Contents lists available at ScienceDirect

Neuropsychologia

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

Combining EEG and 3D-eye-tracking to study the prediction of upcoming speech in naturalistic virtual environments: A proof of principle

Eleanor Huizeling^{a,*}, Phillip M. Alday^b, David Peeters^c, Peter Hagoort^{a,d}

^a Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

^b Beacon Biosignals Inc., Boston, United States

^c Department of Communication and Cognition, TiCC, Tilburg University, Tilburg, the Netherlands

^d Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands

ARTICLE INFO

Keywords: Visual world paradigm Language comprehension Language prediction EEG Eye-tracking Virtual reality

ABSTRACT

EEG and eye-tracking provide complementary information when investigating language comprehension. Evidence that speech processing may be facilitated by speech prediction comes from the observation that a listener's eye gaze moves towards a referent before it is mentioned if the remainder of the spoken sentence is predictable. However, changes to the trajectory of anticipatory fixations could result from a change in prediction or an attention shift. Conversely, N400 amplitudes and concurrent spectral power provide information about the ease of word processing the moment the word is perceived. In a proof-of-principle investigation, we combined EEG and eve-tracking to study linguistic prediction in naturalistic, virtual environments. We observed increased processing, reflected in theta band power, either during verb processing - when the verb was predictive of the noun - or during noun processing - when the verb was not predictive of the noun. Alpha power was higher in response to the predictive verb and unpredictable nouns. We replicated typical effects of noun congruence but not predictability on the N400 in response to the noun. Thus, the rich visual context that accompanied speech in virtual reality influenced language processing compared to previous reports, where the visual context may have facilitated processing of unpredictable nouns. Finally, anticipatory fixations were predictive of spectral power during noun processing and the length of time fixating the target could be predicted by spectral power at verb onset, conditional on the object having been fixated. Overall, we show that combining EEG and eye-tracking provides a promising new method to answer novel research questions about the prediction of upcoming linguistic input, for example, regarding the role of extralinguistic cues in prediction during language comprehension.

1. Introduction

Listeners can process spoken language incredibly quickly. One mechanism that is thought to help with such fast processing is the prediction of upcoming linguistic input. Eye-tracking studies, such as a long tradition of work using the visual world paradigm (VWP), have been fundamental in showing that listeners can use visual and linguistic constraints to predict upcoming referents *prior* to them being mentioned (Allopenna et al., 1998; Altmann and Kamide, 1999). However, a change to the trajectory of eye movements can be difficult to interpret (Huettig et al., 2011), as it could indeed reflect a shift in processing linguistic or semantic input, but also, alternatively, a mere shift in visual attention, without a change in linguistic processing. More specifically, in regards to the prediction of upcoming words, it is possible for eye movement patterns to differ, but the underlying prediction to remain the same. Eye-tracking can reliably tell us that listeners predict upcoming information, but the absence of anticipatory looks cannot with full certainty be taken to mean that there is no prediction. In contrast, while event-related potential (ERP) amplitudes, time-locked to the onset of the predicted word, provide limited information about whether a word has been predicted *before* its onset, they can provide information about the ease of word processing the moment the word is perceived. Combining these measures within a single study and single analysis could therefore help to answer more nuanced theoretical questions about language comprehension (Knoeferle, 2015) and the types of cues used to inform predictions.

* Corresponding author. Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525XD, Nijmegen, the Netherlands. *E-mail address:* eleanor.huizeling@mpi.nl (E. Huizeling).

https://doi.org/10.1016/j.neuropsychologia.2023.108730

Received 31 May 2023; Received in revised form 15 September 2023; Accepted 3 November 2023 Available online 6 November 2023







^{0028-3932/© 2023} The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Listeners may rely on both linguistic and extralinguistic cues to anticipate upcoming speech content. The extent that predictions are informed by extralinguistic cues is unclear. For example, the speaker's facial expressions, eye gaze and gestures could provide useful information about upcoming and concomitant spoken input (Holler and Levinson, 2019; Perniss, 2018; ter Bekke, Drijvers and Holler, 2020). Moreover, disfluencies in speech (filled hesitations, repairs, silent pauses) are thought to provide information about the certainty of the speaker, or the ease of producing the upcoming word (Bortfeld et al., 2001; Brennan and Williams, 1995; Fraundorf and Watson, 2014; Schachter et al., 1991; Smith and Clark, 1993). Listeners have been shown to predict that the speaker will utter less frequent words or discourse-new information after producing a disfluency, rather than highly frequent or discourse-given information (Arnold et al., 2003; Arnold and Tanenhaus, 2011; Arnold et al., 2004). In a recent virtual reality (VR) study, Huizeling et al. (2022) observed a reduced proportion of fixations towards a predictable referent after the constraining verb was followed by a hesitation (e.g. "cutting down uhh the tree"), compared to if the sentence was spoken fluently. Instead, participants' eve gaze moved towards the virtual speaker upon hearing the hesitation. It is unclear whether this shift in eye gaze signifies the listener losing confidence in or even abandoning their prediction altogether. Alternatively, this shift in gaze may indicate a shift in visual attention towards the speaker, either to search for meaningful visual bodily signals to aid in language comprehension or in response to the saliency of the disrupted flow of speech.

Measuring ERPs, such as the N400, in response to the predicted word onset could help to better understand the role of extralinguistic cues, like disfluencies and gestures, on predictive language processing. The N400 is a negative ERP that typically occurs around 400ms after stimulus onset, with a centroparietal topography. It is thought to reflect the amount of new information to be integrated with the prior semantic context (Brothers and Kuperberg, 2021; Hagoort et al., 2009; Hodapp and Rabovsky, 2021; Kutas and Federmeier, 2011; Nieuwland et al., 2020). The N400 amplitude is larger (more negative) when a word is more difficult to integrate with the prior sentential context. A larger N400 amplitude is therefore observed if a word is semantically incongruent with the prior context of the sentence (Hagoort et al., 2004; Kutas and Hillyard, 1980; Nieuwland and van Berkum, 2006). Moreover, as the context provided by the sentence becomes more constraining and upcoming words become more predictable, the N400 amplitude decreases (Dambacher et al., 2006; Federmeier and Kutas, 1999; Kutas and Hillyard, 1984; Terporten et al., 2019). A larger N400 amplitude is also observed when listeners are presented with an image that is incongruent with the coinciding linguistic input (Federmeier and Kutas, 2001; Ganis et al., 1996; Knoeferle et al., 2011; Özyürek et al., 2007; Sitnikova et al., 2008; Tromp et al., 2018; Willems et al., 2008a,b), however, with a more anterior topography. If disfluencies in speech cause listeners to lose confidence in or disregard their prediction, then the N400 amplitude in response to predictable nouns might be expected to be larger if the noun is preceded by a disfluency, relative to if the sentence was spoken fluently (Corley et al., 2007; MacGregor et al., 2010). Before such hypotheses can be tested in more naturalistic virtual environments, we first need to test whether effects of sentence predictability can be measured in the EEG signal while participants are engaged in VR (Tromp et al., 2018).

The event-related signal in a post-N400 time window (e.g. 600–1000ms) has also been found to be sensitive to the constraints employed by the sentence context. An enhanced late *posterior* positivity has been found in response to semantic information that is incongruent with the preceding context, whereas an enhanced late *anterior* positivity has been associated with lexical items that are plausible but unexpected in highly constraining contexts (DeLong et al., 2014; Federmeier et al., 2007; Quante et al., 2018; Thornhill and Van Petten, 2012; Van Petten and Luka, 2012). Where the anterior positivity is thought to reflect either updating, inhibiting an incorrect lexical prediction, or integrating

unexpected semantic information with the prior context, the posterior positivity is thought to reflect the detection of conflict and a failure to integrate new information into the preceding context, leading to a reanalysis of the sentence context and altered expectations about the upcoming content (Brothers et al., 2020; Kuperberg et al., 2020; Rommers and Federmeier, 2018; Van Petten and Luka, 2012). However, it should be noted that evidence that differential anterior and posterior post-N400 effects map onto different functions of language processing has so far been acquired from laboratory-based paradigms. It is not clear to what extent findings generalise to more natural language processing.

ERPs, such as the N400 and the post-N400 positivity, could provide useful information regarding the influence of sentence context constraints on word processing the moment the word is perceived and could provide information about the processing consequences of disfluencies. The brain's response to sentence predictability has also been observed in ongoing neurophysiological oscillatory activity. Unlike ERPs, oscillations are sensitive to effects that are not phase-locked to a stimulus onset. This makes them particularly useful to study on-line predictions, which may form gradually with the build-up of context. Prystauka and Lewis (2019) provide an informative review of oscillatory dynamics of sentence comprehension and anticipatory processing.

There is increasing evidence that alpha (8-12 Hz) and beta (13-30 Hz) power are reduced when the context of a sentence is highly constrained. Broadly, reductions in alpha and beta band power are associated with increased cognitive processing in the cortex, increased attention, memory encoding and memory retrieval (Foxe et al., 1998; Jensen and Mazaheri, 2010; Klimesch et al., 1994). Increased alpha power is thought to reflect the suppression of cognitive processing in the cortex and inhibition (Klimesch et al., 2007; Strauß et al., 2014a,b). Lower alpha and beta power have been observed directly prior to a target word when the target word is predictable compared to unpredictable (Gastaldon et al., 2020; Leon-Cabrera et al., 2022; Molinaro et al., 2017; Piai et al., 2014; Rommers et al., 2017; Roos and Piai, 2020; Terporten et al., 2019; Wang et al., 2018) and in response to highly constraining compared to unconstraining information (Li et al., 2017). Moreover, lower pre-stimulus alpha power has been linked to increased effects of word congruence on N400 amplitudes (Lago et al., 2023). However, alpha/beta power modulation does not seem to be linearly modulated by constraint. Terporten et al. (2019) found the greatest alpha desynchronisation prior to the critical noun in moderately constraining sentences, followed by highly and then weakly constraining sentences. Unexpected words and incongruent visual information have been shown to result in lower alpha and beta power compared to expected words or congruent visual information (Rommers et al., 2017; Wang et al., 2012; Willems et al., 2008a). In contrast, two studies have found no effect of word predictability on alpha or beta power after word onset when the word was consistent with the sentence context (Rommers and Federmeier, 2018; Terporten et al., 2019).

