NEUROSCIENCE
RESEARCH ARTICLE

B. Bianchi et al. / Neuroscience 519 (2023) 131–146

Neural Bases of Predictions During Natural Reading of Known Statements: An Electroencephalography and Eye Movements Co-registration Study

Bruno Bianchi,[a*] Rodrigo Loredo,[b] María da Fonseca,[a,c] Julia Carden,[b] Virginia Jaichenco,[b] Titus von der Malsburg,[d] Diego E. Shalom[e,f] and Juan Kamienkowski[a,g]

[a] *Laboratorio de Inteligencia Artificial Aplicada, Departamento de Computación/Instituto de Ciencias de la Computación, Universidad de Buenos Aires/CONICET, Argentina*

[b] *Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Argentina*

[c] *Universidad Nacional de Río Negro, Argentina*

[d] *Institute of Linguistics at the University of Stuttgart, Germany*

[e] *Instituto de Física de Buenos Aires (CONICET), Universidad de Buenos Aires, Buenos Aires, Argentina*

[f] *Laboratorio de Neurociencia, Universidad Torcuato Di Tella, Buenos Aires, Argentina*

[g] *Maestría en Exploraióón de Datos y Descubrimiento del Conocimiento, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina*

Abstract—Predictions of incoming words performed during reading have an impact on how the reader moves their eyes and on the electrical brain potentials. Eye tracking (ET) experiments show that less predictable words are fixated for longer periods of times. Electroencephalography (EEG) experiments show that these words elicit a more negative potential around 400 ms (N400) after the word onset when reading one word at a time (foveated reading). Nevertheless, there was no N400 potential during the foveated reading of previously known sentences (memory-encoded), which suggests that the prediction of words from memory-encoded sentences is based on different mechanisms than predictions performed on common sentences. Here, we performed an ET-EEG co-registration experiment where participants read common and memory-encoded sentences. Our results show that the N400 potential disappear when the reader recognises the sentence. Furthermore, time–frequency analyses show a larger alpha lateralisation and a beta power increase for memory-encoded sentences. This suggests a more distributed attention and an active maintenance of the cognitive set, in concordance to the predictive coding framework.

Key words: co-registration, EEG, eye movements, reading, cloze-Predictability, N400.

INTRODUCTION

When reading, our eyes move along the text, fixating over the words during short periods of time of around 250 ms (Rayner, 1998). Eye movement experiments show that the duration of each fixation depends on several properties of the fixated word and its context. For example, more frequent or shorter words are fixated for shorter periods of time (Rayner, 1998; Kliegl et al., 2006). These results lead to the hypothesis that the time spent by our eyes on each word is related to the cost of processing the information being read (Just and Carpenter, 1980). Neuroimaging experiments are performed to better understand the cognitive processes that underlie word processing. Electroencephalography (EEG) allows to analyse neural activity by directly recording the scalp electrical potential, that reflects the activation of big neuronal populations.

Studies have found that prediction of upcoming words take place during reading (DeLong et al., 2005; Pickering and Garrod, 2007; Kutas and Federmeier, 2011). Using an estimation of how predictable a word is before (i.e. cloze-Predictability) as a regressor of Fixation and Gaze Duration shows that more predictable words are read for shorter periods of time (Kliegl et al., 2006). At the same time, analyses on the scalp electrical potential shows that cloze-Predictability modulates the amplitude of a negative deflection that appears around 400 ms after the presentation of the word (N400) (Kutas and Hillyard, 1980; Kutas

*Corresponding author.
E-mail address: bbianchi@dc.uba.ar (B. Bianchi).
Abbreviations: EEG, Electroencephalography; ET, Eye Tracking; SVP, Serial Visual Presentation; RP, Recognition Point; ERP, Event Related Potentials; FRP, Fixation Related Potentials; LMM, Linear Mixed Models; SOA, Stimulus onset asynchrony; SD, Standard Deviation; EOG, Electrooculography; CBP, Cluster Based Permutation.

https://doi.org/10.1016/j.neuroscience.2023.03.024
0306-4522/

and Federmeier, 2011). The amplitude of this component is sensible to linguistic and non-linguistic manipulation, and has become a key tool to probe semantic memory, and to study bottom-up and top-down information pathways in the brain. All together, these results suggest that the processing of a word is influenced by how predictable it is. Thus, it is of great interest to better understand how predictions are made, what brain mechanisms are involved on this process, and how EEG and eye tracking results can be linked together. Classical results either from EEG or eye movement studies reflect different processes. In the former ones, readers passively receive one word at a time with an interval between 500 and 1000 ms, centred in the screen, and without further information about following words (Serial Visual Presentation paradigm –SVP–). In the latter ones, readers actively move their eyes across words at their own pace, fixating over each word for around 200 ms (natural reading). Thus, one of the major challenges of the field is to put together the results obtained with these two methodologies/paradigms, and developing new theories that unifies previous and new results (Sereno and Rayner, 2003; Kliegl et al., 2012).

In a previous study, we analysed the cognitive basis of predictions performed while reading *common* and *memory-encoded sentences* (Bianchi et al., 2019). The aim of that study was to analyse the brain electrical potentials evoked by different sources of predictions. Common sentences were sentences made-up for the experiment, that participants could only predict from semantic or syntactical relations within the sentence. Memory-encoded sentences were proverbs, song lyrics, famous phrases, etc. Participants know these sentences from the daily use. Thus, it is possible to find a certain word in the sentence from which participants recognise the whole remaining sentence, and predict all the following words with high probability. We defined this word as the Recognition Point (RP). Using an SVP paradigm, we showed that the classical N400 potential is only present when reading common sentences, and no effect of cloze-Predictability was found on the EEG amplitude for memory-encoded sentences.

Fernández et al. (2014) analysed eye movements during the reading of proverbs, high-predictable, and low-predictable sentences. On this study, they focused on the parafoveal-on-foveal word effect. That is, how the cloze-Predictability of the word after the current fixation (N + 1) impacts on the current gaze duration (N). They found that, on high-predictable sentences the cloze-Predictability of the N + 1 has an impact, the fixation duration on the current word, only before the RP. This effect was not present on these sentences after the RP, on proverbs, nor on low-predictable sentences. Based on their results, they proposed that there is a distributed processing of words, both with respect to visual processing in the perceptual span, and by the generation of accurate representation of the sentence, that leads to predictions performed by retrieving information from memory.

As far as we know, there are few studies analysing the reading of statements stored in memory (Vespignani et al., 2010; Fernández et al., 2014; Canal et al., 2017; Bianchi et al., 2019). Both Vespignani and collaborators (Vespignani et al., 2010) and Canal and collaborators (Canal et al., 2017) analysed idiom reading, and firstly defined the RP. Previously introduced studies from Bianchi and collaborators (Bianchi et al., 2019) and from Fernández and collaborators (Fernández et al., 2014) are the only ones analysing reading of long statements stored in memory. The corpus used by Bianchi et al. (2019) was comprised by the same sentences used by Fernández et al. (2014). Both of them found a limitation when analysing the reading of memory-encoded sentences. These sentences were recognised too quickly. That is, there were no enough words before the RP to analyse eye movements or scalp potentials properly. This hindered the comparison between words before and after the memory recall. In the present study, we aim to analyse the reading of these sentences splitting them intro pre- and post-Recognition Point. Thus, to ensure a sufficient amount of data on preRP split, we modified memory-encoded sentences by adding neutral previous contexts to them.

Predictive coding framework

The predictive coding theoretical framework (Lewis and Bastiaansen, 2015) proposes that during reading there is an ongoing predictive processing whenever possible. Predictions from these processes percolate down from higher-level cognitive areas to the lower-level perceptual areas (top-down feedback transmission). In its way down, these predictions are compared with sensory inputs that come from lower levels (bottom-up or feedforward transmission). This results in a prediction error that is used to update the current and future predictions. Thus, under this framework, predictions are based on information transmission between brain regions and could be reflected on power and synchronisation of oscillations that can be observed through time–frequency analysis (Lewis and Bastiaansen, 2015; Himmelstoss et al., 2020).

This framework proposes several hypotheses on how predictions performed during reading impact on different frequency bands of the EEG. Additionally, each frequency band counts with its own literature about the role on the processing of reading. For these analyses the electrical potential is decomposed into different frequency bands and both their power and the phase synchronisation are analysed (Bastiaansen and Hagoort, 2006; Vignali et al., 2016; Lewis and Bastiaansen, 2015). Throughout the literature, it is possible to find a great variety of processes reflected in different frequency bands, within various scalp locations and latencies. The Theta band, traditionally defined between 4 and 8 Hz (Cohen, 2014), is associated with the synchronisation of large-scale functional networks (von Stein and Sarnthein, 2000; Cohen, 2014). In particular, it is associated with processes related to the encoding of information in memory (Axmacher et al., 2006), and with *top-down* processes (Sauseng et al., 2010; Cavanagh et al., 2012; van Gaal et al., 2012). Regarding reading, phase synchronisation studies show an increase in Theta band synchronisation in early latencies when reading idiomatic expressions (Molinaro et al., 2013). Nevertheless, the predictive coding framework does not have specific hypothesis on this frequency band.

The Alpha band, traditionally defined between 8 and 16 Hz (Lower Alpha 8–12 Hz; Upper Alpha 12–16 Hz), is commonly associated with attention and (semantic) long-term memory performance (Klimesch, 1999; Kornrumpf et al., 2017). On SVP reading, a Lower Alpha power decrease prior to the critical word was found when reading highly constrained sentences (Rommers et al., 2017). Regarding natural reading, Kornrumpf and collaborators (Kornrumpf et al. (2017)) showed a lateralization of posterior Alpha power related to the left to right reading, with a contralateral power decrease and an ipsilateral power increase to the attended space. For the Predictive coding framework, the Alpha band found to be involved in the feedback (top-down) signalling (Lewis and Bastiaansen, 2015).

