nature human behaviour

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41562-019-0548-z

In the format provided by the authors and unedited.

Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries

Adeen Flinker[1,2]*, Werner K. Doyle[3], Ashesh D. Mehta[4], Orrin Devinsky[2] and David Poeppel[1,5]

[1]Department of Psychology, New York University, New York, NY, USA. [2]Department of Neurology, New York University School of Medicine, New York, NY, USA. [3]Department of Neurosurgery, New York University School of Medicine, New York, NY, USA. [4]Department of Neurosurgery, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Manhasset, NY, USA. [5]Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany.
*e-mail: adeen.f@gmail.com

nature human behaviour

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41562-019-0548-z

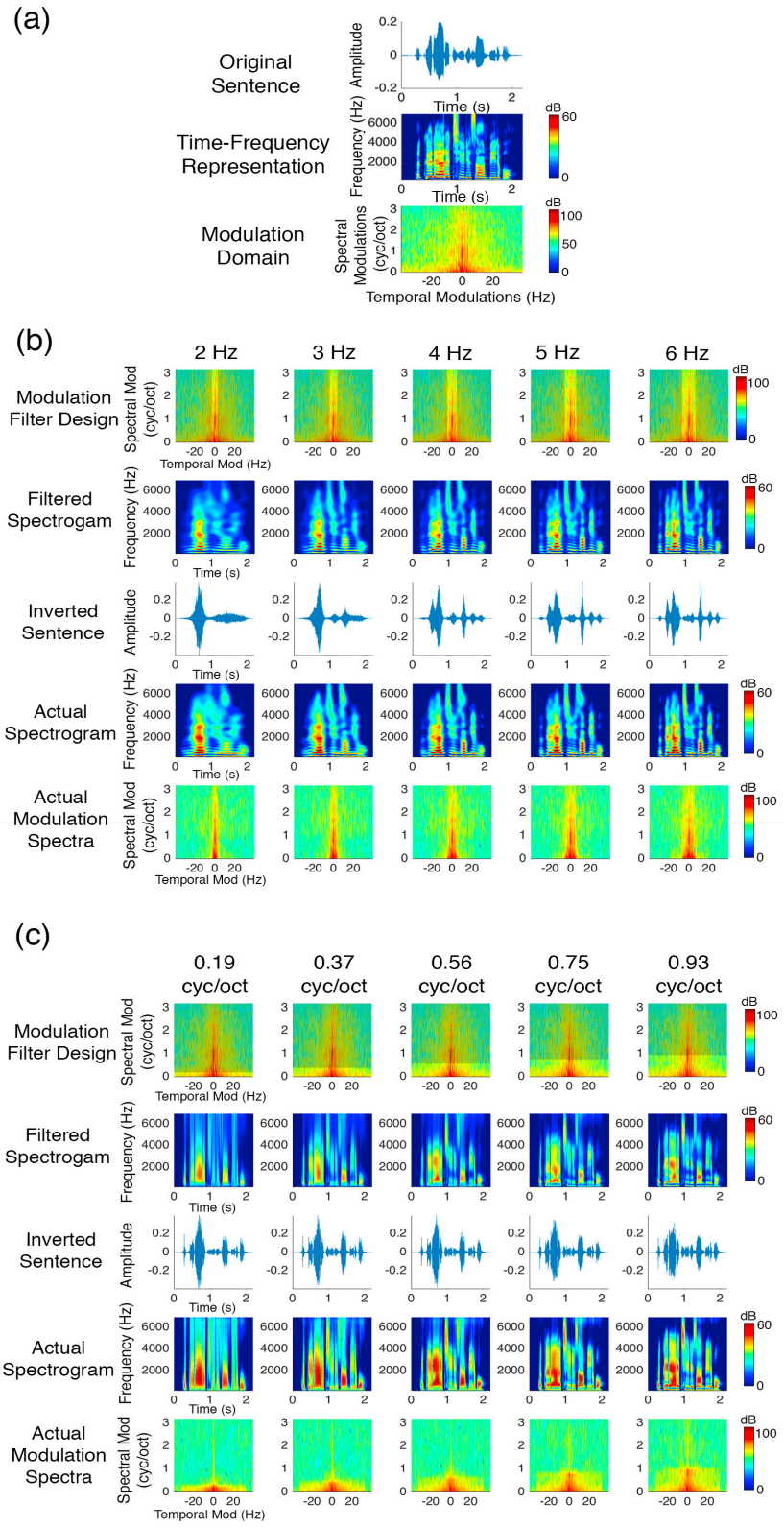
In the format provided by the authors and unedited.

Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries

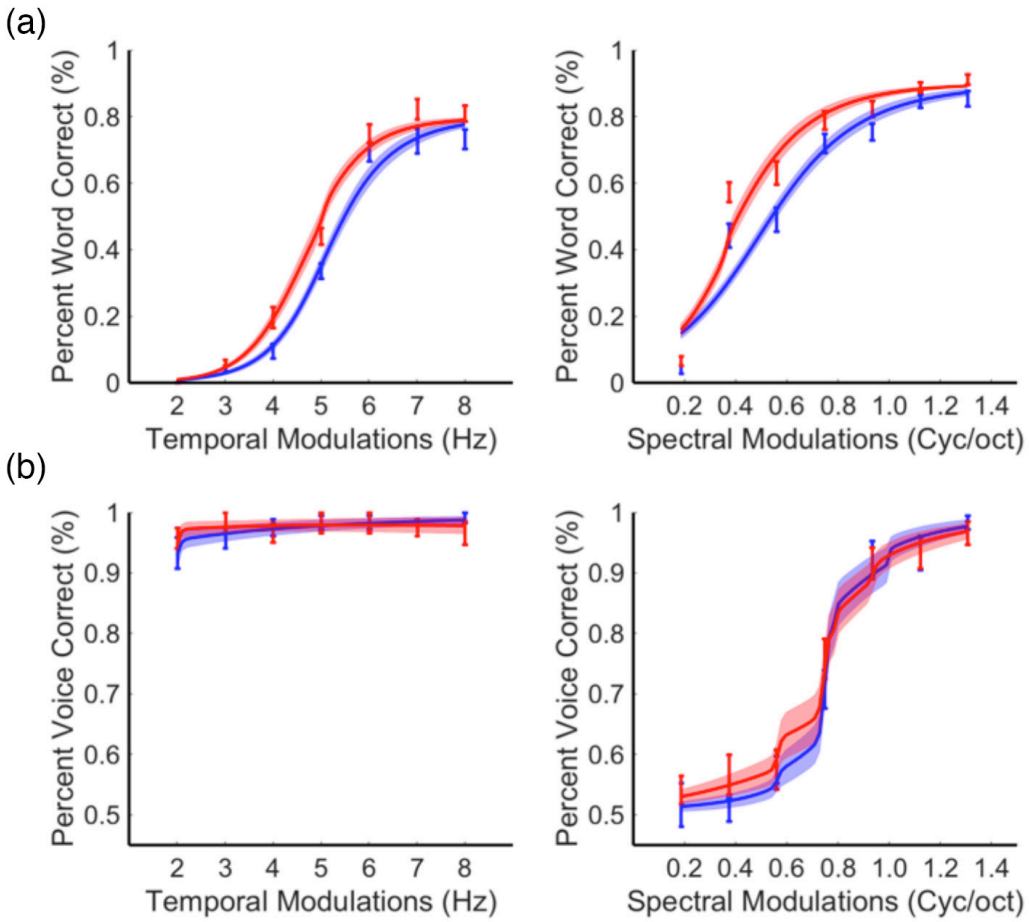
Adeen Flinker[1,2]*, Werner K. Doyle[3], Ashesh D. Mehta[4], Orrin Devinsky[2] and David Poeppel[1,5]