Oscillations in the theta band have been associated with a range of cognitive functions, and tend to increase with increased cognitive processing (Cavanagh and Cohen, 2022; Cavanagh and Frank, 2014; Demiralp and Başar, 1992; Klimesch et al., 1994). Within the language literature, increased theta power has been related to memory retrieval during sentence comprehension and when hearing repeated names or pronouns (Coopmans and Nieuwland, 2020; Heine et al., 2006; Meyer et al., 2015). Less attention has been devoted to reporting effects of predictability in the theta frequency band with EEG data. Prior to the critical noun onset, in response to the constraining information, Li et al. (2017) found no significant differences in theta frequency in response to high compared to low constraining verbs in written Chinese sentences. Higher theta power has been observed, however, in response to unexpected words compared to expected words, for both semantically plausible and anomalous continuations (Bastiaansen and Hagoort, 2015; Li et al., 2017; Rommers et al., 2017; Willems et al., 2008a). It is thought that increased theta power reflects the increased effort to retrieve less predictable words from long term memory.

The large majority of studies discussed above have presented visual and/or auditory stimuli to individual participants in restricted laboratory set-ups that relatively poorly resemble the rich and dynamic situations of language processing in everyday life. An emerging realisation in psycholinguistic research is to move away from artificial laboratory experiments towards more ecologically valid, naturalistic paradigms, where language can be embedded in an enriched context (Hasson et al., 2018; Peeters, 2019). The rapidly growing VR industry means that it is now feasible to conduct psycholinguistic research in increasingly enriched naturalistic contexts, while maintaining the high level of experimental control afforded by laboratory experiments. In our recent work we have shown that people also make predictive eye movements in visually rich, naturalistic environments in the presence of a virtual speaker (Heyselaar et al., 2020; Huizeling et al., 2022). The presence of a virtual speaker shed new light on listeners' behaviour when they hear a disfluency in speech (Huizeling et al., 2022). Rather than the listener's eye gaze moving towards potential upcoming referents in the visual scene, participants looked towards the virtual speaker upon hearing a disfluency. This leaves an open question as to whether listeners lose confidence in their prediction in such instances, or wait for the sentence to become disambiguated. Obtaining the aforementioned eye gaze and electrophysiological measures concomitantly in naturalistic paradigms could provide information about the ease of word processing when listening to disfluent predictable sentences. However, it first needs to be established whether these two measures can be acquired in parallel while participants engage in VR.

Previously, combined EEG-eye-tracking experiments have been largely avoided due to eye movement artifacts, which can correlate with the variables of interest, contaminating the EEG data. One method to overcome this issue is to time-lock EEG analysis to the fixation onset to measure fixation related potentials (Dimigen et al., 2011). However, this is only possible with visually presented linguistic stimuli and not auditory stimuli. Recent advances in data processing have drastically improved the ability to correct for eye movement artifacts in the data (Dimigen, 2020). Although modern artifact correction techniques have been shown to be successful in computer-based reading and scene viewing paradigms, it is not clear to what extent such methods would be successful with EEG data recorded while participants are engaged in VR. Recording EEG in VR faces a number of additional challenges, such as muscle artifacts from head movements and potential interference from electrical noise. Before one can combine EEG and eve-tracking in naturalistic virtual environments to answer more nuanced research questions regarding the extent of extralinguistic cues on predictive processing, it is first vital to test to what extent it is possible to combine EEG, eye-tracking and VR to investigate predictive language processing.

1.1. The current study

The current study was a proof-of-principle investigation into the feasibility of simultaneously collecting EEG and eye-tracking data in VR to investigate the prediction of upcoming speech. EEG has previously been used successfully in VR settings both in our VR laboratory (Peeters, 2020; Raghavan et al., 2023; Tromp et al., 2018) and other laboratories (Badia et al., 2013). Recent studies have also begun to record EEG and eye-tracking simultaneously within free-viewing paradigms (Coco et al., 2020; Dimigen, 2020). Here we go one step further by combining eye-tracking and EEG in VR. We specifically wanted to test whether we could replicate effects of (a) the context constraint, and (b) the congruency of the sentence with the listener's predictions, on eye gaze and electrophysiological measures, including ERPs and frequency power, during a free-viewing VR paradigm. Robust effects of predictability and congruency on the EEG signal would pave the way for future studies to answer new theoretical questions regarding the influence of extralinguistic cues, such as disfluencies, on predictive processing.

To this end, we conceptually replicated Experiment 1 from Huizeling et al. (2022). In a 3D VWP, Huizeling et al. (2022) tracked participants'

eve gaze in VR while they listened to predictable and unpredictable sentences. Here we additionally simultaneously recorded EEG. Sentences (pre-recorded by a Dutch native speaker) were spoken by a virtual agent to the participant during a virtual tour of eight scenes (e.g. a street, a forest, a bathroom). Sentences were either predictable or unpredictable based on the constraints of the verb, where the verb in the sentence could either be constrained towards a single item in the scene, making the sentence conclusion predictable (e.g. sentence 1 in Section 2.5), or unconstrained, related to multiple items in the scene, making the sentence unpredictable (e.g. sentence 2 in Section 2.5). In addition, the noun in the sentence could either refer to an item visible in the scene, confirming the listener's prediction (congruent) or the referent could be absent from the scene, disconfirming the listener's prediction (incongruent). The noun in the sentence was always a plausible continuation of the sentence, but manipulated congruency given that predictions were formed by combining the linguistic context of the sentence with the visual context of the scene.

We hypothesised that, in a critical time window between verb and noun onset, there would be an increased proportion of fixations towards the target object in the predictable but not the unpredictable condition (Heyselaar et al., 2020; Huizeling et al., 2022).

Based on the aforementioned literature, we also hypothesised that, in response to unpredictable relative to predictable nouns and predictionincongruent relative to -congruent nouns, there would be a larger N400 response, a more positive anterior post-N400 effect, lower alpha and beta power, and higher theta power. These effects can be expected to result from the increase in processing required to suppress the predicted word and/or retrieve the unexpected/unpredicted word from long term memory. However, it should be noted that findings as to whether a late positivity is sensitive to mismatches between linguistic and visual information are mixed (Federmeier and Kutas, 2001; Willems et al., 2008b). In addition, it remains an open question as to what extent previous EEG findings actually generalise to more naturalistic and ecologically valid environments.

We additionally expected that, in the critical time window between verb onset and noun onset, there would be lower alpha and beta power in the predictable relative to unpredictable condition. Although there are a limited number of studies investigating the effect of constraining information on theta power, we hypothesised that we would either see no difference between predictable and unpredictable conditions (Li et al., 2017) or that there may be increased theta power in the predictable compared to unpredictable condition, given the increased retrieval of information from long term memory that is associated with making predictions.

Advancing technologies enable the on-line recording and coregistration of several data types at once. For example, EEG has been co-registered together with fMRI, MEG, eye-tracking and behavioural data, and eye-tracking has been co-registered with fMRI (Bonhage et al., 2015). However, these different data types are often analysed separately and only qualitatively compared. Prior research has used the EEG signal to predict RT data, as well as subsequent EEG effects (Alday & Kretzschmar, 2019; Lago et al., 2023; Maess et al., 2016). Here we quantitatively link EEG and eye-tracking data. First, we assessed to what extent the N400 amplitude and frequency power in response to the noun, which may quantify the ease of word processing, could be predicted from the proportion of anticipatory referent fixations prior to noun onset, which quantified noun predictability. We hypothesised that there would be a negative relationship between the proportion of anticipatory fixations prior to noun onset and the N400 amplitude in response to the noun. We additionally expected that anticipatory fixations may predict spectral power, consistent with expected effects of predictability on spectral power. Second, we investigated to what extent oscillatory activity after the constraining verb onset was associated with anticipatory fixations. We hypothesised that predictability-related modulations in alpha/beta and theta power would be associated with an increased proportion of target fixations. We additionally assessed to

what extent N400 amplitude at noun onset could be predicted by spectral power at verb onset.

2. Methods

2.1. Participants

Thirty-eight right-handed native Dutch speakers took part in exchange for a standard fee. Three participants were excluded due to poor EEG data quality, two were excluded due to a technical fault and one was excluded due to poor performance on the comprehension questions and distracted behaviour during the experiment, suggesting they were not paying attention during the experiment. Thirty-two participants remained for the analysis (21 female, 10 male, 1 unreported gender, median age = 23 years, age range = 18-34 years, SD = 4). Participants with epilepsy, uncorrected visual or hearing impairments, colourblindness, language impairments, dyslexia or developmental disorders were excluded from participation. The research was approved by Radboud University's Faculty of Social Sciences (ethics application ECSW-2020-046) and complied with the Declaration of Helsinki.

2.2. Cave system

Stimuli were presented in VR using a cave automatic virtual environment (CAVE) system (Cruz-Neira et al., 1992). An example of the set-up can be seen in Fig. 1. A detailed description of the CAVE environment has previously been described in Eichert et al. (2018). The CAVE system consisted of three 255×330 cm projector screens (VIS-CON GmbH, Neurkirchen-Vluyn, Germany) arranged at right angles. Two vertically displaced, overlapping displays were back-projected onto each projector screen via a mirror by two projectors (F50, Barco N.V., Kortrijk, Belgium).

The experiment was programmed and run in Python through 3D VR software Vizard (Floating Client 5.4, WorldViz LLC, Santa Barbara, CA). Audio was presented through four surrounding speakers (Logitech, US) that were located in the four bottom corners of the CAVE, plus one centred at the bottom of the middle screen.

