Supplementary Figure 1. Differential waveforms of the ERPs for the passive oddball task for the speech (solid line) and the rotated speech (dashed line) calculated by subtracting deviant ERPs from control ERPs across all the electrode sites.
Supplementary Figure 2. Standard and deviant ERPs for the active oddball task for the speech (solid line) and the rotated speech (dashed line) across all the electrode sites.
Supplementary Figure 3. Topographies show the spatial distribution of beta ERDs/ERSs averaged in contiguous 100 ms time windows, characterizing the ERSPs for Control and Deviant events (1st, 2nd, 4th, 5th row) as well as the difference between them (3rd, 6th row) for the Rotated Speech (1st-3rd row) and the Speech condition (4th-6th row) for the passive oddball task.
Supplementary Figure 4. Topographies show the spatial distribution of theta ERDs/ERSs averaged in contiguous 100 ms time windows, characterizing the ERSPs for Standard and Deviant events (1\textsuperscript{st}, 2\textsuperscript{nd}, 4\textsuperscript{th}, 5\textsuperscript{th} row) as well as the difference between them (3\textsuperscript{rd}, 6\textsuperscript{th} row) for the Rotated Speech (1\textsuperscript{st}-3\textsuperscript{rd} row) and the Speech condition (4\textsuperscript{th}-6\textsuperscript{th} row) for the active oddball task.
Supplementary Figure 5. Topographies show the spatial distribution of beta ERDs/ERSs averaged in contiguous 100 ms time windows, characterizing the ERSPs for Standard and Deviant events (1st, 2nd, 4th, 5th row) as well as the difference between them (3rd, 6th row) for the Rotated Speech (1st-3rd row) and the Speech condition (4th-6th row) for the active oddball task.
Supplementary Table 1. *F3, F4, HNR, Shimmer, Jitter values of the experimental stimuli for each talker and each condition.*

| Talker | Vowel | Condition | Speech | Rotated Speech |
|--------|-------|-----------|--------|----------------|
| Male   | a     | F3        | 2447 Hz| 2764 Hz        |
|        |       | F4        | 3223 Hz| 3176 Hz        |
|        |       | HNR       | 7.56 dB| 5.50 dB        |
|        |       | Shimmer   | 0.17   | 0.16           |
|        |       | Jitter    | 0.0031 | 0.0023         |
|        | e     | F3        | 2643 Hz| 1848 Hz        |
|        |       | F4        | 3334 Hz| 3581 Hz        |
|        |       | HNR       | 5.37 dB| 2.43 dB        |
|        |       | Shimmer   | 0.12   | 0.16           |
|        |       | Jitter    | 0.0031 | 0.0530         |
|        | i     | F3        | 2713 Hz| 2031 Hz        |
|        |       | F4        | 3295 Hz| 3619 Hz        |
|        |       | HNR       | 9.12 dB| 6.20 dB        |
|        |       | Shimmer   | 0.14   | 0.15           |
|        |       | Jitter    | 0.0016 | 0.0023         |
|        | ɔ     | F3        | 2861 Hz| 3133 Hz        |
|        |       | F4        | 3294 Hz| 3410 Hz        |
|        |       | HNR       | 7.63 dB| 6.52 dB        |
|        |       | Shimmer   | 0.17   | 0.17           |
|        |       | Jitter    | 0.0021 | 0.0020         |
|        | ε     | F3        | 2268 Hz| 2193 Hz        |
|        |       | F4        | 3037 Hz| 3396 Hz        |
|        |       | HNR       | 8.22 dB| 5.00 dB        |
|        |       | Shimmer   | 0.26   | 0.26           |
|        |       | Jitter    | 0.0018 | 0.0018         |
| Female | a     | F3        | 2019 Hz| 2581 Hz        |
|        |       | F4        | 2726 Hz| 3055 Hz        |
|        |       | HNR       | 11.83 dB| 7.71 dB       |
|        |       | Shimmer   | 0.18   | 0.18           |
|        |       | Jitter    | 0.0035 | 0.0040         |
|        | e     | F3        | 2709 Hz| 2790 Hz        |
|        |       | F4        | 3247 Hz| 3628 Hz        |
|        |       | HNR       | 13.73 dB| 6.01 dB       |
|        |       | Shimmer   | 0.14   | 0.09           |
|        |       | Jitter    | 0.0017 | 0.0421         |
|        | i     | F3        | 2925 Hz| 3138 Hz        |
|        |       | F4        | 3182 Hz| 3645 Hz        |
|        |       | HNR       | 11.25 dB| 5.92 dB       |
|        |       | Shimmer   | 0.17   | 0.25           |
|        |       | Jitter    | 0.003 | 0.0202         |
|        | ɔ     | F3        | 2202 Hz| 2876 Hz        |
|        |       | F4        | 3013 Hz| 3295 Hz        |
|        |       | HNR       | 10.73 dB| 6.95 dB       |
|        |       | Shimmer   | 0.15   | 0.14           |
|        |       | Jitter    | 0.0035 | 0.0091         |
|        | ε     | F3        | 2356 Hz| 2564 Hz        |
|        |       | F4        | 3067 Hz| 3326 Hz        |
|        |       | HNR       | 9.61 dB| 2.50 dB        |
|        |       | Shimmer   | 0.13   | 0.17           |
|        |       | Jitter    | 0.0021 | 0.0056         |
Supplementary Analysis of MMN latency

