

FMRI speech tracking in primary and non-primary auditory cortex while listening to noisy scenes

SUPPLEMENTARY INFORMATION

Lars Hausfeld ^{a,b,*}, Iris M.H. Hamers ^{a,c}, Elia Formisano ^{a,b,d}

^a Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD Maastricht, The Netherlands

^b Maastricht Brain Imaging Centre, 6200 MD Maastricht, The Netherlands

^c current affiliation: Department of Biomedical Sciences of Cells & Systems, Section Cognitive Neurosciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

^d Maastricht Centre for Systems Biology, Faculty of Science and Engineering, 6200 MD Maastricht, The Netherlands

* Corresponding author: lars.hausfeld@maastrichtuniversity.nl

Supplementary Figures

Supplementary Figure 1 – Overview of MRI-Design

Supplementary Figure 2 – Sustained Activation and Deactivation during Listening Blocks

Supplementary Figure 3 – Time-to-Peak Dependency of Speech Tracking

Supplementary Figure 4 – Continuous and Binary Envelope Predictors

Supplementary Figure 5 – Speech Envelopes before and after HRF Convolution

Functional run: single speaker (2x)



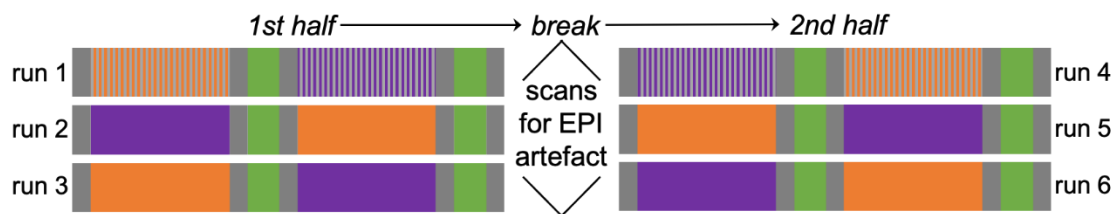
Functional run: auditory scene (4x)



