Perceptual facilitation of word recognition through motor activation during sentence comprehension

Nicola Molinaro[1,2] & Irene F. Monsalve[1]

[1]BCBL, Basque center on Cognition, Brain and Language, Donostia/San Sebastian, Spain

[2]Ikerbasque, Basque Foundation for Science, Bilbao, Spain

Correspondence to:
Nicola Molinaro
BCBL, Basque center on Cognition, Brain and Language
Paseo Mikeletegi, 69, 2º
20009, Donostia/San Sebastian, Spain
Phone: +34 943 309 300 ext. 206
Email: n.molinaro@bcbl.eu
ORCID: 0000-0002-7549-6042

ABSTRACT

Despite the growing literature on anticipatory language processing, the brain dynamics of this high-level predictive process are still unclear. In the present MEG study, we analyzed pre- and post-stimulus oscillatory activity time-locked to the reading of a target word. We experimentally contrasted the processing of the same target word following two highly constraining sentence contexts, in which the constraint was driven either by the semantic content or by the lexical association between words. Previous research suggests the presence of sensory facilitation for expected words in the latter condition but not in the former. We observed a dissociation between beta (~ 20 Hz) and gamma (> 50 Hz) band activity in pre- and post-stimulus time intervals respectively. Both the beta and gamma effects were evident in occipital brain regions, and only the pre-stimulus beta effect additionally involved left pre-articulatory motor regions. Lexically constrained (vs. semantically constrained) words elicited reduced beta power around 400 ms before the target word in motor regions and a functionally related gamma enhancement in occipital regions around 200 ms post-target. The present findings highlight the role of the motor network in word-form prediction and support proposals claiming that low-level perceptual representations can be pre-activated during language prediction.

Keywords: Language prediction, Motor activation, Reading, Magnetoencephalography, Neuronal oscillations

INTRODUCTION

The ease and speed with which language processing unfolds has been explained by the human ability to anticipate information about the incoming input (Federmeier, 2007; Friston & Frith, 2015; Levy, 2008; Lewis & Bastiaansen, 2015). Thus, while listening to or reading a text, the comprehender incrementally integrates internal semantic knowledge that in turn provides constraining information concerning the meaning expressed in the remaining part of the message. This process, however, has been poorly detailed from both cognitive and neurophysiological perspectives, with research often focusing more on the consequences of such anticipatory analysis (for discussion see Molinaro et al., 2013; 2016). The most consistent finding from this literature is that contextual information facilitates the lexical recognition of a predicted word, as evidence by reduced evoked activity around 400 ms post-stimulus onset. This has led to the conclusion that the lowest representational level that top-down predictive processing can affect is the lexical/semantic level, i.e., an abstract representational level independent of the low-level formal properties of a stimulus. Nonetheless, the great majority of these studies employed regular compositional contexts, where predictions would be generated at the semantic level, probably through the pre-activation of specific semantic features or categories. As a consequence, lexical pre-activation could disperse to the semantic neighbors of the expected word (Federmeier et al., 1999) and not be strong enough to activate word-form-related sensory representations. In addition, semantic relatedness and lexical relatedness are not highly correlated, so activating semantic neighbors might be unrelated to activating lexical neighbors. However, in a recent set of studies (Molinaro et al., 2013; Monsalve et al., 2014; Vespignani et al., 2010), we took advantage of a qualitatively different relation between words in natural language, i.e., the lexical association (Hutchison, 2003, for a review). Associative relations reflect links between lexical items due to the number of co-occurrences in natural language. Sequences of words that co-occur in natural language more than would be expected by chance are broadly defined as multi-word units (MWU). These sequences are often recognized before the last word(s) of the string, so that given the initial fragment of the sequence, the predictability of the last word(s) is usually very high. In this scenario, the prediction would be generated directly at the lexical level, pre-activating the corresponding word-form features. No semantic dispersion would take place (if at all, dispersion would in this case involve orthographic neighbors) and the specific lexical item would have the "strength" to pre-activate word-form-related sensory representations. In Molinaro et al. (2013) we employed these stimuli to study language prediction: we compared the electrophysiological correlates of processing expected words following either a compositional (semantic) or a collocational (MWU) context. We reported early post-stimulus evoked electrophysiological responses around 120 ms to be reduced in the MWU condition compared to the semantic condition (even if in both conditions the context was highly constraining), suggesting facilitated visual sensory processing of word-forms in MWU. Furthermore, pre-stimulus oscillatory activity was related to this early post-stimulus effect. Phase-synchronization in the theta range (5-6 Hz) across the EEG signals was stronger for the MWUs compared to the semantic constraining condition. This effect was interpreted as reflecting differential processing cost in the generation of predictions. In a re-analysis of the Molinaro et al. (2013) dataset (Monsalve et al., 2014), we examined the oscillatory post-stimulus activity at the single-trial level, finding a positive correlation between cloze probability and gamma power (> 50 Hz) for both conditions, but overall higher gamma for MWU compared to the semantic condition. This is congruent with other studies finding positive associations between gamma and predictability in language processing, which could result from a stronger perceptual binding between the expected internal representation and the actually encountered stimulus (Hermann et al., 2004; for a discussion Lewis & Bastiaansen, 2015). While this would be the case for any highly constraining condition, the matching is always “greater” for MWU, which could arise from more perceptually-detailed expectations in this case. These constructions thus provide a unique opportunity to explore how sensory predictions unfold in the language domain.

Neurophysiological evidence on sensory processing points to a relevant role of brain oscillations in the interaction between internal expectations and the physical properties of the input. Distinct frequency bands have been implicated in the feedback propagation of predictions and feedforward propagation of prediction error (beta - 13-30 Hz - and gamma - > 30 Hz - respectively) through the processing hierarchy (among others: Arnal et al., 2015; Bastos et al., 2015; Richter et al., 2017). However, in the language literature, support for this scheme is mixed (Lewis & Bastiaansen, 2015). Previous studies of prediction in language have provided support for a top-down role of beta oscillations, in line with the literature from sensory processing. Indeed, in Molinaro et al. (2017) we reported differential beta oscillatory activity (13-30 Hz) in a time interval preceding an expected stimulus. However, the anatomical origins of these predictions are unclear. Wang and colleagues (2017) compared the processing of highly expected vs. low expected words during sentence comprehension. They reported differential effects in the alpha (~10 Hz) and beta (~20 Hz) range involving a left frontal-temporal brain network with lower alpha and beta power in the high vs. the low constraining conditions. In line with this, a large-scale language comprehension study (Schoffelen et al., 2017) reported prefrontal-to-temporal interactions, mainly peaking in the beta frequency channel. These studies support the role of beta activity in top-down generation of predictions, localizing it to a pre-frontal area, as would follow from classic models of language processing (Friederici, 2012; Hagoort, 2017; Lau et al., 2008). On the other hand, Piai and colleagues (2015) localize predictive beta-band activity to motor cortex. They analyzed the oscillatory activity elicited by semantically highly-constraining vs. low-constraining sentence contexts before processing of a target picture. Participants had either to perform a naming (production) or a judgment (comprehension) task. They observed similar beta desynchronization effects involving both left posterior temporal and motor regions, for both tasks, in the highly-constraining condition. The involvement of the motor system in prediction has indeed been put forward in different theoretical proposals (Pickering & Garrod 2007, 2013, Dell & Chang, 2014; Molinaro et al., 2016; Federmeier, 2007).

For what concerns the oscillatory correlates of bottom-up information flow in language processing, experimental evidence seems at odds with what has been reported in the sensory processing literature. As discussed earlier, findings relating to gamma in the language literature typically show a positive relationship between power and predictability, congruent with Hermann et al.'s model (Hermann et al., 2004). In addition, the study by Wang and colleagues (2017) found post-stimulus gamma modulations related to prediction in the left pre-frontal cortex, rather than in perceptual (visual) regions. Furthermore, alpha activity in the temporal lobe modulated gamma oscillations in the prefrontal lobe, an “opposite” dynamic to what sentence processing models hypothesize (prefrontal-to-temporal). In line with this, the data from Schoffelen et al. (2017) also revealed temporal-to-prefrontal interactions in the alpha band.

Nonetheless, these studies typically employed semantically constraining contexts to manipulate target word predictability. The high vs. low constraint manipulation does not elicit differential effects in visual brain regions (e.g., Wang et al., 2017). In contrast, as we discussed earlier, MWU might generate specific word-form sensory expectations and, in this way provide better contexts to study top-down influence of language representations on visual processing. Furthermore, the traditional high and low-constraining sentence conditions can highlight a larger set of processes than the prediction ones. For example, the high vs. low-constraining manipulation calls for different attentional demands. In the eye-movement literature (e.g., Hawelka et al., 2015, Staub, 2015), it has been shown that highly expected words tend to be skipped more during sentence reading compared to low-expected words. Summerfield & Egner (2009; 2016) underscore the importance of dissociating between the cognitive processes involved in perceptual attention from the ones involved in sensory prediction.

In the present study, we aim to take advantage of MWUs and magnetoencephalographic (MEG) recordings to characterize the oscillatory and anatomical profile of the prediction process during language comprehension. We use the same design previously described in the Molinaro et al. (2013) EEG study. Firstly, this will allow us to replicate the early sensory facilitation for MWU processing previously reported. Secondly, using MEG we can determine not only the oscillatory dynamics of prediction, but also the anatomical substrate supporting this process. We focus on pre- and post- stimulus oscillatory power, in order to better compare our results with the literature of sensory processing. This measure (differently from phase-related connectivity measures) has the advantage of providing evidence about which brain regions present specific coordinated activity (synchronized or desynchronized) of the large populations of neurons involved in computing language pre-activation.

We expected beta effects for both the Semantic and the MWU conditions during the pre-target time interval, as an index of generation of prediction. In the post-stimulus interval, the gamma oscillations should reflect perceptual processing of the target word. While the post-stimulus gamma effect should peak in occipital brain regions, it is not clear which brain network should elicit the pre-stimulus beta effect: either a frontal-temporal network (as proposed by sentence processing models) or a sensory-motor network (as proposed by motor theories of prediction) can be hypothesized. It will then be crucial to evaluate if the experimental manipulation here employed differently recruits the beta/gamma oscillatory channels and how such modulation can be interpreted. As a first instance, the MWU-Semantic experimental contrast could produce effects in the beta/gamma frequency ranges. If the pre-stimulus beta effect reflects generation of prediction, we should expect less beta power for the MWU compared to the Semantic condition (see Arnal & Giraud, 2012). In line with previous findings, the post-stimulus gamma power should be higher in the MWU condition, evidencing perceptual binding. However, the emergence of alternative scenarios could call for the proposal of a different oscillatory framework involved in language prediction (see Lewis & Bastiaansen, 2015; Wang et al., 2017).