The Beta band, traditionally defined between 16 and 30 Hz (Lower Beta 16–20 Hz; Middle Beta: 20–24 Hz; Upper Beta 24–30 Hz), can be linked to motor cortex activation (Parkes et al., 2006). The literature of this frequency band on natural reading tasks is scarce (Vignali et al., 2016). Nevertheless, the predictive coding framework also relates this frequency band with feedback (top-down) signalling (Lewis and Bastiaansen, 2015). They proposed that an increase in the beta power is a reflection of the maintenance of the current cognitive set when incoming stimuli agree with ongoing predictions. By contrast, when predictions are found erroneous, a decrease of the Beta power is found.

The Gamma band, traditionally defined between 60 and 90 Hz, is also hypothesised to play an important role on prediction errors transmission. Nevertheless, on free viewing tasks, results in this frequency band are obscured with ocular movements artefacts (Yuval-Greenberg et al., 2008; Dimigen et al., 2009).

To the best of our knowledge, there are few studies in the time–frequency in natural reading. For example, Metzner et al. (2015) compared the time–frequency representations of foveated and natural reading. They concluded that oscillatory brain dynamics differ between these two experimental paradigms, in delta and alpha bands.

Why co-registration and natural stimuli

The debate between the experimental and the observational/correlation approach in psychology is not new. Already in 1957, Cronbach published an article in “American psychologist” about the need to join these two approaches, posting the potential benefits of this convergence (Cronbach, 1957). In 1975 he revisited his own article, after 18 years where he observed no advances on this merge (Cronbach, 1975). Nevertheless, during the last decades, the advances in statistical methodologies and the computational power have helped to better analyse data with many non-controlled variables (Kliegl et al., 2011; Grieve, 2021; Baayen, 2014).

Despite great advances achieved during the last decades on understanding how our brain processes reading, most of the results come from experiments with constrained behaviour and highly controlled stimuli (Hamilton and Huth, 2020). For instance, it is common to observe that eye-movements experiments are preformed on carefully designed sentences, with specific grammatical structures, or EEG experiments where sentences are presented at a slow pace, one word at a time in the centre of the screen to avoid ocular movements (foveated reading). Carefully designed sentences allows simpler analyses and result interpretations. And foveated reading avoids artefacts in EEG signal due to muscular contraction potentials and eyeball movements themselves (Dimigen, 2020). Nevertheless, during the last years experimental paradigms started to release those restrictions, for example designing neuroimaging experiments where eye movements are allowed (Dimigen et al., 2011; Metzner et al., 2015; Metzner et al., 2017). At the same time, it has become increasingly common to find experiments where stimuli are selected from everyday sources, such as literary texts (Bianchi et al., 2019; Bianchi et al., 2020), recordings of radio programs (Huth et al., 2016), film clips (Huth et al., 2012), images of actual crowds (Kaunitz et al., 2014; Kamienkowski et al., 2018), popular television content (Dmochowski et al., 2014), or in educational videos (Madsen et al., 2021).

Nevertheless, this path has its difficulties. Using a corpus made with natural stimuli implies to loose control on confounding variables. For instance, corpus from Bianchi et al. (2019) memory-encoded sentences were collected from the popular knowledge. There, the cloze-Predictability is variable of interest, but other variables like the word lexical frequency or its position in the sentences were not taken into account at the construction stage. This resulted in difficulties to analyse the cloze-Predictability using the classical ERP methodology where the variable of interest is analysed in isolation. To overcome these difficulties, data must be analysed using more powerful statistical techniques. For example, Bianchi (Bianchi et al., 2019) developed a statistical Linear Mixed Model-based non-parametrical technique to analyse several continuous variables in a unified test. It joins the knowledge from eye movements in reading and the electrophysiological fields. The former one has worked out how to deal with multiple independent variables using LMM (Kliegl et al., 2011). And the latter has elaborate on how to perform non-parametrical statistics to solve the multiple comparisons problem. Or, to overcome the difficulties to analyse the EEG data with artefacts from eye movements, Plöchl et al. (Plöchl et al., 2012) and Dimigen (Dimigen, 2020) have been working on pipelines and toolboxes for preanalysis, optimising artefact removal produced by eye movements.

One of the main goals of EEG and ET co-registration experiments is to better understand the timing of the cognitive processes associated with reading. The main issue arises when comparing fixation duration with evoked potentials. While a typical fixation lasts around 200 ms and its duration is influenced by cloze-Predictability, the evoked potentials show that it is not until 300 or 400 ms that this variable has an effect on its amplitude. Along these lines, Dambacher, Kliegl and collaborators analysed two independent sentence reading experiments. In one of them they measured the duration of fixations during natural reading, and in the other, the evoked potentials during foveated reading. They observed that both, the N400 potential and the duration of the fixations, are influenced by the same variables in a very similar way, hypothesising that there are common processes between them (Dambacher and Kliegl, 2007). Moreover, they showed that by modifying the presentation times of words (SOA) in foveated reading, the potentials associated with Lexical Frequency and Cloze Predictability are modified (Dambacher et al., 2012). For the Lexical Frequency they found a significant difference on the effect latency: a short SOA (280 ms) elicited a response 70 ms before than a long SOA (700 ms). For the N400 Predictability effect their analyses suggested a later start (onset latency) with an SOA of 280 than of 700 ms. Finally, they also found a smaller Predictability effect at the SOA of 280 ms compared to the SOA of 700 ms.

In a foundational study, Dimigen, Kliegl and collaborators introduced and proposed fixation-related potentials (FRP) analysis techniques. Here, they observed a significant N400 potential on the expected latency. Nevertheless, they also reported much weaker N400-like effect topographies that did not survive statistical corrections (Dimigen et al., 2011). Previously, Kretzschmar et al. (2009) analysed the parafoveal processing of upcoming words. They found that the N400 arose between 250 and 400 ms after the last pretarget fixation, which lasted about 186 ms. That is, the N400 effect arose earlier that expected by SVP experiments when measured from the first fixation on the target word. A similar result was found recently by Li et al. (2022) in natural reading of Chinese texts. Finally, Metzner et al. (2015) also found shorter N400 latencies when comparing the N400 from foveated and natural reading experiments.

In sum, the neuroscience field is starting to study how information is processed under more ecological conditions.

In the present study, we aim to deepen the understanding of how predictions are performed during reading. First, we plan to extend the results from our previous study to a natural reading paradigm. This implies to analyse if the N400 potential is present when processing words from previously known sentences. To overcome the limitations observed by Bianchi et al. (2019) and Fernández et al. (2014) on their sentences corpus, we add previous neutral contexts to the memory-encoded sentences, enlarging them to balance the position of the Recognition Point. Secondly, we analyse whether eye movements show a differential behaviour between common and memory-encoded sentences. Finally, we explore the electrophysiological bases of the scalp potential recorded during the natural reading of these types of sentences.

EXPERIMENTAL PROCEDURES

Participants

Twenty-eight subjects participated in the experiment (mean 27.1years; range [18–38] years; 11 women), receiving monetary compensation. All the subsejcts had normal or correct to normal vision. Each session took between 1.5 and 2.5 h, including preparation. Eight participants were excluded from final analyses due to noisy EEG data or poorly calibrated eye tracking (see Data pre-processing section for more details). All participants provided written informed consent in agreement with the Helsinki declaration. All the experiments described in this paper were reviewed and approved by the ethics comittee: ''Comité de Ética del Centro de Educación Médica e Investigaciones Clínicas ''Norberto Quirno'' ''(CEMIC)'' and qualified by the Department of Health and Human Services (HHS, USA): IRb00001745 - IORG 0001315 (Protocol 435).

Materials

Sentence corpus was prepared taking into account the limitations observed in the study previously carried out with this type of stimulus (Bianchi et al., 2019). Bianchi et al. observed that a Recognition Point (RP), i.e. a word after which the Predictability rises and does not get low again, was close to the sentences' beginning. And, that was a limitation to compare pre- and post-RP regions. To overcome this issue, one possibility would be to add a short sentence prior to both the unfamiliar and memory-encoded stimuli. But the separated sentence would be read and processed as independent content. Instead, we opted for extending the sentences by adding a neutral context, of the right length to align the RP.

Firstly, to the 50 memory-encoded sentences from the previous study, we added 68 new sentences, obtained from a survey carried out through social networks. These new sentences are proverbs or sayings, verses from song lyrics, popular politicians or TV hosts phrases, etc. Secondly, a cloze-task was performed using all 118 sentences. Results from this experiment were used to manually define the RP on each sentence. The position of the RP was defined semi-automatically as the largest positive difference between the cloze-Predictability of one word with its predecessor. Thirdly, after this RP estimation, 93 memory-encoded sentences were selected. Sentences without a clear RP (i.e. sentences without a jump in the cloze-Predictability or with very-low cloze-Predictability words after it) were discarded. Finally, a neutral context was added to each of them at the beginning, in order to align the position of the RP at the target position of the sixth word. Neutral context discursive type was balanced between direct, indirect and non-referred. For example, the sentence *''Cuatro ojos ven más que dos*'' (a Spanish saying literally translated to ''Four eyes see better than two.'', but also analogous to the English saying ''two heads are better than one'') whose RP is in position 2 (word ''ojos'') was transformed into the sentence **''Quiero tu opinión porque cuatro ojos ven más que dos.**'' (''I want your opinion because four eyes see better than two.'').