RP: response period

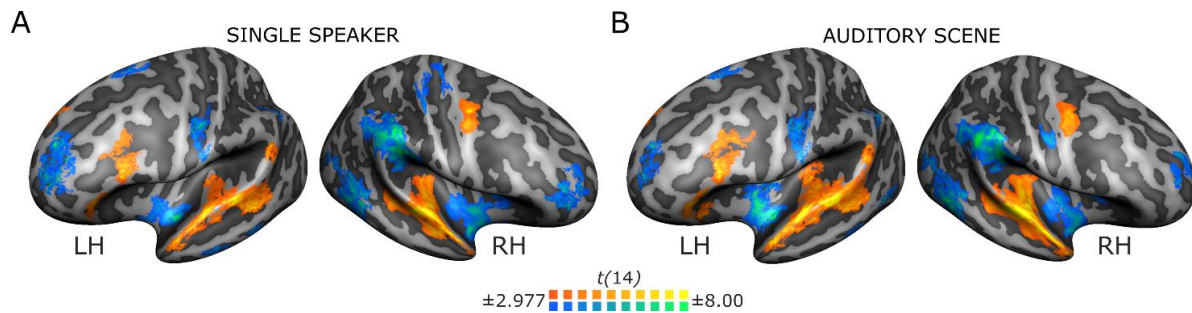
Presentation sequence

presentation order within runs reversed for every other participant



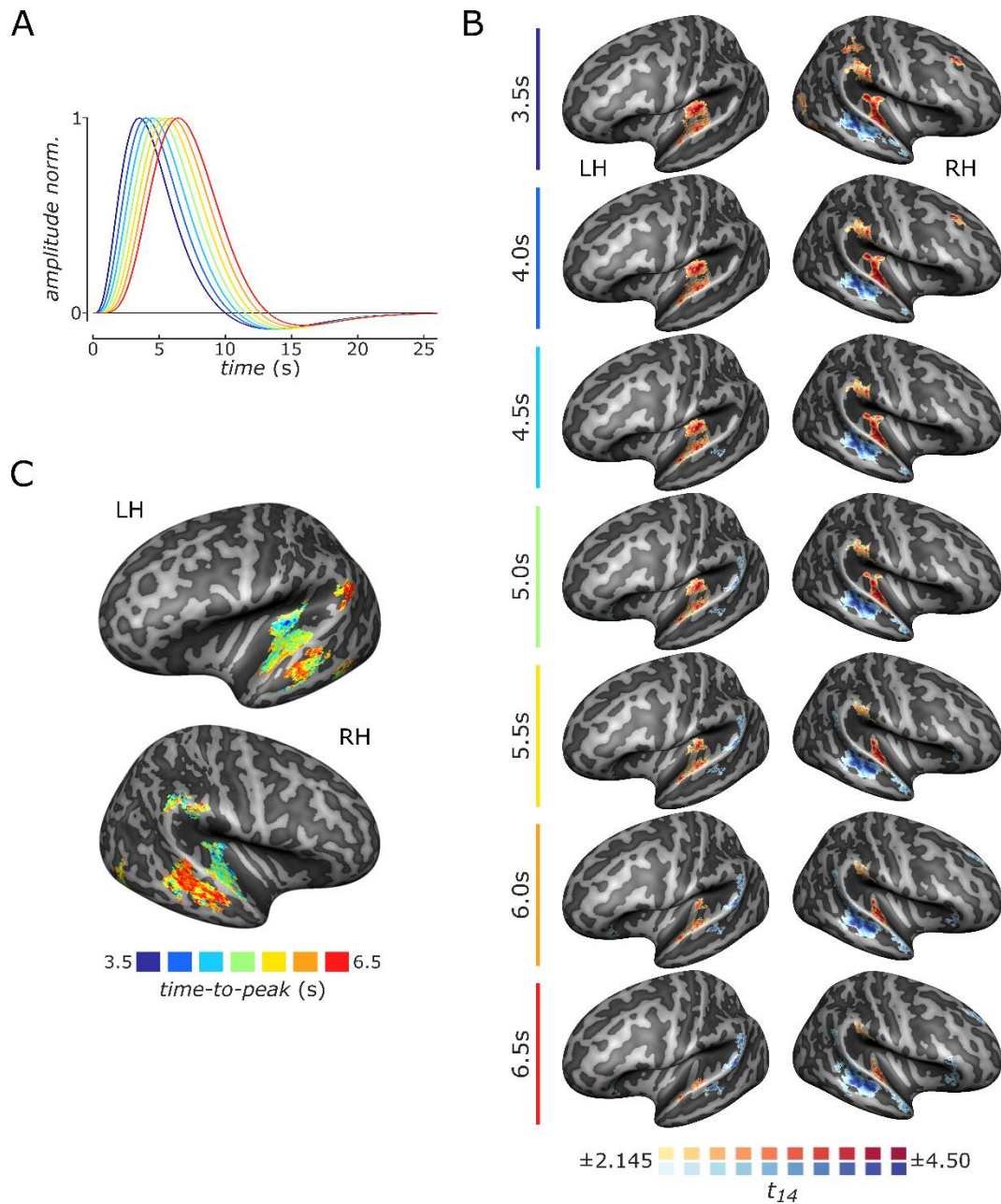
Supplementary Figure 1

Overview of MRI-Design. The upper plot shows examples of functional runs for the single speaker (repeated twice during each scanning sessions) and auditory scene conditions (repeated four times). After presentation of the single speaker or auditory scene stimuli (5 min) and a rest period (15 s), participants were asked to answer 5 questions on the presented content and to rate their listening performance. After another rest period, another single speaker or auditory scene conditions was presented followed by the response period. The lower plot shows the overview one scanning session. These were split in two halves by scans used for EPI artefact correction after 3 functional runs.



