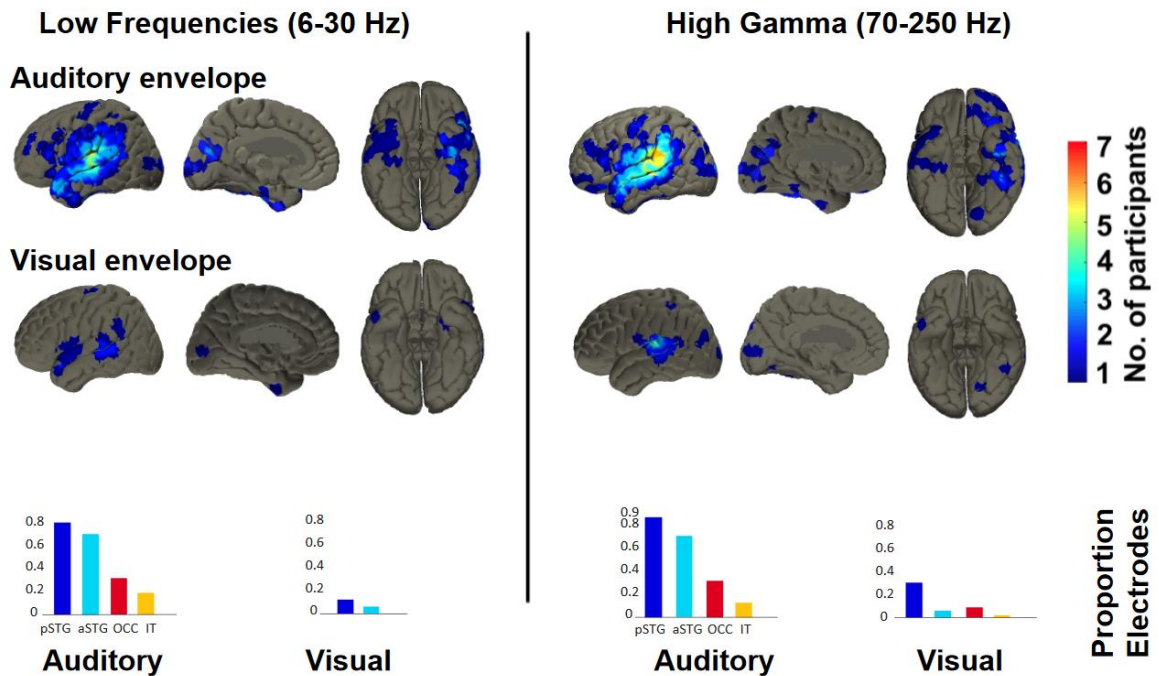


SUPPLEMENTARY FIGURES

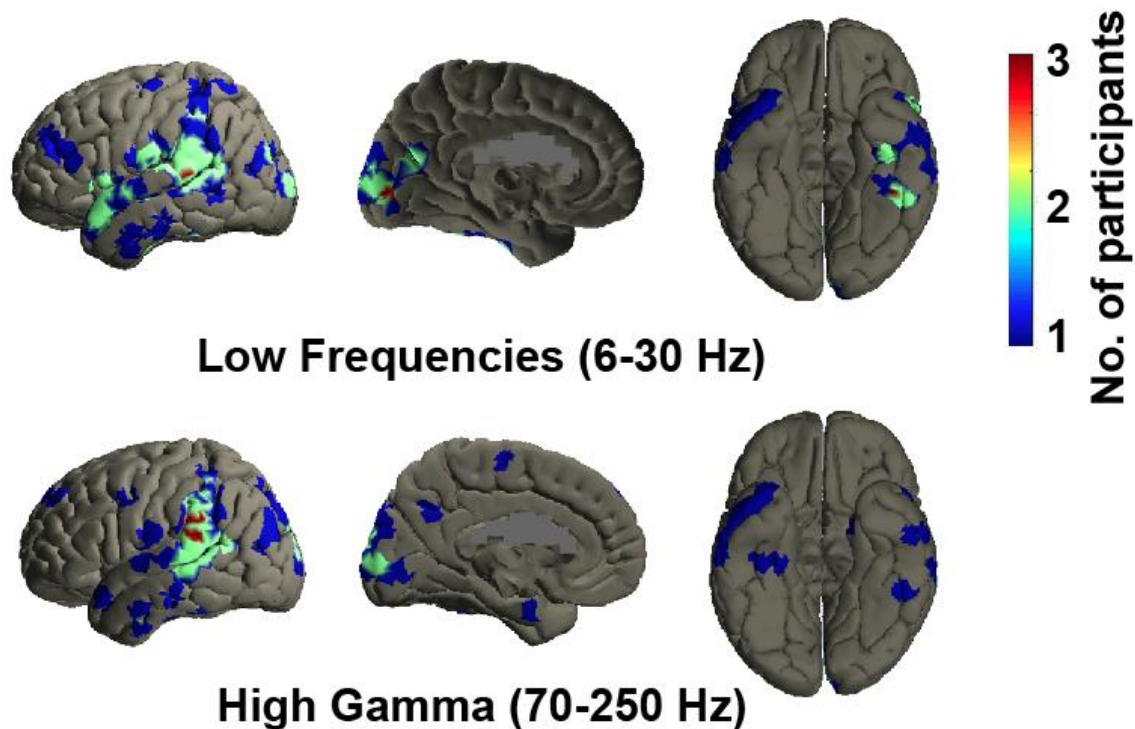
Cumulative areas tracking the speech envelopes



Supplementary Figure 1

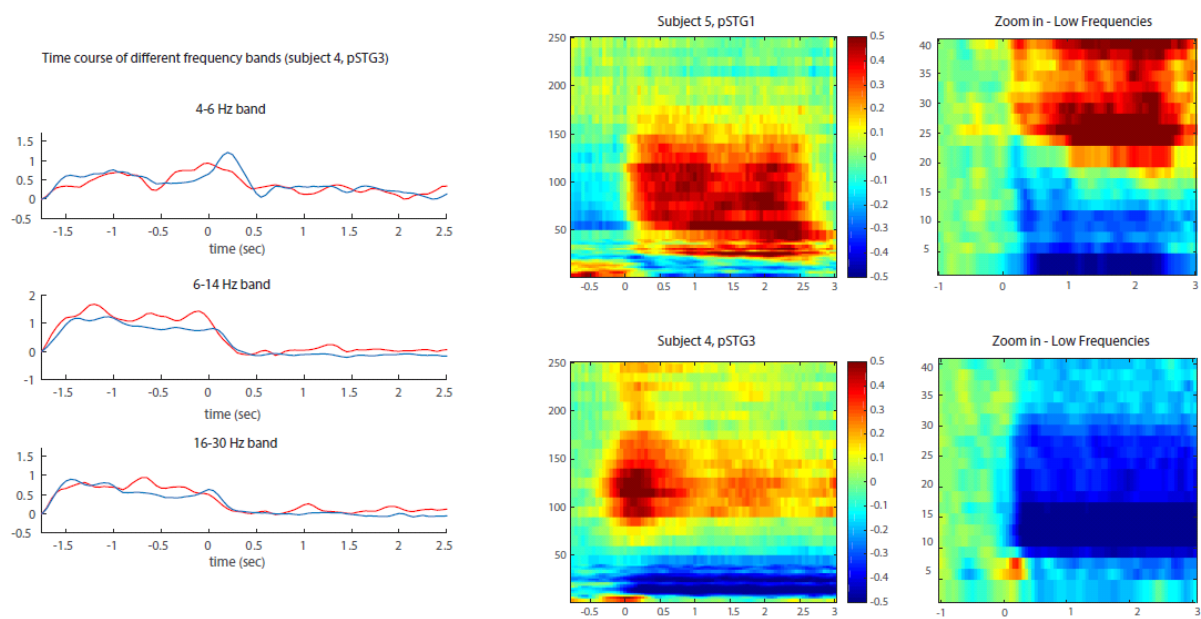
Upper panel) MNI template localization of areas that showed tracking of the auditory speech envelope (upper row) or the envelope of the visible mouth opening (lower row) as revealed by partial correlation. The color coded overlay represents the absolute number of participants per area with significant electrodes over all local electrodes. Note the high density over pSTG and medial occipital cortex, indicating tracking of the auditory and the visual envelope of the speech signal. The effect is more pronounced in the HG- than in the LF-band. See also figure 2 in the main text for a comparison with the proportion of significant electrodes per area.