Control sentences were designed using the same neutral contexts (''I want your opinion because'' from the previous example) and completing with words selected from the memory-encoded sentences. This allowed to

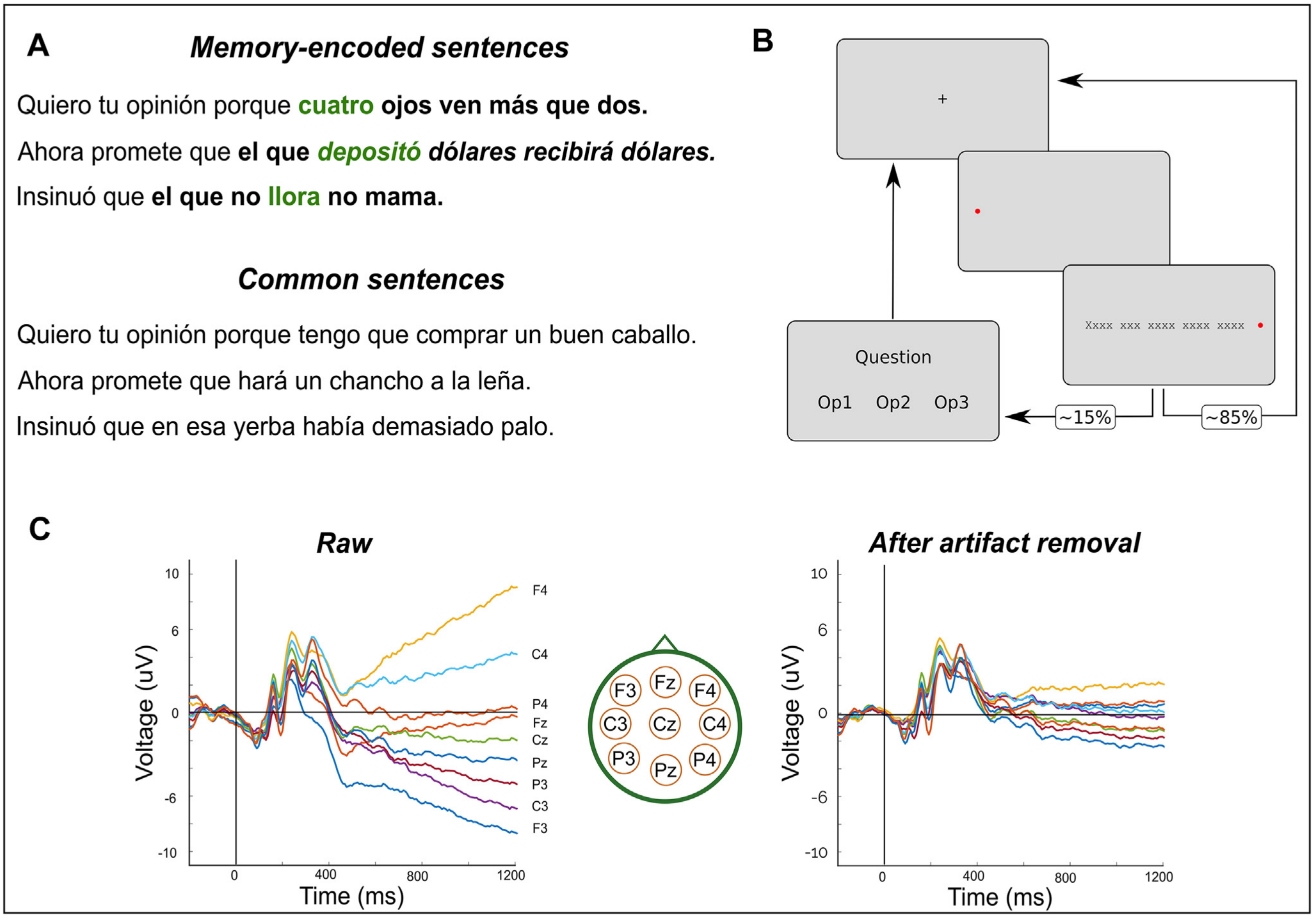

Fig. 1. Experimental design and EEG data pre-processing (A) Examples of common and memory-encoded sentences. Bold words represent the original memory-encoded sentence. The green word represents the Recognition Point; **(B)** Diagram of a trial of the experimental paradigm. A reading comprehension question with 3 options was asked randomly after 15% of the trails; **(C)** Average of all FRP before (left) and after (right) artefact removal for 9 electrodes distributed around the scalp (centre). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

have a balance between the types of sentences. All the stimuli used are presented in B (see Fig. 1A for example sentences).

A new cloze-task was performed for these sentences to obtain the final cloze-Predictability for all the words and the RP for the modified memory-encoded sentences. This new RP definition showed only slight changes relative to the target position (pre-defined: 6, actual mean: 5.78 [SD: 0.08, min: 3, max: 7]). Notably, the cloze-Predictability on memory-encoded sentences after the RP is not completely flat at the maximum value. There is certain variation within high values (2).

Additionally, all the sentences in the corpus were used in an online survey to measure the Familiarity. The survey was answered by 101 participants, 15 of which took part of the EEG-ET experiment more than 12 months before. They rated how familiar each sentence was from 1 (“I have never read it”) to 5 (“This is a very known sentence”) each sentence. The average familiarity for the memory-encoded sentences was 4.61 (SD = 0.41) and for common sentences 1.55 (SD = 0.48) (Fig. A1). These results show that our classification is in line with the proposed familiarity.

EEG

Brain electrical potentials were recorded in a shielded room with a Biosemi Active-Two[1] (Amsterdam, Netherlands) equipment with 128 active electrodes. Four reference electrodes (1 in each ear lobe and 1 on each mastoid) and 4 electrodes around the eyes for EOG recording (1 in each temple and the other 2 above and below the right eye) were added. After placing all the electrodes and ensuring their correct connection, the eye-tracking (ET) equipment was calibrated. EEG data was recorded at 1024 Hz using the built-in references. The EEG signal was visually monitored during the reading session on the ActiView software in a 50 μV Y-scale and with the default lowpass (100 Hz) and highpass filters (0.16 Hz). Noisy electrodes were corrected during the recording session if possible. Data was saved using the default filters of the apparatus [2].

[1] http://www.biosemi.com.
[2] https://www.biosemi.com/activetwo_full_specs.htm.

Eye-tracking

Eye movements were recorded with an EyeLink 1000 equipment (SR Research Ltd., Ottawa, Ontario, Canada[3]). After EEG and EOG electrodes were set in position, participants were placed in front of the stimulus screen. A chin rest was used to ensure a stable head position during the experiment. Participants were instructed to avoid moving their head and not to abandon the position on the chin rest. Eye movements were recorded binocularly at 1 kHz sampling rate. The best calibrated eye on each block was used for all the analyses (right eye was used 85% of the times). Fixation and saccade metrics were calculated by the eye-link software. We used the recommended configuration for Cognitive experiments (recording_parse_type = GAZE, saccade_velocity_threshold = 30, saccade_acceleration_threshold = 8000, saccade_motion_threshold = 0.1, saccade_pursuit_fixup = 60, fixation_update_interval = 0).

Apparatus

Visual stimuli were prepared and presented using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). They were displayed on a 22-inch LCD monitor located 60 cm from the participant's eyes, at a refresh rate of 60 Hz and a resolution of 1920x1080. Sentences were presented on a monospaced font (*courier new*), on a font size corresponding to 0.38° visual angle.

Procedure

Each trial started with a fixation cross in the centre of the screen for a fixed period of time (2 s) (Fig. 1B). Then, a red circle was presented at the place of the first letter of the corresponding sentence. Upon detecting the arrival of the participant's vision to the circle, it disappeared giving place to the presentation of the complete sentence. To minimise fatigue, the experiment was divided into five blocks of 42 sentences, with 4 unlimited-time breaks. During each break, participants were allowed to rest and change their posture. At the beginning of each block, an ET calibration was performed using a 9-point grid.

Participants were instructed to read each sentence as naturally as possible, with the sole task of understanding its meaning. After reading each sentence, participants were instructed to fixate on a red dot placed 5 characters to the right of the end of the sentence. When the gaze reached this point, the sentence disappeared. Once the reading of each sentence was finished, a multiple-choice comprehension question (with 3 options) about the previously read sentence could appear (15% of the times) to verify the engagement of the subjects with the material. The aim of these questions was to ensure that participants were performing a conscious reading. For example, after the sentence "Lara never eats bread when she is at home", the question was "What doesn't Lara eat?", and the options were "Bread", "Banana", and "Rice" (see B). The mean proportion of correct answers was 95% (greater than 88% for all subjects), showing a conscious reading of all participants. Since there was no time limit to respond, participants were also instructed to use this time to rest their eyes, without removing their head from the chin rest. Before starting, the participants performed 3 training trials that were not analysed.

[3] https://www.sr-research.com/.

Eye-map

Eye muscle contraction, eyeball movement, and blinks generate electrical potentials (König et al., 2012) that interfere with the measurement of the brain potentials. An eye-map "pre-experiment" was performed to obtain a record of electrical potentials associated exclusively with eye movements. According to the recommendations in the literature (Dimigen et al., 2011; Plöchl et al., 2012), participants performed two 8-min trials of horizontal and vertical eye movements each, alternating their gaze between two points located in the middle lines of the screen in both directions. These points were placed to measure electrical potentials caused by "short" (50% of the width and height of the screen) and "long" saccades (75% of the width and height of the screen). Additionally, Participants were also instructed to blink at a frequency of approximately 1 Hz. Each trial consisted of 4 measurements of blinks, 4 of horizontal saccades (2 long and 2 short), and 4 of vertical saccades (2 long and 2 short). Each of the measurements was performed for 30 s.

Data pre-processing

EEG and ET computers were synchronised by hardware using marks sent from the stimulus computer at the same time through different ports. EEG computer received the marks via 8-bit parallel port. ET computer received them via an Ethernet connection.

Preprocessing was performed in MATLAB (R2018b, Linux version) using the EEGLAB toolbox (v14.1.1) (Delorme and Makeig, 2004; Makeig et al., 2004), the EYE-EEG plugin (v0.81) (Dimigen et al., 2011), the LMM-CBP package (V0.1) (Bianchi et al., 2019), and in-lab scripts. Preprocessing was performed following previous lab work (Kaunitz et al., 2014; Kamienkowski et al., 2018) and based on the recommendations proposed by Olaf Dimigen for coregistration studies (Dimigen, 2020).

EEG data from all the subjects was re-referenced to mastoids average. Then, following Dimigen's recommendation, a pre-filter was applied using the *pop_eegfiltnew()* function (*lower limit* = 0.1 Hz, *upper limit* = 100 Hz, *order* = 3*round(sampling_rate/low-cut off) = 15360). A notch filter was also used to remove interference from the alternating electric current (*lower limit* = 47.5 Hz, *upper limit* = 52.5 Hz, order: 846). These filters were applied only on the EEG electrodes, and not on the EOG. To finish the EEG data pre-processing, defective channels were selected by visual analysis of potentials. Defective channel was defined as a channel where the raw signal variance clearly exceeds more than 5 standard deviation of the rest of the signal during several trials. When more than 5 channels showed a noisy behaviour during a short period of time only affected trials were removed. Defective channels were eliminated and their signal was estimated by interpolation of the signals registered by neighbouring channels using the *eeg_interp()* EEGLAB function, with the *invdist* method. On average, 1.9 channels were interpolated per subject (range = [0–7]). After EEG preprocessing, data from the ET was incorporated into the EEG data structures using the *pop_importeyetracker()* function of the EYE-EEG plugin. Synchronisation of the marks sent during the acquisition to both equipment was carefully analysed for each subject.