2.3. Eye- and head-tracking

Eye movements were recorded with specialised eye-tracking shutter glasses, which both tracked the participant's eye movements and allowed them to see in 3D through a synchronised shutter mechanism for stereoscopic vision (SMI eye-tracking glasses 2 Wireless, SensoMotoric Instruments GmbH, Teltow, Germany). The interface for the eye tracking calibration and data recording were on a tablet, which transmitted data wirelessly to the tracking software. A camera on the glasses measured 60 Hz binocular recordings with automatic parallax compensation. The accuracy of gaze tracking was 0.5° along each dimension according to manufacturer reports. The eye tracker's latency of 60ms \pm 10ms was corrected for in the analysis.

Head movements were tracked through six passive reflective markers that were placed on the left and right side of the frames of the eye-tracking glasses. Ten infrared cameras (Bonita 10, Vicon Motion Systems Ltd, UK) tracked the position of the reflective markers. Four of the infrared cameras were distributed below the screens at six were distributed above the screens. Head-tracking data were recorded to an accuracy of 0.5 mm with Tracker 3 software (Vicon Motion Systems Ltd, UK) at a sampling rate of 250 frames per second and were continuously combined with the eye-tracking data to determine where the participant was looking in 3D space.

The accuracy of the eye-tracking was determined with a two-step calibration procedure. The first step used the SMI software's One Point Calibration procedure and assessed the accuracy of the eye-tracking alone (without the head-tracking). The participant was asked to look at a specific point in front of them (e.g. the corner of a piece of paper) and the experimenter corrected the gaze location in the software. The second step to the calibration procedure assessed the combined eye-and head-tracking, using an in-house calibration programme in VR, previously described by Eichert et al. (2018). Participants were presented with three spheres in 3D space, which each differed in position along the X, Y and Z axis. Participants were asked to look towards the centre of each sphere, during which the error between the estimated gaze position and the actual gaze position was corrected.

2.4. EEG acquisition

EEG data were continuously recorded with actiCAP EEG caps, which



Fig. 1. An example of the CAVE set up. Six of the ten infrared motion tracking cameras are visible, four above the three projector screens and two below. The scene displays the virtual agent in the street scene with six target objects (lampost, basketball, flag, tree, letterbox, and wheelbarrow).

contained 64 active Ag/AgCl electrodes, arranged in an equidistant montage. We used BrainAmp DC amplifiers (Brain Products, Gilching, Germany). The reference electrode was placed on the left mastoid. The ground electrode was placed on the forehead.

Four EOG electrodes recorded electric potentials caused by eye movements. Two electrodes were placed to the left of the left eye and to the right of the right eye to record horizontal eye movements. One electrode was placed above and one electrode was placed below the left eye to record vertical eye movements.

The data was recorded in BrainVision Recorder 1.2 (Brain Products, Gilching, Germany), at a sampling rate of 500 Hz and filtered online at 0.016–200 Hz. Impedance was kept below 10 k Ω (mean = 2.98, SD = 0.85) and measured with actiCAP 006 (Brain Products, Gilching, Germany).

2.5. Spoken sentence stimuli

Spoken sentence stimuli were adapted from Huizeling et al. (2022). Sentences consisted of 128 Dutch sentence sets that contained a subject-verb-object clause. All four sentences in a set were identical apart from the critical verb and the critical noun in the sentence. Sentences were predictable (50%) or unpredictable (50%) based on the verb constraints, where the verb in the sentence could either be related to a single object in the visual scene (predictive), or related to multiple objects in the scene (unpredictive). Verbs were matched for length (number of letters) and frequency, derived from SUBTLEX-NL (Keuleers et al., 2010). In 50% of sentences the noun referred to an object that was visible in the scene, *congruent* with the participant's prediction, whereas the other 50% of sentences the noun referred to an object absent from the scene, incongruent with the participant's prediction. The sentence always referred to a plausible object, but the constraints of the verb combined with the constraints of the visual scene determined congruency (see Fig. 1 for an example). Participants were presented with one of two sets of objects for any given scene, so that objects that were congruent/incongruent with the participant's prediction were counterbalanced. In other words, referents perceived as congruent in one version of the experiment were perceived as incongruent in the second version of the experiment and vice versa. As a result, there were four versions of each sentence (predictable, object 1; unpredictable, object 1; predictable, object 2; unpredictable, object 2), as shown below (English translations written underneath, with critical verbs and nouns presented in bold font). The full set of sentence stimuli can be found in the supplementary material.

1. Mijn buurman is niet zo goed in het sturen van de kruiwagen.

My neighbour is not very good at steering the wheelbarrow.

2. Mijn buurman is niet zo goed in het fixen van de kruiwagen.

My neighbour is not very good at fixing the wheelbarrow.

3. Mijn buurman is niet zo goed in het sturen van de skelter.

My neighbour is not very good at steering the go-kart.

4. Mijn buurman is niet zo goed in het fixen van de skelter.

My neighbour is not very good at fixing the go-kart.

Sentences were recorded by a trained native Dutch speaker with the aim to sound as natural as possible. During the experiment, sentences appeared to the participant to be spoken by a virtual agent, who matched the recorded voice in apparent age and ethnicity, with lip synchronisation and eye gaze towards the participant. At the beginning of each scene the virtual agent began with an opening sentence about the scene.

2.6. Visual stimuli

Visual stimuli consisted of eight scenes (e.g. street, bathroom, restaurant) in which six critical objects were embedded. Each of the eight scenes were presented twice during the experiment, with different objects (and hence different sentences) embedded on each presentation. Each scene was presented once, followed by a reiteration of all the scenes in the reverse order (e.g. the experiment always started and ended with the street scene). The order of the scenes remained consistent across experiments, but the order of the sentence-object sets was counterbalanced. Note that, due to counterbalancing objects across the condition of congruency (see Section 2.2, Spoken Stimuli), there were in fact four different object sets that were associated with each scene, but each participant only ever saw two of these sets.

For example, for 50% of participants the first street scene contained sentence-object set A (lamppost/traffic light, basketball/volleyball, flag/table cloth, tree/bush, letterbox/dustbin, and wheelbarrow/go-kart) and for 50% of participants the first street scene contained object set B (lolly/donut, hula hoop/marbles, traffic barrier/firework, umbrella/garden chair, balls/marbles, and bucket/watering can).

Objects were placed so as to appear as natural as possible within the scene in regards to scale and position. The virtual agent was positioned in the scene so as to be easy to locate, often towards the centre of the scene. Visual stimuli, including the scenes, the virtual agent and approximately half of the objects were the same as those used in (Huizeling et al., 2022).

In order to counterbalance the objects that would confirm vs disconfirm participants' predictions, for each sentence set there were two possible objects that could be displayed. For example, the scene presented in Fig. 1 could either display the wheelbarrow (as shown) or a go-kart. The former would mean that sentence 1 above (in Section 2.5) confirms the participant's prediction but sentence 3 disconfirms the participant's prediction but sentence 1 disconfirms the participant's prediction but sentence 1 disconfirms the participant's prediction.

2.7. Procedure

Participants listened to sentences spoken by a virtual agent during a virtual tour of eight scenes (e.g., an office, a living room, a canteen). The agent discussed her relation to each scene while participants' eye movements and EEG were continuously recorded. Participants were instructed that they would be taken on a tour through the virtual agent's life, that they only needed to listen to the virtual agent but that they would be asked questions about what she said at the end of the experiment to check that they were paying attention.

Sentence sets were separated into eight lists (4 sentence types \times 2 object sets) that participants were randomly assigned to, so that no participant heard more than one sentence from a set (see Section 2.5 and Section 2.6 for details).

During the experiment participants were seated on a chair in the centre of the CAVE, with the EEG and eye-tracking equipment placed on a table behind them. Participants were instructed that they could make small, gentle head movements, but they should not turn their head fully to the right or left, up or down. The experiment was divided into four blocks of four scenes, allowing the participant opportunities for self-paced breaks. In between each block the experimenter checked the calibration of the eye tracker using the 3D calibration step (see Section 2.6 Eye- and head-tracking).

After the experiment was complete, participants completed two questionnaires on LimeSurvey. The "Object questionnaire" presented participants with a list of all the objects that had been presented during the experiment and, for each object, participants selected either "Ja" (yes) or "Nee" (no) to indicate whether the virtual agent had referred to the object during the experiment. The "Verb questionnaire" presented participants with the list of critical verbs that the virtual agent had spoken within a sentence during the experiment (both the word form spoken in the experiment, and the full verb). Participants selected "Ja" (yes) or "Nee" (no) to indicate whether or not they knew the meaning of the verb. Data are available from the Max Planck Institute for Psycholinguistics' Language Archive (https://hdl.handle.net/1839/607d53a d-9a81-4891-856c-8bfcaff99a8b).

2.8. Data analysis

2.8.1. Eye-tracking

Eye-tracking data were analysed in R (version 4.1.2; R Core Team, 2021). Trials were time locked to 60ms after sentence onset to account for the eye tracker latency. To account for eye blinks or loss of eye tracker connectivity, samples were removed if eye-tracking coordinates along all three dimensions (X-Y-Z) remained the same for more than two samples. Trials were excluded from further analysis if >25% of samples were removed from within the critical time window. There was a maximum exclusion of nine trials per participant (mean = 2.72, SD = 2.22). Eye gaze on an object was considered a fixation if it exceeded 100ms.

