Theta and alpha oscillatory signatures of auditory sensory and cognitive loads during complex listening

Brilliant [a,*,1], Y. Yaar-Soffer [b,c,1], C.S. Herrmann [d], Y. Henkin [b,c,1], A. Kral [a,1]

[a] *Department of Experimental Otology, Hannover Medical School, 30625 Hannover, Germany*
[b] *Department of Communication Disorder, Tel Aviv University, 5262657 Tel Aviv, Israel*
[c] *Hearing, Speech and Language Center, Sheba Medical Center, 5265601 Tel Hashomer, Israel*
[d] *Experimental Psychology Division, University of Oldenburg, 26111 Oldenburg, Germany*

ARTICLE INFO

ABSTRACT

The neuronal signatures of sensory and cognitive load provide access to brain activities related to complex listening situations. Sensory and cognitive loads are typically reflected in measures like response time (RT) and event-related potentials (ERPs) components. It's, however, strenuous to distinguish the underlying brain processes solely from these measures. In this study, along with RT- and ERP-analysis, we performed time-frequency analysis and source localization of oscillatory activity in participants performing two different auditory tasks with varying degrees of complexity and related them to sensory and cognitive load. We studied neuronal oscillatory activity in both periods before the behavioral response (pre-response) and after it (post-response). Robust oscillatory activities were found in both periods and were differentially affected by sensory and cognitive load. Oscillatory activity under sensory load was characterized by decrease in pre-response (early) theta activity and increased alpha activity. Oscillatory activity under cognitive load was characterized by increased theta activity, mainly in post-response (late) time. Furthermore, source localization revealed specific brain regions responsible for processing these loads, such as temporal and frontal lobe, cingulate cortex and precuneus. The results provide evidence that in complex listening situations, the brain processes sensory and cognitive loads differently. These neural processes have specific oscillatory signatures and are long lasting, extending beyond the behavioral response.

1. Introduction

The ability to comprehend spoken language is a fundamental aspect of communication. Listeners are required to incorporate various operations such as parsing the auditory scene, keeping track of who said what, selectively attending to the target speaker, suppressing processing of irrelevant information, extracting meaning and storing it in memory, and utilizing personal knowledge to formulate appropriate responses (Schneider et al., 2010; Schneider, 2011). Multiple neural processes take place during these operations, which occur rapidly, both sequentially and in parallel (Henkin et al., 2002; Hillyard and Kutas, 1983).

Everyday speech interactions often occur in challenging, yet common, listening conditions that may typically include sensory and/or cognitive loads. Sensory load refers to degradation of the acoustic input

which under experimental conditions is frequently induced by presenting speech in background noise, either by broadband or white noise (i.e., "energetic masking") or by competing speech maskers (i.e., "informational masking") (Kaplan-Neeman et al., 2006; Cooke et al., 2008; Getzmann and Falkenstein, 2011; Strauß et al., 2014; Getzmann et al., 2015). There is substantial evidence of decreased speech understanding in background noise, especially in tasks involving informational masking (Sperry et al., 1997; Rajan and Cainer, 2008; Decruy et al., 2019).

Differently, cognitive load refers to additional demands imposed on the listener's attention or memory resources during speech processing (Mattys et al., 2012). An example of such a task is the well-known Stroop task (Stroop, 1935). The task challenges executive functions such as selective attention, inhibition, and conflict resolution (MacLeod, 1992; Kestens et al., 2021). In its auditory version, the task may include words

with two dimensions: a physical and a semantic dimension (Lew et al., 1997). The meaning of the word can be either congruent (e.g., the word "low" in low pitch) or incongruent (e.g., the word "low" in high pitch) (Sharma et al., 2019). The listener is required to selectively attend to a targeted dimension (e.g., pitch) while ignoring the irrelevant dimension (e.g., word meaning). Traditionally, a robust Stroop effect is demonstrated by prolonged response (or reaction) time and reduced performance accuracy to incongruent versus congruent stimuli, indicating failure of selective attention due to processing of conflicting information (Green and Barber, 1981; Morgan and Brandt, 1989; Gregg and Purdy, 2007).

The effect of sensory or cognitive loads on speech processing has also been studied by means of brain-based measures, such as electroencephalogram (EEG) recordings (Martin et al., 2008; Getzmann and Falkenstein, 2011; McCullagh and Shinn, 2013; Getzmann et al., 2015; Dimitrijevic et al., 2017; McHaney et al., 2021). Regardless of whether listeners process sensory information or perform high-demanding cognitive tasks, the activated brain regions exhibit measurable electrical activity, illuminating the underlying neural mechanisms. This electrical activity recorded from multiple-site scalp electrodes can be analyzed in the time (event-related potentials (ERPs); e.g., (Beres, 2017; Martin et al., 2008; Pratt, 2011)) and/or frequency domain (time--frequency representation (TFR); (Buzsaki, 2006; Cohen, 2014)). While the time and frequency domains represent different dimensions of the same signal, the main advantage of the TFR is that at a given time-point, the EEG data are decomposed into multiple oscillatory bands. This allows to identify and differentiate parallel brain processes. Furthermore, non-phase locked EEG activity is not averaged out even if not phase-locked to stimulus onset or to the listeners reaction (behavior response, i.e., button press; (Kalcher and Pfurtscheller, 1995; Herrmann et al., 2004; Siegel et al., 2012). Separation of phase-locked (evoked power, Lakatos et al., 2009) and not phased-locked (induced power, Tallon-Baudry et al., 1996; David et al., 2006) provides a tool for separation of sensory responses and corticocortical interactions taking place at the same time (David et al., 2006; Donner and Siegel, 2011; Chen et al., 2012; Yusuf et al., 2017).

Studies incorporating brain oscillatory activity to investigate load effects in the auditory modality are limited, and often focus on the effect of one specific load (sensory/cognitive). Published results vary and depend on the utilized task, eliciting stimuli, and type of data analysis methods. For example, the effect of increased sensory load, imposed by competing/degraded speech, resulted in increased or decreased alpha power (Obleser and Weisz, 2012; Wöstmann et al., 2015; Dimitrijevic et al., 2017; Wisniewski et al., 2021). The effect of increased cognitive load on neural oscillations has been studied using a variety of tasks (Wilsch et al., 2015; Wilsch and Obleser, 2016; Hjortkjær et al., 2020; Beldzik et al., 2022), however, only few studies utilized the auditory Stroop task (Oehrn et al., 2014; van de Nieuwenhuijzen et al., 2016; Sharma et al., 2021). These studies suggested frontal theta power as an index of conflict processing, showing enhancement to incongruent vs congruent stimuli, even when behavioral manifestations were not observed (Sharma et al., 2021). Additionally, during conflict processing an interplay between theta and gamma oscillations in prefrontal brain regions was evident (Oehrn et al., 2014).

The current study was set to evaluate the effect of both loads on TFRs in a within-subject design. Our working hypothesis was that sensory and cognitive loads will have different oscillatory signatures. Our assumptions relate to the theory of dual control (proactive and reactive) (Braver, 2012; Lieder and Iwama, 2021). Specifically, given that sensory load affects stimulus processing, it may influence reactive control and manifest in the neural activity before the behavioral response, presumably in alpha band activity (Obleser and Weisz, 2012; Wöstmann et al., 2017). While reactive control rests on sensory processing, cognitive load may rely on proactive control in order to maintain the task goals in working memory and enable reaction to the following stimulus based on a pre-prepared motor program. This proactivity may deplete cognitive

resources for action monitoring required to update and execute the motor program for the correct behavioral response. Therefore, we hypothesized that the cognitive load effect will take place mainly after the response time. We further expected, based on previous studies, that cognitive load will manifest in theta oscillations mainly in prefrontal region (Bastiaansen and Hagoort, 2003; Cavanagh et al., 2012; Albouy et al., 2017; Nowak et al., 2021). Finally, we hypothesized that frontal theta would manifest as neural correlates of conflict processing during Stroop task (Kerns et al., 2004; Hanslmayr et al., 2008; Ergen et al., 2014).

For this reason, we set up a paradigm with extended recording time after the behavioral response and separately processed brain signal before (pre-response) and after the response (post-response). Furthermore, we analyzed all data with respect to the stimulus (stimulus time-locked) and to the response (response time-locked). The underlying concept was to differentiate the stimulus-related and the response-related oscillatory activities allowing to distinguish between sensory processing and preparation/execution of the motor response. The main goal of the present study was, therefore, to identify the neural signatures of sensory and cognitive loads. Findings may potentially serve as a baseline for understanding age-related detrimental changes in speech processing in older adults, especially in challenging listening conditions.

2. Materials and methods

2.1. Participants and procedure

Twenty young adult listeners 21.9–28.7 years old ($M = 24.9$, SD $= 1.8$; 10 female) participated in this study. The sample size for this study was based on power analysis according to the effect size reported in previous studies on ERPs (Henkin et al., 2010, 2014). All participants in this study: 1) Reported no history of psychiatric, cognitive illness, brain damage, ear pathology, or any central nervous system disorders; 2) Exhibited right-handedness based on the translated Edinburgh Inventory for Handedness (Oldfield, 1971); 3) Used Hebrew as their dominant language; 4) Had at least 12 years of formal education ($M = 13.4$, SD $= 1.5$). Participants underwent hearing evaluation, including air- and bone-conduction thresholds evaluation in octave frequencies between 250 and 8000 Hz, determination of speech reception thresholds (SRT) and word recognition scores (Hebrew PB monosyllabic words). All participants met the criteria of normal hearing at octave frequencies between 250 and 8000 Hz and demonstrated air-conduction thresholds ≤ 15 dB HL. The mean interaural threshold difference did not exceed 5 dB HL at each frequency, and no clinically significant differences were detected between air- and bone-conduction thresholds (<15 dB HL). Pure tone average (PTA4; average threshold for, 0.5, 1, 2 and 4 kHz in dBHL) were comparable for right and left ears (right: $M = 5.56$, SD $= 2.64$, left: $M = 5.19$, SD $= 3.17$, t(19) $= 0.609$, $p = 0.492$, Cohen's $d = 0.127$). Word recognition scores (WRS) were within the normal range (>88 %) in both ears for all participants. All participants exhibited a score within the normal range in the forward and backward digit span subtest of the Wechsler intelligence scale ($M = 11.3$, SD $= 3.23$) (Wechsler, 1997). Five additional participants were excluded from the analysis due to: 1) extremely prolonged response time (RT) values (>3 standard deviations longer compared to the mean group values; one participant); 2) ERPs recordings significantly contaminated by excessive eye movements and/or myogenic artifacts (four participants). Exclusion of participants was decided only after failure to produce clear waveforms by means of disabling contaminated records and/or utilization of an eye movement correction algorithm. Participants were recruited through personal acquaintance or via internet and social networks. They provided written informed consent and received reimbursement. The study was approved by the Institutional Review Board of Ethics at Tel Aviv University.

2.2. Tasks, conditions and stimuli

The experiment consisted of two tasks, a vowel identification task and an auditory Stroop task, presented under two listening conditions: quiet and background noise. The Stroop task included two types of *stimuli*: congruent and incongruent (details below). In total, there were four task-condition combinations: 1) Vowel identification in quiet, 2) Vowel identification in noise, 3) Stroop task in quiet, 4) Stroop task in noise. The analysis in this study focused the following task-condition-stimuli combinations: 1) Vowel identification in quiet, 2) Vowel identification in noise, 3) Stroop task congruent stimuli in quiet (i.e., Stroop congruent), and 4) Stroop task incongruent stimuli in quiet (Stroop incongruent).

The vowel identification task included the vowels /a/ and /i/ produced by native Hebrew female speaker. The participants were instructed to classify each vowel (/a/ or /i/) with both hands by pressing one of two possible buttons on a response box (left hand operating the left button or right hand operating the right button). Stimuli were digitally recorded at 44.1 kHz sampling rate and 16-bit quantization using Goldwave 6.4 software. From a large sample of naturally produced stimuli, the final vowels set was selected. To minimize potential identification cues related to a specific utterance, four different versions of /a/ as well as four different versions of /i/ were chosen, creating a final set of eight vowels. The vowels were shortened to a duration of 230 ms by windowing the vowel offset, and were subjected to fade in and fade out procedures. The mean fundamental frequency for the four /a/ versions was 206–210 Hz, and for the four /i/ versions 230–235 Hz. In the background noise condition, each one of the stimuli were embedded in a different noise segment in order to avoid identification cues related to specific stimulus-noise combination. The background noise consisted of four talker babble noise (two males/two females) and was presented continuously throughout the task at an SNR level of +3 dB. Using RMS normalization procedure (MATLAB, Mathworks Inc.) the determined SNR was kept constant for each vowel presented in noise.

In the auditory Stroop task, participants were instructed to classify the speaker's gender (male or female) by pressing one of two possible buttons (left or right). The auditory Stroop tasks included two types of stimuli that were previously used (Henkin et al., 2010, 2014): 1) Congruent stimuli – the Hebrew two-syllable words /aba/ (father) and /ima/ (mother) produced by native male and female speakers, respectively; 2) Incongruent stimuli – the word /aba/ produced by a female speaker and the word /ima/ produced by a male speaker. The stimuli were digitally recorded at 44.1 kHz sampling rate and 16-bit quantization using Sound Forge 4.5 software. From a large sample of naturally produced stimuli, the final set of words was selected and shortened to a duration of 374 ms by windowing the final vowel offset. The mean fundamental frequency of the male speaker was 92 Hz (/aba/) and 100 Hz (/ima/), and for the female speaker 180 Hz (/aba/) and 190 Hz (/ima/). In what follows, the congruent Stroop stimuli will be denoted as *Stroop congruent* and the incongruent stimuli as *Stroop incongruent.*

All stimulus amplitudes in both tasks (vowel identification and auditory Stroop task) were calibrated to the same level using overall root-mean-square (RMS) normalization. After electrode application, participants were seated in a comfortable armchair in a sound-treated room. They were instructed to fixate their eyes on a colored circle located on the wall in front and to avoid excessive eye movements while listening to the stimuli and responding. Stimuli were presented every 3000 ms. A total of 200 stimuli were presented in every condition, each lasting approximately 10–12 min. Each of the tasks consisted of 200 stimuli, divided into two blocks of 100 stimuli each. The order of the tasks, as well as the order of the blocks within each task, was randomized. In addition, within each block, stimuli were randomized differently to ensure that no more than four consecutive presentations of the same vowel or two consecutive presentations of the same word (i.e., same combination of word and speaker's gender) occurred. Stimuli were presented via a loudspeaker located at azimuth zero at a distance of 1 m from the participant's head in both tasks. The presentation levels were adjusted individually to 30 dB HL above PTA4. The side of the response buttons was counterbalanced across participants in both tasks. A short practice included 12 stimuli that were presented before each condition. Short intermissions were provided between blocks and conditions.

2.3. Data acquisition

Continuous EEG data was recorded from 64 electrode sites arranged according to the international 10/10 system (Jurcak et al., 2007) using an EEG cap connected to a multichannel amplifier (System Plus Evolution software, Micromed S.p.A). Potentials were amplified from the EEG (100,000 gain) and electroculogram (EOG) (20,000 gain) channels. The reference and ground electrodes were placed on the chin and the right mastoid, respectively. Eye movements were monitored by an EOG, recorded by means of electrodes placed above and below the right eye. The impedance measured for each electrode was kept below 10 kOhm.