[1]Department of Psychology, New York University, New York, NY, USA. [2]Department of Neurology, New York University School of Medicine, New York, NY, USA. [3]Department of Neurosurgery, New York University School of Medicine, New York, NY, USA. [4]Department of Neurosurgery, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Manhasset, NY, USA. [5]Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany.
*e-mail: adeen.f@gmail.com

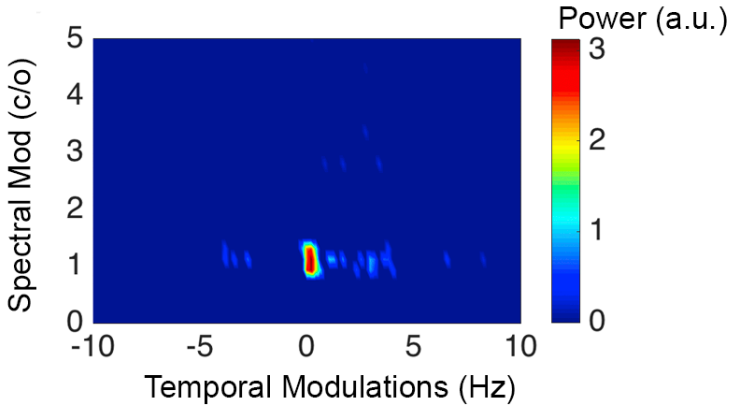
Supplementary Figures



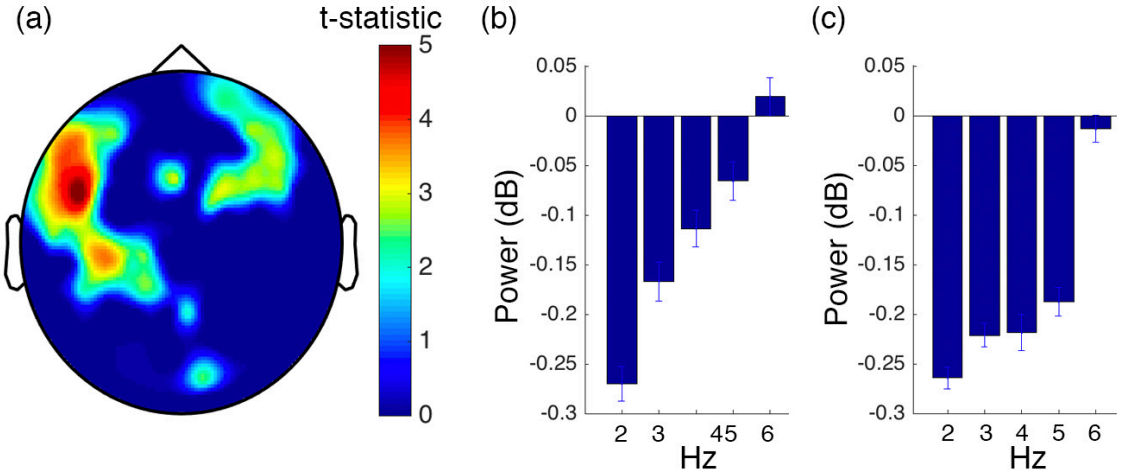
Supplementary Figure 1
Overview of the filtering approach with the original sentence represented in the time-frequency and modulation domains (a) and then filtered according to the different temporal modulation cutoffs (b) and spectral modulation cutoffs (c). Each original sentence produces one new waveform per cutoff with a modulation filter design shown on the top row of panels (b) and (c) followed from top to bottom by: the produced spectrogram using the filter design, the produced waveform following the convex projection iterative procedure, the actual spectrogram of the new waveform and finally the actual modulation spectra. The original and new filtered waveforms for all cutoffs are available as audio file examples Supplementary Audio 1-11. The actual modulation spectra closely match the modulation filter design, which is a prerequisite of any modulation-domain technique that satisfies the Ghitza projection criteria.



Supplementary Figure 2
Full behavioral results for Experiment 1, in which participants (N=20) transcribed, rated and identified voice pitch (male or female) for all stimuli. (a) Percent words correct in the transcriptions of sentences filtered with temporal cutoffs (left) and spectral cutoffs (right) for the first block presented (blue) and the second (red). (b) Percent correct voice pitch identification responses in sentences filtered with temporal cutoffs (left) and spectral cutoffs (right) for the first block presented (blue) and the second (red) for the same participants (N=20). Raw data is shown for each modulation value with mean and standard error of the mean across participants depicted in the error bars. Additionally the within-subject curve fits are shown with a solid curve depicting the mean and standard error of the mean across in the shaded area.

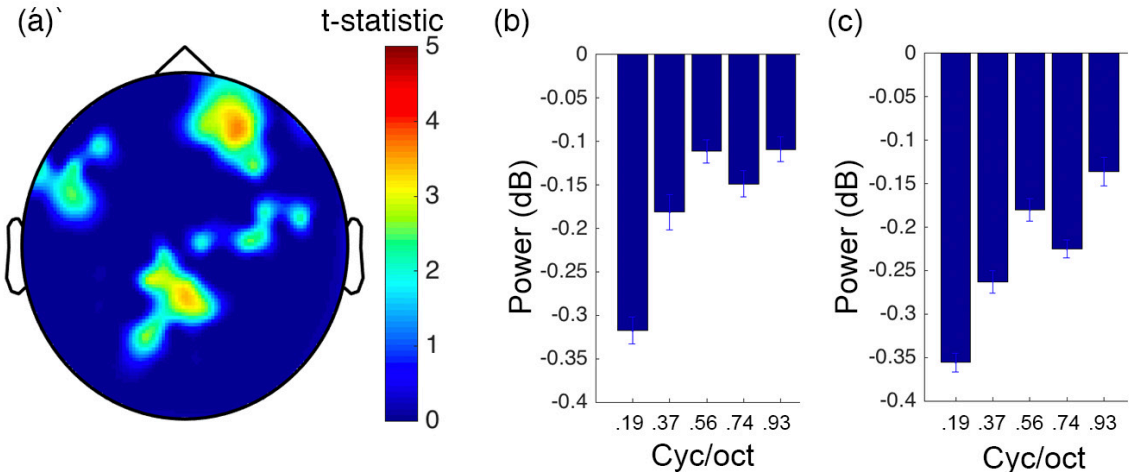


Supplementary Figure 3
Statistically significant differences in modulation spectra between the male and female speaker sentences heard by participants in all Voice Pitch Identification experiments.