We measured the peak latency values of the most negative peak between 110 and 225 ms (Kappenman et al., 2021) of each participant across a subset of channels that were included in the clusters capturing the difference between standard and deviant events for both the Speech and the Rotated Speech conditions (C1, C2, C3, Cz, F1, F2, F3, F4, F5, F6, Fz, AF3, AF4, FC1, FC2, FC3, FC4, FC5, FC6, FCz). The latency values were then averaged across channels for each participant and separately within each condition. A paired-sample t-test \( t(14) = -0.61, p = 0.54 \) showed no difference between the latency values of the MMN in the Speech condition (M = 182 ms, SD = 29 ms) and the Rotated Speech condition M = 175 ms SD = 28 ms).

Supplementary Analyses of phase-locked and non-phase locked power

The non-phase locked power was calculated by recomputing the TFR across the whole spectrum between 4 and 30 Hz within each combinations of the factors condition (speech, rotated-speech) and probability (standard, deviant) after subtracting the ERP (i.e., the mean across trials for each combination of the factors condition and probability) which represents the phase-locked component of the EEG from all trials in the corresponding design cell (e.g., the ERP elicited by standard events in the speech condition was subtracted from the trials of standard events in the speech condition). Secondly, the phase-locked power was computed by subtracting the non-phase locked power from total power (Cohen, 2014).

Results – Non phase locked activity

In the active oddball task, the test if the interaction between the factors condition and probability within the theta band revealed the presence of a positive cluster (\( p = .017, 95 \% CI [.012 .022], d = 1.12 \)) surfacing between 590 and 800 ms on right Central, Centro-Parietal and Parietal electrodes. Post-hoc tests on standard and deviant events within each condition, showed that deviant events yielded a stronger synchronization in the theta band compared to control events, as highlighted by the presence of a positive cluster both in the speech (\( p < .001, 95 \% CI[.0003 .001], d = 1.33 \)) and the rotated speech condition (\( p < .001, 95 \% CI [.0003 .003], d = 1.20 \)), widely distributed from Pre-Frontal to Parietal electrode sites, in the 150-800 ms and in the 150-660 ms time windows, respectively. Therefore, the effect found for the interaction substantially reflected a stronger theta synchronization occurring in deviant events for the speech condition.

In the beta band, the test on the interaction between the factors condition and probability revealed the presence of a positive cluster (\( p = .016, 95 \% CI [.011 .021], d = 1.22 \)), emerging between 280 and 800 ms across Central, Centro-Parietal and Parietal electrode sites. Post-hoc tests operated between standard and deviant events within each condition showed a desynchronization in deviant with respect to standard events both in the speech (\( p = .011, 95 \% CI [.007 .016], d = 1.38 \)) and the rotated speech condition (\( p = .002, 95 \% CI [.0007 .004], d = 1.24 \)), captured by negative clusters unfolding over Central and Centro-Parietal channels, in the 250-590 ms and in the 260-710 ms time windows, respectively. The speech condition was also
characterized by a stronger beta synchronization for deviant events with respect to standard ones, surfacing right after the earlier-occurring desynchronization and widely distributed on the scalp between 570 and 800 ms ($p = .007, 95\% \text{ CI} [.004 .011], d = 1.142$), which presumably induced the effect highlighted by the test on differential ERSPs (see Supplementary Figure 5). The tests on differential ERSPs did not show significant differences between conditions in the alpha frequency band.

