Biomedical Signal Processing and Control 120 (2026) 110149

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Unleashing the potential of OPM-MEG to study event-related fields against low-frequency artifacts: the case of sentence processing

Xinchi Yu [a,b,c,d,1,*], Dongxue Zhang [c,d,1], Thomas Carey [d], Ernst Pöppel [d,e], Ellen Lau [a,b,2], Tilmann Sander [c,2]

[a] Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA
[b] Department of Linguistics, University of Maryland, College Park, MD, USA
[c] Physikalisch-Technische Bundesanstalt, Berlin, Germany
[d] Institute of Medical Psychology, Ludwig Maximilian University Munich, Munich, Germany
[e] Neurocognitive Image Lab, Shanghai International Studies University, Shanghai, China

ARTICLE INFO

Keywords:
OPM-MEG
N400
Sentence processing
DSS (denoising source separation)

ABSTRACT

OPM-MEG (optically pumped magnetometers-magnetoencephalography) offers unprecedented opportunity in its proximity to the brain and in mobility, outperforming other human MEG systems. However, movements induce low-frequency artifacts (up to ~3 Hz) in this system, calling for pipelines that could efficiently reduce these low-frequency noises. Although a high-pass filter of e.g., 4 Hz may minimize these artifacts, it may also eliminate many event-related field (ERF) components, such as the N400 response in sentence processing studies. Moreover, as an emerging technology with expensive sensors, many labs are starting with an experimental setup with fewer sensors (e.g., ~10), making it challenging to reject principal components based on visually inspecting the topographic field maps. In the current paper, we show that a combination of moderate high-pass filtering (1 Hz) and evoked-biased denoising source separation (evoked-biased DSS) can effectively reveal typical N400 deflections and effects (between nouns and verbs), with a nine-sensor OPM-MEG setup. Our current pipeline paves the way to studying ERFs with OPM-MEG in more budget-friendly setups with a small number of sensors.

1. Introduction

Magnetoencephalography (MEG) is a non-invasive neuroimaging technology with both high spatial and temporal resolution. MEG studies therefore offer important complementary evidence in the field of cognitive neuroscience, compared to e.g., EEG (electroencephalography; with good temporal resolution but bad spatial resolution) and fMRI (functional magnetic resonance imaging; with good spatial resolution but bad temporal resolution). MEG also possesses largely underexplored clinical potential including e.g., the diagnosis of concussion [1,2] and dyslexia [3–5]. MEG sites existing for more than 10 years typically make use of SQUID-MEG systems (SQUID: superconducting quantum interference device), but recent years have seen an increasing transition to OPM-MEG systems (OPM: optically pumped magnetometers), which possess several properties that are more desirable than SQUID-MEG systems [6–14]. First, OPM-MEG systems do not require the constant liquid helium supply necessary for SQUIDs, which is expensive and is easily affected by helium shortages. Second, OPM sensors can be much closer to the scalp compared to SQUIDs, allowing for higher signal amplitudes (which is critical for detecting the faint magnetic signals from the brain, generally at the scale of femtoteslas [fT], 10^{-15} T). For example, in the current study the OPM sensors were attached to an EEG cap, which is much closer to neural sources in the brain compared to SQUIDs, which are encapsulated in a dewar that cannot fit flexibly to the scalp (in this case the minimum distance to cortical sources would be ~2–4 cm [15,16]). Third, most current OPM-MEG systems are much more mobile than SQUID-MEG systems, which may allow for mild movements during data collection. Mild movements are generally inevitable, especially for certain populations such as young children, positing challenges to the inclusion of more populations for SQUID-MEG systems. Moreover, a mobile setup allows for MEG recordings in more naturalistic settings.

* Corresponding author at: Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA.
E-mail addresses: xcyu@umd.edu, xinchiyuneuro@gmail.com (X. Yu).
[1] Xinchi Yu and Dongxue Zhang share first-authorship.
[2] Ellen Lau and Tilmann Sander share senior-authorship.

https://doi.org/10.1016/j.bspc.2026.110149
Received 19 March 2025; Received in revised form 22 January 2026; Accepted 18 March 2026
Available online 3 April 2026

However, what comes along with OPM-MEG's mobility is a low-frequency artifact driven by the interaction between the magnetometers and the remnant field gradients present in all magnetically shielded rooms [17]. Although participants are instructed not to move their head during experiments, it is impossible for their heads to be completely motionless due to e.g., breathing. Even seemingly small movements would drive considerable artifacts at the magnetometers, resulting in low-frequency artifacts in the data that are believed to be up to ~3 Hz [7,18]. This contrasts with canonical SQUID-MEG set-ups and some OPM-MEG designs [19], where the SQUID dewar or the OPM helmet are rigidly attached to the ground, and thus could only be affected when the ground moves relative to the Earth's magnetic field (e.g., when someone is stomping on the ground close to the MEG scanner). Several previous studies would therefore apply a high-pass filter with a rather high cutoff of e.g., 4 Hz in order to ensure removal of low-frequency artifacts [7,18,20–22]. While removing signal power below 4 Hz may not cause a problem if the goal is to examine the power of oscillatory bands above this filter, this practice may pose a problem for studying event-related fields (ERFs), as high-pass filters with a high cutoff are known to attenuate or eliminate ERP (event-related potential) amplitudes and ERP effects across conditions in EEG studies [23–26]. Moreover, since visual sentence processing studies usually present each segment (usually a word) for several hundreds of milliseconds, many ERF components may form some regularities in a lower frequency range slightly above 1 Hz [27,28]. Therefore a high-pass filter (HPF) of 4 Hz would especially risk removing these ERF components [7]. To preview, we also observed such effects in the current study. The application of a 4 Hz HPF is thus a compromise between removing slow movement-related signals and preserving all ERFs. As an alternative, some previous studies employed a moderate high-pass filter of e.g., 1 Hz, along with canonical pre-processing methods for SQUID-MEG data including ICA [29]. However, ICA may not be able remove the movement artifacts if they are statistically non-stationary. The fundamental challenge is that, while ICA procedures typically reject undesirable components related to e.g., eye blinks and heart beats through inspection of waveform and topographic field map [30], it is unclear what would be the canonical topography for low-frequency movement artifacts in OPM-MEG. Moreover, with experimental OPM-MEG setups with a handful of (e.g., ~10) sensors, it is challenging to establish full topographic field maps in the first place [29,31]. Furthermore, the ability for ICA to single out "clean" artifact components appears to degrade as the number of sensors decrease [32]. However, the number of OPM sensors available is usually constrained by the nontrivial price of OPM sensors, so that many labs starting on OPM-MEG only possess a limited number of sensors. Therefore, a reasonable pipeline for OPM-MEG needs to be developed to analyze ERFs with the goal of (1) attenuating low-frequency movement artifacts, (2) retaining major ERF deflections and ERF effects, and (3) accommodating OPM-MEG setups with a smaller set of sensors (for which inspecting the topographic field map is challenging). Developing methods for setups with small numbers of OPMs will lead to more cost-effective research tools, as OPM sensors are still very expensive. Such pipelines would greatly unleash the potential of OPM-MEG to study ERFs, including in sentence processing experiments.

One existing denoising method, evoked-biased denoising source separation (evoked-biased DSS [33]), may resolve this challenge, since this method does not rely on visual inspection of topographic field maps. Evoked-biased DSS decomposes data from sensors into PCA (principal component analysis) components, ranking from more time-locked to less time-locked. Assuming that movements are generally not time-locked to the stimulus, by keeping the first few components and excluding the rest of the components, we do not risk removing evoked responses much, while attenuating the effects from movement artifacts. However, evoked-biased DSS has not been tested on OPM-MEG data.

In the current OPM-MEG study, with a canonical visual sentence processing paradigm in Mandarin Chinese, we compared across pre-processing pipelines with various high-pass filtering thresholds and with/without evoked-biased DSS, leveraging the well-studied N400 response (and N400 effects across conditions) observed in SQUID-MEG and EEG [34,35]. In SQUID-MEG, N400 manifests as a deflection around 350 ms which is particularly observable in the sensors above the left hemisphere [36–38]. Apart from the N400 deflection *per se*, we also capitalize on the difference in N400 amplitude between verbs and nouns, which has been observed in SQUID-MEG and EEG systems as well, across a variety of languages including Mandarin Chinese (our stimuli language; [39–43]). To date, few sentence processing experiments have been conducted with OPM-MEG. Although one recent study employed an auditory sentence processing paradigm [44], they primarily employed multivariate analysis methods, and so it was unclear whether the canonical univariate N400 deflection and N400 effects were preserved.

To preview, we observed that while a 4 Hz high-pass filter completely destroyed the N400 deflection and any N400 differences between verbs and nouns, a low (0.1 Hz) or moderate (1 Hz) high-pass filter alone retained too much low-frequency artifacts in ERF data. Combining a moderate high-pass filter (1 Hz) with evoked-biased DSS appeared to offer the best performance in attenuating low-frequency movement artifacts while retaining the N400 deflection and the N400 difference between verbs and nouns.

2. Method

2.1. Participants

11 native speakers of Mandarin Chinese participated in this study (5 females, age M = 25, range = 10, SD = 3). All participants were right-handed, as measured by a translated version of the Edinburgh Handedness Inventory [45]. Informed consent was obtained from all participants and they received monetary reimbursement for their participation. The procedures were approved by the IRB of University of Maryland.

As a proof-of-concept initial study on OPM-MEG, we recruited a smaller sample size compared to SQUID-MEG studies; the current sample size is comparable to other exploratory OPM-MEG studies [46,47]. Meanwhile, we acknowledge that a larger sample size would bring about higher statistical power and would be beneficial as an important future step. As the goal of our current study was not comparing across populations, we did not screen participants based on e.g., psychiatric conditions. N400 components and N400 effects are generally preserved in at least some psychiatric conditions, e.g., schizophrenia [48,49].