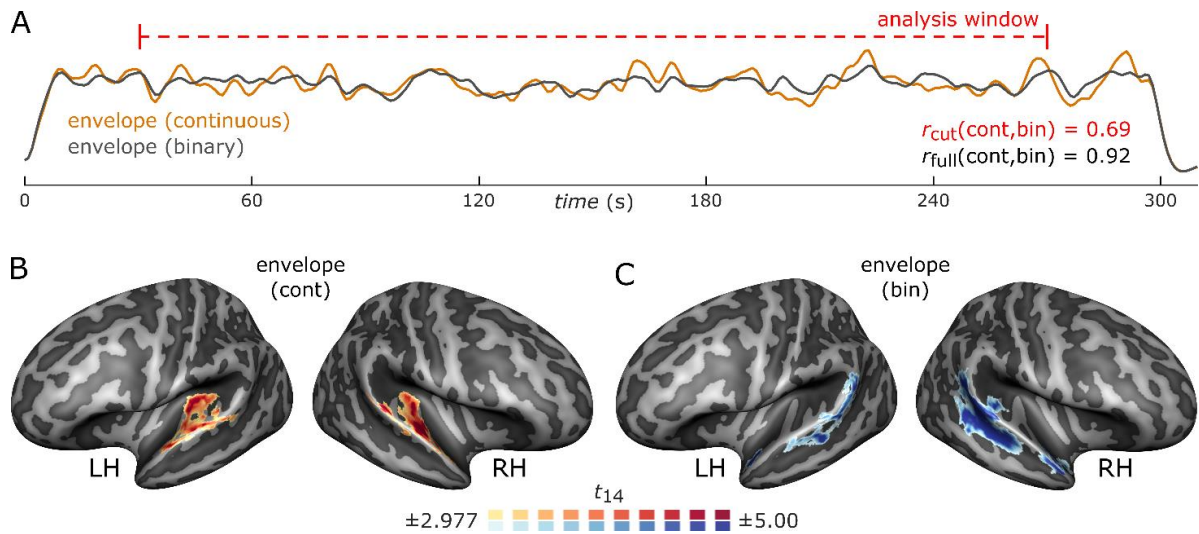
Supplementary Figure 2

Sustained Activation and Deactivation during Listening Blocks. A) Cortical Maps show regions with significant sustained responses during listening blocks as compared to baseline periods (i.e., no sound) during the single speaker condition for the right and left hemisphere. B) Same as A) but for the auditory scene condition. Statistical maps are thresholded at $p < .01$ (two-tailed) and corrected for multiple comparisons by cluster-size ($p < .05$). In addition to the sustained activity predictor, GLMs included predictors coding for onset and offset responses (maps not shown; see Methods - fMRI data analysis - Activation). LH: left hemisphere, RH: right hemisphere.