2.4. Data analysis

The EEG data processing and analysis for the TFR and for the ERPs analysis were performed offline using the FieldTrip toolbox (Oostenveld et al., 2011) and custom-made MATLAB scripts (Mathworks Inc.) (MATLAB scripts are available at: https://github.com/brllant). The raw EEG data was first imported in the Brainstorm software (Tadel et al., 2011) and converted (*.trc to *.eeg, *.vhdr and *.vmrk) for further preprocessing using the FieldTrip toolbox. Out of the 64 electrodes used for recording, 58 were analyzed. The electrodes used in the analysis were: AF3, AF4, Fz, F1, F2, F3, F4, F5, F6, F7, F8, FCz, FC1, FC2, FC3, FC4, FC5, FC6, Fpz, Fp1, Fp2, Cz, C1, C2, C3, C4, C5, C6, CPz, CP1, CP2, CP3, CP4, CP5, CP6, T3, T4, T5, T6, TP7, TP8, Pz, P1, P2, P3, P4, P5, P6, POz, PO1, PO2, PO3, PO4, PO7, PO8, Oz, O1, O2. Six electrodes (a1, a2, VEOR, VEOL, EB and el064) were not analyzed since they did not record brain activity of interest. The time window between the stimulus onset and the mean response time (mean RT, or mean reaction time) will be designated as pre-response period and the time window after the mean response time will be designated as post-response period. Response times were computed from the same trials that were also used for the EEG data analysis, i.e., all trials that were left after artifact removal. Only trials with correct responses were analyzed. The mean RTs for all task-condition-stimuli combinations were plotted with a custom MATLAB script using function from MATLAB Central File Exchange: 'al_goodplot' (Legouhy, 2023) (Fig. 1A).

2.4.1. Preprocessing

The preprocessing procedures were high-pass filtering, segmentation and artifact removal. The imported EEG files were first high-pass filtered at 0.1 Hz. The data were then segmented into 3000 ms epochs and triggered by the stimulus onset (0s: stimulus onset, −1000 ms pre-stimulus, and 2000 ms post-stimulus). For artifact removal, independent component analysis (ICA) weight matrix was calculated with 1Hz-high-pass filtered data and applied to the 0.1 Hz high-pass filtered data to improve the SNR and classification accuracy (Winkler et al., 2015), yet avoid the lower frequency attenuation resulted from the 1Hz-high-pass filter (Rousselet, 2012). Independent components representing eye blinks, horizontal eye movements and technical artifacts were removed from the EEG data. Record by record inspection allowed to identify and manually remove non-stereotyped artifacts. Additionally, the first epoch in each experiment block, and epochs with errors in responses (incorrect respond or no respond) were removed. There were on average 112 trials per task-condition-combination per participants used in the data analyses after preprocessing. For the Stroop task, the epochs were then separated into congruent and incongruent trials.

2.4.2. Time-Domain analysis

The preprocessed trials in each task-condition-combination were

pooled according to the ROI from these following four channels: 1) Frontal: F1, F2, Fz and FCz; 2) Central: C1, C2, Cz and CPz; 3) Parietal: P1, P2, Pz and POz; 4) Occipital: O1, O2, Oz and POz; 5) Temporal Left: T3, T5, TP7 and C5; and 6) Temporal Right: T4, T6, TP8 and C6 (for the temporal ROIs data see Supplementary Fig. 6). The trials in each ROI were baseline (0 – 200 ms) corrected and then averaged to obtain the event-related potentials (ERPs) N1 and P3 of the task-condition-combination in the ROI. For the ERPs in Fig. 1B-E, the previously 0.1 Hz high-pass filtered data were additionally filtered with 30 Hz low-pass filter. Topographical plots were generated with Fieldtrip function 'ft_topoplotER' using the following time window for N1: [2 ms – 120 ms] (Supplementary Fig. 4B) and P3: [200 ms – 400 ms] (Supplementary Fig. 4D).

2.4.3. Time-Frequency analysis

For every trial in each task-condition-combination, complex TFR was calculated using complex Wavelet transformation on the preprocessed data from each channel. For the time-frequency analysis, the continuous data were segmented into 6000 ms epochs, to avoid the edge artifacts resulting from using Wavelet convolution. The decomposition to 6000 ms epochs (−2500 to 3500 ms) into time-frequency space was performed using Morlet Wavelet convolution with the width of 6 cycles in 20 ms steps. Frequencies from 1 to 128 Hz in 1 Hz linear steps were analyzed. The complex TFRs for each trial were averaged throughout all trials. The total power was normalized to the baseline period (−800 to −200 ms) (Cohen, 2014a). The oscillatory activity was defined as *synchronization*, when its power increased relative to baseline, and *desynchronization*, when its power decreased relative to baseline. The TFRs were plotted using MATLAB function: 'contourf' (MATLAB, Mathworks Inc.) for the time-window of interest (3000 ms, −1000 to 2000 ms) logarithmically and covered the frequency range from 2 to 128 Hz (divided into frequency band: delta, theta, alpha, beta and gamma (Table 1)), and all available time points (−1000 to 2000 ms).

Complex TFRs of the trials from the four following channels were pooled in a ROI before later averaged throughout all trials: 1) Frontal: F1, F2, Fz and FCz; 2) Central: C1, C2, Cz and CPz; 3) Parietal: P1, P2, Pz and POz; 4) Occipital: O1, O2, Oz and POz; 5) Temporal Left: T3, T5, TP7 and C5; and 6) Temporal Right: T4, T6, TP8 and C6. Because the central ROIs- and temporal ROIs TFRs turned out to be similar, the temporal ROIs data were included in the supplementary material only (Supplementary Fig. 6). For grand mean TFR (i.e., stimulus time-locked TFR, Figs. 2, 7A, C and 8A, C), complex TFRs for all participants in each ROI were pooled together. For each stimulus time-locked TFR (Fig. 2) and response time-locked TFR (Supplementary Fig. 1), cluster-based permutation test was performed to define statistically significant oscillatory activity relative to the baseline and plotted as black contours on the respective TFR (see *2.4.5. Statistical Tests*).

For the difference TFRs between task-condition-stimuli combinations (Fig. 4I-L, 5I-L, 6I-L and Supplementary Fig. 5I-L), the stimulus time-locked TFRs from one task/condition/stimulus (Fig. 4A-D, 5A-D, 6A-D and Supplementary Fig. 5A-D) were subtracted from stimulus time-locked TFRs from another task/condition/stimulus (Fig. 4E-H, 5E-H, 6E-H and Supplementary Fig. 5E-H). Cluster based permutation test was performed to define statistically significant oscillatory activity differences between the two task-condition-stimuli combinations (e.g., vowel identification in quiet and in noise) and plotted as black contours on the respective difference-TFRs (see *2.4.5. Statistical Tests*). The 0.1 Hz high-pass grand mean averages were plotted on the respective TFRs as blue traces (Fig. 4A-H, 5A-H, 6A-H, Supplementary Fig. 5A-H and Supplementary Fig. 7).

Topographical plots were generated with Fieldtrip function 'ft_topoplotTFR' using the following time-frequency window for delta-theta: [1 – 8 Hz, 0 ms – 600 ms], alpha: [8 – 12 Hz, 200 – 600 ms] and beta: [13 – 30 Hz, 800 – 1600 ms] (Supplementary Fig. 4G).

2.4.4. Response time-locked TFR

To evaluate the response-related activity, the time-domain signals in each trial were triggered by the RT (Fig. 3). The raw EEG data were preprocessed and analyzed in the same manner as described previously. The only differences in analysis pipeline were: 1) In the segmentation stage, the data were then segmented into 3000 ms epochs and triggered not by the stimulus onset, but by the RT (i.e., behavioral reaction onset) (0s: RT or reaction onset, −1500 ms pre-response, and 1500 ms post-response), 2) The baseline-period was −1300 to −700 ms pre-response (instead of −800 to −200 ms pre-stimulus). The same trials as the stimulus-triggered were used for the response-triggered analysis. The 0.1 Hz high-pass grand mean averages were plotted on the respective TFRs as blue traces. For the difference-TFRs between stimulus time-locked- and response time-locked TFRs (Supplementary Fig. 2), the response time-locked TFRs were first aligned so that the mean stimulus onset overlaid the 0 in the stimulus time-locked TFRs (i.e., the mean RT in the stimulus time-locked data overlaid the 0 in the response time-locked TFRs). The aligned response time-locked TFRs were subtracted from stimulus time-locked TFRs.

2.4.5. Statistical tests

The time-frequency data were statistically analyzed using a non-parametric cluster-based permutation approach using Fieldtrip toolbox (Maris and Oostenveld, 2007). For this, independent samples *t*-tests (1) between activation- (−200 to 2000 ms) vs. baseline-period (−200 to-800 ms) and (2) task-condition-stimuli combinations (e.g., vowel identification in quiet and in noise) were calculated for each sample point. Significant values (alpha < 0.05) were clustered based on their adjacency in time, space and frequency. The critical p-value for each cluster was calculated using the Monte Carlo method with 500 random permutations. If the summed t-value of the observed data cluster was higher than 95 % of the random partitions, then the cluster was considered to represent a significant difference between the two compared groups. For comparison between task-condition combinations only the activation-periods were entered into the statistical test. For comparison between activation- vs. baseline-period, because the activation period was longer than the baseline period and the statistical test requires the same data length, the activation period were divided into four sub-periods (−200 ms until 400 ms, 300 ms until 900 ms, 800 ms until 1400 ms, and 1400 until 2000 ms for stimulus time-locked data; −700 ms until 100 ms, −200 ms until 400 ms, 300 ms until 900 ms, and 900 until 1500 ms for response time-locked data) with each having the same data length as the baseline period. Each data from each sub-period were then statistically tested against the baseline-period. The union of the resulted significant clusters from each sub-periods was considered to represent a significant difference between the activation- and the baseline-period. The RTs were statistically analyzed using paired *t*-tests.

2.4.6. Source localization

The forward model was computed using the 3D anatomical dataset (scalp, skull and brain) from the anatomic MRI scans in the Fieldtrip repository. The volume conduction model was calculated with the boundary element method (Oostendorp and Van Oosterom, 1991; Fuchs et al., 2002) using OpenMEEG method. This method ensured a more realistic volume conduction model needed for localizing the source of EEG data (Hamalainen and Sarvas, 1987; Westner et al., 2022). The positions of 58 analyzed electrodes were aligned to 3D anatomical data. To create a source model, volumetric grids inside the 3D brain data were created with one source per cm^3. The source model, along with the volume conduction model and electrodes' positions were then used to compute the forward model.

For the inverse modeling, a dynamic imaging of coherent sources (DICS) beamformer was used to estimate the source power of band-limited activity in the frequency domain (Gross et al., 2001). For the source localization in the time domain (Supplementary Fig. 4C, E) linearly constrained minimum variance (LCMV) beamformer (Van Veen

et al., 1997) was used (for more details of beamformers see (Westner et al., 2022)). The source analysis was computed and plotted for activation period relative to the baseline period. For N1 the activation period was defined from 2 ms until 120 ms, and its baseline period was from −2 ms until −120 ms (Supplementary Fig. 4C). For P3 the activation period was from 200 ms until 400 ms, and the baseline was from −2 ms until −220 ms (Supplementary Fig. 4E). Only activities above 0.7 threshold were plotted.

For frequency domain source analysis preprocessing, time-domain data of each condition/task for each participant was re-referenced to the common average reference and pooled together from all participants. For each condition/task, only the sources of those frequency bands were identified, which were significant in the total-power-TFR. For the analyzed frequency band, an activation period was defined based on the total-power-TFR. For each activation period, a baseline period with the same length was defined, starting at −200 ms.

The baseline- and activation periods for the analyzed frequency bands are summarized in Table 1.

Once the time-domain data were segmented according to the baseline- or activation periods and also frequency of interest, a multitaper frequency transformation was conducted to obtain the power and cross spectral density matrix. A common spatial filter based on both periods (baseline- and activation period) with the regularization parameter (lambda) set to 10 % was calculated. The application of common spatial filter based on both periods was used to prevent the filter from being biased towards one period. The filter was subsequently applied separately to each period, providing the source power estimates. The contrast of the source power estimates of both periods was then calculated by subtracting source power estimate of the baseline period from the source power estimate of the activation period, then divided by the source power estimate of the baseline period. For the time-frequency window, which in the grand mean TFR showed significant synchronization (power increase relative to baseline), negative source power estimate values were masked, while in the time-frequency window, where desynchronization (power decrease relative to baseline) was observed in the grand mean TFR, the positive source power estimates values were masked. This step is commonly done during the plotting step of the source reconstruction result using plotting command in Fieldtrip (Oostenveld et al., 2011) by specifying 'rampup' (only plot positive value) or 'rampdown' (only plot negative value) as parameter. The contrast was then normalized to the maximum (for synchronization) or minimum (for desynchronization) and plotted on the 3D brain surface data (Fig. 7B, D, and 8B, D). The description of anatomical labels of the activated brain regions found in the result section was determined manually based on the 2D-MRI data plotted on three anatomical planes: sagittal-, frontal- and transversal plane. The Brainnetome Atlas (Fan et al., 2016) was used to relate the coordinates to the known anatomical structures (Tables 3–6 and Supplemetary Tables 1–4).

2.4.7. Evoked- and Induced power

The stimulus time-locked TFRs (total-power-TFRs) can be further separated into induced- and evoked-power-TFRs. To compute the induced-power-TFR (Supplementary Fig. 3B, D, F, and H) the time-domain trial average (i.e., the event-related potential (ERP)) was subtracted from each trial before TFR computation. The resulting complex TFRs for each trial were subsequently averaged throughout all trials to obtain the induced-power-TFR. The induced-power-TFR was normalized in the same manner as the total-power-TFR (baseline period: −800 to −200 ms). Finally, the evoked-power-TFR (Supplementary Fig. 3A, C, E, and G) was obtained by subtracting the baseline-normalized induced-power TFR from the baseline-normalized-total-power-TFR (Cohen, 2014a). The 0.1 Hz high-pass filtered ERPs were plotted on the evoked-power TFR as blue traces (Supplementary Fig. 3A, C, E and G).

3. Results

3.1. Mean RTs and ERPs

The mean behavioral RTs (response- or reaction times) (Fig. 1A) in the vowel identification task in quiet, in noise, Stroop congruent and Stroop incongruent are presented in Table 2. RT increased with increasing task complexity. In the vowel identification task, RT in the noise condition was significantly longer than in the quiet condition (two-tailed paired *t*-test, t(19) = −8.7, *p* < 0.001). A significant effect of congruency indicated longer RT to incongruent compared to congruent stimuli (two-tailed paired *t*-test, t(19) = −4.2, *p* < 0.001). The mean RT in the vowel identification in quiet was similar to that in Stroop congruent (two-tailed paired *t*-test, t(19) = −1.1, *p* = 0.2792) and was significantly shorter compared to Stroop incongruent (two-tailed paired *t*-test, t(19) = −2.9, *p* < 0.001). All participants in all task-condition-stimulus combinations achieved performance scores higher than 95 %.