2.2. Stimuli and procedure

The experiment was modeled after the EEG experiment of [50], with more sentences and therefore more trials.

Stimuli. We first came up with two lists of 80 verbs and 80 nouns in Mandarin Chinese, which was a superset of the list used in [50]. The verbs consisted of 40 separable verbs and 40 inseparable verbs; the nouns consisted of 40 compound nouns and 40 simplex nouns. For their detailed syntactic properties, see [50]. Comparing the responses between the two types of verbs or the two types of nouns is beyond the purpose of the current paper. In short, the separable verbs are linguistically more complex than the inseparable verbs, and the compound nouns are linguistically more complex than the simplex nouns. These verbs and nouns were all disyllabic, each corresponding to two Chinese characters (average number of strokes for each verb or noun item: verb: 16.2; noun: 15.3). Mean log frequency in the BCC corpus [51]: verb: 3.9, noun: 3.5. Mean log frequency in the CCL corpus [52]: verb: 4.1, noun: 3.7.

Then we created 160 sentences based on the 160 verb and noun items. The verbs and nouns were embedded in sentences in a certain sentential structure respectively, demonstrated in Table 1. Two counterbalanced lists of sentences for the verbs were created with two

Table 1
Example sentential stimuli for our experiment. The critical segment (i.e., the critical word) for our analysis is marked in bold.

Nouns

Sentence structure:
[因为] **[compound/simplex noun]** [是] [segment 4] [segment 5,] [所以/因此] [continuation.]
English translation:
[Because] **[compound/simplex noun]** [is] [segment 4] [segment 5,] [therefore/so] [continuation.]

Example:
因为/**热狗**/是/一种/高热量的/食物, /所以/不是很健康。
Because/**hotdog**/is/a.kind.of/high-calorie/food,/therefore/not.very.healthy.
"Because hotdog is a kind of high-calorie food, therefore it is not very healthy."

Verbs
Sentence structure:
[Name] [想要/不想/很想] **[separable/inseparable verb,]** [所以/因此] [continuation.]
English translation:
[Name] [wants to/doesn't want to/really wants to] **[separable/inseparable verb,]** [so/therefore] [continuation.]

Example:
小范/很想/**造反**, /所以/他秘密组织了一支军队。
Fan[name]/really.wants.to/rebel,/therefore/he.gathered.an.army.secretly.
"Fan really wants to rebel, therefore he gathered an army secretly."

counterbalanced list of people's names. These two lists were administered to roughly half of the participants respectively. Another 60 filler sentences were adapted from another study [53], rendering in total 220 sentences for each participant.

Procedure. The procedure was also generally consistent with [50]. Participants sat ~90 cm in front of the screen. Due to the setup of the projector in the lab, the screen had to be ~30° tilted from the ground; the participants' high performance suggested that the tilted screen did not affect sentence comprehension. 220 trials were run for each participant, with a different sentence presented in each trial. Each trial began with a fixation cross of 1000 ms with a 200 ms blank screen, followed by presentation of the segments of this sentence in a segment-by-segment (i.e., RSVP, rapid serial visual presentation) manner, with a presentation duration for 600 ms for each segment, and a 200 ms interval between segments, for a total of 800 ms stimulus-onset asynchrony (Fig. 1). The segments were presented in white characters against a black background, in ~ 57pt, Yahei UI font, at the center of the screen. In ~1/3 of the trials (76 out of 220 trials), the last segment of the sentence (i.e., the continuation) was colored in green and the subjects were instructed to respond to whether this continuation was congruent in this sentence, by pressing the left (congruent) or the right button (incongruent) on a response pad (the responses did not have a time-limit). In the other ~2/3 of the trials, the last segment was presented on the screen in white briefly for 1000 ms, and subjects did not need to respond. The inter-trial intervals were self-paced; subjects pressed any button when they were ready to start the next trial. Participants were instructed against reading out aloud prior to the experiment. An intercom connecting the inside of the magnetically shielded room (MSR) to the control room was always on during the experiment, which should capture explicit vocalizations apart from very faint whispers. We did not hear any explicit vocalizations via the intercom for our participants. The participants completed a SQUID-MEG session with similar materials prior to the OPM-MEG visit. Since the two visits were interleaved by on average 10.5 (range 5–18) days, the SQUID-MEG session would unlikely affect the OPM-MEG session. Indeed, participants demonstrated ceiling performance in accuracy for both sessions: SQUID, M = 98.92% (range:

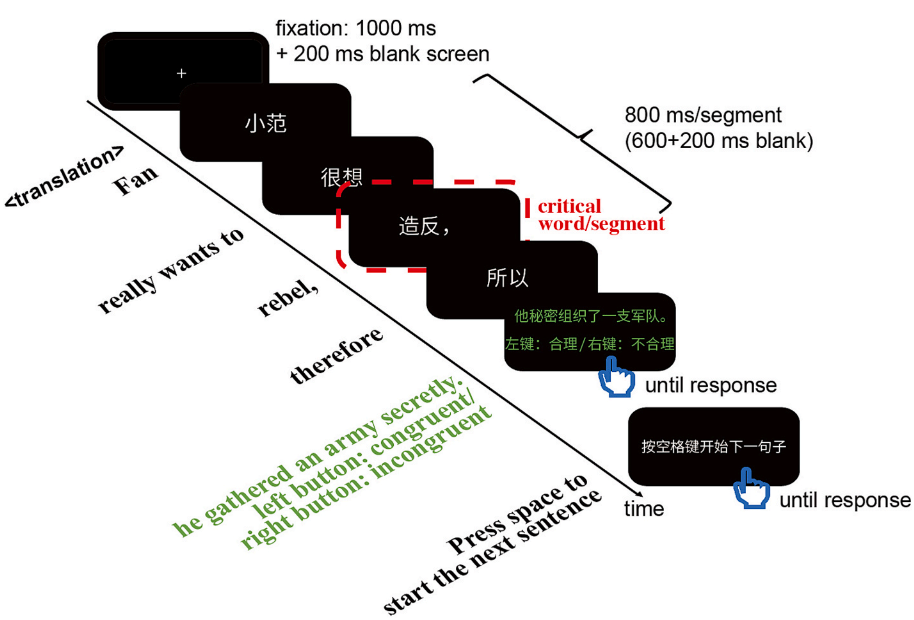

Fig. 1. Illustration of a trial of our sentence processing study. The English translation of the stimuli presented in Chinese is below the arrow.

97.37%-100%); OPM, M = 99.52% (range: 98.68%-100%).

2.3. Behavioral data analysis

Behavioral accuracy for each participant was calculated for the trials with a continuation judgement task, as the proportion of correct responses.

2.4. OPM setup

The OPM measurements were conducted in a two-layered Ak3b magnetically shielded room (Vacuumschmelze GmbH & Co. KG, Hanau, Germany), which had an active shielding to keep the OPMs in the linear regime as described in [17]. Nine OPM magnetometer sensors (QZFM generation 2, QuSpin Inc., Louisville, Colorado, USA) were positioned across the scalp using a modified EEG cap (easycap 21-channel 10–20 cap). The cap had clip in OPM holders attached to the fabric (similar to the approach in https://quspin.com/experimental-meg-cap/). Such fabric caps modified for OPM usage have been used for clinical works as well [54]. OPMs were placed especially on the left hemisphere, at T8, F4, Fpz, F3, F7, C3, T7, P7, O1 and O2 (Fig. 2). This is because in SQUID-MEG, N400 is particularly observable in the sensors above the left hemisphere [36–38]. Because the easycap 21-channel cap doesn't have an Fpz location, we created one at the appropriate location between Fp1 and Fp2. Each sensor recorded from two axes (z and y: radial head field and tangential field). The direction of the y axis was not consistent across all subjects due to rotational uncertainties, but was constant within each subject. In the current paper, we only analyze the z axis (radial direction relative to the head), which is in line with the direction recorded in most existing SQUID-MEG devices. Although SQUID devices mainly consist of gradiometers, these gradiometers are designed to be sensitive to radial magnetic field information. Note that the directions mentioned here refer to the orientation of the magnetic field sensor in relation to the head surface: tangential is parallel to the head surface and radial points into the head. The sensor z axis (the radial axis) captures most of the signal generated by tangential neural dipoles.