**Results – Phase-locked activity**

In the active oddball task, the test on the interaction between the factors condition and probability did not reveal any difference within the theta, alpha and beta band. Upon further exploration of the differences between the phase-locked TFRs of standard and deviant events within each condition, both conditions revealed a similar pattern. Within the theta band, deviant events showed a stronger synchronization both in the speech ($p = .001, 95\% \text{ CI} [.0002 .003], d = 0.92$) and in the rotated speech ($p = .003, 95\% \text{ CI} [.0003 .001], d = 1.09$) conditions as highlighted by the presence of a positive cluster distributed over Frontal, Fronto-Central and Posterior channels between 100-620 ms and 90-800 ms respectively. Similarly, within the alpha band, deviant events showed a stronger synchronization both in the speech ($p = .018, 95\% \text{ CI} [.013 .024], d = 0.96$) and in the rotated speech ($p = .008, 95\% \text{ CI} [.004 .011], d = 1.11$) conditions as highlighted by the presence of a positive cluster distributed over surfacing over Fronto-Central and Central channels between 70-360 ms and 60-410 ms respectively.

**Discussion**

Statistical analyses of non-phase locked power (performed with the same logic used in the analyses of total power reported in the manuscript) showed no substantial differences with respect to the ones performed on total power (see Supplementary Figure below). Therefore, it is safe to assume that the statistical analyses on total power in the main manuscript mostly reflect non-phase locked oscillatory activity in the analysed frequency bands.

The analyses performed on the phase-locked activity highlighted an initial ERS within the alpha band, unfolding approximately in the 100-400 time window in frontal electrodes, and a subsequent ERS, unfolding approximately in the 100-700 time window in pre-frontal, frontal and parietal electrode sites. For both frequency bands, no differences between the conditions emerged.

First, with respect to alpha ERS, considering its spatiotemporal characteristics, it might be linked to ERP components occurring before the P300 (e.g., N2), which were not under investigation in our study. Second, for what concerns the theta ERS, even though the temporal sensitivity for TFR within this band is low (considering the small number of cycles used to compute the TFRs), the temporal unfolding of the ERS is compatible with the one of P300 found in the ERP analysis. This result is in line with previous studies linking the elicitation of the P300 in the ERP domain to a synchronization within the theta band (Başar-Eroglu et al., 1992; Demiralp et al., 2001; Kolev et al., 1997; Yordanova et al., 2000).
It is important to note that no statistically significant differences pertaining theta ERS emerged in the test for the interaction between the factors condition and probability, even though at a qualitative level the speech condition seemed to elicit a stronger theta ERS. This might appear in contrast with the ERP analyses where the speech condition showed a stronger P3b response. The absence of the effect in the phase-locked oscillatory activity might stem from two different sources. First, we analysed only the portion of the spectrum between 4 and 30 Hz, thus it may be possible that oscillatory activity in other frequency bands might have contributed to the generation of the ERP along with the activity within the theta bands. Second, it is possible that our study lacked in statistical power to highlight a between-condition difference in theta ERS, an aspect that directly relates to the strict statistical framework we employed.
Active Oddball

Total Power

Non Phase-locked Power

Phase-locked Power
**Supplementary Figure 6.** Time-Frequency results for the active oddball tasks of total power (1st row), non-phase-locked power (2nd row) and phase-locked power. The time-frequency power spectra show the power modulations (% change) characterizing the differential ERSPs for each condition (1st and 2nd columns) as well as the difference between them, corresponding to the interaction effect (3rd column). Spectra were obtained by averaging activity for the electrodes F5, F3, F1, Fz, F2, F4, F6, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P5, P3, P1, Pz, P2, P4, P6, PO5, PO3, PO1, POz, PO2, PO4, PO6. Black squares represent the temporal distribution of the significant clusters within theta (4-7 Hz) and beta (13-30 Hz) bands. The mean number of channels included in each cluster represented in the power spectra was calculated across all time-samples and only the time-bins including at least half of the mean number of channels are enclosed in black squares. Topographies in the lower and higher row show the spatial distribution of theta and beta ERDs/ERSs characterizing the differential ERSPs for each condition (1st and 2nd columns) as well as the difference between them, corresponding to the interaction effect (3rd column). Electrodes that were included in the clusters for more than 50% of the samples within the cluster time windows (reported below each topography) are represented by black asterisk marks superimposed to the maps. Black squares on topographies represent the channels that were included in the averaged spectral plots.