The ERP components N1 and P3 (Fig. 1B-E, shown for the central ROI) were elicited in all task-condition-stimuli combinations at the expected latencies (Table 2). Sustained DC potentials modulated with oscillatory pattern were observed throughout the time window of 0.5 – 2.0 sec after stimulus presentation, indicative of ongoing neural activity. This shift exceeded the mean behavioral RT (the post-response period) and was found in the vowel identification task in quiet, in noise, Stroop congruent and Stroop incongruent.

3.2. Stimulus time-locked TFRs

Time-frequency analysis of the stimulus time-locked data was performed to reveal the underlying oscillatory activity related to Fig. 1. The stimulus time-locked TFRs revealed significant power changes (relative to the baseline) in all task-condition-stimuli combinations (Fig. 2, for TFR with grand mean averages plotted see Supplementary Fig. 7). Both the pre- and post-response periods included significant oscillatory activity, depending on task, condition and stimulus. We observed in general a high delta-theta power right after stimulus onset in all task-condition-stimuli combinations, often followed by alpha desynchronization. Interestingly, there was significant long-lasting post-response oscillatory activity in the theta band observable with highest power in parietal and occipital ROIs, especially in the Stroop task. Additionally, in

Table 1
Time-frequency windows used to localize the sources of frequency bands activities. The baseline- and activation periods are relative to stimulus onset (stimulus onset is 0 ms).

| Frequency band	Frequency range [Hz]	Baseline period [ms re stimulus onset]	Activation period [ms re stimulus onset]
Pre-response Theta	4 – 8	(−200) – (−900)	0 – 700
Pre-response Alpha	8 – 12	(−200) – (−600)	200 – 600
Pre-response Beta Synchronization	13 – 18	(−200) – (−600)	200 – 600
Pre-response Beta Desynchronization	13 – 30	(−200) – (−800)	0 – 600
Post-response Theta	4 – 8	(−200) – (−800)	1200 – 1800
Post-response Alpha	8 – 12	(−200) – (−600)	1200 – 1600
Post-response Beta	13 – 30	(−200) – (−600)	800 – 1400

Event Related Potential (ERP)
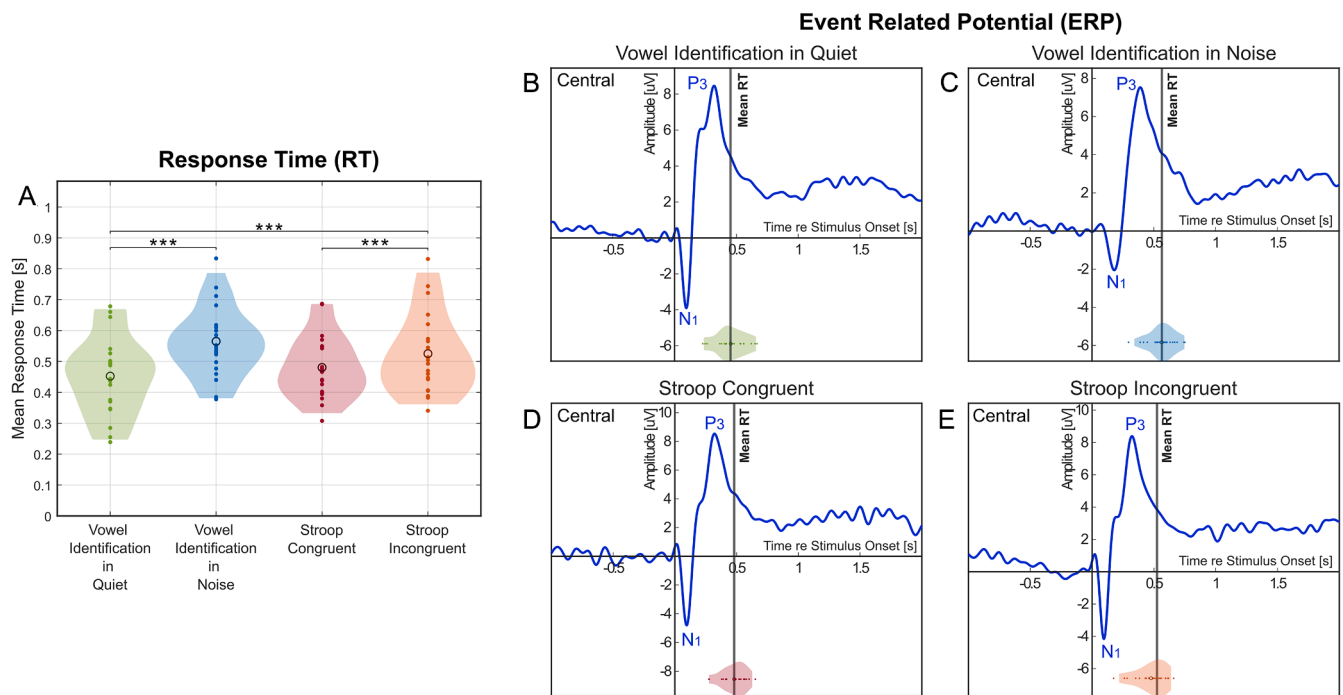
Fig. 1. A. Mean RTs (response- or reaction times) from all participants in the vowel identification task in quiet, in noise, Stroop congruent and incongruent. Filled circles indicate individual data, and open circles indicate the grand mean RT from all participants. B-E. Grand mean ERPs (event-related potentials) in the central ROI for B. vowel identification task in quiet; C. vowel identification task in noise; D. Stroop congruent; E. Stroop incongruent. *** ~ $p < 0.001$.

Table 2
Response times, N1- and P3-amplitudes and latencies for the vowel identification and the Stroop task for central ROI, shown as means and standard deviations.

Task/Condition	Response Time [ms]	N1 Amplitude [μV]	N1 Latency [ms]	P3 Amplitude [μV]	P3 Latency [ms]
Vowel identification in Quiet	452.3 ± 126.2	−4 ± 2.9	94.5 ± 16.4	8 ± 3.6	322.5 ± 79.8
Vowel identification in Noise	565.1 ± 116.2	−2.5 ± 2	182.5 ± 21.8	7 ± 3.6	390.5 ± 42.5
Stroop congruent	480.5 ± 102	−4.4 ± 2.9	96.5 ± 15	8.6 ± 5.4	330.5 ± 66.5
Stroop incongruent	525.3 ± 114.9	−4 ± 2.7	96.5 5 ± 15.3	7.9 ± 4	328.5 ± 50.4

all task-condition-stimuli combinations, there was beta-band activity observed in the post-response period in the frontal and central ROIs. Some transient changes of oscillatory power were also observed in the gamma and beta range.

3.3. Response time-locked TFRs

To separate the response-related from stimulus-related activity, the recordings were averaged with respect to the RT (Fig. 3, for cluster-based permutation test results see Supplementary Fig. 1). This means that the stimulus onset was no longer synchronized among trials (i.e., jitters) but the responses were aligned at the exact same time point in each trial (0 sec. on the abscissa in Fig. 3). In the response time-locked grand mean averages (blue traces in the panels of Fig. 3), there was no observable early component (i.e., stimulus time-locked N1 component, see Fig. 1B-E and compare blue traces in the panels of Supplementary Fig. 7 with those in Fig. 3). The P3 component was, however, still well discernible in the response time-locked data (blue traces in the panels of Fig. 3B-D, F-G, J-K and N-P at around 0 s.). Interestingly, there were two distinct peaks/components shortly before and after the 0 s. observable in frontal-ROI (Fig. 3A, E, I, M).

In the time-frequency domain, following the stimulus presentation, delta-theta synchronization was found in all response-time locked TFRs. This pre-response delta-theta power in the vowel identification task in quiet, Stroop congruent, and Stroop incongruent were weaker in comparison to the stimulus time-locked TFRs (Fig. 2, for difference-TFRs between stimulus- and response time-locked TFRs see Supplementary

Fig. 2). Additionally, alpha desynchronization overlapping the response time in all response time-locked TFRs (more prominent in Stroop task) was evident.

After the behavioral response, post-response alpha synchronization was found in the parietal and occipital regions (Fig. 3G, H) in the noise condition. This was stronger in response time-locked TFRs than compared to stimulus time-locked TFRs (Figs. 2, 5E-H). Finally, the beta activity in the post-response period was observed in both stimulus- and response time-locked TFRs, mainly in frontal-ROIs (Figs. 2, 5A, E, I, M).

Taken together, the stimulus time-locked and response time-locked TFRs shared most of the properties with some minor differences (Fig. 2, Supplementary Fig. 1). Difference TFRs (Supplementary Fig. 2) revealed that the main differences were observed in the time between stimulus onset and response time (pre-response period), especially in theta band. The post-response period appeared similar in stimulus time-locked and response time-locked TFRs.

3.4. Sensory load: vowel identification in noise vs. quiet

To reveal the effect of sensory load on the TFRs, we subtracted the stimulus time-locked TFRs in noise (Fig. 4E-H) from those in quiet (Fig. 4A-D). Significant effects of sensory load were found (Fig. 4I-L) (cluster-based permutation test, noise vs. quiet, $p < 0.05$): (1) less delta and theta activity in noise compared to the quiet condition in all ROIs, most prominent in the frontal ROI (Fig. 4I-L, this effect was mostly due to reduced evoked power in noise compared to the quiet condition, Supplementary Fig. 3), (2) more alpha power in noise compared to the
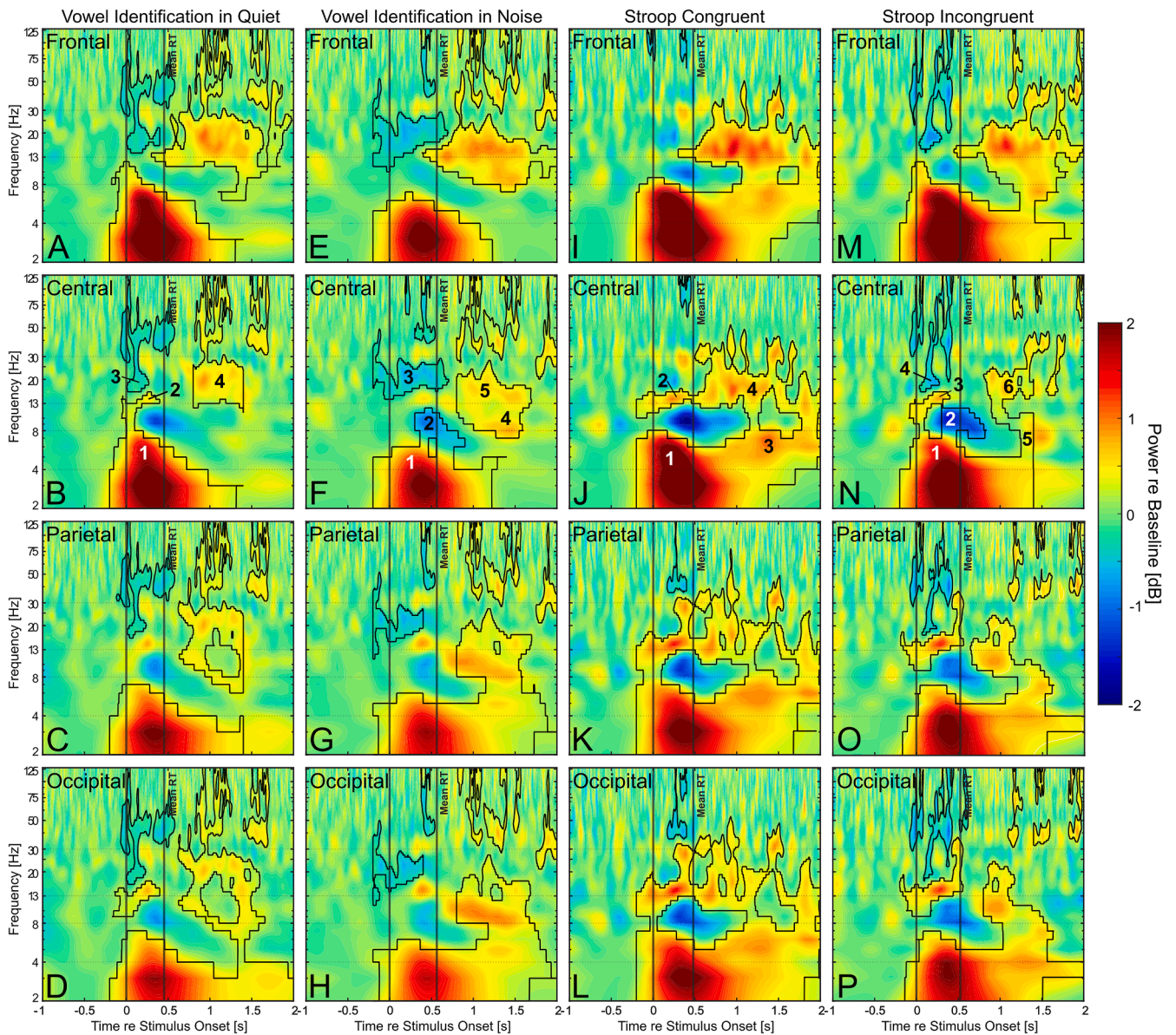
Fig. 2. Stimulus time-locked TFRs with black contours, showing the cluster-based permutation test result compared to the baseline (cluster-based permutation test, *p* < 0.05) for: A-D. Vowel identification in quiet, E-H. Vowel identification in noise, I-L. Stroop congruent and O-P. Stroop incongruent. Stimulus onset (0 ms) and mean RT (response- or reaction time) are shown as grey vertical lines. The index shown in the central ROI-TFRs (B, F, J, and N) correspond to the index of significant activity used in the following text. Significant oscillatory activity in the vowel identification task in quiet (A-D) included delta-theta- synchronization (index 1), beta synchronization (index 2), and beta-gamma desynchronization (index 3) after stimulus onset. After the mean response time (post-response period) there was a significant beta synchronization (index 4) in the vowel identification task in quiet. In the vowel identification task in noise (E-H) the significant activities were delta-theta synchronization (index 1), alpha desynchronization (index 2), and beta desynchronization (index 3). In the post-response period, there were significant theta-alpha synchronization (index 4) and beta synchronization (index 5). In the Stroop congruent (I-L) significant activity included delta-theta synchronization (index 1), and alpha-beta synchronization (index 2). After the response (post-response period) activity included theta synchronization (index 3) and beta synchronization and (index 4). The significant activities in Stroop incongruent (M-P) included delta-theta synchronization (index 1), alpha desynchronization (index 2), beta synchronization (index 3) and beta-gamma desynchronization (index 4). In the post-response period, significant delta-theta synchronization (index 5) and beta synchronization (index 6) were evident.

quiet condition, around and after the RT in all ROIs, but predominantly in the occipital ROI; (3) more theta power in noise compared to the quiet condition in the parietal and occipital ROIs; (4) transient differences in beta-gamma bands.