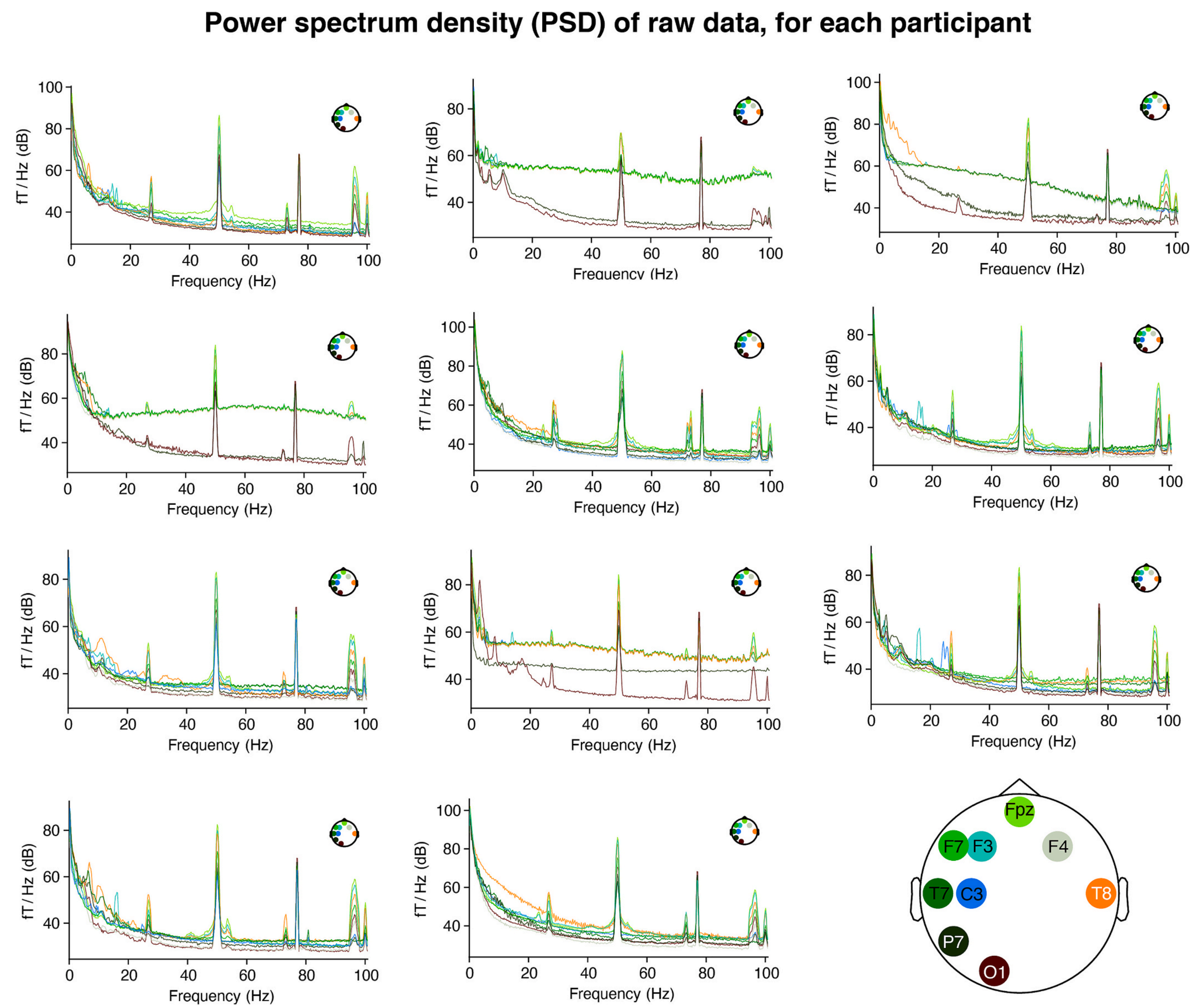

Fig. 2. Power spectrum density (PSD) of raw OPM-MEG data for each of the 11 participants. The OPM sensors were attached to an EEG cap, whose layout is illustrated in the bottom right corner. Consistent with prior OPM-MEG studies, we observed two spectral peaks around 27 Hz (cause yet unknown) and 50 Hz (line noise) across most participants.

2.5. OPM data analysis

OPM data analyses was conducted with customized code in Python and MATLAB, with MNE-Python [55] and MEEGkit [56]. Aligned Rank Transform ANOVA [57] was administered in R with the ARTool package [58]. Wilcoxon tests, Bayes Factor analysis and Shapiro-Wilk normality checks were administered in JASP 0.18.1 [59]; Bayes Factors were calculated based on the default Cauchy prior. A Bayes Factor (BF_{10}) of 3–10 indicate moderate evidence for H_1, 1–3 for anecdotal evidence for H_1, 1/3–1 for anecdotal evidence for H_0, and 1/10–1/3 for moderate evidence for H_0 [60]. The p-values reported are two-tailed and uncorrected. The effect size *r* for Wilcoxon signed-rank tests was quantified by rank-biserial correlation.

Preprocessing. Upon inspecting the power spectrum (Fig. 2) for individual participants, there were generally a 50 Hz component as well as a ~27 Hz component. The 50 Hz component was most likely line noise. A 20–30 Hz component is generally present in other OPM systems as well [7,19,61] with a yet-unclear source. Therefore, we applied notch filters at 50 and 27 Hz (transitional bandwidth = 4 Hz) to remove these components. Then we tested six different treatments: (1) Data was passed with a 0.1–40 Hz IIR (infinite impulse response) filter (i.e., high-pass filter, HPF = 0.1 Hz) only, (2) Data was passed with a 0.1–40 Hz IIR filter, then an evoked-biased denoising source separation (evoked-biased DSS), (3) a 1–40 Hz IIR filter only, (4) a 1–40 Hz IIR filter followed by evoked-biased DSS, (5) a 4–40 Hz IIR filter only, (6) a 4–40 Hz IIR filter followed by evoked-biased DSS. The other parameters for the IIR filters were determined by default algorithms in MNE-Python.

A sketch of the procedure of evoked-biased DSS [33] is as follows (see Section 2.6 for a more detailed description of the algorithm). Evoked-biased DSS transforms data from 9 sensors into 9 components, ranking from more stimulus-evoked-like to more noise-like. By keeping the first few components and excluding the rest of the components, we do not risk removing evoked responses much, while attenuating the effects from movement artifacts. The final epoch size that we wanted to have is −200:800 ms, time-locked to critical word onset. Each epoch was baselined based on the −200:0 ms baseline. In order to reduce the influence of the baseline period in DSS (note: in the current paper, "DSS" refers to evoked-biased DSS), we ran DSS on longer epochs of −200:1600 ms [62,63]. Evoked-biased DSS was conducted with all epochs across the four conditions (epoch data were normalized for each sensor prior to DSS), resulting in 9 DSS components. We projected the first two DSS components back to the sensor space to create corrected epochs for analysis.

ERFs (event-related fields) were computed as the median of the epochs for each participant and for each condition respectively: separable verb, inseparable verb, compound noun, simplex noun, aggregated verb condition across separable and inseparable verbs, as well as aggregated noun condition across compound and simplex nouns. Because the mean is more subject to the influence of outlier data points, computing mean ERFs would require us to apply an additional outlier cutoff processing step which would likely result in different numbers of excluded trials across preprocessing pipelines. Median ERFs are less sensitive to outliers, therefore computing the ERFs with the median mitigates the need for applying a cutoff.

Calculation of cross-trial variance for ERFs. As a means of quantifying overall noise level in the results of each pipeline, we computed a cross-trial variance measure for the ERFs. For each pipeline, for each participant and at each time point, we first calculated *one* variance value across all verb epochs and all noun epochs respectively. This variance value was calculated in the following way. First, we calculated the standard deviation at each sensor. Then, we averaged the standard deviations across all 9 sensors. After that, for each time point, we averaged the standard deviations across all participants, for the noun epochs and verb epochs respectively. This way, for noun epochs and verb epochs respectively, we derived an 1D array across time, with one averaged variance value at each time point.

Calculation of RMS (root mean square). To determine whether the pipelines adequately captured the typical peaks in the ERF, we used an RMS measure. We calculated the RMS from ERFs that were first averaged across the verb and noun conditions, then averaged across participants. RMS was calculated as the root mean square across all 9 sensors.

Calculation of N400 amplitude. As described below, ERFs collapsed across conditions indicated a clear N400 peak between 300–400 ms that was maximal at sensor F7. Therefore, we calculated N400 amplitude for each condition and for each participant at this sensor as the mean response amplitude across the 300–400 ms time window upon segment onset.

2.6. The DSS algorithm

Here we provide a more detailed description of the algorithm of evoked-biased DSS [33], implemented in the dss1() function of MEEGkit [56].

For each participant, the algorithm worked as follows. The input data was all epoched trials **S** with shape: number of time points (T) × number of sensors (K) × number of trials/epochs (N_m). A time-shift "baseline" covariance matrix $\mathbf{C_0}$ was derived from S and normalized by total weight; another "biased" covariance matrix $\mathbf{C_1}$ was derived from the average across all epochs $\frac{1}{N_m}\sum_{n=1}^{N_m}\mathbf{S}^n(t)$ (i.e., this is the "bias function" of DSS) and normalized by total weight as well. Eigenvector $\mathbf{E_0}$ and eigenvalues $\mathbf{\Lambda_0}$ were derived from a PCA upon $\mathbf{C_0}$:

$$\mathbf{E_0}, \mathbf{\Lambda_0} = \mathrm{PCA}(\mathbf{C_0})$$

For all PCAs, components $<10^{-12}$ were discarded following MEEGkit defaults, to save computation and avoid numerical problems. Then a covariance matrix $\mathbf{C_2}$ was derived by "biasing" whitened data:

$$\mathbf{W} = \mathrm{diag}\left(\mathbf{\Lambda_0}^{-1/2}\right)$$

$$\mathbf{C_2} = (\mathbf{E_0}\mathbf{W})^{\mathrm{T}}\mathbf{C_1}(\mathbf{E_0}\mathbf{W})$$

Then we can obtain **T**, the matrix to convert **S** to normalized DSS components. First, $\mathbf{T_{raw}}$ was derived as follows, which involves a PCA on $\mathbf{C_2}$:

$$\mathbf{E_2}, \mathbf{\Lambda_2} = \mathrm{PCA}(\mathbf{C_2})$$

$$\mathbf{T_{raw}} = \mathbf{E_0}\mathbf{W}\mathbf{E_2}$$

$\mathbf{T_{raw}}$ was subsequently normalized to **T**:

$$\mathbf{N} = \sqrt{\mathrm{diag}\left(\mathbf{T_{raw}}^{\mathrm{T}}\mathbf{C_0}\mathbf{T_{raw}}\right)}$$

$$\mathbf{T} = \mathbf{T_{raw}}/\mathbf{N}$$

The matrix **F** to convert DSS components back to the sensor space is then:

$$\mathbf{F} = \mathrm{pinv}(\mathbf{T})$$

T and its pseudoinverse **F** were then used to project original data **S** into the DSS space, then the first x components (in our case x = 2) were projected back to the sensor space:

$$\mathbf{S_{denoised}} = \mathbf{S}\mathbf{T}_{[:,1:x]}\mathbf{F}_{[1:x,:]}$$

3. Results

3.1. Behavioral results

Accuracy. Participants generally achieved ceiling accuracy (range 98.68%-100%) in judging the semantic congruity of sentence continuations. Therefore, all trials were included in the MEG analysis.