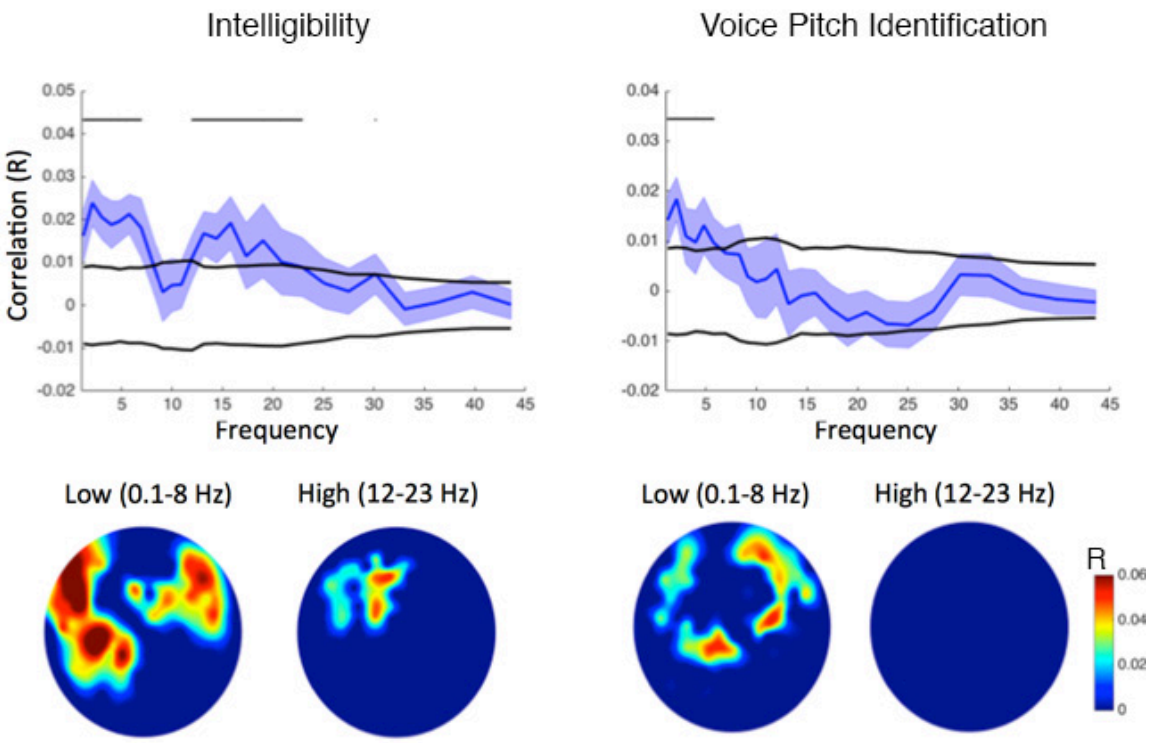
Supplementary Figure 4

(a) Significant power differences between the lowest (2 Hz) and highest (6 Hz) cutoff in Experiment 3 (MEG) computed with a t-test and quantified as a t-statistic. (b) Average power estimates across left sensors (N=76) for each temporal cutoff. (c) Average power estimates across right sensors (N=81) for each temporal cutoff. Error bars denote standard error of the mean across sensors.

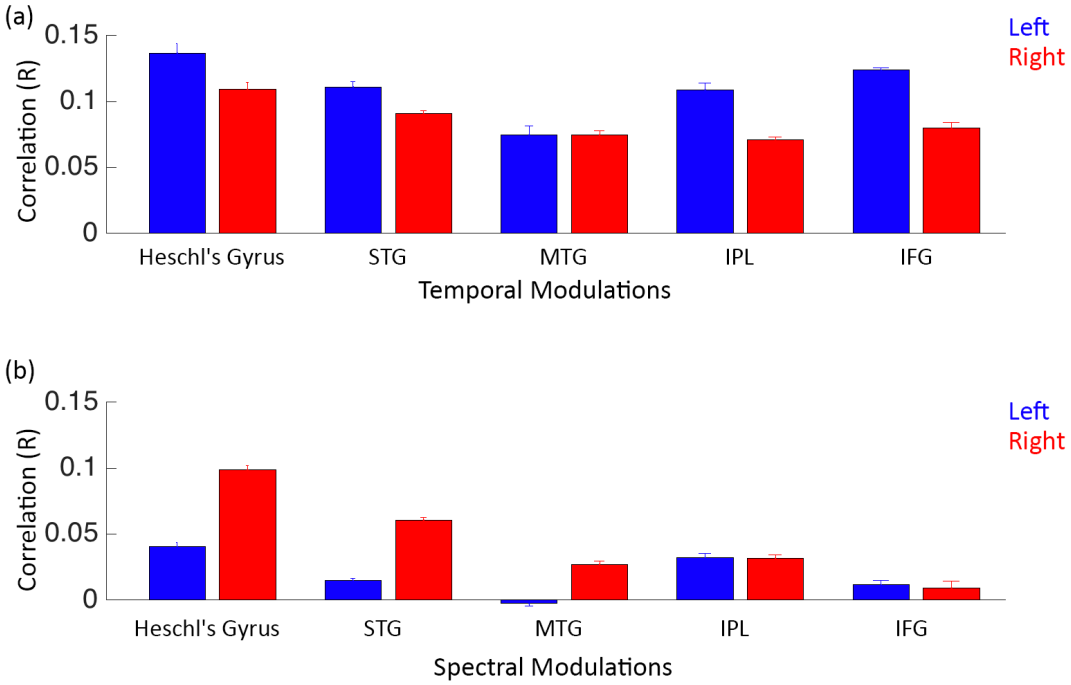


Supplementary Figure 5

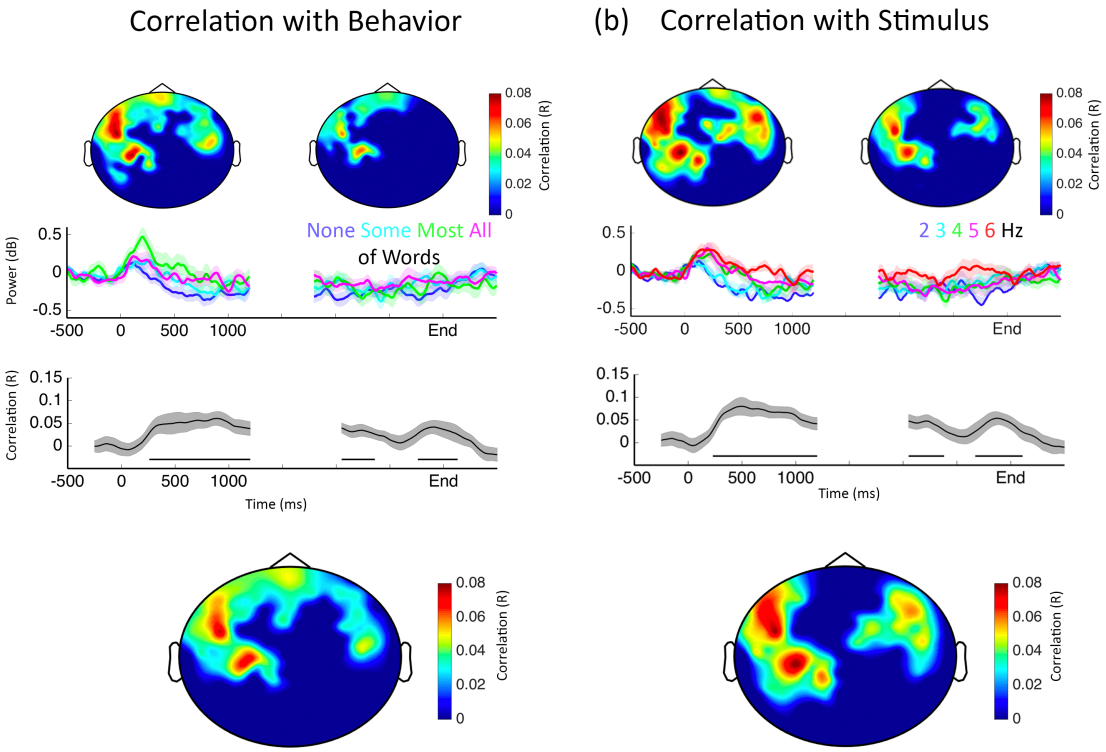
(a) Significant power differences between the lowest (0.1867 cyc/oct) and highest (0.9333 cyc/oct) cutoff in Experiment 3 (MEG) computed with a t-test and quantified as a t-statistic. (b) Average power estimates across left sensors (N=76) for each spectral cutoff. (c) Average power estimates across right sensors (N=81) for each spectral cutoff. Error bars denote standard error of the mean across sensors.