3.5. Cognitive load: stroop incongruent vs. vowel identification in quiet

The task-condition-stimuli combinations allowed to study different aspects of cognitive load. To quantify the cognitive load we compared Stroop incongruent, involving the maximal cognitive load (including conflict processing), with vowel identification in quiet, involving the lowest cognitive load. That Stroop incongruent was the most challenging is supported by the significantly prolonged response time compared to the Stroop congruent (and vowel identification task, Fig. 1A). This comparison (Stroop incongruent vs. vowel identifcation task in quiet) allowed us to further use the same baseline task (vowel identification task in quiet), and thus enabled a direct comparison between cogntive load effects with sensory load effects.

The mean TFR of the vowel identification task (Fig. 5A-D) was subtracted from the mean TFR of Stroop incongruent (Fig. 5E-H) to study
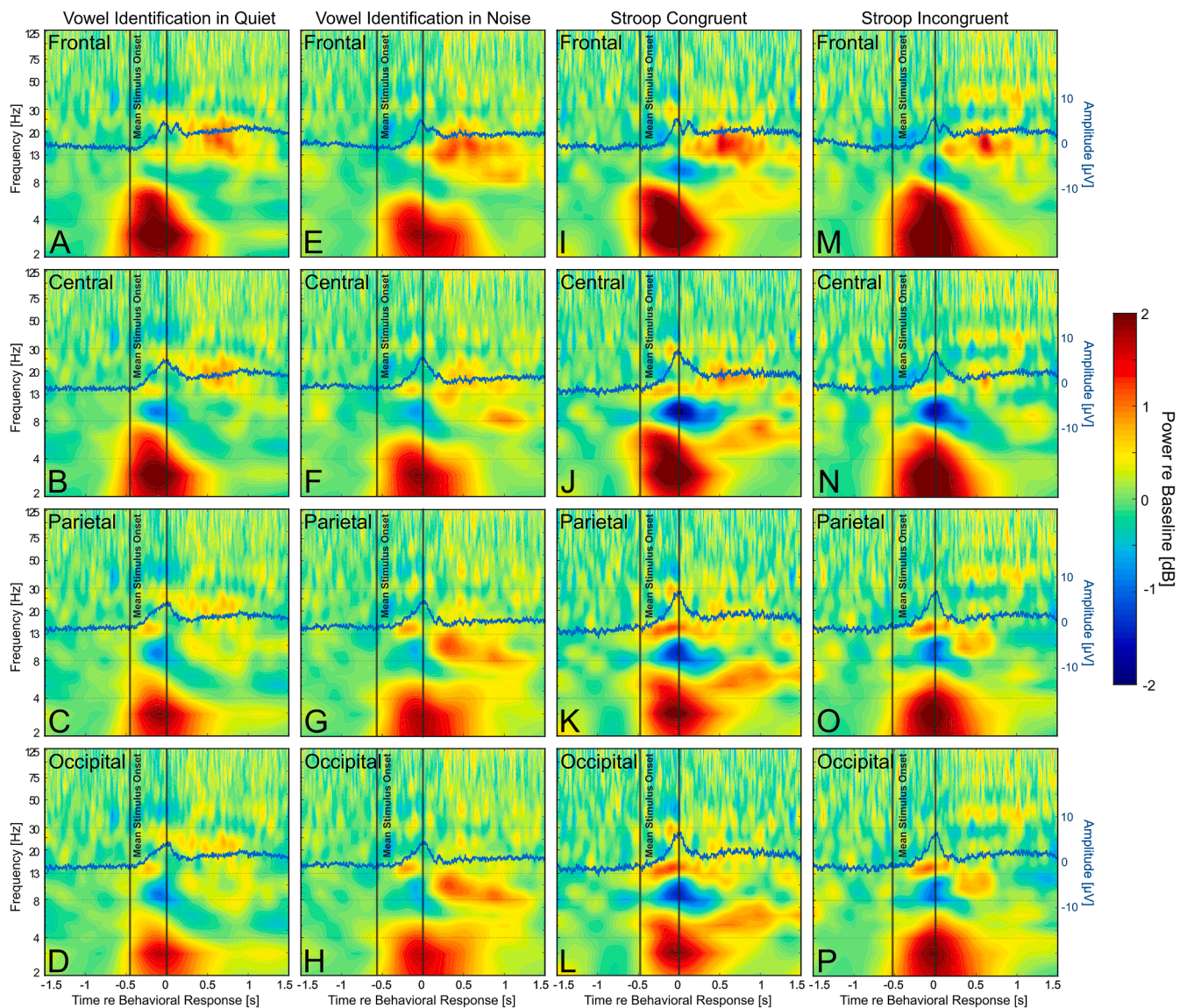
Fig. 3. Response time-locked TFRs for: A-D. Vowel identification task in quiet, E-H. Vowel identification task in noise, I-L. Stroop congruent, M-P. Stroop incongruent. Behavioral response onset (0 ms) and mean stimulus onset are shown as grey vertical lines. In the grand mean averages (blue trace), the N1 component disappeared in all panels. Only the P3 component remained discernible, especially in frontal ROI with a bimodal peak. In these response time-locked TFRs, the post-response activity is still well discernible, suggesting that both the stimulus and the response are involved in its generation (for difference-TFRs between stimulus- and response time-locked TFRS see Supplementary Fig. 2).

cognitive load effects. The difference TFRs revealed several significant effects (cluster-based permutation test, noise vs. quiet, $p < 0.05$): (1) theta synchronization was more pronounced in the Stroop incongruent than in the vowel identification task at around behavioral response time and lasted for about 500 ms. This significant difference in theta between the two tasks was observable in all ROIs and was found in all difference plots (Fig. 5I-L); (2) larger alpha power in the Stroop incongruent compared to the vowel identification at around 1 second post-stimulus, most prominent in parietal and occipital ROIs (Fig. 5K-L); and (3) there were differences in beta-gamma transients (Fig. 5I-L).

3.6. Conflict processing: stroop incongruent vs. stroop congruent

A way to quantify conflict processing is by the comparison between Stroop incongruent and Stroop congruent, known also as the congruency effect (or Stroop effect). Within the Stroop task, both these stimuli impose cognitive load, however, the incongruent stimuli involves more conflict processing. The mean TFR of the congruent stimuli (Fig. 6A-D)

was subtracted from the mean TFR of the incongruent stimuli (Fig. 6E-H) to reveal the neural correlates of conflict processing in absence of the shared aspects of cognitive load involved in the processing of both stimuli (neuronal activities related to these aspects will be eliminated due to the subtraction).

The subtraction revealed significant effects (cluster-based permutation test, incongruent vs. congruent, $p < 0.05$) including: (1) statistically higher theta power in the congruent vs. incongruent trials in both pre-response and post-response periods (Fig. 6I-L). In other bands differences were not significant with the exception of (2) small islands in beta-gamma range, and (3) alpha at around 1 second in parietal and occipital ROIs (Fig. 6K-L).

3.7. Sources of oscillatory activities

We localized the significant oscillatory activations to their estimated sources in the brain (for more details of the time- and frequency-windows used for the analysis see *Methods Source localization,*
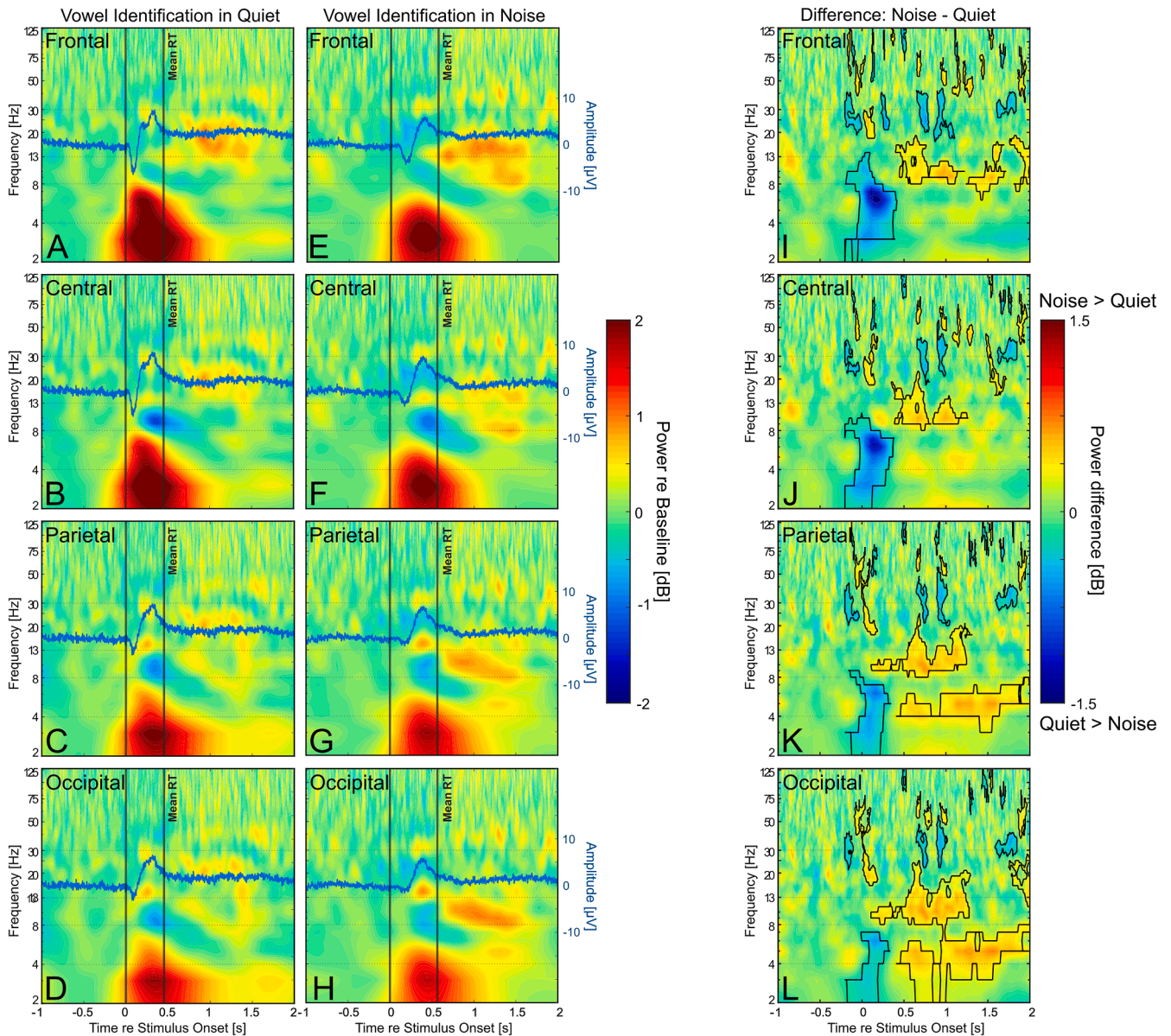
Fig. 4. Effects of sensory load on oscillatory activities. A-D. TFRs of the vowel identification task in quiet. E-H. TFRs of the vowel identification task in noise. I-L. Difference-TFRs between TFRs of the vowel identification in quiet and in noise to observe sensory load effects. In the difference plots (I-L), red color means more power in noise condition, while blue means more power in quiet condition. Noise resulted in decreased pre-response theta activity. Whereas the pre-response theta power was weaker, in the post-response period higher theta and alpha activities were observed in noise compared to the quiet condition. Stimulus onset (0 ms) and mean RT (response- or reaction time) are shown as grey vertical lines. Blue trace in A-H denotes the less- filtered ERPs from the same ROI. Black contours in I-L denotes significant differences between the two conditions (cluster-based permutation test, $p < 0.05$).

Table 1). We focused on activity that was significantly different from baseline (Fig. 2). Only sources in cerebral cortex were considered and transient activities were not considered. In what follows we highlighted similarities and differences between the sources of the specific task/ condition/stimulus (for details see Supplementary Tables 1–4).

3.7.1. Sources comparison: vowel identification task in quiet and in noise

In the pre-response activity (Fig. 7 and Tables 3, 4 Pre-Response), sources often had a bihemispheric activation pattern. In the pre-response period, the oscillatory activity was dominated by theta response that had sources in middle and inferior frontal gyrus, superior temporal gyrus, as well as precuneus. This was more localized in the noise condition. Furthermore, in the noise condition we observed alpha desynchronization. The alpha desynchronization had sources in superior and inferior temporal gyrus, inferior and middle frontal gyrus, and

precuneus. Post-response activity (Fig. 7 and Tables 3, 4 Post-Response) was mainly in the beta band that had widespread sources in frontal gyrus, cingulate gyrus, and precuneus on both hemispheres. Additionally, there was alpha synchronization in the noise condition that had sources on the left hemisphere, mainly in frontal lobe (widespread) and lingual gyrus.

3.7.2. Sources comparison: vowel identification in quiet and stroop incongruent

Stroop incongruent showed more significant activities than vowel identification in quiet in the pre-response period (Figs. 7, 8 and Tables 3, 6 Pre-Response). Pre-response theta synchronization had sources in middle frontal gyrus and cingulate gyrus, but were localized to the left hemisphere in Stroop incongruent. Beta synchronization sources were also more localized mostly to left temporal gyrus in Stroop incongruent,
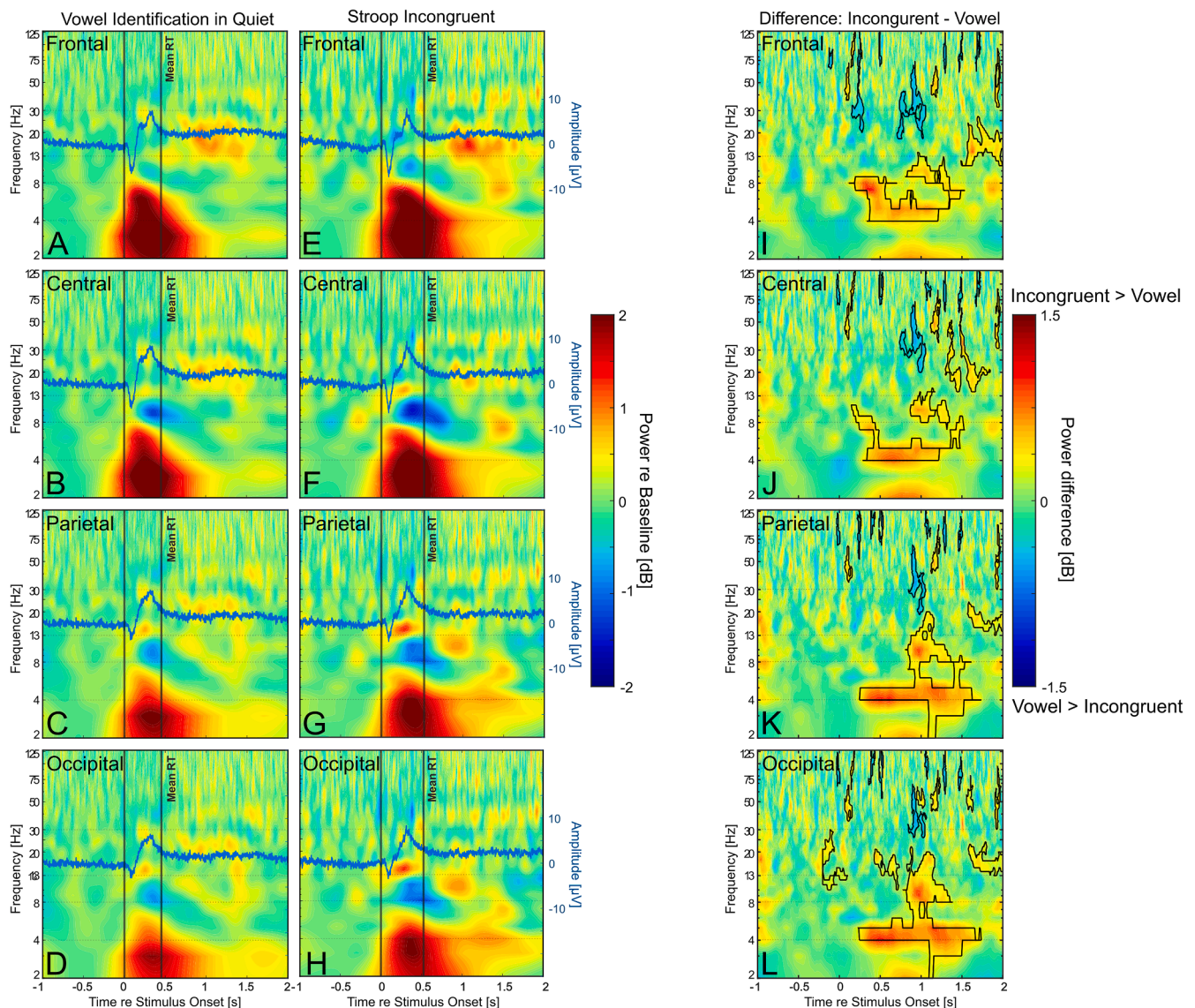
Fig. 5. Effects of cognitive load on oscillatory activities. A-D. TFRs of the vowel identification task in quiet. E-H. TFRs of Stroop incongruent. I-L. Difference-TFRs between TFRs of the vowel identification in quiet and Stroop incongruent to study the cognitive load effects. In the difference plots (I-L), red color means more power in Stroop incongruent, while blue means more power in vowel identification task. The difference is most pronounced in the theta and delta range at around the RT and lasted for approximately 500 ms. At around 1 second post-stimulus, there was also stronger alpha power in Stroop incongruent. Stimulus onset (0 ms) and mean RT (response- or reaction time) are shown as grey vertical lines. Blue trace in A-H denotes the ERPs from the same ROI. Black contours in I-L denotes significant differences between the two conditions (cluster-based permutation test, $p < 0.05$). .