3.2. MEG results

In this section we evaluate the processing pipelines based on several qualitative and quantitative aspects.

Shape of ERF. In pipelines utilizing a 1 Hz high-pass filter (HPF), we were able to identify a clear deflection around 400 ms in left centro-frontal sensors, including F7, T7 and C3 (Fig. 3). The deflection had the highest amplitude at F7, from which we calculate the amplitude of N400 responses for further analysis. This deflection was faintly identifiable when using a 0.1 Hz HPF, and was practically invisible when applying a 4 Hz HPF. The same observation generally holds when inspecting individual participants' data (Fig. 4). Notably, the N400 deflection we observed was at the magnitude of hundreds of fT, which is higher compared to the typical magnitude observed in SQUID-MEG systems, at the magnitude of tens of fT [35,64], as expected given the much closer proximity of OPM-MEG sensors to the underlying neural generators. Additionally, this difference in magnitude may also be relevant to the fact that, while OPM sensors are magnetometers, SQUIDs are often designed in a gradiometer configuration.

Cross-trial variance for ERFs. Across the entire epoch, the pipeline with HPF = 4 Hz, with DSS had the lowest cross-trial variance, which holds for both verb trials and noun trials (Fig. 5). This was followed by the pipeline with HPF = 1 Hz, with DSS. The pipeline with HPF = 0.1 Hz without DSS had the highest variance among the pipelines.

Shape of RMS. Previous MEG sentence or single-word processing studies tend to observe a ~400 ms peak, along with earlier peak(s) around ~200 ms in RMS [42,65,66]. The ~200 ms peak(s) likely reflect visual-orthographic processing [5,67]. The ~400 ms peak or bump was clearly visible when we employed a 1 Hz HPF, while being faint or invisible for 0.1 or 4 Hz HPFs (Fig. 6). A peak ~200 ms can only be clearly visually identified when a 1 Hz HPF is combined with DSS, but not when only applying a 1 Hz HPF.

N400 effect between verbs and nouns. We started with an omnibus repeated-measures ANOVA with all factors: 2 (linguistic category: noun vs. verb) × 2 (linguistic complexity: simpler vs. more complex) × 3 (HPF: 0.1, 1, 4 Hz) × 2 (DSS: yes, no). As not all condition's data passed the Shapiro-Wilk normality test ($p < 0.05$), we conducted the non-parametric equivalent (Aligned Rank Transform ANOVA) instead [57]. Notably, there was a significant interaction effect between linguistic category and DSS, $F(1,230) = 6.49$, $p = 0.011$, suggesting the application of DSS significantly affects the N400 effect across verbs and nouns. There was also a significant interaction effect between HPF and DSS, $F(2,230) = 6.00$, $p = 0.003$, suggesting that we cannot simply average all HPF conditions within a certain DSS condition together. Furthermore, neither the main effect of linguistic complexity or the interaction effects involving linguistic complexity was statistically significant (p's > 0.096), suggesting that we can safely average the ERFs of separable and inseparable verb conditions, and compound and simplex noun conditions, respectively. The full results are summarized in Table 2.

Following our goal to test the verb-noun N400 effects for each HPF-DSS combination, corroborated by the interaction effects observed in the omnibus ANOVA, we administered planned, uncorrected Wilcoxon signed-rank tests (two-tailed) across the N400 amplitudes for verbs and nouns within each HPF-DSS combination. These Wilcoxon signed-rank tests only revealed a significant and reliable difference ($p < 0.05$, $BF_{10} > 3$) between the N400 amplitudes of verbs and nouns for the pipeline with HPF = 1 Hz and DSS, $z = -2.05$, $p = 0.042$, effect size $r = -0.70$, 95% CI [−0.91,-0.19], $BF_{10} = 4.39$ [>3, moderate evidence for H_1] (Fig. 7). This difference was not significantly different for the other pipelines, although the effects were generally in the same direction. HPF = 0.1 Hz without DSS: $|z| < 0.01$, $p = 1.00$, $|\text{effect size } r| < 0.01$, 95% CI [−0.58,0.58], $BF_{10} = 0.32$; HPF = 0.1 Hz with DSS: $z = 1.33$, $p = 0.21$, effect size $r = 0.46$, 95% CI [−0.18,0.82], $BF_{10} = 0.70$; HPF = 1 Hz without DSS: $z = -1.87$, $p = 0.067$, effect size $r = -0.64$, 95% CI [−0.89,-0.08], $BF_{10} = 0.86$; HPF = 4 Hz without DSS: $z = -1.16$, $p = 0.28$, effect size $r = -0.39$, 95% CI [−0.80,0.45], $BF_{10} = 0.57$; HPF = 4 Hz with DSS: $z = -0.53$, $p = 0.64$, effect size $r = -0.18$, 95% CI [−0.69,0.45], $BF_{10} = 0.40$.

4. Conclusion

In the current paper, by comparing different preprocessing pipelines for OPM-MEG data, we demonstrated that combining a moderate high-pass filter (HPF) of 1 Hz with evoked-biased DSS preserves the N400 deflection and allowed us to successfully replicate prior observations of an N400 difference between nouns and verbs. Moreover, this pipeline revealed canonical RMS peaks, and had a relatively low cross-trial variance among the pipelines. This pipeline thus unleashes the potential of OPM-MEG to study ERFs against low-frequency artifacts, especially for OPM-MEG setups with fewer sensors, where methods like ICA would be inapplicable. This can allow a beneficial 'convergent methods' approach for labs to obtain high-quality MEG data in closer proximity to the neural generators than traditional systems, even in the face of current constraints on the number of OPM sensors.

5. Future directions

5.1. On applying higher high-pass filters

Although a 4 Hz HPF largely destroyed the N400 deflection and N400 effects in our study, such higher HPFs may still be applicable in studying some other ERFs. For example, the auditory evoked responses of M100, which is a peak ~100 ms upon auditory stimuli onset, appeared to survive such higher HPFs in previous studies [18]. Notably, in [18], auditory pure tones were presented at a duration of 500 ms, with intervals of 1200 ms between the tones. Thus, the stimuli presentation would form regularities below the 4 Hz HPF. Nevertheless, M100 was observed in that study. This is not surprising since M100 is a transient response that only lasted for ~50 ms, compared to our N400 deflections, which appeared to extend for ~300 ms.

In all, a high HPF may still be applicable depending on the component of interest, albeit with care. For ERPs or ERFs, deflections closer to the stimulus onset (the time point that epochs are time-locked to) tend to be narrower, corresponding to higher frequencies; while deflections later from the stimulus onset tend to be wider and sustained, corresponding to lower frequencies. Thus, a certain HPF may be admissible if the interest is on earlier components such as M100, but not for later ones such as N400 or P300 [24]. An often overlooked benefit of high pass filtering is a more stable PCA result since the high pass removes the high power low frequency noise typical for a 1/f noise behavior.

5.2. Other pipelines

Beyond our current pipeline, some other pipelines may also achieve satisfactory performances, and future studies could work on comparing across these pipelines. In our current case, we worked with a setup with a very limited number of sensors and a lack of geometrical information (e.g., the sensors' locations and anatomical information), which is rather plausible in experimental or clinical settings. In this type of scenario, DSS appears to be one of the only available preprocessing options. While ICA is widely used in SQUID-MEG data preprocessing in rejecting for example ocular and cardiac artifacts [55], it falls short in rejecting motion-related artifacts as these components are temporally non-stationary (see Introduction). ICA as such is a blind method based on the assumption that the number of independent sources (artifacts and brain signals) equals the number of measured data channels. This is violated in practice quite often, but the severity is higher for a low number of sensors. DSS in contrast is semi-blind [33,68] since the trial structure is used to improve the signal-to-noise ratio. Our work demonstrates that the inclusion of the known trial structure improves the detectability of physiological effects. Evoked-biased DSS (which

