i>

List of Publications

1. L'Hermite, S., and **Zoefel, B. (2022)**. Rhythmic entrainment echoes in auditory perception. *bioRxiv*, doi: 10.1101/2022.12.07.519456
2. **Zoefel, B.**, and Kösem, A. (2022). Neural Dynamics: Speech is special. *PsyArXiv*, doi: 10.31234/osf.io/eyrtb
3. **Zoefel, B.**, Gilbert, R.A., and Davis, M.H. (2023). Intelligibility improves the perception of speech rhythm. *PLOS One*, 18, e0279024.
4. Ekhtiari, H., Ghobadi-Azbari, P., Thielscher, A., [...], **Zoefel, B.**, [...] (2022). A checklist for assessing the methodological quality of concurrent tES-fMRI studies (ContES checklist): a consensus study and statement. *Nat. Protoc.* 17, 596–617.
5. van Bree, S., Alamia, A., and **Zoefel, B. (2022)**. Oscillation or not-Why we can and need to know (commentary on Doelling and Assaneo, 2021). *Eur. J. Neurosci.* 55, 201–204.
6. van Bree, S. van, Sohoglu, E., Davis, M.H., and **Zoefel, B. (2021)**. Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLOS Biol.* 19, e3001142.
7. Reddy, L., Self, M.W., **Zoefel, B.**, Poncet, M., Possel, J.K., [...] (2021). Theta-phase dependent neuronal coding during sequence learning in human single neurons. *Nat. Commun.* 12, 4839.
8. Reddy, L., **Zoefel, B.**, Possel, J.K., Peters, J., [...] (2021). Human Hippocampal Neurons Track Moments in a Sequence of Events. *J. Neurosci.* 41, 6714–6725.
9. **Zoefel, B. (2021)**. Visual speech cues recruit neural oscillations to optimise auditory perception: Ways forward for research on human communication. *Curr. Res. Neurobiol.* 2, 100015.
10. **Zoefel, B. (2019)**. Auditory Perception: A Rhythm Reflecting the Past. *Curr. Biol.* 29, R1318–R1320.
11. **Zoefel, B.**, Allard, I., Anil, M., and Davis, M.H. (2019). Perception of Rhythmic Speech Is Modulated by Focal Bilateral Transcranial Alternating Current Stimulation. *J. Cogn. Neurosci.* 32, 226–240.
12. **Zoefel, B.**, Davis, M.H., Valente, G., and Riecke, L. (2019). How to test for phasic modulation of neural and behavioural responses. *NeuroImage* 202, 116175.
13. **Zoefel, B. (2018)**. Speech Entrainment: Rhythmic Predictions Carried by Neural Oscillations. *Curr. Biol.* 28, R1102–R1104. 10.1016/j.cub.2018.07.048.
14. Riecke, L., and **Zoefel, B. (2018)**. Conveying Temporal Information to the Auditory System via Transcranial Current Stimulation. *Acta Acust. United Acust.* 104, 883–886.
15. **Zoefel, B.**, Archer-Boyd, A., and Davis, M.H. (2018). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr. Biol.* 28, 401-408.e5.

16. **Zoefel, B.**, ten Oever, S., and Sack, A.T. (2018). The Involvement of Endogenous Neural Oscillations in the Processing of Rhythmic Input: More Than a Regular Repetition of Evoked Neural Responses. *Front. Neurosci.* 12.
17. **Zoefel, B.**, Costa-Faidella, J., Lakatos, P., Schroeder, C.E., and VanRullen, R. (2017). Characterization of neural entrainment to speech with and without slow spectral energy fluctuations in laminar recordings in monkey A1. *NeuroImage* 150, 344–357.
18. **Zoefel, B.**, and Davis, M.H. (2017). Transcranial electric stimulation for the investigation of speech perception and comprehension. *Lang. Cogn. Neurosci.* 32, 910–923.
19. **Zoefel, B.**, and VanRullen, R. (2017). Oscillatory Mechanisms of Stimulus Processing and Selection in the Visual and Auditory Systems: State-of-the-Art, Speculations and Suggestions. *Front. Neurosci.* 11.
20. **Zoefel, B.**, and VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage* 124, 16–23.
21. **Zoefel, B.**, and VanRullen, R. (2015). The Role of High-Level Processes for Oscillatory Phase Entrainment to Speech Sound. *Front. Hum. Neurosci.* 9.
22. **Zoefel, B.**, and VanRullen, R. (2015). Selective Perceptual Phase Entrainment to Speech Rhythm in the Absence of Spectral Energy Fluctuations. *J. Neurosci.* 35, 1954–1964.
23. **Zoefel, B.**, Reddy Pasham, N., Brüers, S., and VanRullen, R. (2015). The ability of the auditory system to cope with temporal subsampling depends on the hierarchical level of processing. *Neuroreport* 26, 773–778.
24. VanRullen, R., **Zoefel, B.**, and Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130214.
25. **Zoefel, B.**, and Sokoliuk, R. (2014). Investigating the Rhythm of Attention on a Fine-Grained Scale: Evidence from Reaction Times. *J. Neurosci.* 34, 12619–12621.
26. Heil, P., Verhey, J.L., and **Zoefel, B.** (2013). Modelling detection thresholds for sounds repeated at different delays. *Hear. Res.* 296, 83–95.
27. **Zoefel, B.**, and Heil, P. (2013). Detection of Near-Threshold Sounds is Independent of EEG Phase in Common Frequency Bands. *Front. Psychol.* 4.
28. **Zoefel, B.**, Huster, R.J., and Herrmann, C.S. (2011). Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *NeuroImage* 54, 1427–1431.

Institutional Responsibilities, Organisation & Outreach

- 2020** – Active participation in **institutional organisation** (e.g., main organiser of *CerCo Project Updates*, member of committee for *CerCo Support Fund*)
- 2020** **Invited member of PhD committee**, Max-Planck-Institute for Psycholinguistics, Nijmegen, NL
- 2019** – **Member** of the *International Network of tES-fMRI*, a consortium to standardize protocols combining brain stimulation and brain imaging
- 2018** **Main organiser** of the **international conference** “*Perturbing and Enhancing Perception and Action using Oscillatory Neural Stimulation*”, Cambridge, UK
- 2018** – **Founding Member and Co-organiser** of the “*tACS challenge*”, a multi-lab project designed to standardize and replicate brain stimulation protocols
- 2017** – Regular participation in **public events** such as the **Brain Awareness Week** in Cambridge and Toulouse
-

Reviewing and Editorial Activities

- 2022** – **Associate Editor** and **Reviewer Mentor** for The Journal of Neuroscience
- 2020** – **Associate Editor** for Frontiers in Neuroscience
- 2020** – **Reviewer** for multiple international funding bodies, including German Research Foundation (DFG), Human Frontier Science Program (HFSP), Swiss National Science Foundation (SNSF), National Science Foundation (NSF)
- 2011** – **Reviewer** for many international journals, including Sciences Advances, Nature Human Behaviour, Trends in Cognitive Sciences, PNAS, Current Biology, PLOS Biology, eLife, The Journal of Neuroscience, Brain Stimulation
-

Teaching

- planned** **Lecture** “*Méthodologie et analyse de données en neurosciences*” for Master Programme in Cognitive and Integrative Neuroscience (upcoming), Paul Sabatier University, Toulouse
- 2023** **Lecture** “*Fundamentals and practical applications of tACS*” University of Helsinki, Finland
- 2022** Member of **evaluation committee** for Master Programme (**M2**) “*Neuroscience, Behaviour, and Cognition*”, Paul Sabatier University Toulouse
- 2016 – 2019** **Lecture** “Brain Stimulation”, University of Cambridge, UK
- 2012 – 2015** **Lecture** “Electroencephalography”, Paul Sabatier University Toulouse
-

Supervision

Training

- 2018 – 2019** completed training in **team management and leadership** (University of Cambridge, UK)

Post-Doc

Jules Erkens (2022-2025), **Florian Kasten** (2022-2024), **Ka-Yan Lui** (2022-2024)

Research Engineer (Ingénieur d'Etude)

Sander van Bree (2019), **Ram Kumar Pari** (2023-2024)

Master Students

Christoph Daube (M2, 2014), **Naveen Reddy Pasham** (M1, 2014), **Sylvain L'Hermite** (M2 and M1, 2021/22), **Damien Ricci** (M1, 2022), **Eva Boglietti** (M1, planned for 2023), **Quentin Busson** (M1, planned for 2023), **Zoé Darrasse** (M1, planned for 2023)

Bachelor and Medical Students

Romain Brousse (2013), **Megha Anil** (2018), **Isabella Allard** (2018 and 2019), **Ambre Besse** (L3, 2021), **Victorina de Combles de Nayves** (L3, 2022)

I will be attached to the doctoral school CLESCO (Comportement, Langage, Education, Socialisation, Cognition).

3 INTRODUCTION

3.1 Auditory Oscillations – an attentional filter that operates in time

"You can't do everything at once" – what is true for the career of a scientific researcher is all the truer for the human brain. Continuously bombarded with sensory input, it needs to filter the information it receives to avoid interference and overload. This function – *selective attention* – operates in both space and time, but requirements differ across modalities. **Acoustic information is defined in the temporal domain**, as regular changes in sound pressure over time that arrive at the ear, conventionally expressed in cycles per second (*Hertz*). We would not perceive sound without integrating information over time. Already in the cochlea, each hair cell responds selectively to sound waves at a certain frequency; this *frequency tuning* is preserved across the hierarchy up to (at least) primary auditory cortex (Saenz and Langers, 2014), a notion I will come back to in later chapters. We are also remarkably good in processing the temporal features of acoustic information. Two short auditory clicks can be distinguished if they are only 2 ms apart, a precision that is ten times that found in the visual system (Exner, 1875). Even the detection of sound *in space* relies on time: Humans use interaural time differences to localise sound and can distinguish timing differences between the two ears that are as low as ten μs (Bonham and Lewis, 1999).

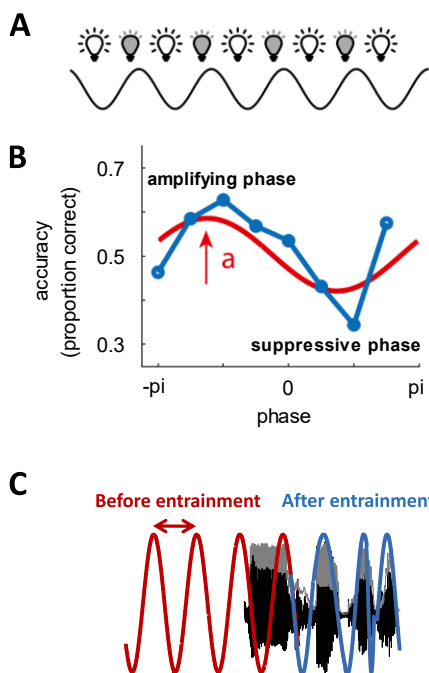
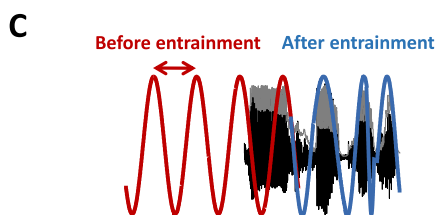


Figure 1. Neural Oscillations. **A.** Neural oscillations underlie rhythmic changes in perception, alternating between amplifying and suppressive phases for stimulus input. From VanRullen, 2016. **B.** Accordingly, perception (e.g., target detection) covaries with oscillatory phase. The blue line illustrates such a phasic effect. The red line shows a sine function fitted to the data. The amplitude a of the sine fit reflects the magnitude of the phasic modulation of performance and is often used to quantify such phasic effects. **C.** The brain is able to adjust neural oscillations (red) so that they are aligned (blue) to a rhythmic stimulus, such as human speech (black). This phenomenon is often termed “neural entrainment” (e.g., Lakatos et al., 2008).



Given this strong link with the temporal domain, it is clear that the **auditory system has a particular need to select information in time**. Although multiple processes exist to achieve this goal, throughout my scientific career I have extensively explored one of them: *neural oscillations*.

There is little doubt that neural oscillations structure information processing in time (Jensen and Mazaheri, 2010; VanRullen, 2016). These oscillations reflect rhythmic fluctuations in neural excitability (Buzsáki and Draguhn, 2004) and cycle, several times per second, between (high-excitability) phases during which input is amplified and (low-excitability) phases during which input is suppressed (**Fig. 1A,B**). **This property makes oscillations a neural substrate of information selection in time:** High-excitability phases can amplify important input while low-excitability phases can inhibit irrelevant, potentially distracting events (Schroeder and Lakatos, 2009). The inhibition of incoming information during the low-excitability phase might also prevent interference with other internal processes, such as memory consolidation. Thus,

neural oscillations might not only gate information uptake, but coordinate various neural functions in time.

One important consequence of this neural rhythm is that **stimulus processing is not uniform over time, but alternates between optimal and non-optimal moments for a given function**, depending on the instantaneous phase of the oscillatory cycle. Indeed, a seminal study demonstrated that the probability of detecting a visual target co-varies with the phase of an oscillation fluctuating at ~ 7 Hz (Busch et al., 2009). Since then, the large amount of studies demonstrating “rhythmic attention” leaves little doubt that neural oscillations are a driving force behind numerous perceptual and attentional functions (Landau and Fries, 2012; Fiebelkorn et al., 2013; VanRullen, 2016; Helfrich et al., 2018; Nobre and van Ede, 2018; Fiebelkorn and Kastner, 2019). Notably, however, a vast majority of these rhythmic effects has been discovered in the visual domain. This is surprising, as the auditory system is confronted with information that changes rapidly in time and therefore seems to be in a stronger need of a temporal filter than the visual one, which is not exposed to such rapid fluctuations. Is this a mere reflection of the fact that vision is studied more extensively than audition? As I will detail below, it is most likely not the only reason: **As a central thread that runs through this thesis, oscillations in the auditory system can be a curse or a blessing** – they can be harmful or useful, depending on the context they are confronted with.

This notion is a consequence of a property of neural oscillations that has been ignored in most of previous work: **Neural oscillations will inevitably entail moments during which a stimulus is relatively unlikely to be perceived and processed**. These low-excitability phases are necessary to ensure an effective attentional filter as they inhibit the processing of irrelevant, to-be-ignored events that occur at the same time. However, such stimulus-suppressive phases become problematic if they coincide with important events, leading to a potential loss of information.

Logically, a temporary inhibition of stimulus processing will be particularly harmful if that stimulus is present only briefly. The transient nature of typical auditory stimuli (e.g., speech, music) therefore suggests that the **auditory system is particularly vulnerable to information loss during low-excitability parts of the oscillatory cycle**. In contrast, visual information is relatively stable over time, making the visual system more robust against such a loss of information.

Does this mean that oscillations do more harm than good in the auditory system, as they might suppress important information? Indeed, we have previously reported that auditory detection in quiet does not co-vary with the phase of neural oscillations (Zoefel and Heil, 2013), a finding which was confirmed by other groups at the same time (summarized in VanRullen et al., 2014). This is in contrast to the visual system, where oscillations are omnipresent (VanRullen, 2016). On the other hand, **there is no doubt that neural oscillations do play an important role in audition**. A rapidly growing field of research is focused on the alignment between neural oscillations and rhythmic auditory stimuli (**Fig. 1C**), termed “*neural entrainment*” (Lakatos et al., 2008; Thut et al., 2011a; Lakatos et al., 2019; Obleser and Kayser, 2019). Entrainment is at the heart of current theories of speech perception (Giraud and Poeppel, 2012), and my own work has played a role in this development.

Until now, **my scientific career was dedicated to these two sides of the coin** of oscillations in audition. During my PhD, I tested the idea that a random oscillatory “sampling” of acoustic information can indeed lead to information loss and a corresponding deterioration of auditory perception. I am currently testing the hypothesis that the degree of potential information loss is a decisive factor for whether neural oscillations are present in electrophysiological or behavioural data. During my PhD and post-doc, I extensively studied neural entrainment. Due to the temporal regularity (and hence predictability) in the stimulus, neural oscillations can align their high-excitability phase to important events, thereby reducing the risk of information loss. **Rhythmicity is therefore an ideal stimulus property for oscillations**, and neural entrainment is the prime example of oscillatory information selection in time. At the same time, the field of entrainment is prone to responses that only seem oscillatory, a point I will come back to below. My work continues to be shaped by the theory of neural entrainment that I investigate from various angles and describe in detail in various chapters of this thesis.

Together, neural oscillations can act as an attentional filter that operates in time, but stimuli can be missed if they occur during the inhibitory part of the oscillation. This issue is most critical for audition as it needs to process brief, rapidly changing input. Nevertheless, oscillations can be a powerful tool for audition if amplifying, high-excitability phases are synchronized with important events. This brain-stimulus synchronization is easiest to achieve when stimuli are rhythmic, an observation that led to the field of neural entrainment. My previous studies and their findings that ultimately led to this conclusion will be presented in the following chapter. It will be followed by a chapter dedicated to open questions and hypotheses that I will address in the future. Of course, scientific research is never done well when it is done alone, and I am very grateful to advisors, collaborators and students involved in both past and future research. Those contributing to a specific study are mentioned explicitly at the beginning of each section.

4 PREVIOUS RESEARCH

4.1 PhD work (2012-2015)

I have completed my PhD thesis at the *Centre de Recherche Cerveau & Cognition (CerCo)*, CNRS, under supervision of Dr Rufin VanRullen in Toulouse, France. During this time, I was also a visiting scholar at the *Nathan Kline Institute for Psychiatric Research (NKI)* in Orangeburg, NY, USA, in collaboration with Dr Peter Lakatos and Prof Charles E Schroeder. The PhD thesis was titled "*Phase entrainment and perceptual cycles in audition and vision*". This work was supported by a doctoral fellowship awarded by one of the most prestigious research foundations in Germany (Studienstiftung des Deutschen Volkes).

4.1.1 Perceptual impact of information loss during rhythmic sampling of stimulus input

Advisors: Rufin VanRullen, Peter Heil

Collaborators: Barkin Ilhan, Saskia Brüers

Supervision: Naveen Reddy Pasham (B.Sc.)

Publications:

Zoefel, B., and Heil, P. (2013). Detection of Near-Threshold Sounds is Independent of EEG Phase in Common Frequency Bands. *Front. Psychol.* 4.

VanRullen, R., **Zoefel, B.**, and Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130214.

Zoefel, B., Reddy Pasham, N., Brüers, S., and VanRullen, R. (2015). The ability of the auditory system to cope with temporal subsampling depends on the hierarchical level of processing. *Neuroreport* 26, 773–778.

Zoefel, B., and VanRullen, R. (2017). Oscillatory Mechanisms of Stimulus Processing and Selection in the Visual and Auditory Systems: State-of-the-Art, Speculations and Suggestions. *Front. Neurosci.* 11.

The first study of my PhD thesis was inspired by the previous observation – made during my M.Sc. project and confirmed in parallel by other research groups – that auditory target detection in quiet does not depend on the phase of neural oscillations (Zoefel and Heil, 2013; VanRullen et al., 2014), whereas corresponding phase effects in the visual domain seem robust (Busch et al., 2009; Busch and VanRullen, 2010). We speculated that this is due to the fact that vision, but not audition, can tolerate information loss during the low-excitability part of the oscillation (see **Introduction**) to a certain degree, a hypothesis we aimed to test during my PhD work.

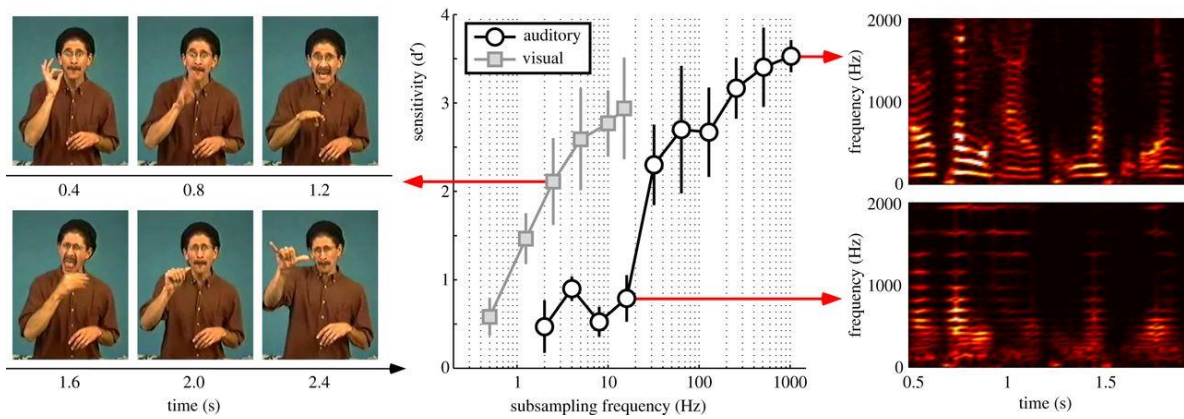


Figure 2. Videos of sign language (left) and speech sounds (right) were subsampled at different frequencies. Recognition performance (in a 2-back task; middle) in the auditory condition declined faster with decreasing subsampling frequency than in the visual condition. Error bars show standard error of the mean (SEM) across participants. From VanRullen et al. (2014).

In psychophysical experiments, we tested this hypothesis by mimicking the rhythmic modulation of stimulus processing that goes along with neural oscillations. This was done by temporarily subsampling auditory (speech sounds) and visual stimuli (videos of sign language) and measuring the impact of subsampling frequency on recognition performance (2-back task) in human subjects. As shown in **Fig. 2**, performance declined in both modalities with a decrease in subsampling frequency. This is not very surprising as more information, and in particular temporal patterns (“fine structure”), is lost when the stimulus is sampled more coarsely. Importantly, however, performance declines much faster for the auditory system (black line/circles in **Fig. 2**). For a sampling frequency of 15 Hz, for example, participants performed the visual task with ease (d' > 3), whereas they failed to do so in the auditory case (d' < 1). We concluded, therefore, that **subsampling input to the visual and auditory systems is indeed more disruptive for the auditory one**. In principal, this result might indicate that oscillatory processes – at least at commonly considered frequencies – result in a loss of information for the auditory system that is too large to be affordable – and therefore speak against the use of neural oscillations for auditory stimulus processing.

There is, however, an alternative that we described in several publications (VanRullen et al., 2014; Zoefel et al., 2015; Zoefel and VanRullen, 2017): Whereas auditory input fluctuates rapidly at early levels of auditory processing, the representation of input is more abstract (and potentially more stable in time) on higher levels of the auditory hierarchy (Davis and Johnsrude, 2003; Edwards and Chang, 2013). We designed another psychophysical experiment (Zoefel et al., 2015) for which we constructed speech stimuli that were subsampled in two different ways, mimicking processing at two different levels of processing: In the first condition, sounds were subsampled directly in the time domain (i.e. the input waveform; mimicking processing at early auditory levels such as the cochlea). This condition corresponds to the auditory condition of the first experiment described. In the second condition, sounds were subsampled at the level of auditory features (obtained by a vocoder using linear predictive coding; mimicking cortical levels of auditory processing). As shown in **Fig. 3**, performance in the same recognition task indeed improves when subsampling is applied on the feature level (grey line). This finding suggests that **auditory recognition is more robust**

to subsampling on a relatively high level of auditory processing than to subsampling in the time domain.

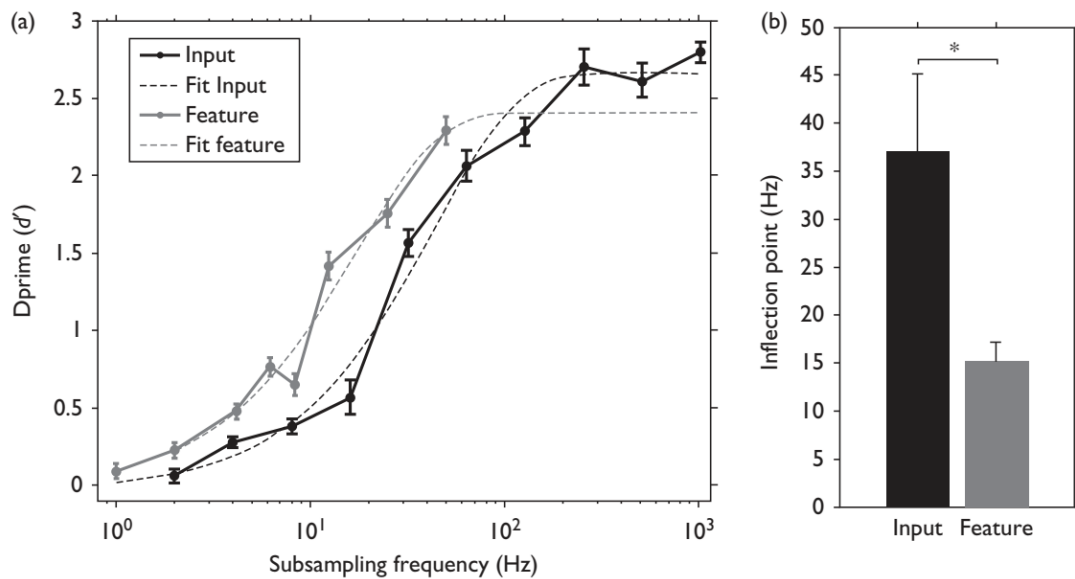


Figure 3. A. Recognition performance in a two-back task for two experimental conditions (subsampling on the input or auditory feature level, respectively). Dashed lines show sigmoidal fits to the data. **B.** Inflection points of the sigmoidal curves shown in A. The higher inflection point for the input condition indicates that subsampling has more detrimental effects when done on the level of the input waveform (mimicking cochlear processing). Error bars show SEM across participants. From Zoefel et al. (2015).

Together, our results confirmed that a rhythmic inhibition of acoustic information, at most rates (< 40 Hz) typically associated with neural oscillations, can disrupt perception. However, on a relatively high hierarchical level, neural oscillations could operate in the auditory system without critically disrupting temporal information.

4.1.2 Neural entrainment to speech without slow changes in spectral energy

Advisors: Rufin VanRullen, Peter Lakatos

Collaborators: Jordi Costa-Faidella, Charles E Schroeder

Publications:

Zoefel, B., and VanRullen, R. (2015a). Selective Perceptual Phase Entrainment to Speech Rhythm in the Absence of Spectral Energy Fluctuations. *J. Neurosci.* 35, 1954–1964.

Zoefel, B., and VanRullen, R. (2015b). The Role of High-Level Processes for Oscillatory Phase Entrainment to Speech Sound. *Front. Hum. Neurosci.* 9.

Zoefel, B., and VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage* 124, 16–23.

Zoefel, B., Costa-Faidella, J., Lakatos, P., Schroeder, C.E., and VanRullen, R. (2017). Characterization of neural entrainment to speech with and without slow spectral energy fluctuations in laminar recordings in monkey A1. *NeuroImage* 150, 344–357.

A second, not mutually exclusive, possibility exists that would enable oscillatory processes in the auditory domain: Loss of information during non-optimal oscillatory phases would be reduced if the auditory system actively adjusted optimal phases to particularly relevant moments in time. Indeed, it has been shown repeatedly that the phase of neural oscillations can be aligned to rhythmic (hence predictable) input, a mechanism termed neural entrainment (Lakatos et al., 2008) and that was already described in the Introduction. Thus, **neural entrainment seems to be a mechanism that allows auditory oscillations to play to their strengths while reducing harmful effects at the same time**; its further investigation was an essential part of my PhD work.

It has been shown that neural entrainment to speech (**Fig. 1C**) – arguably the most important (quasi-)rhythmic stimulus in the auditory environment – correlates with successful speech perception (Luo and Poeppel, 2007; Peelle and Davis, 2012; Gross et al., 2013; Peelle et al., 2013). Speech perception was more successful (e.g., participants responded more accurately) when rhythmic brain responses, measured using electro- or magnetoencephalography (EEG/MEG), aligned to the rhythm of speech more reliably. However, speech is a complex stimulus, and it remained unclear to which features of speech the brain actually adjusts to. For instance, in standard speech, **changes in sound amplitude or spectral energy always go along changes in phonetic or linguistic features** that could possibly be “tracked” or followed by the brain. As long as certain acoustic and linguistic features co-vary in speech, their effects on speech tracking cannot be disentangled.

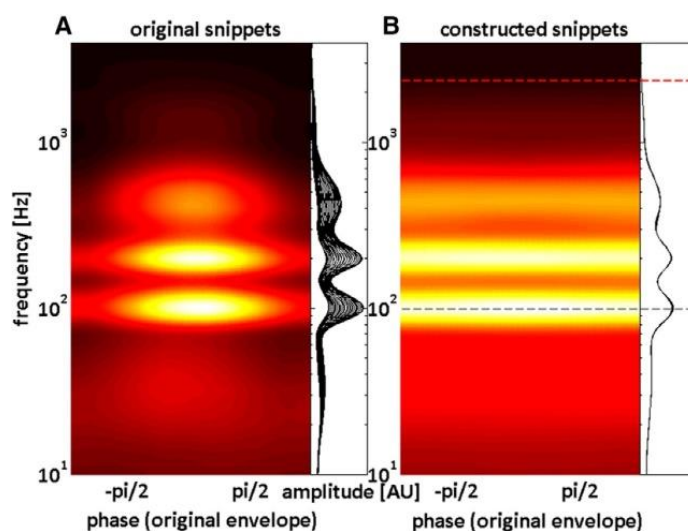


Figure 4. In everyday speech (A), spectral energy (color-coded) is concentrated at certain phases of the speech envelope. It is therefore unsurprising that neural activity follows these pronounced energy fluctuations. In our constructed speech/noise stimuli (B), these differences in spectral energy were counterbalanced (by injection of noise). From Zoefel and VanRullen (2015a).