Secondly, ocular artefacts were selected and removed. Independent Component Analysis (ICA) was performed for each subject on the experiment and the 2 eye-map sessions using EEGLAB's *pop_runica()* function to obtain 64 independent components (function parameters: 'icatype','binica';'extended',1;'pca',64). Independent Components associated with Eye Movements were automatically selected using the *pop_eyetrackerica()* function (function parameters: 'sactol', [2 0]; 'threshratio', 1.1). This function implements the criteria presented by Plöchl et al. (Plöchl et al. (2012)). Additionally, all components were manually analysed to ensure the best possible artefact selection, selecting not only eye artefact, but also independent components that could come from head movements and other interference sources. All the selected components were removed from the EEG data (Fig. 1C).

Finally, experiment data was epoched on fixations. For this, fixations were assigned for each trial to its corresponding word. Fixations above or below 15% of the screen height from its centre were removed. Subjects with fixations outside this threshold on more than 20% of the trials were marked as poorly calibrated. Epochs were defined around the first fixation of each word. Since in these experiments there was no stimulus start, fixation onset was used as reference. Thus, instead of defining Event-Related Potentials (ERP), Fixation-Related Potentials (FRP) were analysed. FRP were defined from 200 ms prior to the beginning of fixation up to 600 ms after it. Epoch baseline was calculated from −200 to −50 ms to avoid the saccadic spike potential that occurs from −50 ms to fixation onset. Finally, data were down-sampled to 128 Hz. This procedure was performed for the original scalp potentials and for the frequency bands power previously estimated.

For the Frequency Analysis EEG data was filtered before epoching. Frequency band power was calculated for Theta (4–8 Hz), Alpha (8–16 Hz), and lower Beta (16–20 Hz) bands. For this, EEG data was filtered using the *pop_eegfiltnew()* function. Then, the power was calculated as the absolute value of the Hilbert transform (Cohen, 2014) for each electrode.

Statistics

LMM-CBP analysis: Then, we performed the analysis presented in the previous study based on Linear Mixed Models (LMM) and the non-parametric procedure for correction for multiple comparisons (LMM-CBP) (Bianchi et al., 2019). This method consists on fitting an LMM for each time-electrode sample of the whole FRP from all the subjects (i.e. 90 time samples by 128 electrodes, 11,520 LMM). The t-values for each fixed term were analysed to find when and where locally significant effects appear (abs(t-val) > 2). Since multiple LMM were run, a non-parametrical correction was performed by permuting the epoch information, running all the LMM again and selecting significant samples by a cluster analysis. For each fitted model (see below), 3,000 data permutations were performed (a posterior analysis showed that p-values vary less than 5% after 500 permutations and no significant variations after 1000 permutations). This procedure is based on Maris and Oostenveld (2007). The toolbox was used with the parameters proposed in the tutorial: sample significance threshold was set at 1.86, according with a p-value of 0.5 of a normal distribution; a sample was part of a cluster if it was significant and has at least 1 temporal and 2 spatial significant neighbours; significant clusters were those with a t-value sum greater or lower than 95% of the random-generated clusters (two-tails).

In the present study, three different models were fitted (Eq. M0, Eq. M1, and Eq. M2). Here, *freq* represents the logarithm of the lexical frequency of the fixated word, *pos* is the ordinal position of the word in the sentences, *pred* is the logit of the cloze-Predictability of the fixated word, *type* is the sentence type (0 for memory-encoded, and 1 for common sentences), *typeRP* is the sentence type splitting memory-encoded sentences in before and after RP (0 for preRP words from memory-encoded, 1 for postRP words from memory-encoded, and 2 for common sentences), *fixDur* is the current fixation duration, and *saccDur* is the inward saccade duration.

$$y = freq + pos : type + pred : type + (1|subj) \tag{M0}$$

$$y = freq + pos : type + pred : typeRP + (1|subj) \tag{M1}$$

$$y = freq + pos : type + pred : typeRp + fixDur + saccDur + (1|subj) \tag{M2}$$

All the variables were centred and scaled before fitting the models. All the models were run on the exact same data. Due to convergence failures, only subject was used as random variable. Nevertheless, whenever a model failed to converge additional iterations were run. In case of no convergence, all the co-variables were set as no-significant This happened on less than 0.7% of the models. See Bianchi et al. (2019) for more details. Both *type* and *typeRP* co-variables were coded as treatment contrasts. Main effects of *pred* and *pos* were exclude from the models, but in case of including them, they would have the exact same results as the first level of the categorical co-variables (i.e. the level coded as 0). All the necessary functions needed to perform these analyses are available in the LMM-CBP toolbox[4].

[4] https://github.com/brunobian/LMM-CBP.

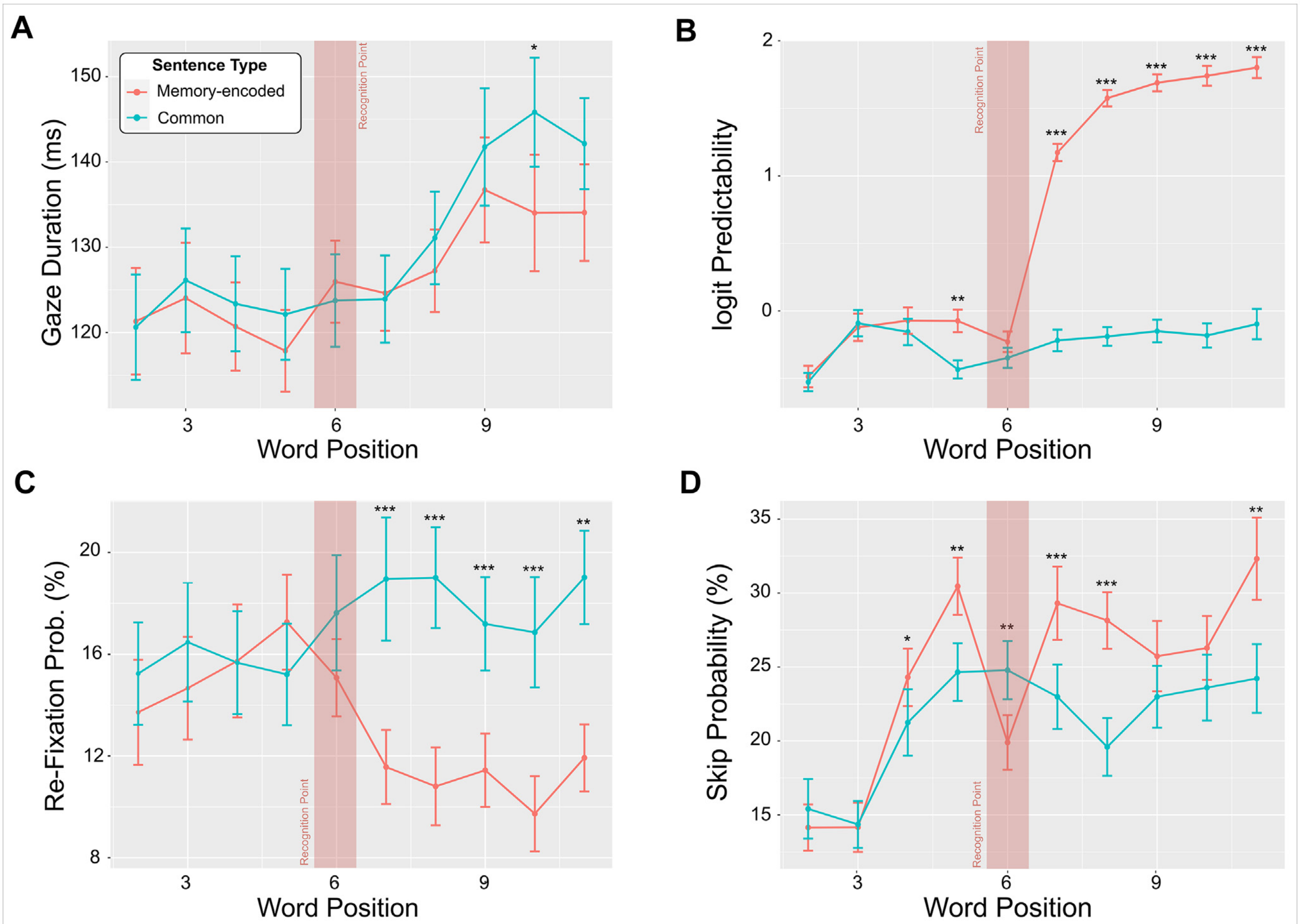

Fig. 2. Behavioural analyses along the sentence Blue lines correspond to common sentences. Red lines correspond to memory-encoded sentences. Error bars correspond to standard deviation between subjects. Red shaded area indicates the approximate position of the RP on memory-encoded sentences **(A)** Gaze Duration calculated as the sum of all the fixations on the first pass; **(B)** logit of cloze-Predictability; **(C)** Re-fixation Probability, calculated as the proportion of subjects that re-fixated the word; **(D)** Skipping Probability for each subject was calculated as the proportion of skipped words in the first pass. **Statistical Analyses:** Linear Mixed Models were fitted between word position (as factor) and each one of the 4 other variables interacting with sentence type (as factors). On Gaze Duration, Skip probability and Re-fixation probability analyses Subject and Sentence ids were used as random factor. On cloze-Predictability analysis, only Sentence id was used as every subject read all the sentences. Pairwise contrast were performed using the *emmeans* package for R to obtain a comparison between type for each word position. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

RESULTS

Reading behaviour

Gaze Duration along the absolute word position follows the expected dynamic for both types of sentences (Fig. 2A). Reading is performed at a constant speed for the first seven words and then starts to slow down toward the sentence end, where the so called wrap-up effect is observed (Kuperman et al., 2010). Significant differences were found between words from common and from memory-encoded sentences only for words on the 10th position ($p_{10th} = 0.0271$), despite significant differences on cloze-Predictability from 5th word up to the sentence end (Fig. 2B, $p_{5th} = 0.0012, p_{6th} = 0.0185, p_{+7th} < .0001$). Note that for memory-encoded sentences the Recognition Point (RP) is usually located at the 6th word.