while in vowel identification the sources were postcentral gyrus, parietal lobule and precuneus. Beta desynchronization in both Stroop incongruent and vowel identification were widely distributed, including precuneus and lingual gyrus. Additionally, in Stroop incongruent alpha desynchronization was observed, with sources including middle and inferior temporal gyrus as well as pre- and postcentral gyrus.

In the post-response period (Figs. 7, 8 and Tables 3, 6 Post-Response), beta activity was observed in both vowel identification and Stroop incongruent, with their sources being superior frontal gyrus and cingulate gyrus. In Stroop incongruent, we furthermore observed significant post-response theta, with sources in middle frontal gyrus and bilateral paracentral gyri.*3.7.3 Sources Comparison: Stroop Congruent and Stroop Incongruent*

Stroop incongruent showed more significant activities than Stroop congruent in the pre-response period (Fig. 8 and Tables 5, 6 Pre-Response). In the pre-response period, theta synchronization activated sources on the left hemisphere in Stroop congruent, whereas it activated

sources mainly in the right hemisphere in Stroop incongruent. The sources were mainly localized in the middle frontal gyrus. Furthermore, compared to Stroop congruent, in the Stroop incongruent we observed less temporal activation and more activation in cingulate gyrus. The frontal activation was also more widespread in the Stroop incongruent. Additionally, we observed alpha and beta desynchronization in Stroop incongruent, with sources in the temporal lobe, pre- and postcentral gyrus, and precuneus.

In the post-response period (Fig. 8 and Tables 5, 6 Post-Response), theta and beta activity were observed in both Stroop congruent and incongruent. The sources of theta activity were in the frontal gyrus. In the Stroop congruent, the inferior frontal gyrus and the superior temporal gyrus were activated, together with lingual and fusiform gyrus. In the Stroop incongruent, it was the middle frontal gyrus and bilateral paracentral gyri. The post-response beta activity had sources for both types of Stroop stimuli in superior frontal and cingulate gyrus, similar to the sources of post-response beta in the vowel identification task both in
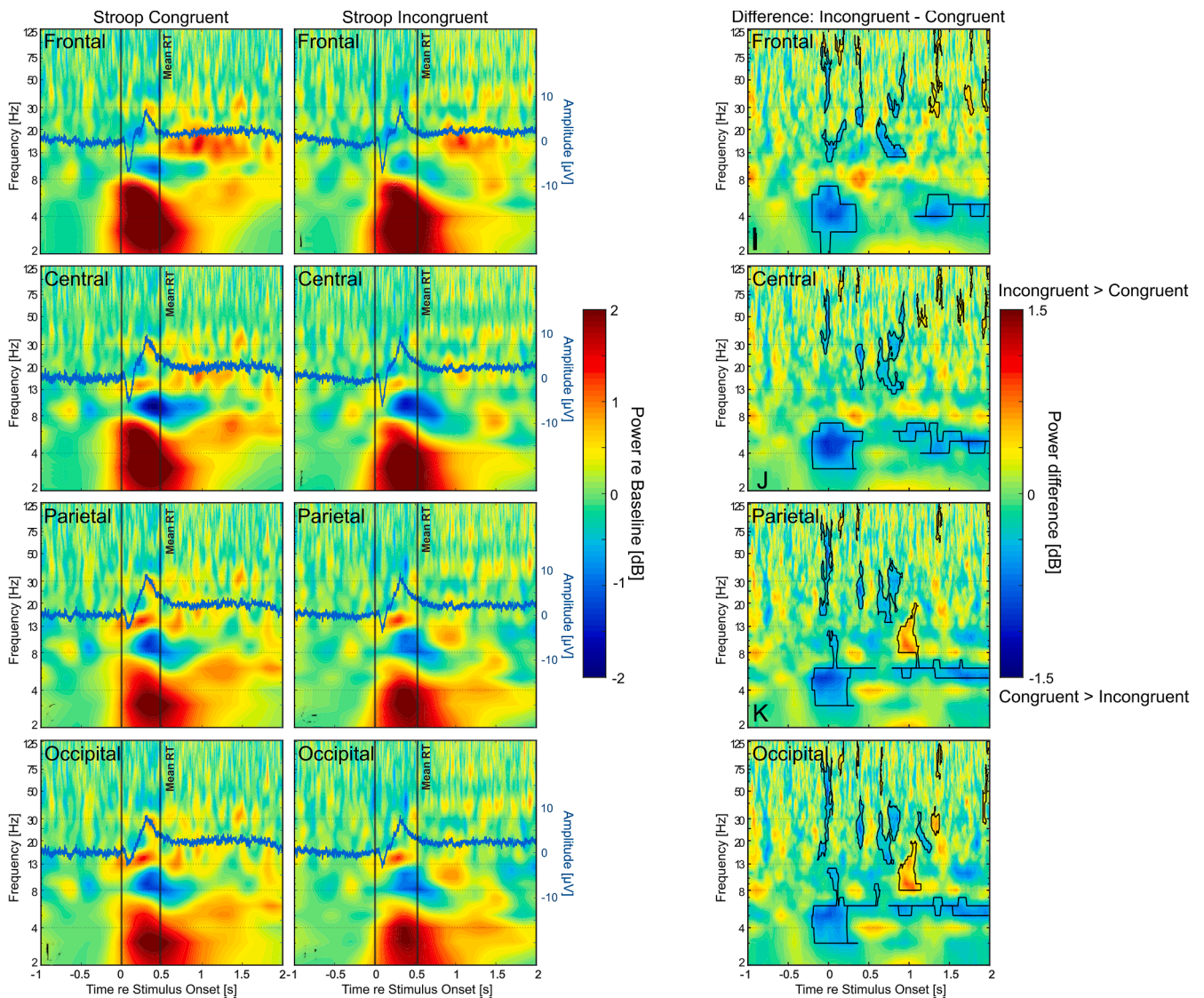
Fig. 6. Congruency effect (or Stroop effect) on oscillatory activities as signatures of conflict processing. A-D. TFRs of the Stroop congruent. E-H. TFRs of the Stroop incongruent. I-L. Difference-TFRs between TFRs of the Stroop congruent and Stroop incongruent to observe the congruency effect. In the difference plots (I-L), red color means more power in Stroop incongruent, while blue means more power in Stroop congruent. In congruent stimuli, pre- and post-response activity is stronger in both theta band. Stimulus onset (0 ms) and mean RT (response- or reaction time) are shown as grey vertical lines. Blue trace in A-H denotes the ERPs from the same ROI. Black contours in I-L denotes significant differences between the two conditions (cluster-based permutation test, $p < 0.05$). .

quiet and in noise.

4. Discussion

The present study, for the first time, directly dissociates between the oscillatory neuronal signatures of sensory and cognitive loads. The oscillatory activities provided distinct information which is not observable in the ERPs. We identified the oscillatory activities for each task, condition, and stimuli, localized their sources in the brain, and compared between them to study the neuronal signatures of each load. We analyzed activities in both pre- and post-response periods and observed signatures of executive control, especially in the post-response period.

In general, sensory and cognitive loads showed different effects. Higher sensory load resulted in lower theta-delta power in the time window between stimulus presentation and behavioral response (pre-response period). Differently, higher cognitive load increased theta oscillatory activity in the second half of the pre-response period that further continued in the post-response period. Furthermore, sensory

load resulted in prolonged increased alpha power in the post-response period, whereas the increased alpha power in cognitive load was less robust. In addition to load-specific theta and alpha oscillatory effects, we observed task-condition-stimulus specific pre- and post-response beta activity and beta-gamma transients. This demonstrates the significant amount of postprocessing during challenging listening situations.

4.1. Sensory load effects

In all tasks and conditions, after stimulus onset, the largest power change was in the delta-theta band (Başar et al., 2001; Demiralp et al., 2001; Huster et al., 2014), with its timing matching the N1-component (Fig. 6A-D). In higher sensory load (noise compared to quiet), theta power (mainly evoked, Supplementary Fig. 3) overlapping the early ERP component was smaller (Fig. 6E-L), presumably related to the effect of energetic masking on sensory-perceptual processing (Riecke et al., 2009; Hsiao et al., 2009; Hickok et al., 2015; Niemczak and Vander Werff, 2019; Yarali, 2020). We furthermore observed increased theta activity in noise after the behavioral response in parietal and occipital ROIs
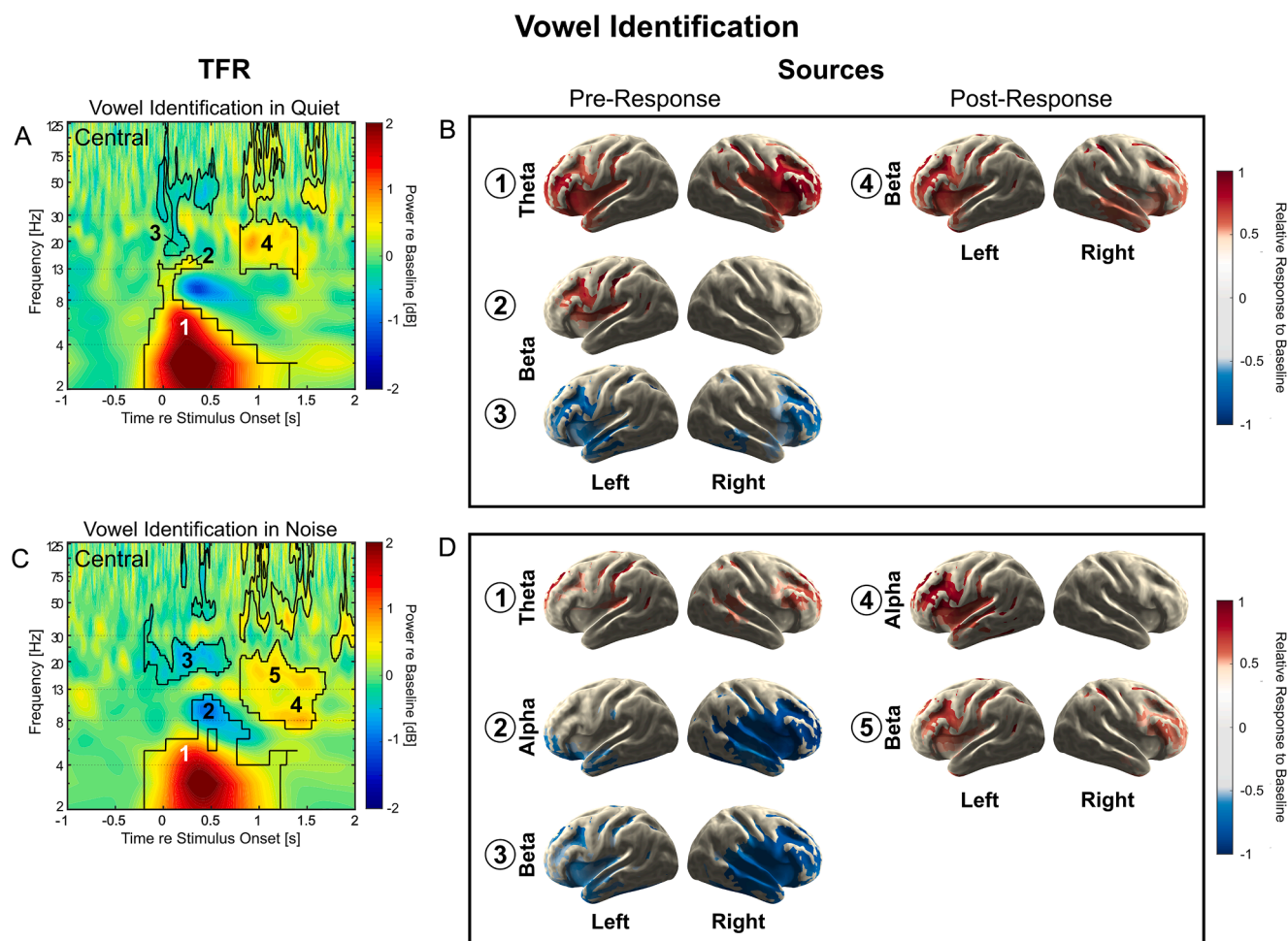
Vowel Identification

Fig. 7. Sources of oscillatory activities in the Vowel identification task. A. The central ROI TFRs for vowel identification in quiet with indices corresponding to the sources. B. Sources of oscillatory activities in vowel identification in quiet (see Table 3, for the nomenclature of the brain areas and their coordinates see Supplementary Table 1). C. The central ROI TFRs for vowel identification in noise with indices corresponding to the sources. D. Sources of oscillatory activities in vowel identification in noise (see Table 4, for the details see Supplementary Table 2). The specific time-frequency windows for the source localization of each oscillatory activity are listed in the Methods section (Table 1).