METHODS

Participants

Thirty-five participants took part in the experiment (twenty-one females; mean age: 26.54 years, Standard Deviation: 3.44). None of them took part in a pilot experiment discussed in Molinaro et al. (2016). Due to excessive noise in the recording, two participants were not considered in the following analyses. All were paid 10€ per hour for taking part in the experiment. All of them were right-handed and reported Spanish as the language they acquired first. When presented with pictures of increasing difficulty (decreasing their frequency of use) their average naming score in Spanish was 64.75/65. It should be noted that participants reported knowing some Basque (34.96/65) and English (35.36/65), as is common in the young adult population of Donostia/San

Sebastian. None of the participants reported any neurological or visual impairments. The present study was approved by the BCBL Ethics Committee.

Materials

We selected a set of sentences from a previous study (Molinaro et al., 2013). To avoid semantic compositionality issues, we avoided using multi-word expressions with a figurative meaning: the meaning of the whole multi-word unit could be derived by the composition of the meaning of the individual words. We then adapted the sentences for the goals of the present experiment. Overall, we employed 60 sentence pairs in which the same target word could follow (i) either a sentence context in which no fixed expression was present (as reported by five independent native speakers) and the constraint was mainly driven by the semantics of the overall message (ii) or the initial fragment of a multi-word unit. The multi-word units were three-to-six words long and extracted by the CESS-ECE corpus (Martí & Taulé, 2007). On a 1-to-7 scale the MWU were all highly familiar: In a questionnaire administered to 10 independent native Spanish speakers no string was rated with less than 6. The main difference between the two constraining contexts was the lexical association strength. To make the sentence set as natural as possible we employed sentences of highly variable length (16-to-27 words), so that participants could not develop unnatural reading strategies during the experiment due to the "fixed" structure of our stimuli (see Appendix A). Nonetheless, we kept differences between the two experimental conditions as minimal as possible. Cloze-probability of the target word (tested in a group of 40 independent native Spanish speakers, asked to report the most natural continuation of a sentence in a questionnaire) was balanced (Semantic: 85.78, Standard Deviation: 15.20; MWU: 85.75, SD: 18.01), so that in each pair, scores were comparable. Sentence position was the same for each pair (18.33, SD: 2.46; never at sentence ending). Importantly, the lexical and sentence-level parameters of the word preceding the target were controlled. In the majority of cases, this was a function word (Semantic: 31; MWU: 34). Lexical parameters of the words preceding the target (word length: Semantic: 3.65, SD: 1.88; MWU: 3.21, SD: 1.79; logFrequency: Semantic: 3.19, SD: 1.41; MWU: 3.54, SD: 1.18; Levenstein distance: Semantic: 1.20, SD: 0.34; MWU: 1.15, SD: 0.24) did not differ between groups (all p-values >0.15). Importantly, the cloze-probability of these words did not differ (Semantic: 37.83, SD: 14.03; MWU: 39.67, SD: 14.49; p=0.5). As already noticed in Molinaro et al. (2013), an important difference emerged in the "conditional cloze-probability" of reporting the target word while considering only the cases in which the pre-target word was reported. For the Semantic condition, the target word was reported in 60.42% (SD: 23.12) of the cases in which the pre-target was reported; for the MWU condition the conditional cloze probability was significantly higher (92.56, SD: 13.67; t(118)=9.2690, p< 0.01). This indicates that when correctly identifying the MWU before the target, participants probably have already pre-activated the target word (the final word of the unit) unambiguously. On the other hand, when reporting the word preceding the target in the Semantic condition, a set of lexical candidates (including the target) were pre-activated. Overall, this indicates that a more lexically-specific prediction is active in the MWU condition compared to the Semantic.

In addition to these 120 sentences, a further 206 filler sentences were employed. These sentences also included items with a lower contextual constraint so that participants were presented with a more natural set of stimuli. Without fillers, it could be argued that a more strategic language prediction process could have been developed by participants during the experiment. A total of 326 sentences was read by each participant.

Experimental procedure

Prior to recording each participant, we digitized the positions of three landmarks (nasion and pre-auricular points) and four head-position indicator (HPI) coils, placed on the high forehead and above the ears. We also digitized the head shape to improve the co-registration accuracy between MEG and individual structural MRI. During the MEG recordings, the participants sat on a comfortable chair in the magnetically shielded room with their head in the MEG sensor helmet. Participants were instructed to limit head and face movements as much as possible and to fixate on the center of the screen.

The sentences were presented visually (one word at a time) through a projector onto a back-projection screen (white letters on a dark gray background, no word exceeded a visual angle of about 5°). Each trial started with a fixation cross at the center of the screen for 500 ms followed by a 300 ms blank. Words were displayed in white letters on a dark gray background. Each word was presented for 300 ms followed by a 300-ms blank screen, the inter-trial interval was variable (500-1700 ms). Sentence order was fully randomized and every five sentences on average, participants were asked to answer a YES/NO comprehension question by pressing the corresponding button on a joypad with their left/right index finger; comprehension questions appeared randomly across the whole experiment. YES/NO button position appeared randomly either on the left (half of the times) or on the right (half of the times). Participants across the experiments showed a very good average level of accuracy (94.3%), varying between 88% and 97%.

To familiarize participants with the experimental procedure, before the experimental session participants had to perform 12 practice trials in the presence of the experimenter. The experiment lasted around 1 h and 30 min, with a break every 15 min.

At the end of the MEG experimental session, participants were administered a questionnaire in which they were presented with a random set of 20% of the MWU used in the experiment without the last target word.

They had to complete it. All the participants correctly completed more than 90% of the stimuli showing that they were familiar with the MWUs used in the experiment.

Data Acquisition

Brain activity was recorded using a 306-channel Vectorview system (Elekta, Helsinki, Finland) at BCBL (Basque Center on Cognition, Brain and Language, Donostia, Spain). Ocular activity was monitored with four electrodes arranged in bipolar montages: two electrodes (on top and below the left eye) constituted the vertical EOG channel and two electrodes placed on the external canthi of both eyes constituted the horizontal EOG.

MEG was acquired at 1 kHz sampling rate and on-line filtered to 0.1–330 Hz. Head movements were monitored continuously using the four HPI coils attached to the participant's head. For each participant, a high-resolution 3D structural MRI (T1-weighted MPRAGE sequence) was acquired with a 3T Trio scanner (Siemens, Munich, Germany).

Data Analysis

Data were preprocessed off-line using the Signal-Space-Separation method (Taulu et al., 2005) implemented in Maxfilter 2.1 (Elekta-Neuromag) to subtract external magnetic noise from the MEG recordings. The MEG data were also corrected for head movements and bad channels were substituted using interpolation algorithms implemented in the software.

The subsequent analyses were performed using Matlab R2012 (Mathworks, Natick, MA, USA) and toolboxes such as Fieldtrip (Oostenveld et al., 2011) and SPM8 (Wellcome Department of Cognitive Neurology, London, UK). The recordings were segmented for each trial (from –1.2 to 1 s) time-locked to the fixation point onset - to be used as a baseline period - and time locked to the target word. Data were filtered with a low-pass filter (cutoff: 150 Hz) and a DFT filter to remove line noise. A semi-automatic procedure was then employed to remove epochs with muscular and jump artifacts and epochs with flat signal. Eye movements, blinks and electrocardiographic artifacts were reduced using independent component analysis (Jung et al., 2000). Further sensor-data analysis was performed using only gradiometers, but both magnetometers and gradiometers were employed during source-localization.

Sensor-level analyses

Evoked response. Data concerning the time-course for each epoch were low-passed filter at 35 Hz and then averaged independently for each condition and participant. The signals were baseline corrected to the pre-sentence interval (–1000 to 0 ms prior to fixation point). In order to evaluate the difference between conditions, we did not combine the two gradiometer directions to preserve information about the directionality of a magnetic response. This was also done to compare the magnetic responses at the sensor-level to the EEG responses from Molinaro et al. (2013). The statistical comparison of the two conditions focused on the time interval 100-150 ms post-stimulus onset (based on previous EEG evidence). Signals for each gradiometer were compared with a parametric dependent-sample t-test (p-values were fdr corrected).

Time-frequency. Two separate analyses were performed for the low and high portion of the frequency spectrum. Time-frequency representations over 2-30 Hz range were obtained using Hanning tapers and a fixed window length of 500 ms advancing in 10 ms steps, giving rise to a 2 Hz frequency resolution. In the 30-100 Hz frequency range a multitaper approach was employed with a window length of 400 ms (10 ms steps, 200 ms time-smoothing) and the frequency resolution of 2.5 Hz. Due to filter edge artifacts, we further segmented the epochs of interest in the –700 to 600 ms time locked to the target word. Single subject estimates were then obtained by averaging power estimates for each trial separately for the different conditions. Power estimates were calculated separately for each orthogonal direction of a gradiometer pair and then averaged, resulting in 102 measurement channels. Subsequently, power was expressed as relative change with respect to a ~1 sec baseline interval (–1000 to –50 ms prior to fixation point) before sentence presentation.

Differences between conditions (MWU vs. Semantic) were assessed using cluster-based permutation tests (Maris & Oostenveld, 2007), that control for multiple comparisons while maintaining sensitivity by taking into account the temporal, spatial and frequency dependency of neighbouring samples. First, data were clustered by performing pairwise comparisons (MWU vs. Semantic) between each sample (time-frequency-sensor point) per condition. Contiguous values exceeding a p=0.05 threshold were grouped in a cluster, and a cluster-based statistic was derived by adding the t-values within that cluster. Then, a null distribution assuming full exchangeability (i.e. no difference between conditions) was approximated by drawing 1000 random permutations of the observed data and calculating the cluster-level statistics for each randomization. Finally, the cluster-level statistics observed in the actual data were evaluated under this null distribution. Two independent analyses were carried out for low (2-30 Hz) and high (30-100 Hz) frequency ranges in the whole interval (–700 to 600 ms).