When analysing variables from eye movements, significant differences are mainly observe on those variables that reflect late processing (Clifton et al., 2007; Demberg and Keller, 2008). That is, those variables that reflects the processing of the words on a first pass, like First Fixation Duration (Fig. A2A) or Gaze Duration (Fig. 2A) does not shows any significant differences. But, significant differences emerge when analysing variables that reflects processes that are performed late on the reading dynamic. For example, a significant decrease on the probability of re-fixation is observed for the postRP words (Fig. 2C, $p_{7-10th} < 0.0001, p_{11th} = 0.0016$). These differences are also observed on Re-reading Time (Fig. A2D), Total Regressions In (Fig. A2E), and Total Number of Fixations (Fig. A2G). Additionally, the skipping probability (Fig. 2D) shows an increase in skips on the word immediately prior to the RP (5th word, $p_{5th} = 0.0062$), a decrease on the RP (6th word, $p_{6th} = 0.0046$) and a new increase on the following two words (7th $p_{7th} = 0.0004$ and 8th words $p_{8th} < 0.0001$). This suggests that the RP is influencing the reading at

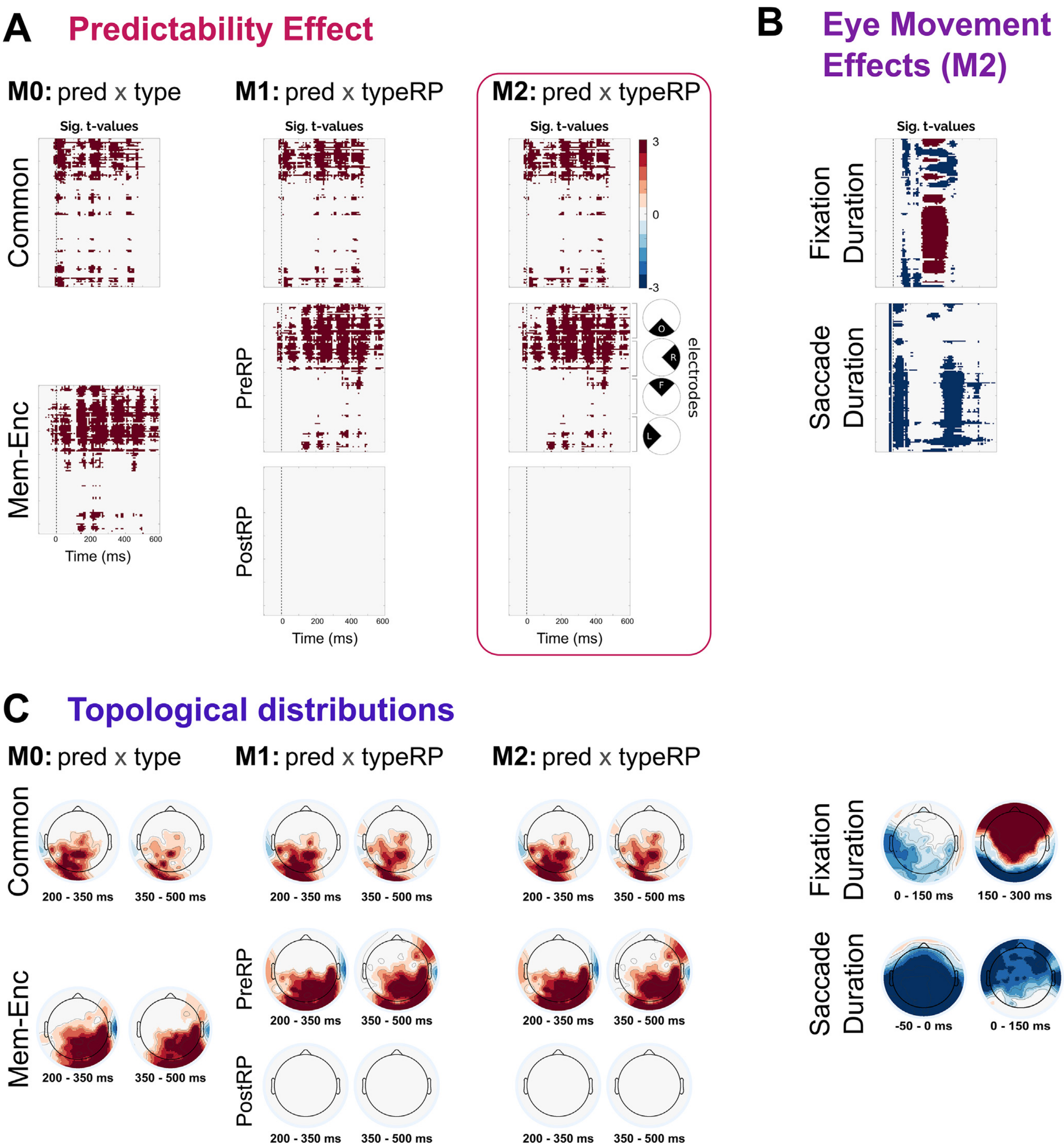

Fig. 3. Main effects from fitted LMMs (A) Predictability effect from Models M0, M1, and M2 (the ultimate model). Panels show the significant t-values after the non-parametrical correction using a Cluster-based permutation test. On M1 and M2, memory-encoded sentences (Mem-Enc) were divided into pre and postRP words; **(B)** Eye movement effects from M2. Panels show the significant t-values after the non-parametrical correction using a Cluster-based permutation test. Duration of the current Fixation (top) and of the inward saccade (bottom); **(C)** Topological distributions of the significant t-values shown on (A) and (B) for the [200–350 ms] and the [350–500 ms] time windows.

least from its previous word. Also, that the RP is a key point from where to analyse the rest of the sentence.

Fixation Related Potential (FRP) analyses

Since we based this study on a previous work on SVP reading of common and memory-encoded sentences (Bianchi et al., 2019), we fitted the same Linear Mixed Model (LMM) than Bianchi et al. (Model 0). These results serve as a baseline to compare to the analyses specifically defined for the present study. In short, the new set of stimuli is comprised by modified memory-encoded sentences, that count with neutral previous contexts (see Methods for more details). These contexts allow to split memory-encoded sentences in two sections (before and after the recognition point), and both sections are large enough to perform comparisons between them (Model 1).

In this analysis, LMMs were fitted for each epoch (i.e. fixation in each word) against the word lexical frequency, the absolute position in the sentence, and the cloze-Predictability. Also, two interaction terms of these latter variables with the sentence type (common or memory-encoded) were included (Bianchi et al., 2019) (Model 0 –M0–, see Eq. M0). Results from this model for the Predictability term (Pred x Type in Fig. 3A) shows a significant positive parieto-occipital effect, lasting from the fixation onset (0 ms) to the FRP end (600 ms) in both types of sentences. This Predictability effect in the FRP signal was very similar in scalp distribution to the effect observed for common-sentences in the foveated Serial Visual Presentation paradigm by Bianchi et al. (2019). Thus, we interpreted it as a N400-like effect although its peak appeared earlier (between 150 and 250 ms).

In contrast to our previous results, where there was no N400-like effect for memory-encoded sentences, the same model (M0) shows little difference between the effect for common and memory-encoded sentences (Fig. 3A, left). This is an expected result, since in the present study, memory-encoded sentences included a longer neutral context before the RP. Thus, the observed N400-like effect could be due to this preRP words. To take into account this difference in the used stimuli, we split the memory-encoded sentences in words before or after the RP. We called this co-variable typeRP (Model 1 –M1–, see Eq. M1).

Results for the cloze-Predictability term from this model (Fig. 3A, middle) shows that after recognising the memory-encoded sentences (postRP), the N400-like effect is not present. This result is in line with our previous study replicating the main finding and reinforcing the hypothesis that, after recognising and retrieving the sentence from memory, mechanisms that elicits the N400 component reduce their influences in the word processing.

A possible confound could be that the distribution of cloze-Predictability for postRP words has a low variance, as compared to preRP and words from common sentences (Fig. 2B). The low variance on cloze-Predictability could be responsible for the absence of observed N400 on these words. To discard this, we fitted the M1 model to a subset, comprised by postRP words and words from common sentences resampled to match the variance (Fig. A5). On this model the interaction between cloze-Predictability and common sentences reproduced effect from the whole dataset, while no effect was found for the interaction between cloze-Predictability and postRP words (Fig. A6). We conclude that, the absence of the N400 component on the postRP words is not due to the differences in variance on cloze-Predictability.

Eye movement co-variables

In addition to understanding how word-level variables modulate brain potentials, LMMs can be useful to capture and isolate effects from other confounding variables. For instance, it is interesting to use this statistical approach to model the influence of eye movements variables that could interfere on the analyses. For this, a new model was fitted (Model 2, –M2–, see Eq. M2) adding the Fixation Duration and the (inward) Saccade Duration to M1 as fixed variables. The addition of these variables did not modify the results obtained on M1 for the Predictability term (Fig. 3A, right) and the rest of the fixed terms (Figs. A3 and A4).

Observed effects for Fixation Duration show a left parietal negative effects around 50 ms after the fixation onset and a large fronto-central positive effect around 200 ms (Fig. 3B,C). Altogether these effects resemble to the lambda potential observed in free eye movement experiments (Yagi, 1981; Dimigen et al., 2011). Similarly, the inward saccade duration variable generates an early effect, even with a significant cluster some milliseconds before the fixation onset. This brief and widely spread effect could be linked to the spike potential (Keren et al., 2010; Plöchl et al., 2012; Kliegl et al., 2014). Also, a left temporal negative effect lasting for around 100 ms is observed twice: the first time starting on the fixation onset; and the second time starting at 250 ms. We interpret the second cluster as a result of the outward saccade, which will typically have the same direction and similar amplitude (Fig. A7). Overall, this model outperformed the M1, and will be used for further analyses.