For my PhD work, we therefore developed **novel speech/noise stimuli without systematic fluctuations in amplitude and spectral content, but with conserved intelligibility** and speech characteristics (Zoefel and VanRullen, 2015a). This was done by constructing noise that was tailored to counterbalance the rhythmic fluctuations in spectral energy of everyday speech – mixing speech with the constructed noise yielded speech/noise sounds without systematic slow fluctuations in spectral energy (**Fig. 4**). These stimuli allowed us to test whether neural entrainment to speech rhythm persists, even when these acoustic variables are controlled for. In several experiments, we demonstrate that the brain can indeed follow the rhythm of speech that is conserved in these stimuli. This outcome demonstrates

that neural entrainment is more than a passive following of fluctuations in acoustic energy – it must involve some higher-level processes, necessary to distinguish speech from the spectrally counterbalanced noise (reviewed in Zoefel and VanRullen, 2015b).

Psychophysics

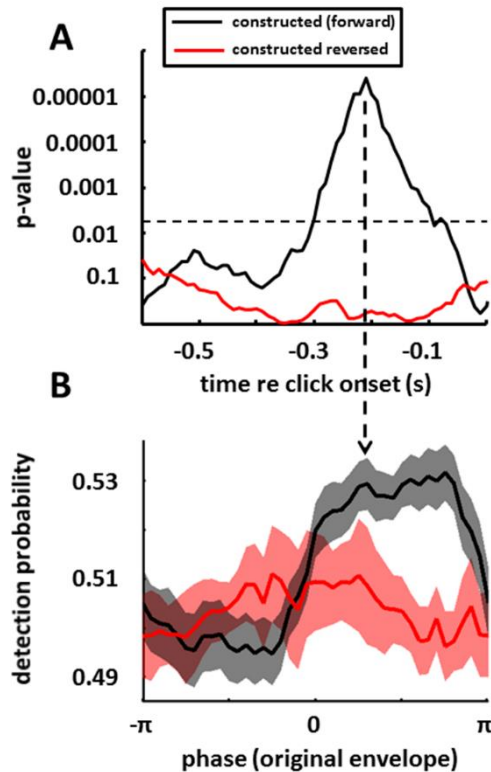
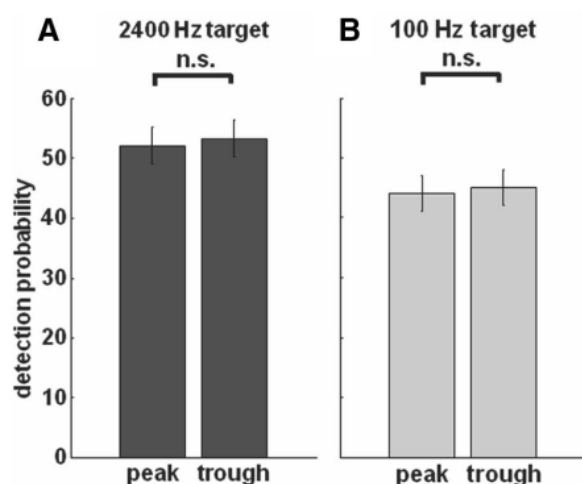


Figure 5. The detection of a short pure tone (click) was significantly modulated by the rhythm of speech, inherent in the constructed speech/noise stimuli. **A** shows the p-value of the modulation (comparison with a simulated null distribution), **B** the actual performance as a function of original envelope phase (which reflects the conserved speech rhythm). Note that the significant modulation of performance is absent when the stimuli were presented in reverse (abolishing intelligibility), indicating a potential dependence on linguistic features. Shaded areas show SEM across participants. From Zoefel and VanRullen (2015a).

We first carried out a psychophysical experiment (Zoefel and VanRullen, 2015a) in which human participants were asked to detect short pure tones, embedded in the constructed speech/noise stimuli. If neural entrainment to speech persisted, despite counterbalanced spectral energy fluctuations (**Fig. 4**), then target detection should depend on the phase of the original speech envelope (which reflects the conserved speech rhythm). Perceptually, the latter reflects an alternation of speech and noise (stimuli were constructed so that speech dominates perception at some envelope phases, while noise dominates at others), and **a phasic modulation of target detection would demonstrate that participants can entrain to the speech rhythm hidden in noise.** Such a phasic modulation was indeed what we found (**Fig. 5**, black). We also reported two additional observations: First, the entrained phase (leading to most accurate target detection) was dependent on the sound frequency of the target, and opposite for the two target frequencies tested (100 Hz and 2400 Hz). This result implies that neural entrainment follows



the tonotopical organization of the auditory system, a concept that is discussed in detail below (non-human primates) and in **Section**

Figure 6. Control experiment. Probability of detecting a pure tone target in noise (**A** and **B** show results for the two target frequencies separately). The noise had the spectral profile of the constructed speech/noise stimuli, at the peak or trough of the original speech envelope (cf. Fig. 4). Error bars show SEM across participants. From Zoefel and VanRullen (2015a).

4.3.3. Second, we found that the phasic modulation was abolished when stimuli were presented in reverse (**Fig. 5**, red). This result is important as it suggests that neural entrainment includes a tracking of linguistic features of speech that are removed by the temporal reversal.

In a control experiment, we tested whether remaining spectral differences at different phases of the original speech envelope can explain our effects. We again asked subjects to detect short pure tones. These were presented in noise that was designed to have the average spectral profile of the constructed speech/noise snippets, either at the peak (**Fig. 4B**, phase 0) or the trough (**Fig. 4B**, phase $-\pi/\pi$) of the original signal envelope. We found no differences in target detection between the two noises (**Fig. 6**). This result shows that the **perceptual effects described are not due to small residual spectral differences during stimulus construction.**

EEG

We then carried out an EEG experiment in human subjects to follow up on these perceptual results. In addition to the constructed speech/noise stimuli and their time-reversed version, we also presented the original speech. We found that EEG responses align to the speech rhythm hidden in the constructed stimuli (**Fig. 7**), corroborating the perceptual effects described. However, we also found entrained EEG responses to the time-reversed version. In the corresponding publication (Zoefel and VanRullen, 2016) and a following review article (Zoefel and VanRullen, 2015b), we discussed this discrepancy to our perceptual study. We proposed that neural entrainment to sound occurs to any pattern in acoustic information (note that reversed speech and noise alternated perceptually in the time-reversed speech/noise stimuli), independent of intelligibility, explaining our EEG results. However, this entrainment has stronger behavioural consequences in the presence of an intelligible stimulus, explaining the absence of a perceptual modulation for the time-reversed stimuli. Irrespective of this speculative explanation, results from this EEG experiment confirmed that, in addition to perceptual responses, **neural responses can align to speech rhythm in the absence of spectral energy fluctuations.**

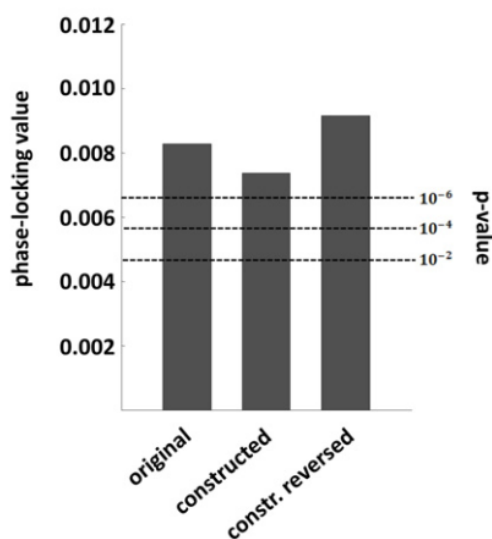


Figure 7. Alignment (phase-locking or entrainment) between EEG responses and the original speech envelope in three stimulus conditions (averaged across EEG channels). P-values were obtained from a permutation procedure. From Zoefel and VanRullen (2016).

Non-human primates

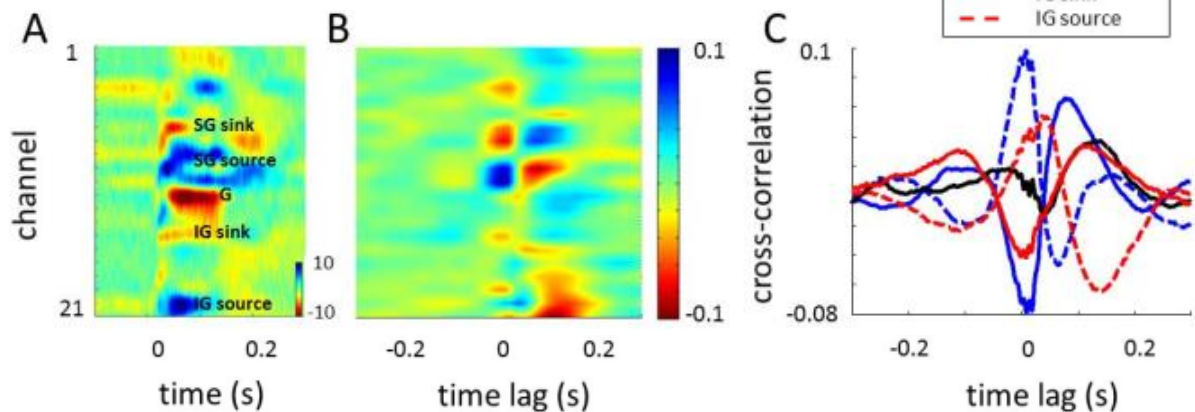
In a third experiment, and in collaboration with researchers at the Nathan Kline Institute (Orangeburg, New York), the same constructed speech/noise stimuli were presented during intracranial recordings in the auditory cortex of non-human primates. This experiment had several aims: (1) to test whether entrainment effects described above are restricted to humans (given the EEG results, this seemed unlikely); (2) to characterize entrainment on the laminar level; (3) to test for tonotopic entrainment effects.

The third aim was based on previous research and requires a more detailed explanation. As described in the Introduction, the auditory system is organized according to sound frequency: A mechanism that is conserved across species, neurons respond preferentially to certain sound frequencies but not to others, and neurons that are spatially close to each other are more likely to respond to similar frequencies than those further apart (Römer, 1983; Lippe, 1995; Lee et al., 2004). This *tonotopy* is found throughout the auditory hierarchy, up to (at least) primary auditory cortex (Saenz and Langers, 2014). Interestingly, **neural entrainment seems to follow a similar tonotopical organization**. Previous studies in non-human primates have reported that neural entrainment depends on the sound frequency of the presented stimulus: Only those parts of the auditory cortex that are tuned to that sound frequency align the high-excitability phase of their oscillations to the presented sounds – other parts are set to their low-excitability phase (O’Connell et al., 2011, 2014). This suggests that **entrainment is not only a temporal filter (see Introduction), but also a spectral one: At a given point in time, only sound frequencies considered to be important are amplified, whereas others are suppressed** (Lakatos et al., 2013a). This is a notion that I will come back to in **Section 4.3.3** of this thesis. In the experiment described here, we tested for similar tonotopical effects during the presentation of human speech.

We quantified oscillatory responses in non-human primate auditory cortex using *current-source density* (CSD) estimation. CSD profiles are the second spatial derivative (Nicholson and Freeman, 1975) of the more commonly measured *local field potential* (LFP). They have the advantage of being less affected by volume conduction than the LFP, allowing more straightforward interpretation of the location and direction (Mitzdorf, 1985; Schroeder et al., 1998). **Fig. 8** shows an example of such a laminar profile (**A**) and how it changes in response to (clear) human speech (**B,C**). We used cross-correlation to quantify how and when a change in the speech envelope leads to corresponding changes in the CSD profile. As shown in **Fig. 8B,C**, the time lag and magnitude of the neural response to speech depends on the layer in which the signal is recorded. Note that the sign of cross-correlation also depends on the polarity of the layer (*sink vs source*).

We next examined how this response depends on the sound frequencies present in speech, relative to the “*best*” frequency of the recording site (i.e. the sound frequency that produces the strongest response). That is, we again cross-correlated speech envelope and CSD profiles, but instead of using the broadband speech envelope (as for **Fig. 8B,C**), we first filtered the (clear) speech stimulus into narrower frequency bands and used their envelopes for cross-correlation. For **Fig. 9**, we contrasted results for recording sites tuned to low (≤ 1000 Hz) and high (≥ 8000 Hz) sound frequencies.

Figure 8. **A.** Example CSD profile measured in response to a pure tone in primary auditory cortex. The sound frequency of the tone corresponded to the “best” frequency (producing the strongest response) of the recording site. The profile reveals sinks and sources of current flow (referenced to a pre-stimulus interval) in various cortical layers. SG: Supra-granular, G: granular, IG: infra-granular. **B** and **C** show the equivalent profile in response to human speech. The amplitude envelope of the speech signal was cross-correlated with CSD responses to obtain the patterns shown. For **C**, recording channels that best capture sinks and sources were selected. From Zoefel et al. (2017).



This contrast revealed that the response pattern (i.e. the sign of cross-correlation) depends on the best frequency of the recording site relative to that of the stimulus. For instance, cross-correlation peaked around 50 ms for the infragranular (IG) sink, but the direction of this peak depended on whether the frequency of speech used for the analysis matches the best frequency of the site (negative correlation) or not (positive correlation). In other words, neurons responded to speech even if they were not tuned to its frequency – however, their response was *opposite* as compared to that to the preferred frequency.

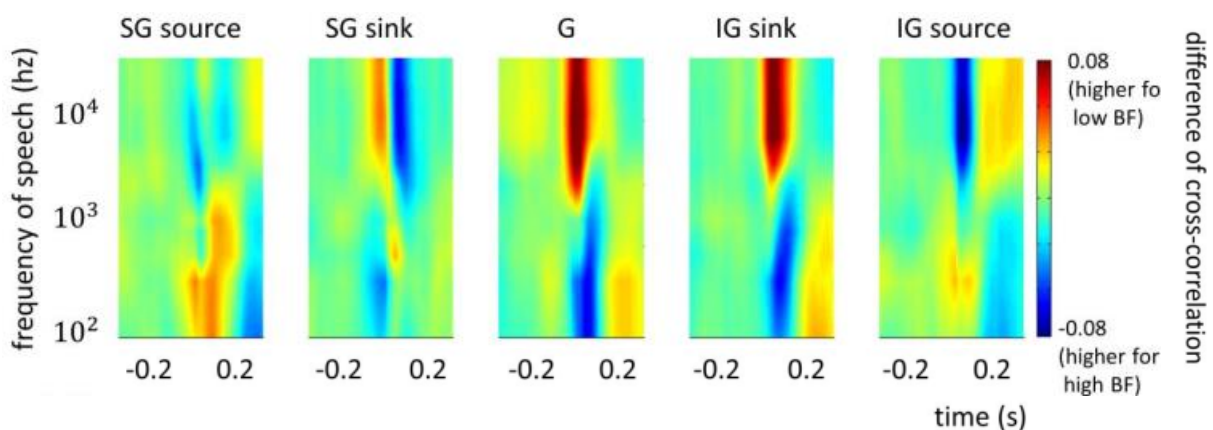


Figure 9. Difference of cross-correlation (recording sites tuned to sound frequencies ≤ 1000 Hz minus recording sites tuned to frequencies ≥ 8000 Hz) between CSD and the speech envelope filtered into narrow frequency bands. For other conventions, see Fig. 8. From Zoefel et al. (2017).

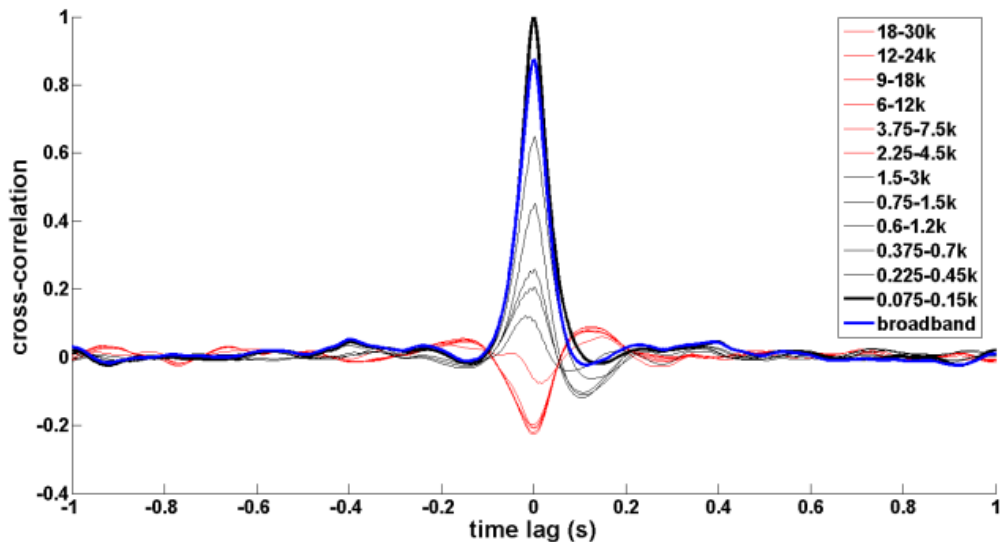
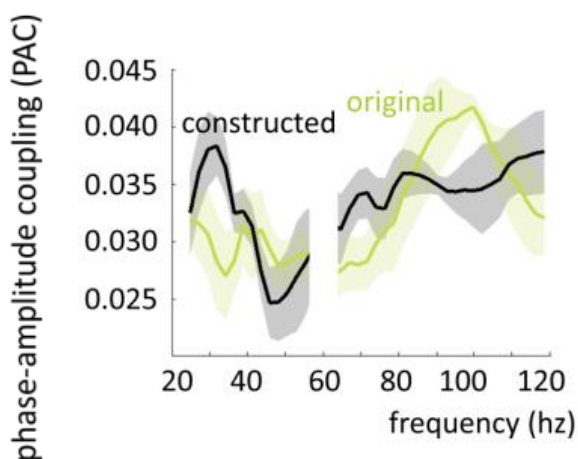


Figure 10. Cross-correlation between the amplitude envelope of one selected frequency band in human speech (here 75-150 Hz; thick black line) and the envelopes of other frequency bands. Note that, at time lag 0, this cross-correlation is positive for most bands of relatively low frequency, including the broadband envelope (thick blue line). However, the sign of correlation becomes negative for higher frequencies. Note also a change of sign at a time lag that corresponds to approximately one cycle of the broadband envelope ($\sim \pm 200$ ms). This pattern can be explained by alternating low- and high-frequency components. From Zoefel et al. (2017).

These results indicate that neural entrainment to human speech in auditory cortex of non-human primates is indeed tonotopically organized. This observation is particularly important when combined with a property of human speech that we described in the same publication (Zoefel et al., 2017) and that is illustrated in **Fig. 10**: High- and low-frequency components in speech, dominated by certain fricatives and vowels (Shamma, 1988), respectively, often alternate. **The two findings combined – tonotopical organization of neural entrainment and alternation of high- and low-frequency sounds in speech – point to a spectrotemporal filter mechanism as proposed in previous work** (Lakatos et al., 2013a). If neural oscillations entrained tonotopically, then aligning their high-excitability phase to moments of high spectral energy in one particular frequency band would automatically align low-excitability phases to high spectral energy in other bands (as these occur, on average, half a cycle later; **Fig. 10**). Such a mechanism would not only boost neural activity precisely *when* “needed” and suppress it otherwise (temporal filter), but also



depending on the spectral tuning of the neural population (spectral filter) – activity in regions with different spectral tuning can be amplified and suppressed at the same time.

Figure 11. Coupling between the phase of neural responses aligned to speech rhythm and the amplitude of a higher frequency. The two lines correspond to the stimulus presented (original clear speech or constructed speech/noise stimuli). Shaded areas show SEM across nine recordings in one subject. From Zoefel et al. (2017).

In addition to these neural responses to clear human speech, we also tested whether CSD responses align to the rhythm of the constructed speech/noise stimuli. We found that such entrained responses were present in all cortical layers, without significant differences between them (not shown). In addition, these stimulus-aligned responses were locked to changes in amplitude at higher frequencies, a phenomenon commonly termed *phase-amplitude coupling* (Lisman and Jensen, 2013). We found that the phase of the entrained neural response was coupled to amplitude changes around 30 Hz during the presentation of speech/noise stimuli (black in **Fig. 11**), but to ~100 Hz during original, clear speech (green). This result was reminiscent of a previous study in human subjects (Fontolan et al., 2014) that observed similar spectral components, and associated them with top-down (~30 Hz) and bottom-up processing (~95 Hz), respectively. Together, we found (1) that neural entrainment to human speech persists in the absence of slow changes in spectral energy, even in non-human primates; and (2) evidence for a top-down process involved in separating speech and noise (of course, the latter needs to be tested more explicitly in future experiments).

4.1.3 Neural oscillations structure acoustic information

Advisors: Rufin VanRullen, Peter Lakatos

Collaborators: Monica N O’Connell, Annamaria Barczak

Another question that I addressed during my PhD work is the impact of oscillatory processes on auditory perception. As mentioned above, it is likely that oscillatory activity is structured by stimulus input in the auditory system – but **is it possible that oscillatory activity provides structure to the input as well?** We tested this using the so-called “ABA paradigm”, a

popular stimulus in auditory research (van Noorden, 1975): In this paradigm, triplets (“ABA”) of tones are presented, with different sound frequencies for A and B

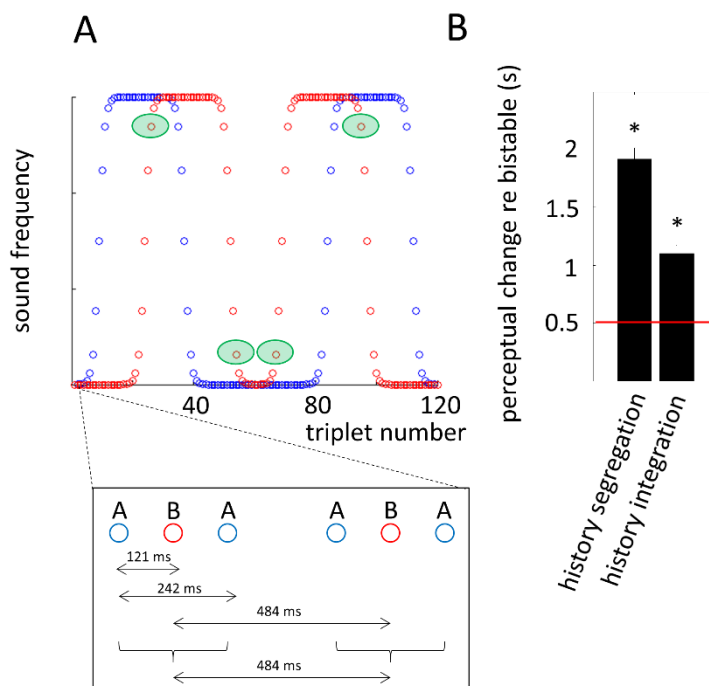


Figure 12. *A.* Sound frequencies and timing of A and B tones (blue and red dots, respectively) in our ABA paradigm. Bistable ABA constellations are indicated in green. *B.* Based on human psychophysical data, their perception is biased by the “stimulation history” (i.e. towards the preceding percept): A change in percept after the bistable constellation occurred significantly later than a typical reaction time, shown as red bar.

tones. The ABA triplet can be perceived as one integrated stream when the sound frequencies for A and B tones are similar, and as two segregated streams (one stream of A tones and another stream of B tones with half the presentation rate) when the sound frequencies for A and B tones are far apart (Moore and Gockel, 2012). Although it seems to be a straightforward idea to compare oscillatory phases for integrated vs segregated stimulus streams, it should be noted that the stimuli physically differ in these two cases – a potential phase difference might thus reflect a mere consequence of these stimulus differences, but not a change in percept. We therefore developed a version of the ABA paradigm in which an identical stimulus can evoke different percepts. In the course of our experimental protocol (**Fig. 12A**), bistable constellations of A and B tones were present, i.e. at this constellation of sound frequencies, it has been reported that listeners sometimes perceive stream integration, and sometimes stream segregation (Deike et al., 2012). Importantly, these occurrences of bistability were either preceded by stream integration or segregation; it might thus be that the “history” of stimulation affects perception at the time of bistability (towards the percept of the preceding trials) if the stimulus streams are presented at a rapid pace. We verified this hypothesis in a psychophysical experiment on human subjects indicating that perception can indeed be biased towards the percept evoked by the preceding stimulation (**Fig. 12B**).

We then applied this paradigm during electrophysiological measurements in the primary auditory cortex of non-human primates, and focused again on CSD responses. We found a difference in oscillatory phase when comparing “stimulation history” (integration vs segregation) at the time of bistable ABA constellation (**Fig. 13**). The effect was strongest in supragranular layers, the cortical layers in which the alignment of neural oscillations to rhythmic input is typically observed (Lakatos et al., 2008, 2009, 2013a). This finding supports the notion that our observed effect reflects a direct impact of oscillatory processes on auditory perception. Together, we were able to confirm the idea that, in an ambiguous auditory scene, it is the neural phase that “decides” whether two simultaneous auditory inputs are grouped into a single stream or segregated into two separate streams. This finding indicates that **the phase of neural oscillations indeed has consequences on perception and stimulus processing**, and changing it actively might be an important tool to control or “gate” stimulus input.

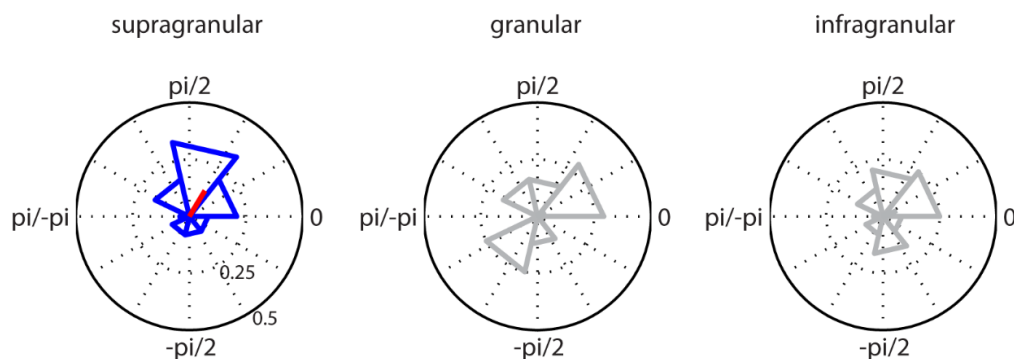


Figure 13. Circular phase differences between trials with different “stimulation history” but identical stimulation (green patches in Fig. 12). There is a significant phase difference between (assumed) stream integration and segregation, but only in supragranular layers, where oscillatory effects are typically strongest. Non-significant results are shown in grey.

4.1.4 Summary: PhD work

Whereas the rhythmic, oscillatory component in visual processing is relatively established (VanRullen and Koch, 2003), the amount of negative results in the auditory system seemed surprising at first glance. My findings demonstrate that **this apparent absence of oscillatory processes in the auditory domain does not necessarily mean that they do not exist**. Rather, due to a higher risk of information loss, oscillations might be harmful or helpful in audition, and the system needs to develop ways to minimize such information loss during oscillatory sampling. The entrainment of neural oscillations to rhythmic input might fulfill this role in the auditory system; oscillations might also operate on a hierarchically high level of stimulus processing. Both processes ensure that auditory oscillations do not entail a critical loss of information. At the same time, my PhD results demonstrate that neural entrainment is a critical mechanism for auditory processing, in particular for that of human speech. The development of methods to enhance neural entrainment during my post-doc was therefore the direct consequence of my PhD work and is described in the next section.

4.2 Post-doc work (2016-2019)

My post-doctoral work was carried out at the *MRC Cognition and Brain Sciences Unit (CBU)*, University of Cambridge, UK, in the group of Dr Matthew H Davis. This work was supported by a post-doctoral fellowship awarded by the German Academic Exchange Service (DAAD) and a Marie-Curie Individual Fellowship.

4.2.1 Causal role of neural entrainment for speech processing

Advisor: Matthew H Davis

Collaborators: Alan Archer-Boyd, Lars Riecke

Supervision: Megha Anil (medical student), **Isabella Allard** (medical student)

Publications:

Zoefel, B., and Davis, M.H. (2017). Transcranial electric stimulation for the investigation of speech perception and comprehension. *Lang. Cogn. Neurosci.* 32, 910–923.

Riecke, L., and **Zoefel, B.** (2018). Conveying Temporal Information to the Auditory System via Transcranial Current Stimulation. *Acta Acust. United Acust.* 104, 883–886.

Zoefel, B., Archer-Boyd, A., and Davis, M.H. (2018a). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr. Biol.* 28, 401-408.e5.

Zoefel, B., Allard, I., Anil, M., and Davis, M.H. (2019). Perception of Rhythmic Speech Is Modulated by Focal Bilateral Transcranial Alternating Current Stimulation. *J. Cogn. Neurosci.* 32, 226–240.