Source-level analyses

Source reconstruction was carried out in order to examine the spatial specificity of the experimental effects obtained at the sensor level. The rationale of this analysis was to first source-reconstruct the main effects of interest for the two experimental conditions compared to baseline. This was done to identify the spatial location of the peaks of activity in which oscillatory activity differed from baseline. Afterwards, we sourcereconstructed the time-course of the frequency band of interest for each condition at the previously identified peak location. Thus, rather than comparing anatomical maps of activation in each condition, we focused on the peaks of activity for all conditions pooled together. In MEG research, contrast maps confound amplitude and location differences, thus suffering the mislocation problem, i.e., differences may peak away from the true generating sources (Bourguignon et al., 2018).

The forward solution was based on the anatomical image (T1) of each individual participant. MRIs were segmented using Freesurfer software (Dale & Sereno, 1993; Fischl et al, 1999). The forward model was based on a one-shell boundary element model of the intracranial space. It was computed for three orthogonal directions of sources, which were placed on a 5-mm grid covering the whole brain using MNE suite (Martinos Center for Biomedical Imaging, Massachusetts, USA). For each grid point (three directions), the forward model was then reduced to its two principal components of highest singular value, which closely correspond to sources tangential to the skull. We used both gradiometers and magnetometers in the source estimation, normalizing each sensor signal by its noise variance estimated from the baseline prior to the fixation point in all conditions. For the evoked activity, we computed the evoked responses and the covariance matrix in the time-window of interests and in a equal-size baseline interval. Linearly Constrained Minimum Variance Beamformer (LCMVB) focused on the real part of the covariance matrix as inverse model of the evoked response at ~130 ms. For the time-frequency estimates, power and cross-spectral density matrices were estimated in the selected time-frequency windows and in an equal-size baseline period prior to fixation point. LCMVB was applied to the real part of the cross-spectral density matrix as inverse model restricted to the frequency of interest identified in the sensor-level analyses. For both evoked responses and time-frequency estimates, power in the windows of interest was normalized by power in the baseline and log-transformed.

A brain map was then obtained for each possible combination of participant and time-frequency window of interest. A non-linear transformation from individual MRIs to the standard Montreal Neurological Institute (MNI) brain was first computed using the spatial-normalization algorithm implemented in Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK). This was then applied to every individual map. Statistical significance was computed employing the maximum statistic permutations approach (Nichols & Holmes, 2001). The sampling distribution of the maximal difference of power values was evaluated using the exhaustive permutation test. Source-level grid points for which the non-permuted maximal difference exceeded the 95 percentile of this permutation distribution were defined as sources of interest.

We then determined the local maxima (or minima) showing maximal average power change values in the source maps of interest. For the evoked responses, the power in the peak location of the maxima was contrasted with a parametric t-test. For time-frequency estimates, time-courses from these grid points were obtained with the LCMV beamformer by employing a forward solution restricted to the local maxima(/minima). Here, a single time-course was obtained for each of these grid points by computing the average of the two resulting orthogonal tangential dipoles. The time course reflects the power change for each condition (MWU and Semantic) in the frequency of interest. Shaded error bars (considering the standard error) were then plotted to evaluate the differences between the time courses of the two conditions.

RESULTS

Evoked responses

Evoked responses time-locked to the presentation of each word showed a clear peak around 130 ms in the posterior sensor regions. This peak following the target word was larger for the Semantic compared to the MWU condition. Parametric statistics showed that this increased response for the Semantic condition was mainly evident in a set of posterior sensors (Figure 1A). Source reconstruction of this effect highlighted that the peak (for both conditions) was mainly driven by activity in the occipital cortex (Figure 1B). Two local maxima were identified, one for each hemisphere. Power differences between the two conditions were mainly evident in the right occipital peak with power being larger for the Semantic compared to the MWU condition (Figure 1C). The present findings are in line with effects observed in a previous EEG study (Molinaro et al., 2013).

Time-frequency: Sensor-level effects

In the lower frequency range (2-30 Hz), the reading of each word (for all conditions) elicited an early (100-200 ms) positive theta burst (5-8 Hz) in posterior brain. In the beta range (13-30 Hz) we observed a power reduction that was more pronounced in the –600 up to –200 ms interval. This beta effect was evident mainly in the posterior and in the left lateralized sensors (Figure 2A). Cluster-based permutations (MWU vs. Semantic) in the lower frequency range (2-30 Hz) highlighted a significant difference between the two experimental conditions (Figure 2B). The beta range (18-22 Hz) negative effect was significant in the pre-target word time interval starting around –600 up to –330 ms ($p = 0.028$). This comprised a period of reduced power for both the experimental conditions compared to the pre-sentence baseline. In this time period, the MWU condition triggered significant less power compared to the Semantic mainly in the left-anterior sensors (Figure 2B).

In the higher frequency range (30-100 Hz) there was a positive gamma effect peaking around 300 ms post-stimulus onset. This effect was more pronounced in the posterior sensors. Here, the analyses highlighted a significant difference between the two experimental conditions in the post-target word time interval. In the high-gamma range (70-95 Hz) a positive effect emerged in the posterior sensors from 200 to 400 ms (p = 0.026), that was larger for the MWU condition, with a slightly right-lateralized distribution (Figure 2C).

Time-frequency: Source-level effects

Firstly, we performed source reconstruction of the time-frequency windows of the effects identified in the sensor-level analysis. Secondly, we located the anatomical peaks of activity of both conditions pooled together with respect to baseline, for each of the identified effects. Thirdly, we compared the time-courses for each condition, at each identified peak, for each frequency band.

Pre-target word beta: The beta band (18-22 Hz) power decrease in the time interval preceding the target word involved the occipital regions bilaterally and the left primary motor cortex (Figure 3A), congruent with the topographies observed at sensor level. Local minima were detected in each of the occipital cortices and in the left motor cortex. The time courses revealed differential trends for the two conditions in the left motor cortex in a time interval coinciding with the cluster identified in the sensor analysis (between –600 and –370 ms, see Figure 3B). Power was lower for the MWU compared to the Semantic condition.

Post-target word gamma: In the gamma band (70-95 Hz) the increased activity compared to baseline observed in the post-target word time interval was located in the occipital cortex. A local maximum was detected in each of the two cortices (Figure 4A). The time course of the effect revealed increased gamma activity for the MWU condition compared to the Semantic in the right occipital regions between 200 and 400 ms (Figure 4B).

Relation between pre- and post-stimulus effects: We evaluated the relation between these two effects ([MWU – Semantic] for beta and gamma) across participants. Given the large individual variability of the MEG recordings (due to multiple factors, such as brain morphology, different SNRs, variable distance from the sensors, etc.), we focused on the differential effects between the two experimental conditions. We correlated the beta and gamma effects (MWU minus Semantic; left motor source for beta averaged over a –600 and –370 ms time window, and right occipital for gamma over 200 to 400 ms) and observed a significant negative correlation (r = –0.39, p = 0.024). This indicates that participants showing a larger beta power decrease in the MWU condition compared to the Semantic also showed larger gamma power increase for the same experimental comparison.

DISCUSSION

In the present MEG sentence comprehension experiment, we observed oscillatory cortical effects that provide relevant hints concerning the neurophysiological dynamics supporting language anticipatory phenomena. By means of a reading paradigm, we here report (i) in the time interval preceding an expected target word, there is reduced oscillatory power (with respect to baseline) in the beta frequency range that involves both left motor and bilateral occipital cortical regions; power is lower in the ventral motor regions for more accurate lexical predictions (see also Molinaro et al., 2017); (ii) in the post-target word time interval, there is increased gamma band power for expected stimuli in bilateral occipital regions that is even larger for more lexically specific predictions (as reported in Monsalve et al., 2014); this gamma-band occipital modulation is related to the pre-stimulus motor beta effect. The post-stimulus gamma effect observed in occipital regions (see Hoogenboom et al., 2006, 2010) is in line with the perceptual facilitation effect observed in Molinaro et al. (2013) and replicated here. Indeed, the reduced evoked response around 120-130 ms for the MWU condition compared to the Semantic condition reflects facilitated perception of the expected stimulus. It thus seems that the experimental design employed here is confirmed as successful in identifying perceptual facilitation of the expected stimulus. This places us in the best scenario to compare the present language prediction effects with the literature on sensory prediction. In the following sections, we discuss the implications of the beta and gamma effects observed here.

Pre-stimulus beta

The beta oscillatory channel has been previously associated with two distinct domains in the neuroscientific literature: planning and execution of motor commands (for a review Kilavik et al., 2013), and more abstract cognitive domains (Engel & Fries, 2010; Weiss & Mueller, 2012) including predictive coding across different modalities (e.g., Arnal et al., 2015; Bastos et al., 2015; Richter et al., 2017). However, the beta modulations in these two research lines can be linked through the involvement of the motor system in the generation of sensory predictions (as we proposed in Molinaro et al., 2016). The beta evidence observed in the present study can be interpreted in this light. On the one hand, the two predictive conditions trigger strong beta desynchronization in occipital regions (congruent with a visual expectation, as expected in a reading experiment). On the other, left motor regions also exhibit strong beta desynchronization during this period, peaking in the more ventral portion of the primary motor cortex, in the vicinity of the regions involved in pre-articulatory planning (Murphy et al., 1997). Furthermore, the condition in which lexical pre-activation is more specific elicits less beta power, as observed in previous studies (Molinaro et al., 2017; Wang et al., 2017). This suggests that the generation of sensory word-form predictions might involve motor articulatory areas (as proposed by motor theories of language prediction: Federmeier, 2007; Pickering & Garrod, 2007), recruiting less beta power when the prediction is stronger (MWU).

The activation of the articulatory motor cortical regions in a reading experiment has been already observed in the neuroimaging literature. Kell et al. (2017) in fact reported an fMRI study in which the brain networks activated during overt speech and during silent reading are largely overlapping, and, more importantly, they similarly elicit increase of activation in left motor regions. The authors discuss the motor activation observed during silent reading as evidence of inner speech processes. Crucially, fine-grained speech information such as consonant voicing can be similarly decoded during both overt and inner speech (silent reading) conditions. It is thus reasonable to assume that the language production system is active during language comprehension. This is important since Clark (2015) discusses that the ability of performing mental simulations is consequence of predictive processing skills. In the language domain, the possibility of simulating speech during reading can thus support pre-activation of word-form representations that can in turn facilitate perception of the incoming perceptual input.