Time–frequency analyses

We hypothesise that different sources of predictions have different correlates in the time–frequency domain (Lewis and Bastiaansen, 2015; Himmelstoss et al., 2020). Thus, to further analyse the observed effects, we calculated the power for different frequency bands: theta band from 4 to 8 Hz, alpha band from 8 to 16 Hz, and beta band from 16 to 20 Hz. The same LMM-CBP analyses (M2) performed on FRP was applied on each band. First, no significant effects were found for any of the analysed variables on the theta band.

On the alpha band, a negative central effect was found for Position on memory-encoded sentences (Fig. 4A). Additionally, a positive frontocentral effect was observed for Predictability on common sentences and pre-RP words from memory-encoded sentences (Fig. 4B). The continuous data shows a larger occipital alpha lateralization for memory-encoded sentences (Fig. 4C).

On the beta band, a right occipital positive effect was found for the word Position, only for words from memory-encoded sentences (Fig. 5A). Additionally, a right occipital negative effect is observed for Predictability only on post-RP words of memory-encoded sentences (Fig. 5B). The difference on the power for the continuous data topological distribution also shows a greater occipital beta power for memory-encoded sentences in comparison to common sentences (Fig. 5C). Similar to the observed effect on the alpha band, differences are visible from the sentence beginning. This difference increases from very early in the reading time until the mean RP latency (~850 ms). Then it decreases at the RP and starts rising again. This non-linear behaviour

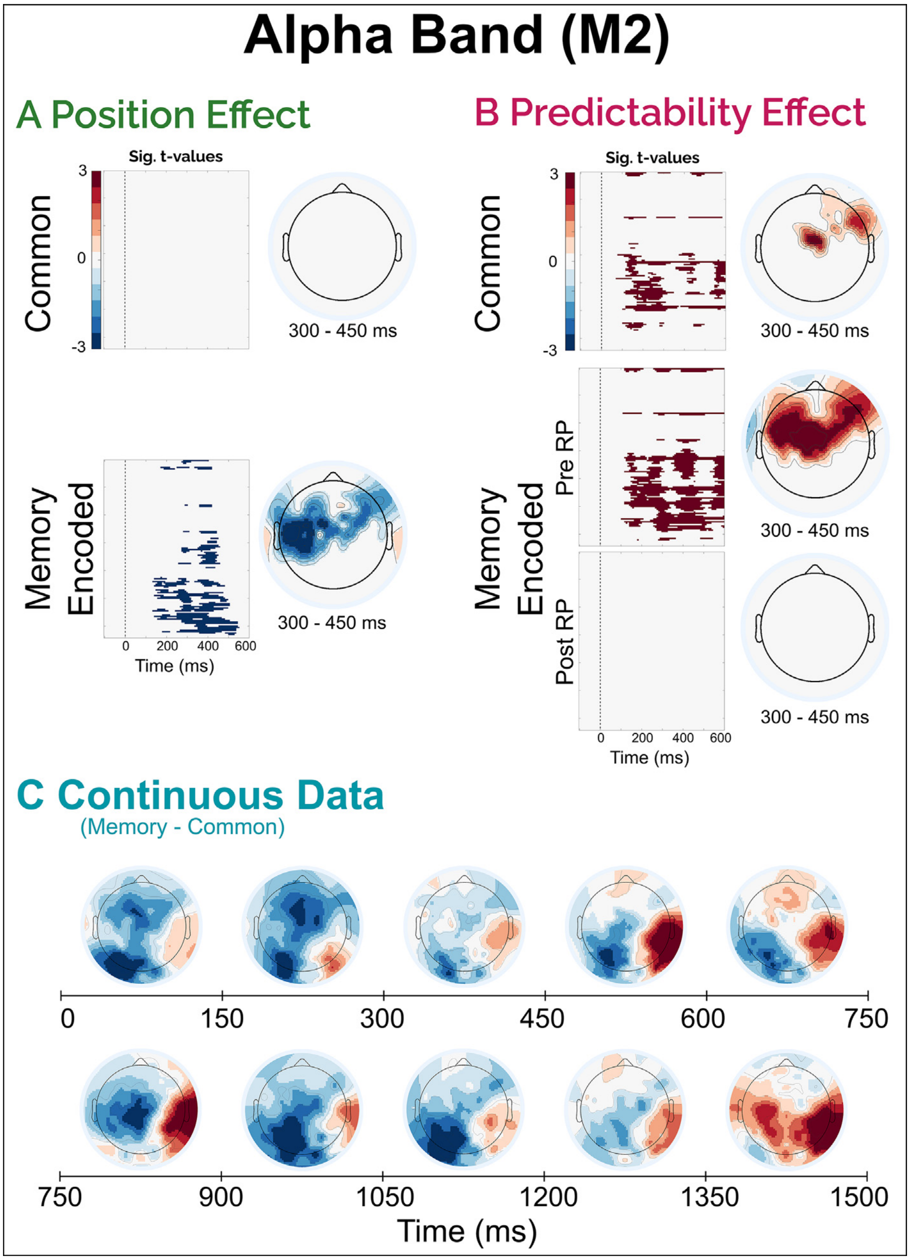

Fig. 4. Results for M2 on alpha band (8–16 Hz) (A) Position effect: t-values and topological distribution on the 300–450 ms time windows; **(B)** Predictability effect: t-values and topological distribution on the 300–450 ms time windows; **(C)** Mean alpha power on continuous data from the sentences start. RP mean latency: 877 ms.

could be explained by the combination of the Position effect and the Predictability effect after the RP in the right parieto-occipital regions.

DISCUSSION

Most of the knowledge about the brain mechanisms that underlie the reading process acquired during the last decades is based on foveated reading (Kutas and Federmeier, 2011; Dambacher and Kliegl, 2007; Dambacher et al., 2012). Thus, the study of natural reading will shed light on the actual processes that occur when the reader is free to move his/her eyes across the text. The main limitation to examine this type of reading is dealing with electrical artefacts from eye movements (EM) (Keren et al., 2010; Plöchl et al., 2012; Dimigen, 2020). More recently, and thanks to the rise in the computational power, EEG and ET co-registration studies have become more tractable and common in the neurolinguistic field (Dimigen et al., 2011; Kornrumpf et al., 2017; Dimigen, 2020; Himmelstoss et al., 2020).

The goal of the present study was to analyse how predictions are performed during natural reading of sentences, allowing participants to read them moving their eyes across the words at their own pace. In particular, we aimed to analyse how predictions are performed when readers know how sentences continue from a certain word up to the end. For this, sentences from popular culture were used. This sentence corpus was comprised by sayings, song lyrics, famous phrases, etc. In a previous study, we defined them as "memory-encoded sentences" (Bianchi et al., 2019). There, we analysed foveated reading and found no N400 effect for this type of sentences. Additionally, Fernández et al. (2014) have also used this type of sentences in an eye-tracking (ET) experiment. They found differences in eye movements variables between previously known and highly predictive sentences and low predictable sentences. For example, the former ones showed a high skip probability than latter.

Memory-encoded sentences used in the present study are similar to those used by Bianchi et al. (2019) and Fernández et al. (2014). On these studies, they reported that memory-encoded sentences were recognised on words too close to the sentence beginning. The addition of these neutral contexts allowed us to align the Recognition Point (RP) between sentences and to perform analyses on preRP words.

Eye movement variables that reflect late processing showed differences in the reading behaviour between both types of sentences. In particular, we observed a rise in the skip probability in the words before the RP, and a decrease on the RP word. This suggests that the

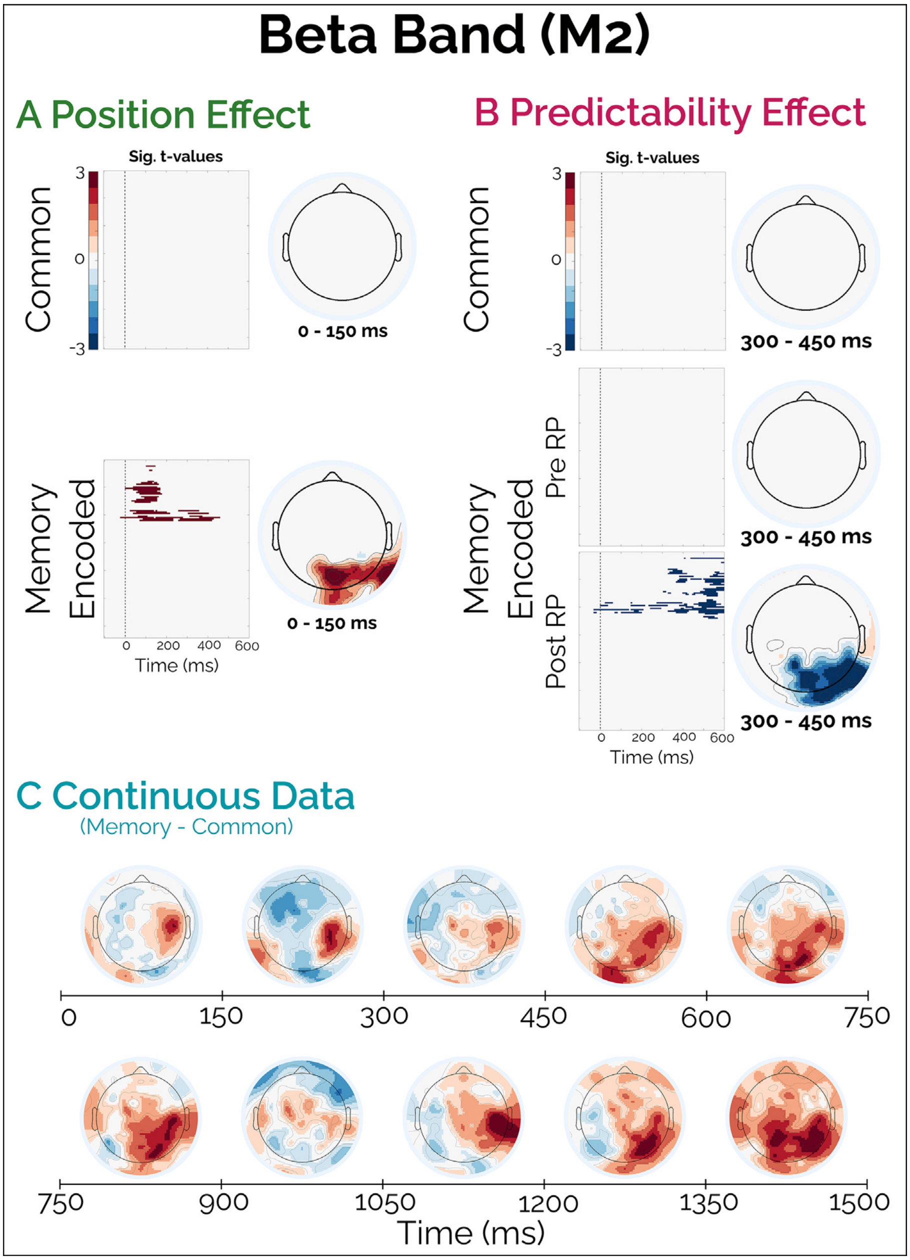

Fig. 5. Results for M2 on beta band (16–20 Hz) (A) Position effect: t-values and topological distribution on the 300–450 ms time windows; **(B)** Predictability effect: t-values and topological distribution on the 300–450 ms time windows; **(C)** Mean alpha power on continuous data from the sentences start. RP mean latency: 877 ms.