In line with our previous work and with recommendations of analysing VWP data (Porretta et al., 2018) we analysed the data with generalized additive mixed effects models (GAMMs) using the bam function from mgcv (version 1.8-26; Wood, 2017) and using itsadug (version 2.3; Van Rij, Wieling, Baayen and Van Rijn, 2017) to interpret results. GAMMs extend the linear mixed-effects regression framework to include smooth curves and so can detect both linear and non-linear effects. Detailed discussions of the advantages of using GAMMs can be found in the previous literature (Porretta et al., 2018; Van Rij, Hendriks, Van Rijn, Baayen and Wood, 2019; Wieling, 2018; Winter and Wieling, 2016). In sum, GAMMs allowed us to model non-linear effects of continuous variables, which enabled us to enter time as a continuous predictor in our model to investigate whether target fixations changed dependent on time and condition. Moreover, as GAMMs can model both linear and nonlinear effects we were not required to make assumptions about the linearity of our data.

Data were analysed during a critical window of 200ms after verb onset until the mean noun onset to account for the time it takes to programme an eye-movement (Rayner et al., 1983). Binomial responses of whether the target object was fixated or not were entered into the model as a dependent variable, applying a logit link function. Sentence Predictability (predictable/unpredictable) was entered into the model as a parametric effect along with factor smooth interactions of Time \times Predictability, Time \times Sentence, and Time \times Subject. Predictability was coded with a deviation contrast-coding scheme (with contr. sum), which compares the mean proportion of fixations for each level of predictability with the overall mean across levels. To avoid overfitting, the parameter k (which limits the number of basis functions used to fit the model) was limited to five. To remain consistent with (Huizeling et al., 2022), Maximum Likelihood estimation was selected to smooth parameter estimates (Wieling, 2018). Segments of time with significant differences between conditions were estimated with the plot_diff function. Analysis scripts for modelling eye-tracking and EEG data are available from the Max Planck Institute for Psycholinguistics' Language Archive (https://hdl.handle.net/1839/607d53ad-9a81-4891-856c-8bf caff99a8b).

2.8.2. EEG

2.8.2.1. Preprocessing. Data were analysed in MATLAB toolbox fieldtrip-20200,831 (Oostenveld et al., 2011). Data were band pass filtered between 0.1 and 35 Hz, using a forward-backward zero-phase 4th order Butterworth filter with a Hamming window, and re-referenced to the average of the left and right mastoids. For the optimal detection of eye artifacts we followed the recommendations for free-viewing

paradigms outlined in Dimigen (2020). Specifically, the independent component analysis (ICA) training data (henceforth called ICA data) were band-pass filtered between 2 and 100 Hz using a forward-backward zero-phase finite impulse response window sync filter with a Hamming window. The ICA data was demeaned and line noise was removed using a discrete Fourier transform.

Both the ICA data and "true" data were then epoched from -1000-2000ms relative to verb and noun onset. Note that this meant that the baseline window of the noun often overlapped with the verb and thus differed between conditions. This was controlled for in the analysis by including the baseline and an interaction between baseline and condition as predictors in the model (see Section 2.8 ERPs and Spectral analysis). However, it meant that the ICA assumption of independent samples was violated. On visual inspection of the ICs, this did not appear to affect the performance of the ICA. Trials were then visually inspected and noisy trials were removed prior to running ICA (with fieldtrip's *runica*). Components were visually inspected and those resembling artifacts caused by eye blinks, horizontal eye movements and spike potentials were then removed from the main data. Noisy sensors were interpolated with the average signal of neighbouring sensors.

2.8.2.2. ERPs. For the analysis of the N400 in response to the critical noun, the EEG signal was averaged over time, from 300 to 500ms, and over space across centroparietal sensors, plotted in Fig. 4. Averaged data were entered into a linear mixed effects model, using lmer from lme4 (Bates et al., 2015), with baseline (centred), condition (predictable vs unpredictable/congruent vs incongruent) and a baseline \times condition interaction as covariates, and random intercepts for subject and item. The baseline was defined as the average amplitude in the 100ms prior to noun onset and was included in the model to control for fluctuations in pre-stimulus EEG amplitude (Alday, 2019). A model comparison procedure revealed no significant difference between a model containing random slopes of condition and random intercepts for subject and item $(\chi^2 = 6.10, df = 4, p > .10)$. The model was therefore simplified to include only random intercepts in order to maintain parsimony and computational efficiency (Bates, 2019; Matuschek et al., 2017). Separate models were used to investigate the effect of congruency (congruent vs incongruent) and the effect of predictability (predictable vs unpredictable). The degrees of freedom (df) of fixed effects were computed with Satterthwaite approximation using the lmerTest package (version 0.9-40; Kuznetsova et al., 2017).

For the analysis of the post-N400 time window, time-locked to the critical noun, a time window of 700-1700ms was selected to avoid overlap with the N400 itself. Data were down-sampled to 125 Hz. The paradigm used in the current study was different to those used in most previous studies that investigate the post-N400 time window and we had no clear hypothesis about the specific latency of expected effects. Data were therefore analysed with GAMMs (k limited to 10), allowing us to include time as a continuous predictor in the models (see Section 2.8 Data Analysis, Eve-tracking). GAMMs have been used to model EEG data in past research (Hendrix et al., 2017; Lago et al., 2023; Tremblay and Newman, 2015). Previous investigations of the post-N400 have found differential anterior and posterior effects (DeLong et al., 2014; Federmeier et al., 2007; Quante et al., 2018; Thornhill and Van Petten, 2012; Van Petten and Luka, 2012). Data were averaged over sensors encompassed in two regions of interest (ROI), including an anterior and a posterior ROI (see selected electrodes for each ROI plotted in Fig. 5). In addition to time (in steps of 8ms and centred) and ROI, condition, baseline (centred), and an interaction between condition and ROI were entered into the model as parametric effects, as well as smooth terms for time by condition and time by condition aggregated with ROI. Random smooths were entered for subject and item. Separate models were fit to investigate the effect of congruency (congruent vs incongruent) and the effect of predictability (predictable vs unpredictable). For each model, condition was coded with a deviation contrast-coding scheme (with

contr. sum).

2.8.2.3. Spectral analysis. Time frequency analysis was performed from -1000 - 1500ms relative to verb and noun onset, from 2 to 30 Hz for every 1 Hz, with a sliding window moving in stages of 50ms with four cycles per time window and applying a Hann taper (as specified by "Hanning" in fieldtrip).

A time window of 0–1000ms relative to verb/noun onset was selected for further investigation, as well as theta (4–6 Hz), alpha (7–11 Hz) and beta (14–30 Hz) frequency bands, within which frequency power was averaged for further analysis. The specific alpha range was selected based on visual inspection of Figure SM1, which displays a time frequency representation (TFR) averaged across all conditions and all electrodes (Luck and Gaspelin, 2017). As no specific frequency effects are visible in the theta or beta ranges in Figure SM1, standard theta and beta ranges of 4–6 Hz and 14–30 Hz, respectively, were selected, while avoiding overlap with the selected alpha band.

As there was no clear hypothesis about the latency at which effects may be observed in theta and alpha frequencies, we again kept time as a continuous predictor (in steps of 50ms) and analysed the data with GAMMs. Data were averaged across sensors within two ROIs, an anterior and a posterior ROI (see Fig. 8) and log transformed for the analysis. Posterior and anterior sensors were selected as ROIs in accordance with previous effects of context constraints on pre-stimulus alpha power recorded with EEG (Rommers et al., 2017). In addition to time (centred) and ROI, condition and a condition \times ROI interaction were entered into the model as parametric effects. Smooth terms for time by condition and time by condition aggregated with ROI were added to the model. Random smooths were entered for subject and item. Separate models were used to investigate the effect of congruency (congruent vs incongruent) and the effect of predictability (predictable vs unpredictable).

The baseline was not entered into models analysing oscillatory power, as the time-frequency decomposition does not suffer from the slow drifts in EEG data that is present in ERPs.

2.8.3. Predicting N400 amplitude and frequency power from anticipatory fixations

We hypothesised that the prediction of the referent before noun onset would influence the ease of noun processing and that this would be reflected in the EEG data. The proportion of fixations towards the referent prior to noun onset was taken as a metric of prediction. This differs from using the condition of predictability as a (binary) predictor, as there may be variability within conditions and across participants as to whether the referent was predicted or not.

To this end, N400 amplitude and log frequency power (alpha and theta, 0-1000ms) were exported from MATLAB, read into R with R. matlab (Bengtsson, 2017) and then entered into three separate linear mixed effects model (lmer from lme4) as dependent variables, with the time (number of samples) spent fixating the target as a predictor variable and random intercepts for subject and item. Random slopes of condition for subject and item were included in the frequency models but omitted from the N400 model because the resultant fit was singular (Bates, 2019). Baseline amplitude and ROI (anterior/posterior) were entered as additional predictor variables for the N400 and frequency models respectively. A model comparison procedure showed that adding an interaction between baseline amplitude and target fixations did not improve the model fit ($\chi^2 = 2.36$, df = 1, p > .10). Similarly, whereas including ROI as a predictor improved the model fit for both theta ($\chi^2 =$ 228.60, df = 1, p < .001) and alpha ($\chi^2 = 26.49$, df = 1, p < .001) models, adding an interaction between ROI and target fixations did not improve the model fit for either theta ($\chi^2 = 1.28$, df = 1, p > .10) or alpha ($\chi^2 = 0.03$, df = 1, p > .10). Nevertheless, we report the models containing the interaction in Table 8 in Section 3.3, as the theta model without the interaction failed to converge. This did not change the overall pattern of results and made only minor differences to the models'

estimates.

The time spent fixating the target in the time window between 490ms after the verb onset until mean noun onset was selected as a metric for predictability, consistent with the time window in which a significant effect of predictability on target fixations was found (see Section 3.1). Trials were restricted to those in which the referent was visible in the scene, confirming the participant's prediction, rather than those in which the referent was not present in the scene, as we were only interested in effects of noun predictability and not congruence of the referent with the prediction.