Another interesting aspect concerns the time-course of the beta modulation. Kilavik et al. (2013) highlight how beta power activity can be modulated by the predictive value of a warning cue before performance of a task-related motor command. The authors point out that beta power (that peaks at cue onset and later decreases) during motor preparation is lower if the cue is less informative compared to a more informative cue. In the present experiment, we observed a beta effect (MWU < Semantic) that peaked when participants were reading the word preceding the target and rapidly decreased afterwards (Figures 2B and 3B). This parallel provides support for the claim that the oscillatory beta activity in motor regions observed during silent reading can be associated to predictive processing. While reading silently for comprehension, the neurocognitive system recruits the language-related motor regions to a larger extent when the incoming input is more predictable. Internal knowledge-based expectations are top-down “converted” into more specific form-related representations by employing the same internal cognitive mechanisms at work for language production, possibly mediated by a fine-grained phonological code. The important difference is that during “silent” language prediction, the motor command is not implemented, while the sensory component of the motor command would be available to constrain (visual) perception. Classic theoretical models of reading emphasise that the comprehension system activates language-related phonological representations (e.g., McClelland & Rumelhart, 1981). During reading acquisition, phonological representations mediate the mapping of visual codes onto meaning (Ans et al., 1998). Phonological information can thus be available for mediation between the speech production system and the reading system. This working hypothesis highlights a relevant role of phonology even for predictive processing in silent reading.

Post-stimulus gamma

The gamma modulation (MWU > Semantic) that we here report does not align with what would be expected by the predictive coding hypothesis. This effect involves primary visual regions, as would be expected for a reading experiment in which expectation of the visual stimulus is more or less specific. However, in the predictive coding view, prediction error would be higher for the less expected stimulus (Friston, 2005), leading to increased gamma power for the Semantic compared to the MWU condition (and this logic better applies to the evoked response around 130 ms, Figure 1). However, we observed the opposite effect, with increased gamma response for the more detailed prediction. This aligns with a number of previous studies reporting increased gamma effect for more expected items (among others, Canal et al., 2017; Monsalve et al., 2014; Wang et al., 2012). To account for the contrasting gamma evidence emerging from the language prediction literature, compared to the predictive coding proposals (Bastos et al., 2015; Richter et al., 2017), Lewis and Bastiaansen (2015) proposed a division of labour between higher (> 50 Hz) and lower (30-50 Hz) gamma frequency effects. While the higher frequency oscillations would reflect predictive coding-related modulations, the lower component would reflect perceptual binding between an internally activated low-level representation and the properties of the external input. Hermann et al. (2004) in fact proposed that the occipital gamma response observed from intracranial recordings in monkeys reflects the parallel firing of neurons representing the input in a specific receptive field and neurons representing top-down expectations about the sensory input. Synchronous firing would result in higher gamma power, while more asynchronous firing in reduced gamma power. Our findings are better explained by the perceptual binding hypothesis, in which a more detailed pre-activated representation determines a better match with stimulus-related properties. The fact that we observe such effects for the higher gamma range (> 70 Hz) extends the frequency range considered by Lewis and Bastiaansen (2015, i.e., the lower gamma), confirming that the research on the relation between gamma oscillations and reading requires further scrutiny. Important to note, we are not claiming that no gamma-related prediction error effects can be observed during language prediction. To test this hypothesis, a different experimental design should be employed in which different amounts of deviation from a predicted linguistic representation are employed. In the present design, we compare two conditions in which the prediction strength is different but no "violation" of a prediction is employed. This was strategically done to avoid task-related inhibition of ecologically natural predictive processes at work during language comprehension.

The location of the EEG gamma effects discussed by Herrmann et al. (2004) have a posterior, slightly right-lateralized distribution. Similarly, in the present MEG study the differential effect in the gamma range emerged in the right occipital cortex. This right-lateralization is somewhat unexpected for language tasks (interestingly, the evoked response around 130 ms is also right-lateralized), as effects mainly emerge in the left hemisphere. From a morphological perspective however, lexical information is unambiguously represented in the initial part of words in Spanish (word stem), while the word endings are morphologically more variable. Internal automatic expectations could thus prioritize the left visual hemifield during reading and this process could be stronger for the MWU condition. This hypothesis deserves further attention by possibly designing adequate cross-linguistic tests.

It can be noticed that we did not find differential effects (MWU vs Semantic) in primary visual regions in the time interval preceding the target word. In our opinion, the relation between pre-stimulus beta and post-stimulus gamma effects show that the generation of predictions is pursued at a different (more abstract) cognitive level and that it has an influence later on lower-level perception. It should be noticed, however, that in the MWU and Semantic conditions the participants were predicting exactly the same words but in different sentence contexts. Differential pre-stimulus visual activity can possibly be studied by evaluating the prediction of visual stimuli with different physical properties, that recruit different neural populations of the visual cortex. It is then still to be seen whether it is possible to decode the formal properties of a predicted representation from pre-stimulus activity in sensory cortices.

Language prediction processing

An important aspect of the present data concerns the type of representation that is pre-activated during linguistic prediction. The idea that low-level form-related representations can be accessed during language processing is under debate (for discussions see Huettig, 2015; Huettig & Mani, 2016; Luke & Christianson, 2016). The present data speak for an involvement of primary motor (articulatory) and sensory (visual) regions sensitive to contextual expectations. This is the case of the MWUs (which are very frequent in natural language[1]), in which the strong lexical association between words can concentrate the amount of pre-activation of the predicted item. On the other hand, when the prediction is not specific enough onto a specific lexical item (but disperses across the semantic neighbours) the prediction of more abstract semantic representations can lead

---

[1] Jackendoff (1997) estimates that the number of MWUs in a speaker's lexicon is of the same order of magnitude as individual words

to the recruitment of more classic language-related regions (prefrontal and middle temporal regions of the left hemisphere). The study by Wang et al. (2017), that focuses on the high vs. low semantically constraining contrast, provides evidence in this direction. The present study shows that differential pre-activation effects *can* be mediated via lower-level linguistic representations. It is worth mentioning, however, that Piai and colleagues (2015) suggested that the beta-band motor activity could be associated with some aspects of lexical retrieval. Beta band effects should thus be further investigated in the future to better understand their relations with lexical processing (Schoffelen et al., 2017; for a more general "beta oscillations proposal" on memory processes see Hanslmayer et al. (2012).

A final aspect of the present beta findings is related to the timing of the effects. Predictive processing is often described as a drifting process that slowly increases until the onset of the expected stimulus (Walter et al., 1964; for a review Chennu et al., 2013). In our opinion, this description better represents an attentional effect related to temporal expectation (in other words, active expectation that an "unspecified-event" is going to happen). Such a process would be under attentional control and associated to the task goals. In the present experiment, participants did not pay "special" attention to the target words (as compared to any other stimulus in the experiment)[2]. Prediction correlates (pre-activating a specific representation), on the other hand, would not necessarily follow such a ramp-like time course. During sentence reading, the system dynamically integrates information that constrains the properties of incoming stimuli. There is, thus, a continuous update of the internal priors: importantly, when such priors are retrieved or updated, there is no need for further processing costs until those priors are compared with the incoming stimulus. In the present study (and also Molinaro et al., 2017), we observed that the beta-effect emerges early and is not maintained until the target word. In our opinion, as soon as the word preceding the target has been integrated, the system adapts its internal priors - largely before (~300 ms) the appearance of the target. Such processing is related to the perception of the target, as evidenced by the inverse relation between beta and gamma effects.

Conclusions

In the present experiment, we evaluated the cortical networks supporting word pre-activation during sentence processing and show that the pre-articulatory motor regions support this process even during a reading task. This pre-stimulus beta modulation is related to the perceptual analysis of the pre-activated word in primary visual regions, as evidenced by the relationship between pre-stimulus beta and post-stimulus gamma power

---

[2] Thanks also to the use of natural heterogeneous material, post-experiment debriefings confirmed that participants were not aware of the experimental manipulation.

increase. The present findings thus align more with motor theories of “language prediction” as we highlight the important role of the language production system also during reading. We also show that predictions are updated as soon as contextual information is available, in a time frame that is relatively independent of the appearance of the expected stimulus.

References

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological review*, *105*(4), 678. doi: 10.1037/0033-295X.105.4.678-723

Arnal, L. H. (2012). Predicting “when” using the motor system’s beta-band oscillations. *Frontiers in Human Neuroscience, 6*, 225. doi: 10.3389/fnhum.2012.00225

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta–beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex, 25*(9), 3077–3085. doi:10.1093/cercor/bhu103

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences, 16*(7), 390– 398. doi: 10.1016/j.tics.2012.05.003

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., ... Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron, 85*(2), 390–401. doi: 10.1016/j.neuron.2014.12.018

Bourguignon, M., Molinaro, M., & Wens, V. (2018). Contrasting functional imaging parametric maps: the mislocation problem and alternative solutions. *Neuroimage, 169*, 200–211. Doi: 10.1016/j.neuroimage.2017.12.033

Canal., P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2017). Basic Composition and Enriched Integration in Idiom Processing: An EEG Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(6), 928-943. doi: 10.1037/xlm0000351

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., ... & Bekinschtein, T. A. (2013). Expectation and attention in hierarchical auditory prediction. *Journal of Neuroscience*, *33*(27), 11194-11205. doi: 10.1523/JNEUROSCI.0114-13.2013

Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience, 5*, 162–176. doi: 10.1162/jocn.1993.5.2.162.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634), 1–9. doi: 10.1098/rstb.2012.0394

Engel, A. K., & Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current opinion in neurobiology, 20*(2), 156-165. doi: 10.1016/j.conb.2010.02.015

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491-505. doi: 10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495. doi: 10.1006/jmla.1999.2660

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*, 195–207. doi: 10.1006/nimg.1998.0396

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in cognitive sciences, 16*(5), 262-268. doi: 10.1016/j.tics.2012.04.001

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836. doi: 10.1098/rstb.2005.1622

Friston, K. J., & Frith, C. D. (2015). Active inference, communication and hermeneutics. *Cortex, 68*, 129-143. doi: 10.1016/j.cortex.2015.03.025

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience and Biobehavioral Reviews, 81* Part B, 194-204. doi: 10.1016/j.neubiorev.2017.01.048

Hanslmayr, S., Staudigl, T., & Fellner, M. C. (2012). Oscillatory power decreases and long-term memory: The information via desynchronization hypothesis. *Frontiers in Human Neuroscience, 6*, 74. doi:10.3389/fnhum.2012.00074