Lower panel) Proportion of significant electrodes in different cortical areas differentiated per band (left right) and per modality of tracking (Auditory and Visual envelopes). Note that these results are equivalent to the results reported in table 1 in the main text. Acronyms: pSTG - posterior STG, aSTG - anterior STG, IT - inferior temporal + fusiform gyrus + parahippocampal cortex, OCC - occipital cortex.



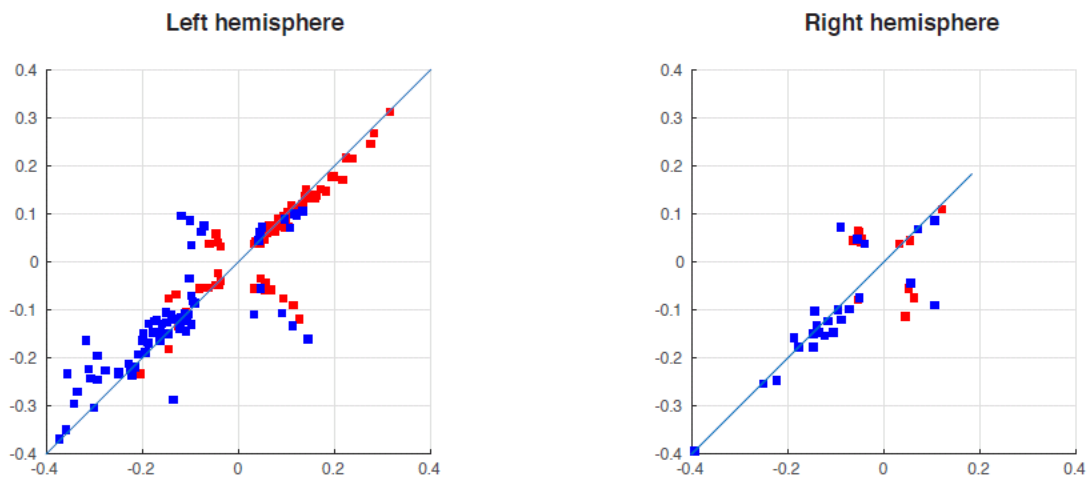
Supplementary Figure 2

Anatomical distribution of participants that show significant effects (corrected for multiple comparisons) of adding speaker mouth movements (AVdyn) to auditory speech (AVstatic) calculated from spectrograms of neural activation in two dynamic bands and aggregated over subjects. Activation in pSTG and in medial visual cortex was modulated by additional visual speech information in both dynamic bands (the overlay is the absolute number of participants cumulated per cortical area). Figure 3 in the main text reports the proportion of participants with significant electrodes, obtained dividing the result of supplementary figure 2 and the density of participants per area (shown in figure 1C).