2.8.4. Predicting N400 amplitude from frequency power

We hypothesised that frequency power at verb onset would be predictive of N400 amplitude at noun onset. Frequency power (theta, 4-6 Hz; alpha 7-11 Hz) averaged over time (0-1000ms relative to verb onset) and electrodes within two ROIs (anterior and posterior; see Fig. 8) was centred and entered as a predictor in two linear mixed effects models with additional predictors of baseline N400 amplitude (centred), condition (predictable-congruent, unpredictable-congruent, predictable-incongruent, and unpredictable-incongruent) and condition \times anterior frequency power and condition \times posterior frequency power interactions. The models contained random intercepts for sentence and subject. Models containing additional random slopes did not converge. A model comparison procedure revealed no significant difference between models containing only random intercepts and models containing random slopes of frequency power for subject and item (p > .10). Separate models were fit for alpha and theta power. There was some collinearity between frequency power in the two ROIs (the generalized variance-inflation factors, GVIFs ^(1/(2*Df)), for the two respective interaction terms were alpha: 1.79 and 1.50; theta: 2.19 and 2.23) and so caution should be taken when making inferences about the topography of effects.

2.8.5. Predicting anticipatory fixations from frequency power

For anticipatory fixations to take place, the listener must first combine information from the constraining verb with the constraints of the visual scene. Anticipatory processing may therefore be reflected in the brain signal prior to fixations towards the predicted item. We therefore tested whether anticipatory fixations could be predicted by frequency power after verb onset, with the length of time spent fixating the referent between 490ms relative to verb onset until noun onset as the dependent variable.

Frequency power (theta, 4–6 Hz; alpha 7–11 Hz) averaged over time (0–1000ms relative to verb onset) and electrodes within two ROIs (anterior and posterior; see Fig. 8) was a predictor in the models. Due to the collinearity between the data within the two ROIs, fixation data were modelled as a function of frequency power in each ROI in separate models. Alpha and theta frequency bands were additionally run as separate models. Multiple comparisons of two tests for the two separate ROIs were corrected for with Bonferroni correction, adjusting the significance threshold to p < .025.

Due to the large number of trials in which the target was not fixated, there was a disproportionately large number of zeros in the data, which could not be accommodated by distributions typically used to model count data, such as Poisson or negative binomial distributions. The pattern of fixations can be separated into a) the decision to fixate on the target object and b) the length of time (number of samples) the item is fixated for during the critical time window. Such differences could reflect different underlying cognitive processes. We therefore implemented a hurdle model to model fixations. A hurdle model fits the model in two stages, thereby accounting for the large number of zeros. The first part of the model fits the probability of the target being fixated. The second part of the model describes the length of time (number of samples) the object was fixated, given that it was fixated, with a Poisson distribution.

To allow for modelling a nested random effects design, in order to

control for items and subjects, and for continuity across analyses, data were modelled using the *gam* function from the *mgcv* package, with a *ziplss* family. For both stages of each hurdle model, average frequency power was entered as a parametric predictor and as a smooth term. Random smooths were entered for subject and item. Here, GAMMS allowed the number of fixations to vary non-linearly as a function of frequency power, but could also detect linear effects if present. GAMMs have previously been used to co-register EEG and eye-tracking data (Kretzschmar et al., 2015; Nikolaev et al., 2016; Van Humbeeck, Meghanathan, Wagemans, van Leeuwen and Nikolaev, 2018).

3. Results

All participants achieved over 71% accuracy on the Object questionnaire (mean = 86.69%, SD = 6.43%), suggesting that all participants were attentive. For each verb, we calculated the percentage of participants who knew the meaning of the verb. Each individual verb was known by an average of 98.62% participants (SD = 5.71). No further analysis was conducted on the questionnaire data.

3.1. Eye gaze

Fig. 2 presents the proportion of fixations towards the target, distractors and virtual agent while listening to predictable and unpredictable sentences.

The model (see Table 1) revealed a parametric effect of condition, with a greater proportion of target fixations in the predictable compared to unpredictable condition. There was a significant smooth for time for predictable but not unpredictable sentences, which resulted from an increase in fixations over time in the predictable but not the unpredictable condition (see smooth plotted in Fig. 3).

Model estimated difference curves are presented in Fig. 3. As the model estimated the effect to be linear, it estimated the proportion of fixations to be greater in the unpredictable relative to predictable condition from the beginning of the analysis window (200ms after verb onset) until 380ms after verb onset and greater in the predictable relative to unpredictable condition from 480ms after the verb onset until mean noun onset (unadjusted for the time to programme a fixation).

3.2. EEG

3.2.1. N400

ERPs in response to noun onset, averaged over centroparietal electrodes, are presented for each condition in Fig. 4. The waveform of a single centroparietal electrode is presented in Figure SM2 in the



Figure SM3 in the supplem

A later, post-N400, component is thought to reflect ongoing processing after the word has been perceived, for example, through reanalysis or updating of the sentence interpretation, or integrating new

alysis or updating of the sentence interpretation, or integrating new semantic information to the prior context. Based on the aforementioned literature, we hypothesised that there would be a more positive post-N400 in response to unpredictable compared to predictable nouns and in response to incongruent relative to congruent nouns. To investigate amplitude modulations in a post-N400 time window, we used GAMMs to model the change in amplitude over time 700–1700ms after noun onset, in a posterior and an anterior ROI. ERP waveforms for the two ROIs are presented in Fig. 5 and the model summaries can be found in Table 4.

There was a significant parametric effect of predictability on post-N400 amplitude, as well as ROI and the baseline amplitude. Model estimated post-N400 amplitude was more positive in the predictable compared to unpredictable condition, contrary to our hypotheses (see Figure SM3 in the supplementary material for time windows of significant differences). There was a significant interaction between predictability and ROI. Inspection of the smooth terms revealed that the change in amplitude during the post-N400 time window reversed in direction for anterior and posterior ROIs but was qualitatively similar across predictable and unpredictable conditions (see Fig. 6).

We found a significant parametric effect of congruency on the post-N400 amplitude in response to the noun, as well as effects of ROI and the baseline amplitude. Consistent with our hypotheses, model estimated post-N400 amplitude was overall more positive in the incongruent compared to congruent condition. There was a significant interaction between congruency and ROI. The smooth terms presented in Fig. 7 revealed that the change in amplitude over time predominantly differed across conditions in the posterior ROI, where there was a decrease in amplitude over time in the congruent but not the incongruent condition (also shown in the model estimated difference curves presented in Figure SM4 in the supplementary material).

In summary, we found effects of both predictability and congruence on post-N400 components, where amplitude was overall greater in the predictable and incongruent conditions relative to the unpredictable and congruent conditions, respectively. The effect of time on EEG amplitude was qualitatively similar across predictable and unpredictable conditions. In contrast, an effect of congruency over time was predominantly seen in the posterior ROI, where amplitude further decreased over time in the congruent but not the incongruent condition.

Fig. 2. Proportion of target fixations. Vertical lines represent critical time points (mean verb onset and mean noun onset). Shaded ribbons display the standard error of the mean. Zero indicates 500ms prior to verb onset.

Table 1

Model summary for target fixations in Predictable vs Unpredictable conditions after verb onset.

Parametric coefficients	Estimate	SE	Z value	Р
Intercept Condition Smooth terms	-3.28 0.13	0.13 0.01 Pof. df	-25.60 14.19 Chi sa	<.001 <.001 P
Smooth for Time: Predictable Smooth for Time: Unpredictable Random effect for Subjects Random effect for Sentence	1.01 1.00 147.08 598.13	1.02 1.01 287.00 1151.00	60.83 2.48 3099.43 8520.55	P <.001 .116 <.001 <.001

edf, effective degrees of freedom; Ref. df, reference degrees of freedom; SE, standard error. Deviation contrast-coding: Predictable (1); Unpredictable (-1).

supplementary material. To investigate effects of predictability and congruence on the N400 amplitude, two linear mixed effects models were conducted on the average data from a group of centroparietal electrodes between 300 and 500ms (see Fig. 4 panel A for a schematic of the selected electrodes). The model (see Table 3) revealed a significant effect of congruency on the N400 amplitude, in which the N400 timelocked to noun onset was larger for prediction-incongruent compared to congruent nouns. However, contrary to our hypotheses, there was no significant effect of predictability on N400 amplitude (see Table 2).

3.2.2. Post-N400

8



Fig. 3. Model estimated smooths and difference curve for target fixations: Model estimated smooths for the proportion of fixations in predictable (red) and unpredictable (teal) sentences (panel A) and the difference between the model-estimated smooth splines of the predictable and unpredictable conditions (panel B). Red dashed lines mark windows of significant differences. Time is relative to verb onset.



Fig. 4. ERP amplitudes (μ V) in response to noun onset averaged over centroparietal sensors. A baseline of 100ms prior to the noun onset was subtracted from the amplitude. Topographical plots display the difference in amplitude (congruent – incongruent left, predictable – unpredictable right) averaged between 300 and 500ms after noun onset. Time (s) is relative to noun onset. The grey shaded area highlights the time segment that was averaged over and entered into the linear model. **p < .01, ns p > .10.

3.2.3. Frequency

3.2.3.1. Verbs. TFRs and topographical plots presenting frequency power in the time window following verb onset are presented in Fig. 8. To investigate theta power modulations in response to (un)predictive verbs, we used GAMMs to model the change in theta power over time 0–1000ms after critical verb onset in a posterior and an anterior ROI. Effects in theta power in response to constraining information have not often been reported in the previous literature. In one study, Li et al. (2017) reported no significant difference in theta power in response to constraining and unconstraining verbs in written Chinese sentences.

We found a significant parametric effect of verb predictivity on theta power after verb onset, where theta power was higher in response to predictive relative to unpredictive verbs. There was a significant effect of ROI, with higher theta power in the anterior relative to posterior ROI (see also TFR in Fig. 8 panel A), but no significant interaction between ROI and condition.