(Fig. 6K-L). Although this effect in the post-response period was not as robust as during cognitive load (e.g., compare Fig. 2F and J or N), increased theta suggests involvement of cognitive processing under sensory load (compare *4.2 Cognitive Load Effects*). The most relevant theta-related cognitive process under sensory load may be related to active listening, especially extraction of features from background noise (Alain et al., 2002; Ciocca, 2008; Tóth et al., 2016) and input comparison to target templates in working memory (Bastiaansen and Hagoort, 2003; Albouy et al., 2017; Nowak et al., 2021). Taken together, the present study suggests that there are theta effects of sensory load in the pre-response time period and to some degree in the post-response period.

In addition to theta effects, post-response alpha power was higher in noise (Fig. 2I-L). Such alpha power increase is likely a marker of functional inhibition of irrelevant information (Sauseng et al., 2005; Klimesch, 2012; Strauß et al., 2014; Wöstmann et al., 2015, 2017). The balance between top-down processes of template matching and functional inhibition (Gazzaley and Nobre, 2012; Obleser et al., 2012), reflected by theta and alpha oscillations, is modulated by attentional control (Klimesch et al., 1999; Kerlin et al., 2010; Keller et al., 2017; Fiebelkorn and Kastner, 2019; Cona et al., 2020). Such attentional control requires more cognitive resources when there are competing speakers (as in babble noise), thus increasing listening effort (Picou et al., 2016; Krueger et al., 2017; Dimitrijevic et al., 2019). Therefore,

listening in noise can be regarded as an attentional interplay between early sensory-perceptual processing, template matching in working memory, and functional inhibition of the distractor.

4.2. Cognitive load effects

The effect of cognitive load (Stroop incongruent compared to vowel identification) was manifested in theta. This difference in theta started within the pre-response period (in its second half) and outlasted the behavioral response (Fig. 3I-L). It is known that the processing of conflict information during the Stroop task occurs in initial perceptual and mainly in later post-perceptual stages (Lew et al., 1997; Boenke et al., 2009; Henkin et al., 2010). The current data further emphasize that cognitive processing lasts even beyond the behavioral response. This also confirms the role of theta as a common substrate for cognitive control (Cavanagh and Frank, 2014). While in general, theta is signaling cognitive control, there may be multiple types of executive functions that come into play at a given time point (Eisma et al., 2021; Xiao et al., 2023). The most relevant aspect of cognitive control in the current study is conflict processing (Hanslmayr et al., 2008; Ergen et al., 2014; Li et al., 2021; Heidlmayr et al., 2020; Sharma et al., 2021; Beldzik et al., 2022; see *4.3. Conflict Processing*).

In our interpretation, increased theta in Stroop incongruent compared to vowel identification is a signature of proactive control
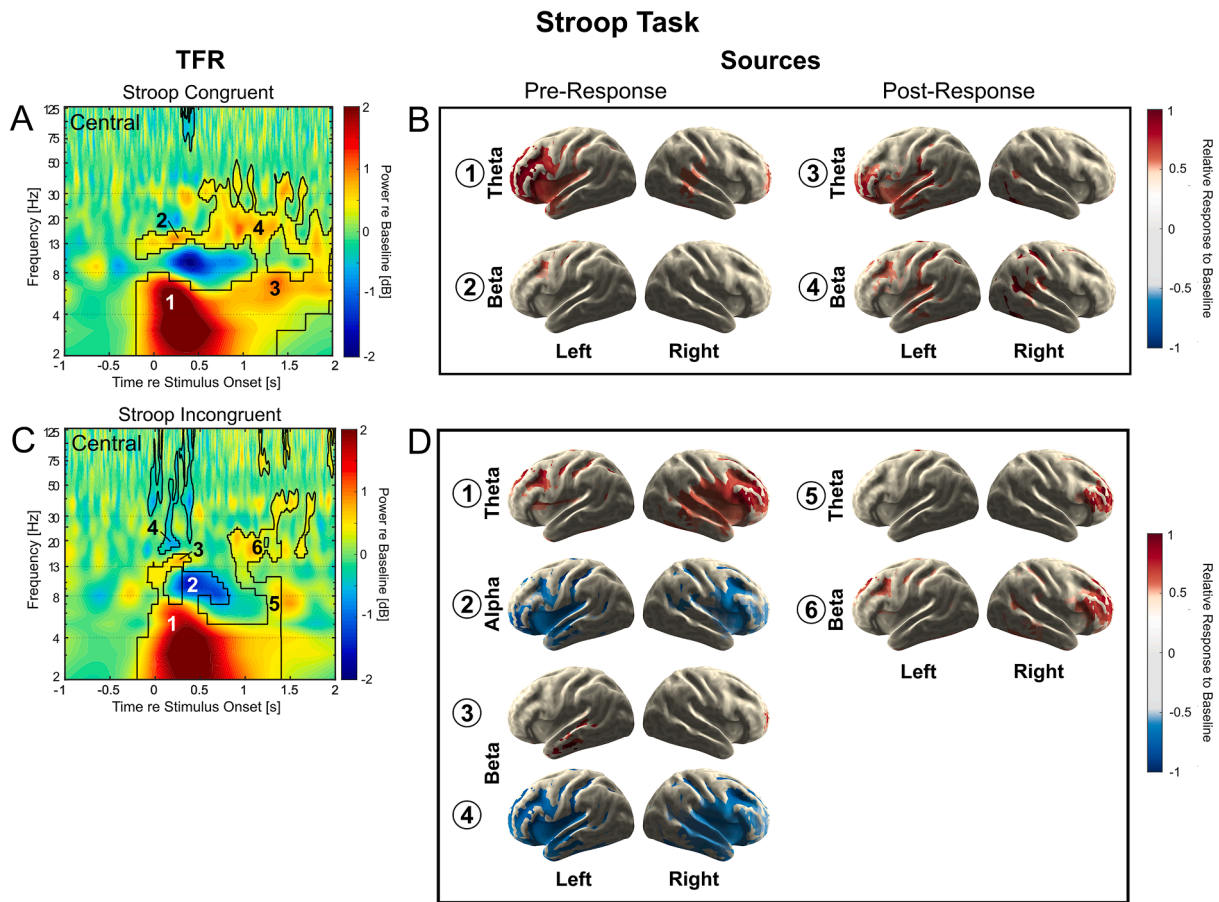
Fig. 8. Sources of oscillatory activities in the Stroop task. A. The central ROI TFRs for Stroop congruent with indices corresponding to the sources. B. Sources of oscillatory activities in Stroop congruent (see Table 5, for the nomenclature of the brain areas and their coordinates see Supplementary Table 3). C. The central ROI TFRs for Stroop incongruent with indices corresponding to the sources. D. Sources of oscillatory activities in Stroop incongruent (see Table 6, for the details see Supplementary Table 4). The specific time-frequency windows for the source localization of each oscillatory activity are listed in the Methods section (Table 1).

Table 3

List of source localization results in pre- and post-response periods for vowel identification in quiet. (For details of the coordinates, see Supplementary Table 1.1).

Vowel Identification in Quiet	
Frequency Band	Sources
1. Theta (pre-response)	Middle frontal gyrus (BA 9, 46)
	Inferior frontal gyrus (BA 44, 45)
	Superior temporal gyrus (BA 22, 38, 41, 42)
	Precentral gyrus (BA 6)
	Cingulate gyrus (BA 23)
	Superior parietal lobule (BA 5)
	Precuneus
2. Beta synch. (pre-response)	Postcentral gyrus (BA 1, 2, 3)
	Precuneus (BA 31)
	Inferior parietal lobule (BA 39, 40)
3. Beta desynch. (pre-response)	MedioVentral occipital cortex
	Cingulate gyrus (BA 23)
	Precuneus (BA31)
	Superior temporal gyrus (BA 22, 38)
4. Beta (post-response)	Cingulate gyrus (BA 23, 24)
	Superior frontal gyrus (BA 8)
	Precuneus (BA31)
	MedioVentral occipital cortex

(Cooper et al., 2015; van Driel et al., 2015; Littman et al., 2019; Eisma et al., 2021), in particular action monitoring in the post-response period (Cavanagh et al., 2012). Action monitoring is a crucial cognitive process for maintaining task requirements throughout the session and for

Table 4

List of source localization results in pre- and post-response periods for vowel identification in noise. (For details of the coordinates, see Supplementary Table 1.2).

Vowel Identification in Noise	
Frequency Band	Sources
1. Theta (pre-response)	Inferior parietal lobule (BA 39, 40)
	Superior frontal gyrus (BA 8, 9)
	MedioVentral occipital cortex
	Precuneus
2. Alpha desynch. (pre-response)	Superior temporal gyrus (BA 38, 41, 42)
	Inferior temporal gyrus (BA 37)
	Middle frontal gyrus (BA 9, 46)
	Inferior frontal gyrus (BA 44, 45)
	Precuneus (BA 31)
	Orbital gyrus (BA 11)
3. Beta desynch. (pre-response)	MedioVentral occipital cortex
	Precuneus (BA 31)
	Superior frontal gyrus (BA 8)
4. Alpha synch. (post-response)	Superior frontal gyrus (BA 8, 9, 10)
	Middle frontal gyrus (BA 9, 46)
	Inferior frontal gyrus (BA 44, 45)
	MedioVentral occipital cortex
5. Beta (post-response)	Superior frontal gyrus (BA 6, 8)
	Middle frontal gyrus (BA 9, 46)
	Inferior frontal gyrus (BA 44, 45)
	Precuneus (BA31)
	Cingulate gyrus (BA 23, 24)
	MedioVentral occipital cortex

Table 5

List of source localization results in pre- and post-response periods for Stroop congruent. (For details of the coordinates, see Supplementary Table 1.3).

Stroop Congruent	
Frequency Band	Sources
1. Theta (pre-response)	Middle frontal gyrus (BA 9, 46)
	Inferior frontal gyrus (BA 44, 45)
	Superior temporal gyrus (BA 38)
	Middle temporal gyrus
	Inferior parietal lobule (BA 40)
2. Beta synch. (pre-response)	Superior frontal gyrus (BA 8)
	Precentral gyrus (BA 6)
	Middle frontal gyrus
3. Theta (post-response)	Superior temporal gyrus (BA 38)
	Inferior frontal gyrus (BA 44, 45)
	Fusiform gyrus (BA 37)
	Orbital gyrus (BA 12, 47)
4. Beta (post-response)	Superior frontal gyrus (BA 8)
	Precentral gyrus A6cvl
	Middle frontal gyrus
	Cingulate gyrus (BA 23)
	Precuneus (BA31)
	MedioVentral occipital cortex
	Fusiform gyrus (BA 37)

Table 6

List of source localization results in pre- and post-response periods for Stroop incongruent. (For details of the coordinates, see Supplementary Table 1.4).

Stroop Incongruent	
Frequency Band	Sources
1. Theta (pre-response)	Superior frontal gyrus (BA 8, 9)
	Middle frontal gyrus (BA9, 46)
	Cingulate gyrus (BA 24, 32)
2. Alpha desynch. (pre-response)	Postcentral gyrus (BA 2)
	Precentral gyrus (BA 6)
	Inferior parietal lobule (BA 40)
	Middle temporal gyrus (BA 21)
	Inferior temporal gyrus (BA 20)
3. Beta synch. (pre-response)	Middle temporal gyrus (BA 21)
	Inferior temporal gyrus (BA 20, 37)
	Middle temporal gyrus (BA 37)
	Superior temporal gyrus (BA 38)
4. Beta desynch. (pre-response)	Precuneus (BA31)
	MedioVentral occipital cortex
5. Theta (post-response)	Paracentral lobule (BA 4)
	Middle frontal gyrus (BA9, 46)
6. Beta (post-response)	Superior frontal gyrus (BA 8, 9)
	Middle frontal gyrus (BA9, 46)
	Inferior frontal gyrus (BA 44)
	Cingulate gyrus (BA 23, 24)

planning of subsequent responses (Gratton et al., 2018; Cohen, 2016), and is sensitive to attentional modulation (Fiebelkorn and Kastner, 2019).

4.3. Conflict processing

The Stroop congruency effect, i.e., comparison between Stroop congruent and incongruent stimuli, is considered a gold-standard measure of selective attention and inhibition abilities (Stroop, 1935; Ergen et al., 2014; Kestens et al., 2021; Haciahmet et al., 2023). Previous studies in the visual modality demonstrated increased theta power during conflict processing (Kerns et al., 2004; Hanslmayr et al., 2008; Ergen et al., 2014). Reports regarding the auditory version of the Stroop task show similar effects, although to a much lesser extent (Oehrn et al., 2014; van de Nieuwenhuijzen et al., 2016).

In the present study, we observed increased theta power approximately 300 ms post-stimulus to incongruent compared to congruent, but this effect was not significant (Fig. 8I-J). The sensor level TFRs in this study suggested a more robust effect in a lower delta-theta power to incongruent vs. in congruent, occurring already in the pre-response time. In this version of the auditory Stroop task, the stimulus identification may have already occurred during the first syllable of the presented word (Henkin et al., 2002). Furthermore, the source reconstruction of theta suggested a hemisphere difference in the conflict processing (left in congruent and right in incongruent stimulation). This may relate to the specific difference in the first vowels that can be resolved by pitch comparison (differences in speaker's gender) more pronounced in the right hemisphere. In Stroop congruent, pitch and semantics are consistent, thus more pronounced in the left hemisphere (Zatorre and Belin, 2001). This suggests that the signature of conflict processing is crucially dependent on the exact stimulus.

Finally, the present data showed increased theta power to congruent than to incongruent stimuli in the post-response period (Fig. 8I-L), which have been similarly observed during intracranial recordings in frontal cortex in an auditory Stroop task (Oehrn et al., 2014). This theta difference is more likely related to a difference in proactive cognitive control more than to conflict processing per se (see *4.2, Cognitive Load Effects*).

4.4. Pre-Response beta and alpha, post-response beta, and gamma activity

In addition to the sensory/cognitive load effects, there were other observable oscillatory activities (Fig. 2). A comparison between the response time-locked (Fig. 5) and the stimulus time-locked TFRs (Fig. 2), revealed that while pre-response delta-theta activity was related to the stimulus, other activities were likely related to a combination of both stimulus- and response-components (Supplementary Fig. 2). Comparing the response- and the stimulus time-locked grand mean averages, the early component (i.e., N1 in stimulus time-locked data) was stimulus related, while the later component (i.e., P3, specifically P3b) was common to both stimulus- and response time-locked data (for details see: Polich, 2007). Finally, the second peak/component (Fig. 3A, E, I, M) shortly after the response in the response time-locked data may be primarily motor response-related (Fogarty et al., 2020).

Beta synchronization is known as signature of motor preparation (Spitzer and Haegens, 2017). We observed beta synchronization in the pre-response period (Fig. 2J-L, 3J-L) that may be related to motor function. Another cognitive process reflected in beta might be lexical retrieval (Signoret et al., 2013). Given the beta was also observable in the vowel identification task that did not require lexical retrieval, and not only in the Stroop task in the present setting, we assume that it likely reflects motor function. Along with pre-response beta, alpha desynchronization occurring around the response time is also observable in both stimulus- and response time-locked TFRs, especially in central ROI (Fig. 2B, F, J, N and 5B, F, J, N). Alpha desynchronization might be related to attentional control (Sauseng et al., 2005; Klimesch, 2012; Wöstmann et al., 2015, 2017; Sharma et al., 2021) and beta activity to the motor preparation (Spitzer and Haegens, 2017).