Supplementary Figure 3

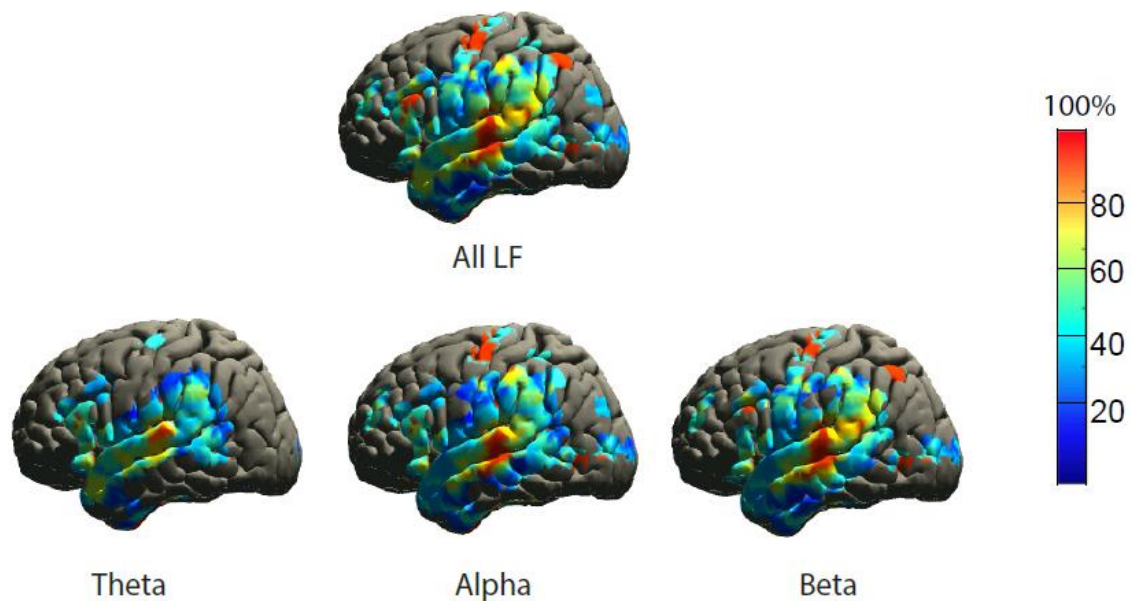
Left: time course of a pSTG electrode (subject 4) in different frequency bands. **Center**: two time-frequency plots for different pSTG electrodes (subject 4 and 5). The 0 on the x-axis denotes speech onset. **Right: zoom in of the LF band (same plots as Center)**. Note that the decrease in magnitude after speech onset shows a consistent pattern in the LF and HG bands. The effect of the LF band is clearly band-limited and distinguishable from a (non-functional) $1/f$ effect magnitude decay along the frequency axis.



Supplementary Figure 4

The panels show the maxima of partial correlations for all significant electrodes in the left hemisphere (left) and in the right hemisphere (right). Each point corresponds to one electrode.

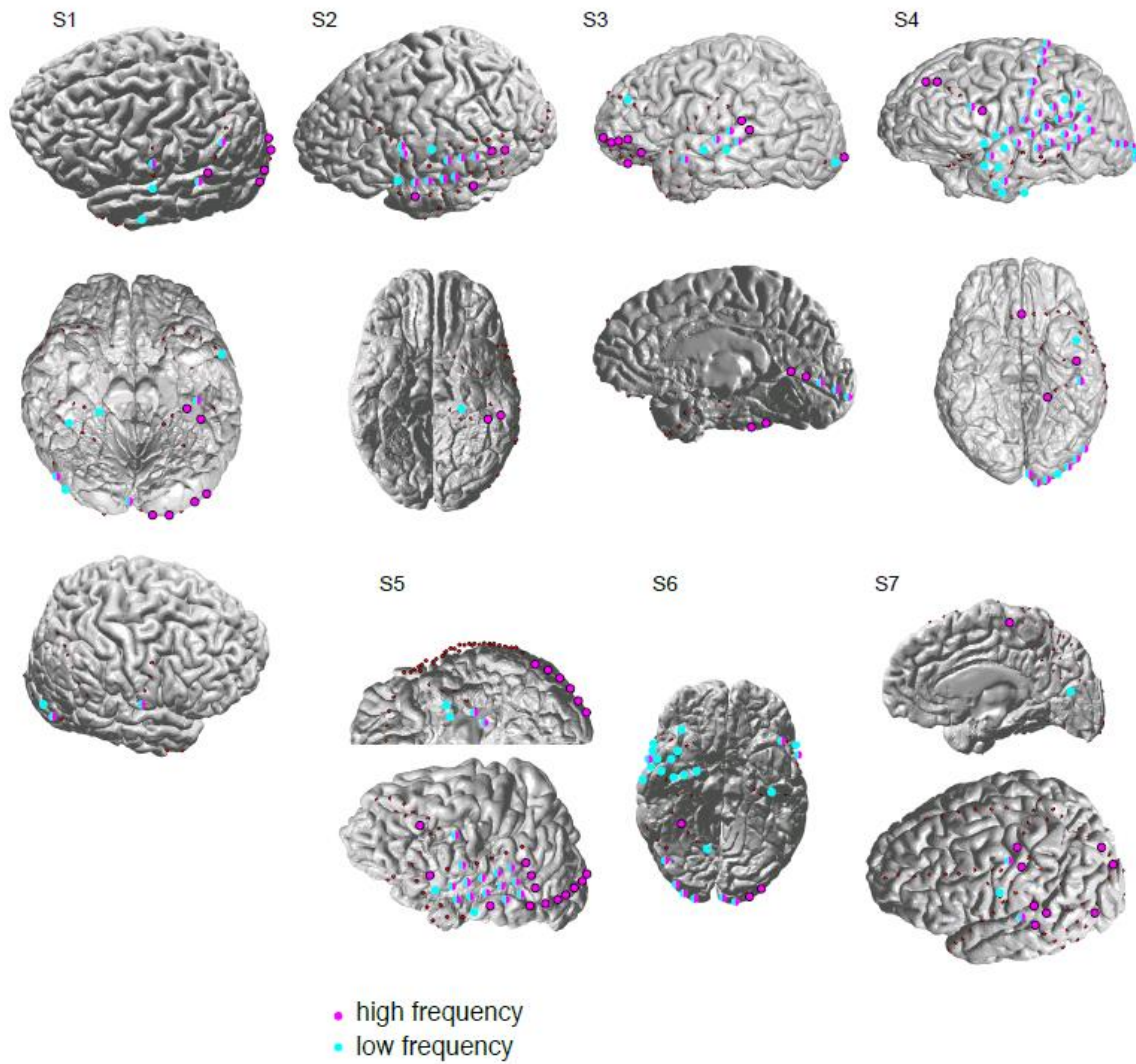
Subdivision of cumulative audio tracking per frequency band