Neural entrainment has been associated with improved speech comprehension (Luo and Poeppel, 2007; Peelle and Davis, 2012). This conclusion mainly resulted from studies reporting that the time-reversal of speech, masking with background noise, or noise vocoding (all of which disrupt intelligibility) impacts neural entrainment (Gross et al., 2013; Peelle et al., 2013; Hauswald et al., 2022). However, typical manipulations of speech intelligibility might alter the acoustic properties of the sound that the brain entrains to (such as “acoustic edges”; Doelling et al., 2014), leading to a decline in neural entrainment and speech intelligibility at the same time. Consequently, the link between neural entrainment and speech comprehension could be merely *epiphenomenal*. **Modulating neural entrainment with brain stimulation and observing consequences for speech processing can therefore provide us with a decisive answer to the question of whether these two factors are indeed causally linked.**

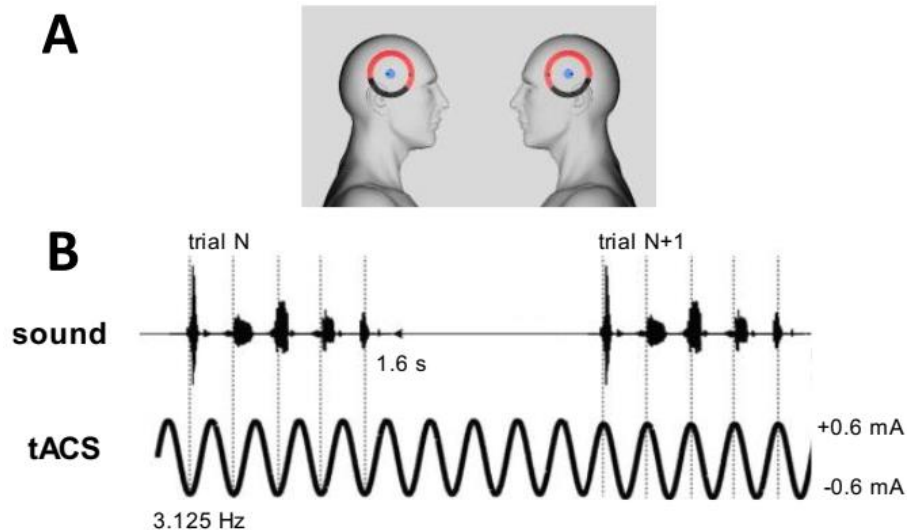


Figure 14. Modulating speech processing with tACS. **A.** Position of tACS ring electrodes to target auditory brain regions in Zoefel et al. (2020). Ring-shaped electrodes are assumed to improve the focality of the stimulation as compared to standard rectangular ones (Saturnino et al., 2017). **B.** Neural entrainment (cf. Fig. 1C) can be manipulated by presenting speech (top) at different phases (here: trough and peak, or phase $\pm\pi$ and 0) of the alternating current (bottom), applied simultaneously. If perceptual or neural responses to the sound co-vary with the phase relation between tACS and sound (cf. Fig. 1B), this suggests a causal role of neural entrainment for the responses tested.

One method to manipulate the ongoing rhythm of neural oscillations is **transcranial alternating current stimulation (tACS)**. This method played a crucial role for my post-doc and continues to do so in my current research. During tACS, rubber electrodes (**Fig. 14A**) are used to apply alternating current to the scalp, which then propagates through the skull to interact with underlying cortical tissue and leads to oscillating membrane potentials at the stimulated frequency (Ali et al., 2013; Fröhlich, 2015). Importantly, previous research has suggested that using tACS we can “impose” a rhythm on neural oscillations (Herrmann et al., 2013b). For my post-doctoral work, I manipulated neural entrainment by **systematically varying the phase relation between tACS and speech rhythm (Fig. 14B)**, and measured the consequences for speech processing in several experiments. I also summarized and discussed the application of transcranial current stimulation techniques for the investigation of speech processing (Zoefel and Davis, 2017; Riecke and Zoefel, 2018).

Experiment 1: Neural entrainment causally modulates fMRI responses to intelligible speech

In a first study (Zoefel et al., 2018a), we applied tACS at 3.125 Hz with electrodes attached over left auditory cortex (inset of **Fig. 15**). We combined tACS with concurrent functional magnetic resonance imaging (fMRI) to measure BOLD responses to rhythmic speech stimuli. Stimuli consisted of sequences of rhythmically spoken one-syllable words, presented at specific phases of tACS (**Fig. 15**). The “*perceptual centre*”, or *p-centre* (Scott, 1998) is the part of the word that is aligned to the beat when someone is asked to speak rhythmically (e.g., to a metronome). The *p-centre* is defined perceptually, but not necessarily acoustically as it is not straightforward to define an acoustic correlate (Fowler et al., 1988; Howell, 1988). By defining

the stimulus rhythm (and its phase relation to tACS) based on the p-centre of the spoken words (vertical lines in **Figs. 14B and 15**), we therefore constructed a stimulus that is *perceptually* rhythmic.

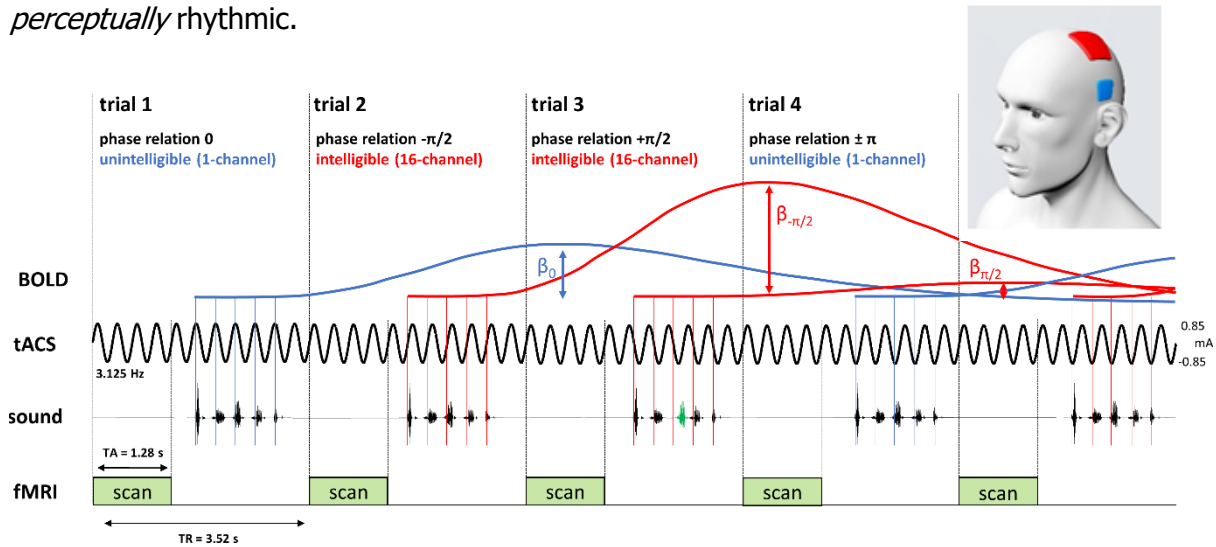


Figure 15. Rhythmic speech was presented at different phases of the underlying alternating current (tACS). We expected that the magnitude of the BOLD response to these speech sounds (shown as beta value) depends on the phase relation between tACS and speech. The panel in the top right corner shows the electrode positions and shapes during the applied tACS. From Zoefel et al. (2018a).

We used noise-vocoding (Shannon et al., 1995) to manipulate the intelligibility of speech sounds. During noise-vocoding, speech is filtered into N frequency bands and the amplitude envelopes of these bands are imposed onto noise, filtered likewise, before noises are re-combined to form a single sound (**Fig. 17**). Depending on N , the resulting sound can be intelligible or not, while preserving the overall speech envelope. In this experiment we used 16-channel vocoded speech (**Fig. 17** with $p = 1$), which is clearly intelligible, and 1-channel vocoded speech (**Fig. 17** with $p = 0$), which sounds like unintelligible noise. We then tested whether the BOLD response to speech depends on the tACS phase the speech is presented at, and whether this effect depends on intelligibility. The hypothesized phasic modulation of neural activity was quantified by fitting a sine function to the BOLD response as a function of tACS phase. The amplitude of the sine fit reflects the magnitude of phasic modulation (labeled a in **Fig. 1B**). These amplitude values are shown, for a specific region of interest (roughly corresponding to *Superior Temporal Gyrus (STG)*, a region involved in speech processing; Davis and Johnsrude, 2003), in **Fig. 16A**, along with their statistical quantification (relative to a simulated null distribution) in **Fig. 16B**. We found **a reliable tACS-induced modulation of the BOLD response, but only when speech was intelligible**. This phasic effect corresponded to a suppressed BOLD response, relative to a sham condition, that was present only at some tACS phases and only during intelligible speech (**Fig. 16C**). A suppressed BOLD response does not necessarily correspond to a decrease in performance. As a next step, we therefore tested whether tACS leads to corresponding changes in speech perception (that was not measured here).

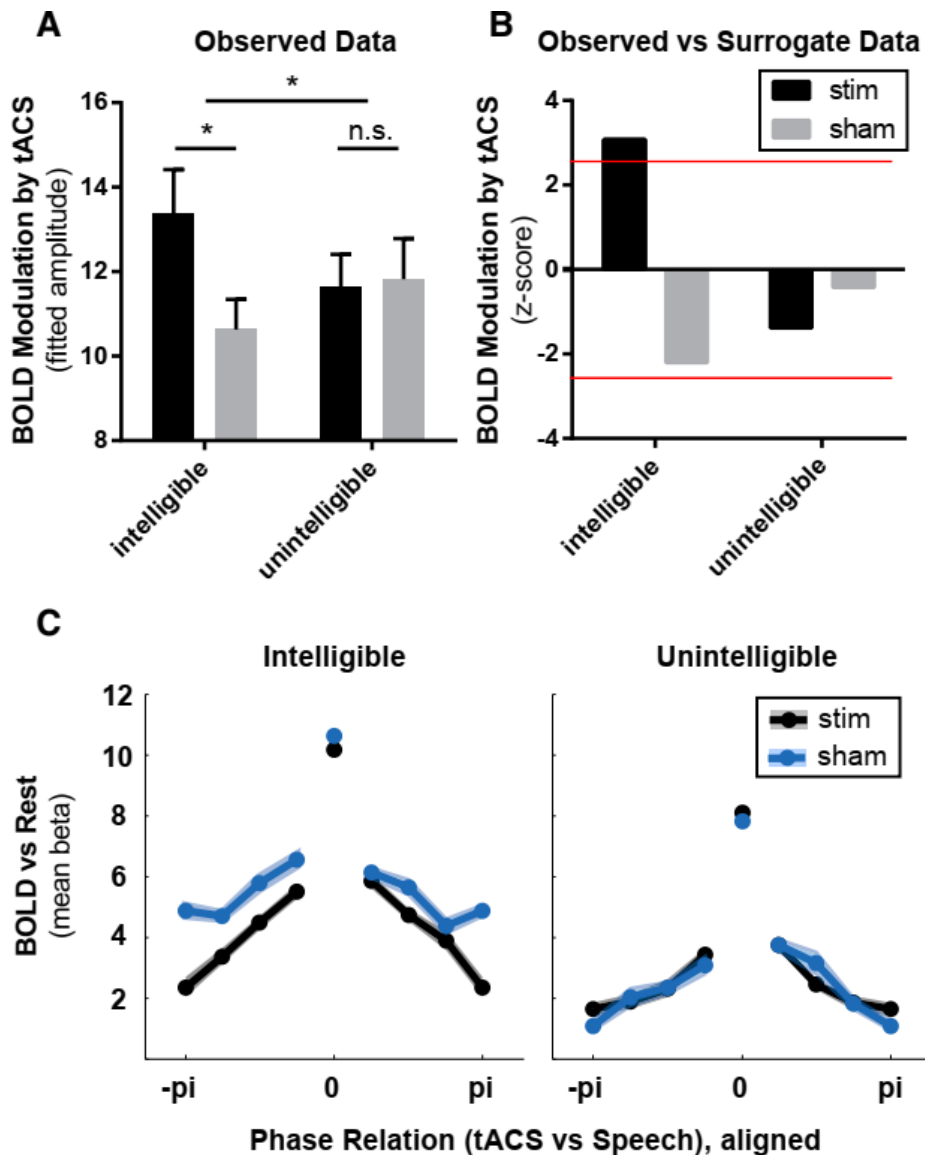


Figure 16. *A. BOLD modulation by the phase relation between tACS and speech in four experimental conditions (stimulation vs sham, intelligible vs unintelligible), quantified as the amplitude of a sine function fitted to the data (cf. Fig. 1B). B. Same as A, but expressed relative to a surrogate distribution (z-score). The significance threshold is shown by a red line. C. BOLD response as a function of the phase relation between tACS and speech. In order to account for individual “best” phase relations (leading to strongest BOLD response), the maximum response of each participant was aligned at the centre bin before averaging across participants. This bin is therefore trivially a maximum and is shown separately. Error bars and shaded areas show SEM across participants. From Zoefel et al. (2018a).*

Experiment 2: Neural entrainment causally modulates speech perception

In a second study (Zoefel et al., 2020), participants performed a word report task while tACS was again applied at different phase relations to the speech rhythm. 16- and 1-channel vocoded speech stimuli were mixed (**Fig. 17**) so that participants reported on average 50% of the presented words correctly. They were asked to click on an image that corresponds to the word they heard (8-alternative forced choice). Performance was quantified relative to a sham stimulation condition. We compared two tACS electrode setups: One that corresponded to the fMRI study (Zoefel et al., 2018a) and used unilateral stimulation with rectangular

electrodes (**Fig. 18**, top left); and another that used bilateral stimulation with “ring” electrodes, assumed to improve focality of the stimulation (Saturnino et al., 2017).

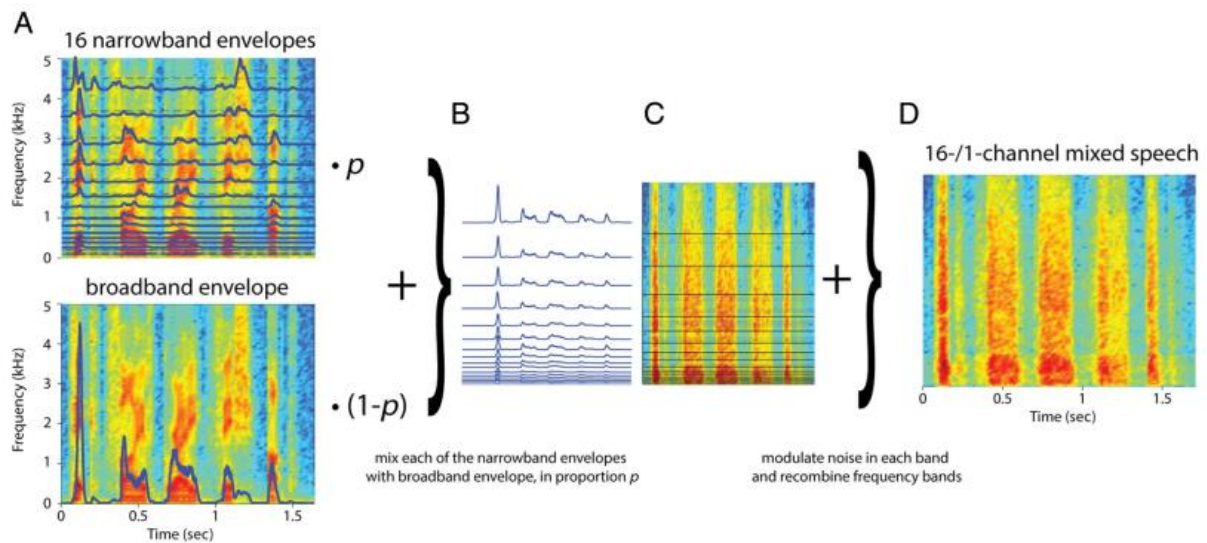


Figure 17. Noise-vocoding. **A.** For noise-vocoding, speech is filtered into various frequency bands and their envelope is extracted (blue). The top panel shows envelopes for 16 frequency bands whereas the bottom panel only uses one (corresponding to the broadband envelope). **B-D.** These envelopes are then used to modulate noise (filtered into the same frequency bands) and later recombined to form noise-vocoded speech. For the experiment described in Zoefel et al. (2020), we mixed 16-channel and 1-channel speech. Here, p corresponds to the ratio between the two. Note that p 's of 1 and 0 yield conventional 16-channel and 1-channel noise-vocoded speech, respectively.

It is a common finding in the field of tACS that **the “best” tACS phase – e.g., the phase that yields highest accuracy in a task – varies across individuals** (Riecke and Zoefel, 2018; Riecke et al., 2018; Wilsch et al., 2018). To account for these individual differences, the maximal or minimal response of each participant can be aligned at an arbitrary (e.g., centre) phase bin before averaging across participants. This has been done to produce results shown in **Fig. 18**, where open circles show bins used for the alignment. These bins cannot be analyzed as they are trivially an extremum (maximum or minimum, depending on the alignment procedure). However, in the presence of a phasic modulation, the rest of the data points should follow a circular shape. For this study, we compared performance in the two bins adjacent to the bin used for alignment with those opposite to it (green vs blue and yellow vs orange in **Fig. 18**). In the presence of a phasic effect, this difference should be positive when performance is aligned to the maximum, or negative when it is aligned to the minimum. **We found such a pattern, but only when tACS was applied bilaterally and with ring electrodes** (right column in **Fig. 18**), not unilaterally with rectangular electrodes (left column). This result does not only highlight tACS as a promising tool to modulate entrainment and speech perception, but also supports the notion that the **choice of stimulation protocol is critical** in brain stimulation research (Zoefel and Davis, 2017). Together, results from these two experiments suggest that **entrainment has a specific, causal influence on speech processing**.

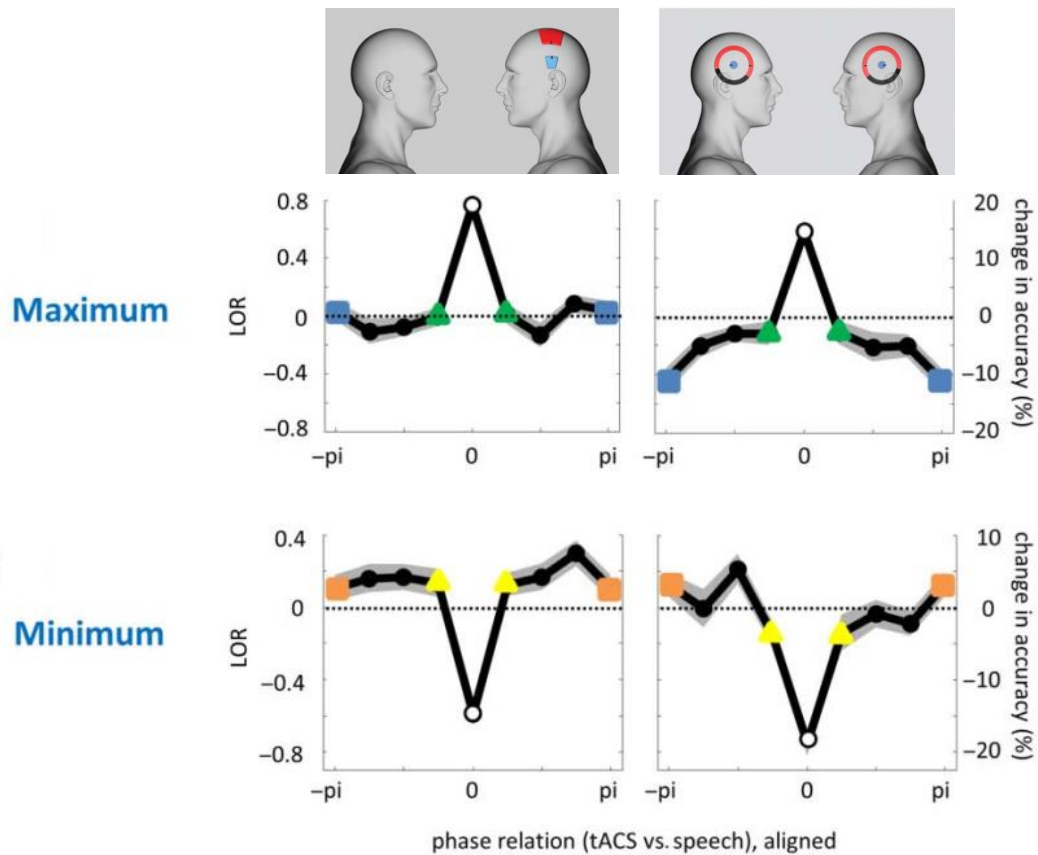


Figure 18. Performance in a word report task, expressed relative to a sham condition (as Log Odds Ratio, LOR). Results are shown for two different tACS electrode setups, and aligned relative to maximal or minimal performance (open circles). Shaded areas show SEM. From Zoefel et al. (2020).

4.2.2 Optimised methods to test for oscillatory phase effects

Advisor: Matthew H Davis

Collaborators: Giancarlo Valente, Lars Riecke

Publications:

Zoefel, B., Davis, M.H., Valente, G., and Riecke, L. (2019). How to test for phasic modulation of neural and behavioural responses. *NeuroImage* 202, 116175.

A crucial question for my research, and for the field of neural oscillations in generally, is whether perception or various neural measures depend on the oscillatory phase. Indeed, many of the results shown in this thesis depend on the outcomes of such analyses. It is all the more surprising that it remained unknown which statistical method is optimal to detect such a phasic modulation of perception or neural activity. In collaboration with a group in Maastricht, I conducted Monte-Carlo simulations (Zoefel et al., 2019) to determine **how reliably each of several different statistical methods can detect a true oscillatory modulation of**

performance or reject an absent one. We also tested how these results depend on two different classes of parameters (**Fig. 19**): Parameters that describe the nature of the phase effect and reflect underlying neural processes (effect size, width, and asymmetry), and parameters that depend on the specific experimental design (e.g., number of trials, number of phase bins used in the analysis).

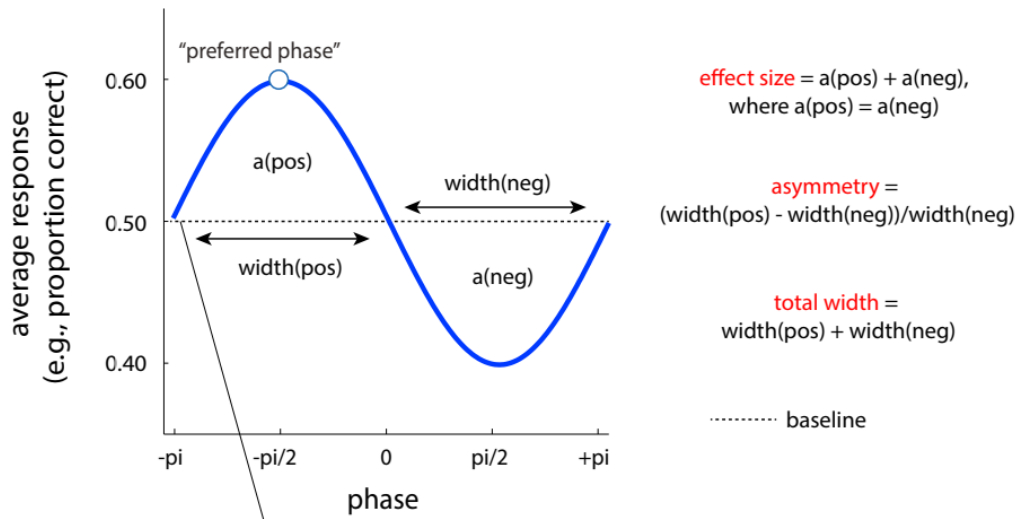


Figure 19. Modelling phasic effects. We simulated data with a phasic modulation of task performance. We varied asymmetry (relative width of positive and negative deflections), total width (proportion of the cycle covered by negative and positive deflections) and effect size of this phasic effect. From Zoefel et al. (2019).

We made various observations that continue to guide my research (**Fig. 20**). (1) Methods that **regress single-trial responses on circular predictors** (sine and cosine) generally outperform other methods. The fit of a sinusoidal function to data, as described for the tACS-fMRI study above (Zoefel et al., 2018a), is part of this class of methods. (2) Sensitivity of methods to detect phasic effects **increases rapidly with the number of trials** used, whereas the number of phase bins does not play an important role. (3) The question of whether targets are randomly presented at different (e.g., neural) phases and sorted post-hoc, or imposed specifically (e.g., with tACS) does not affect sensitivity to detect phase effects. (4) These observations were not strongly affected by the shape of the effect (**Fig. 19**), although symmetric effects were easier to detect than asymmetric ones.

In sum, **this study lays a foundation for optimized experimental designs and analyses in future studies** – including ours described in many of the following sections – investigating the role of neural phase for perception and behavior.

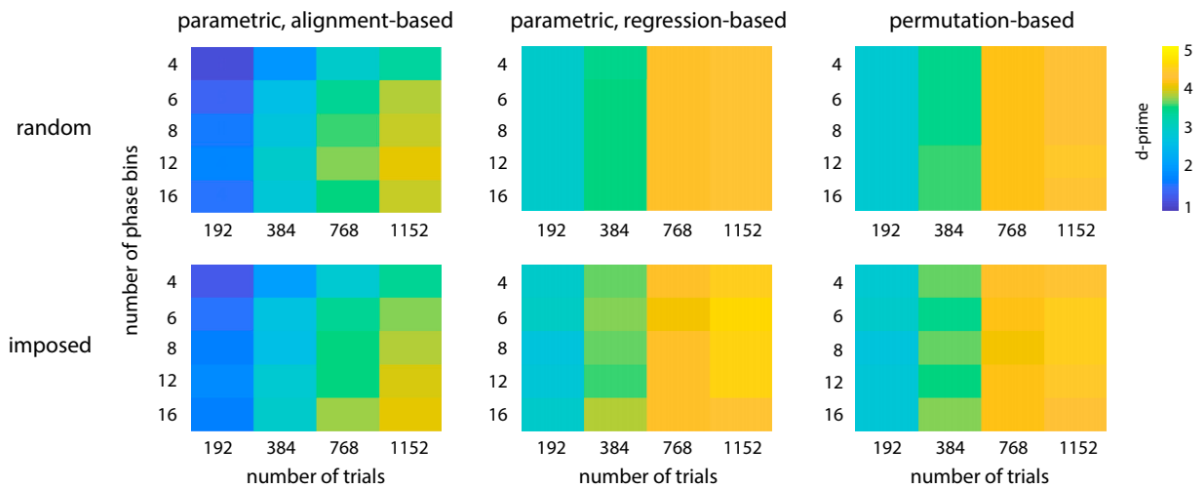


Figure 20. Results from simulation study. Sensitivity of methods to detect a phasic effect is color-coded. Random and imposed designs refer to the possibility of sorting neural phases post-hoc or imposing a limited number of phases during data collection, respectively. The three columns refer to three method categories tested. Each category comprised several methods (not shown); the one with highest sensitivity was selected for this illustration. From Zoefel et al. (2019).

4.2.3 Intelligibility improves the perception of speech rhythm

Advisor: Matthew H Davis

Collaborator: Rebecca A Gilbert

Publications:

Zoefel, B., Gilbert, R.A., and Davis, M.H. (2023). Intelligibility improves the perception of speech rhythm. *PLOS One*. 18: e0279024.

In a final post-doc project (Zoefel et al., 2023), I examined the role of intelligibility for the perception of speech rhythm. Again, we relied on the fact that noise-vocoding can manipulate speech intelligibility without affecting broadband changes in amplitude, or changes in rhythm perception that such amplitude changes would cause (**Fig. 17**). In the tACS-fMRI experiment described above (Zoefel et al., 2018a), participants were asked to detect an irregularity in the rhythm of the presented (intelligible or unintelligible) speech sequences. This irregularity was introduced by shifting one of the words towards another. Interestingly, **irregularity detection was reliably better for intelligible as compared to unintelligible speech** (not shown). For this project, we designed a series of experiments to follow up on this finding. Most of these were run online using Prolific, a web-based participant recruitment and payment platform (<https://www.prolific.ac/>).

In a first (lab) experiment, we replicated the original effect (better rhythm perception for 16-channel than for 1-channel speech) in a forced-choice paradigm (participants had to choose between “regular” and “irregular”) and with more irregular trials as compared to the original

paradigm (Zoefel et al., 2018a). In a second (online) experiment, we addressed the fact that 16-channel speech is spectrally more complex than 1-channel speech. By including a condition that has the same spectral complexity, but is unintelligible (16-channel rotated speech), we found that **improved performance for intelligible speech cannot be explained by differences in spectral complexity**, as performance was also disrupted for 16-channel spectrally rotated speech (**Fig. 21A**).

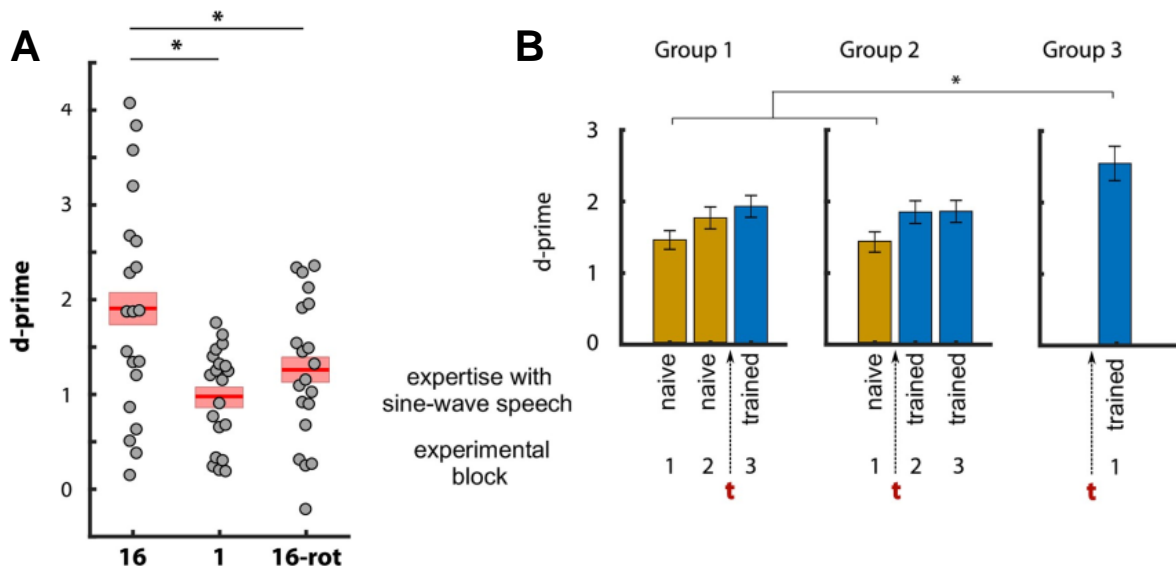


Figure 21. A. Sensitivity (d -prime) to detect rhythmic irregularities in 16-channel, 1-channel or 16-channel rotated noise-vocoded speech. Points represent data from individual participants, mean and SEM are shown as a red line and colored area, respectively. **B.** Sensitivity in detecting rhythmic irregularities in sine-wave speech. Bars shows average performance for three experimental blocks (only one for group 3), a red t indicates when participants were trained to perceive the stimulus as speech. Note that participants' improvements are larger when they are trained prior to the block as compared when they are not (across groups, for the same experimental blocks). Error bars show SEM across participants. From Zoefel et al. (2023).