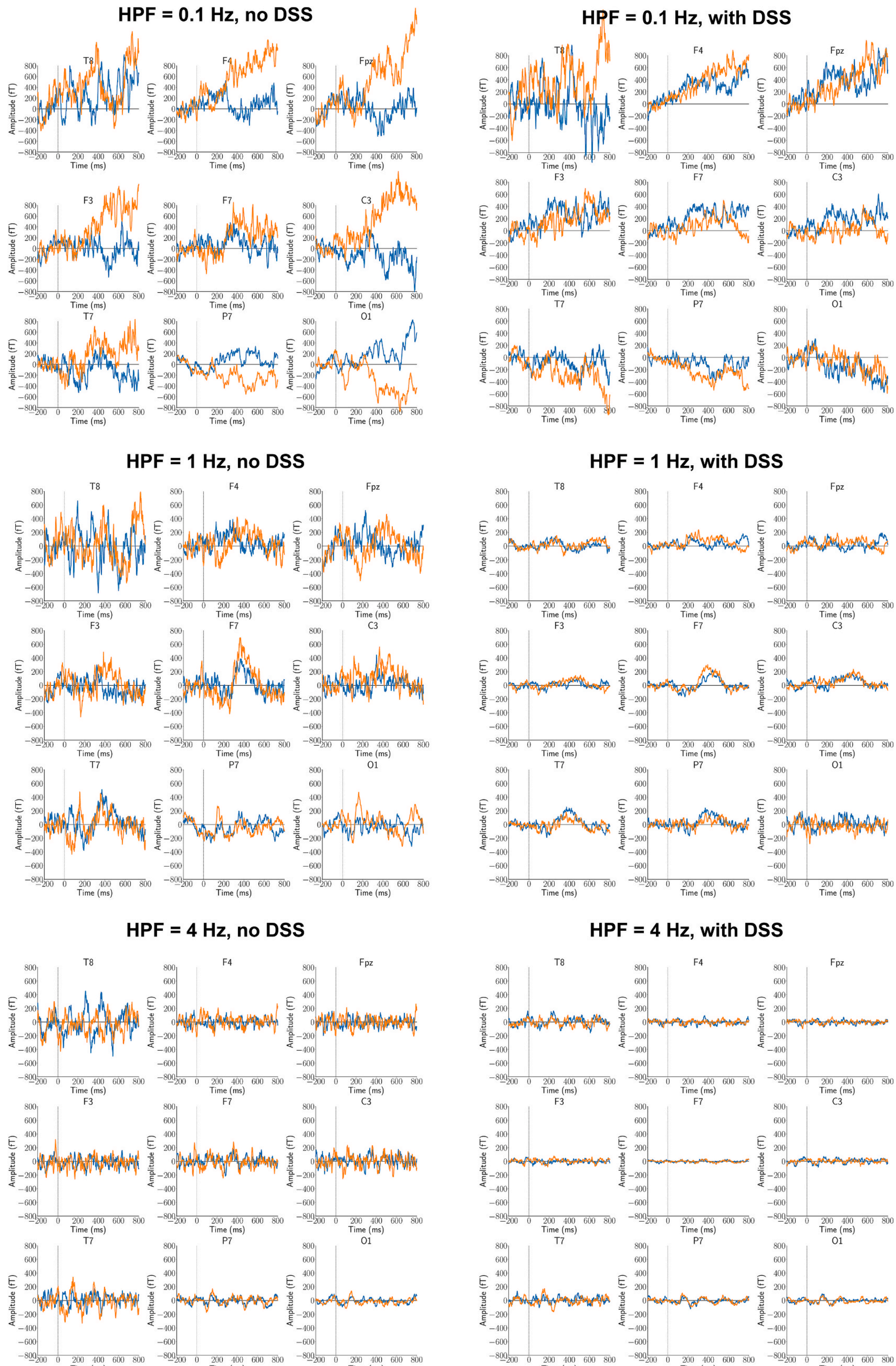

Fig. 3. Grand average ERFs across all participants (blue line: verb; orange line: noun). ERF: event-related field.

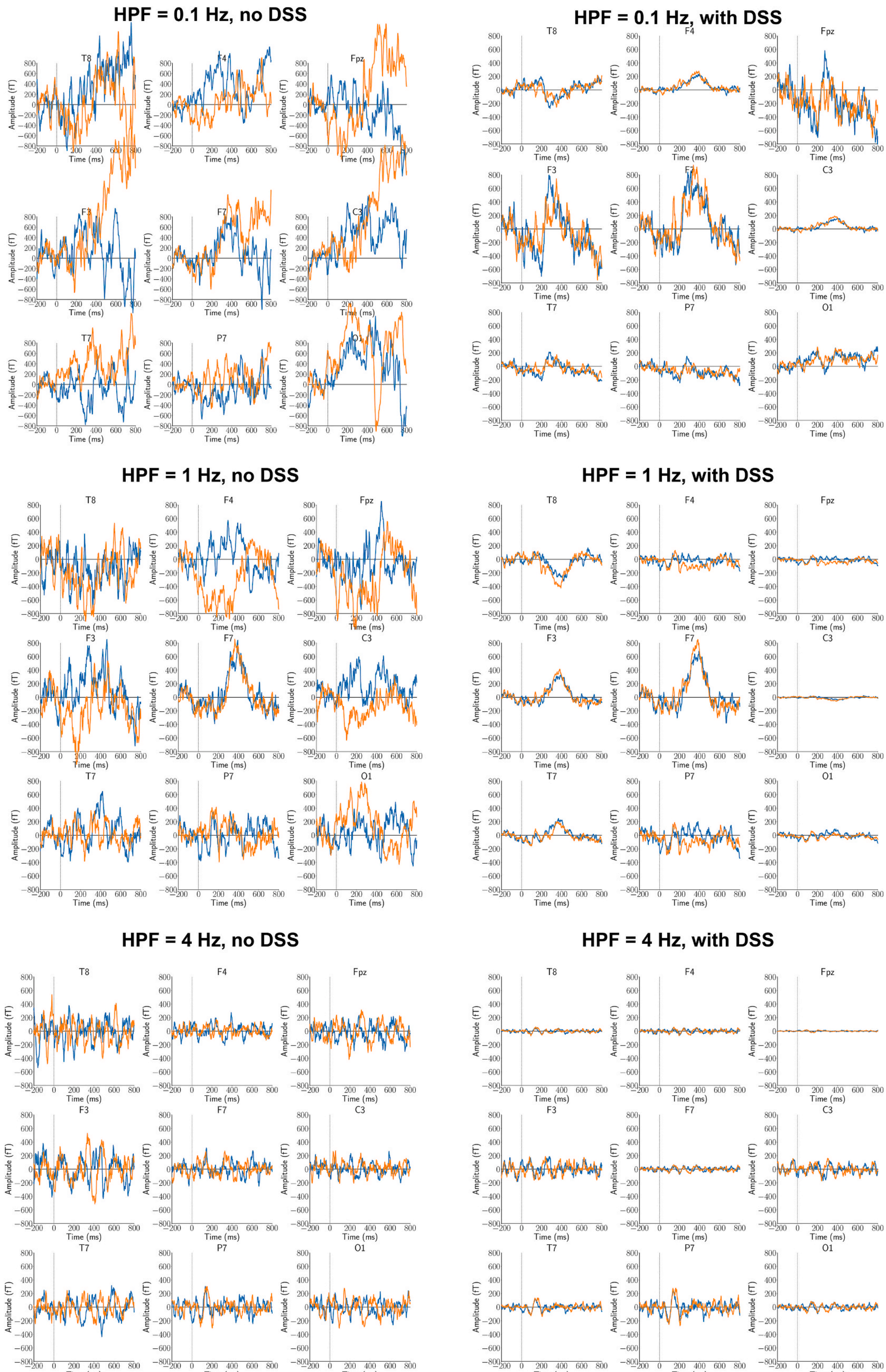

Fig. 4. Median ERFs for noun and verb epochs in one participant (blue line: verb; orange line: noun). ERF: event-related field. The figures for every individual participant are shared openly on OSF.

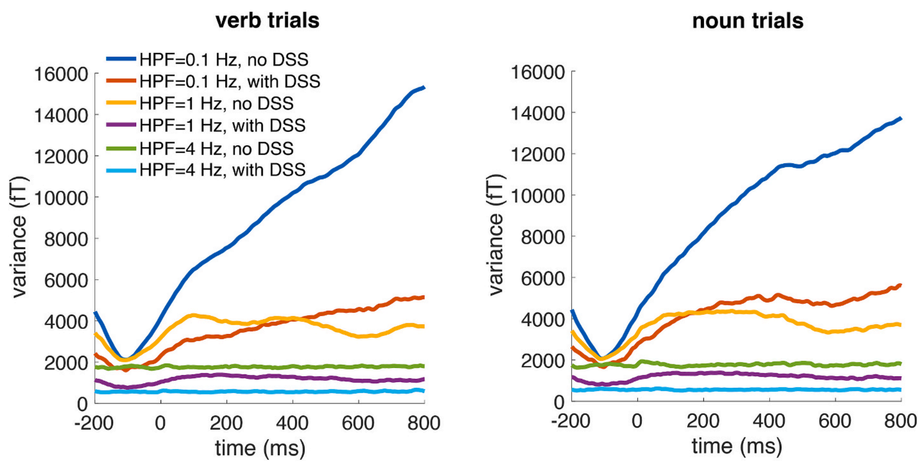

Fig. 5. Cross-trial variance for each pipeline among verb trials (left) and noun trials (right), averaged across all participants. HPF: high-pass filter; DSS: denoising source separation.

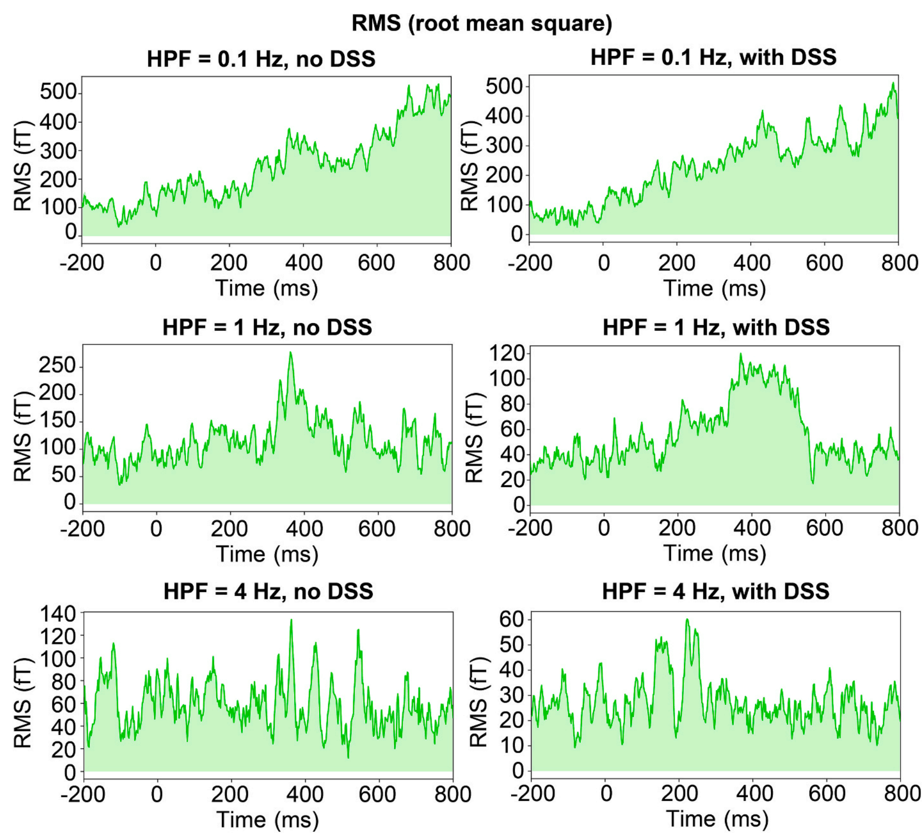

Fig. 6. RMS (root mean square) across all sensors, based on the ERFs averaged across verbs and nouns. Only when combining a 1 Hz HPF and DSS are we able to clearly visually identify peaks around 200 ms and 400 ms. ERF: event-related field; HPF: high-pass filter; DSS: denoising source separation.

involves PCA in its procedure) is also favorable over a simple PCA pipeline, as PCA components are ordered by power instead of cross-trial reproducibility, therefore raising a challenge in choosing the right principal components (see [33]). Of note, HFC (homogeneous field correction) may be a promising method for this purpose [69]. We did not perform HFC for our current data because it requires the sensor orientation angles, which we did not record in this study. Future studies could compare across our current pipeline and HFC for the metrics that we used in the current work, or even combine them together.