RP is generating an attraction of reader's gaze, even before fixating it. Then, after the RP, we observed a decrease on re-fixations, the re-reading time, the total fixation time, and the Total Number of Fixations, and more skips. From these results we can conclude that reading becomes easier after recognising and the sentence, and it is performed in a more straight forward way. Since Gaze Duration showed very low values for all the words (around 130 ms) we assumed that the lack of effects is due to a floor effect. That is, Gaze Duration is so short that there is no place for shorter gazes, even if reading becomes easier. We hypothesise that this floor effect is due to a sentence corpus that did not raise any particular difficulty for the readers. This is supported by the high performance on the comprehension questions (95% of correct answers). Note that the corpus was constructed with 50% of highly popular sentences, and the other 50% were designed to avoid difficulties during reading and undesirable effects on eye movements and scalp potentials.

Analyses of the Fixation-Related Potentials (FRP) aligned at the beginning of the first fixation on each word, showed the typical shape observed for the Event-Related Potentials (ERP) from foveated reading experiments (Bianchi et al., 2019). Nevertheless, when analysing these FRP by Predictability terciles, Sentence Type, or their interactions, no differences were observed (Fig. A5). This may be striking, since in previous EEG and ET co-registration studies the classic effect of cloze-Predictability was found (Dimigen et al., 2011). However, those studies used stimuli carefully designed, sampling target words from defined categories of Predictability (for example low- and high-Predictable words), with all the non-target words matching on various lexical properties that could interfere on the analysis. As discussed by Hamilton and Huth (2020), these stimuli could exacerbate and isolate the desired effects from the rest of cognitive processing to be evaluated with simple statistics. In contrast, in experiments performed under environments with more variability, similar to the daily life experience, data should be analysed with more robust and multivariate statistical methods in order to isolate the effects. For example, in the LMM-CBP procedure (Bianchi et al., 2019), a Linear Mixed Models (LMM) with all the co-variables of interest is fitted for each electrode-time sample. This allows to use continuous variables, to test interactions, and to avoid averaging in pre-established time windows or regions (Dimigen et al., 2011). Using this method, we analysed the exact same model used by Bianchi et al.

(2019) (Model 0 –M0–), and extended it to a model where memory-encoded sentences were splitted in words before and after the RP (Model 1 –M1–). On Model 2 (M2) eye movement variables (First Fixation Duration and inward Saccade Duration) were added. We found that only words from common sentences and preRP elicited an N400-like effect. That is, the N400 potential is only sensible to the cloze-Predictability before readers recognise the sentence. This analysis was only possible due to the addition of a neutral context to these sentences. Furthermore, this result is not due to the low cloze-Predictability variance of the postRP given that words from common-sentences with a similar variance did show the N400-like effect.

On the M2, we observed that the inclusion of eye movement variables on the LMM did not interfere with the results obtained on the relevant experimental co-variables. This implies that the effects from all the analysed co-variables on M1 are not influenced by ocular movements. Additionally, our results strengthen the observations of sustained responses to cloze-Predictability during more than one fixation. On future analyses this could serve as a model to understand how the ocular artefacts interact with lexical properties.

Fitting EM variables along variables of interest into the same statistical model is a great advantage of the LMM-CBP procedure. Nevertheless, this analysis could not be enough for modelling and reducing all the artefacts present on natural reading EEG data. That is, effects found in the present study exceed the fixation duration, in latencies where the eye are already over one (o more) words ahead. Thus, it could be possible that, at some extent, part of the observed effects are confounds due to differential overlap between conditions. At the time of this publication, methods that account for these overlaps are not able to perform such complex analyses as those presented here. For example, the Unfold toolbox (Dimigen and Ehinger, 2021) models the overlap by fitting a Lineal Model with information about fixations and co-variables of interest. Nevertheless, since this method uses the continuous EEG data and fits many co-variables for each model (number of effects by number of time samples on the FRP window) it is not computationally possible to use the data from all the subjects (or any other random variable in general), and it is typically implemented on a individual-subject basis.

Altogether, these results extend previous conclusions: when the readers find out that they are reading a sentence that they know how it continues, the brain mechanisms responsible for the N400 potential cease to have a role on future predictions.

Although the exact brain bases of the N400 component is not yet established, two classical views are found in literature. On the one side, the integration interpretation posits that the N400 reflects the effort on integrating the current word with the context. On the other side, the lexical-access interpretation posits that it reflects the pre-activation of lexical features of words (See Lau et al. (2008) for a further discussion). Regardless which interpretation is correct, the lack of N400 effect after recognising the sentence implies that after the RP the sentence is read without neither a combinatorial nor a facilitation process (Bianchi et al., 2019).

To follow up this hypothesis, analyses on the power of theta, alpha, and beta frequency bands were performed using the same LMM-CBP methodology. Results showed significant effects for Position and Predictability on alpha and beta bands. No significant effects were found for theta band, which is commonly linked to memory processes (Cohen, 2014), although it is more associated with encoding and not recall processes.

On the alpha band, an occipital lateralisation was found on the continuous data, with a power increase ipsilateral to the direction of reading. Interestingly, this lateralisation was greater for the memory-encoded than for the common sentences. According to previous studies on natural reading, the lateralisation on this frequency band reflects the attention to future words (Kornrumpf et al., 2017). Thus, our results suggest that when reading memory-encoded sentences, readers have their attention more distributed on future words than when reading a common sentence. This is in line with the EM behaviour results, where we found that these sentences are read more straightforward, with less refixations and regressions. Furthermore, results from previous EM studies on these sentences showed enhanced parafoveal-on-foveal effects after recognising the sentences (Fernández et al., 2014).

On the beta band, an occipital positive effect was found for the position co-variable on the LMM analyses only for memory-encoded sentences. According to the Predictive Coding framework, the beta band reflects the feedback connection between hierarchical layers of cognitive processing (Lewis and Bastiaansen, 2015). An increase in the beta band power signals the maintenance of the neural configurations produced by future words predictions. In reading, this corresponds to a mental model that is constructed from the incoming stimuli. This model is constantly updated after each new incoming word, and in some cases, a word can disrupt this model. But, when reading a memory-encoded sentence, a memory retrieval is performed around the RP. At this point, the rest of the sentence is retrieved and the model is updated for all the upcoming words. Since successive stimuli will reinforce the model, this will be constantly maintained.

In conclusion, our results suggest that during natural reading there are continuous processing of the prediction of future words, which produce a model about the upcoming words. These processes generate a cognitive set that is actively maintained as the predictions are corroborated, but that can be modified if the predictions are not correct. The maintenance of this set is reflected mainly in increase of the power of the beta frequency band, and generates the decrease of the N400 potential. Even more, under an extreme situation, where the predictions are performed based on knowing exactly how the sentences continue, the N400 potential is suppressed. Finally, this established cognitive set allows a faster processing of incoming words, generating a more straight forward reading.

AUTHOR CONTRIBUTIONS

Bruno Bianchi: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - Original Draft, Visualization; **Rodrigo Loredo:** Investigation; **María Da Fonseca:** Formal analysis, Writing - Review & Editing; **Julia Carden:** Data Curation; **Virginia Jaichenco:** Data Curation; **Titus von der Malsburg:** Conceptualization, Writing - Review & Editing; **Diego E. Shalom:** Methodology, Resources, Writing - Review & Editing; **Juan Kamienkowski:** Conceptualization, Methodology, Resources, Writing - Original Draft, Supervision, Funding acquisition.

FUNDING SOURCE

The authors were supported by the National Science and Technology Research Council (CONICET) and the University of Buenos Aires (UBA). This study was supported by the following grants: UBA grant for short-term stay abroad 2018; National Agency of Promotion of Science and Technology grant (PICT 2016–1256); National Scientific and Technical Research Council grant (PIP 11220150100787CO).

DECLARATION OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AND CODE AVAILABILITY STATEMENT

The datasets analysed for this study can be found in https://github.com/brunobian/coregistration2022-analyses and in http://reading.liaa.dc.uba.ar. More detailed or complementary data are available on request.

ETHICS STATEMENT

All participants provided written informed consent in agreement with the Helsinki declaration. All the experiments described in this paper were reviewed and approved by the ethics comittee: "Comité de Ética del Centro de Educación Médica e Investigaciones Clínicas "Norberto Quirno" "(CEMIC)" and qualified by the Department of Health and Human Services (HHS, USA): IRb00001745 - IORG 0001315 (Protocol 435). Footer.

ACKNOWLEDGMENTS

We thank Marcos L. Pietto, for his assistance during data acquisition. We thank Luz Bavassi and Laura Kaczer for their thoughtful help on the time–frequency analysis.

REFERENCES

Axmacher N, Mormann F, Fernández G, Elger CE, Fell J (2006) Memory formation by neuronal synchronization. Brain Res Rev 52 (1):170–182.