Supplementary Figure 5

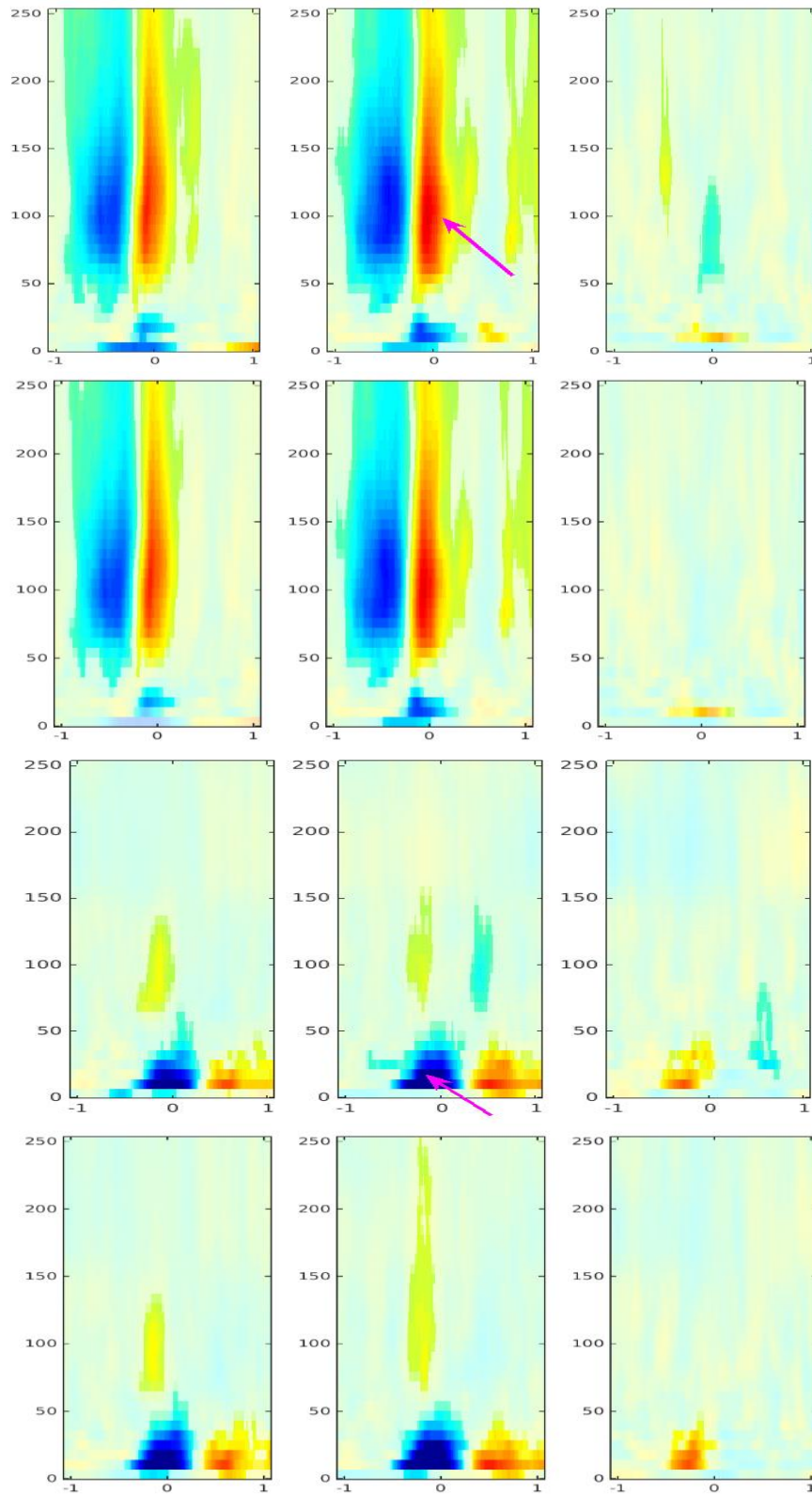
Cumulative plot representing the proportion of subjects with audio-tracking areas, with respect to the total number of electrodes in that area. Above: plot representing the lower frequencies (LF: 6-30 Hz, see also figure 2A). Below: the same overlay is subdivided into three different bands: theta (6-7 Hz), alpha (7-15 Hz), beta (15-30 Hz).