Hawelka, S., Schuster, S., Gagl, B., & Hutzler, F. (2015). On forward inferences of fast and slow readers. An eye movement study. *Scientific reports, 5*. doi: 10.1038/srep08432

Herrmann, C. S., Munk, M. H. J., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences, 8*(8), 347e355. doi: 10.1016/j.tics.2004.06.006

Hoogenboom, N., Schoffelen, J. M., Oostenveld, R., & Fries, P. (2010). Visually induced gamma-band activity predicts speed of change detection in humans. *Neuroimage, 51*(3), 1162-1167. doi: 10.1016/j.neuroimage.2010.03.041

Hoogenboom, N., Schoffelen, J. M., Oostenveld, R., Parkes, L. M., & Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *Neuroimage, 29*(3), 764-773. doi: 10.1016/j.neuroimage.2005.08.043

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 1626*, 118-35. doi: 10.1016/j.brainres.2015.02.014

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language,* *Cognition and Neuroscience, 31*(1), 19-31. doi: 10.1080/23273798.2015.1072223

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*(4), 785-813. doi: 10.3758/BF03196544

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*, 163–178. doi: 10.1111/1469-8986.3720163

Kell, C. A., Darquea, M., Behrens, M., Cordani, L., Keller, C., & Fuchs, S. (2017). Phonetic detail and lateralization of reading-related inner speech and of auditory and somatosensory feedback processing during overt reading. *Human brain mapping, 38*(1), 493-508. doi: 10.1002/hbm.23398

Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., & Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Experimental neurology, 245*, 15-26. doi: 10.1016/j.expneurol.2012.09.014

Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de)constructing the N400. *Nature Reviews: Neuroscience, 9*, 920–933. doi: 10.1038/nrn2532

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177. doi: 10.1016/j.cognition.2007.05.006

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex, 68*, 155-168. doi: 10.1016/j.cortex.2015.02.014

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology, 88*, 22-60. doi: 10.1016/j.cogpsych.2016.06.002

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. doi: 10.1016/j.jneumeth.2007.03.024

Martí, M.A., & Taulé, M. (2007). CESS-ECE: Corpus anotados del español y catalán. Arena Romanística. *Corpus and text linguistics Romance studies*, 1. Department of Foreign Languages, Bergen, Norway.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*(5), 375. doi: 10.1037/0033-295X.88.5.375

Molinaro, N., Barraza, P., & Carreiras, M. (2013). Long-range neural synchronization supports fast and efficient reading: EEG correlates of processing expected words in sentences. *Neuroimage, 72*, 120–132. doi:

10.1016/j.neuroimage.2013. 01.031

Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition, 164*, 61-73. doi: 10.1016/j.cognition.2017.03.012

Molinaro, N., Monsalve, I. F., & Lizarazu, M. (2016). Is there a common oscillatory brain mechanism for producing and predicting language? *Language, Cognition and Neuroscience, 31*(1), 145–158. doi: 10.1080/23273798.2015.1077978

Monsalve, I. F., Pérez, A., & Molinaro, N. (2014). Item parameters dissociate between expectation formats: A regression analysis of time-frequency decomposed EEG data. *Frontiers in Psychology, 5*, 847. doi:10.3389/fpsyg.2014.00847

Murphy, K., Corfield, D. R., Guz, A., Fink, G. R., Wise, R. J. S., Harrison, J., & Adams, L. (1997). Cerebral areas associated with motor control of speech in humans. *Journal of Applied Physiology, 83*(5), 1438-1447.

Nichols, T. E., & Holmes, A. P. (2001). Non-parametric permutation tests for functional neuroimaging: A Primer with examples. *Human Brain Mapping, 15*, 1–25. 10.1016/j.jneumeth.2007.03.024

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 156869. doi: 10.1155/2011/156869.

Piai, V., Roelofs, A., Rommers, J., & Maris, E. (2015). Beta oscillations reflect memory and motor aspects of spoken word production. *Human Brain Mapping, 36,* 2767–2780. doi: 10.1002/hbm.22806

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105–110. doi:10.1016/j. tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329–347. doi: 10.1017/S0140525X12001495

Richter, C. G., Thompson, W. H., Bosman, C. A., & Fries, P. (2017). Top-down beta enhances bottom-up gamma. *Journal of Neuroscience, 7*, 3771-16. doi: 10.1523/JNEUROSCI.3771-16.2017

Schoffelen, J. M., Hultén, A., Lam, N., Marquand, A. F., Uddén, J., & Hagoort, P. (2017). Frequency-specific directed interactions in the human brain network for language. *Proceedings of the National Academy of Sciences USA, 114*(30), 8083-8088. doi: 10.1073/pnas.1703155114.

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9*(8), 311-327. doi: 10.1111/lnc3.12151

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in cognitive sciences, 13*(9), 403-409. doi: 10.1016/j.tics.2009.06.003

Summerfield, C., & Egner, T. (2016). Feature-based attention and feature-based expectation. *Trends in cognitive sciences, 20*(6), 401-404. doi: 10.1016/j.tics.2016.03.008

Taulu, S., Simola, J., & Kajola, M. (2005). Applications of the signal space separation method. *IEEE Transactions on Signal Processing, 53*, 3359–3372. doi: 10.1109/TSP.2005.853302

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience, 22*(8), 1682–1700. doi: 10.1162/jocn.2009.21293

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature*, *203*(4943), 380-384. doi: 10.1038/203380a0

Wang, L., Hagoort, P., & Jensen, O. (2017). Language Prediction Is Reflected by Coupling between Frontal Gamma and Posterior Alpha Oscillations. *Journal of Cognitive Neuroscience*, 1-16. doi: 10.1162/jocn_a_01190

Wang, L., Zhu, Z., & Bastiaansen, M. C. M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Frontiers in Psychology, 3*(187). doi: 10.3389/fpsyg.2012.00187

Weiss, S., & Mueller, H. M. (2012). "Too many betas do not spoil the broth": the role of beta brain oscillations in language processing. *Frontiers in Psychology, 3*. doi: 10.3389/fpsyg.2012.00201

Acknowledgements

This work was partially supported by the Spanish Ministry of Economy and Competitiveness (MINECO), the Agencia Estatal de Investigación (AEI) and the Fondo Europeo de Desarrollo Regional FEDER) (grant PSI2015-65694-P, “Severo Ochoa” programme SEV-2015-490 for Centres of Excellence in R&D), and by the Basque government (grant PI_2016_1_0014).

Further support derived from the AThEME project funded by the European Commission 7th Framework Programme, the ERC- 2011-ADG-295362 from the European Research Council.

We would like to thank Margaret Gillon-Dowens and the Proactive group for comments on previous versions of this manuscript. We also thank Mikel Lizarazu and Mathieu Bourguignon for their enlightening comments on our MEG data analyses.

Interest statement

The authors have no competeing interests to declare

Figure captions

Figure 1: A. Occipital evoked responses (sensor-level) for the two experimental conditions in the time interval of interest time-locked to the expected word. B. On the left, source reconstruction of the evoked response for the two experimental conditions pooled together. On the right, histograms showing the magnitude of the evoked response in the two conditions for the peaks of power emerged in each occipital lobe.

Figure 2: (A) Power estimates (expressed in relative change compared to the pre-sentence baseline) for each condition for the whole time interval of interest: the panels on top present the lower frequency range (1-35 Hz) in left-lateralized sensors, the middle panels the same frequency range in the posterior sensors and the panels at the bottom the higher frequency range (60-100 Hz) in the posterior sensors. (B-C) Sensor-level oscillatory effects (expressed in relative change compared to the pre-sentence baseline) in the beta (B) and gamma (C) range. For each case, we report at the left-top, the oscillatory pattern for both experimental conditions in which we highlight (dashed-line white square) the time-frequency window showing the differential effect (MWU vs. Semantic) in a relevant sensor (white circle highlight in the plots below). Below, we report the topographical distribution of the effect for each of the two experimental conditions and, on the right, the topographical distribution of the significant cluster expressed in t-values. The beta effect emerged in the pre-target word time interval and the gamma effect in the post-target time interval.

Figure 3: Source-level beta-band (~ 20 Hz) effect emerged in the pre-stimulus time interval. (A) Brain maps representing the statistically significant relative power change compared to a pre-sentence baseline for both the experimental condition together. White numbers represent local minima of power. (B) Time course (with relative shaded error bars) of the of the beta power in each condition (MWU and Semantic) in the three local minima (see corresponding numbers) identified at the group level. The grey shaded sections represent the time interval in which the effect was statistically significant at the sensor level. The horizontal black bar highlights the time intervals in which the two time-courses differ. Below we report the timing of each word presentation.

Figure 4: Source-level gamma-band (> 70 Hz) effect emerged in the post-stimulus time interval. (A) Brain maps representing the statistically significant relative power change compared to a pre-sentence baseline for both the experimental condition together. White numbers represent local maxima of power. (B) Time course (with

relative shaded error bars) of the of the beta power in each condition (MWU and Semantic) in the two local maxima (see corresponding numbers) identified at the group level. The grey shaded sections represent the time interval in which the effect was statistically significant at the sensor level. The horizontal black bar highlights the time intervals in which the two time-courses differ. Below we report the timing of each word presentation.