In the post-response period, beta synchronization lasted around a second and was strongest in frontal ROI (Fig. 2A, E, I, M and 5A, E, I, M). This might be further related to motor function (Pfurtscheller and Da Silva, 1999; Engel and Fries, 2010). This beta activity was observed in both stimulus- and response time-locked TFRs, indicating the role of frontal circuitry in integrating sensory-related and response-related aspects. A similar increase in beta power, occurring as repeating bursts, has been documented during working memory operation (Siegel et al., 2009; Lundqvist et al., 2016), likely reflecting preservation and updating of the current information in working memory (Engel and Fries, 2010; Spitzer and Haegens, 2017; Coleman et al., 2023).

Gamma transients were also identified during the whole trial duration (i.e., Fig. 2A). Gamma activity reflects both sensory (Schadow et al., 2007) and cognitive processing (Herrmann et al., 2010). It has been recorded oftentimes in invasive studies (Fontolan et al., 2014; Nourski et al., 2022), and also during the Stroop task (Oehrn et al., 2014; Tang

et al., 2016; van de Nieuwenhuijzen et al., 2016). In auditory cortex (Yusuf et al., 2017; Nourski et al., 2022), gamma-band responses were found along with alpha activity (Weisz et al., 2011), suggesting alpha-gamma coupling (see also Roux and Uhlhaas, 2014). Due to their transient characteristics, significant effects in gamma band were rare in our scalp data, but the data suggests dominant occurrence of gamma desynchronizations during the pre-response period and gamma synchronizations in the post-response period.

4.5. Source localization

While topographic plots provide some spatial information of the EEG activity, their sources can be further estimated using beamforming (Debener et al., 2005; Hauthal et al., 2013). Sources of N1 covered the expected auditory cortex bilaterally, but also included sources beyond auditory cortex (Supplementary Fig. 4). Sources of P3 were previously reported to be localized in superior temporal, medial frontal- and inferior frontal gyrus (Knight et al., 1995; Opitz et al., 2002; Doeller et al., 2003; Garrido et al., 2009), corresponding to the present results (Supplementary Fig. 4). Significant activities from the TFRs (Fig. 3, 4) were also localized to temporal and frontal lobes, as well as cingulate gyrus and precuneus.

The activation of temporal lobe was expected, given the tasks were auditory, while frontal lobe activation has been attributed to executive control and working memory (Braver, 2012; Cristofori et al., 2019). Involvement of cingulate gyrus is related to conflict processing (Haupt et al., 2009; Christensen et al., 2011), action selection (Akam et al., 2021) and translation of intentions into motor actions (Hoffstaedter et al., 2014; Holroyd and Verguts, 2021). The connection between cingulate gyrus and dorsolateral prefrontal cortex has been reported in conflict tasks (Kerns et al., 2004, see also Oehrn et al., 2014). Similarly, theta related coupling between anterior cingulate and left prefrontal cortex was also found in a previous Stroop task study (Hanslmayr et al., 2008). This connection between cingulate gyrus and frontal lobe might be part of the saliency networks (Seeley, 2019; Uddin et al., 2019; Uddin, 2021), which activated to route information further to the motor cortex.

Precuneus (Cavanna and Trimble, 2006) was a main source of the oscillatory activities in this study. This area is known to be involved in memory functions (Baird et al., 2013; Ye et al., 2018) and information integration (Lyu et al., 2021). Recently, proactive control has been similarly associated with activation of both precuneus as well as cingulate cortex (Sznabel et al., 2023). In sum, the active network found in this study is consistent with areas known to be responsible for sensory processing of auditory stimulus as well as higher-level cognitive functions, such as memory, action monitoring, and conflict processing.

4.6. Limitation and future directions

The present study shows that sensory- and cognitive loads specifically modulate brain processing before and after the behavioral response. However, we did not modify the demand for load processing (e.g., working memory- or conflict loads or demand for more attentional control), nor studied the linguistic processing as well the circuitry of these oscillatory activities. Such studies, together with investigations of the effects of declined cognitive performance and sensory processing, as present in aging and hearing impairment, are crucial to apply the full potential of these oscillatory markers.

5. Conclusion

Oscillatory activities allow differentiating listening under auditory sensory and cognitive load. Sensory load initially reduced theta-activity and increased alpha activity. Cognitive load, on the other hand, increased theta power over long periods of time after behavioral response. In this post-response period, beta activity was observed in all conditions and tasks. Combining the results of the present study with previous literatures suggests that the increased alpha power in sensory load signifies an attention-modulated inhibitory process and the theta increase in cognitive load reflects conflict processing, proactive control, and action monitoring. The sources of the corresponding neuronal activities included the frontal lobe, temporal lobe, motor cortex, cingulate gyrus and precuneus. Together with their sources, the observed oscillatory activities in this study could be exploited as signatures/biomarkers of the brain processing related to auditory sensory and cognitive load in complex listening situations.

Data/Code availability statement

MATLAB scripts are available at: https://github.com/brllant

CRediT authorship contribution statement

Brilliant: Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Y. Yaar-Soffer:** Conceptualization, Investigation, Data curation, Writing – review & editing. **C.S. Herrmann:** Data curation, Supervision, Writing – review & editing. **Y. Henkin:** Conceptualization, Investigation, Resources, Supervision, Writing – review & editing. **A. Kral:** Conceptualization, Visualization, Resources, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare no competing financial interest.

Data availability

I have shared the link to my code at the "Attach File" Step

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2024.120546.

References

Akam, T., Rodrigues-Vaz, I., Marcelo, I., Zhang, X., Pereira, M., Oliveira, R.F., Dayan, P., Costa, R.M., 2021. The anterior cingulate cortex predicts future states to mediate model-based action selection. Neuron 109, 149–163 e7.

Alain, C., Schuler, B.M., McDonald, K.L., 2002. Neural activity associated with distinguishing concurrent auditory objects. J. Acoust. Soc. Am. 111, 990–995.

Albouy, P., Weiss, A., Baillet, S., Zatorre, R.J., 2017. Selective entrainment of theta oscillations in the dorsal stream causally enhances auditory working memory performance. Neuron 94, 193–206 e5.

Baird, B., Smallwood, J., Gorgolewski, K.J., Margulies, D.S., 2013. Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. J. Neurosci. 33, 16657–16665.

Başar, E., Schürmann, M., Demiralp, T., Başar-Eroglu, C., Ademoglu, A., 2001. Event-related oscillations are 'real brain responses'—Wavelet analysis and new strategies. Int. J. Psychophysiol. 39, 91–127.

Bastiaansen, M., Hagoort, P., 2003. Event-induced theta responses as a window on the dynamics of memory. Cortex 39, 967–992.

Beldzik, E., Ullsperger, M., Domagalik, A., Marek, T., 2022. Conflict- and error-related theta activities are coupled to BOLD signals in different brain regions. Neuroimage 256, 119264.

Beres, A.M., 2017. Time is of the essence: a review of electroencephalography (EEG) and event-related brain potentials (ERPs) in language research. Appl. Psychophysiol. Biofeedback 42, 247–255.

Boenke, L.T., Ohl, F.W., Nikolaev, A.R., Lachmann, T., Leeuwen, C., 2009. Different time courses of Stroop and Garner effects in perception–an event-related potentials study. Neuroimage 45, 1272–1288.

Braver, T.S., 2012. The variable nature of cognitive control: a dual mechanisms framework. Trends Cogn. Sci. (Regul. Ed.) 16, 106–113.

Buzsaki, G., 2006. Rhytms of the Brain. Oxford University Press, Oxford.

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. (Regul. Ed.) 18, 414–421.

Cavanagh, J.F., Zambrano-Vazquez, L., Allen, J.J., 2012. Theta lingua franca: a common mid-frontal substrate for action monitoring processes. Psychophysiology 49, 220–238.

Cavanna, A.E., Trimble, M.R., 2006. The precuneus: a review of its functional anatomy and behavioural correlates. Brain 129, 564–583.

Chen, C.-C., Kiebel, S.J., Kilner, J.M., Ward, N.S., Stephan, K.E., Wang, W.-J., Friston, K. J., 2012. A dynamic causal model for evoked and induced responses. Neuroimage 59, 340–348.

Christensen, T.A., Lockwood, J.L., Almryde, K.R., Plante, E., 2011. Neural substrates of attentive listening assessed with a novel auditory Stroop task. Front. Hum. Neurosci. 4, 236.

Ciocca, V., 2008. The auditory organization of complex sounds. Front. Biosci. 13, 148–169.

Cohen, M.X., 2014. Analyzing Neural Time Series Data. MIT Press, Cambridge, Mass.

Coleman, S.C., Seedat, Z.A., Whittaker, A.C., Lenartowicz, A., Mullinger, K.J., 2023. Beyond the beta rebound: post-task responses in oscillatory activity follow cessation of working memory processes. Neuroimage 265, 119801.

Cona, G., Chiossi, F., Di Tomasso, S., Pellegrino, G., Piccione, F., Bisiacchi, P., Arcara, G., 2020. Theta and alpha oscillations as signatures of internal and external attention to delayed intentions: a magnetoencephalography (MEG) study. Neuroimage 205, 116295.

Cooke, M., Garcia Lecumberri, M.L., Barker, J., 2008. The foreign language cocktail party problem: energetic and informational masking effects in non-native speech perception. J. Acoust. Soc. Am. 123, 414–427.

Cooper, P.S., Wong, A.S., Fulham, W.R., Thienel, R., Mansfield, E., Michie, P.T., Karayanidis, F., 2015. Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes. Neuroimage, 108, 354–363.

Cristofori, I., Cohen-Zimerman, S., Grafman, J., 2019. Executive functions. Handb. Clin. Neurol. 163, 197–219.

David, O., Kilner, J.M., Friston, K.J., 2006. Mechanisms of evoked and induced responses in MEG/EEG. Neuroimage 31, 1580–1591.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D.Y., Engel, A.K., 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J. Neurosci. 25, 11730–11737.

Decruy, L., Vanthornhout, J., Francart, T., 2019. Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. J. Neurophysiol. 122, 601–615.

Demiralp, T., Ademoglu, A., Comerchero, M., Polich, J., 2001. Wavelet analysis of P3a and P3b. Brain Topogr. 13, 251–267.

Dimitrijevic, A., Smith, M.L., Kadis, D.S., Moore, D.R., 2019. Neural indices of listening effort in noisy environments. Sci. Rep. 9, 11278.

Dimitrijevic, A., Smith, M.L., Kadis, D.S., Moore, D.R., 2017. Cortical Alpha Oscillations Predict Speech Intelligibility. Front. Hum. Neurosci. 11, 88.

Doeller, C.F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., Schröger, E., 2003. Prefrontal cortex involvement in preattentive auditory deviance detection:: neuroimaging and electrophysiological evidence. Neuroimage 20, 1270–1282.

Donner, T.H., Siegel, M., 2011. A framework for local cortical oscillation patterns. Trends. Cogn. Sci. 15, 191–199.

Eisma, J., Rawls, E., Long, S., Mach, R., Lamm, C., 2021. Frontal midline theta differentiates separate cognitive control strategies while still generalizing the need for cognitive control. Sci. Rep. 11, 14641.

Engel, A.K., Fries, P., 2010. Beta-band oscillations–signalling the status quo? Curr. Opin. Neurobiol. 20, 156–165.

Ergen, M., Saban, S., Kirmizi-Alsan, E., Uslu, A., Keskin-Ergen, Y., Demiralp, T., 2014. Time-frequency analysis of the event-related potentials associated with the Stroop test. Int. J. Psychophysiol. 94, 463–472.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., 2016. The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb. Cortex 26, 3508–3526.

Fiebelkorn, I.C., Kastner, S., 2019. A rhythmic theory of attention. Trends Cogn. Sci. (Regul. Ed.) 23, 87–101.

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., Giraud, A.-L., 2014. The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. Nat. Commun. 5, 4694.

Fogarty, J.S., Robert, J.B., Genevieve, Z.S., 2020. Auditory stimulus- and response-locked ERP components and behavior. Psychophysiology 57, e13538.

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., Ebersole, J.S., 2002. A standardized boundary element method volume conductor model. Clin. Neurophysiol. 113, 702–712.

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. 120, 453–463.

Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: bridging selective attention and working memory. Trends. Cogn. Sci. 16, 129–135.

Getzmann, S., Falkenstein, M., 2011. Understanding of spoken language under challenging listening conditions in younger and older listeners: a combined behavioral and electrophysiological study. Brain Res. 1415, 8–22.

Getzmann, S., Wascher, E., Falkenstein, M., 2015. What does successful speech-in-noise perception in aging depend on? Electrophysiological correlates of high and low performance in older adults. Neuropsychologia 70, 43–57.

Gratton, G., Cooper, P., Fabiani, M., Carter, C.S., Karayanidis, F., 2018. Dynamics of cognitive control: theoretical bases, paradigms, and a view for the future. Psychophysiology 55, e13016.

Green, E.J., Barber, P.J., 1981. An auditory Stroop effect with judgments of speaker gender. Percept. Psychophys. 30, 459–466.

Gregg, M.K., Purdy, K.A., 2007. Graded auditory Stroop effects generated by gender words. Percept. Mot. Skills 105, 549–555.

Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., Salmelin, R., 2001. Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc. Natl. Acad. Sci. 98, 694–699.

Haciahmet, C.C., Frings, C., Beste, C., Münchau, A., Pastötter, B., 2023. Posterior delta/theta EEG activity as an early signal of Stroop conflict detection. Psychophysiology 60, e14195.

Hamalainen, M.S., Sarvas, J., 1987. Feasibility of the homogeneous head model in the interpretation of neuromagnetic fields. Phys. Med. Biol. 32, 91.

Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., Klimesch, W., 2008. The electrophysiological dynamics of interference during the Stroop task. J. Cogn. Neurosci. 20, 215–225.

Haupt, S., Axmacher, N., Cohen, M.X., Elger, C.E., Fell, J., 2009. Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. Hum. Brain Mapp. 30, 3043–3056.

Hauthal, N., Thorne, J.D., Debener, S., Sandmann, P., 2013. Source localisation of visual evoked potentials in congenitally deaf individuals. Brain Topogr. 27, 412–424.

Heidlmayr, K., Kihlstedt, M., Isel, F., 2020. A review on the electroencephalography markers of Stroop executive control processes. Brain Cogn. 146, 105637.

Henkin, Y., Kishon-Rabin, L., Gadoth, N., Pratt, H., 2002. Auditory event-related potentials during phonetic and semantic processing in children. Audiol. Neurotol. 7, 228–239.

Henkin, Y., Yaar-Soffer, Y., Gilat, S., Muchnik, C., 2010. Auditory conflict processing: behavioral and electrophysiologic manifestations of the Stroop effect. J. Am. Acad. Audiol. 21, 474–486.

Henkin, Y., Yaar-Soffer, Y., Steinberg, M., Muchnik, C., 2014. Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. Audiol. Neurotol. 19, 21–26.

