

Temporal and Hemispheric Dynamics in Neural Processing of Auditory Stimuli Across Linguistic Complexity: A MEG Source Space Study

Patrick Krauss^{1,2,*,#}, Nikola Kölbl^{1,2,*}, Nadia Müller-Voggel³, Stefan Rampp³, Martin Kaltenhäuser³, Konstantin Tzirdis¹, Achim Schilling^{1,2}

¹ Neuroscience Lab, University Hospital Erlangen, Erlangen, Germany

² CCN Group, Pattern Recognition Lab, University Erlangen-Nürnberg

³ MEG Center, University Hospital Erlangen, Erlangen, Germany

* authors contributed equally

patrick.krauss@fau.de

Abstract

In this study we investigated the neural processing of auditory stimuli of varying complexity: a non-linguistic (pure tone), a simple linguistic (phoneme) and a complex linguistic (word) stimulus. We recorded brain activity of 30 healthy, right-handed participants using magnetoencephalography (MEG), and compared the resulting evoked fields (ERFs) in source space in three different time intervals, i.e. early (0-250ms), mid (250-500ms) and late (500-750ms) responses. Our results reveal a bilateral activation during early response and right-lateralized activation in the mid-phase for all stimuli. However, the late response exhibited lateralization variations. The pure tone predominantly activated the right hemisphere, consistent with pitch processing theories. The phoneme primarily engaged the left hemisphere, supporting its role in phonemic processing. Notably, the word elicited activation in both hemispheres, reflecting phonemic processing on the left and stress patterns on the right. These findings highlight the intricate interplay between temporal processing and hemispheric lateralization in speech perception, emphasizing the importance of stimulus complexity and temporal dynamics in understanding auditory and speech processing.

Keywords: MEG; event related fields (ERFs); source space analysis; language; speech; auditory perception

Introduction

The brain's capacity to process auditory stimuli, from basic tones to complex speech, showcases the intricate neural mechanisms involved in auditory perception and language processing. This ability is crucial for effective communication, enabling both sound recognition and interpretation within linguistic contexts. Through many functional MRI (Goucha & Friederici, 2015), but also (invasive) electroencephalography (EEG) (Metzger et al., 2020; Alday, 2019) and magnetoencephalography (MEG) (Tavabi, Obleser, Dobel, & Pantev, 2007; Hickok & Poeppel, 2007; Poeppel & Assaneo, 2020)

studies, we know which brain regions are active in language processing, but the exact neuronal circuits and their temporal dynamics are still not fully understood. To unravel these neuronal processes, in recent years, neurolinguistic research has increasingly used natural continuous speech such as audio books as a stimulus (Schilling et al., 2021; Garibyan, Schilling, Boehm, Zankl, & Krauss, 2022; Koelbl, Schilling, & Krauss, 2023; Schüller, Schilling, Krauss, Rampp, & Reichenbach, 2023; Schüller, Schilling, Krauss, & Reichenbach, 2024). In this preliminary study, we want to decode human speech by analysing data from 30 healthy, right-handed subjects recorded with MEG, in an effort to apply the resulting knowledge to further studies of speech analysis. Given the high temporal resolution of MEG, we want to analyse not only localisation but also temporal dynamics of brain activity using event-related-fields (ERFs) in source space. For this purpose, we have chosen three auditory stimuli of different complexity: a non-linguistic pure tone (5kHz), a simple linguistic phoneme ("i"), and a more complex linguistic stimulus, i.e. a single word ("gut"; German for "good"). In our analysis we find statistically significant differences in the spatio-temporal patterns evoked by the different stimuli in the primary auditory cortex, e.g. Heschl Gyrus (HG).

Methods

Each stimulus was presented bilateral 30 times in a randomized sequence with a 1-second interval between onsets to 30 healthy, right-handed participants (15 females, average age: 22.9 ± 3) using a 248-channel MEG system (Magnes 3600WH, 4D-Neuroimaging). All measurements were approved by the Ethics Committee of the University Hospital Erlangen (No: 22-361-2). Data preparation involved using *MNE* software to identify and remove bad sensors, applying band-pass filters (1-20 Hz), down-sampling the data to 200 Hz and conducting independent component analysis to remove artifacts (Koelbl et al., 2023). Finally, the data were segmented into intervals from -0.5s to 1s corresponding to stimulus presentations. We calculated a vector volume source estimate (SE) for each subject and each stimulus utiliz-