Supplementary Figure 6
MEG mean correlations across the frequency spectrum (shaded error bars denote standard error of the mean across participants, N=19). Average correlations are shown for both tasks as a function of neural frequency power. Permutation based significance thresholds (95% Confidence Interval of the null hypothesis statistic) are shown in the black lines, and top horizontal straight lines denote frequency regions exceeding significance. Correlation topographies for the two frequency ranges of interest are depicted in the bottom panel for each task.

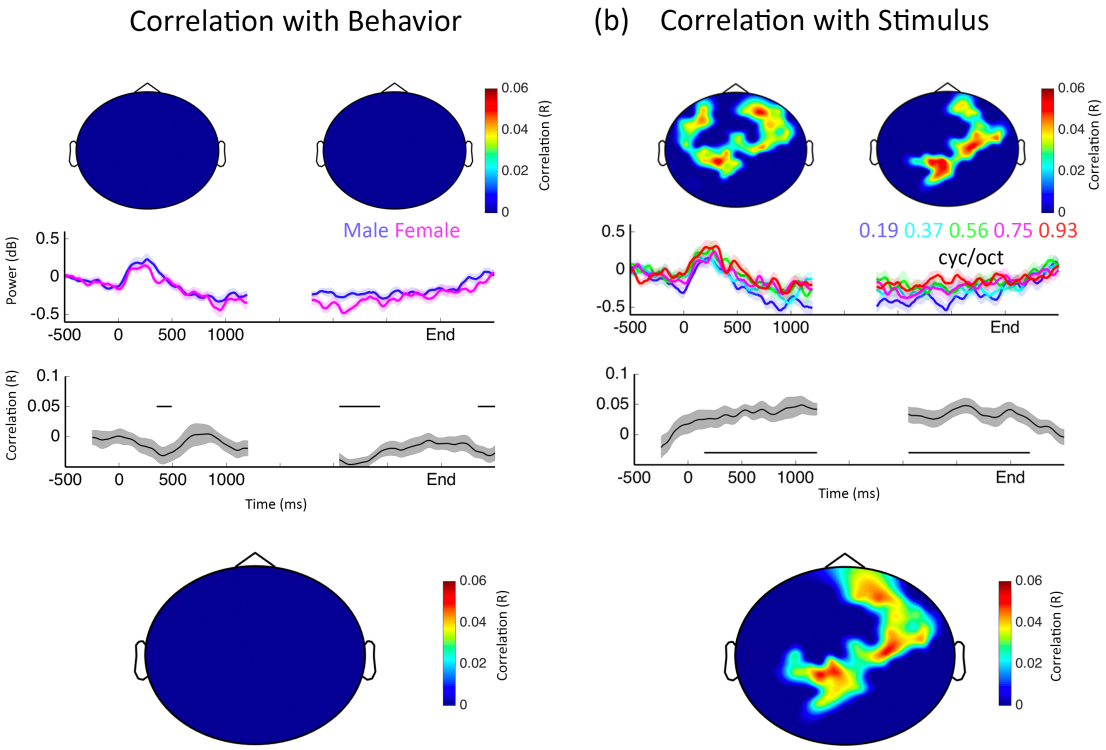


Supplementary Figure 7
Average correlation values of source space results in Experiment 3 (MEG) for selected anatomical ROIs with left hemisphere structures in blue and right hemisphere structures in red. The error bars reflect standard error of the mean across source space voxels. A statistical comparison using a non-parametric permutation test (shuffling labels) across participants (N=19) for left and right power shows significant results as follows for Temporal Modulations in Heschl's Gyrus ($P_{differenc\_of\_mean\_R}$ = 0.039545, p=0.018000, CI=[-0.004113 0.000000]), STG ($P_{differenc\_of\_mean\_R}$ = 0.027690, p=0.024000, CI=[-0.009962 0.012655]), MTG ($P_{differenc\_of\_mean\_R}$ = 0.017936, p=0.028000, CI=[-0.010348 0.010315]), IPL ($P_{differenc\_of\_mean\_R}$ =0.068682, p=0.003000, CI=[-0.013067 0.015313]), IFG ($P_{differenc\_of\_mean\_R}$ = 0.071809, p=0.004000, CI=[-0.010108 0.010523]) and for Spectral Modulations in Heschl's Gyrus ($P_{differenc\_of\_mean\_R}$ = -0.058460, p=0.017000, CI=[-0.003926 0.004063]), STG ($P_{differenc\_of\_mean\_R}$ = -0.045938, p=0.013000, CI=[-0.012126 0.012375]), MTG ($P_{differenc\_of\_mean\_R}$ = -0.029701, p=0.013000, CI=[-0.012894 0.013945]), IPL ($P_{differenc\_of\_mean\_R}$ = 0.000445, p=0.182000, CI=[-0.012774 0.011546]), IFG ($P_{differenc\_of\_mean\_R}$ = -0.016789. p=0.033000, CI=[-0.009785 0.009314]).



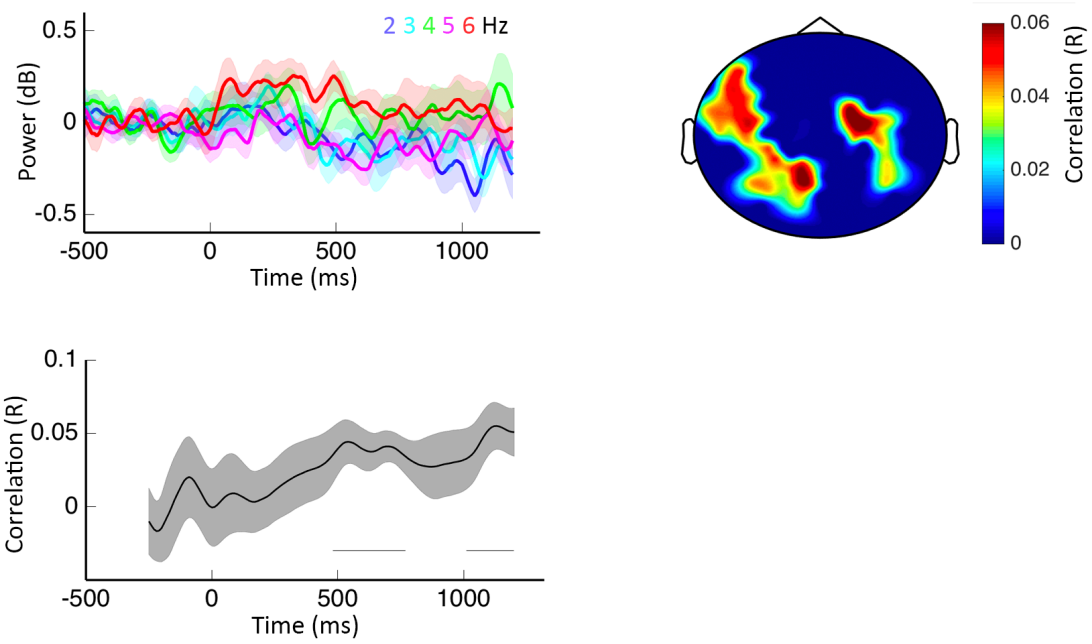
Supplementary Figure 8

Comparison of correlating power estimates in Experiment 3 (MEG) with behavior (actual participant response, N=19) compared with the stimulus *temporal modulation* cutoff. (a) Correlations with participant's responses (button press corresponding to comprehending None, Some, Most or All of the words). Top row topographies show correlations with the first half of the sentence and last half of the sentence, with the mean power and correlations shown beneath (shaded plots reflect standard error of the mean across participants, N=19). Bottom topography represents correlations with power collapsed across the entire sentence. (b) Correlations with stimulus cutoff values. Top row topographies show correlations with the first half of the sentence and last half of the sentence, with the mean power and correlations shown beneath (shaded plots reflect standard error of the mean across participants, N=19). Bottom topography represents correlations with power collapsed across the entire sentence.