Tracking of audio envelope (per subject/frequency band)



Supplementary Figure 6

Individual subject's brains (S1 to S7) representing the electrodes tracking the audio envelope. Such electrodes are evidenced as colored dots: for the LF band in cyan and for the HF band in magenta color. Note that many electrodes track both bands.



Supplementary Figure 7

Tracking correlograms (partial correlation) for the AUDIO envelope. The x axis represents envelopes' lags with respect to the speech onset (scale from -1 to 1 second lag). The y-axis indicates frequencies (2-250 Hz). Significant clusters are colored and non-significant areas are masked by a higher opacity. The three plots indicate (left to right) AVstatic trials' average, AVdyn trials average, difference AVstatic-AVdyn.

From top to bottom: 1) temporal electrode for subject 3 WITHOUT CAR referencing, 2) same electrode WITH CAR referencing, 3) temporal electrode of subject 4 WITHOUT CAR referencing, 4) same electrode WITH CAR referencing. The arrows indicate the maximum value of correlation.

SUPPLEMENTARY TABLE

| Occipital Audio tracking (AVdyn/AVstatic) | Also Video tracking | Higher visual areas (MIT, PIT) | Also Video tracking |
|---|---------------------|--------------------------------|---------------------|
| 16 (s3,s4,s5,s6) | 2 (s4,s6) | 8 (s1,s3,s4,s6) | 1 (s3) |

Supplementary table 1. Number of visual cortex electrodes (and subject number in parentheses) that present audio envelope tracking in both the AVstatic and the AVdyn conditions. The leftmost column indicates how many electrodes track audio envelopes in cortical regions belonging to the occipital lobe. Note that the tracking has been observed both in AV static and in AVdyn conditions. The second column indicates how many out of the previous set of electrodes also track the visual envelope. The third column indicates the number of electrodes concurrently tracking audio intensity for higher visual areas (such as middle and posterior IT). Finally, the fourth column reports that one electrode belonging to the previous set also tracks the video feature.