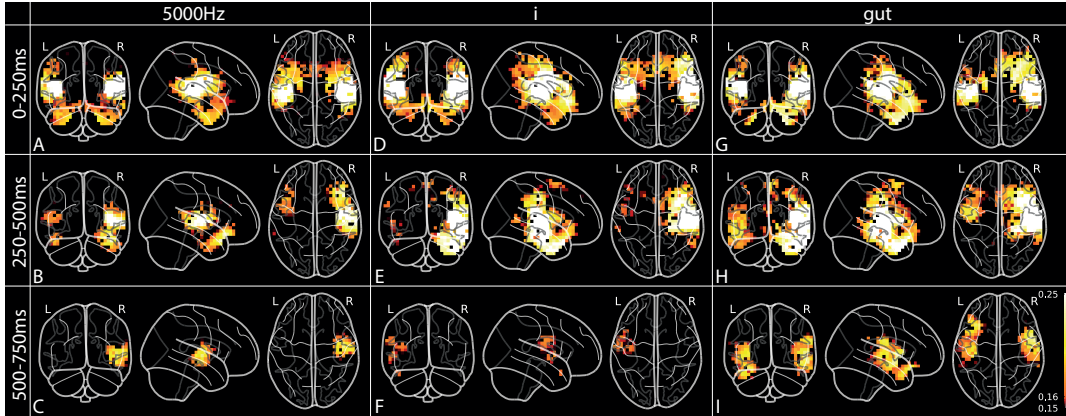


Figure 1: Grand Average Source Estimates over 30 participants for three stimuli (each presented 30 times) averaged over three time intervals. (A-C): Average brain response for pure tone 5kHz from 0-250ms, 250-500ms and 500-750ms. (D-F): Responses for phoneme "i". (G-I): Responses for word "gut".

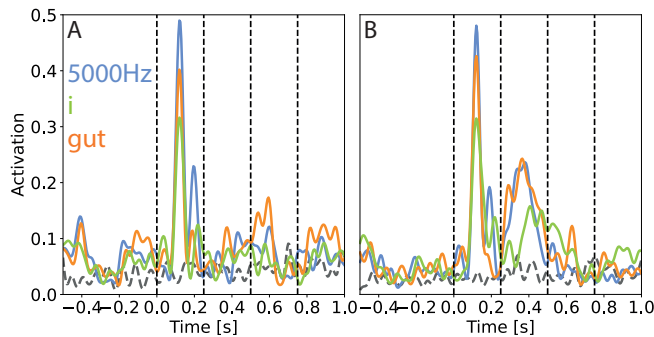


Figure 2: (A): Grand Average activity in left HG for three stimuli. (B): Same as (A), but in the right HG. Grey signal: Activity of random time-points.

ing a cortical source space (5mm grid) and the boundary element model of the template brain *fsaverage* from *FreeSurfer* (<https://surfer.nmr.mgh.harvard.edu/>) with method="sLORETA". We calculated the Grand Average for each stimulus over all subjects SEs, time-averaged the resulting Grand Average SEs for three time intervals (early (0-250ms), mid (250-500ms), and late response (500-750ms)) and visualized them (Figure 1). Additionally, we examined activation differences in the primary auditory cortex, focusing on the HG, using label-time-courses within this region (Figure 2). For statistical evaluation we calculated paired Wilcoxon tests using the root-mean-square (RMS) amplitudes in the left and right HG (Table 1).

Results

Our findings (Figure 1) indicate bilateral activation during the early response phase, despite amplitude differences possibly linked to audio volume. In the mid response phase, right-lateralized activation is prominent for all three stimuli, with larger amplitudes for the two linguistic stimuli. The late response phase unveiled dis-

tinct patterns: right-lateralized for the pure tone, left-lateralized for the phoneme and bilateral for the word, implying varied processing mechanisms. The average time courses in the left and right HG show similar patterns (Figure 2). The corresponding Wilcoxon tests of response amplitudes affirmed Bonferroni-corrected significant p-values (colored in red in Table 1), among others, between the non-linguistic and the complex linguistic stimuli in the mid time interval in both HGs and between the two linguistic stimuli in the late interval in the right HG.

left HG	0-250ms	250-500ms	500-750ms
5kHz vs. i	0.0004	0.0155	0.0136
5kHz vs. gut	0.0577	0.0016	0.0047
i vs. gut	0.0497	0.3387	0.0327
right HG	0-250ms	250-500ms	500-750ms
5kHz vs. i	0.0234	0.0050	0.9354
5kHz vs. gut	0.3707	0.0001	0.0137
i vs. gut	0.0522	0.1981	0.0040

Table 1: p-values of differences in activities in the left and right HG. Evaluated using a paired Wilcoxon Test with the RMS-amplitudes in three time intervals of three stimuli (5kHz, "i", "gut") of 30 participants. Bonferroni-corrected significant p-values shown in red.