Herrmann, C.S., Munk, M.H.J., Engel, A.K., 2004. Cognitive functions of gamma-band activity: memory match and utilization. Trends. Cogn. Sci. 8, 347–355.

Herrmann, C.S., Fründ, I., Lenz, D., 2010. Human gamma-band activity: a review on cognitive and behavioral correlates and network models. Neurosci. Biobehav. Rev. 34, 981–992.

Hickok, G., Farahbod, H., Saberi, K., 2015. The rhythm of perception: entrainment to acoustic rhythms induces subsequent perceptual oscillation. Psychol. Sci. 26, 1006–1013.

Hillyard, S.A., Kutas, M., 1983. Electrophysiology of cognitive processing. Annu. Rev. Psychol. 34, 33–61.

Hjortkjær, J., Märcher-Rørsted, J., Fuglsang, S.A., Torsten, D., 2020. Cortical oscillations and entrainment in speech processing during working memory load. Eur. . Neurosci. 51, 1279–1289.

Hoffstaedter, F., Grefkes, C., Caspers, S., Roski, C., Palomero-Gallagher, N., Laird, A.R., Fox, P.T., Eickhoff, S.B., 2014. The role of anterior midcingulate cortex in cognitive motor control: evidence from functional connectivity analyses. Hum. Brain Mapp. 35, 2741–2753.

Holroyd, C.B., Verguts, T., 2021. The best laid plans: computational principles of anterior cingulate cortex. Trends Cogn. Sci. (Regul. Ed.) 25, 316–329.

Hsiao, F.J., Wu, Z.A., Ho, L.T., Lin, Y.Y., 2009. Theta oscillation during auditory change detection: an MEG study. Biol. Psychol. 81, 58–66.

Huster, R.J., Plis, S.M., Lavallee, C.F., Calhoun, V.D., Herrmann, C.S., 2014. Functional and effective connectivity of stopping. Neuroimage 94, 120–128.

Jurcak, V., Tsuzuki, D., Dan, I., 2007. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. Neuroimage 34, 1600–1611.

Kalcher, J., Pfurtscheller, G., 1995. Discrimination between phase-locked and non-phase-locked event-related EEG activity. Electroencephalogr. Clin. Neurophysiol. 94, 381–384.

Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., Muchnik, C., 2006. Identification of syllables in noise: electrophysiological and behavioral correlates. J. Acoust. Soc. Am. 120, 926–933.

Keller, A.S., Payne, L., Sekuler, R., 2017. Characterizing the roles of alpha and theta oscillations in multisensory attention. Neuropsychologia 99, 48–63.

Kerlin, J.R., Shahin, A.J., Miller, L.M., 2010. Attentional gain control of ongoing cortical speech representations in a "cocktail party. J. Neurosci. 30, 620–628.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior cingulate conflict monitoring and adjustments in control. Science 303, 1023–1026.

Kestens, K., Degeest, S., Miatton, M., 2021. An auditory Stroop test to implement in cognitive hearing sciences: development and normative data. Int. J. Psychol. Res. (Medellin) 14, 37–51.

Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn. Sci. (Regul. Ed.) 16, 606–617.

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., Winkler, T.H., 1999. 'Paradoxical' alpha synchronization in a memory task. Cogn. Brain Res. 7, 493–501.

Knight, R.T., Grabowecky, M.F., Scabini, D., 1995. Role of human prefrontal cortex in attention control. Adv. Neurol. 66, 21–34 discussion 34.

Krueger, M., Schulte, M., Zokoll, M.A., Wagener, K.C., Meis, M., Brand, T., Holube, I., 2017. Relation between listening effort and speech intelligibility in noise. Am. J. Audiol. 26, 378–392.

Lakatos, P., O'Connell, M.N., Barczak, A., Mills, A., Javitt, D.C., Schroeder, C.E., 2009. The leading sense: supramodal control of neurophysiological context by attention. Neuron 64, 419–430.

Lew, H., Chmiel, R., Jerger, J., Pomerantz, J.R., Jerger, S., 1997. Electrophysiologic indices of Stroop and Garner interference reveal linguistic influences on auditory and visual processing. J. Am. Acad. Audiol. 8, 104–118.

Li, Z., Yang, G., Wu, H., Li, Q., Xu, H., Göschl, F., Nolte, G., Liu, X., 2021. Modality-specific neural mechanisms of cognitive control in a Stroop-like task. Brain Cogn. 147, 105662.

Lieder, F., Iwama, G., 2021. Toward a formal theory of proactivity. Cogn., Affect., Behav. Neurosci. 21, 490–508.

Littman, R., Keha, E., Kalanthroff, E., 2019. Task conflict and task control: a mini-review. Front. Psychol. 10, 1598.

Legouhy A. (2023) al_goodplot - boxblot & violin plot (https://www.mathworks.com/matlabcentral/fileexchange/91790-al_goodplot-boxblot-violin-plot), MATLAB Central File Exchange. Retrieved November 5, 2023.

Lundqvist, M., Rose, J., Herman, P., Brincat, S.L., Buschman, T.J., Miller, E.K., 2016. Gamma and beta bursts underlie working memory. Neuron 90, 152–164.

Lyu, D., Pappas, I., Menon, D.K., Stamatakis, E.A., 2021. A precuneal causal loop mediates external and internal information integration in the human brain. J. Neurosci. 41, 9944–9956.

MacLeod, C.M., 1992. The Stroop task: the" gold standard" of attentional measures. J. Exp. Psychol. 121, 12.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190.

Martin, B.A., Tremblay, K.L., Korczak, P., 2008. Speech evoked potentials: from the laboratory to the clinic. Ear Hear. 29, 285–313.

Mattys, S.L., Davis, M.H., Bradlow, A.R., Scott, S.K., 2012. Speech recognition in adverse conditions: a review. Lang. Cogn. Process 27, 953–978.

McCullagh, J., Shinn, J.B., 2013. Auditory cortical processing in noise in younger and older adults. Hear., Balance Commun. 11, 182–190.

McHaney, J.R., Gnanateja, G.N., Smayda, K.E., Zinszer, B.D., Chandrasekaran, B., 2021. Cortical tracking of speech in delta band relates to individual differences in speech in noise comprehension in older adults. Ear Hear. 42, 343–354.

Morgan, A.L.R., Brandt, J.F., 1989. An auditory Stroop effect for pitch, loudness, and time. Brain Lang. 36, 592–603.

Niemczak, C.E., Vander Werff, K.R., 2019. Informational masking effects on neural encoding of stimulus onset and acoustic change. Ear Hear. 40, 156–167.

Nourski, K.V., Steinschneider, M., Rhone, A.E., Kovach, C.K., Kawasaki, H., Howard, M. A., 2022. Gamma activation and alpha suppression within human auditory cortex during a speech classification task. J. Neurosci. 42, 5034–5046.

Nowak, K., Costa-Faidella, J., Dacewicz, A., Escera, C., Szelag, E., 2021. Altered event-related potentials and theta oscillations index auditory working memory deficits in healthy aging. Neurobiol. Aging 108, 1–15.

Obleser, J., Weisz, N., 2012. Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. Cereb. Cortex. 22, 2466–2477.

Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., Maess, B., 2012. Adverse listening conditions and memory load drive a common alpha oscillatory network. J. Neurosci. 32, 12376–12383.

Oehrn, C.R., Hanslmayr, S., Fell, J., Deuker, L., Kremers, N.A., Do Lam, A.T., Elger, C.E., Axmacher, N., 2014. Neural communication patterns underlying conflict detection, resolution, and adaptation. J. Neurosci. 34, 10438–10452.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Oostendorp, T., Van Oosterom, A., 1991. The potential distribution generated by surface electrodes in inhomogeneous volume conductors of arbitrary shape. IEEE Trans. Biomed. Eng. 38, 409–417.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869.

Opitz, B., Rinne, T., Mecklinger, A., Von Cramon, D.Y., Schröger, E., 2002. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. Neuroimage 15, 167–174.

Pfurtscheller, G., Da Silva, F.H.L., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. 110, 1842–1857.

Picou, E.M., Gordon, J., Ricketts, T.A., 2016. The effects of noise and reverberation on listening effort in adults with normal hearing. Ear Hear. 37, 1–13.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118, 2128–2148.

Pratt H. (2011) Sensory ERP components. The Oxford handbook of event-related potential components, 89–114.

Rajan, R., Cainer, K.E., 2008. Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. Neuroscience 154, 784–795.

Riecke, L., Esposito, F., Bonte, M., Formisano, E., 2009. Hearing illusory sounds in noise: the timing of sensory-perceptual transformations in auditory cortex. Neuron 64, 550–561.

Rousselet, G.A., 2012. Does filtering preclude us from studying ERP time-courses. Front. Psychol. 3, 131.

Roux, F., Uhlhaas, P.J., 2014. Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information. Trends Cogn. Sci. (Regul. Ed.) 18, 16–25.

Sauseng, P., Klimesch, W., Schabus, M., Doppelmayr, M., 2005. Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. Int. J. Psychophysiol. 57, 97–103.

Schadow, J., Lenz, D., Thaerig, S., Busch, N., Frund, I., Herrmann, C., 2007. Stimulus intensity affects early sensory processing: sound intensity modulates auditory evoked gamma-band activity in human EEG. Int. J. Psychophysiol. 65, 152–161.

Schneider, B.A., 2011. How age affects auditory-cognitive interactions in speech comprehension. Audiol. Res. 1, e10.

Schneider, B.A., Pichora-Fuller, K., Daneman, M., 2010. Effects of senescent changes in audition and cognition on spoken language comprehension. Aging Auditory Syst. 167–210.

Seeley, W.W., 2019. The salience network: a neural system for perceiving and responding to homeostatic demands. J. Neurosci. 39, 9878–9882.

Sharma, V.V., Thaut, M., Russo, F., Alain, C., 2019. Absolute pitch and musical expertise modulate neuro-electric and behavioral responses in an auditory Stroop paradigm. Front. Neurosci. 13, 932.

Sharma, V.V., Thaut, M., Russo, F.A., Alain, C., 2021. Neural dynamics of inhibitory control in musicians with absolute pitch: theta synchrony as an oscillatory signature of information conflict. Cereb. Cortex. Commun. 2, tgab043.

Siegel, M., Donner, T.H., Engel, A.K., 2012. Spectral fingerprints of large-scale neuronal interactions. Nat. Rev. Neurosci. 13, 121–134.

Siegel, M., Warden, M.R., Miller, E.K., 2009. Phase-dependent neuronal coding of objects in short-term memory. Proc. Natl. Acad. Sci. U S A 106, 21341–21346.

Signoret, C., Gaudrain, E., Perrin, F., 2013. Similarities in the neural signature for the processing of behaviorally categorized and uncategorized speech sounds. Eur. J. Neurosci. 37, 777–785.

Sperry, J.L., Wiley, T.L., Chial, M.R., 1997. Word recognition performance in various background competitors. J.-Am. Acad. Audiol. 8, 71–80.

Spitzer, B., Haegens, S., 2017. Beyond the status quo: a role for beta oscillations in endogenous content (re) activation. eNeuro 4.

Strauß, A., Wöstmann, M., Obleser, J., 2014. Cortical alpha oscillations as a tool for auditory selective inhibition. Front. Hum. Neurosci. 8, 350.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18, 643.

Sznabel, D., Land, R., Kopp, B., Kral, A., 2023. The relation between implicit statistical learning and proactivity as revealed by EEG. Sci. Rep. 13, 15787.

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011, 1–13.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40Hz visual responses in human. J. Neurosci. 16, 4240–4249.

Tang, H., Yu, H.-Y., Chou, C.-C., Crone, N.E., Madsen, J.R., Anderson, W.S., Kreiman, G., 2016. Cascade of neural processing orchestrates cognitive control in human frontal cortex. Elife 5, e12352.

Tóth, B., Kocsis, Z., Urbán, G., Winkler, I., 2016. Theta oscillations accompanying concurrent auditory stream segregation. Int. J. Psychophysiol. 106, 141–151.

Uddin, L.Q., 2021. Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. Nat. Rev. Neurosci. 22, 167–179.

Uddin, L.Q., Yeo, B.T.T., Spreng, R.N., 2019. Towards a universal taxonomy of macro-scale functional human brain networks. Brain Topogr. 32, 926–942.

van de Nieuwenhuijzen, M.E., Axmacher, N., Fell, J., Oehrn, C.R., Jensen, O., van Gerven, M.A.J., 2016. Decoding of task-relevant and task-irrelevant intracranial EEG representations. Neuroimage 137, 132–139.

van Driel, J., Swart, J.C., Egner, T., Ridderinkhof, K.R., Cohen, M.X., 2015. No) time for control: frontal theta dynamics reveal the cost of temporally guided conflict anticipation. Cogn. Affect. Behav. Neurosci. 15, 787–807.

Van Veen, B.D., Van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880.

Wechsler, D., 1997. Wechsler Adult Intelligence Scale-III. Psychological Corp, London.

Weisz, N., Hartmann, T., Müller, N., Lorenz, I., Obleser, J., 2011. Alpha rhythms in audition: cognitive and clinical perspectives. Front. Psychol. 2, 73.

Westner, B.U., Dalal, S.S., Gramfort, A., Litvak, V., Mosher, J.C., Oostenveld, R., Schoffelen, J.M., 2022. A unified view on beamformers for M/EEG source reconstruction. Neuroimage 246, 118789.

Wilsch, A., Obleser, J., 2016. What works in auditory working memory? A neural oscillations perspective. Brain Res.

Wilsch, A., Henry, M.J., Herrmann, B., Maess, B., Obleser, J., 2015. Alpha oscillatory dynamics index temporal expectation benefits in working memory. Cereb. Cortex. 25, 1938–1946.

Winkler, I., Debener, S., Müller, K.-R., Tangermann, M., 2015. On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4101–4105.

Wisniewski, M.G., Zakrzewski, A.C., Bell, D.R., Wheeler, M., 2021. EEG power spectral dynamics associated with listening in adverse conditions. Psychophysiology 58, e13877.

Wöstmann, M., Herrmann, B., Wilsch, A., Obleser, J., 2015. Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. J. Neurosci. 35, 1458–1467.

Wöstmann, M., Lim, S.-J., Obleser, J., 2017. The human neural alpha response to speech is a proxy of attentional control. Cereb. Cortex. 27, 3307–3317.

Xiao, Y., Chou, C.C., Cosgrove, G.R., Crone, N.E., Stone, S., Madsen, J.R., Reucroft, I., Shih, Y.C., Weisholtz, D., Yu, H.Y., Anderson, W.S., Kreiman, G., 2023. Cross-task specificity and within-task invariance of cognitive control processes. Cell Rep. 42, 111919.

Yarali, M., 2020. Varying effect of noise on sound onset and acoustic change evoked auditory cortical N1 responses evoked by a vowel-vowel stimulus. Int. J. Psychophysiol. 152, 36–43.

Ye, Q., Zou, F., Lau, H., Hu, Y., Kwok, S.C., 2018. Causal evidence for mnemonic metacognition in human precuneus. J. Neurosci. 38, 6379–6387.

Yusuf, P.A., Hubka, P., Tillein, J., Kral, A., 2017. Induced cortical responses require developmental sensory experience. Brain 140, 3153–3165.

Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11, 946–953.