In a third (online) experiment, we used sine-wave speech (Remez et al., 1981) that is often heard as a series of whistles by naïve listeners but can be understood when listeners are trained to identify it as speech sounds. We trained three different groups of subjects to perceive sine-wave speech as speech, but, crucially, at different moments during the experiment. We were thus able to contrast irregularity detection between groups that had already learnt to perceive the sine-wave stimulus as human speech with others that were still naïve at the same moment in time. We found that the **ability to detect an irregularity in the stimulus rhythm is indeed improved when the sounds are identified as speech** (in trained participants) as compared to when they are not (in naïve participants) (**Fig. 21B**). As the stimulus is physically identical for these contrasts, this result confirms that linguistic information contributed to auditory rhythm perception. This result also supports our previous finding that intelligible speech contributes specifically to neural entrainment and oscillatory dynamics (Zoefel and VanRullen, 2015a; Zoefel et al., 2018a). Nowadays, there is converging evidence that human speech is a stimulus that plays a particularly important role for auditory neural dynamics, and produces neural activity that is distinct from that in response to sounds

in general. In collaboration with Dr Anne Kösem (Lyon), I have recently summarized such evidence (Zoefel and Kösem, 2022).

4.2.4 Summary: Post-doc work

Most of my post-doc time was dedicated to the development of tACS protocols that aim to manipulate neural entrainment underlying speech perception. In several experiments, we were able to demonstrate a causal role of neural entrainment for speech processing. We also found that this role might be specific, or at least stronger, for the processing of intelligible speech and the perception of its rhythm. Finally, we developed statistical analyses that are tailored to detect the corresponding phasic effects with high sensitivity. All of these results continue to influence my research, as described in the next chapters. Meanwhile, we were able to show that the effects observed are indeed due to an entrainment of endogenous neural oscillations (**Section 4.3**), and other work highlights the involvement of the cerebellum in these processes (**Section 5.1.1**). In parallel, we are working on novel paradigms to manipulate neural oscillations (**Section 5.3**) and to understand their role for auditory and speech perception (**Section 5.2**). All of these together form the next part of this thesis.

4.3. First steps as principal investigator (2019 – 2022)

During my final post-doc year in Cambridge (2019), I started to run my own studies as principal investigator. For this purpose, I acquired a small grant from the Leverhulme Trust, and was able to employ Sander van Bree as my first trainee beyond student supervision (**Sections 4.3.1 and 4.3.2**). In 2020, when I joined the *CNRS-CerCo* in Toulouse as a *chargé de recherche* (roughly equivalent to an assistant professor), my work was, like that of many of us, temporarily disrupted by the Covid pandemic. Together with Master student Sylvain L’Hermite, we used recent technological improvements to run an extensive online study on the role of neural oscillations for auditory perception (**Section 4.3.3**). The following sections also include the theoretical background for a line of research that I will work on in the coming years, and therefore set the stage for **Chapter 5** on my future research.

Rhythmic entrainment echoes: A window into endogenous neural oscillations

The field of *neural entrainment* is fundamentally based on the assumption that **rhythmic electro- or neurophysiological responses to a rhythmic stimulus reflect endogenous neural oscillations aligned to the stimulation** (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Thut et al., 2011a; Giraud and Poeppel, 2012; Obleser and Kayser, 2019). However, this assumption is rarely tested.

This is because **each event in a rhythmic sequence will evoke a neural response**; due to the regularity of the stimulus, the evoked responses will appear regular as well (middle row in **Fig. 22**), without necessarily involving endogenous neural oscillations (bottom row in **Fig. 22**) (Capilla et al., 2011; Thut et al., 2011a; Keitel et al., 2014; Zoefel et al., 2018b). Strikingly, this problem has already been described by Walter and Walter (1949), but continues to affect our field today. In their paper, they explain that rhythmic brain responses could be produced, among several alternatives, by a “*fusion of evoked responses giving an accidental appearance of rhythmicity*” or by a “*true augmentation or driving of local rhythms at the frequency of the stimulus*”.

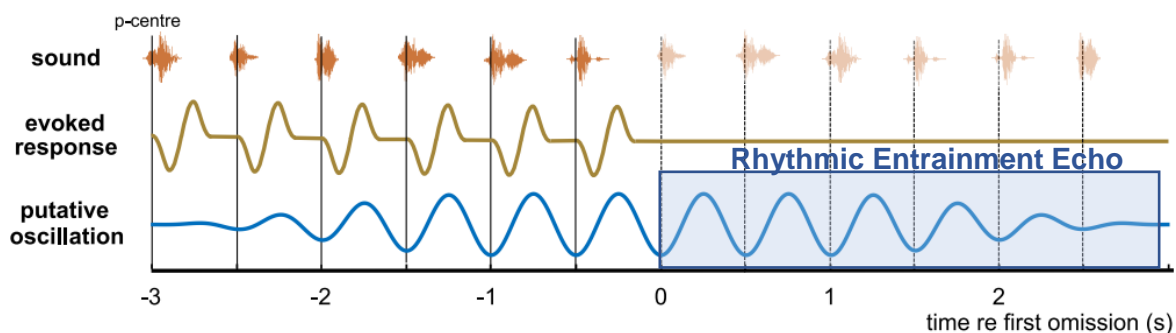


Figure 22. Rhythmic entrainment echoes. A rhythmic neural response that is measured during rhythmic sounds (here speech; top) cannot distinguish endogenous neural oscillations (bottom) aligned to the stimulus rhythm from regular stimulus-evoked responses (middle). However, only endogenous oscillations outlast the rhythmic stimulus (leading to a rhythmic entrainment echo). Modified from van Bree et al. (2021).

One solution to this problem is based on the notion that **endogenous oscillations should linger for some time** after having been entrained, similar to a swing that has been pushed, whereas other, evoked brain activity will disappear rapidly when no stimulus is present. **These rhythmic entrainment echoes (Fig. 22)** – defined as rhythmic brain responses that are produced by a rhythmic stimulus and persist after its offset – therefore play a crucial role in distinguishing entrained neural oscillations from other brain responses that are not endogenously rhythmic.

Rhythmic entrainment echoes are also interesting for other reasons. For example, they might give us insights into whether and when participants expect a stimulus to occur, and how properties of the preceding rhythmic stimulus affect this expectation. **Entrainment echoes might therefore be intimately linked to temporal expectation**, a notion that I will come back to below and that is much in line with previous models considering neural oscillations a mechanism of attentional selection in time (Large and Jones, 1999; Lakatos et al., 2008; Schroeder and Lakatos, 2009; Lakatos et al., 2013a). My first years as principal investigator were dedicated to the investigation of rhythmic entrainment echoes, and this research is explained in the following.

4.3.1 Entrainment echoes in the MEG

Collaborators: Ediz Sohoglu, Matthew H Davis, Andrea Alamia

Supervision: Sander van Bree (research engineer)

Publications:

van Bree, S. van, Sohoglu, E., Davis, M.H., and **Zoefel, B. (2021)**. Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLoS Biol.* 19, e3001142.

van Bree, S., Alamia, A., and **Zoefel, B. (2022)**. Oscillation or not-Why we can and need to know (commentary on Doelling and Assaneo, 2021). *Eur. J. Neurosci.* 55, 201–204.

We first tested whether rhythmic speech produces rhythmic entrainment echoes in the MEG (van Bree et al., 2021). As in previous experiments, participants were presented with sequences of intelligible (16-channel) or unintelligible (1-channel) noise-vocoded speech and were asked to detect rhythmic irregularities (**Fig. 23**). Sequences were presented at 2 Hz or 3 Hz. We then quantified rhythmic entrainment echoes as rhythmic brain responses that follow and outlast the stimulus rhythm. To do so, we contrasted MEG responses at a frequency that corresponds to the stimulus rate (e.g., 2-Hz responses after a 2-Hz stimulus) with responses at the same frequency but after a different stimulus rate (e.g., 2-Hz responses after a 3-Hz stimulus). This contrast allowed us to remove generic effects that are produced by stimulus offset (e.g., omission response) and affect the outcomes of spectral analyses, but are not

specific to the rate of the stimulus. We labelled the resulting measure “*rate-specific response*” index (RSR).

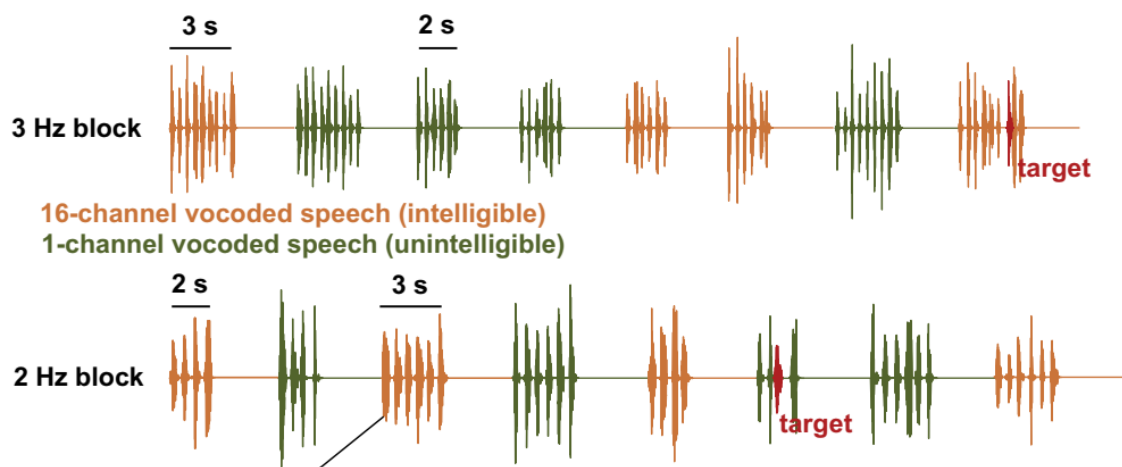


Figure 23. Experimental design. Participants were presented with rhythmic sequences of words that were 2 or 3 seconds long, intelligible (16-channel) or not (1-channel vocoded speech) and presented at 2 or 3 Hz. They were asked to detect an irregularity in the rhythm (red). We tested whether rhythmic MEG responses to these rhythmic stimuli persist in the silent periods between sequences. From van Bree et al (2021).

We found **reliable rate-specific responses outlasting the stimulus rhythm only after intelligible (but not after unintelligible) speech**. **Fig. 24A** shows the topography of these rhythmic entrainment echoes (plus signs). **Fig. 24B** shows that these were localised to parietal areas, although a more thorough analysis revealed an involvement of the cerebellum (see **Section 5.1.1**). **Fig. 24C** illustrates the time course of these echoes (0 corresponds to the first omitted word after a rhythmic speech sequence), for selected sensors with the strongest RSR during (green) or after (red) the stimulus, respectively. Importantly, the entrainment echo seems to cover most of the post-stimulus time period rather than being dominated by a single outlier.

Together, we reported evidence for an **involvement of endogenous neural oscillations in neural entrainment to human speech**. The fact that entrainment echoes were only found after intelligible speech underlines the **unique role that speech might play for auditory neural oscillations** and that was already suggested by our previous work (Zoefel and VanRullen, 2015a; Zoefel et al., 2018a, 2023; Zoefel and Kösem, 2022).

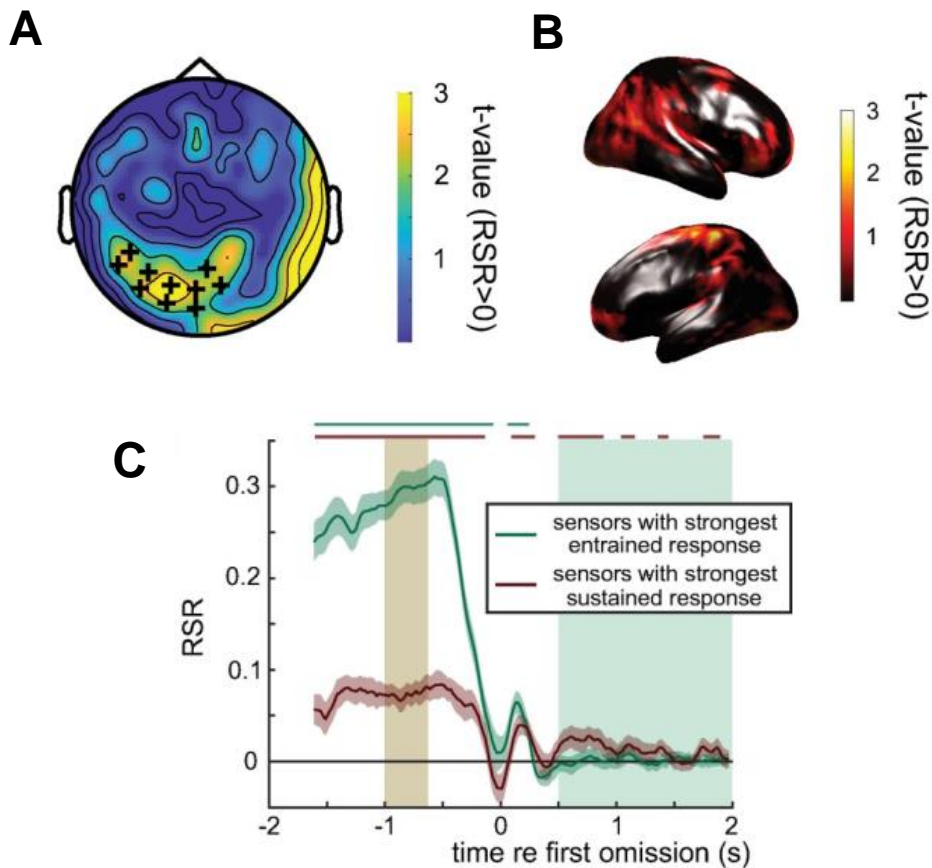


Figure 24. **A.** Rate-specific responses (RSR) after intelligible speech sequences. Plus signs show a cluster of MEG sensors with statistically reliable RSR (t-test against 0). **B.** Estimated neural sources for results shown in A. **C.** Time course of RSR, shown for selected sensors with the strongest RSR during or after the stimulus, respectively. Sensors selected for the red curve correspond to those shown with plus signs in A. Shaded areas show SEM. From van Bree et al (2021).

4.3.2 Entrainment echoes after tACS

Collaborators: Ediz Sohoglu, Matthew H Davis

Supervision: Sander van Bree (research engineer)

Publications:

van Bree, S. van, Sohoglu, E., Davis, M.H., and **Zoefel, B. (2021)**. Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLoS Biol.* 19, e3001142.

During the presentation of a rhythmic stimulus, neural responses fluctuate at the frequency of the stimulus, without necessarily involving genuine endogenous oscillations. What was described above for neurophysiological recordings also holds for tACS, commonly assumed to entrain neural oscillations (Herrmann et al., 2013b): Rhythmic changes in speech perception, produced by the stimulation (Riecke et al., 2018; Wilsch et al., 2018; Zoefel et al., 2018a,

2020; Keshavarzi et al., 2020), might simply reflect the rhythmicity of the current, rather than a rhythmic neural process. However, not only the issue but *also the solution holds* for tACS: **If rhythmic effects are measured after the offset of the electric stimulation, when oscillations are assumed to persist, but effects produced by the current *per se* can be ruled out (as no stimulation is applied), then we can more confidently assume that endogenous oscillations underlie these effects.**

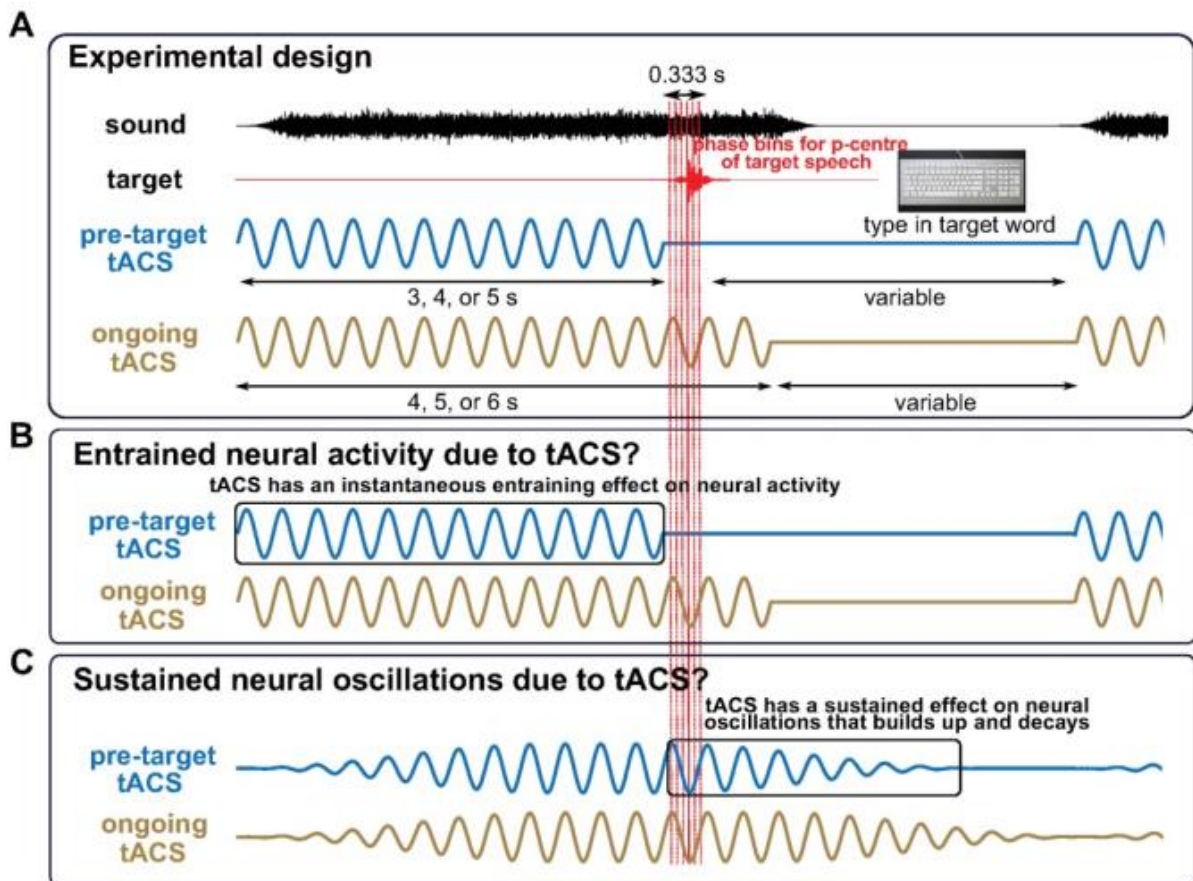


Figure 25. Testing for entrainment echoes produced by tACS. **A.** tACS was applied while participants listened to and typed in target words embedded in noise. tACS was either turned off immediately prior to (pre-target tACS) or after tACS (ongoing tACS). **B.** If rhythmic effects are exclusively due to the current applied, phasic effects on word report accuracy should only be visible in the ongoing tACS condition. **C.** If they are also caused by entrained neural oscillations, phasic effects should also be visible in the pre-target tACS condition. From van Bree et al (2021).

We therefore tested for rhythmic entrainment echoes in speech perception after tACS. The experimental paradigm is shown in **Fig. 25**. Participants were asked to listen to and type in acoustically presented target words (red in **Fig. 25A**) embedded in noise (black in **Fig. 25A**). Accuracy in this task was quantified using *Levenshtein distance*, which is the minimum number of edits (deletions, insertions etc.) necessary to change response into target word. While participants completed this task, tACS was applied at 3 Hz in two different conditions, and using the ring electrode configuration shown in **Fig. 18** (right, top). In one condition ("ongoing tACS" in **Fig. 25**) it was applied so that the target word was presented at one out of six

possible tACS phases in each trial, resembling our previous work. In the second condition (“pre-target tACS”), tACS was turned off immediately before the target word, and the latter was presented so that it falls at one out of six tACS phases, *had it continued*. If tACS produced entrainment echoes, then accuracy in reporting the target word should still depend on the rhythm of tACS even after it has been turned off (**Fig. 25C**) whereas no such effect would be expected if rhythmic effects in perception are exclusively due to changes in current applied (**Fig. 25B**). tACS was applied for 3, 4, or 5 seconds before presentation of the target word to examine the role of stimulation duration for the hypothesized effect.

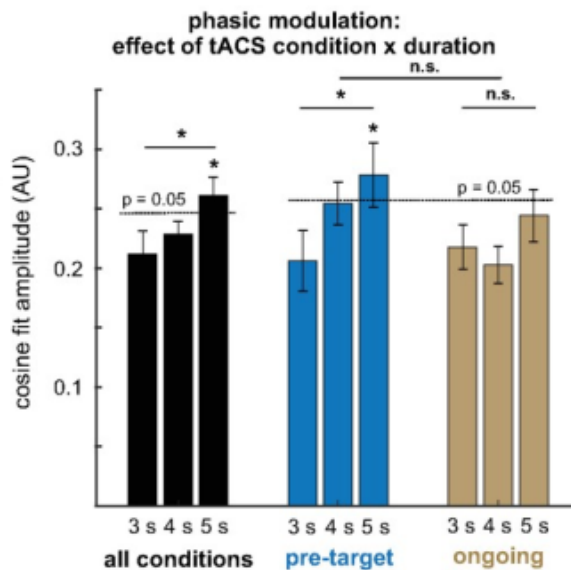


Figure 26. Phasic modulation of word report accuracy across and separately for the two tACS conditions, and how it depends on tACS duration. Error bars show SEM. From van Bree et al (2021).

We used statistical methods that were revealed to be optimal to detect phasic effects in the simulation study described above (**Section 4.2.2**; Zoefel et al., 2019). Here, the amplitude of a sine function fitted to the data (cf. **Fig. 1B**) reflects how well perceptual outcomes can be predicted by tACS phase (or its imaginary continuation) and therefore the strength of the phasic modulation. These amplitudes are shown in **Fig. 26**. We found an overall effect of tACS (irrespective of experimental condition) that increases with stimulation duration, in line with entrained neural oscillations (that should take some time to align to the current; Fröhlich, 2015). Most importantly, we found a phasic modulation of word report accuracy in the condition where tACS was turned off before the target was presented (pre-target tACS). Although the effect was not statistically reliable in the ongoing tACS condition, it was also not reliably different from the pre-target tACS condition. Together, the most important conclusion is that **tACS produces entrainment echoes in speech perception and can therefore be assumed to entrain endogenous neural oscillations.**

4.3.3 Entrainment echoes in auditory perception

Supervision: Sylvain L’Hermite (M.Sc.)

Publications:

L’Hermite, S., and Zoefel, B. (2022). Rhythmic entrainment echoes in auditory perception. BioRxiv. doi: <https://doi.org/10.1101/2022.12.07.519456>

In a series of lab and online experiments, we examined rhythmic entrainment echoes in (human) auditory perception. Experiments followed a relatively simple paradigm, illustrated in **Fig. 27**. Participants were presented with a pure tone that was *amplitude-modulated* (AM) rhythmically, and that was followed by a target tone they were asked to detect. The role of the AM tone was to entrain oscillations and produce a rhythmic entrainment echo (red in **Fig. 27**). The target was presented at different delays relative to the AM tone and thus used to “sample” the entrainment echo.

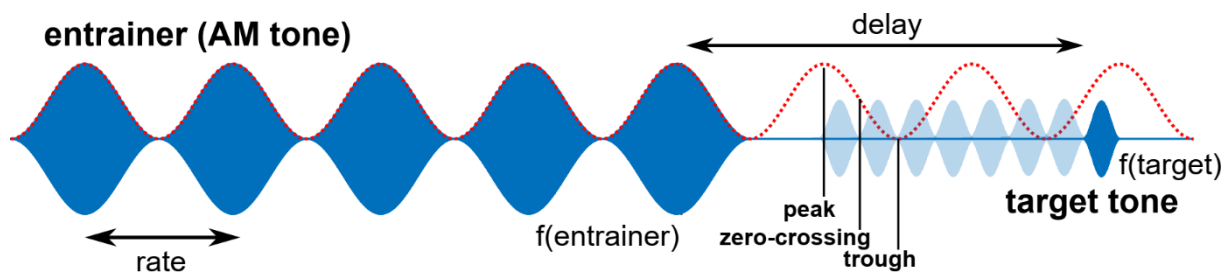


Figure 27. *Experimental paradigm. A target pure tone was presented at variable delays after another pure tone that was rhythmically modulated in amplitude (AM), and presented at a certain rate. The hypothesized entrainment echo, produced by the AM tone, is shown as red dotted line. All possible delays for the target tone (one per trial) are shown. These delays corresponded to the peak, zero-crossing, or trough of the preceding AM rhythm (four delays per cycle). From L’Hermite and Zoefel (2022).*

This paradigm was inspired by a study by Hickok and colleagues (2015) that had already demonstrated entrainment echoes in auditory perception. In that study, a 3-Hz AM noise produced entrainment echoes in the detection of a subsequently presented target tone. Targets were most likely to be detected when they occurred in anti-phase with the preceding AM noise. Since then, several research groups have tried to replicate the results, with mixed outcomes (Lin et al., 2022; Saberi and Hickok, 2022a; Sun et al., 2022). This led to a current debate on the effectiveness of the paradigm (Saberi and Hickok, 2022b, 2022c; Sun et al., 2022) and even to a multi-lab effort with the aim of replicating results across labs (Henry et al., 2022). In our study, we tried to improve the original paradigm by including several modifications.

First, we replaced the entraining AM noise with an AM tone. **This allowed us to test whether entrainment echoes in auditory perception are tonotopically organised**

(which is not possible with noise, as it contains acoustic energy across a wide range of frequencies). As explained in **Section 4.1.2**, previous work in non-human primates (O’Connell et al., 2011; Lakatos et al., 2013a; O’Connell et al., 2014), including ours (Zoefel et al., 2017), showed that entrainment of neural activity in auditory cortex is organized according to sound frequency. That is, the phase of stimulus-aligned activity is opposite depending on whether it was measured in parts of auditory cortex that are tuned to the sound frequency of the stimulus or not (cf. **Fig. 9**). It is therefore plausible that entrainment echoes are organised tonotopically as well, and the use of noise might have led to unclear results in the past because it is tonotopically hard to define. In this study, we independently varied the sound frequencies of AM tone and target tone to test for tonotopic entrainment echoes. Based on results in non-human primates, we expected that targets are more likely to detect if they are presented in phase with the preceding AM tone *and* if the two sound frequencies match (orange in **Fig. 28**), but in anti-phase if they mismatch (black in **Fig. 28**).

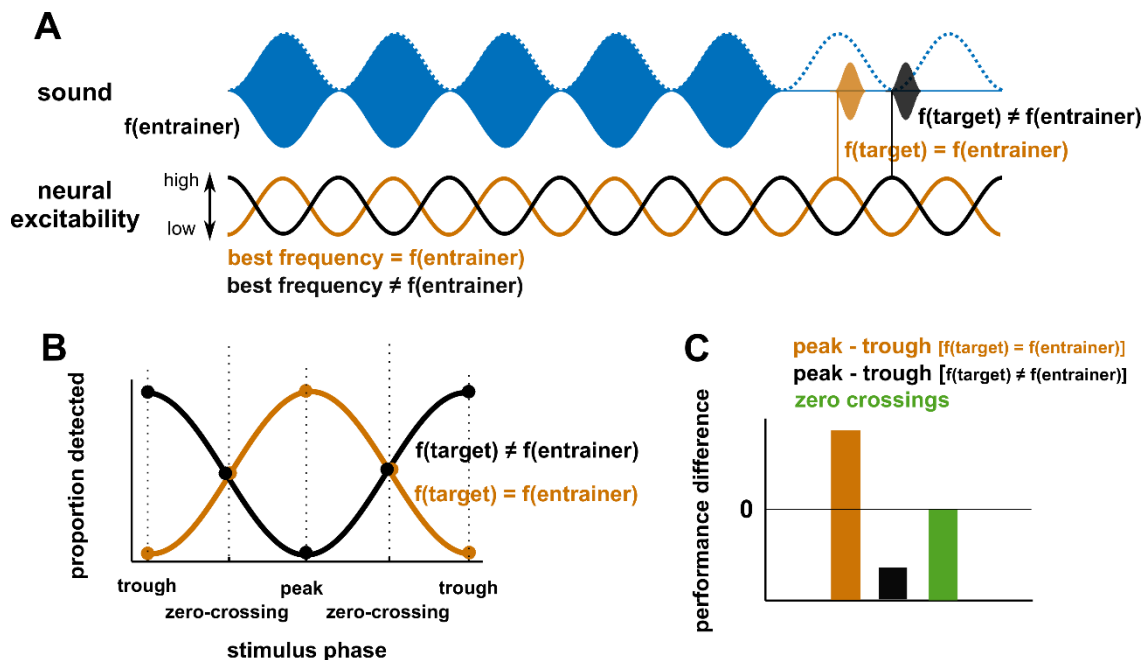


Figure 28. A. We assumed that a rhythmic (AM) stimulus entrains fluctuations in neural excitability (i.e. oscillations) differently, depending on whether they occur in an auditory region that is tuned to the sound frequency of that entraining stimulus (orange) or not (black). In our paradigm, the role of the target tone was to “probe” these oscillations: The likelihood of target tone detection should depend on neural excitability in the region processing its sound frequency. **B,C.** If sound frequencies of AM tone and target tone match, tonotopic entrainment echoes would lead to a highest proportion of detected targets at the peak of the preceding AM rhythm (B), and (C) to a positive difference for peak vs trough performance. If the sound frequencies mismatch, tonotopic entrainment echoes would lead to a highest proportion of detected targets at the trough of the AM tone (B), and (C) a negative difference for peak vs trough performance. In both cases, we expected a(n absolute) peak-trough performance difference that is larger than that between the two zero-crossings tested (green in C). From L’Hermite and Zoefel (2022).