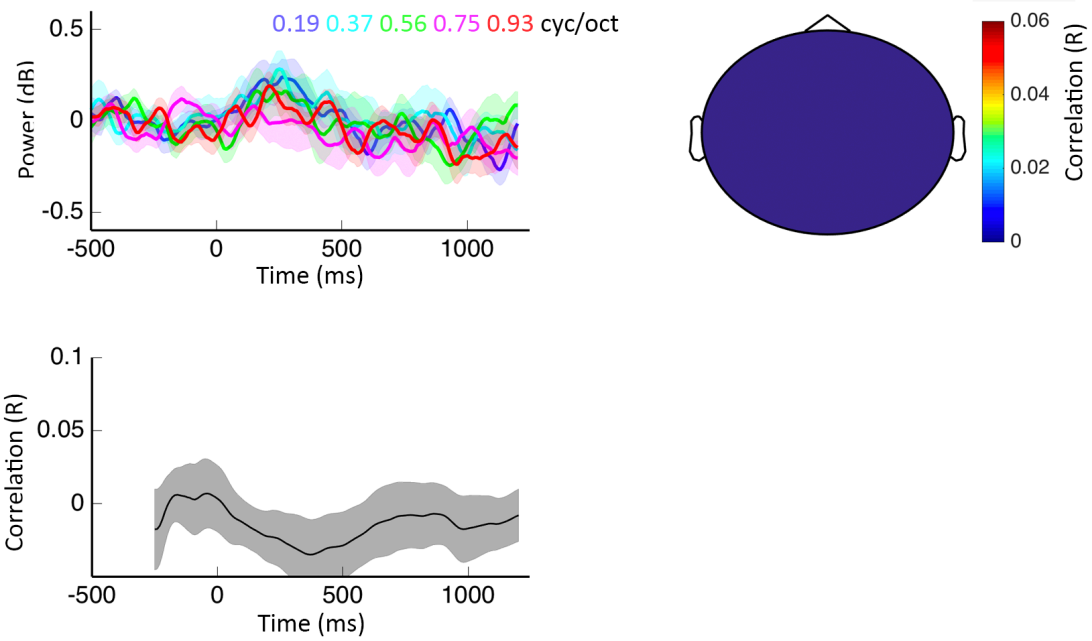


Supplementary Figure 9

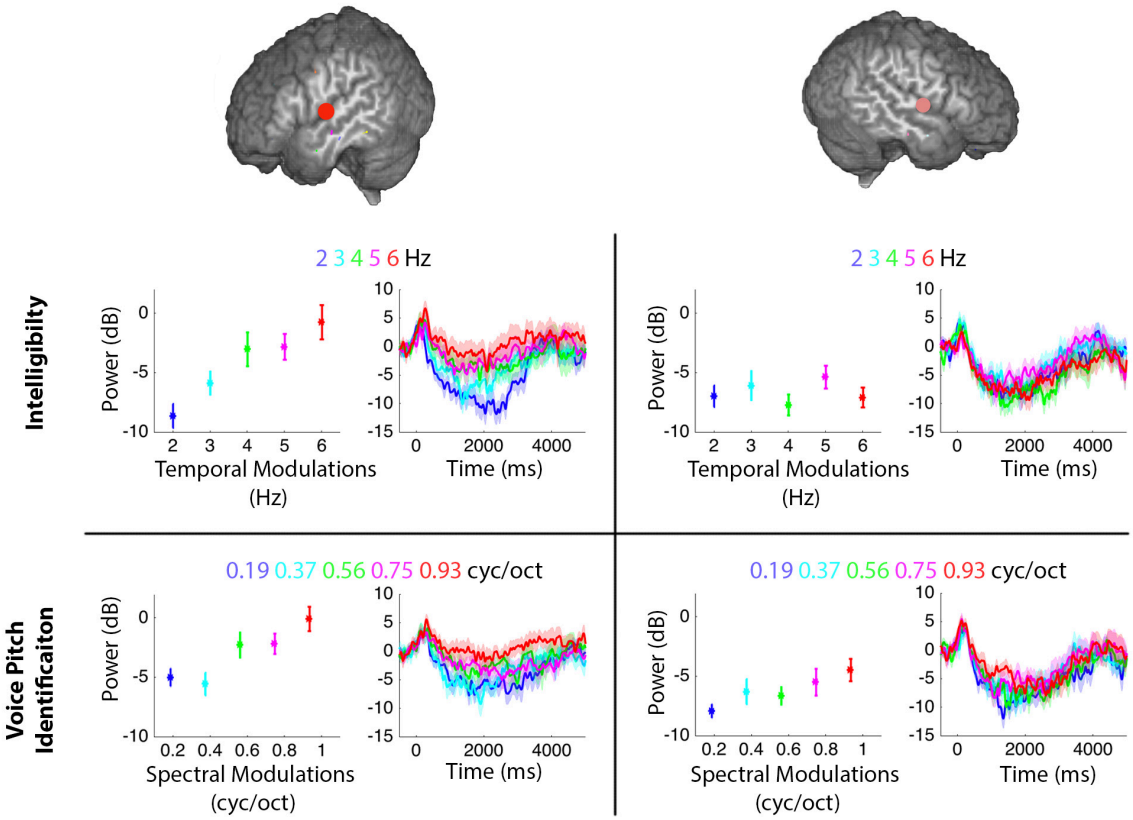
Comparison of correlating power estimates in Experiment 3 (MEG) with behavior (actual participant response, N=19) compared with the stimulus spectral modulation cutoff. (a) Correlations with participant's responses (button press corresponding to male or female speaker). Top topographies show correlations (none passed significance) with the first half of the sentence and last half of the sentence, with the mean power and correlations shown beneath (shaded plots reflect standard error of the mean across participants, N=19). Bottom topography represents correlations (none passed significance) with power collapsed across the entire sentence. (b) Correlations with stimulus cutoffs. Top topographies show correlations with the first half of the sentence and last half of the sentence with the mean power and correlations shown beneath (shaded plots reflect standard error of the mean across participants, N=19). Bottom topography represents correlations with power collapsed across the entire sentence.



Supplementary Figure 10
Analysis results of Experiment 4 (MEG), using reversed speech stimuli filtered with temporal modulation cutoffs (N=10). On the left, the low frequency power (0.1-8 Hz) is plotted for each cutoff (top) as well as average correlations across sensors (bottom). Shaded plots reflect standard error of the mean across participants (N=10). On the right, the topography of significant correlations is shown.