Appendix A

Multi Word Expressions *constraint*

1	Mis amigos me dijeron que Susana salió muy guapa en televisión, a cara **descubierta**, ayer por la tarde.
2	El jueves por la noche, las enfermeras me dijeron que la operación se realizó a corazón **abierto** por el cirujano.
3	A pesar de que en el pasado los ignoraba, Juan supo sus objetivos a corto **plazo** cuando empezó el curso.
4	Luego de reunir todos los ingredientes necesarios para su marmitako, Marina cocina el pescado a fuego **lento** en su cocina.
5	Mostrando una habilidad extraordinaria para enfrentar retos difíciles ella terminó su tarea a pesar de **todo** lo que pasó.
6	En el convenio que las personas interesadas tendrán que firmar, se establece claramente que la hipoteca se pagará a plazo **fijo** durante diez años.
7	El pequeño pájaro que todas las mañanas trina junto a mi ventana voló a ras del **suelo** hasta que aterrizó.
8	Cumpliendo las expectativas de todos, él imaginó al personaje a su imagen y **semejanza** cuando escribía la historia.
9	Aunque siempre se ha caracterizado por ser muy responsable, Pablo no irá a la escuela bajo ningún **concepto** teniendo tanta fiebre.
10	Todos queremos ver terminada la biblioteca, pero entendemos que los albañiles no pueden acabar de hoy para **mañana** la construcción del edificio.
11	Como era de esperar, Julián terminó el trabajo final del semestre de prisa y **corriendo** una hora antes de la entrega.
12	No entiendo por qué Mauro gasta dinero inútilmente si el consejo del médico es de sentido **común** para todos.
13	Diez días después del crimen, el comisario confirmó la detención del principal sospechoso en pocas **palabras** ante la prensa.
14	En el zoológico principal de la ciudad, el mono araña saltaba alegremente de rama en **rama** cuando le grababan en video.
15	Aquel día que fuimos de camping al parque más bonito de la ciudad, parecía que empezaría a llover de un momento a **otro**, pero no fue así.
16	Luis, el hijo de Patricia, se ha esforzado mucho y quiere aprobar de una manera o de **otra** el dichoso examen.
17	Con la persistencia que lo caracteriza cuando está muy entusiasmado con algo, Pedro leyó el libro de Borges del principio al **fin** en menos de dos días.
18	Cuando Martín le preguntó qué había pasado, la niña confesó en honor a la **verdad** haber roto el plato.
19	La hija menor de Jorge, Luisa, casi enfermó de pulmonía por salir en mangas de **camisa** al colegio en enero.
20	Después de la forma como se comportó, decidimos que Ana no vendrá a casa en modo **alguno** mientras no pida perdón.
21	En la fiesta que le organizamos a nuestra hija pequeña, el pastel se repartió en partes **iguales** para los niños.
22	El admirado héroe de la película de acción se encontraba entre la vida y la **muerte**, pero consiguió reanimarse.
23	Ya te conté que Carmen estaba tan emocionada con su boda que anunció su enlace a voz en **grito** en su trabajo.

24	Mi prima Laura viene de vacaciones a nuestra ciudad este verano y ni que **decir** que se quedará aquí.
25	Como mi esposo y yo hicimos las compras por separado, tenemos galletas en casa para dar y **tomar** durante más de un mes.
26	A partir de su divorcio, Susana sale de noche a tomar algo y a bailar con sus amigas por norma **general** los viernes.
27	La playa está abarrotada de adolescentes surfistas porque hoy no hay clase y por si fuera **poco** hace buen tiempo.
28	Estoy convencida de que mi mejor amigo, Felipe, aprobará el examen de matemáticas sin duda **alguna** por lo mucho que estudio.
29	Tras mucho pensarlo durante toda la semana, Clara se compró finalmente el coche de segunda **mano** porque es más barato.
30	Todos sus enemigos la acusan duramente, pero quienes la conocemos bien sabemos que ella no lo hizo adrede sin ninguna **duda**, sino que fue un accidente.
31	Piensa lo que quieras, pero en verdad te digo que Alicia no robó tu cartera, todo lo **contrario**, la guardó para ponerla a salvo.
32	Sus bromistas amigos la tenían retorciéndose de la risa con sus cosquillas, pero Cecilia les chillaba a más no **poder** hasta que la soltaron.
33	Aunque son gemelas idénticas, Susana y Raquel no tienen absolutamente nada en común, pelean frecuentemente y se llevan como el día a la **noche** desde pequeñas.
34	Había un gran desorden en todas las habitaciones y mis sobrinos debieron recoger la casa a contra **reloj** cuando vino una visita.
35	El paciente que llegó anoche al hospital lloraba de dolor, pero ya será **menos** porque no tenía nada aparentemente.
36	David, luego de tanto estudiar en las clases de física, no consiguió aprobar ni a la de **tres** el examen de conducir.
37	No cabe duda que Alejandro ha entregado su trabajo a tiempo, pero decir que es listo ya es mucho **decir** porque no es un genio.
38	Tras recibir algunas monedas de los transeúntes, los chicos tuvieron que irse con la música a otra **parte** porque molestaban.
39	Cuando fui a su casa de verano en la costa, el matrimonio González me trató a cuerpo de **rey** mientras estuve allí.
40	Ante los regaños de la directora por lo que había sucedido, los niños llegaron a clase en fila **india** sin hacer ruido.
41	A pesar de lo que decidió el magistrado González, el asesino lo hizo con conocimiento de **causa** porque la ley es clara.
42	Tras retirar el sobre de la urna y abrirlo, él pronunció el nombre del ganador a viva **voz** en la gala de premios.
43	Al abrir el testamento de la abuela, Sara no recibió una herencia, pero ella tiene mucha suerte y cada dos por **tres** gana la lotería.
44	El día de la presentación, tras haber pasado frío el fin de semana, estaba tan acatarrado que a duras **penas** podía hablar.
45	Como me decía siempre mi abuelo Carlos cuando yo era muy pequeño, el ser humano es curioso desde que el mundo es **mundo** y así seguirá.
46	Ya se lo dije a la tripulación y no saltaré de un avión por nada del **mundo**, ni que se esté quemando.

47	Estefanía es una verdadera atleta profesional que entrena hace muchos años y pudo levantar 100 kilos como si tal **cosa** varias veces.
48	En el pueblo, esa era una fiesta muy esperada y se celebró el cumpleaños por todo lo **alto** el viernes por la noche.
49	Después de un sueño horroroso Lisa se despertó en la sala, y el homicida serial le pasó el cuchillo a flor de **piel** para asustarla.
50	De la misma manera que hace siempre con todas las personas, ella abusó de su confianza más de la **cuenta** y ahora están enfadados.
51	Por la tarde llegó una fuerte galerna en la costa y los barcos tuvieron que luchar contra viento y **marea** para volver al puerto.
52	Como él pensó que llovería mucho esa noche, antes de irse cerró todas las ventanas a cal y **canto** por lo que hacía calor.
53	A diferencia de cuando vivía en casa de mis padres, en mi pequeña habitación tengo todo al alcance de la **mano** y me gusta.
54	Me han dicho los del barrio que los dos hombres se pelearon ayer con uñas y **dientes** fuera del bar.
55	Sandra no era buena cocinera, pero en esa ocasión siguió la receta al pie de la **letra** para hacer el pavo.
56	El carpintero que vive arriba tiene mucha experiencia en su trabajo y consiguió hacer la tarea con los ojos **cerrados** dejándonos sin palabras.
57	Después del accidente el seguro ya pagó el arreglo total del coche porque a fin de **cuentas** la culpa no fue suya.
58	A pesar de que es el examen más importante de la carrera, no voy a dejar de dormir ni mucho **menos** por tener que estudiar.
59	Como en otros aspectos de su vida se sentía muy descontenta y frustrada, se entregó a su trabajo en cuerpo y **alma** durante mucho tiempo.
60	Muchas veces hemos leído en los cuentos de los niños que dos caballeros se enfrentaron sin miedo a vida o **muerte** por el amor de la princesa.

Semantic *constraint*

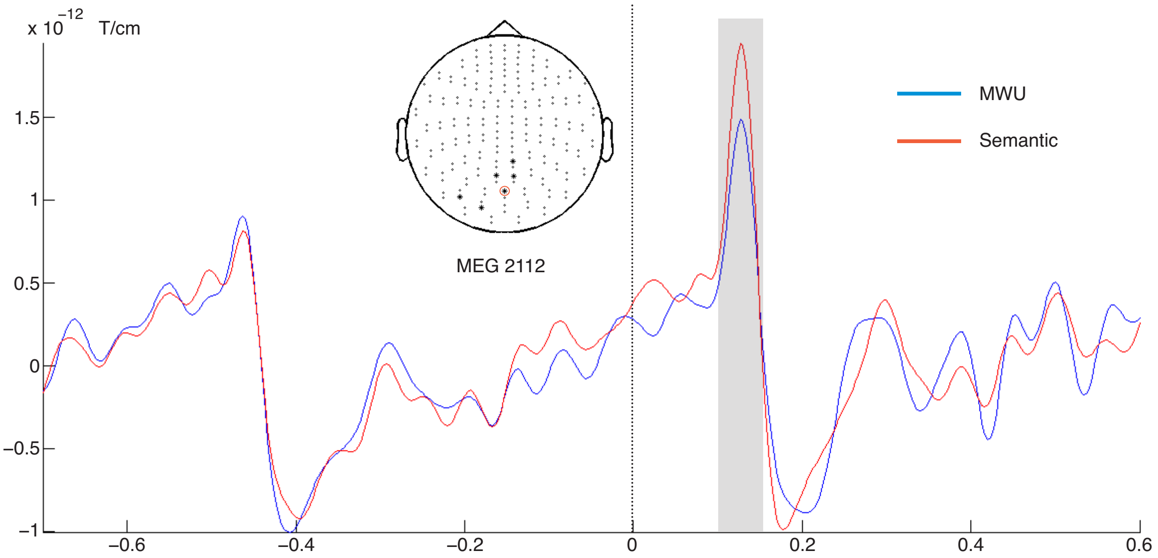
1	Los famosos investigadores han comunicado que una nueva vacuna contra el cáncer ha sido **descubierta** recientemente.
2	Los chicos de la cuadrilla de José no piensan retirarse a casa mientras el pub siga **abierto** hasta la mañana.
3	Laura no pudo realizar la matrícula en el curso de natación porque había finalizado el **plazo** de inscripción.
4	El entrenador de atletismo dejó fuera a Daniel para la competición porque era el corredor más **lento** con diferencia.
5	Por perseguir su sueño ha dejado su casa, su trabajo, vamos que lo ha dejado **todo** en un momento.
6	Después de cinco años trabajando en la empresa como eventual le han propuesto quedarse y le han hecho un contrato **fijo** por fin.
7	Andrés tuvo que recoger con la escoba todos los trozos de cristal que había en el **suelo** sin rechistar.
8	No son gemelos, pero siempre confunden a los dos hermanos por la gran **semejanza** que comparten.
9	Como parecía que algunos alumnos no entendían el significado de "democracia" el profesor tuvo que explicar ese **concepto** otra vez.