Discussion

Our research reveals the complexity of auditory and speech processing, showing specific activation patterns which differ particularly in the late processing phase depending on stimulus complexity. The predominance of right hemisphere activation could be related to the rhythmic presentation of stimuli, potentially reflecting the right hemisphere's involvement in processing pitch and melody (Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002). Conversely, left hemisphere activation for the simple linguistic stimulus (phoneme) dur-

ing the late phase aligns with its role in phoneme processing (Desai, Liebenthal, Waldron, & Binder, 2008; Blumstein, Myers, & Rissman, 2005; Hyde, Peretz, & Zatorre, 2008) in particular in HG (Chait, Poeppel, De Cheveigné, & Simon, 2007). Bilateral activation for the more complex linguistic stimulus (word) underscores the intricate nature of language processing, requiring both interpretation of meaning and stress recognition. To deepen our understanding of speech perception's neural underpinnings, future studies will have to explore additional brain areas. Additionally, machine learning could aid in identifying distinct activation clusters, offering deeper insights into the networks facilitating auditory processing and language comprehension.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): grants KR5148/2-1 (project number 436456810), KR5148/3-1 (project number 510395418) and GRK2839 (project number 468527017) to PK, grant TZ100/2-1 (project number 510395418) to KT, and grant SCHI1482/3-1 (project number 451810794) to AS.

References

- Alday, P. M. (2019). M/EEG analysis of naturalistic stories: a review from speech to language processing. *Language, Cognition and Neuroscience*, 34(4), 457–473.
- Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: an fMRI investigation of phonetic category structure. *Journal of cognitive neuroscience*, 17(9), 1353–1366.
- Chait, M., Poeppel, D., De Cheveigné, A., & Simon, J. Z. (2007). Processing asymmetry of transitions between order and disorder in human auditory cortex. *Journal of Neuroscience*, 27(19), 5207–5214.
- Desai, R., Liebenthal, E., Waldron, E., & Binder, J. R. (2008). Left posterior temporal regions are sensitive to auditory categorization. *Journal of Cognitive Neuroscience*, 20(7), 1174–1188.
- Garibyan, A., Schilling, A., Boehm, C., Zankl, A., & Krauss, P. (2022). Neural correlates of linguistic collocations during continuous speech perception. *Frontiers in Psychology*, 13, 1076339.
- Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's area. *NeuroImage*, 114, 294–302.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, 8(5), 393–402.
- Hyde, K. L., Peretz, I., & Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*, 46(2), 632–639.
- Koelbl, N., Schilling, A., & Krauss, P. (2023). Adaptive ICA for speech EEG artifact removal. In *2023 5th international conference on bio-engineering for smart technologies (biosmart)* (pp. 1–4).
- Metzger, B. A., Magnotti, J. F., Wang, Z., Nesbitt, E., Karas, P. J., Yoshor, D., & Beauchamp, M. S. (2020). Responses to visual speech in human posterior superior temporal gyrus examined with iEEG deconvolution. *Journal of Neuroscience*, 40(36), 6938–6948.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36(4), 767–776.
- Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature reviews neuroscience*, 21(6), 322–334.
- Schilling, A., Tomasello, R., Henningsen-Schomers, M. R., Zankl, A., Surendra, K., Haller, M., ... Krauss, P. (2021). Analysis of continuous neuronal activity evoked by natural speech with computational corpus linguistics methods. *Language, Cognition and Neuroscience*, 36(2), 167–186.
- Schüller, A., Schilling, A., Krauss, P., Rampp, S., & Reichenbach, T. (2023). Attentional modulation of the cortical contribution to the frequency-following response evoked by continuous speech. *Journal of Neuroscience*, 43(44), 7429–7440.
- Schüller, A., Schilling, A., Krauss, P., & Reichenbach, T. (2024). The early subcortical response at the fundamental frequency of speech is temporally separated from later cortical contributions. *Journal of Cognitive Neuroscience*, 36(3), 475–491.
- Tavabi, K., Obleser, J., Dobel, C., & Pantev, C. (2007). Auditory evoked fields differentially encode speech features: an MEG investigation of the p50m and n100m time courses during syllable processing. *European journal of neuroscience*, 25(10), 3155–3162.