Significant smooth terms show a linear (edf is close to 1; see Table 5)

change in theta power over time for predictive verbs in the anterior ROI. Theta power was greater in response to predictive compared to unpredictive verbs from 650ms (see differences between predictive and unpredictive smooth terms presented in Fig. 9 panel A and B). There was also a significant smooth for time for the unpredictive verbs in the anterior ROI, however, this was close to the threshold of significance (p = .043). There were no other significant smooths for time.

It was hypothesised that there would be lower alpha and beta power in response to predictive compared to unpredictive verbs (Gastaldon et al., 2020; Leon-Cabrera et al., 2022; Li et al., 2017; Molinaro et al., 2017; Piai et al., 2014; Rommers et al., 2017; Roos and Piai, 2020; Terporten et al., 2019; Wang et al., 2018). The model summary for alpha power can be found in Table 5. Only tentative effects of verb predictivity were found in beta power. Results for beta power can be found in the supplementary material (see pable SM1).

There was a significant parametric effect of verb predictivity on alpha power after verb onset, as well as a significant effect of ROI and a significant interaction between ROI and condition. Contrary to our



Fig. 5. ERP amplitudes (μ V) for frontal (panel A) and posterior (panel B) electrodes. The electrodes that were selected are presented in the schematic in the top right of each panel. A baseline of 100ms prior to the noun onset was subtracted from the amplitude. Topographical plots present the difference between conditions (predictable congruent – predictable incongruent, predictable congruent – unpredictable congruent). The grey shaded area highlights the time segment that was averaged over and entered into the linear model.

Model summary for N400 amplitude in response to Predictable vs Unpredictable nouns.

Random effects:					
Groups	Name	Variance	SD		
Sentence	(Intercept)	0.79	0.89		
Subject	(Intercept)	1.34	1.16		
Residual		42.99	6.56		
Fixed effects:					
	Estimate	SE	df	t	Pr(> t)
Intercept	-1.43	0.27	36.90	-5.39	<.001
Predictability	0.19	0.15	1762.06	1.27	.205
Baseline	0.06	0.02	1883.74	2.84	.005
$\textbf{Condition} \times \textbf{Baseline}$	-0.06	0.02	1878.95	-3.11	.002

df, degrees of freedom; Estimate, beta coefficient; SD; standard deviation; SE, standard error.

Table 3

Model summary for N400 amplitude in response to Congruent vs Incongruent nouns.

Random effects:					
Groups	Name	Variance	SD		
Sentence	(Intercept)	0.47	0.69		
Subject	(Intercept)	1.11	1.05		
Residual		38.96	6.24		
Fixed effects:					
	Estimate	SE	df	t	Pr(> t)
Intercept	-1.91	0.24	34.68	-7.86	<.001
Congruence	0.67	0.14	1775.70	4.66	<.001
Baseline	0.04	0.02	1887.17	2.13	.033
$Condition \times Baseline \\$	-0.05	0.02	1861.81	-2.36	.018

For abbreviations, see legend of Table 2.

Model summary for post-N400 in response to Predictable vs Unpredictable and Congruent vs Incongruent nouns.

	Predictability			Congruency				
Parametric coefficients	В	SE	t	Р	В	SE	t	Р
Intercept	0.50	0.12	4.23	<.001	0.52	0.14	3.67	<.001
Condition	0.06	0.01	5.84	<.001	-0.03	0.01	-3.19	.001
ROI	-0.30	0.01	-28.47	<.001	-0.39	0.01	-37.18	<.001
Baseline	-0.34	< 0.01	-196.25	<.001	-0.32	< 0.01	-178.60	<.001
$\textbf{Condition} \times \textbf{ROI}$	0.10	0.01	9.77	<.001	0.19	0.01	17.90	<.001
Smooth terms	edf	Ref. df	F	Р	edf	Ref. df	F	Р
S: Cong/Pred	1.04	1.05	0.26	.624	1.15	1.19	0.95	.327
S: Incong/Unpred	1.41	1.53	0.32	.780	1.42	1.53	0.05	.958
S: Cong/Pred A	6.21	7.19	9.72	<.001	6.43	7.35	8.28	<.001
S: Cong/Pred P	1.05	1.06	2.87	.089	1.17	1.22	4.25	.033
S: Incon/Unpred A	5.68	6.75	8.18	<.001	3.74	4.53	3.38	.010
S: Incon/Unpred P	3.29	4.00	4.29	.002	4.86	5.95	6.64	<.001
RS: subject	230.94	287.00	21.92	<.001	240.72	287.00	25.65	<.001
RS: sentence	915.34	1151.00	12.02	<.001	945.63	1151.00	13.53	<.001

A, anterior ROI; P, posterior ROI; Pred, predictable condition; Unpred, unpredictable condition; Cong, congruent; Incong, incongruent; B, beta estimate; edf, effective degrees of freedom; RS, random smooth for time; Ref. df, reference degrees of freedom; SE, standard error; S, smooth for time. Deviation contrast-coding: Predictable (1); Unpredictable (-1).



Fig. 6. Model estimated smooths for amplitude in response to predictable and unpredictable nouns in anterior (panel A) and posterior (panel B) ROIs. Time is relative to noun onset.



Fig. 7. Noun congruency model estimated smooths for amplitude in response to congruent and incongruent nouns in anterior (panel C) and posterior (panel D) ROIs. Time is relative to noun onset.



Fig. 8. Panel A: Verbs' TFRs for the time window following predictive and unpredictive verbs. Frequency power (μV^2 /Hz for 4–30 Hz) plotted for predictive and unpredictive conditions represents the absolute difference from baseline (-100-0ms). No baseline correction was applied when plotting the difference between conditions. The vertical black line overlaying TFRs marks the verb onset. The variable black line overlaying TFRs represents the proportion of fixations towards the target object. Panel B: Topographical plots of differences in frequency power 0–1000 ms after verb onset for theta (4–6 Hz), alpha (7–11 Hz) and beta (14–30 Hz) frequency bands. No baseline correction was applied. Significance levels are based on parametric effects: ***p < .001, *ns*: not significant.

hypotheses, alpha power was higher in response to predictive compared to unpredictive verbs, specifically in the anterior ROI throughout the entire analysis window (see estimated differences presented in Fig. 9 panels C–D).

There was a linear increase in alpha power over time in the anterior ROI for the predictive condition only, reflected in a significant smooth for time for this condition (see Table 5). There were no other significant smooths for time.

3.2.3.2. Nouns. TFRs and topographical plots presenting frequency power in the time window following noun onset are presented in Fig. 10. Greater theta responses to both unpredictable and incongruent nouns relative to predictable and congruent nouns were expected (Bastiaansen and Hagoort, 2015; Li et al., 2017; Rommers et al., 2017; Willems et al., 2008a), reflecting an increased effort to retrieve lexical and semantic information from long term memory.

There was a significant parametric effect of predictability and ROI on theta power after noun onset, where theta power was lower in response to predictable compared to unpredictable nouns. There was no significant interaction between condition and ROI.

There was a significant smooth for time for unpredictable nouns, which stemmed from the posterior ROI (see Table 6). Accordingly, a peak in theta power in response to unpredictable nouns can be seen in the posterior ROI's TFR in Fig. 10 and the smooths in Fig. 11 panel F, peaking at around 600–700ms after noun onset. From the model estimated difference curves presented in Fig. 11 panel A–B, it can be seen that theta was greater for unpredictable relative to predictable nouns from 220 to 880ms in the posterior ROI and from 250 to 860ms in the anterior ROI. There were no other significant smooths for time (p > .10).

Supporting our hypotheses, there was a significant parametric effect of congruency on theta power, where theta power was higher in response to incongruent compared to congruent nouns (see Table 7). There was a significant effect of ROI, with higher theta power in the anterior compared to posterior ROI, and a significant interaction between condition and ROI.

Theta power significantly changed over time in the posterior ROI in response to incongruent nouns only (see smooth terms in Table 7). In both posterior and anterior ROIs theta was higher in the incongruent relative to congruent nouns from around 400ms (340ms and 420ms for posterior and anterior ROIs respectively), as can be seen in the model estimated difference curves presented in Fig. 12 panel A and B. Differences peaked at around 600ms after noun onset.

We expected lower alpha and beta power for incongruent compared to congruent nouns (Rommers et al., 2017; Wang et al., 2012; Willems et al., 2008a). On the other hand, we had no clear evidence to expect a difference in alpha or beta power between predictable and unpredictable nouns (Rommers and Federmeier, 2018; Terporten et al., 2019). Only tentative effects of noun predictability and congruence were found in beta power. Results for beta power can be found in the supplementary material (see Tables SM2 and SM3).

There was a significant parametric effect of noun predictability and ROI on alpha power in response to noun onset, but no interaction between predictability and ROI (see Table 6). Alpha power was lower in response to predictable compared to unpredictable nouns, around 200–720ms relative to noun onset (see model estimated difference curves in Fig. 11 panels C and D).

There was a significant smooth for time for unpredictable nouns when collapsed across ROIs, suggesting that alpha power changed over time in this condition. However, when separated across ROIs the smooths for time did not reach significance (p = .089).

There were significant parametric effects of congruence and ROI on alpha power, but no significant interaction between condition and ROI. Alpha power was higher in the incongruent relative to congruent nouns from 350 to 1000ms in the anterior ROI and between 410 and 1000ms in the posterior ROI (see model estimated differences in Fig. 12). The smooth for time for alpha power in the congruent condition did not reach significance (p = .057). There were no other significant smooths for time (p > .10).

Summary of frequency results.

In summary, effects of verb predictivity and noun predictability were observed most robustly in theta and alpha power. Higher theta and alpha power were both seen either after hearing a predictive verb (relative to unpredictive) or an unpredictable or incongruent noun (relative to predictable or incongruent respectively).