Another opportunity of future research lies in exploring the usage of evoked-biased DSS in the source space. For our current setup, source space analysis cannot be conducted due to a very limited number of sensors and a lack of geometrical information. Future studies could explore whether DSS could improve data quality in the source space, complementary to source-localization methods such as beamforming [70], MNE (minimum-norm estimate) [71,72] and dSPM (dynamic statistical parametric mapping) [73]. Future studies could also investigate the extent to which our current method and other methods are robust against different magnetic shielding conditions [74,75], which may correspond to different landscapes of noise. The shielding used in this study [17] represents the majority of the roughly 200 magnetically shielded rooms used for MEG worldwide. The only addition in our setup is the active shielding (see Section 2.4), which is not necessary anymore if a closed-loop OPM system is available.

Table 2
Results of the omnibus Aligned Rank Transform ANOVA: 2 (linguistic category: noun vs. verb) × 2 (linguistic complexity: simpler vs. more complex) × 3 (HPF: 0.1, 1, 4 Hz) × 2 (DSS: yes, no). p-values <0.05 are marked in bold.

	F-value	*p*-value
linguistic category	3.47	0.064
linguistic complexity	0.76	0.38
HPF	16.61	**<0.001**
DSS	18.39	**<0.001**
linguistic category × linguistic complexity	0.11	0.75
linguistic category × HPF	2.84	0.061
linguistic complexity × HPF	0.48	0.62
linguistic category × DSS	6.49	**0.011**
linguistic complexity × DSS	2.78	0.097
HPF × DSS	6.00	**0.003**
linguistic category × linguistic complexity × HPF	0.05	0.96
linguistic category × linguistic complexity × DSS	0.05	0.83
linguistic category × HPF × DSS	1.88	0.15
linguistic complexity × HPF × DSS	1.85	0.16
linguistic category × linguistic complexity × HPF × DSS	0.41	0.67

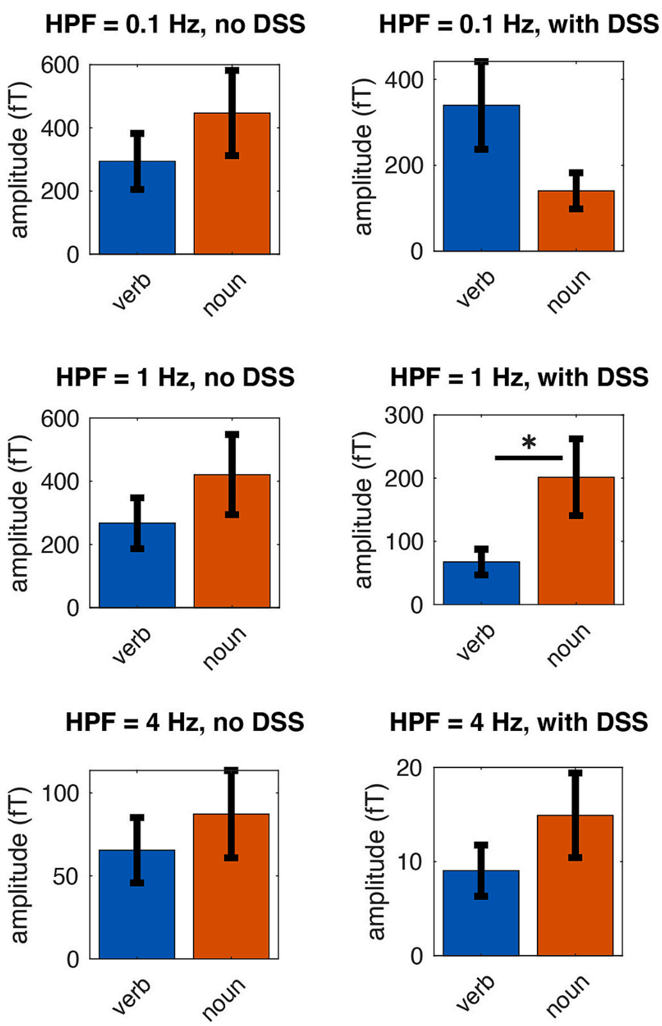

Fig. 7. Comparison of N400 response (300–400 ms) amplitude at sensor F7 across verbs and nouns. Only when an HPF (high-pass filter) of 1 Hz was combined with DSS, were we able to observe a significant difference across verbs and nouns. Error bars stand for standard error. *: $p < 0.05$. HPF: high-pass filter; DSS: denoising source separation.

It is also important to note that, as some participants may involuntarily whisper as they read despite explicit instructions against speech production, evoked-biased DSS may not be able to disentangle time-locked neural responses and time-locked speech production motor responses. As whispering likely only happen for a small proportion of trials if it does happen for a certain participant, we would not expect speech-related motor responses to be picked up by the top one or two components, as evoked-biased DSS was designed to pick out signals that are reproducible across trials. Nevertheless, experimentally quantifying this would be an interesting direction for future exploration.

5.3. Exploring OPM-MEG recording with a limited number of sensors

In the current study, we demonstrated a pipeline that accommodates OPM-MEG systems with a limited number of sensors (in our case, nine sensors), and obtained meaningful ERF responses in a sentence processing study. In practice, the number of sensors may not only be constrained by financial costs, but also by preparation time. For example, attaching more OPM sensors to the cap increases setup time, which may be a limiting factor in certain clinical contexts where shorter preparation times are preferred. Notably, previous research has shown that as few as four dual-axis OPM sensors can be sufficient for a brain-computer interface for motor imagery decoding [76]. This, along with our findings, suggests that even with a reduced sensor count, OPM-MEG can still yield meaningful results. Taken together, our study marks one of the initial steps in validating OPM-MEG systems with fewer sensors, paving the way for more cost-effective solutions which would enable wider clinical usage of OPM-MEG.

CRediT authorship contribution statement

Xinchi Yu: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Dongxue Zhang:** Writing – review & editing, Software, Investigation. **Thomas Carey:** Writing – review & editing, Investigation. **Ernst Pöppel:** Writing – review & editing, Supervision, Resources. **Ellen Lau:** Writing – review & editing, Supervision, Conceptualization. **Tilmann Sander:** Writing – review & editing, Supervision, Resources, Methodology.

Funding

This study was supported by Dean's Research Initiative, College of Behavioral and Social Sciences, University of Maryland (to X.Y.).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Jonathan Simon for discussions.

Data availability

Data, stimuli and processing scripts for the current study are openly shared on the Open Science Framework (link: https://osf.io/cnbs9/).

References

[1] R.R. Lee, M. Huang, Magnetoencephalography in the diagnosis of concussion, Concussion (2012) 94–111, https://doi.org/10.1159/000358768.
[2] E.M. Davenport, J.E. Urban, C. Vaughan, J.C. DeSimone, B. Wagner, M. A. Espeland, A.K. Powers, C.T. Whitlow, J.D. Stitzel, J.A. Maldjian, MEG measured delta waves increase in adolescents after concussion, Brain Behav. 12 (2022) 1–9, https://doi.org/10.1002/brb3.2720.
[3] S. Dikker, M.F. Assaneo, L. Gwilliams, L. Wang, A. Kösem, Magnetoencephalography and language, Neuroimaging Clin. N. Am. 30 (2020) 229–238, https://doi.org/10.1016/j.nic.2020.01.004.

[4] A. Tarkiainen, P. Helenius, R. Salmelin, Category-specific occipitotemporal activation during face perception in dyslexic individuals: an MEG study, Neuroimage 19 (2003) 1194–1204, https://doi.org/10.1016/S1053-8119(03)00161-7.

[5] R. Salmelin, Clinical neurophysiology of language: the MEG approach, Clin. Neurophysiol. 118 (2007) 237–254, https://doi.org/10.1016/j.clinph.2006.07.316.

[6] M. Brickwedde, P. Anders, A.A. Kühn, R. Lofredi, M. Holtkamp, A.M. Kaindl, T. Grent-'t-Jong, P. Krüger, T. Sander, P.J. Uhlhaas, Applications of OPM-MEG for translational neuroscience: a perspective, Transl. Psychiatry 14 (2024) 1–12, https://doi.org/10.1038/s41398-024-03047-y.

[7] R.A. Seymour, N. Alexander, S. Mellor, G.C. O'Neill, T.M. Tierney, G.R. Barnes, E. A. Maguire, Interference suppression techniques for OPM-based MEG: opportunities and challenges, Neuroimage 247 (2022) 118834, https://doi.org/10.1016/j.neuroimage.2021.118834.

[8] M.J. Brookes, J. Leggett, M. Rea, R.M. Hill, N. Holmes, E. Boto, R. Bowtell, Magnetoencephalography with optically pumped magnetometers (OPM-MEG): the next generation of functional neuroimaging, Trends Neurosci. 45 (2022) 621–634, https://doi.org/10.1016/j.tins.2022.05.008.