Second, we asked whether some stimulus rates produce stronger rhythmic entrainment echoes than others. A **general property of neural oscillations is that they have an eigenfrequency**, i.e. a frequency they oscillate at in the absence of stimulus input, or a

stimulus rate that they respond to most strongly (Thut et al., 2011a; Fröhlich, 2015; van Bree et al., 2022). The *eigenfrequency* of audition, and in particular that of entrainment echoes, remained to be examined in detail. One previous study (Farahbod et al., 2020) found strongest echoes for relatively slow rates (~ 2 -3 Hz), but the number of subject tested was low ($N = 3$ -5). Another important addition to the field was therefore the sample size used in our study, which was considerably higher (total $N = 154$).

To quantify perceptual entrainment echoes, we contrasted the proportion of detected targets presented at the peak and trough of the preceding AM rhythm, respectively (**Fig. 28B,C**). In the presence of entrainment echoes, this difference should be positive or negative, depending on whether sound frequencies of AM tone and target tone match or mismatch, respectively (**Fig. 28C**). In contrast, target detection should not differ much between the two zero-crossings of the rhythm. We therefore also compared the peak-trough performance difference with the corresponding difference between zero crossings (**Fig. 28C**).

In a first (lab) experiment ($N = 16$), we found **entrainment echoes in auditory perception** – reflected by a *peak-trough performance difference* that was reliably different from 0 and from the corresponding difference between zero-crossings (**Fig. 29A**) – but **exclusively after 6-Hz stimulation**. This result supports the notion that the auditory system is tuned to process sounds that mimic certain characteristics of human speech (Poeppl and Assaneo, 2020; Zoefel and Kösem, 2022). Across languages, speech contains amplitude modulations between ~ 2 and 8 Hz (Ding et al., 2017; Varnet et al., 2017); the 6-Hz *eigenfrequency* revealed here falls into that range. Strikingly, a similar 6-Hz rhythm was recently demonstrated even at the cochlear level (Köhler et al., 2021). In this experiment, the proportion of detected targets peaked in anti-phase with the preceding AM tone (**Fig. 29B**). This result is **in line with tonotopic entrainment echoes**, as sound frequencies of AM and target tones mismatched in this experiment (matching sound frequencies were not tested).

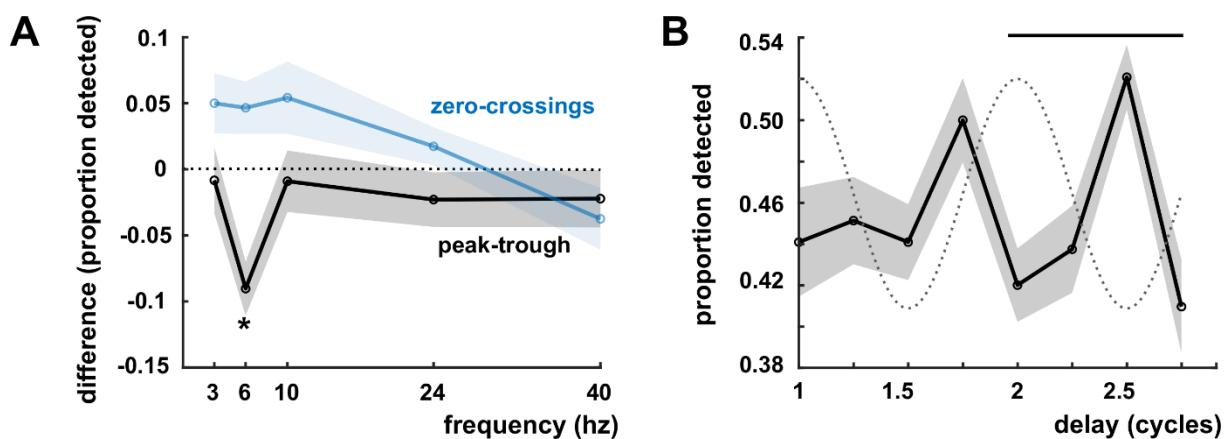


Figure 29. Results from Experiment 1. A Differences in the proportion of detected targets that would have occurred at the peak and trough of the preceding AM tone (had it continued), or at the two zero-crossings (cf. Figs. 27,28). **B.** Performance as a function of the delay of target presentation, relative to the offset of a 6-Hz AM tone. The dashed line shows the rhythm of the AM tone (had it continued). Continuous lines represent delays with statistically reliable entrainment echoes. Shaded areas show SEM. From L’Hermite and Zoefel (2022).

In a second (online) experiment (N = 47), we fixed the stimulus rate to 6 Hz and tested the hypothesized tonotopical organisation in more detail. We again found **rhythmic entrainment echoes in auditory perception, however in opposite phase** to what was observed in Experiment 1: Most targets were detected in phase with the AM rhythm when sound frequencies of AM tone and target tone mismatched (black in **Fig. 30**), and in anti-phase when they did match (orange).

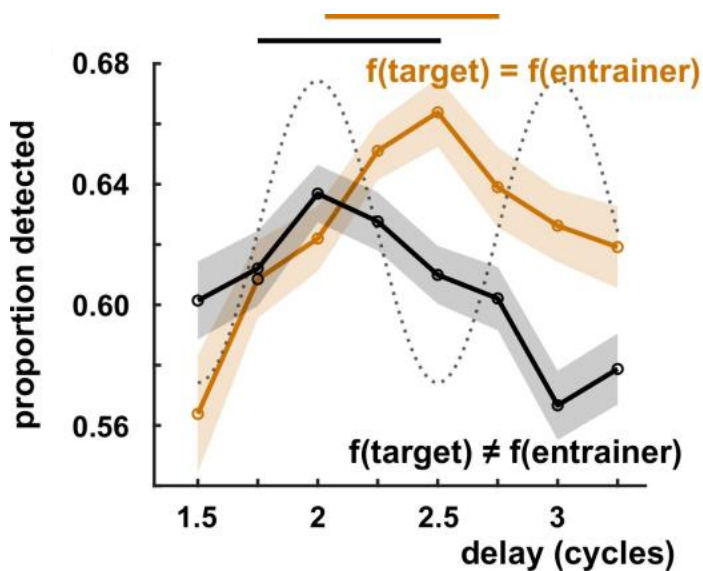


Figure 30. Same as Fig. 29B, but for Experiment 2. The two lines correspond to conditions where sound frequencies of AM tone and target tone matched (orange) or not (black), respectively. Note that only the latter case was tested in Experiment 1 (Fig. 29).

This result was **not in line with our expectations** (Fig. 28), and we therefore designed a third (online) experiment (N = 49) to reveal potential reasons behind this discrepancy. Participants were able to predict the sound frequency of the target in Experiment 1 (because it

was identical in each trial) but not in Experiment 2 (where it was randomly selected in each trial). In Experiment 3, some blocks used predictable combinations of sound frequencies for AM and target tones, whereas other block used unpredictable, random ones. We found, however, results that closely resembled those from Experiment 2, irrespective of the predictability of sound frequencies (not shown).

In a fourth (online) experiment (N = 42), we therefore explored another hypothesis to explain differences between experiments. If a stimulus, such as a pure tone, is repeated, then neural responses to this stimulus decrease over time (Lange, 2009; Costa-Faidella et al., 2011; Herrmann et al., 2013a). This is a phenomenon that can be described as **neural adaptation or habituation** (sometimes also called repetition suppression) (Bäuerle et al., 2011; Segaert et al., 2013; Prado-Gutierrez et al., 2015). After habituation, any deviance in the expected presentation time or stimulus identity produces a stronger neural response, and can lead to a higher probability of the stimulus to be detected (Ulanovsky et al., 2003; Khouri and Nelken, 2015). Importantly, **such a scenario predicts entrainment echoes that are in opposite phase as compared to those predicted from the conventional theory of neural entrainment.** Targets would be least expected in anti-phase with the preceding rhythm (only) if their sound frequency matches that of the AM stimulus, and vice versa for a mismatch. Such a pattern is precisely what we observed in Experiments 2 and 3. It is possible that in those experiments habituation was particularly strong, as the rate of the AM stimulus was constant across trials, whereas it varied across trials in Experiment 1. In Experiment 4, we tested this hypothetical explanation by again varying stimulus rate across trials (this time with a finer resolution of stimulus rates, centred at 6 Hz). As in Experiments 2 and 3, AM and target tones either had the same or different sound frequencies, randomly selected in each trial.

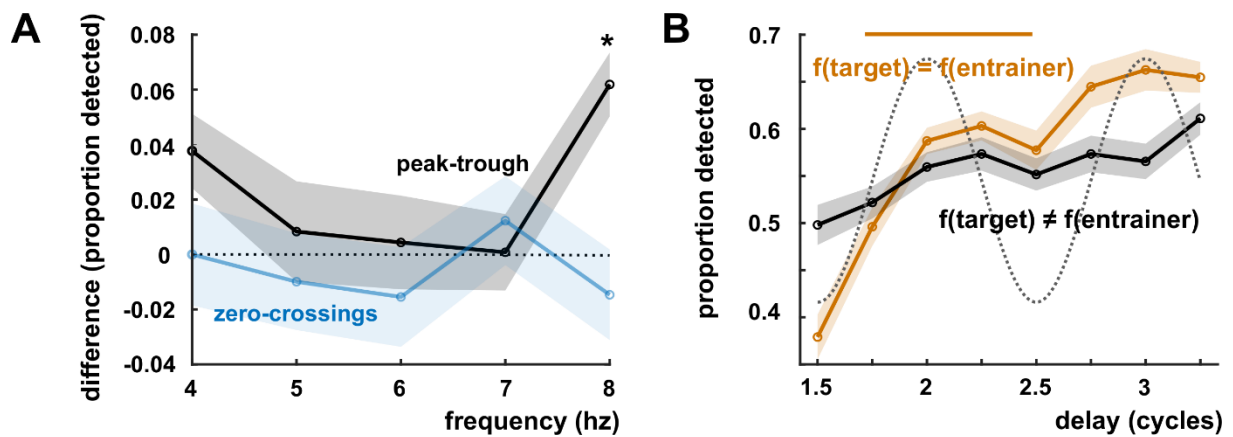


Figure 31. Same as Fig. 29A,B, but for Experiment 4.

We found statistically reliable entrainment echoes, but only after 8-Hz stimulation (Fig. 31A). Importantly, **results again resembled those predicted from tonotopic entrainment:** When the two sound frequencies matched, performance was best in phase with the preceding stimulus rhythm (orange in **Fig. 31B**). However, no such effect was found for the mismatch condition (black). Together, these results can be seen as tentative evidence that repetition-related habituation – most prominent in Experiments 2 and 3 – and neural entrainment – most prominent in Experiments 1 and 4 – exercise **competing, opposite influences on perception**. So far, this notion is speculative and we will aim to confirm it in future work. This could be done by using EEG to measure neural adaptation during the rhythmic stimulus, or by replacing the rhythmic stimulus with an irregular one that should not lead to strong adaptation. If confirmed, it would have important consequences for experimental design, as mixing two counteracting processes might lead to falsely negative or conflicting outcomes. It might also explain why previous studies on entrainment echoes have not always produced the same outcomes, and contribute to resolving the resulting debate (Lin et al., 2022; Saberi and Hickok, 2022b, 2022c; Sun et al., 2022).

5 FUTURE RESEARCH

My future research is a logical continuation of the previous work described in the preceding chapter. As I will describe in the following, **I am now leading a group of people with the common aim of understanding neural oscillations in the auditory system, and financed by several grants** (ANR, Fondation pour l'Audition, Fondation Fyssen). Inspired from previous results, our overarching aims are to reveal (1) in which situations neural oscillations modulate auditory perception and in which they do not ("*harmful vs helpful*"), (2) how neural entrainment shapes auditory perception, and (3) how neural entrainment can be manipulated with various techniques and experimental protocols to ultimately change perception.

5.1 Rhythmic entrainment echoes

5.1.1 Entrainment echoes in the cerebellum

Collaborators: Sonja Kotz, Omid Abbasi, Joachim Gross

As described in detail in **Section 4.3**, entrainment echoes are rhythmic brain responses that are driven by the rhythm of the stimulus and reverberate afterwards. It is therefore plausible that these echoes reflect participants anticipating the stimulus rhythm. This is in line with initial theories of *neural entrainment* or *temporal attending* that considered temporal prediction a primary function of neural oscillations (Large and Jones, 1999; Schroeder and Lakatos, 2009). Entrainment echoes might be a first step towards substantiating this assumption that was previously difficult to verify.

Revealing the neural sources of rhythmic entrainment echoes was not the focus of the MEG study described above (van Bree et al., 2021) and was therefore restricted to the neocortical surface (**Fig. 24B**). In an unpublished re-analysis of the MEG data, we found that **strongest entrainment echoes are present in the cerebellum (Fig. 32)**. This is a striking observation, given that **patients with cerebellar lesions struggle to adapt to changes in stimulus rhythm** (Schwartz et al., 2016; Stockert et al., 2021), and the role of the cerebellum for temporal sensory predictions in general (Roth et al., 2013; Cao et al., 2017). It is possible that entrainment echoes are causally relevant for temporal predictions and, consequently, an important variable for both the diagnosis and treatment of pathological conditions that are associated with corresponding deficits.

In collaboration with Prof Sonja Kotz (Maastricht), we plan to test the hypothesis that entrainment echoes are both a marker and valuable treatment approach for alterations in temporal prediction abilities. In a first experiment, we will use a specifically developed EEG setup to measure cerebellar activity in patients with cerebellar lesions. We expect that they show reduced entrainment echoes as compared to a control group without lesions, and that these go along with a deficit in adapting to changes in stimulus rate. In other words, **we**

expect a neural marker of temporal prediction to be absent in a population that struggle with this very same cognitive function.

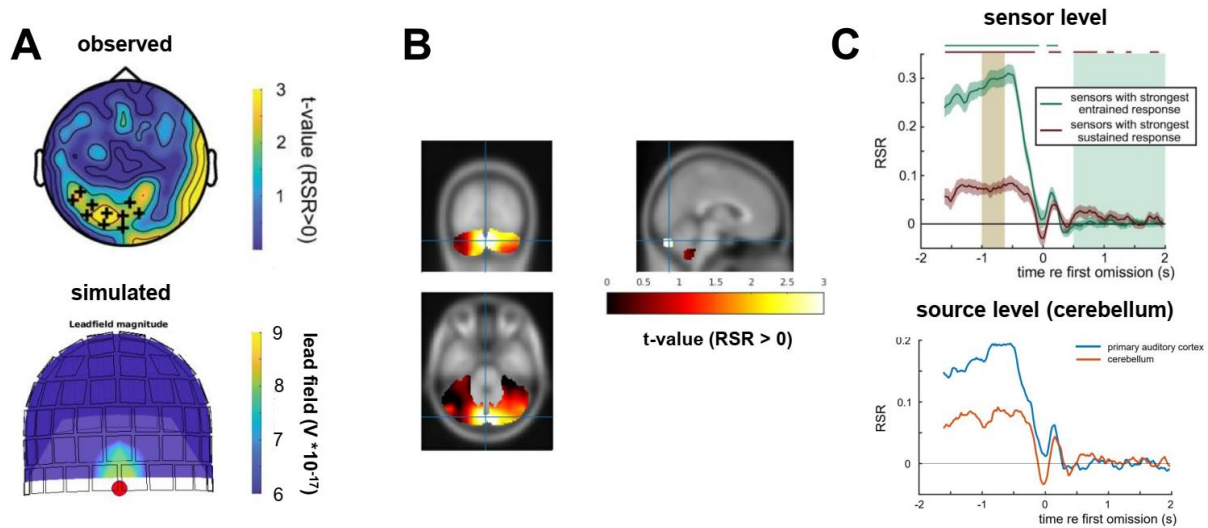


Figure 32. Rhythmic entrainment echoes stem from the cerebellum. **A.** The top panel shows the topography of entrainment echoes from the original study (van Bree et al., 2021), also depicted in Fig. 24A. The bottom panel shows where activity from the cerebellum would be captured with MEG sensors (simulated lead field). Note that the topography shown on top has been “flattened” so that the lowest sensors correspond to the back of the head. The similarity between the two graphs is evidence that the topography observed might indeed stem from the cerebellum. **B.** Entrainment echoes (as RSR after the first omission) in the cerebellum. One panel is shown for each of three possible orientations. The colour illustrates the reliability of the rate-specific response, corresponding to Fig. 24B (cortical surface). **C.** The top panel shows the entrainment echo as a function of time, for selected MEG sensors (also depicted in Fig 24C). The bottom panel shows the echo for estimated sources from primary auditory cortex (blue) and the cerebellum (red). Note that activity from these sources closely resembles that in MEG sensors selected for their strong response during (green in top panel) and after (red in top panel) the rhythmic stimulus, respectively.

In a second experiment, we will apply tACS to manipulate neural entrainment in healthy participants. We will use a tACS electrode configuration designed to stimulate the human cerebellum, based on previous work (Wessel et al., 2023) and modelling of the electric field (<https://simnibs.github.io/simnibs/build/html/index.html>). We will then **enhance or disrupt entrainment echoes with tACS (Fig. 33)**, and hypothesize that this leads to changes in participants’ timing abilities. Participants will be asked to tap along with an isochronous stimulus and to adjust their tapping to the stimulus rate when the latter changes, a task that requires temporal anticipation and involves the cerebellum (Schwartz et al., 2016). We expect entrainment echoes to reflect and “store” previous stimulus rates (**Fig. 33**) that are needed for the interpretation of subsequent information and, consequently, to adapt motor responses accordingly.

We define the stimulus rate before a change as r_1 , the rate after the change as r_2 , and an unrelated rate as r_3 . Our hypothesis states that, when the stimulus changed to r_2 , successful adaptation requires an entrainment echo at r_1 (**Fig. 33A**). We will apply tACS at a frequency that corresponds to r_1 and vary the timing of tACS (i.e. its phase) relative to the stimulus. We

expect that rate adaptation performance depends on the phase relation between tACS and stimulus: tACS either boosts performance (if it is applied at an “optimal” phase relation to the stimulus, presumably in phase with the oscillatory echo produced by the stimulus; **Fig. 33B**) or disrupts it (if is applied in anti-phase to the optimal phase relation; **Fig. 33C**). We will compare this tACS-induced modulation in performance with another condition in which tACS is applied at $r3$. This controls for effects of potential sensations evoked by tACS which might influence performance and can be used to define “baseline” performance. We will further test whether performance is better when tACS is applied at $r1$ but turned off (Fig. 33D) when the stimulus changes to $r2$ (producing an echo that might resemble or boost that produced by the isochronous stimulus) or when it is applied at $r1$ even after the rate change (**Fig. 33B**; potentially boosting the echo further, but also leading to a disequilibrium between internal models of $r1$ and $r2$).

Together, these experiments will examine the **causal role of cerebellar entrainment echoes for temporal prediction and adaptation**, and might reveal a potential approach to improve temporal predictions (via a manipulation of entrainment echoes) in pathological conditions in the near future.

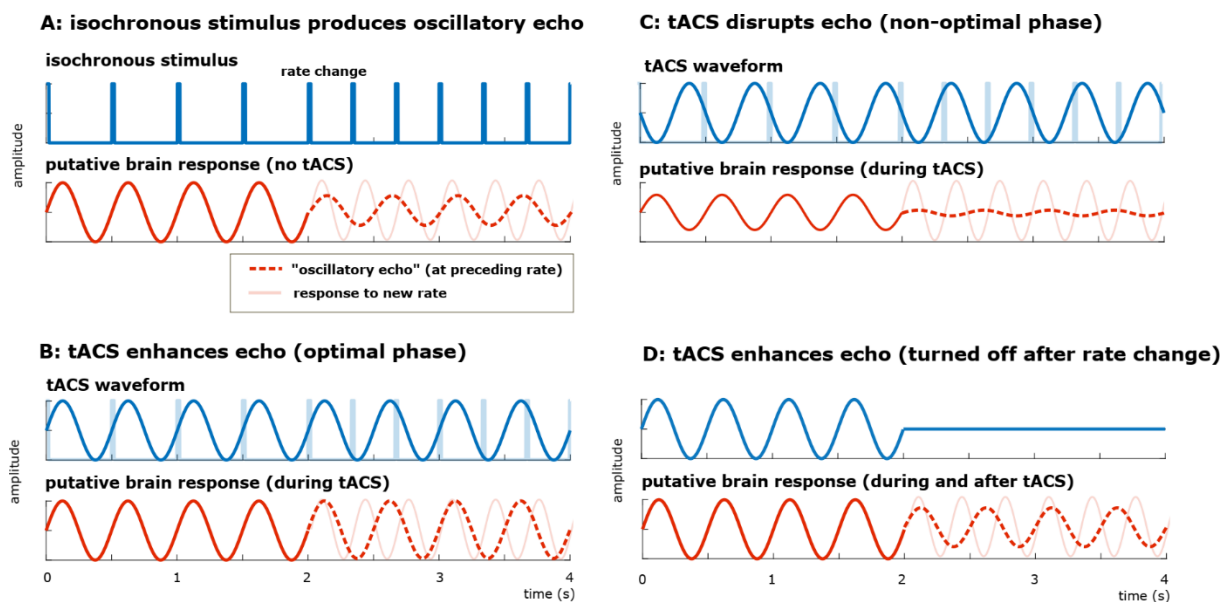


Figure 33. Overview of tACS protocol and hypothesized brain responses in different experimental conditions. **A.** We assume that an isochronous stimulus (top) produces an oscillatory response and corresponding echo (bottom). After a change in stimulus rate, this echo reverberates, by definition, at the preceding stimulus rate. It might allow a more reliable detection of changes in stimulus rate as old (in form of the echo; dashed line) and new rates (faint continuous line) can be compared more directly. **B-D.** We will test whether tACS can enhance (B,D) or disrupt (C) the entrainment echo, and lead to corresponding changes in performance in a rate adaptation task.

In parallel, in collaboration with Prof Kotz and researchers from Münster (Prof Gross, Dr Abbasi), we will analyse neural connectivity profiles between the cerebellum and other cortical regions during neural entrainment and its echo. Much progress has been done on the development of such analyses (e.g., multivariate *Granger causality*; Chalas et al., 2022), and we will apply those to the MEG dataset collected by van Bree et al. (2021). This will reveal

“who drives whom”: If rhythmic cerebellar activity consistently precedes the cortical one (e.g., from auditory cortex), then cerebellar entrainment echoes might reflect rhythmic predictions about stimulus onsets that are relayed to cortical areas. If the cerebellum lags cortical activity, then the cerebellum might instead process the discrepancy between predicted information (with predictions relayed, e.g., from auditory cortex) and the actual outcomes. In this case, entrainment echoes might be particularly prominent in the cerebellum as prediction errors are largest when the stimulus has stopped.

5.1.2 Entrainment echoes after tACS

Supervision: Florian Kasten (post-doc), **Jules Erkens** (post-doc)

We will run an experiment that follows up on the finding that **tACS produces rhythmic echoes in speech perception**, described in **Section 4.3.2**. Participants will be asked to detect and report a target word, embedded in noise, while tACS is applied at 3 Hz (**Fig. 34**). In each trial, tACS will be applied for 5 seconds and then turned off. The target word will be presented at various delays after tACS offset, corresponding to different tACS phases and covering the full 3-Hz cycle of 333.3 ms. This part of the design is identical to the one used in the original study (**Section 4.3.2**). However, in addition, (1) we will record participants' EEG to measure **neural oscillations assumed to underlie the sustained tACS effect**; (2) the timing of the target will cover three instead of one cycle post-tACS (red in **Fig. 34**), allowing us to **test how long the sustained effect lasts**; and (3) in a separate condition, participants will be asked to detect a simple auditory target (e.g., pure tone), presented at the same delays relative to tACS offset. In this condition, we can test whether tACS induces sustained rhythmic changes in the perception of a wider range of (acoustic) input, or **whether the effect is specific to speech sounds**.

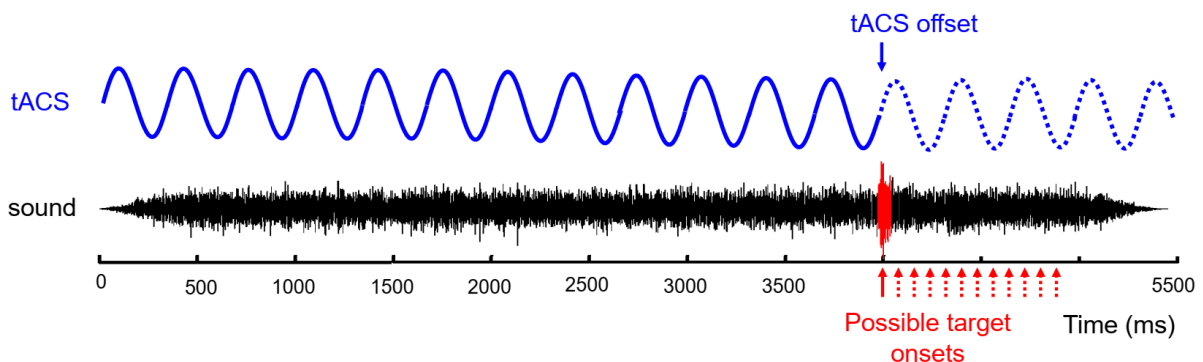


Figure 34. In each trial, tACS will be applied at 3 Hz and then turned off. A target word will be presented at various delays after tACS offset, covering three full 3 Hz-cycles (4 delays per cycle). EEG will be recorded in parallel, but data will only be analysed after tACS offset to avoid artefacts produced by the applied current.

For (1), phases in the EEG will be extracted using established methods such as *Fast Fourier Transformation* (FFT). Rhythmic sustained EEG responses after tACS offset will be quantified using *inter-trial coherence* (ITC), a standard measure quantifying phase consistency across

trials at a given frequency and already used in our previous work (van Bree et al., 2021). A simulated null (surrogate) distribution will be constructed by adding a random phase to each trial and repeating the analysis, thereby abolishing the hypothesized phase consistency across trials. A high ITC in the observed data (at 3 Hz), as compared to the null distribution, reflects a sustained rhythmic response in the EEG. By analysing EEG data after tACS offset, we also avoid artefacts in the EEG that are produced by the applied current (Noury et al., 2016). Auditory responses in the EEG are typically strongest over central sensors (Parker et al., 2019). The placement of tACS ring electrodes over auditory cortex (**Fig. 18**, top right), preventing EEG recordings for certain temporal sensors, will therefore not hinder the measurement of the hypothesized effect.

For (2), we will test for how long, relative to tACS offset, ITC remains higher than that seen in the surrogate distribution. Using the sine fit procedure, we will quantify how strongly word report accuracy depends on the timing of the target word relative to the preceding tACS, reflecting sustained rhythmic changes in speech perception as in the original study. By gradually shifting the time window in which the tACS-induced modulation of speech perception is analysed (from earlier to later delays), we will be able to test for how long the effect lasts. Finally, we will correlate the sustained rhythmic modulation of perception with that measured in the EEG (i.e. the ITC), linking perceptual and neural measures of sustained oscillations.

For (3), we will compare the degree of sustained rhythmic modulation of perception (i.e. the amplitude of the fitted sine functions) across the two conditions (speech vs tone target).

Together, we expect to demonstrate, and further characterize, **how tACS manipulates endogenous oscillatory activity**. Given the relative lack of such a demonstration in the literature, this study will be of critical importance to the field, as one of the few investigating neural mechanisms underlying tACS.

5.2 Auditory Oscillations: Useful or harmful?

I now come back to the idea that neural oscillations can be useful or harmful in the auditory system, depending on the amount of information lost at phases of low excitability of the oscillation (see Introduction and PhD work). During neural entrainment, oscillations are useful as the timing of events is known and oscillations can be aligned accordingly.

But what happens if stimulus timing is unknown? **According to the hypothesis presented here, auditory oscillations are only helpful if a loss of important information can be ruled out**, for example, if the low-excitability parts of the oscillation do not coincide with relevant input. Only in this case, I expect auditory oscillations to operate, as evidenced by a rhythmic modulation of perception and neural processes – **otherwise they are suppressed**, resulting in input processing that is more stable over time. As explained in the following, this is a hypothesis that will guide our research in the coming years.