3.3. Combined analysis

3.3.1. Predicting N400 amplitude from anticipatory fixations

We hypothesised that the prediction of the referent before noun onset would influence the ease of processing once the noun was perceived. In the analysis of the eye-tracking data (see Section 3.1), an effect of predictability on target fixations was found from 490ms after the verb onset. We therefore investigated to what extent the proportion of fixations from 490ms post-verb onset until mean noun onset could

Model summary for theta and alpha power in response to Predictive vs Unpredictive verbs.

	Theta				Alpha				
Parametric coefficients	В	SE	t	Р	В	SE	t	Р	
Intercept	0.79	0.07	10.93	<.001	1.14	0.10	12.02	<.001	
Condition	0.01	< 0.01	3.36	.001	0.01	< 0.01	4.35	<.001	
ROI	0.11	< 0.01	65.38	<.001	-0.06	< 0.01	-31.50	<.001	
$\textbf{Condition} \times \textbf{ROI}$	< 0.01	< 0.01	0.14	.889	<0.01	< 0.01	2.50	.013	
Smooth terms	edf	Ref. df	F	P	edf	Ref. df	F	Р	
S: Pred	1.00	1.01	26.95	<.001	1.00	1.01	1.73	.189	
S: Unpred	2.44	2.98	2.24	.080	1.00	1.01	1.23	.269	
S: Pred Ant	1.01	1.01	7.58	.006	1.02	1.03	8.26	.004	
S: Pred Post	< 0.01	< 0.01	< 0.01	.999	< 0.01	0.01	0.01	.993	
S: Unpred Ant	1.36	1.92	2.93	.043	1.01	1.01	0.56	.452	
S: Unpred Post	1.02	1.03	0.19	.670	< 0.01	< 0.01	< 0.01	.999	
RS for subject	140.40	287.00	219.73	<.001	130.00	287.00	285.90	<.001	
RS for sentence	504.70	1151.00	3.09	<.001	478.50	1151.00	2.70	<.001	

Ant, anterior ROI; Post, posterior ROI; Pred, predictive condition; Unpred, unpredictive condition; B, beta estimate; edf, effective degrees of freedom; RS, random smooth for time; Ref. df, reference degrees of freedom; SE, standard error; S, smooth for time. Deviation contrast-coding: Predictive (1); Unpredictive (-1).



Fig. 9. Verb modulated frequency power difference curves: The difference between the model-estimated smooth splines of the predictive and unpredictive conditions for log theta (panels A_B) and alpha (panels C_D) power, in anterior (panels A and C) and posterior (panels B and D) ROIs in the 1000 ms following verb onset. Red dashed lines mark windows of significant differences.

predict N400 amplitude in response to the noun. As we were only interested in sentence predictability and not congruence, for this analysis we only included trials where the referent was visible in the scene, confirming the participant's prediction.

This question was investigated with a linear mixed effects model with N400 amplitude as a dependent variable, the proportion of target fixations and baseline EEG amplitude as predictor variables (see Methods Section 2.8 for details). Target fixations did not significantly predict N400 amplitude (B = 1.32, SE = 0.86, t = 1.54, p > .10). The model summary can be found in the supplementary material (Table SM4).

3.3.2. Predicting N400 amplitude from frequency power

The processing of the constraining information at verb onset was expected to influence the ease of subsequent noun processing. We therefore expected frequency (theta/alpha) power at verb onset to



Fig. 10. Panel A: Nouns' TFRs for the time window following predictable congruent (PC), predictable incongruent (PI) and unpredictable congruent (UC) nouns. Frequency power (μ V²/Hz for 4–30 Hz) plotted in PI, PC, and UC TFRs represents the absolute difference from baseline (-100-0ms). No baseline correction was applied when plotting the differences between conditions. Vertical black lines overlaying the TFRs mark the noun onset. Panel B: Topographical plots of differences in frequency power 0–1000ms after verb onset for theta (4–6 Hz), alpha (7–11 Hz) and beta (14–30 Hz) frequency bands. No baseline correction was applied. Significance levels are based on parametric effects: ***p < .001, *p < .05, *ns*: not significant.

predict the N400 amplitude at noun onset. N400 amplitude was entered into two linear mixed effects models as a dependent variable, with anterior frequency power, posterior frequency power, baseline amplitude, condition, a condition \times anterior frequency power interaction and a condition \times posterior frequency power interaction as predictors (see Methods Section 2.8 for details). Here we focused only on theta and alpha frequencies, as it was in these frequency bands that we predominantly observed effects of noun predictability (see Results, Frequency, Section 3.2.3).

The model summaries are presented in Tables 8 and 9. There was no main effect of theta power on N400 amplitude (p > .05). However, there was a significant interaction between theta power and condition, in both predictable congruent and predictable incongruent conditions in the anterior ROI. An estimate of 0.24 for the predictable congruent condition interaction (see Table 8) suggests that, relative to the average of all other conditions, the N400 amplitude was more positive with increased theta power. In other words, when the noun was predictable and *congruent* with the prediction, the N400 amplitude at noun onset decreased with increased theta power after verb onset. An estimate of -0.17 suggests that the opposite pattern was true for the predictable *incongruent* condition, where increased anterior theta power at verb onset resulted in a more negative (larger) N400 amplitude at noun onset.

There was a main effect of alpha power on N400 amplitude for the unpredictable congruent condition and the predictable incongruent condition. There was additionally a significant interaction between condition and alpha power in the anterior ROI, where N400 amplitude at noun onset decreased (became more positive) with increased alpha power at verb onset when the noun was predictable and congruent, as shown by the estimate of 0.15 (see Table 9).

3.3.3. Predicting frequency power from anticipatory fixations

We investigated to what extent the proportion of fixations before noun onset could predict log frequency power in the 0–1000ms after noun onset. To this end, log theta power and log alpha power were entered into two linear mixed effects models as dependent variables, with target fixations, ROI and a target fixations \times ROI interaction as predictors (see Methods Section 2.8 for details).

The proportion of fixations prior to noun onset was negatively associated with both theta and alpha power in the time period after the noun onset (see Tables 10 and 11, respectively for model summaries). A greater proportion of anticipatory target fixations was therefore associated with a weaker theta/alpha increase at noun onset. There was a significant effect of ROI, but no interaction between ROI and target fixations.

3.3.4. Predicting anticipatory fixations from frequency power

We hypothesised that oscillatory activity time-locked to the verb onset would predict anticipatory target fixations. We therefore investigated the extent that spectral power, averaged over 0–1000ms after verb onset, could predict the proportion of fixations in the time window from 490ms after verb onset until noun onset.

Due to the large number of zeros in the data, hurdle GAMM models were used (see Methods, Data Analysis Section 2.8). The proportion of target fixations was entered into the model as a dependent variable, with frequency power as a predictor and subject and sentence as random smooths. Data from anterior and posterior ROIs, and for alpha and theta frequencies, were entered into separate models. Model summaries are presented in Tables 12 and 13.

3.3.4.1. Theta anterior ROI. The zero part of the model tested the extent to which each variable predicted the binary outcome of whether or not the target object was fixated during the specified time window. There was no parametric effect of theta power and no significant smooth for theta power. These findings demonstrate that frontal theta power is not a good predictor of whether the participant will fixate on the target.

Model summary for theta and alpha power in response to Predictable vs Unpredictable nouns.

	Theta	Theta				Alpha			
Parametric coefficients	В	SE	t	Р	В	SE	t	Р	
Intercept	0.84	0.07	12.31	<.001	1.11	0.09	12.10	<.001	
Condition	-0.02	< 0.01	-9.84	<.001	-0.01	< 0.01	-4.16	<.001	
ROI	0.10	< 0.01	41.39	<.001	-0.05	< 0.01	-18.94	<.001	
$\textbf{Condition} \times \textbf{ROI}$	<0.01	<0.01	0.91	.365	< 0.01	< 0.01	0.37	.711	
Smooth terms	edf	Ref. df	F	P	edf	Ref. df	F	Р	
S: Pred	2.08	2.41	1.79	.122	1.01	1.01	2.13	.144	
S: Unpred	3.84	4.62	9.24	<.001	3.50	4.25	12.24	<.001	
S: Pred Ant	1.01	1.49	1.31	.205	1.01	1.01	0.09	.777	
S: Pred Post	1.00	1.01	0.95	.329	0.68	1.04	0.21	.761	
S: Unpred Ant	0.01	0.02	0.10	.967	< 0.01	< 0.01	0.04	.992	
S: Unpred Post	3.74	4.59	7.38	<.001	1.00	1.00	2.89	.089	
RS for subject	144.69	287.00	96.23	<.001	143.40	287.00	145.91	<.001	
RS for sentence	534.24	1151.00	3.29	<.001	528.60	1151.00	2.56	<.001	

For abbreviations, see legend of Table 5.

The non-zero part of the model tested the extent to which theta power predicted the total fixation time during the selected time window, conditional on having fixated the object. There was no parametric effect of theta power but there was a significant smooth for theta power.

3.3.4.2. Theta posterior ROI. There was no parametric effect of posterior theta power on whether or not the target object was fixated during the critical window. The smooth for theta power in the zero part of the model did not reach significance after Bonferroni correction for multiple tests (p = .044 uncorrected). There was, however, a significant parametric effect of theta power on the duration of fixations given the target was fixated and a significant smooth for posterior theta power on the total duration of fixations during the critical window.

For abbreviations, see legend of Table 5.

3.3.4.3. Alpha anterior ROI. Neither the zero nor non-zero parts of the model revealed significant parametric effects nor any significant smooth for anterior alpha power.

3.3.4.4. Alpha posterior ROI. In the zero part of the model, there was no significant parametric effect of posterior alpha power on the number of fixations and no significant smooth for alpha power. In the non-zero part of the model, there was no significant parametric effect of alpha power on total fixation duration, but a significant smooth for posterior alpha power.