10	Ya es muy tarde para llamarle y además ya estará dormido así que le llamaré por la **mañana** sin falta.
11	Ana hace footing cerca de mi casa y desde mi ventana veo cómo va **corriendo** todos los días.
12	La pareja ha tenido que arreglar sus diferencias por el bien del hijo que tienen en **común** desde hace un año.
13	Cuando tiene que hablar en público, se pone muy nervioso y no le salen las **palabras** que quiere decir.
14	Lorea subió confiada al árbol, pero se cayó al poner el pie en una seca **rama** sin saberlo.
15	El niño quería que le regalaran dos juguetes y cuando el padre le dio solo uno, no calló hasta que le compró **otro** más.
16	El cree que en aquella ocasión no supo demostrar todo lo que sabía, por eso ha pedido **otra** oportunidad.
17	Hoy he visto una película preciosa que ha conseguido atrapar mi interés hasta que ha acabado y ha aparecido la palabra **fin** en la pantalla.
18	Aunque la mentira no es ética, a veces podemos hacer más daño con la **verdad** si no tenemos cuidado.
19	No podía atarse bien la corbata porque olvidó abrocharse el botón del cuello de la **camisa** con las prisas.
20	No hay manera de que la gente se vaya a casa a su hora, siempre queda **alguno** dando vueltas.
21	Los dos se parecen muchísimo y por mucho que intenten confundirnos, a mí nunca me parecerán **iguales** del todo.
22	La vida es maravillosa, pero yo pienso que no hay que tenerle miedo a la **muerte** para disfrutar de verdad.
23	Mi amiga Amaia es aracnofoba así que si ve una araña es capaz de dejarte sordo con un **grito** de los suyos.
24	Con lo habladora que es Edurne, es raro que se quede sin saber qué **decir** en algún momento.
25	A estas alturas de la película y con las dos opciones delante, todavía no sabe qué decisión **tomar** para seguir adelante.
26	El mando superior de la armada, el que se encuentra por encima del comandante, teniente, coronel, capitán etc. es el **general** siempre.
27	Digan lo que digan sigo pensando que para hacer esa tarea tan larga un minuto es **poco** tiempo.
28	La búsqueda de la niña desaparecida se prolongó durante meses, aunque no contaban con pista **alguna** para seguir.
29	Me he hecho una herida en un dedo, pero no del pie, sino de la **mano** derecha.
30	El profesor antes de empezar el examen explicó a los alumnos que debían levantar la mano si querían preguntar alguna **duda** que les hubiera surgido.
31	Mikel es un ecologista empedernido y sobre la cuestión de edificar en zonas protegidas se muestra muy **contrario** sin ninguna duda.
32	Hasta hace pocas décadas, la iglesia influía muchísimo en las decisiones del gobierno, se puede decir que tenía mucho **poder** sobre los políticos.
33	Cuando llega el fin de semana, mi hermano se dedica a dormir todo el día y a salir de juerga durante toda la **noche** con sus amigos.
34	Dime por favor qué hora es porque no tengo donde mirar y se me ha parado el **reloj** nuevo.

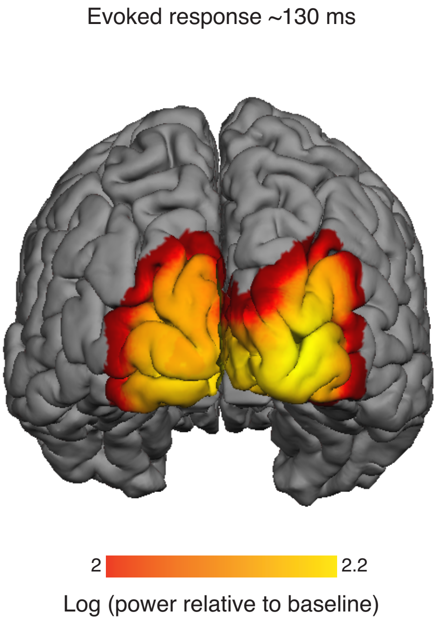
35	Aunque nunca lo vaya a dejar, el médico le ha aconsejado que fume **menos** por su bien.
36	Los valientes mosqueteros que acompañaban a Dartañan en sus aventuras no eran dos personas, sino que **tres** en total.
37	Cuando le han dado la noticia se ha quedado sin palabras, esto es, se ha quedado sin saber qué **decir** del susto.
38	No sabes ni la mitad de lo ocurrido, lo que te han contado es sólo una pequeña **parte** de la verdad.
39	En la ceremonia de entrega de premios presidían la mesa la reina y su marido el **rey** de España.
40	En carnavales mi amiga se pintó la cara y llevaba plumas, un arco y flechas, parecía una autentica **india** de película.
41	Pedro siempre lucha por lo que cree y más aún si piensa que es una justa **causa** que defender.
42	Cuando hablamos con ancianos nos debe de parecer que están todos medio sordos porque siempre alzamos la **voz** un poco.
43	Los reyes magos de oriente que llegan a visitar a Jesús en Belén son Melchor, Gaspar y Baltasar en total son **tres** ni más ni menos.
44	Mi amiga Edurne sólo me llama cuando le pasa algo malo, parece que solo me quiere para contarme sus **penas** de vez en cuando.
45	Cuando éramos niños veíamos en la tele las aventuras de Willi Fog, que en 80 días tenía que dar la vuelta al **mundo** por una apuesta.
46	Es un bicho muy extraño que, aunque es de este planeta, parece que viene de otro **mundo** diferente.
47	Aunque he intentado razonar y dar mil explicaciones está convencida de lo que vio y no quiere escuchar otra **cosa** diferente.
48	Mi hermano no es el más bajito como se estaba diciendo, todo lo contrario, es el más **alto** de la clase.
49	Antes de tomar el sol es muy importante darse crema solar con protección más alta de 50 para proteger de quemaduras la **piel** del cuerpo.
50	El camarero ya nos ha traído el postre y el café, así que si quieres marcharte puedes pedirle la **cuenta** ya.
51	Hoy por la mañana han encontrado el cuerpo de un pescador en la playa, ha debido de llegar arrastrado por la **marea**, que desgracia.
52	Su ilusión es ser cantante, por lo que para trabajar su voz a decidido apuntarse a clases de **canto** en la academia.
53	El cocinero cuando va a sacar la bandeja del horno se pone el guante de cocina para no quemarse la **mano** porque suele estar muy caliente.
54	El dentista le ha aconsejado que deje de fumar si no quiere estropear más sus blancos **dientes** que tanto cuida.
55	En la ikastola de María Reina el niño practica caligrafía fuera de horas para mejorar su **letra** porque nadie la entiende.
56	Lo que te voy a enseñar es una sorpresa, así que hasta que yo te diga mantén los ojos **cerrados** por favor.
57	Ana y Alberto se han casado hace varios meses, pero en el banco mantienen separadas sus **cuentas** de ahorro.
58	Las madres siempre llenan el plato hasta arriba y al contario, los hijos siempre piden que les echen un poco **menos**, aunque sin resultado.

59	El poeta, aunque no era muy espiritual, decía que al igual que las personas todos los seres de este mundo tienen **alma** y viven.
60	El soldado luchó hasta el final y consiguió sobrevivir a la batalla, pero perdió a su mejor amigo, presenciando su **muerte** en sus brazos.

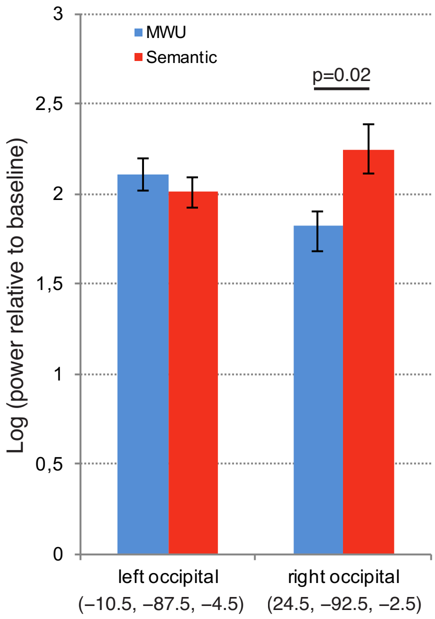
A.

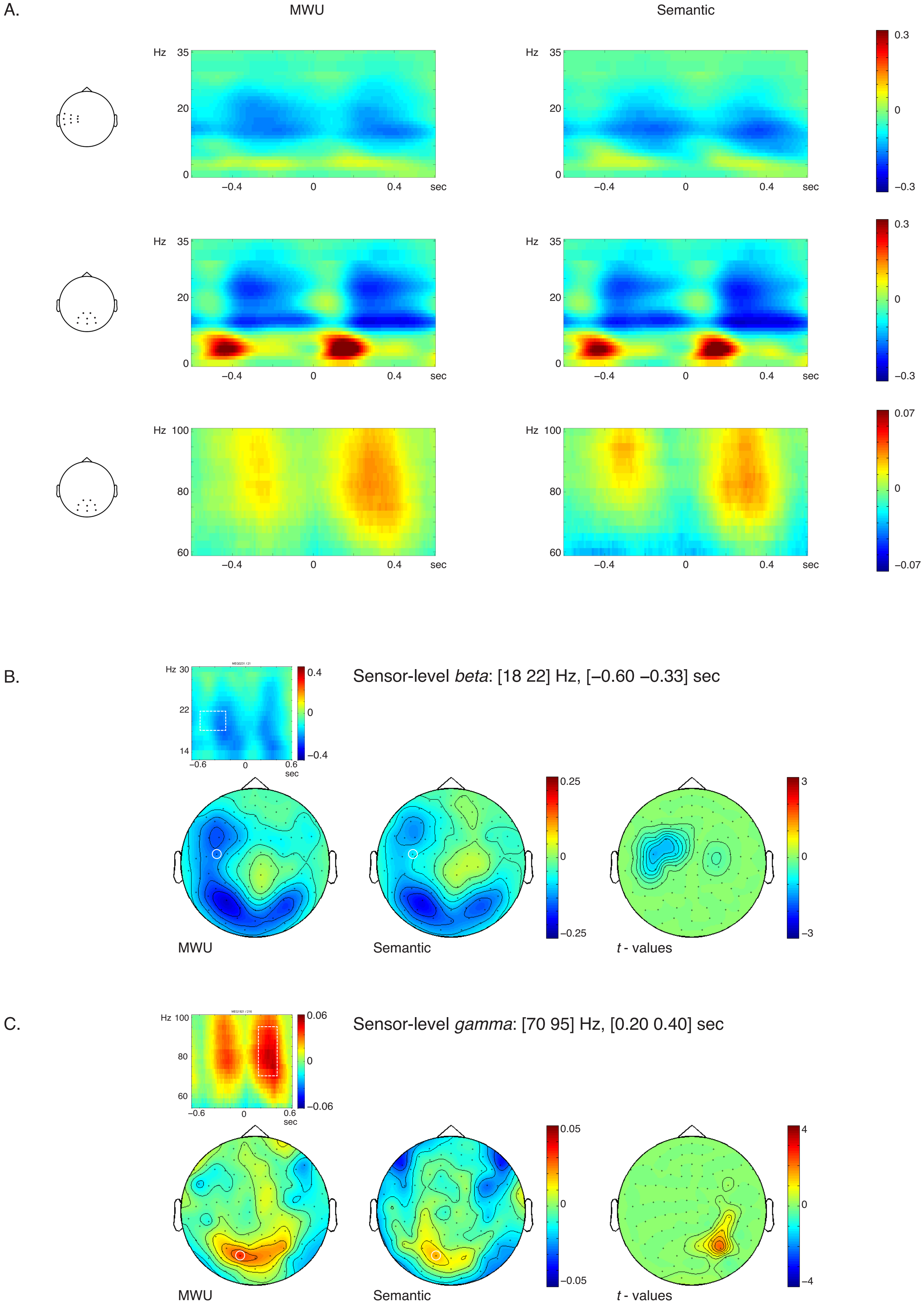

x 10^{-12} T/cm
1.5
1
0.5
0
−0.5
−1
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
MEG 2112
MWU
Semantic