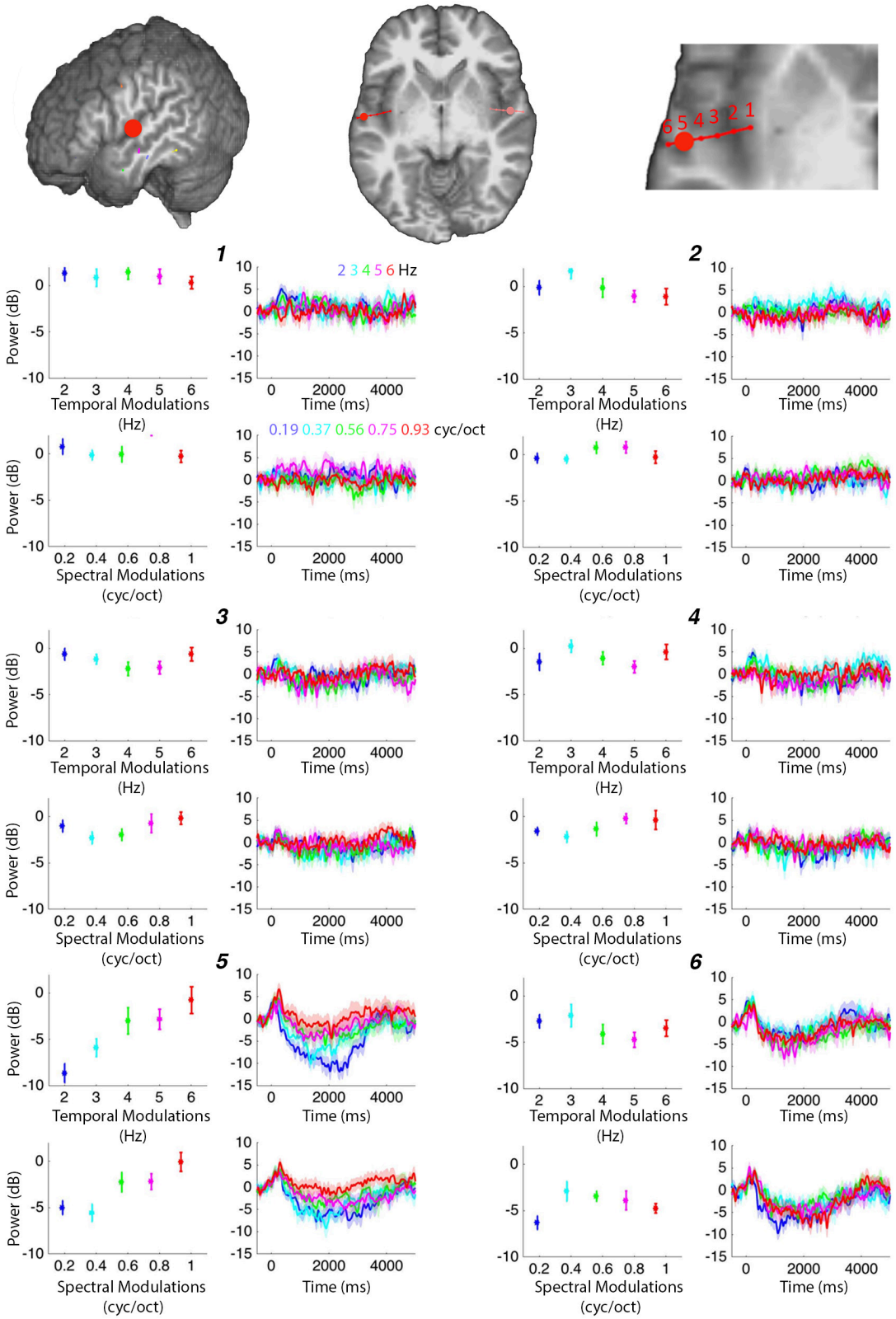


Supplementary Figure 11
Analysis results of Experiment 4 (MEG, N=10) reversed speech stimuli filtered with spectral modulation cutoffs. On the left the low frequency power (0.1-8 Hz) is plotted for each cutoff (top) as well as average correlations across sensors (bottom). Shaded plots reflect standard error of the mean across participants (N=10). On the right the topography of correlations is shown, none passed significance.



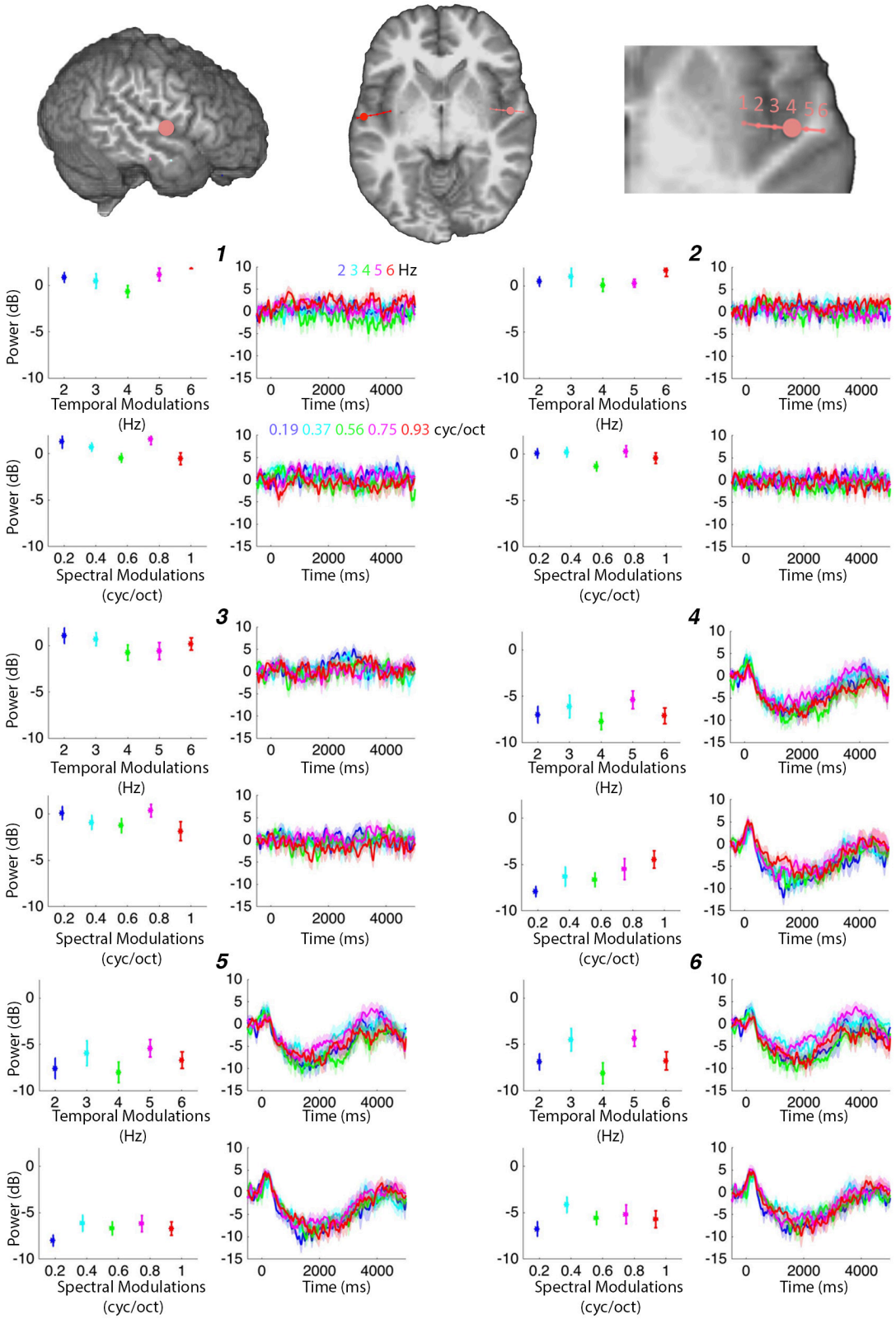
Supplementary Figure 12

Intracranial ECoG recordings in a patient with bilateral stereotactic depth electrodes (MNI coordinates = [-60.17, -7.76, 2.82] , [52.81, -7.95, 2.64]) sampling superior temporal cortices (same as Figure 6 in the main text). Electrode locations (top) are shown for two electrodes with low frequency (0.1-8 Hz) power results shown for each task (top and bottom rows) and electrode (left and right columns). For each electrode and task the mean power and standard error of the mean (SEM, across trials, N=16) collapsed across time is shown (left) as well as mean and shaded SEM (across trials, N=16) power traces for each modulation filter (right). The left electrode showed a significant correlation with both temporal modulations (permutation test, $P_{correlation}$=0.413, p<0.001, CI =[-0.173279 0.196382], R^2=0.1708) as well spectral modulations (R=0.381 p=<0.001, CI=[-0.182786 0.192845], R^2= 0.1449). The right electrode showed a significant correlation with spectral modulations (permutation test, $P_{correlation}$=0.241, p=0.015, CI =[-0.171243 0.193584], R^2= 0.05831) but not with temporal modulations (R=0.015, p=0.441 permutation test, $P_{correlation}$=0.015185 p=0.432, CI=[-0.189196 0.200897], R^2= 0.000231). The relative difference in correlation values for the left electrode is 0.032 and for the right electrode is 0.226.