[9] L. Qin, J.H. Gao, New avenues for functional neuroimaging: ultra-high field MRI and OPM-MEG, Psychoradiology 1 (2021) 165–171, https://doi.org/10.1093/psyrad/kkab014.

[10] H. Cheng, K. He, C. Li, X. Ma, F. Zheng, W. Xu, P. Liao, R. Yang, D. Li, L. Qin, S. Na, B. Lyu, J.H. Gao, Wireless optically pumped magnetometer MEG, Neuroimage 300 (2024) 120864, https://doi.org/10.1016/j.neuroimage.2024.120864.

[11] E. Boto, S.S. Meyer, V. Shah, O. Alem, S. Knappe, P. Kruger, T.M. Fromhold, M. Lim, P.M. Glover, P.G. Morris, R. Bowtell, G.R. Barnes, M.J. Brookes, A new generation of magnetoencephalography: room temperature measurements using optically-pumped magnetometers, Neuroimage 149 (2017) 404–414, https://doi.org/10.1016/j.neuroimage.2017.01.034.

[12] A. Borna, T.R. Carter, J.D. Goldberg, A.P. Colombo, Y.Y. Jau, C. Berry, J. McKay, J. Stephen, M. Weisend, P.D.D. Schwindt, A 20-channel magnetoencephalography system based on optically pumped magnetometers, Phys. Med. Biol. 62 (2017) 8909–8923, https://doi.org/10.1088/1361-6560/aa93d1.

[13] T.H. Sander, J. Preusser, R. Mhaskar, J. Kitching, L. Trahms, S. Knappe, Magnetoencephalography with a chip-scale atomic magnetometer, Biomed. Opt. Express 3 (2012) 981, https://doi.org/10.1364/BOE.3.000981.

[14] I.K. Kominis, T.W. Kornack, J.C. Allred, M.V. Romalis, A subfemtotesla multichannel atomic magnetometer, Nature 422 (2003) 596–599, https://doi.org/10.1038/nature01484.

[15] L.M. Andersen, C. Pfeiffer, S. Ruffieux, B. Riaz, D. Winkler, J.F. Schneiderman, D. Lundqvist, On-scalp MEG SQUIDs are sensitive to early somatosensory activity unseen by conventional MEG, Neuroimage 221 (2020), https://doi.org/10.1016/j.neuroimage.2020.117157.

[16] J. Iivanainen, R. Zetter, L. Parkkonen, Potential of on-scalp MEG: robust detection of human visual gamma-band responses, Hum. Brain Mapp. 41 (2020) 150–161, https://doi.org/10.1002/hbm.24795.

[17] V. Jazbinšek, U. Marhl, T. Sander, SERF-OPM Usability for MEG in Two-Layer-Shielded Rooms, in: Flex. High Perform. Magn. F. Sensors, Springer International Publishing, Cham, 2022: pp. 179–193. Doi: 10.1007/978-3-031-05363-4_10.

[18] U. Marhl, A. Jodko-Władzińska, R. Brühl, T. Sander, V. Jazbinšek, Transforming and comparing data between standard SQUID and OPM-MEG systems, PLoS One 17 (2022) 1–22, https://doi.org/10.1371/journal.pone.0262669.

[19] O. Alem, K.J. Hughes, I. Buard, T.P. Cheung, T. Maydew, A. Griesshammer, K. Holloway, A. Park, V. Lechuga, C. Coolidge, M. Gerginov, E. Quigg, A. Seames, E. Kronberg, P. Teale, S. Knappe, An integrated full-head OPM-MEG system based on 128 zero-field sensors, Front. Neurosci. 17 (2023) 1–8, https://doi.org/10.3389/fnins.2023.1190310.

[20] R.A. Seymour, N. Alexander, S. Mellor, G.C.O. Neill, T.M. Tierney, G.R. Barnes, E. A. Maguire, NeuroImage using OPMs to measure neural activity in standing, mobile participant, Neuroimage 244 (2021) 118604, https://doi.org/10.1016/j.neuroimage.2021.118604.

[21] A. Gialopsou, C. Abel, T.M. James, T. Coussens, M.G. Bason, R. Puddy, F. Di Lorenzo, K. Rolfs, J. Voigt;, T. Sander, M. Cercignani, P. Krüger, Improved spatio - temporal measurements of visually evoked fields using optically - pumped magnetometers, Sci. Rep. (2021) 1–11, https://doi.org/10.1038/s41598-021-01854-7.

[22] A. Jodko-Władzińska, T. Sander, Emotionally charged visually evoked magnetic fields, Acta Phys. Pol. A 146 (2024) 521–525, https://doi.org/10.12693/APhysPolA.146.521.

[23] G. Zhang, D.R. Garrett, S.J. Luck, Optimal filters for ERP research I: a general approach for selecting filter settings, Psychophysiology 61 (2024), https://doi.org/10.1111/psyp.14531.

[24] G. Zhang, D.R. Garrett, S.J. Luck, Optimal filters for ERP research II : Recommended settings for seven common ERP components, Psychophysiology 61 (2024) 1–24, https://doi.org/10.1111/psyp.14530.

[25] D. Tanner, K. Morgan-Short, S.J. Luck, How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition, Psychophysiology 52 (2015) 997–1009, https://doi.org/10.1111/psyp.12437.

[26] J. van Driel, C.N.L. Olivers, J.J. Fahrenfort, High-pass filtering artifacts in multivariate classification of neural time series data, J. Neurosci. Methods 352 (2021) 109080, https://doi.org/10.1016/j.jneumeth.2021.109080.

[27] N. Ding, L. Melloni, H. Zhang, X. Tian, D. Poeppel, Cortical tracking of hierarchical linguistic structures in connected speech, Nat. Neurosci. 19 (2015) 158–164, https://doi.org/10.1038/nn.4186.

[28] Y. Lu, P. Jin, X. Pan, N. Ding, Delta-band neural activity primarily tracks sentences instead of semantic properties of words, Neuroimage (2022) 118979, https://doi.org/10.1016/j.neuroimage.2022.118979.

[29] J. Iivanainen, T.R. Carter, M.C.S. Trumbo, J. McKay, S. Taulu, J. Wang, J. M. Stephen, P.D.D. Schwindt, A. Borna, Single-trial classification of evoked responses to auditory tones using OPM- and SQUID-MEG, J. Neural Eng. 20 (2023) 056032, https://doi.org/10.1088/1741-2552/acfcd9.

[30] A. Puce, M.S. Hämäläinen, A review of issues related to data acquisition and analysis in EEG/MEG studies, Brain Sci. 7 (2017), https://doi.org/10.3390/brainsci7060058.

[31] T. Sander, A. Jodko-Władzińska, S. Hartwig, R. Brühl, T. Middelmann, Optically pumped magnetometers enable a new level of biomagnetic measurements, Adv. Opt. Technol. 9 (2020) 247–251, https://doi.org/10.1515/aot-2020-0027.

[32] M. Klug, K. Gramann, Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments, Eur. J. Neurosci. 54 (2021) 8406–8420, https://doi.org/10.1111/ejn.14992.

[33] A. de Cheveigné, J.Z. Simon, Denoising based on spatial filtering, J. Neurosci. Methods 171 (2008) 331–339, https://doi.org/10.1016/j.jneumeth.2008.03.015.

[34] E. Lau, C. Phillips, D. Poeppel, A cortical network for semantics: (De)constructing the N400, Nat. Rev. Neurosci. 9 (2008) 920–933, https://doi.org/10.1038/nrn2532.

[35] E. Lau, K. Weber, A. Gramfort, M.S. Hämäläinen, G.R. Kuperberg, Spatiotemporal signatures of lexical-semantic prediction, Cereb. Cortex 26 (2016) 1377–1387, https://doi.org/10.1093/cercor/bhu219.

[36] L. Pylkkänen, A. Stringfellow, A. Marantz, Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density, Brain Lang. 81 (2002) 666–678, https://doi.org/10.1006/brln.2001.2555.

[37] L. Pylkkänen, R. Llinás, G.L. Murphy, The representation of polysemy: MEG evidence, J. Cogn. Neurosci. 18 (2006) 97–109, https://doi.org/10.1162/089892906775250003.

[38] L. Stockall, A. Stringfellow, A. Marantz, The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision, Brain Lang. 90 (2004) 88–94, https://doi.org/10.1016/S0093-934X(03)00422-X.

[39] P. Khader, A. Scherag, J. Streb, F. Rösler, Differences between noun and verb processing in a minimal phrase context: a semantic priming study using event-related brain potentials, Cogn. Brain Res. 17 (2003) 293–313, https://doi.org/10.1016/S0926-6410(03)00130-7.

[40] Q. Xia, G. Peng, The roles of object and action, and concreteness and imageability, in the distinction between nouns and verbs: an ERP study on monosyllabic words in chinese, J. Neurolinguistics 61 (2022) 101026, https://doi.org/10.1016/j.jneuroling.2021.101026.

[41] Q. Xia, L. Wang, G. Peng, Nouns and verbs in Chinese are processed differently: evidence from an ERP study on monosyllabic and disyllabic word processing, J. Neurolinguistics 40 (2016) 66–78, https://doi.org/10.1016/j.jneuroling.2016.06.002.

[42] S. Tsigka, C. Papadelis, C. Braun, G. Miceli, Distinguishable neural correlates of verbs and nouns: a MEG study on homonyms, Neuropsychologia 54 (2014) 87–97, https://doi.org/10.1016/j.neuropsychologia.2013.12.018.

[43] K.D. Federmeier, J.B. Sega, T. Lombrozo, M. Kutas, Brain responses to nouns, verbs and class-ambiguous words in context, Brain 123 (2000) 2552–2566, https://doi.org/10.1093/brain/123.12.2552.