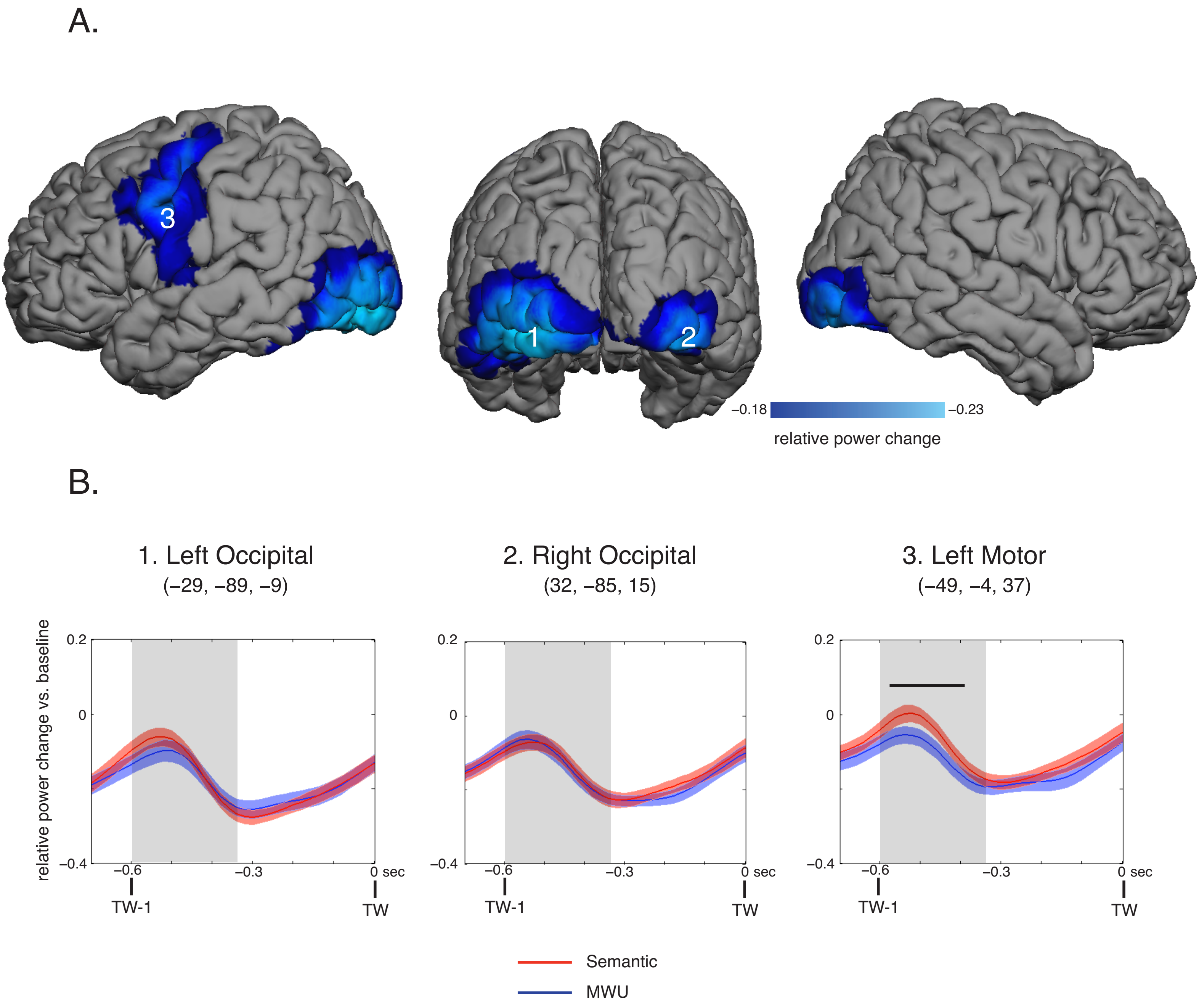
B.

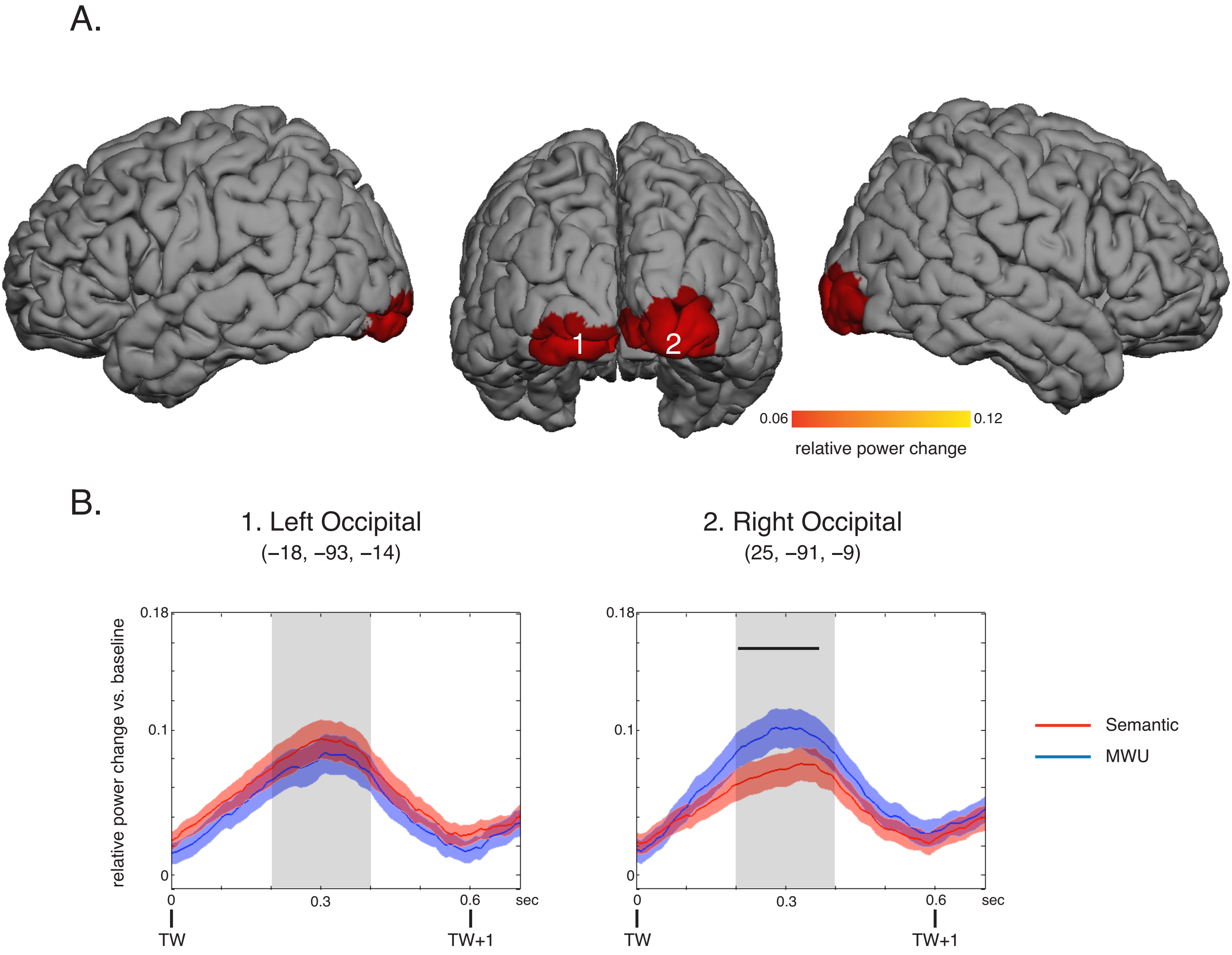

Evoked response ~130 ms
2
2.2
Log (power relative to baseline)

C.


3
2,5
2
1,5
1
0,5
0
■ MWU
■ Semantic
p=0.02
Log (power relative to baseline)
left occipital
(−10.5, −87.5, −4.5)
right occipital
(24.5, −92.5, −2.5)


A.
MWU
Semantic
Hz 35
20
0
–0.4
0
0.4
sec
0.3
0
–0.3
Hz 100
80
60
0.07
–0.07
B.
Sensor-level *beta*: [18 22] Hz, [–0.60 –0.33] sec
Hz 30
22
14
–0.6
0.6
sec
0.4
–0.4
0.25
–0.25
3
–3
MWU
Semantic
t - values
C.
Sensor-level *gamma*: [70 95] Hz, [0.20 0.40] sec
Hz 100
80
60
0.06
–0.06
0.05
–0.05
4
–4
MWU
Semantic
t - values


A.
3
1
2
−0.18
−0.23
relative power change
B.
1. Left Occipital
(−29, −89, −9)
2. Right Occipital
(32, −85, 15)
3. Left Motor
(−49, −4, 37)
relative power change vs. baseline
0.2
0
−0.4
−0.6
−0.3
0 sec
TW-1
TW
Semantic
MWU


A.
1
2
0.06
0.12
relative power change
B.
1. Left Occipital
(−18, −93, −14)
2. Right Occipital
(25, −91, −9)
relative power change vs. baseline
0.18
0.1
0
0
0.3
0.6
sec
TW
TW+1
0.18
0.1
0
0
0.3
0.6
sec
TW
TW+1
Semantic
MWU