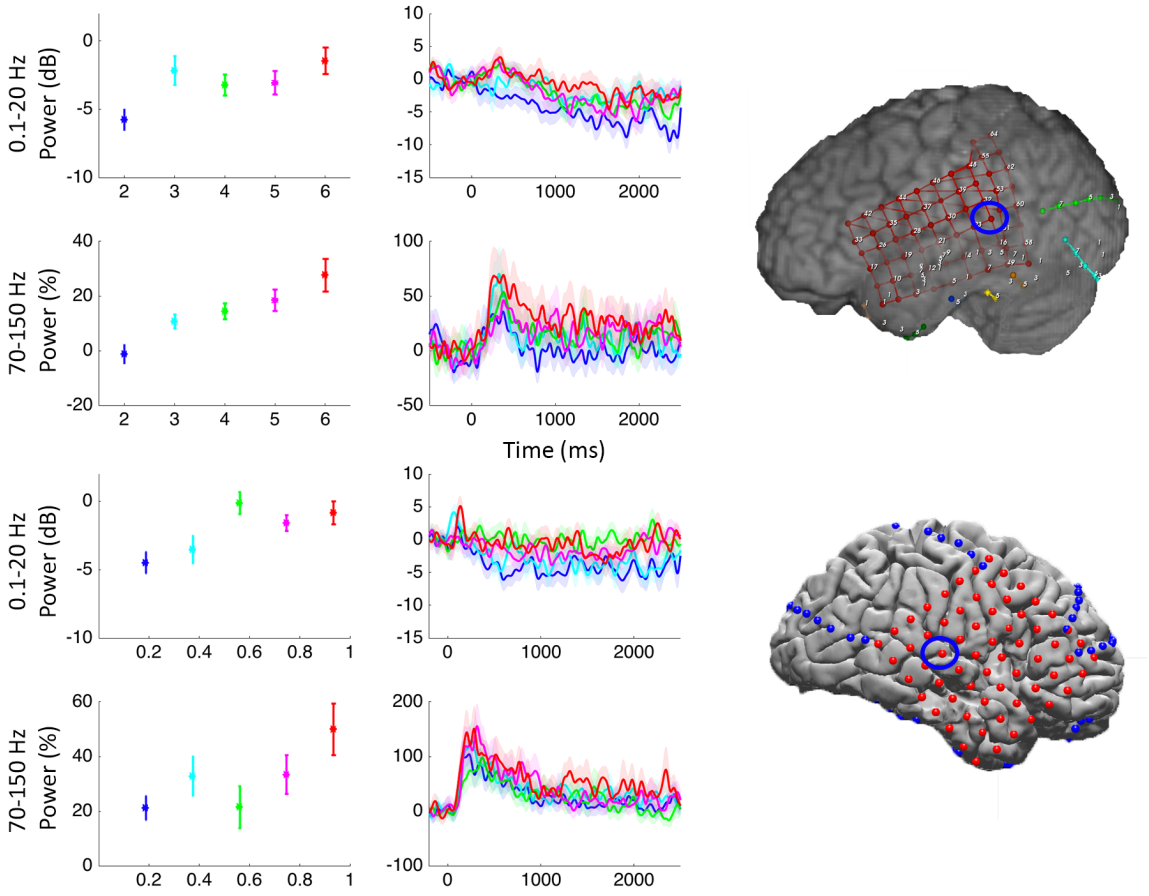
Supplementary Figure 13

Intracranial ECoG recordings for all left depth contact in a patient with bilateral sterotactic depth electrodes sampling superior temporal cortices (same as Figure S12 and Figure 6, main text). Each contact is numbered from 1-6 (MNI coordinates = [-39.68,-4.58,2.92],[-44.89,-5.49,2.88],[-49.93,-6.56,2.63],[-55.05,-7.16,2.73],[-60.17,-7.77,2.83],[-65.30,-8.56,2.80]) and neurophysiology is depicted for each electrode as well as corresponding anatomy in the top panel (axial slice Z=3). Each contact is shown as in Supplementary Figure 12: the top graphs depict temporal modulations and the bottom depicts spectral modulations. For each electrode and task the average mean power and SEM (across trials, N=16) collapsed across time is shown (left) as well as mean and shaded SEM (across trials, N=16) power traces for each modulation filter (right). For each panel the correlation and p value (permutation test) as well as a test if total average power (all trials collapsed across time) was significantly different from zero (ttest p value) is listed in Supplementary Table 1.