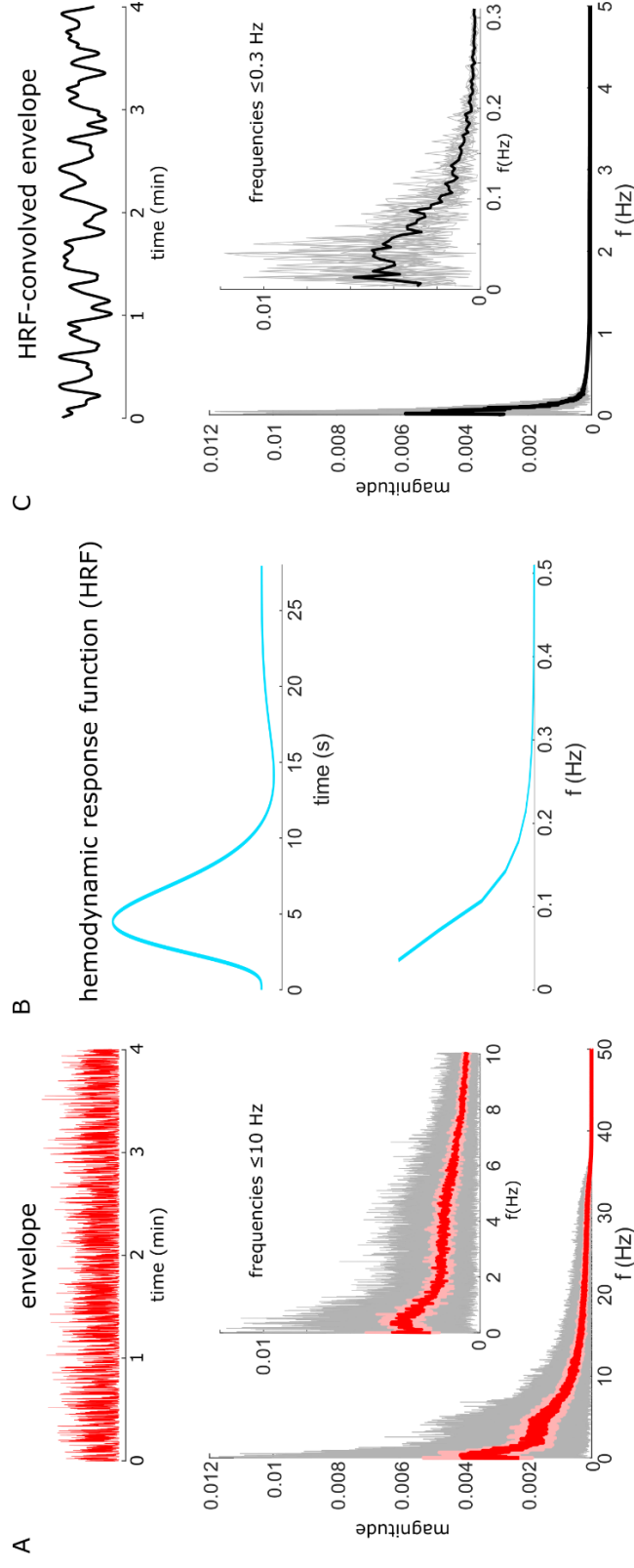
Supplementary Figure 3

Time-to-Peak Dependency of Speech Tracking. A) The hemodynamic response functions (HRFs) created by varying the time-to-peak latency. Colors indicate the HRFs with different time-to-peak parameters (see legend in panel C) for color coding). B) Results of speech tracking for the single speaker condition for analyses with the different HRFs. Color-coded regions show significant speech tracking across participants (multiple comparison corrected by cluster size; $p < .05$; initial threshold $p < .05$, more liberal to show effect of parameter for larger extent). Maps are presented on inflated reconstructions of the average grey-white matter boundary after cortex-based alignment. C) Overview of the time-to-peak parameter resulting in best speech tracking performance. These maps (upper panel: left hemisphere; lower panel: right hemisphere) are restricted to loci showing significant speech tracking for at least one model. LH: left hemisphere, RH: right hemisphere.



Supplementary Figure 4

Continuous and Binary Envelope Predictors. A) HRF-convolved predictors using the continuous (orange) or binary predictor (grey). The binary predictors were computed by thresholding continuous envelopes before convolution and stress sound vs baseline effects. The correlation of the HRF-convolved predictors for the full and cut trial period of this example are indicated on the right in red and black font. B) Areas significantly correlated with the continuous predictor. C) same as in B) but for the binary envelope predictor. Maps are multiple comparison corrected by cluster size ($p < .05$, initial threshold $p < .01$). LH: left hemisphere, RH: right hemisphere.



Supplementary Figure 5

Speech Envelopes before and after HRF Convolution. A) Top: Example time course of the acoustic envelopes used for analyses. Bottom: Spectral magnitude of the long-term envelope of stimuli. Grey lines show individual spectra, light red lines show their average and red lines show the 5-bin running average of the average spectra. B) Amplitude and spectral magnitude of the HRF. C) Same as A) but after convolution with the HRF. Grey lines show individual spectra, the black line shows their average. For comparison, the red line shows the average before convolving with the HRF (same as in panel A). Insets in bottom panels of A) and C) show the spectral magnitude in the lower frequency ranges. Note that the sampling rate was 100Hz throughout; decimation to 1Hz sampling rate for analysis (to match fMRI data) was performed after convolution.