5.2.1 Effects of relevance and predictability on phasic modulation of auditory perception

Supervision: Ka-Yan Lui (post-doc)

Unpredictable targets

In a first EEG experiment, participants will be presented with two pure tones at two different sound frequencies, several seconds long and presented simultaneously. In each trial, participants will be instructed to attend to one or both of the tones. They will be asked to detect a gap that can occur in any of the two tones, and at a random time. As an incentive to follow instructions, they will receive more “points” if they detect a gap in the attended tone than in the unattended one. According to our hypothesis, a potential loss of irrelevant (but not relevant) information can be tolerated by the auditory system. **We therefore expect the detection of the unattended (but not the attended) target to co-vary with the phase of oscillations**, measured in the EEG (**Fig. 35A**). Analytically, such a result will be reflected in significantly predictive value of oscillatory phase for perceptual outcomes in the unattended (blue in **Fig. 35A**) but not in the attended condition (orange), evaluated using regression models with circular predictors as described in **Section 4.2.2**.

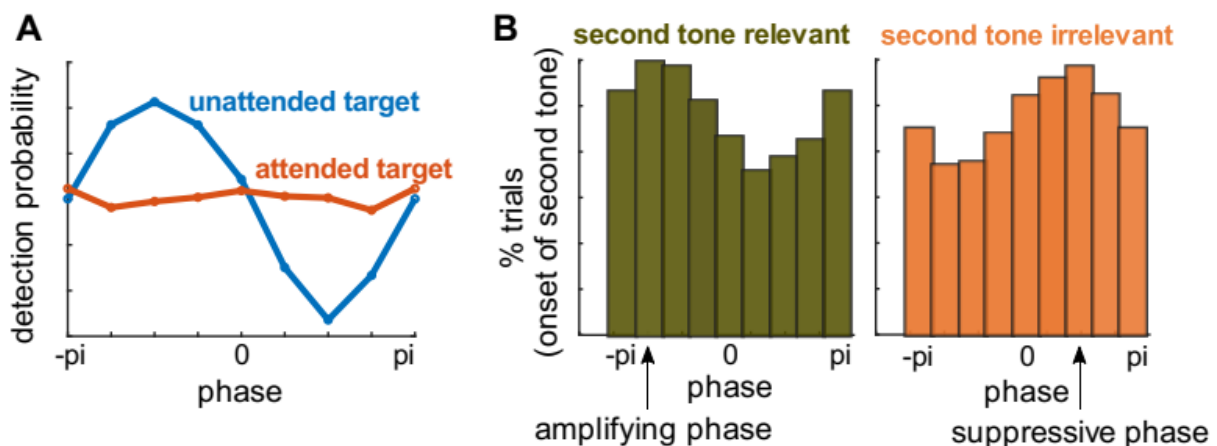


Figure 35. **A.** We expect the detection of unattended (blue), but not attended (red) targets to be modulated by EEG phase. **B.** We expect amplifying and suppressive oscillatory phases to be aligned at the onset of the second tone if the latter is relevant (brown) or irrelevant (orange) for the task, respectively.

Of note is that our hypothesized result is in stark contrast to previous findings in the visual domain, where oscillations phasically modulate perception only for attended stimuli (Busch and VanRullen, 2010). Such a result would again demonstrate that findings obtained for the visual system cannot necessarily be generalized to the auditory one.

Additionally, we will test for oscillatory effects when both tones are equally important and attention needs to be divided. Intuitively, it seems difficult to attend more than one sound frequency at a time, making a continuous mode of processing impossible to sustain. Studies

on the visual domain have suggested that oscillatory sampling of the environment switches rhythmically between multiple objects in the attentional focus (Landau and Fries, 2012; Fiebelkorn et al., 2013; Fiebelkorn and Kastner, 2019). Given the equivalence between visual space and auditory sound frequency (Kubovy, 1988; VanRullen et al., 2014), divided attention might evoke a rhythmic alternation of the attentional focus between sound frequencies (see also next section). In this case we would expect to find a phasic modulation of perception when both tones are attended. Alternatively, it could be that audition cannot afford such regular switches in which case we would not find any phasic modulation, similar to the condition with a single attended tone.

Predictable, non-rhythmic targets

The field of neural oscillations in the auditory domain is heavily focused on “neural entrainment” and corresponding stimulus-aligned responses to rhythmic stimuli (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Calderone et al., 2014; Zoefel and VanRullen, 2017; Obleser and Kayser, 2019). However, perhaps due to this current focus on rhythmic stimuli, it remains largely unknown how neural oscillations operate in audition when such external rhythms are absent. **Even in the absence of rhythmic input, as long as the timing of information is known, the adjustment of high-excitability (amplifying) and low-excitability (suppressive) phases to expected important and irrelevant events, respectively, would be a powerful way of reducing information loss during oscillatory mechanisms.** Whereas such an effect has been reported in the visual (Bonfond and Jensen, 2012) and olfactory (Arabkheradmand et al., 2020) domains, we currently lack evidence that auditory oscillations adjust their phase to expected, non-rhythmic (e.g., single) events. This will be tested in a second experiment, by presenting participants with stimuli which occur at a predictable time and need to be detected or suppressed to perform well in a given task, respectively. In each trial, participants will be asked to retain the pitch of a pure tone in memory while a second tone is presented at a predictable time. They will then have to decide whether the pitch of a third tone (“target”), presented subsequently, is higher or lower than the pitch of one of the other two tones (“reference”). Importantly, prior to each trial, participants will be informed about the identity of the reference tone. If the first tone is the reference, the second tone will **act as an irrelevant distractor, and we expect to see the suppressive phase (measured with EEG) aligned** to the expected onset of the second tone (reflected by a consistent phase across trials, i.e. by a high ITC relative to a null distribution; **Fig. 35B**, orange). If the second tone is the reference, **it becomes relevant, and we expect to see the amplifying phase aligned** to its onset (**Fig. 35B**, brown). Thus, in addition to a high ITC in both conditions (irrelevant and relevant second tone), we also expect opposite phases aligned to the tone, depending on whether it is relevant or not.

5.2.2 Auditory phase-reset

Collaborator: Céline Cappe

Supervision: Ka-Yan Lui (post-doc), Eva Boglietti (M1)

It has been shown previously that neural signals and perception in the visual system fluctuate rhythmically after a cue, and this finding has been interpreted as an external manipulation of endogenous oscillations that are “reset” by the cue (Landau and Fries, 2012; Fiebelkorn et al., 2013). Another study suggested similar mechanisms in the auditory system (Ho et al., 2017). However, such studies remain sparse; it is therefore unknown under which conditions such an auditory “phase-reset” occurs, and the consequences it has for perceptual and neural processes. On the one hand, auditory oscillations need to be flexible enough to respond and adapt (or “entrain”) to stimulus input – **a phase-reset might therefore be indispensable for an efficient adjustment to the rhythm of upcoming input** (Ghitza, 2011; Giraud and Poeppel, 2012; Thorne and Debener, 2014). On the other hand, **such a mechanism might be harmful when the timing of subsequent information is unknown**. This apparent discrepancy will be addressed in this experiment. Participants will be asked to detect target tones, presented at variable intervals after a cue. These targets will have two possible sound frequencies, and participants will be cued to attend to one or both of the target frequencies. Given its importance for neural adjustment mechanisms, we expect a cue-induced phase-reset in both conditions (attended and unattended sound frequencies). This phase-reset will lead to rhythmic fluctuations in perception after the cue (**Fig. 36**), which will be evaluated statistically using regression models (with the delay between cue and target as a predictor for perceptual outcome). However, due to the unpredictable timing of the targets (making oscillations potentially harmful to their detection), **we expect those “perceptual oscillations” to disappear more quickly for attended (i.e. relevant) than for unattended (i.e. irrelevant) target frequencies (Fig. 36, top).**

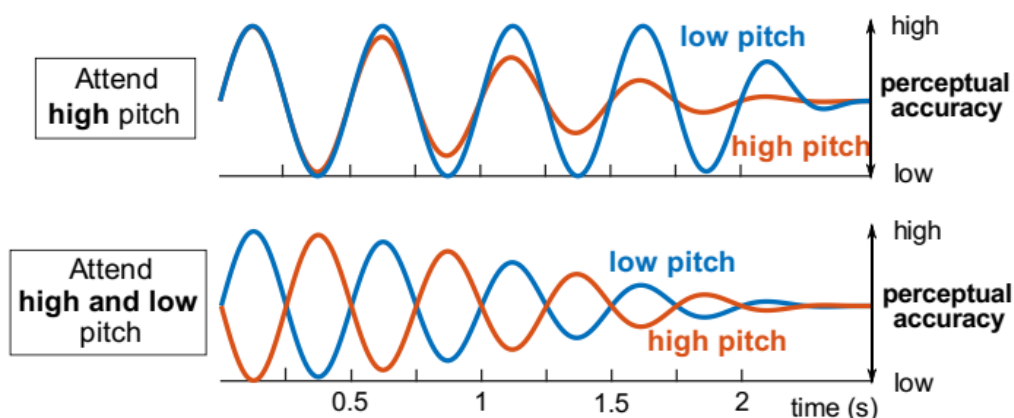


Figure 36. We expect that a cue “resets” neural oscillations and produces rhythmic fluctuations in perceptual accuracy that are longer if the stimulus is irrelevant (blue in top panel). Amplifying/suppressive phases alternate between two possible targets if both are relevant (bottom panel).

Using similar paradigms, previous research has shown that visual attention can alternate rhythmically between the two hemifields or different visual objects (Landau and Fries, 2012; Fiebelkorn et al., 2013; Fiebelkorn and Kastner, 2019). In the auditory system, such attentional rhythms were investigated by testing whether moments of highest perceptual accuracy alternate between the two ears, claiming equivalence to the spatial alternation observed in vision (Ho et al., 2017, 2019). Although the study revealed some evidence for oscillatory mechanisms underlying auditory attention, effects sizes were small, the experimental scenario relatively artificial (targets were presented only in one ear), and results difficult to interpret (as they varied with outcome measure and ear). This might have been due to the fact that visual space has its equivalent in frequency (and not space) in audition (Kubovy, 1988; VanRullen et al., 2014). This experiment addresses this fact by manipulating sound frequency rather than location of the targets. More precisely, **we will test whether audition switches rhythmically between multiple sound frequencies** by asking participants to attend to both sound frequencies. Such an alternation would be visible as a rhythmic pattern in target detection for both possible target frequencies, but also as a phase opposition between the two (**Fig. 36**, bottom).

Simultaneous EEG recordings will be used to measure the neural underpinnings of the behavioural effects – in particular, we will test whether a phasic modulation of perception goes along with an oscillatory response in the EEG.

Finally, we will run modified versions of this paradigm to address additional research questions:

- We will replace a single cue with a sequence of tones. This sequence can be rhythmic or irregular. If it is rhythmic, we expect to find a rhythmic entrainment echo in the detection of the subsequent target as described in **Section 4.3.3**. In the rhythmic condition, these echoes should follow the rate of the tone sequence that is used as a cue. By comparing results with those obtained after a single cue (i.e. in the original experiment, without rhythmic component), we can test whether and to which degree the **rhythm in the sequence can “overwrite” the natural frequency** of audition. For example, a single cue might produce perceptual oscillations at 6 Hz (**Fig. 36**) whereas an 8-Hz tone sequence would be expected to produce an 8-Hz echo. In addition, we can address the dichotomy (*entrainment vs adaptation*) described in **Section 4.3.3**: If a rhythmic stimulus leads to neural adaptation at specific moments in time, but an irregular one (with the same average rate) does not, then best moments for target detection in the subsequent echo should differ between the two.
- Auditory stimuli in everyday life are often coupled with those from other modalities (e.g., speech and music also involve the visual and motor systems). Together with Dr Céline Cappe (Toulouse), we will **address such cross-modal interactions**: We will test whether a visual cue can produce oscillations in auditory perception, and whether voluntary actions are followed by similar rhythmic changes in perceptual accuracy, such as previously reported for the visual system (Benedetto et al., 2016).

5.3 Manipulating oscillations to control auditory perception and attention

If neural oscillations acted as an attentional filter, any alteration in oscillatory activity – irrespective of whether these are induced pathologically or experimentally – should have consequences for various perceptual and attentional processes. In line with this notion, a malfunctioning of neural oscillations has been associated with pathological conditions such as attention-related deficits (Calderone et al., 2014) or schizophrenia (Lakatos et al., 2013b). Changes in neural entrainment have been demonstrated in hearing-impaired individuals (Petersen et al., 2016; Fuglsang et al., 2020), suggesting neural oscillations underlying basic auditory abilities required in everyday life. Logically, developing tools to manipulate neural oscillations in a controlled fashion (Thut et al., 2012) is of utmost importance. Successful manipulation of neural oscillations leading to changes in perception and attention would not only establish a causal link between oscillations and such functions; it would also make oscillations an important target for clinical interventions and demonstrate how such interventions could work in practice. As described extensively in **Section 4.2.1**, we have repeatedly used tACS to modulate auditory processing via a manipulation of neural oscillations. I will continue this line of research in the future and complement it with novel experimental and methodological protocols. All of these are designed in the aim of manipulating oscillations to improve hearing and attention.

5.3.1 Modulating speech perception with tACS

Adapting tACS protocols to individual participants

Collaborators: Matthew H Davis, Mathieu Marx

Supervision: Jules Erkens (post-doc), **Ram Kumar Pari** (research engineer)

Although statistically reliable and replicable (Zoefel et al., 2020), tACS effect sizes are often small and fall short of those needed for clinical utility. There is, however, important evidence that participants differ in their susceptibility to a given stimulation protocol (Laakso et al., 2015). Individualizing brain stimulation protocols has therefore been proposed as a crucial step to advance this field of research (Romei et al., 2016; Zoefel and Davis, 2017). Indeed, anatomical factors partly explain how strongly 10-Hz oscillations in visual regions can be modulated by tACS in individual participants (Kasten et al., 2019). However, it remains unknown if similar factors determine success in tACS-induced modulation of speech processing. **Revealing parameters that are critical to adapt tACS to individual brains would make the stimulation more efficient, increasing its potential to play an important role in both research and everyday life applications.**

We will apply the tACS-fMRI protocol that was described in **Section 4.2.1** to tackle this aim (following the paradigm described in Zoefel et al., 2018a). In short, participants will listen to

rhythmic speech while tACS is applied over auditory brain regions during simultaneous fMRI (using bilateral ring configurations shown in **Fig. 14A**). In each trial, the speech stimulus will be presented at one out of six possible tACS phases, and participants will be asked to repeat the words they have heard. In parallel, participants' BOLD responses to the presented speech will be measured. This will allow us to determine how strongly the BOLD response (for each fMRI voxel), as well as word report accuracy, is modulated by the phase relation between tACS and speech.

We will use available state-of-the-art computer simulations (Thielscher et al., 2015; Puonti et al., 2020; Saturnino et al., 2020; Salvador et al., 2021) to estimate the electric field induced by tACS in each fMRI voxel, based on individual, structural MRI data. We will then **use the individual electric fields to "normalise" the phasic effects observed**. For instance, we might find the strongest phasic modulation in a voxel that is also exposed to the strongest electric field, but another voxel might show a phasic modulation that is nearly as strong despite being exposed to a much weaker electric field. "Correcting" for differences in electric field between voxels will **reveal brain regions that respond – given the *same* electric field – most readily to tACS**. Regions that respond most readily across participants will be **considered an optimal target** for tACS-induced modulation of speech processing.

We will then use computer simulations again, but this time to **predict optimal electrode positions to target these regions in individual participants**. For example, we might find that tACS is particularly successful in modulating speech processing when applied to target Superior Temporal Gyrus. We would confirm with computer simulations where electrodes need to be attached to optimally target this region in individual participants, given their structural MRI data.

In a second experimental session, participants will again be asked to complete the word report task. tACS will be applied (without simultaneous fMRI), using two different protocols: (1) Our standard protocol, identical for each participant (electrodes attached over auditory regions) and (2) the optimised, individualized protocol. We expect that **the optimised tACS protocol produces a stronger modulation of speech perception than a conventional protocol**. This result would be important as it shows that we can tailor protocols to individual brains and increase tACS effect sizes by doing so. This experiment would therefore have implications for both the scientific field and future practical applications.

Effects of tACS in complex listening scenarios

Collaborator: Mathieu Marx

Supervision: Jules Erkens (post-doc)

It has been shown repeatedly that tACS can be used to change speech perception. However, these findings are often based on relatively artificial scenarios in which participants listen to simple, isolated words (Wilsch et al., 2018; Zoefel et al., 2018a, 2020; van Bree et al., 2021). In one study (Riecke et al., 2018) with a more sophisticated experimental setting (listeners attended to one of two simultaneously presented speech streams), the role of the distracting stimulus remained untested. This is important, as **tACS might enhance perception by amplifying attended speech, or by suppressing distracting speech** (or both). Understanding whether tACS should be timed relative to the attended or distracting input is of crucial relevance, not only to understand underlying neural effects, but also from an applied perspective and for future clinical applications.

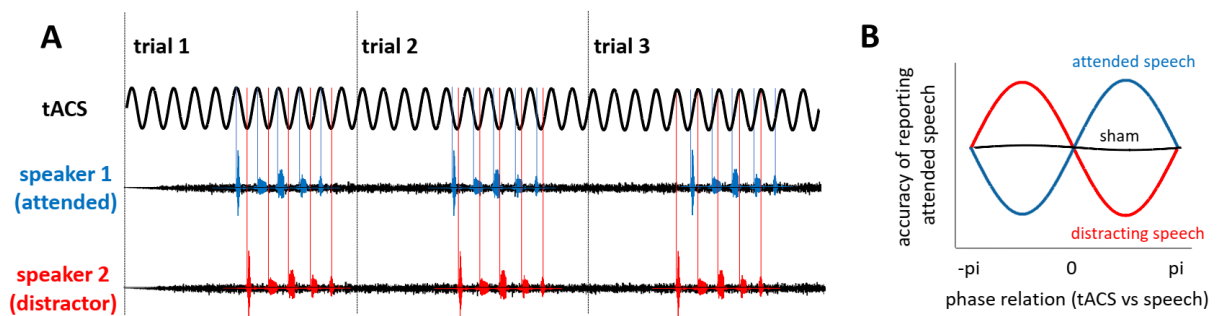


Figure 37. A. Two rhythmic speech streams will be presented simultaneously, both with a well-defined phase relation to tACS in each trial, but with a random phase relation to each other. **B.** We will test whether tACS can enhance speech perception by boosting attended speech (blue), relative to a sham condition (black), and/or by suppressing distracting speech (red). In the example shown here, enhancing attended speech (at phase relation $\pi/2$) and suppressing distracting speech (at phase relation $-\pi/2$) is equally effective.

In this experiment, participants will be asked to listen to one of two rhythmic speech streams, presented simultaneously, while tACS is applied (**Fig. 37A**). After each trial, they will be asked to repeat the words they have identified from the attended speech. The timing of the two speech streams relative to each other (i.e. their phase relation) will be determined randomly in each trial. Both (attended and distracting) speech streams will be presented so that they fall at one out of six possible tACS phases in each trial. It will therefore be possible to determine how accuracy in reporting words from the attended speech depends on the phase relation between tACS and attended speech, but also on that between tACS and distracting speech (**Fig. 37B**).

In a given trial, word recognition accuracy might be a result from the phase relation between tACS and both attended and distracting speech. However, on average, due to the random phase relation between speech streams, it will be possible to disentangle the two: Averaged across trials, word report accuracy for a given phase relation between tACS and attended (or distracting) speech will not be affected by the phase relation between tACS and distracting (or attended) speech, as the latter is random, and potential phase effects will cancel out. This is the critical difference to previous work using tACS in relatively complex listening scenarios (Riecke et al., 2018), where the phase relation between the two speech streams was fixed, and effects of tACS on attended vs distracting speech could not be disentangled. Our design

also controls for the effect of temporal delay between attended and distracting speech (e.g., perception might be more accurate if the two do not overlap), as this delay is varied independently of the phase relation between tACS and speech, and will therefore not affect the trial average shown in **Fig. 37B**.

We will quantify the hypothesized tACS-induced modulation of word report accuracy as explained in detail in **Sections 4.2.1, 4.2.2, and 4.3.2**. The degree of phasic modulation will then be compared: (1) across conditions (attended vs distracting speech), to **test whether tACS is more effective in modulating attended or distracting speech**; (2) against a null distribution (obtained with a surrogate procedure), separately for each of the two conditions, to **test whether tACS is effective in modulating speech processing** in the respective condition; (3) against a sham condition to **test whether tACS can enhance or disrupt speech perception**, as compared to a scenario without stimulation. We will also extract the “best” phase relations (leading to most accurate perception) between tACS and speech in the two conditions (attended vs distracting speech). We expect to find opposite best phase relations between the two conditions (**Fig. 37B**): The phase associated with an amplification of input would lead to highest accuracy if centred on attended speech, but to lowest accuracy if centred on distracting speech.

Together, this experiment will **reveal whether tACS works more efficiently by enhancing relevant information, or by suppressing distracting ones, both of which is predicted to lead to improved speech perception** in a multi-speaker scenario.

Predicting individual best tACS phases for speech perception

Collaborator: Mathieu Marx

Supervision: Florian Kasten (post-doc), **Zoé Darrasse** (M1)

A common finding across studies is that the “best” tACS phase (that leads to most accurate perception or strongest neural responses) **differs across individuals** (Riecke et al., 2018; Wilsch et al., 2018; Zoefel et al., 2018a, 2020; van Bree et al., 2021). From a practical

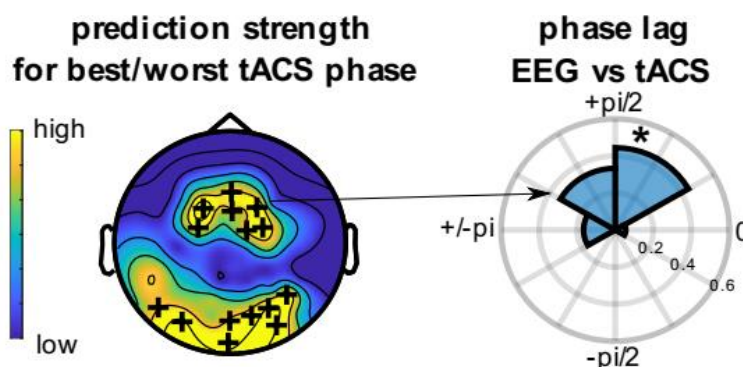


Figure 38. We found that the phase of EEG responses relative to rhythmic intelligible speech predicts individual best tACS phases for speech perception. Crosses show EEG channels with significant predictive value. The circular distribution shows the phase lag between EEG and tACS, yielding a $\sim 90^\circ$ lag for most participants. Modified from van Bree et al (2021).

perspective, this means that in typical tACS experiments, all phases have been sampled for each subject, leading to **relatively inefficient procedures and long experiments**. From a theoretical perspective, reasons for this inter-individual variability remained unclear. In the study described in **Sections 4.3.1 and 4.3.2** (van Bree et al., 2021), we found that **the individual best tACS phase can be predicted from the individual lag between EEG responses and rhythmic intelligible speech**, obtained in a separate dataset and collected on a different day (**Fig. 38**). This is an important step towards practical applications of tACS: If perception is optimal at different tACS phases for each individual, then extensive testing might be required to determine this optimal phase before further interventions. Based on our results, however, it should be possible to predict optimal tACS phase from individual EEG responses to rhythmic speech.

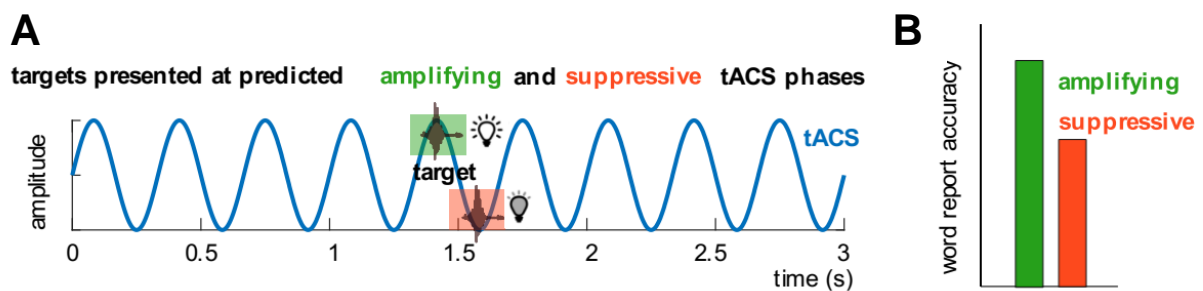


Figure 39. *A. We will control perception by presenting targets (grey) at tACS phases that were predicted (based on separate EEG recordings) to be “best” (green) and “worst” (red) for perception in individual participants, respectively. B. We expect more accurate perception at the predicted best than at the predicted worst tACS phases.*

We will build on this finding in a follow-up experiment. In a first session, we will record participants’ EEG responses to rhythmic speech and use them to predict, for each participant, their individual best tACS phase (as well as their worst tACS phase, 180° away). In a second session, we will then apply tACS and present target words at the predicted best (i.e. amplifying) and worst (i.e. suppressive) phases for speech perception, respectively (**Fig. 39A**). We expect that **reporting target words is easier for participants** when these were presented at the predicted best (as compared to predicted worst) tACS phase (**Fig. 39B**). This finding would have important implications for both research and practical applications, as it suggests that **we can use tACS to control perception in individual participants**.

5.3.2 Neural entrainment vs alpha oscillations: Two competing “modes” of auditory stimulus processing?

Collaborator: Mathieu Marx

Supervision: Florian Kasten (post-doc), **Quentin Busson** (M1)

A previous study in non-human primates (Lakatos et al., 2016) reported the important observation that not all oscillations in the auditory system are directly linked to the processing of incoming information. While subjects attended a rhythmic stimulus, the authors observed oscillations entrained to the stimulus. At certain times, however, neural entrainment was reduced and *alpha oscillations* (~10 Hz) dominated processing in the auditory system. Importantly, while alpha oscillations rhythmically modulated neuronal firing, they also reduced neural and behavioural responses to stimulus input. This finding was interpreted as a reflection of “internal attention”, characterized by strong alpha oscillations and reduced sensitivity to external information. In humans, strong alpha oscillations are similarly associated with an inhibition of incoming information (Klimesch et al., 2007; Romei et al., 2008; Jensen and Mazaheri, 2010), and specific auditory ones do exist (Weisz et al., 2011).

We will develop a **novel approach that will allow us to automatically detect moments of presumed internal attention in human participants and “rescue” task performance by triggering a “switch” to external attention (Fig. 40)**. We will test the same participants that volunteered for the experiment described in the preceding section (5.3.1); we will therefore already have identified their “best” tACS phase for speech perception. The same recordings will be used to identify EEG sensors reflecting alpha oscillations in the auditory system. Identification of auditory alpha oscillations is feasible with EEG (Obleser and Weisz, 2012; Lim et al., 2015; Wöstmann et al., 2017).

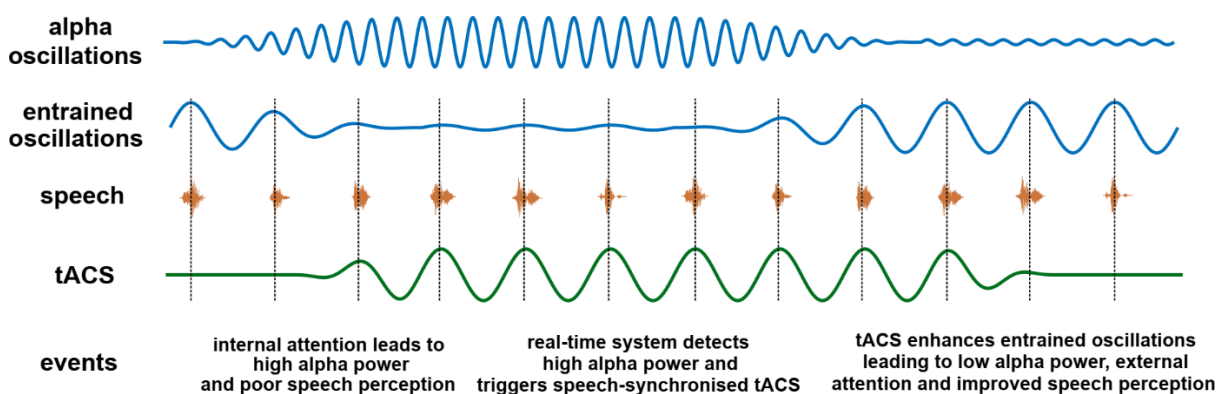


Figure 40. While participants listen to rhythmic speech (orange), our real-time system will detect periods of strong alpha oscillations (blue, first row) which have been shown to reflect states of internal attention with corresponding decline in speech perception and entrained activity (blue, second row). When such an “alpha-mode” is detected, the system will trigger tACS (green), which we expect to re-synchronize neural oscillations with the speech sounds, and lead to a suppression of alpha oscillations and improved speech perception.

In this experiment, while participants listen to rhythmic speech sounds and complete simple tasks measuring speech perception, we will use EEG to identify – in “real-time” – times with strong alpha oscillations in auditory regions (at the EEG sensors selected for their alpha activity). When processing in such an “alpha mode” is detected, we will apply tACS to boost the processing of the rhythmic speech – i.e. at the phase relation, relative to the presented speech sounds, that is predicted to lead to most accurate speech perception. **We expect that the electrical stimulation will lead to a switch from internal to external attention and, consequently, to reduced alpha oscillations and improved task performance** (as compared to a control condition during which the applied tACS is unrelated to the presented speech stimuli). Such a finding would suggest new exciting technological applications for combined EEG-tACS, in situations in which sustained attention is critical (e.g., for pilots).