Baayen RH (2014) Experimental and psycholinguistic approaches to studying derivation. In: Handbook of derivational morphology. p. 95–117.

Bastiaansen M, Hagoort P (2006) Oscillatory neuronal dynamics during language comprehension. Progress Brain Res 159:179–196.

Bianchi B, Monzón GB, Ferrer L, Slezak DF, Shalom DE, Kamienkowski JE (2020) Human and computer estimations of predictability of words in written language. Sci Rep 10(1):1–11.

Bianchi B, Shalom DE, Kamienkowski JE (2019) Predicting known sentences: Neural basis of proverb reading using non-parametric statistical testing and mixed-effects models. Front Human Neurosci 13:82.

Brainard DH (1997) The psychophysics toolbox. Spatial Vision 10:433–436.

Canal P, Pesciarelli F, Vespignani F, Molinaro N, Cacciari C (2017) Basic composition and enriched integration in idiom processing: An eeg study. J Exp Psychol: Learn Memory Cognit 43(6):928.

Cavanagh JF, Zambrano-Vazquez L, Allen JJ (2012) Theta lingua franca: A common mid-frontal substrate for action monitoring processes. Psychophysiology 49(2):220–238.

Clifton Jr C, Staub A, Rayner K (2007) Eye movements in reading words and sentences. Eye Movem:341–371.

Cohen MX (2014) Analyzing neural time series data: theory and practice. MIT Press.

Cronbach LJ (1957) The two disciplines of scientific psychology. Am Psychol 12(11):671.

Cronbach LJ (1975) Beyond the two disciplines of scientific psychology. Am Psychol 30(2):116.

Dambacher M, Dimigen O, Braun M, Wille K, Jacobs AM, Kliegl R (2012) Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. Neuropsychologia 50(8):1852–1870.

Dambacher M, Kliegl R (2007) Synchronizing timelines: Relations between fixation durations and n400 amplitudes during sentence reading. Brain Res 1155:147–162.

DeLong KA, Urbach TP, Kutas M (2005) Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nature Neurosci 8(8):1117–1121.

Delorme A, Makeig S (2004) Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. J Neurosci Methods 134(1):9–21.

Demberg V, Keller F (2008) Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. Cognition 109(2):193–210.

Dimigen O (2020) Optimizing the ica-based removal of ocular eeg artifacts from free viewing experiments. Neuroimage 207:116117.

Dimigen O, Ehinger BV (2021) Regression-based analysis of combined eeg and eye-tracking data: Theory and applications. J Vision 21(1). 3–3.

Dimigen O, Sommer W, Hohlfeld A, Jacobs AM, Kliegl R (2011) Coregistration of eye movements and eeg in natural reading: analyses and review. J Exp Psychol: General 140(4):552.

Dimigen O, Valsecchi M, Sommer W, Kliegl R (2009) Human microsaccade-related visual brain responses. J Neurosci 29 (39):12321–12331.

Dmochowski JP, Bezdek MA, Abelson BP, Johnson JS, Schumacher EH, Parra LC (2014) Audience preferences are predicted by temporal reliability of neural processing. Nat Commun 5(1):1–9.

Fernández G, Shalom DE, Kliegl R, Sigman M (2014) Eye movements during reading proverbs and regular sentences: The

incoming word predictability effect. Language Cognit Neurosci 29 (3):260–273.
Grieve J (2021) Observation, experimentation, and replication in linguistics. Linguistics 59(5):1343–1356.
Hamilton LS, Huth AG (2020) The revolution will not be controlled: natural stimuli in speech neuroscience. Language Cognit Neurosci 35(5):573–582.
Himmelstoss NA, Schuster S, Hutzler F, Moran R, Hawelka S (2020) Co-registration of eye movements and neuroimaging for studying contextual predictions in natural reading. Language Cognit Neurosci 35(5):595–612.
Huth AG, De Heer WA, Griffiths TL, Theunissen FE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532(7600):453–458.
Huth AG, Nishimoto S, Vu AT, Gallant JL (2012) A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76 (6):1210–1224.
Just MA, Carpenter PA (1980) A theory of reading: From eye fixations to comprehension. Psychol Rev 87(4):329.
Kamienkowski JE, Varatharajah A, Sigman M, Ison MJ (2018) Parsing a mental program: Fixation-related brain signatures of unitary operations and routines in natural visual search. NeuroImage.
Kaunitz LN, Kamienkowski JE, Varatharajah A, Sigman M, Quiroga RQ, Ison MJ (2014) Looking for a face in the crowd: fixation-related potentials in an eye-movement visual search task. NeuroImage 89:297–305.
Keren AS, Yuval-Greenberg S, Deouell LY (2010) Saccadic spike potentials in gamma-band eeg: characterization, detection and suppression. Neuroimage 49(3):2248–2263.
Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C (2007) What's new in psychtoolbox-3. Perception 36(14):1–16.
Kliegl R, Dambacher M, Dimigen O, Jacobs AM, Sommer W (2012) Eye movements and brain electric potentials during reading. Psychol Res 76(2):145–158.
Kliegl R, Dambacher M, Dimigen O, Sommer W (2014) Oculomotor control, brain potentials, and timelines of word recognition during natural reading. In: Current trends in eye tracking research. Springer. p. 141–155.
Kliegl R, Nuthmann A, Engbert R (2006) Tracking the mind during reading: The influence of past, present, and future words on fixation durations. J Exp Psychol: General 135(1):12.
Kliegl R, Wei P, Dambacher M, Yan M, Zhou X (2011) Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. Front Psychol 1:238.
Klimesch W (1999) Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Rev 29(2–3):169–195.
König P, Plöchl M, Ossandón JP (2012) Combining eeg and eye tracking: Identification, characterization and correction of eye movement artifacts in electroencephalographic data. Biomed Eng/Biomedizinische Technik 57(SI-1-Track-F). 392–392.
Kornrumpf B, Dimigen O, Sommer W (2017) Lateralization of posterior alpha eeg reflects the distribution of spatial attention during saccadic reading. Psychophysiology 54(6):809–823.
Kretzschmar F, Bornkessel-Schlesewsky I, Schlesewsky M (2009) Parafoveal versus foveal n400s dissociate spreading activation from contextual fit. NeuroReport 20(18):1613–1618.
Kuperman V, Dambacher M, Nuthmann A, Kliegl R (2010) The effect of word position on eye-movements in sentence and paragraph reading. Quart J Exp Psychol 63(9):1838–1857.
Kutas M, Federmeier KD (2011) Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp). Ann Rev Psychol 62:621–647.
Kutas M, Hillyard S (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427):203–205.
Lau EF, Phillips C, Poeppel D (2008) A cortical network for semantics:(de) constructing the n400. Nat Rev Neurosci 9 (12):920–933.
Lewis AG, Bastiaansen M (2015) A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. Cortex 68:155–168.
Li N, Wang S, Kornrumpf F, Sommer W, Dimigen O (2022) Parafoveal and foveal n400 effects in natural sentence reading: Evidence from overlap-corrected fixation-related potentials. bioRxiv:12.
Madsen J, Júlio SU, Gucik PJ, Steinberg R, Parra LC (2021) Synchronized eye movements predict test scores in online video education. Proc Nat Acad Sci 118(5). e2016980118.
Makeig S, Debener S, Onton J, Delorme A (2004) Mining event-related brain dynamics. Trends Cognit Sci 8(5):204–210.
Maris E, Oostenveld R (2007) Nonparametric statistical testing of eeg-and meg-data. J Neurosci Methods 164(1):177–190.
Metzner P, von der Malsburg T, Vasishth S, Rösler F (2015) Brain responses to world knowledge violations: A comparison of stimulus-and fixation-triggered event-related potentials and neural oscillations. J Cognit Neurosci 27(5):1017–1028.
Metzner P, Von Der Malsburg T, Vasishth S, Rösler F (2017) The importance of reading naturally: Evidence from combined recordings of eye movements and electric brain potentials. Cognit Sci 41:1232–1263.
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.
Parkes LM, Bastiaansen MC, Norris DG (2006) Combining eeg and fmri to investigate the post-movement beta rebound. Neuroimage 29(3):685–696.
Pelli DG (1997) The videotoolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision 10(4):437–442.
Pickering MJ, Garrod S (2007) Do people use language production to make predictions during comprehension? Trends Cognit Sci 11 (3):105–110.
Plöchl M, Ossandón JP, König P (2012) Combining eeg and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. Front Human Neurosci 6:278.
Rayner K (1998) Eye movements in reading and information processing: 20 years of research. Psychol Bull 124(3):372.
Rommers J, Dickson DS, Norton JJ, Wlotko EW, Federmeier KD (2017) Alpha and theta band dynamics related to sentential constraint and word expectancy. Language Cognit Neurosci 32 (5):576–589.
Sauseng P, Griesmayr B, Freunberger R, Klimesch W (2010) Control mechanisms in working memory: a possible function of eeg theta oscillations. Neurosci Biobeh Rev 34(7):1015–1022.
Sereno SC, Rayner K (2003) Measuring word recognition in reading: eye movements and event-related potentials. Trends Cognit Sci 7 (11):489–493.
van Gaal S, De Lange FP, Cohen MX (2012) The role of consciousness in cognitive control and decision making. Front Human Neurosci 6:121.
Vespignani F, Canal P, Molinaro N, Fonda S, Cacciari C (2010) Predictive mechanisms in idiom comprehension. J Cognit Neurosci 22(8):1682–1700.

Vignali L, Himmelstoss NA, Hawelka S, Richlan F, Hutzler F (2016) Oscillatory brain dynamics during sentence reading: a fixation-related spectral perturbation analysis. Front Human Neurosci 10:191.

von Stein A, Sarnthein J (2000) Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. Int J Psychophysiol 38(3):301–313.

Yagi A (1981) Visual signal detection and lambda responses. Electroencephalog Clin Neurophysiol 52(6):604–610.

Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY (2008) Transient induced gamma-band response in eeg as a manifestation of miniature saccades. Neuron 58(3):429–441.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.neuroscience.2023.03.024.

(Received 4 August 2022, Accepted 22 March 2023)
(Available online 31 March 2023)