[44] H. Wu, X. Liang, R. Wang, Y. Ma, Y. Gao, X. Ning, A Multivariate analysis on evoked components of Chinese semantic congruity: an OP-MEG study with EEG, Cereb. Cortex 34 (2024), https://doi.org/10.1093/cercor/bhae108.

[45] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory, Neuropsychologia 9 (1971) 97–113, https://doi.org/10.1016/0028-3932(71)90067-4.

[46] Y. Wang, P. De Weerd, A.T. Sack, V. Van De Ven, A sensitivity analysis of preprocessing pipelines: toward a solution for multiverse analyses Open Access, Imaging Neurosci. 2 (2024) 1–20, https://doi.org/10.1162/imag.

[47] Z.A. Seedat, K. St, N. Pier, M. Holmes, L. Rea, T.M. Al-Hilaly, C.M. Tierney, R. Embury, K.J. Pardington, J.H. Mullinger, E. Cross, M.J.B. Boto, Simultaneous whole-head electrophysiological recordings using EEG and OPM-MEG, Imaging Neurosci. 2 (2024), https://doi.org/10.1162/imag_a_00179.

[48] G.R. Kuperberg, K. Weber, N. Delaney-busch, C. Ustine, B. Stillerman, M. Hämäläinen, E. Lau, Neuropsychologia Multimodal neuroimaging evidence for looser lexico-semantic networks in schizophrenia : evidence from masked indirect semantic priming, Neuropsychologia 124 (2019) 337–349, https://doi.org/10.1016/j.neuropsychologia.2018.10.024.

[49] S. Hirano, K.M. Spencer, T. Onitsuka, Y. Hirano, Language-related neurophysiological deficits in Schizophrenia, Clin. EEG Neurosci. 51 (2020) 222–233, https://doi.org/10.1177/1550059419886686.

[50] X. Yu, S. Mancha, X. Tian, E. Lau, Shared neural computations for syntactic and morphological structures: evidence from Mandarin Chinese, Cognitive Science (in press).

[51] E. Xun, G. Rao, X. Xiao, J. Zang, Dashuju beijing xia BCC yuliaoku de yanzhi [the construction of the BCC corpus in the age of big Data], Yuliaoku Yuyanxue [Corpus Linguistics] 3 (2016) 93–109.

[52] W. Zhan, R. Guo, B. Chang, Y. Chen, L. Chen, Beijing Daxue CCL Yuliaoku de Yanzhi [the building of the CCL corpus: its design and implementation], Yuliaoku Yuyanxue [Corpus Linguistics] 6 (2019) 71–86.

[53] C.H. Liao, E. Lau, Enough time to get results? An ERP investigation of prediction with complex events, Lang. Cogn. Neurosci. 35 (2020) 1162–1182, https://doi.org/10.1080/23273798.2020.1733626.

[54] O. Feys, P. Corvilain, A. Aeby, C. Sculier, F. Christiaens, N. Holmes, M. Brookes, S. Goldman, V. Wens, X. De Tiège, On-scalp optically pumped magnetometers versus cryogenic magnetoencephalography for diagnostic evaluation of epilepsy in school-aged children, Radiology 304 (2022) 429–434, https://doi.org/10.1148/radiol.212453.
[55] A. Gramfort, M. Luessi, E. Larson, D.A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, M. Hämäläinen, MEG and EEG data analysis with MNE-Python, Front. Neurosci. 7 (2013) 1–13, https://doi.org/10.3389/fnins.2013.00267.
[56] N. Barascud, P. Alcocer, P. Roujansky, L. de Clermont, J.K. Cooper, J.J.T. Tresols, M. Szul, M. Pals, R. Quentin, V.P. Kumaravel, C. Royen, G. Ferraro, python-meegkit: v0.1.7, (2023). Doi: 10.5281/zenodo.10303516.
[57] J.O. Wobbrock, L. Findlater, D. Gergle, J.J. Higgins, The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures, Conf. Hum. Factors Comput. Syst. - Proc. (2011) 143–146, https://doi.org/10.1145/1978942.1978963.
[58] M. Kay, L.A. Elkin, J.J. Higgins, J.O. Wobbrock, ARTool: Aligned Rank Transform for Nonparametric Factorial ANOVAs. R package version 0.11.2., (2025).
[59] JASP Team, JASP 0.18.1, (2023).
[60] E.J. Wagenmakers, J. Love, M. Marsman, T. Jamil, A. Ly, J. Verhagen, R. Selker, Q. F. Gronau, D. Dropmann, B. Boutin, F. Meerhoff, P. Knight, A. Raj, E.J. van Kesteren, J. van Doorn, M. Šmíra, S. Epskamp, A. Etz, D. Matzke, T. de Jong, D. van den Bergh, A. Sarafoglou, H. Steingroever, K. Derks, J.N. Rouder, R.D. Morey, Bayesian inference for psychology. Part II: example applications with JASP, Psychon. Bull. Rev. 25 (2018) 58–76, https://doi.org/10.3758/s13423-017-1323-7.
[61] X. Zhang, C. Chen, M. Zhang, C. Ma, Y. Zhang, H. Wang, Q. Guo, T. Hu, Z. Liu, Y. Chang, K. Hu, X. Yang, Detection and analysis of MEG signals in occipital region with double-channel OPM sensors, J. Neurosci. Methods 346 (2020) 108948, https://doi.org/10.1016/j.jneumeth.2020.108948.
[62] S.G. Kim, D. Poeppel, T. Overath, Modulation change detection in human auditory cortex: evidence for asymmetric, nonlinear edge detection, Eur. J. Neurosci. (2020) 1–16, https://doi.org/10.1111/ejn.14707.
[63] S.G. Kim, T. Overath, W. Sedley, S. Kumar, S. Teki, Y. Kikuchi, R. Patterson, T. D. Griffiths, MEG correlates of temporal regularity relevant to pitch perception in human auditory cortex, Neuroimage 249 (2022), https://doi.org/10.1016/j.neuroimage.2022.118879.
[64] E. Lau, E. Nguyen, The role of temporal predictability in semantic expectation: an MEG investigation, Cortex 68 (2015) 8–19, https://doi.org/10.1016/j.cortex.2015.02.022.
[65] B. Maess, F. Mamashli, J. Obleser, L. Helle, A.D. Friederici, Prediction signatures in the brain: semantic pre-activation during language comprehension, Front. Hum. Neurosci. 10 (2016) 1–11, https://doi.org/10.3389/fnhum.2016.00591.
[66] R. Assadollahi, F. Pulvermüller, Early influences of word length and frequency: a group study using MEG, Neuroreport 14 (2003) 1183–1187, https://doi.org/10.1097/00001756-200306110-00016.
[67] L. Gwilliams, G.A. Lewis, A. Marantz, Functional characterisation of letter-specific responses in time, space and current polarity using magnetoencephalography, Neuroimage 132 (2016) 320–333, https://doi.org/10.1016/j.neuroimage.2016.02.057.
[68] J. Särelä, H. Valpola, Denoising source separation, J. Mach. Learn. Res. 6 (2005) 233–272.
[69] T.M. Tierney, N. Alexander, S. Mellor, N. Holmes, R. Seymour, G.C. O'Neill, E. A. Maguire, G.R. Barnes, Modelling optically pumped magnetometer interference in MEG as a spatially homogeneous magnetic field, Neuroimage 244 (2021) 118484, https://doi.org/10.1016/j.neuroimage.2021.118484.
[70] A. Hillebrand, G.R. Barnes, Beamformer analysis of MEG data, Int. Rev. Neurobiol. 68 (2005) 149–171, https://doi.org/10.1016/S0074-7742(05)68006-3.
[71] J.-Z. Wang, S.J. Williamson, L. Kaufman, Magnetic source images determined by a lead-field analysis: the unique minimum-norm least-squares estimation, IEEE Trans. Biomed. Eng. 39 (1992) 665–675, https://doi.org/10.1109/10.142641.
[72] M.S. Hämäläinen, R.J. Ilmoniemi, Interpreting magnetic fields of the brain: minimum norm estimates, Med. Biol. Eng. Comput. 32 (1994) 35–42, https://doi.org/10.1007/BF02512476.
[73] A.M. Dale, A.K. Liu, B.R. Fischl, R.L. Buckner, J.W. Belliveau, J.D. Lewine, E. Halgren, Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity, Neuron 26 (2000) 55–67, https://doi.org/10.1016/S0896-6273(00)81138-1.
[74] N. Holmes, M. Rea, J. Chalmers, J. Leggett, L.J. Edwards, P. Nell, S. Pink, P. Patel, J. Wood, N. Murby, D. Woolger, E. Dawson, C. Mariani, T.M. Tierney, S. Mellor, G. C. O'Neill, E. Boto, R.M. Hill, V. Shah, J. Osborne, R. Pardington, P. Fierlinger, G. R. Barnes, P. Glover, M.J. Brookes, R. Bowtell, A lightweight magnetically shielded room with active shielding, Sci. Rep. 12 (2022) 1–13, https://doi.org/10.1038/s41598-022-17346-1.
[75] M. Władziński, A. Jodko-Władzińska, T.H. Sander, Active Compensation for OPM-MEG Inside a Two-Layer Magnetically Shielded Room, in: 9th Eur. Med. Biol. Eng. Conf. EMBEC 2024. IFMBE Proc., 2024: pp. 48–54. Doi: 10.1007/978-3-031-61625-9_6.
[76] G. Tan, J. Gai, F. Li, Q. Huang, W. Yu, Y. Huang, G.Y. Zhang, X. Zhao, Q. Lin, Z. Hu, A Brain-Computer Interface System based on Magnetoencephalography with Optically Pumped Magnetometers, IEEE Trans. Instrum. Meas. PP (2025) 1. Doi: 10.1109/TIM.2024.3522701.