5.3.3 Modulating speech perception with neurofeedback

In addition to a passive (from a participant’s point of view) modulation of neural oscillations via tACS, **humans can also learn to actively control their own oscillations via neurofeedback** (Herrmann et al., 2016). Neurofeedback works by extracting a certain measure of interest (e.g., oscillations) from a neurophysiological (e.g., EEG) signal and by providing participants with an arbitrary means of feedback that reflects the status of this measure (**Fig. 41A**). For example, in previous work we have used a square whose colour changes with the amplitude of oscillations (Zoefel et al., 2011). Participants were only told to change the colour, but were unaware that it is linked to their oscillations. In this way, they implicitly learnt to regulate their oscillations which led to improvements in a cognitive task. Nowadays, neurofeedback is routinely used in clinical practice (Marzbani et al., 2016). We will use novel applications of neurofeedback, with the common aim of modulating neural oscillations and speech perception.

Neurofeedback to regulate neural entrainment

Collaborator: Stefanie Enriquez-Geppert

Supervision: Jules Erkens (post-doc), **Ram Kumar Pari** (Research Engineer), **Damien Ricci** (M1)

In a first experiment, we will **examine whether participants can learn to change their own neural entrainment, and consequently, speech perception**. We will also test whether an external manipulation of entrainment via tACS can help participants in unsuccessful phases of the neurofeedback training.

The experimental design is illustrated in **Fig. 41**. Participants will listen to rhythmic speech embedded in noise, and their EEG will be recorded simultaneously. For each participant, their “baseline” entrainment will be measured, i.e. the degree to which EEG oscillations are aligned

to speech without neurofeedback. Subsequently, participants will listen to the same stimuli again, but this time they will receive feedback about their own entrainment. Entrainment will be measured over central EEG sensors as auditory responses are typically largest there (Parker et al., 2019). Feedback will be given visually, as the colour of a square on the screen (**Fig. 41A**), inspired by previous successful studies (Zoefel et al., 2011; Herrmann et al., 2016). The square will be red (blue) if the instantaneous entrainment is above (below) that measured during baseline, with several colour increments between brightest red and brightest blue. Participants will be asked to make the square as red as possible, but no specific instructions will be given of how this can be achieved. To make the visual stimulation during baseline comparable with the neurofeedback condition, during the former a square will be presented as well, but with colour changes unrelated to the participants' entrainment. In this case, participants will be asked to attend to both speech and square.

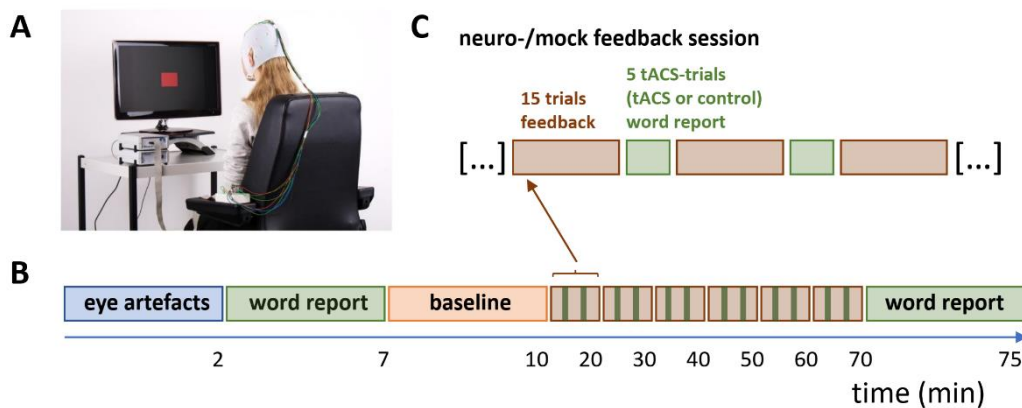


Figure 41. **A.** During neurofeedback, participants receive feedback about their own oscillatory activity, for example reflected by the colour of a square, which they are asked to change. From Herrmann et al. (2016). **B.** Experimental design. First, a template for eye artefacts will be estimated in individual participants' EEG. This will be used to make mock feedback more realistic (by changing the colour of the square after eye blinks). Second, participants will be asked to repeat words they have heard while their EEG is recorded. Third, for individual participants, baseline entrainment to rhythmic speech will be measured. Fourth, in six blocks of neurofeedback or mock feedback (in two separate groups), participants will receive feedback about their own entrainment, relative to their baseline (or feedback from another block as mock feedback). Neuro-/mock feedback will be interrupted by trials in which tACS is applied and participants are asked to report words they have heard (detailed in **C**). Entrainment will either be boosted by presenting target speech at the same tACS phase in five consecutive trials, or, in a control condition, the phase relation between tACS and speech will be randomized, abolishing such effects on EEG-speech alignment. Fifth, participants will again complete the word report task while EEG is recorded.

A separate group will undergo the same procedure, but receive mock feedback (as the colour sequence from another experimental block) instead of true neurofeedback. With help from Dr Stefanie Enriquez-Geppert (Groningen), the mock feedback will be made more realistic by changing the colour of the square immediately after participants' eye blinks (Enriquez-Geppert et al., 2014b, 2014a). For this purpose, prior to the experiment, artefacts in the EEG, produced by eye blinks, will be estimated for each participant (Enriquez-Geppert et al., 2017). This

makes it difficult for participants to distinguish mock feedback from neurofeedback (Enriquez-Geppert et al., 2014b, 2014a, 2017).

In certain trials ("tACS-trials"; **Fig. 41C**), while speech is presented, the square will become grey. In this case, participants will be instructed to ignore the square. Only during these trials, tACS will be applied, and the speech sounds will be presented at one out of four possible tACS phases. tACS will always be applied for five consecutive trials, with speech presented at the same tACS phase, to increase the effect of tACS on the subsequent neurofeedback trials. In a control condition, speech will be presented at a random tACS phase in each of the five trials. Due to this random phase relationship, tACS cannot increase subsequent EEG-speech alignment and therefore acts as a control. To avoid tACS-related artefacts in the EEG, tACS will not be applied in other trials. After each tACS-trial (including controls), participants will be asked to repeat the words they have heard in that trial. In addition, their word report accuracy will be tested before and after the experiment. In this case, their EEG will be recorded but no tACS will be applied.

In each part of the experiment (word report, baseline, neurofeedback/mock sessions), the alignment between EEG and speech will be quantified using the *phase-locking value* (PLV; Lachaux et al., 1999). The PLV is a standard measure in the field of entrainment (Peelle et al., 2013), and quantifies the phase consistency between EEG and stimulus at a given frequency (typically that corresponding to the stimulus rate). It was, for instance, used to quantify entrainment results shown in **Fig. 7** (Zoefel and VanRullen, 2016). We will then test the following hypotheses: (1) We will see an **improvement in word report accuracy** (after vs before the experiment), which is **higher in the neurofeedback than in the mock feedback group**. (2) This improvement goes along with an **increase in PLV** (after vs before the experiment). (3) PLVs during neurofeedback differ when comparing those obtained after tACS with those obtained after control trials (within subjects), indicating that **tACS can be used to influence participants' success** during neurofeedback. More precisely, we expect higher (lower) PLVs for individual participants after tACS was applied with their "best" ("worst") phase relation to speech. These phase relations will be determined post-hoc. (4) Both PLV and word report accuracy will **increase gradually during the experiment**, with more pronounced increases in the neurofeedback as compared to the mock feedback group. This can be quantified as the gradient of a regression line fitted to PLV/accuracy as a function of experimental time, compared across groups (Zoefel et al., 2011). This experiment will therefore show us whether participants can regulate their own entrainment to speech, whether this leads to benefits for speech perception, and whether tACS can support them during the neurofeedback training.

Neurofeedback to regulate entrainment echoes

Collaborator: Burkard Maess

Supervision: Florian Kasten (post-doc)

Rhythmic entrainment echoes and their importance for both fundamental and clinical research have already been described in detail in **Sections 4.3 and 5.1**. In this experiment, **we will teach participants to control their own entrainment echoes**. Given the involvement of the cerebellum in these echoes (**Section 5.1.1**), we will use MEG rather than EEG, due to its higher spatial resolution and ability to localise cerebellar activity (Andersen et al., 2020). Data will be collected in collaboration with Dr Burkhard Maess (Leipzig).

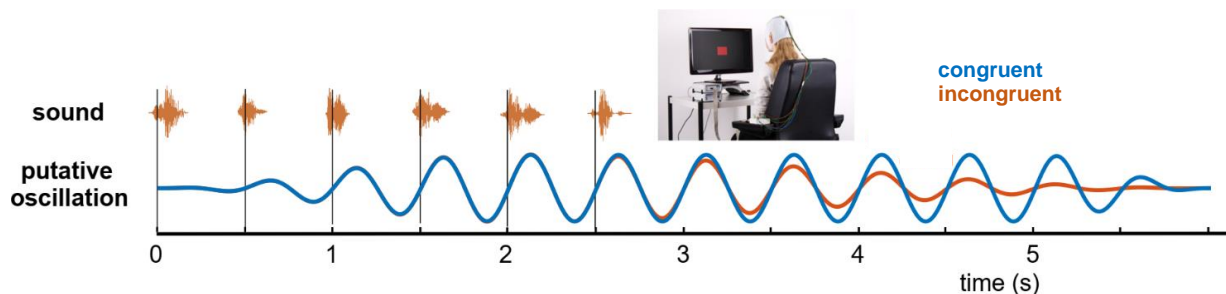


Figure 42. Participants will be presented with a square that changes in colour depending on the phase coherence (phase locking) between stimulus rhythm and MEG signal. We expect to find entrainment echoes that are stronger in the congruent neurofeedback condition (blue) than in an incongruent control condition (orange). In the latter, neurofeedback is given for a frequency that is unrelated to the stimulus rate.

We will use real-time MEG signal analyses to calculate participants' entrainment echo, produced by rhythmic speech sounds and extracted from sensors shown in **Fig. 42**. PLV will be used to quantify neural entrainment and its echoes as described in the preceding section. For entrainment echoes, PLV will be calculated as the phase coherence between MEG signal and the stimulus rhythm, *had it continued*. In each trial, participants will see a square that changes colour according to this PLV. In other words, we will visualize how strongly their rhythmic neural activity at those sensors is phase-locked to the stimulus rhythm (and to the imaginary continuation of this rhythm after stimulus offset). When participants learn to control the square's colour, **we expect to see entrainment echoes that become stronger during the course of the neurofeedback training**, and that these training-induced changes are larger than those in a control condition (tested in the same participants but on a different day). For the latter, participants receive feedback on a frequency in the MEG signal that is unrelated to the stimulus rate. We label this condition incongruent, as stimulus rate and neurofeedback frequency mismatch. This allows us to **isolate effects of neurofeedback on the entrainment echo from neurofeedback *per se***, as participants are only trained to enhance the entrainment echo in the congruent condition, but exposed to neurofeedback in both conditions.

Our neurofeedback protocol has implications for several lines of research. First, it allows us to test the hypothesis that the auditory system only relies on neural oscillations when they are “useful” whereas oscillations are suppressed in other situations (see **Sections 4.1.1 and 5.2**). In the incongruent condition, rhythmic entrainment echoes are not particularly “useful” and should therefore be relatively short. In the congruent condition, echoes are necessary to complete the task and therefore become “useful”. **Our expected finding – stronger or longer echoes during neurofeedback – would therefore support the hypothesis that neural oscillations are only sustained in the auditory system when they are needed.**

Second, given the link between entrainment echoes and the cerebellum (**Section 5.1.1**), our findings will have implications for our understanding of temporal predictions and associated pathologies. Patients with cerebellar lesions do not perform well when they have to tap along with a beat that changes in rate (Stockert et al., 2021). Other pathologies, such as dyslexia or schizophrenia have been linked with a deficit in the ability to predict stimulus timing (Lakatos et al., 2013b; Casini et al., 2018). Together with the fact that neural oscillations are often seen as a mechanism of temporal prediction (Lakatos et al., 2008, 2019), **an active training of entrainment and its echo might lead to improvements in temporal stimulus processing** – a hypothesis that needs to be tested in follow-up studies – and open up new opportunities to treat such pathological conditions.

Neurofeedback at the “cocktail party”

Supervision: Ram Kumar Pari (research engineer)

When humans age, it is extremely common that their hearing declines: This is one of the most prominent health issues of an aging world. This decline is not necessarily caused by actual hearing loss but often by a **reduced ability to suppress distracting speakers or noise** (Füllgrabe et al., 2015). Indeed, impaired speech perception in multi-speaker scenarios seems to be an early warning sign for dementia (Livingston et al., 2020). At the same time, **suppression of background noise is a task that current hearing aids often fail at**. It is therefore important to find alternative ways of improving distractor suppression in ageing listeners. In this experiment, we aim to find such a way by **using neurofeedback to control neural oscillations involved in the suppression of distracting information**. Oscillations can suppress information that occurs at their low-excitability phase (see **Introduction**), and recent research suggested that this role is particularly important in multi-speaker scenarios (Zion Golumbic et al., 2013). Crucially, neural oscillations and their synchronisation to sounds are also altered in ageing listeners (Henry et al., 2017), which makes them a promising target for such manipulative interventions.

In this experiment, we will present participants with two simultaneous speech streams and ask them to attend to one of them – this is a classical **task that requires distractor suppression** and becomes more difficult with age. Using EEG, we will estimate participants’ oscillations synchronised to speech in “real-time”, an approach already described above. We

will present participants with a square whose colour reflects this synchronisation and ask them to keep it at a certain colour without further instructions.

In two different conditions, this colour will either reflect the synchronisation with the attended speech stream or with the distracting one, respectively. Thus, depending on the condition, **participants will either be trained to increase brain-speech synchronisation with the attended stream or to decrease synchronisation with the distractor** (the colour of the square will be adapted accordingly). At random times during training, participants will be asked to repeat words from the attended stream which allows us to measure their speech comprehension and its evolution during neurofeedback. Ageing subjects (without hearing impairment) will be contrasted with younger ones with matched hearing in quiet and other cognitive abilities.

Our hypotheses are the following:

- As shown previously (Henry et al., 2017), **ageing listeners show overall differences in brain-speech synchronisation** as compared to younger ones.
- Both groups (young and old) can learn to regulate their own brain-speech synchronisation, leading to **improvements in a speech perception task**. These improvements are larger than those in a control group that receives mock (arbitrary) neurofeedback (to control for placebo effects).
- As it is mostly the suppression of distractors that ageing listeners struggle with, (1) their speech perception improves more from reducing synchronisation with distracting speech than from increasing synchronisation to attended speech, and (2) this difference is larger than for younger subjects (i.e. we expect an **interaction between age and condition**).

Together, this project will establish a novel approach to manipulate neural substrates of successful speech perception and reveal new ways of improving distractor suppression in ageing participants.

5.3.4 Modulating speech perception with AM-kTMP

Collaborator: Ludovica Labruna

Supervision: Ram Kumar Pari (research engineer)

Given results presented in this thesis, there is little doubt that tACS can produce robust, replicable effects when applied to modulate neural oscillations and speech processing. However, its **fundamental issue is that most of the current applied is shunted by skin and skull** (Lafon et al., 2017). Due to tolerance limits imposed by peripheral nerve stimulation, conventional tACS intensities (1-2 mA) only produce weak electric fields of $\sim 0.2-0.5$ V/m (Vöröslakos et al., 2018). **Current intensities that are sufficiently strong to have more**

important effects on the neural level would cause skin sensations that exceed a tolerable level for participants (Vöröslakos et al., 2018).

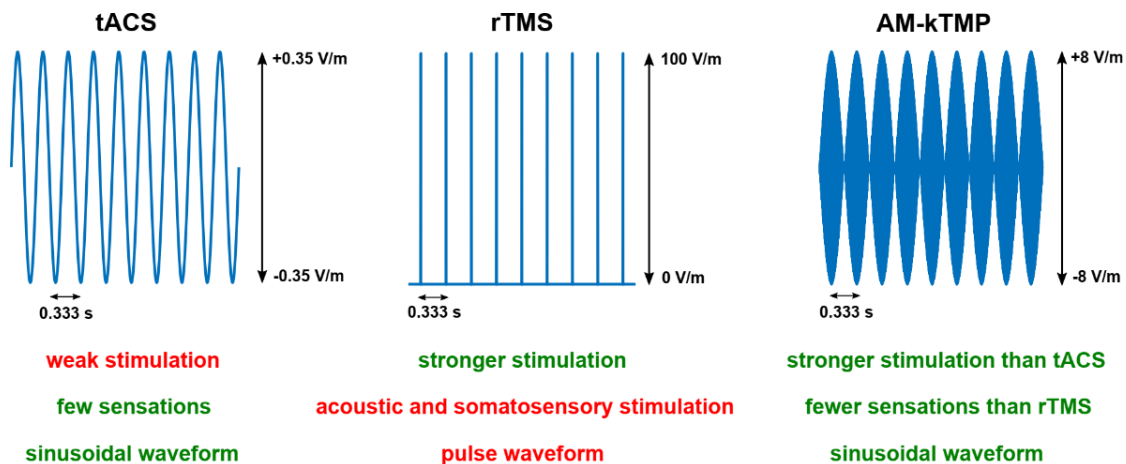


Figure 43. Comparison of transcranial stimulation methods to manipulate brain oscillations. Note that all of the stimulation waveforms shown are designed to entrain oscillations at 3 Hz. Note also that AM-kTMP is a promising “compromise” between tACS and rTMS that combines benefits (green) and avoids drawbacks (red) of these two more established methods.

Another established technique is repetitive transcranial magnetic stimulation (rTMS; **Fig. 43**, middle). This approach is based on rhythmically applied magnetic pulses to which brain oscillations align (Thut and Miniussi, 2009; Thut et al., 2011b). The electric field produced (as a consequence of the magnetic field applied) is significantly stronger than that produced by tACS and spatially more focused. However, several drawbacks exist that hinder successful application in research on neural entrainment and speech perception. (1) It produces rhythmic clicks; effects of rhythmic auditory and magnetic stimulation therefore cannot be distinguished. (2) It uses pulsed waveforms with limited potential for probing frequency-specific effects on neural oscillations (a pulse contains power in a wide range of frequencies). (3) It produces relatively strong, rhythmic sensations (leading to rhythmic somatosensory stimulation in addition to auditory and magnetic ones).

In recent years, a technique has been developed that combines benefits of tACS (sinusoidal waveform) and rTMS (stronger stimulation and higher focality) while avoiding the issues described (Labruna et al., 2021). This technique, called *kilohertz transcranial magnetic perturbation*, is a magnetic induction method that can produce cortical electric fields at kHz frequencies. In one variant, the high-frequency carrier waveform is rhythmically modulated in amplitude (AM-kTMP; **Fig. 43**, right). Due to their non-linear properties, **neural populations follow the amplitude modulation (AM) of higher-frequency carrier signals**, leading to brain oscillations entrained to the rhythmic AM (Grossman et al., 2017; Negahbani et al., 2018; Haslacher et al., 2021). For tACS, this approach has been used, but seems too weak to have important effects (Negahbani et al., 2018). **Relying on the stronger electric fields and higher focality of magnetic stimulation, AM-kTMP has therefore the potential to manipulate neural oscillations and their synchronisation to speech to a degree that is clinically relevant, while avoiding previous drawbacks of rTMS described above.**

In a previous study, Labruna and colleagues (2021) showed that AM-kTMP over participants' motor cortex leads to an increase in neural excitability (measured as the TMS-evoked motor potential) that is significantly stronger than after sham stimulation. This shows that AM-kTMP indeed modulates neural activity. In addition, annoyance, pain, and muscle twitches during kTMP were low (all between 0 and 0.5 on a subjective report scale between 0 and 10 where 0 = "not at all" and 10 = "extremely"). Sensations were not rated as significantly stronger than during sham stimulation, and were significantly weaker than during conventional TMS. kTMP produces a faint tone at the carrier frequency which has been masked successfully in this study.

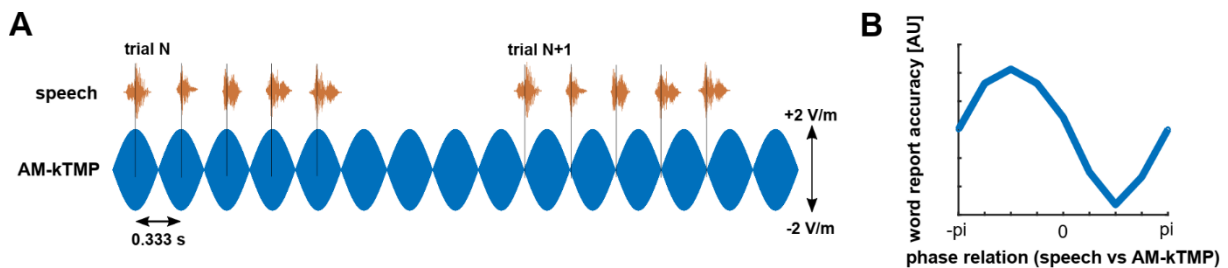


Figure 44. *A. Experimental paradigm. In all experiments, we will present rhythmic speech sounds (orange) at different phases of the amplitude-modulated kTMP signal (blue). Participants will be asked to repeat the words. B. We expect that accuracy in reporting words depends on the phase relation between speech and AM-kTMP.*

In collaboration with Dr Ludovica Labruna (Berkeley), we will apply AM-kTMP in three experiments, all of them designed to manipulate neural oscillations and speech perception. All experiments rely on a paradigm that has been described repeatedly in this thesis (**Sections 4.2.1 and 5.3.1**), and used to reveal an important role of neural oscillations for speech processing. The three experiments differ in that they test **auditory-only, auditory-visual, and auditory-motor (top-down) contributions** to speech perception. Their common aim is to **demonstrate that AM-kTMP can indeed manipulate neural entrainment and speech perception**, and does so to a degree that goes beyond that achieved with more established techniques. Such a result would have important consequences: First, it would reveal a method for the manipulation of neural oscillations that is more efficient than current ones that often struggle with small effect sizes (Riecke and Zoefel, 2018). Second, it would suggest new treatments for conditions associated with altered neural oscillations and their synchronisation to input (Calderone et al., 2014; Petersen et al., 2016; Henry et al., 2017). If AM-kTMP can improve neural entrainment, it is possible that these conditions will also improve following (or during) the stimulation. Exploring such a possibility is a realistic endeavour for the years following the project.

We will acquire data from healthy participants who will complete all of the three experiments in randomised order. Participants will be asked to listen to and report sequences of one-syllable words that are spoken rhythmically at 3 Hz and degraded in intelligibility (**Fig. 44A**, orange). As in other experiments described above, participants' perceptual accuracy will be quantified using Levenshtein distance. During the task, AM-kTMP will be applied at the corresponding frequency (AM at 3 Hz; **Fig. 44A**, blue), and the phase of the 3-Hz AM will be varied relative to that of the speech stimulus. Based on previous work and theoretical considerations

(Fröhlich, 2015; Negahbani et al., 2018; Haslacher et al., 2021; Labruna et al., 2021), we expect neural oscillations to follow the AM. By changing the phase relation between 3-Hz AM and speech rhythm, we will thus alter neural entrainment and, consequently, speech perception. **We therefore expect word report accuracy to depend on the phase relation between speech rhythm and AM-kTMP (Fig. 44B).**

The three experiments differ in the brain region targeted with AM-kTMP while participants complete the word report task (auditory: left Superior Temporal Gyrus [STG], visual: left middle occipital gyrus [MOG], motor: left inferior frontal gyrus [IFG]). As participants complete all experiments, we will be able (1) to determine whether AM-kTMP modulates speech perception in each experiment separately, but also (2) to compare results between them, thereby contrasting rhythmic auditory, visual and motor/top-down contributions to speech perception. In all experiments, we will also compare results with unmodulated kTMP, which produces similar sensations (e.g., a faint tone) but does not entrain neural oscillations. This comparison will enable us to test whether AM-kTMP improves speech perception (relative to unmodulated kTMP) at some phase relations, disrupts it, or both. We will produce an estimated electric field of 8 V/m at the cortical surface. This exceeds a field strength that an influential study (Vöröslakos et al., 2018) proposed to be necessary to mimic physiological conditions in oscillatory networks and that tACS cannot achieve.

In our previous tACS work, we targeted auditory regions including (left) STG, known to contribute to speech processing (Davis and Johnsrude, 2003). Although we did find reliable phasic modulation of speech processing induced by tACS (**Section 4.2.1**), effect sizes were low, in the range of 2-3 % modulation of accuracy (or a 0.3 dB change in hearing, summarized in Riecke and Zoefel, 2018), and only in one case up to 8 % (Zoefel et al., 2020). In Experiment 1, we will target the same auditory regions but replace tACS with AM-kTMP. As explained above, **we expect to reproduce our previous findings, but with a larger effect size.**

In Experiment 2, we aim to manipulate the impact of neural activity in the visual system on speech perception. Recent studies have shown such an impact (Hauswald et al., 2018). For example, visual speech cues robustly precede corresponding auditory information with a relatively constant delay (Chandrasekaran et al., 2009), and the brain seems to use these visual cues to “prepare” oscillations in the auditory system for upcoming auditory events (Biau et al., 2021). We will use kTMP to mimic and systematically vary visual information. We will use the same paradigm (**Fig. 44**), i.e. rhythmic speech sounds will be presented acoustically, with no corresponding visual cues, and participants will repeat the words they heard. We will then apply AM-kTMP over a visual region highlighted by previous research (left MOG in Biau et al., 2021) so that – if timed correctly – **the stimulation can prepare auditory regions just as the (absent) visual cues would do.** We expect that speech perception is optimal when the timing (i.e. phase) of AM-kTMP precedes the speech stimulus with a certain delay, but non-optimal at other delays. Such a result would not only demonstrate a causal role of audio-visual neural interactions for speech perception, but also suggest novel approaches to improve speech perception in populations that cannot use visual cues during speech processing (e.g., blind individuals).

In Experiment 3, we aim to manipulate interactions between auditory and “higher-order” brain regions during speech processing. A recent, impactful study has shown that some participants spontaneously align their speech production with a rhythmic, acoustically presented stream of syllables (Assaneo et al., 2019). This motor-speech entrainment was correlated with performance in a speech perception task and reflected in anatomical differences to participants that did not synchronise spontaneously. Most notably, “high synchronisers” exhibited stronger white matter connections between auditory and inferior frontal regions. We will use AM-kTMP to manipulate participants’ motor-speech entrainment. We will use the same paradigm, but AM-kTMP will be applied over the left IFG. The IFG is involved in motor production and top-down modulation of auditory processing (Davis and Johnsrude, 2003; Poeppel and Assaneo, 2020), in particular in the left hemisphere, and had a distinctive role for motor-speech entrainment in the study by Assaneo and colleagues (2019). Again, we expect task accuracy to depend on the phase relation between AM-kTMP and speech rhythm. This phase relation reflects the timing of oscillations in IFG (entrained by AM-kTMP) relative to those in auditory regions (entrained by speech). The hypothesized phasic effect would therefore **support a causal role of auditory-frontal synchronisation for speech perception** and suggest that those participants that do not spontaneously exhibit such synchronisation might particularly benefit from AM-kTMP in follow-up studies.

6 CONCLUSION: PERSPECTIVES IN FUNDAMENTAL, APPLIED, AND CLINICAL RESEARCH

In this thesis I have given an overview of my scientific career, dedicated to the investigation of neural oscillations underlying auditory perception and attention. This career had its beginning during my Bachelor and Master projects, where I explored neurofeedback to manipulate oscillations (Zoefel et al., 2011), and tested whether auditory target detection depends on EEG phase (Zoefel and Heil, 2013), respectively. The “negative” results obtained in the latter study inspired much of my PhD work where I developed, together with Dr VanRullen, hypotheses to explain this observation. I also started to work on neural entrainment, which then became the main focus of my Post-Doc work. Together with Dr Davis, I developed brain stimulation methods to manipulate neural entrainment and speech processing. Today, I continue this work in the form of several lines of research. This is not possible alone; I am grateful to the group of students and researchers I work with, and my local and international collaborators. Each of these research lines will open up new questions that we will aim to answer in the future, or possibly lead to new directions in our research. **I hope that, within the next decade, our results can critically contribute to our understanding of how neural oscillations can help us cope with the vast amount of information we continuously receive. I also foresee that brain stimulation methods will gain in importance, as these can be used to regulate neural information flow in case it fails.** I hope that the corresponding line of research describe can also contribute to this development. The following paragraphs summarise the different paths our research might follow, and of the impact it might have from a fundamental, applied, or clinical perspective.

Current work on neural oscillations is heavily focused on vision, leading to the implicit assumption that similar mechanisms are at work in audition. This approach might not be valid, given that audition (but not vision) crucially relies on information unfolding over time. In parallel, **auditory oscillation research is dominated by “neural entrainment”** and therefore almost exclusively based on rhythmic stimuli. So far, this has prevented us from understanding how auditory oscillations operate when confronted with rapidly changing information whose **timing cannot be predicted**. By addressing these gaps in our knowledge (**Section 5.2**), we might find **a highly flexible attentional filter** that relies on continuous predictions of the timing of upcoming input and its relevance for current goals, sometimes operating rhythmically, sometimes in a more constant fashion.

In addition, **the tonotopical organization of the auditory system is long known** and thoroughly investigated (Brewer and Barton, 2016), but it has been **mostly ignored in the field of neural oscillations**. Previous research in non-human primates has demonstrated its importance (O’Connell et al., 2011, 2014; Lakatos et al., 2013a), however corresponding results in human participants are almost completely lacking. One reason for this is the current focus on the visual system and the common assumption that results are directly transferable to other modalities (VanRullen et al., 2014; Ho et al., 2017, 2019; Zoefel and VanRullen, 2017). By testing whether rhythmic fluctuations in attention and perception are organized according to sound frequency (**Sections 4.3.3 and 5.2**), we aim to demonstrate **tonotopy structuring neural oscillations in humans**.