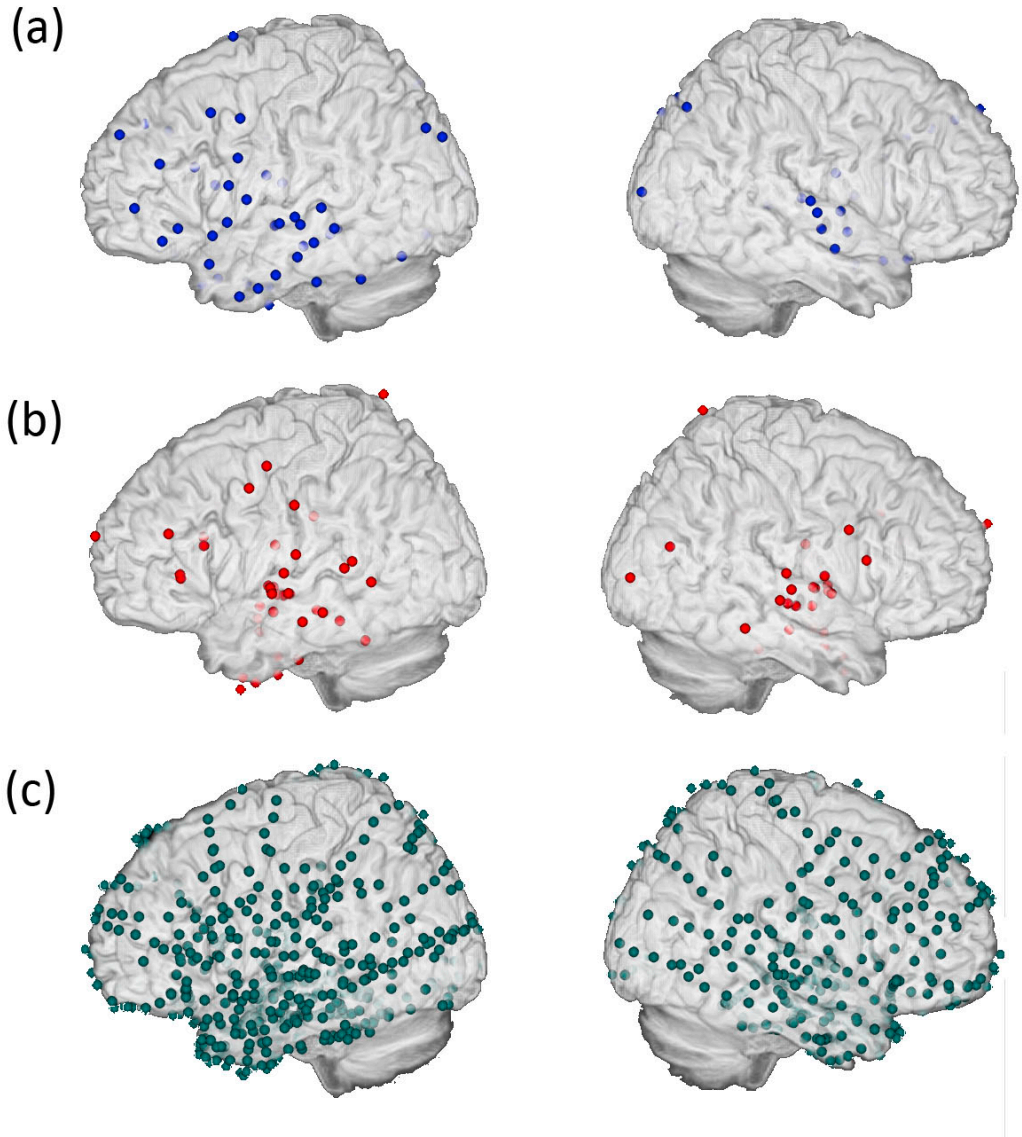


Supplementary Figure 14

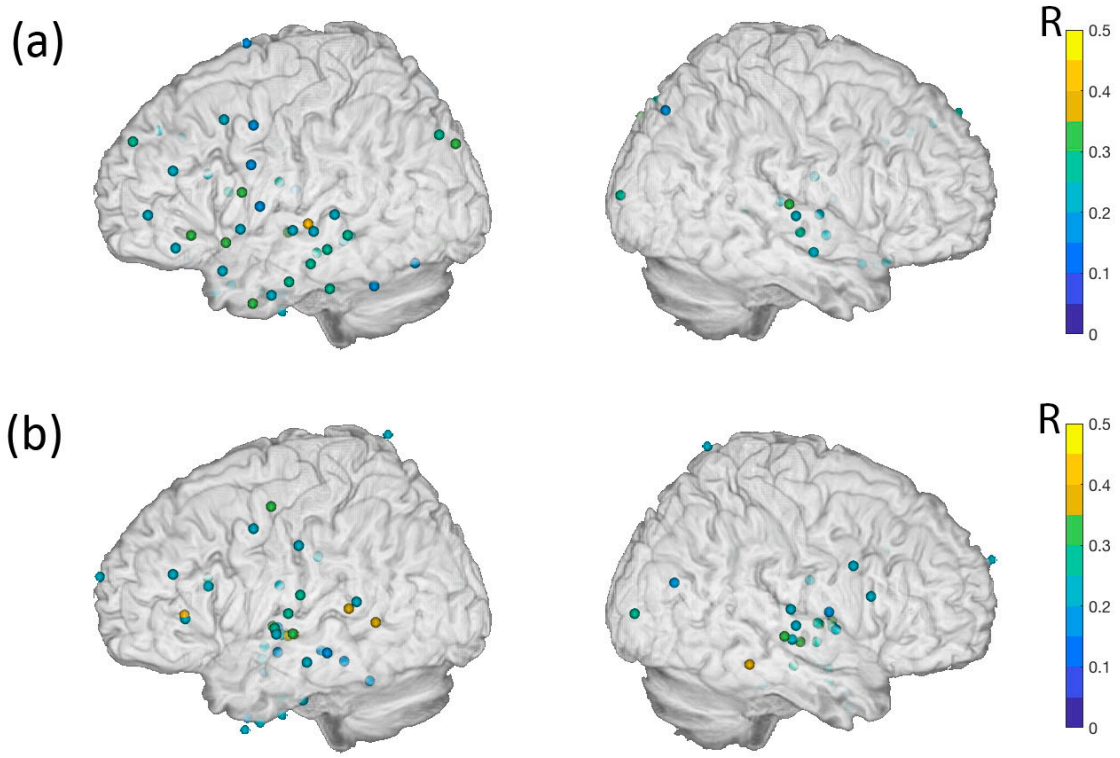
Intracranial ECoG recordings for all right depth contact in a patient with bilateral sterotactic depth electrodes sampling superior temporal cortices (same as Supplementary Figure 12 and Figure 6). Each contact is numbered from 1-6 (MNI coordinates = [37.98,-5.58,-1.21],[42.51,-6.54,0.24],[47.70,-7.22,1.58],[52.81,-7.95,2.64],[57.68,-8.43,3.46],[62.95,-9.19,4.88]) and neurophysiology is depicted for each electrode as well as corresponding anatomy in the top panel (axial slice Z=3). Each contact is shown as in Figure S12 and S13: the top graphs depict temporal modulations and the bottom depicts spectral modulations. For each electrode and task the average mean power and SEM collapsed across time is shown (left) as well as mean and shaded SEM power traces for each modulation filter (right). For each panel the correlation and p value (permutation test) as well as whether the average power (all trials collapsed across time) was significantly different from zero (ttest p value) is listed in Supplementary Table 1.



Supplementary Figure 15
Two representative patients with the most significant electrodes over the lateral surface of STG, showing both low frequency and high gamma power ranges. Top panel depicts average low frequency power in patient S2 (left) as well as the time course of power locked to hearing sentences (middle), color coded by temporal cutoffs (blue, cyan, green, red, magenta corresponding to 2,3,4,5,6 Hz). Similarly, power traces in the high gamma (70-150) range are depicted locked to hearing sentences. Third and fourth row panels show low frequency and high gamma power in patient S5 for spectral modulation cutoffs. Error bars and shaded error bars represent SEM across trials (N=16).



Supplementary Figure 16
Significant electrodes across patients in MNI coordinates for the intelligibility task (blue), the gender voice identification task (red) and all implanted electrodes (green). (a) Significant electrodes correlating with temporal modulation cutoffs ($p<0.05$, permutation test). (b) Significant electrodes correlating with spectral modulation cutoffs ($p<0.05$, permutation test). (c) Locations of all implanted electrodes. Across patients, there were significant correlations with temporal modulations in 8.042% of total left electrodes and 3.218% of total right electrodes producing a laterality index of 0.428. Correlations with spectral modulations were evident in 6.818% of total left electrodes and 6.188% of total right producing a laterality index of 0.048.



Supplementary Figure 17
Significant electrodes across patients in MNI coordinates with correlation values color coded. (a) Significant electrode correlations with temporal modulation cutoffs ($p<0.05$, permutation test). (b) Significant electrodes correlations with spectral modulation cutoffs ($p<0.05$, permutation test).

Supplementary Tables

Supplementary Table 1
For each electrode (related to Figure 6 and Supplementary Figures 12-14) the correlation and p value (permutation test) as well as a test if total average power (all trials collapsed across time) was significantly different from zero (ttest p value) is listed in Supplementary Table 1. Rows in bold represent the electrodes highlighted in Figure 6.

Electrode	Temporal modulations			Spectral modulations		
	R	p-value	ttest p-value	R	p-value	ttest p-value
Left 1	-0.072	0.251	0.024	0.021	0.417	0.148
Left 2	0.162	0.073	0.762	0.076	0.238	0.751
Left 3	-0.038	0.360	0.000	0.127	0. 131	0.004
Left 4	0.003	0.486	0.028	0.181	0.056	0.005
Left 5	**0.413,**	**0.000**	**0.000**	**0.381**	**0.001**	**0.000**
Left 6	0.125	0.146	0.000	0.074	0.247	0.000
Right 1	0.162	0.078	0.021	-0.134	0.105	0.114
Right 2	0.067	0.301	0.070	0.048	0.317	0.473
Right 3	-0.112	0.150	0.765	0.093	0.220	0.116
Right 4	**0.015**	**0.441**	**0.000**	**0.241**	**0.014**	**0.000**
Right 5	0.062	0.290	0.000	0.095	0.193	0.000
Right 6	0.009	0.471	0.000	0.034	0.409	0.000