Finally, our research can improve our **ability to identify endogenous neural oscillations**. Distinguishing “true” rhythmic activity in the brain from other, simultaneously active processes is not always straightforward (Thut et al., 2011a; Keitel et al., 2014) and involves multiple pitfalls during signal processing (Widmann et al., 2015). In our research, we use various paradigms (**Sections 4.3 and 5.1**) that avoid a contamination of the measured oscillatory responses with those which **only seem rhythmic** due to the regularity of the stimulus (e.g., regular sequences of stimulus-evoked responses).

Our research has implications that are not restricted to fundamental science. **If an attentional filter fails, important consequences are easy to imagine**: Irrelevant information might be prioritized, or important events ignored. Indeed, altered oscillatory activity is present in various pathological conditions, such as attention-related disorders (Calderone et al., 2014), schizophrenia (Lakatos et al., 2013b), or hearing impairments (Petersen et al., 2016). Deficits in the ability to predict the timing of upcoming events, potentially linked to oscillatory activity (Lakatos et al., 2013a) and its echo after entrainment (**Section 5.1**), have been observed as well, e.g. for dyslexia (Casini et al., 2018) or cerebellar lesions (Schwartz et al., 2016).

By developing ways to manipulate neural oscillations, we might also develop tools that can be used for future clinical research or applications. This includes, for instance, neurofeedback protocols that can help participants sustain their own oscillations or regulate neural entrainment (**Section 5.3.3**). Neurofeedback might help listeners who struggle at present to extract speech features necessary for optimal entrainment (such as hearing aid users; Petersen et al., 2016); once these have been extracted, they might be able to do so even in absence of neurofeedback. Other neurofeedback protocols are aimed to improve distractor suppression, a task ageing listener typically struggle with (**Section 5.3.3**).

During tACS, the electrical stimulation can control which information is enhanced, and which is suppressed (**Section 5.3.1**). Consequently, **the stimulation mimics the role of an attentional filter**, which might be impaired in the clinical conditions described above, and therefore improved by the stimulation in clinical follow-up studies. Our efforts to individualise the stimulation and adapt them to more realistic situations (**Section 5.3.1**) are designed under this perspective. Similarly, controlling whether information is perceived or not is one of the most important tasks for hearing aids. **Our research might contribute to advancing current technologies to improve hearing, with important implications for an aging world**. Finally, some of our protocols are designed to detect lapses in external attention, or “correct” them using electrical stimulation (**Section 5.3.2**). Such attentional lapses can be dangerous in certain situations (e.g., flying an airplane) – a possible outcome of our research is therefore an **integration of the obtained results in existing technologies that aim to prevent “human errors”** (Reason, 2000) – our proximity to potential industrial partners (e.g., Airbus in Toulouse) would support such a perspective.

I will end this thesis by insisting, once more, that research needs collaboration. None of the previous, current, and future work described would have been possible without all the people involved, and I thank all of them for joining with me on this journey.

7 REFERENCES

- Ali MM, Sellers KK, Fröhlich F (2013)** Transcranial alternating current stimulation modulates large-scale cortical network activity by network resonance. *J Neurosci* 33:11262–11275.
- Andersen LM, Jerbi K, Dalal SS (2020)** Can EEG and MEG detect signals from the human cerebellum? *NeuroImage* 215:116817.
- Arabkheradmand G, Zhou G, Noto T, Yang Q, Schuele SU, Parvizi J, Gottfried JA, Wu S, Rosenow JM, Koubeissi MZ, Lane G, Zelano C (2020)** Anticipation-induced delta phase reset improves human olfactory perception. *PLOS Biol* 18:e3000724.
- Assaneo MF, Ripollés P, Orpella J, Lin WM, de Diego-Balaguer R, Poeppel D (2019)** Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nat Neurosci* 22:627–632.
- Bäuerle P, Behrens W von der, Kössl M, Gaese BH (2011)** Stimulus-Specific Adaptation in the Gerbil Primary Auditory Thalamus Is the Result of a Fast Frequency-Specific Habituation and Is Regulated by the Corticofugal System. *J Neurosci* 31:9708–9722.
- Benedetto A, Spinelli D, Morrone MC (2016)** Rhythmic modulation of visual contrast discrimination triggered by action. *Proc R Soc B Biol Sci* 283:20160692.
- Biau E, Wang D, Park H, Jensen O, Hanslmayr S (2021)** Auditory detection is modulated by theta phase of silent lip movements. *Curr Res Neurobiol* 2:100014.
- Bonham BH, Lewis ER (1999)** Localization by interaural time difference (ITD): Effects of interaural frequency mismatch. *J Acoust Soc Am* 106:281–290.
- Bonnefond M, Jensen O (2012)** Alpha Oscillations Serve to Protect Working Memory Maintenance against Anticipated Distracters. *Curr Biol* 22:1969–1974.
- Brewer AA, Barton B (2016)** Maps of the Auditory Cortex. *Annual Rev Neurosci* 39, 385–407.
- Busch NA, Dubois J, VanRullen R (2009)** The Phase of Ongoing EEG Oscillations Predicts Visual Perception. *J Neurosci* 29:7869–7876.
- Busch NA, VanRullen R (2010)** Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc Natl Acad Sci* 107:16048–16053.
- Buzsáki G, Draguhn A (2004)** Neuronal oscillations in cortical networks. *Science* 304:1926–1929.
- Calderone DJ, Lakatos P, Butler PD, Castellanos FX (2014)** Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn Sci* 18:300–309.
- Cao L, Veniero D, Thut G, Gross J (2017)** Role of the Cerebellum in Adaptation to Delayed Action Effects. *Curr Biol* 27:2442–2451.e3.
- Capilla A, Pazo-Alvarez P, Darriba A, Campo P, Gross J (2011)** Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS One* 6:e14543.
- Casini L, Pech-Georgel C, Ziegler JC (2018)** It's about time: revisiting temporal processing deficits in dyslexia. *Dev Sci* 21.
- Chalas N, Daube C, Kluger DS, Abbasi O, Nitsch R, Gross J (2022)** Multivariate analysis of speech envelope tracking reveals coupling beyond auditory cortex. *NeuroImage* 258:119395.
- Chandrasekaran C, Trubanova A, Stillitano S, Caplier A, Ghazanfar AA (2009)** The Natural Statistics of Audiovisual Speech. *PLOS Comput Biol* 5:e1000436.

- Costa-Faidella J**, Baldeweg T, Grimm S, Escera C (2011) Interactions between “What” and “When” in the Auditory System: Temporal Predictability Enhances Repetition Suppression. *J Neurosci* 31:18590–18597.
- Davis MH**, Johnsrude IS (2003) Hierarchical processing in spoken language comprehension. *J Neurosci* 23:3423–3431.
- Deike S**, Heil P, Böckmann-Barthel M, Brechmann A (2012) The Build-up of Auditory Stream Segregation: A Different Perspective. *Front Psychol* 3:461.
- Ding N**, Patel AD, Chen L, Butler H, Luo C, Poeppel D (2017) Temporal modulations in speech and music. *Neurosci Biobehav Rev* 81:181–187.
- Doelling KB**, Arnal LH, Ghitza O, Poeppel D (2014) Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage* 85 Pt 2:761–768.
- Edwards E**, Chang EF (2013) Syllabic (~2-5 Hz) and fluctuation (~1-10 Hz) ranges in speech and auditory processing. *Hear Res* 305:113–134.
- Enriquez-Geppert S**, Huster RJ, Figge C, Herrmann CS (2014a) Self-regulation of frontal-midline theta facilitates memory updating and mental set shifting. *Front Behav Neurosci* 8:420.
- Enriquez-Geppert S**, Huster RJ, Herrmann CS (2017) EEG-Neurofeedback as a Tool to Modulate Cognition and Behavior: A Review Tutorial. *Front Hum Neurosci* 11.
- Enriquez-Geppert S**, Huster RJ, Scharfenort R, Mokom ZN, Zimmermann J, Herrmann CS (2014b) Modulation of frontal-midline theta by neurofeedback. *Biol Psychol* 95:59–69.
- Exner S** (1875) Experimentelle Untersuchung der einfachsten psychischen Prozesse. *Arch Für Gesamte Physiol Menschen Tiere* 11:403–432.
- Farahbod H**, Saberi K, Hickok G (2020) The rhythm of attention: Perceptual modulation via rhythmic entrainment is lowpass and attention mediated. *Atten Percept Psychophys* 82:3558–3570.
- Fiebelkorn IC**, Kastner S (2019) A Rhythmic Theory of Attention. *Trends Cogn Sci* 23:87–101.
- Fiebelkorn IC**, Saalman YB, Kastner S (2013) Rhythmic sampling within and between objects despite sustained attention at a cued location. *Curr Biol* 23:2553–2558.
- Fontolan L**, Morillon B, Liegeois-Chauvel C, Giraud A-L (2014) The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat Commun* 5:4694.
- Fowler CA**, Whalen DH, Cooper AM (1988) Perceived timing is produced timing: A reply to Howell. *Percept Psychophys* 43:94–98.
- Fröhlich F** (2015) Experiments and models of cortical oscillations as a target for noninvasive brain stimulation. *Prog Brain Res* 222:41–73.
- Fuglsang SA**, Märcher-Rørsted J, Dau T, Hjortkjær J (2020) Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention. *J Neurosci* 40:2562–2572.
- Füllgrabe C**, Moore BCJ, Stone MA (2015) Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci* 6:347.
- Ghitza O** (2011) Linking Speech Perception and Neurophysiology: Speech Decoding Guided by Cascaded Oscillators Locked to the Input Rhythm. *Front Psychol* 2:130.
- Giraud A-L**, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci* 15:511–517.

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S (2013) Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol* 11:e1001752.

Grossman N, Bono D, Dedic N, Kodandaramaiah SB, Rudenko A, Suk H-J, Cassara AM, Neufeld E, Kuster N, Tsai L-H, Pascual-Leone A, Boyden ES (2017) Noninvasive Deep Brain Stimulation via Temporally Interfering Electric Fields. *Cell* 169:1029-1041.e16.

Haslacher D, Nasr K, Robinson SE, Braun C, Soekadar SR (2021) A set of electroencephalographic (EEG) data recorded during amplitude-modulated transcranial alternating current stimulation (AM-tACS) targeting 10-Hz steady-state visually evoked potentials (SSVEP). *Data Brief* 36:107011.

Hauswald A, Keitel A, Chen Y-P, Rösch S, Weisz N (2022) Degradation levels of continuous speech affect neural speech tracking and alpha power differently. *Eur J Neurosci* 55:3288–3302.

Hauswald A, Lithari C, Collignon O, Leonardelli E, Weisz N (2018) A Visual Cortical Network for Deriving Phonological Information from Intelligible Lip Movements. *Curr Biol* 28:1453-1459.e3.

Helfrich RF, Fiebelkorn IC, Szczepanski SM, Lin JJ, Parvizi J, Knight RT, Kastner S (2018) Neural Mechanisms of Sustained Attention Are Rhythmic. *Neuron* 99:854-865.e5.

Henry MJ, Herrmann B, Kunke D, Obleser J (2017) Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nat Commun* 8:15801.

Henry MJ, Obleser J, Crusey M, Peelle JE (2022) How strong is the rhythm of perception? A registered replication of Hickok, Farahbod, and Saberi (2015). Available at: <https://psyarxiv.com/62r3u/>.

Herrmann B, Henry MJ, Obleser J (2013a) Frequency-specific adaptation in human auditory cortex depends on the spectral variance in the acoustic stimulation. *J Neurophysiol* 109:2086–2096.

Herrmann CS, Rach S, Neuling T, Strüber D (2013b) Transcranial alternating current stimulation: a review of the underlying mechanisms and modulation of cognitive processes. *Front Hum Neurosci* 7.

Herrmann CS, Strüber D, Helfrich RF, Engel AK (2016) EEG oscillations: From correlation to causality. *Int J Psychophysiol* 103:12–21.

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

Ho HT, Burr DC, Alais D, Morrone MC (2019) Auditory Perceptual History Is Propagated through Alpha Oscillations. *Curr Biol* 29:4208-4217.e3.

Ho HT, Leung J, Burr DC, Alais D, Morrone MC (2017) Auditory Sensitivity and Decision Criteria Oscillate at Different Frequencies Separately for the Two Ears. *Curr Biol* 27:3643-3649.e3.

Howell P (1988) Prediction of P-center location from the distribution of energy in the amplitude envelope: I. *Percept Psychophys* 43:90–93.

Jensen O, Mazaheri A (2010) Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Front Hum Neurosci* 4.

Kasten FH, Duecker K, Maack MC, Meiser A, Herrmann CS (2019) Integrating electric field modeling and neuroimaging to explain inter-individual variability of tACS effects. *Nat Commun* 10:1–11.

Keitel C, Quigley C, Ruhnau P (2014) Stimulus-Driven Brain Oscillations in the Alpha Range: Entrainment of Intrinsic Rhythms or Frequency-Following Response? *J Neurosci* 34:10137–10140.

Keshavarzi M, Kegler M, Kadir S, Reichenbach T (2020) Transcranial alternating current stimulation in the theta band but not in the delta band modulates the comprehension of naturalistic speech in noise. *NeuroImage* 210:116557.

Khoury L, Nelken I (2015) Detecting the unexpected. *Curr Opin Neurobiol* 35:142–147.

Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Res Rev* 53:63–88.

Köhler MHA, Demarchi G, Weisz N (2021) Cochlear activity in silent cue-target intervals shows a theta-rhythmic pattern and is correlated to attentional alpha and theta modulations. *BMC Biol* 19:48.

Kubovy M (1988) Should we resist the seductiveness of the space:time::vision:audition analogy? *J Exp Psychol Hum Percept Perform* 14:318–320.

Laakso I, Tanaka S, Koyama S, De Santis V, Hirata A (2015) Inter-subject Variability in Electric Fields of Motor Cortical tDCS. *Brain Stimulat* 8:906–913.

Labruna L, Merrick C, Inglis B, Ivry R, Sheltraw D (2021) kTMP: A New Non-invasive Magnetic Induction Method to Modulate Cortical Excitability. *BioRxiv*, available at: <https://www.biorxiv.org/content/10.1101/2021.11.17.465477v1>.

Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8:194–208.

Lafon B, Henin S, Huang Y, Friedman D, Melloni L, Thesen T, Doyle W, Buzsáki G, Devinsky O, Parra LC, A. Liu A (2017) Low frequency transcranial electrical stimulation does not entrain sleep rhythms measured by human intracranial recordings. *Nat Commun* 8:1199.

Lakatos P, Barczak A, Neymotin SA, McGinnis T, Ross D, Javitt DC, O’Connell MN (2016) Global dynamics of selective attention and its lapses in primary auditory cortex. *Nat Neurosci* 19:1707–1717.

Lakatos P, Gross J, Thut G (2019) A New Unifying Account of the Roles of Neuronal Entrainment. *Curr Biol* 29:R890–R905.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320:110–113.

Lakatos P, Musacchia G, O’Connell MN, Falchier AY, Javitt DC, Schroeder CE (2013a) The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77:750–761.

Lakatos P, O’Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The leading sense: supramodal control of neurophysiological context by attention. *Neuron* 64:419–430.

Lakatos P, Schroeder CE, Leitman DI, Javitt DC (2013b) Predictive suppression of cortical excitability and its deficit in schizophrenia. *J Neurosci* 33:11692–11702.

Landau AN, Fries P (2012) Attention samples stimuli rhythmically. *Curr Biol* 22:1000–1004.

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

Large EW, Jones MR (1999) The dynamics of attending: How people track time-varying events. *Psychol Rev* 106:119–159.

Lee CC, Imaizumi K, Schreiner CE, Winer JA (2004) Concurrent Tonotopic Processing Streams in Auditory Cortex. *Cereb Cortex* 14:441–451.

Lim S-J, Wöstmann M, Obleser J (2015) Selective Attention to Auditory Memory Neurally Enhances Perceptual Precision. *J Neurosci* 35:16094–16104.

Lin WM, Oettringer DA, Bakker-Marshall I, Emmerzaal J, Wilsch A, ElShafei HA, Rassi E, Haegens S (2022) No behavioural evidence for rhythmic facilitation of perceptual discrimination. *Eur J Neurosci* 55:3352–3364.

Lippe WR (1995) Relationship between frequency of spontaneous bursting and tonotopic position in the developing avian auditory system. *Brain Res* 703:205–213.

- Lisman** JE, Jensen O (2013) The Theta-Gamma Neural Code. *Neuron* 77:1002–1016.
- Livingston** G et al. (2020) Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet* 396:413–446.
- Luo** H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54:1001–1010.
- L’Hermite, S.**, and Zoefel, B. (2022). Rhythmic entrainment echoes in auditory perception. *BioRxiv*, available at: <https://doi.org/10.1101/2022.12.07.519456>
- Marzbani** H, Marateb HR, Mansourian M (2016) Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications. *Basic Clin Neurosci* 7:143–158.
- Mitzdorf** U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol Rev* 65:37–100.
- Moore** BCJ, Gockel HE (2012) Properties of auditory stream formation. *Philos Trans R Soc Lond B Biol Sci* 367:919–931.
- Negahbani** E, Kasten FH, Herrmann CS, Fröhlich F (2018) Targeting alpha-band oscillations in a cortical model with amplitude-modulated high-frequency transcranial electric stimulation. *NeuroImage* 173:3–12.
- Nicholson** C, Freeman JA (1975) Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. *J Neurophysiol* 38:356–368.
- Nobre** AC, van Ede F (2018) Anticipated moments: temporal structure in attention. *Nat Rev Neurosci* 19:34–48.
- Noury** N, Hipp JF, Siegel M (2016) Physiological processes non-linearly affect electrophysiological recordings during transcranial electric stimulation. *NeuroImage* 140:99–109.
- Obleser** J, Kayser C (2019) Neural Entrainment and Attentional Selection in the Listening Brain. *Trends Cogn Sci* 23:913–926.
- Obleser** J, Weisz N (2012) Suppressed Alpha Oscillations Predict Intelligibility of Speech and its Acoustic Details. *Cereb Cortex* 22:2466–2477.
- O’Connell** MN, Barczak A, Schroeder CE, Lakatos P (2014) Layer Specific Sharpening of Frequency Tuning by Selective Attention in Primary Auditory Cortex. *J Neurosci* 34:16496–16508.
- O’Connell** MN, Falchier A, McGinnis T, Schroeder CE, Lakatos P (2011) Dual Mechanism of Neuronal Ensemble Inhibition in Primary Auditory Cortex. *Neuron* 69:805–817.
- Parker** DA, Hamm JP, McDowell JE, Keedy SK, Gershon ES, Ivleva EI, Pearlson GD, Keshavan MS, Tamminga CA, Sweeney JA, Clementz BA (2019) Auditory steady-state EEG response across the schizobipolar spectrum. *Schizophr Res* 209:218–226.
- Peelle** JE, Davis MH (2012) Neural Oscillations Carry Speech Rhythm through to Comprehension. *Front Psychol* 3:320.
- Peelle** JE, Gross J, Davis MH (2013) Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb Cortex* 23:1378–1387.
- Petersen** EB, Wöstmann M, Obleser J, Lunner T (2016) Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J Neurophysiol* 117:18–27.
- Poeppel** D, Assaneo MF (2020) Speech rhythms and their neural foundations. *Nat Rev Neurosci* 21:322–334.

- Prado-Gutierrez P**, Castro-Fariñas A, Morgado-Rodríguez L, Velarde-Reyes E, Martínez AD, Martínez-Montes E (2015) Habituation of Auditory Steady State Responses Evoked by Amplitude-Modulated Acoustic Signals in Rats. *Audiol Res* 5:113.
- Puonti O**, Van Leemput K, Saturnino GB, Siebner HR, Madsen KH, Thielscher A (2020) Accurate and robust whole-head segmentation from magnetic resonance images for individualized head modeling. *NeuroImage* 219:117044.
- Reason J** (2000) Human error. *West J Med* 172:393–396.
- Remez RE**, Rubin PE, Pisoni DB, Carrell TD (1981) Speech perception without traditional speech cues. *Science* 212:947–949.
- Riecke L**, Formisano E, Sorger B, Başkent D, Gaudrain E (2018) Neural Entrainment to Speech Modulates Speech Intelligibility. *Curr Biol* 28:161-169.e5.
- Riecke L**, Zoefel B (2018) Conveying Temporal Information to the Auditory System via Transcranial Current Stimulation. *Acta Acust Unit Acust* 104, 883-886.
- Romei V**, Brodbeck V, Michel C, Amedi A, Pascual-Leone A, Thut G (2008) Spontaneous Fluctuations in Posterior α -Band EEG Activity Reflect Variability in Excitability of Human Visual Areas. *Cereb Cortex* 18:2010–2018.
- Romei V**, Thut G, Silvanto J (2016) Information-Based Approaches of Noninvasive Transcranial Brain Stimulation. *Trends Neurosci*, 782–795.
- Römer H** (1983) Tonotopic organization of the auditory neuropile in the bushcricket *Tettigonia viridissima*. *Nature* 306:60–62.
- Roth MJ**, Synofzik M, Lindner A (2013) The Cerebellum Optimizes Perceptual Predictions about External Sensory Events. *Curr Biol* 23:930–935.
- Saberi K**, Hickok G (2022a) Forward entrainment: Psychophysics, neural correlates, and function. *Psychon Bull Rev*. doi: 10.3758/s13423-022-02220-y
- Saberi K**, Hickok G (2022b) Confirming an antiphase bicyclic pattern of forward entrainment in signal detection: A reanalysis of Sun et al. (2021). *Eur J Neurosci* 56:5274–5286.
- Saberi K**, Hickok G (2022c) A critical analysis of Lin et al.'s (2021) failure to observe forward entrainment in pitch discrimination. *Eur J Neurosci* 56:5191–5200.
- Saenz M**, Langers DRM (2014) Tonotopic mapping of human auditory cortex. *Hear Res* 307:42–52.
- Salvador R**, Biagi MC, Puonti O, Splittgerber M, Moliadze V, Siniatchkin M, Thielscher A, Ruffini G (2021) Personalization of Multi-electrode Setups in tCS/tES: Methods and Advantages. In: *Brain and Human Body Modeling 2020: Computational Human Models* (Makarov SN, Noetscher GM, Nummenmaa A, eds). Cham: Springer.
- Saturnino GB**, Madsen KH, Siebner HR, Thielscher A (2017) How to target inter-regional phase synchronization with dual-site Transcranial Alternating Current Stimulation. *NeuroImage* 163:68–80.
- Saturnino GB**, Madsen KH, Thielscher A (2020) Optimizing the electric field strength in multiple targets for multichannel transcranial electric stimulation. *J Neural Eng* 18.
- Schroeder CE**, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18.
- Schroeder CE**, Mehta AD, Givre SJ (1998) A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb Cortex* 8:575–592.

- Schwartz** M, Keller PE, Kotz SA (2016) Spontaneous, synchronized, and corrective timing behavior in cerebellar lesion patients. *Behav Brain Res* 312:285–293.
- Scott** SK (1998) The point of P-centres. *Psychol Res* 61:4–11.
- Segaert** K, Weber K, de Lange FP, Petersson KM, Hagoort P (2013) The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia* 51:59–66.
- Shamma** S (1988) The acoustic features of speech sounds in a model of auditory processing: vowels and voiceless fricatives. *J Phon* 16:77–91.
- Shannon** RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304.
- Stockert** A, Schwartz M, Poeppel D, Anwander A, Kotz SA (2021) Temporo-cerebellar connectivity underlies timing constraints in audition *eLife* 10:e67303.
- Sun** Y, Michalareas G, Poeppel D (2022) The impact of phase entrainment on auditory detection is highly variable: Revisiting a key finding. *Eur J Neurosci* 55:3373–3390.
- Thielscher** A, Antunes A, Saturnino GB (2015) Field modeling for transcranial magnetic stimulation: A useful tool to understand the physiological effects of TMS? *IEEE Eng Med Biol*: 222–225.
- Thorne** JD, Debener S (2014) Look now and hear what’s coming: On the functional role of cross-modal phase reset. *Hear Res* 307:144–152.
- Thut** G, Miniussi C (2009) New insights into rhythmic brain activity from TMS-EEG studies. *Trends Cogn Sci* 13:182–189.
- Thut** G, Miniussi C, Gross J (2012) The Functional Importance of Rhythmic Activity in the Brain. *Curr Biol* 22:R658–R663.
- Thut** G, Schyns P, Gross J (2011a) Entrainment of Perceptually Relevant Brain Oscillations by Non-Invasive Rhythmic Stimulation of the Human Brain. *Front Psychol* 2.
- Thut** G, Veniero D, Romei V, Miniussi C, Schyns P, Gross J (2011b) Rhythmic TMS Causes Local Entrainment of Natural Oscillatory Signatures. *Curr Biol* 21:1176–1185.
- Ulanovsky** N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398.
- van Bree** S, Alamia A, Zoefel B (2022) Oscillation or not-Why we can and need to know (commentary on Doelling and Assaneo, 2021). *Eur J Neurosci* 55:201–204.
- van Bree** S, Sohoglu E, Davis MH, Zoefel B (2021) Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLOS Biol* 19:e3001142.
- van Noorden** LPAS (1975) Temporal coherence in the perception of tone sequences. *PhD thesis*, University of Technology, Eindhoven, The Netherlands.
- VanRullen** R (2016) Perceptual Cycles. *Trends Cogn Sci* 20:723–735.
- VanRullen** R, Koch C (2003) Is perception discrete or continuous? *Trends Cogn Sci* 7:207–213.
- VanRullen** R, Zoefel B, Ilhan B (2014) On the cyclic nature of perception in vision versus audition. *Philos Trans R Soc Lond B Biol Sci* 369:20130214.
- Varnet** L, Ortiz-Barajas MC, Erra RG, Gervain J, Lorenzi C (2017) A cross-linguistic study of speech modulation spectra. *J Acoust Soc Am* 142:1976.

Vöröslakos M, Takeuchi Y, Brinyiczki K, Zombori T, Oliva A, Fernández-Ruiz A, Kozák G, Kincses ZT, Iványi B, Buzsáki G, Berényi A (2018) Direct effects of transcranial electric stimulation on brain circuits in rats and humans. *Nat Commun* 9:483.

Walter VJ, Walter WG (1949) The central effects of rhythmic sensory stimulation. *Electroencephalogr Clin Neurophysiol* 1:57–86.

Weisz N, Hartmann T, Müller N, Obleser J (2011) Alpha Rhythms in Audition: Cognitive and Clinical Perspectives. *Front Psychol* 2.

Wessel MJ, Draaisma LR, Hummel FC (2023) Mini-review: Transcranial Alternating Current Stimulation and the Cerebellum. *Cerebellum* 22:120–128.

Widmann A, Schröger E, Maess B (2015) Digital filter design for electrophysiological data – a practical approach. *J Neurosci Methods* 250:34–46.

Wilsch A, Neuling T, Obleser J, Herrmann CS (2018) Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *NeuroImage* 172:766–774.

Wöstmann M, Lim S-J, Obleser J (2017) The Human Neural Alpha Response to Speech is a Proxy of Attentional Control. *Cereb Cortex* 27:3307–3317.

Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, Mckhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a 'Cocktail Party.' *Neuron* 77:980–991.

Zoefel B, Allard I, Anil M, Davis MH (2020) Perception of Rhythmic Speech Is Modulated by Focal Bilateral Transcranial Alternating Current Stimulation. *J Cogn Neurosci* 32:226–240.

Zoefel B, Archer-Boyd A, Davis MH (2018a) Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr Biol* 28:401-408.e5.

Zoefel B, Costa-Faidella J, Lakatos P, Schroeder CE, VanRullen R (2017) Characterization of neural entrainment to speech with and without slow spectral energy fluctuations in laminar recordings in monkey A1. *NeuroImage* 150:344–357.

Zoefel B, Davis MH (2017) Transcranial electric stimulation for the investigation of speech perception and comprehension. *Lang Cogn Neurosci* 32:910–923.

Zoefel B, Davis MH, Valente G, Riecke L (2019) How to test for phasic modulation of neural and behavioural responses. *NeuroImage* 202:116175.

Zoefel B, Gilbert RA, Davis MH (2023) Intelligibility improves perception of timing changes in speech. *PLOS ONE* 18:e0279024.

Zoefel B, Heil P (2013) Detection of Near-Threshold Sounds is Independent of EEG Phase in Common Frequency Bands. *Front Psychol* 4.

Zoefel B, Huster RJ, Herrmann CS (2011) Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *NeuroImage* 54:1427–1431.

Zoefel B, Kösem A (2022) Neural Dynamics: Speech is Special. *PsyArXiv*, available at: <https://psyarxiv.com/eyrtb/>

Zoefel B, Reddy Pasham N, Brüers S, VanRullen R (2015) The ability of the auditory system to cope with temporal subsampling depends on the hierarchical level of processing. *Neuroreport* 26:773-778.

Zoefel B, ten Oever S, Sack AT (2018b) The Involvement of Endogenous Neural Oscillations in the Processing of Rhythmic Input: More Than a Regular Repetition of Evoked Neural Responses. *Front Neurosci* 12.

Zoefel B, VanRullen R (2015a) Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *J Neurosci* 35:1954–1964.

Zoefel B, VanRullen R (2015b) The Role of High-Level Processes for Oscillatory Phase Entrainment to Speech Sound. *Front Hum Neurosci* 9.

Zoefel B, VanRullen R (2016) EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage* 124:16–23.

Zoefel B, VanRullen R (2017) Oscillatory Mechanisms of Stimulus Processing and Selection in the Visual and Auditory Systems: State-of-the-Art, Speculations and Suggestions. *Front Neurosci* 11.