3.3.5. Summary of combined analysis results

Anticipatory target fixations predicted spectral power (alpha and theta) but not N400 amplitude after noun onset. Spectral power after verb onset did not predict the binary outcome of whether or not the target was fixated prior to noun onset. However, posterior theta power after verb onset predicted the amount of time spent fixating on the target prior to noun onset given the target was fixated to begin with. No such relationship was found with alpha power. Theta and alpha power at verb onset interacted with condition when predicting N400 amplitude at noun onset, where an increase in anterior alpha and theta was associated with a smaller N400 in the predictable congruent condition, but there was a larger N400 in the predictable incongruent condition with increased anterior theta power.

4. Discussion

We tested the extent to which EEG and eye-tracking data could be simultaneously recorded and combined to investigate predictive processing of upcoming speech in naturalistic environments, in VR. Advancements in recording and data processing techniques mean that it is now possible to record EEG during free-viewing paradigms, uncovering a wide range of possibilities to study important theoretical questions in naturalistic environments in the lab. When studying prediction during spoken language comprehension, although eye gaze allows us to see whether a prediction can be made before the verbal onset of the referent, it is not possible to determine whether a reduction in the proportion of fixations is due to a shift in visual attention or an actual change in the listener's prediction. On the other hand, although the EEG signal at the onset of a spoken word provides an indication of the ease of word processing the moment it is perceived, we are not yet able to use the EEG signal alone to determine whether a noun was predicted or not. These two methods therefore provide complementary information that can be combined to answer important theoretical research questions (Knoeferle, 2015). Here we present a proof-of-principle investigation to test whether these methods can indeed be successfully combined to study the prediction of speech in naturalistic environments.

In summary, we replicated increased anticipatory fixations towards a referent when the verb in the sentence was predictive compared to unpredictive. We also found that the predictiveness of the verb and predictability of the noun were related to modulations in theta (4-6 Hz) and alpha (7-11 Hz) frequency bands at verb and noun onsets. Only tentative modulations were found in the beta frequency band (14-30 Hz). Increased theta power, arguably reflecting increased processing, occurred either at the constraining verb onset in predictable sentences or at noun onset in unpredictable sentences. Accordingly, anterior theta and alpha power after verb onset predicted subsequent N400 amplitude at the noun onset. Although spectral power (alpha and theta) at verb onset did not predict whether or not the target was fixated prior to noun onset, posterior theta power at verb onset did predict the amount of time spent fixating on the target prior to noun onset given the target was fixated to begin with. No such relationship was found with alpha power. The proportion of anticipatory fixations was found to predict theta and alpha power after the noun onset, possibly reflecting the need for fewer processing resources to process a word that has been predicted. The predictability of the noun was also associated with modulations in post-N400 ERP amplitudes, but not with the N400 amplitude itself, contrasting with the previous literature. Nouns congruent with the prediction were, however, associated with a smaller N400 amplitude relative to incongruent nouns. Noun congruence was additionally associated with modulations in the post-N400 ERP amplitude, as well as in alpha and theta band power. Below we discuss these results, and their theoretical and methodological implications, in more detail.

4.1. Noun predictability

We replicated the finding of increased anticipatory target fixations prior to noun onset in the predictable relative to the unpredictable



Fig. 11. Noun predictability modulated frequency power difference curves (panels A–D) and model estimated smooths (panels E–H): The difference between the model-estimated smooth splines of the predictive and unpredictive conditions for log theta (panels A–B) and alpha (panels C–D) power, and model estimated smooths for log theta (panels E and F) and alpha (panels G and H) power in anterior (panels A, C, E and G) and posterior (panels B, D, F and H) ROIs in the 1000 ms following noun onset. Red dashed lines mark windows of significant differences.

Model summary for theta and alpha power in response to Congruent vs Incongruent nouns.

	Theta	Theta				Alpha			
Parametric coefficients	В	SE	t	Р	В	SE	t	Р	
Intercept	0.83	0.07	12.08	<.001	1.11	0.09	12.42	<.001	
Condition	-0.02	< 0.01	-7.26	<.001	-0.01	< 0.01	-3.90	<.001	
ROI	0.11	< 0.01	46.82	<.001	-0.05	< 0.01	-18.33	<.001	
$\textbf{Condition} \times \textbf{ROI}$	-0.01	< 0.01	-4.41	<.001	< 0.01	< 0.01	-0.32	.747	
Smooth terms	edf	Ref. df	F	Р	edf	Ref. df	F	Р	
S: Cong	2.31	2.68	1.97	.087	1.01	1.01	3.61	.057	
S: Incong	4.93	5.89	8.72	<.001	1.01	1.02	0.81	.360	
S: Cong Ant	0.96	1.43	1.07	.255	1.84	2.27	1.87	.151	
S: Cong Post	1.01	1.01	1.01	.315	< 0.01	0.01	0.01	.994	
S: Incong Ant	< 0.01	0.01	0.13	.973	1.01	1.01	0.65	.414	
S: Incong Post	1.00	1.00	13.92	<.001	< 0.01	0.01	0.21	.971	
RS for subject	136.30	287.00	95.77	<.001	147.30	287.00	138.84	<.001	
RS for sentence	547.90	1151.00	3.35	<.001	498.00	1151.00	2.42	<.001	

For abbreviations, see legend of Table 5.



Fig. 12. Noun congruence modulated frequency power difference curves: The difference between the model-estimated smooth splines of the predictive and unpredictive conditions for log theta (panels A–B) and alpha (panels C–D) power, in anterior (panels A and C) and posterior (panels B and D) ROIs in the 1000 ms following noun onset. Red dashed lines mark windows of significant differences.

condition in a naturalistic virtual environment (Heyselaar et al., 2020; Huizeling et al., 2022). Fixations towards the target increased 480ms after verb onset to a proportion of \sim 0.12, which is consistent with Huizeling et al. (2022) and Heyselaar et al. (2020), where anticipatory fixations were found to increase at around 400ms after the constraining verb onset to a proportion of \sim 0.15. Importantly, these findings confirm that the referent was predicted before noun onset.

Both theta and alpha power were found to be greater in response to unpredictable compared to predictable nouns. Additionally, an increase in the proportion of anticipatory target fixations prior to the noun

Model summary: Predicting N400 amplitude at noun onset from theta power at verb onset.

Random effects:					
Groups	Name	Variance	SD		
Sentence	(Intercept)	0.18	0.43		
Subject	(Intercept)	0.99	1.00		
Residual		40.56	6.37		
Fixed effects:					
	Estimate	SE	df	t	Pr(> t)
(Intercept)	-2.01	0.21	33.04	-9.64	<.001
Pred cong	0.85	0.18	3622.00	4.67	<.001
Unpred cong	0.46	0.18	3616.00	2.55	.011
Pred incong	-0.62	0.18	3620.00	-3.41	<.001
Ant theta power	0.10	0.06	2695.00	1.81	.070
Post theta power	0.00	0.08	2644.00	-0.04	.970
Baseline	0.07	0.01	3722.00	5.08	<.001
Pred cong:ant	0.24	0.10	3702.00	2.41	.016
Unpred cong:ant	-0.03	0.09	3712.00	-0.28	.777
Pred incong:ant	-0.17	0.09	3714.00	-1.99	.047
Pred cong:post	-0.15	0.13	3696.00	-1.22	.222
Unpred cong:post	0.00	0.14	3725.00	-0.02	.983
Pred incong:post	0.10	0.12	3715.00	0.84	.403

For abbreviations, see legend of Table 5.

Table 9

Model summary: Predicting N400 amplitude at noun onset from alpha power at verb onset.

Random effects:					
Groups	Name	Variance	SD		
Sentence	(Intercept)	0.16	0.40		
Subject	(Intercept)	0.98	0.99		
Residual		40.57	6.37		
Fixed effects:					
	Estimate	SE	df	t	Pr(> t)
(Intercept)	-2.01	0.21	33.25	-9.71	<.001
Pred cong	0.84	0.18	3622.00	4.66	<.001
Unpred cong	0.46	0.18	3617.00	2.54	.011
Pred incong	-0.61	0.18	3621.00	-3.38	<.001
Ant theta power	0.01	0.03	2275.00	0.32	.746
Post theta power	0.02	0.01	2633.00	1.23	.220
Baseline	0.07	0.01	3723.00	5.05	<.001
Pred cong:ant	0.15	0.05	3718.00	2.81	.005
Unpred cong:ant	-0.06	0.05	3712.00	-1.27	.204
Pred incong:ant	-0.01	0.04	3732.00	-0.37	.713
Pred cong:post	-0.03	0.02	3717.00	-1.17	.242
Unpred cong:post	0.02	0.03	3721.00	0.79	.430
Pred incong:post	0.00	0.02	3718.00	-0.07	.946

For abbreviations, see legend of Table 5.

significantly predicted reduced theta and alpha power at the subsequent noun onset, independently of condition. Prediction of the referent therefore seems to reduce the amount of processing effort required when processing the noun, as is reflected in weaker alpha/theta increases. A relationship between the constraints provided by the verb in a sentence and the ease of subsequent noun processing has previously been found by Maess et al. (2016), who observed a negative correlation between the N400 amplitude at the onset of a constraining verb and the N400 amplitude at the predictable noun. Here the N400 could not be predicted by anticipatory fixations, in line with the absence of an effect of predictability on N400 amplitude. N400 amplitude was, however, related to frequency power at verb onset, dependent on condition. As anterior theta and alpha power in response to the verb increased, N400 amplitude decreased in the predictable congruent condition but increased with anterior theta power in the predictable incongruent condition. These findings corroborate that processing of the preceding contextual constraints facilitated subsequent noun processing when the noun was

Table 10

Predicting theta power at noun onset from fixations prior to noun onset.

Random effects:									
Groups	Name	Variance	SD	Corr					
Sentence	(Intercept)	0.01	0.08						
	Target fixations	0.06	0.24	-0.11					
Subject	(Intercept)	0.15	0.39						
	Target	0.02	0.16	0.38					
	fixations								
Residual		0.19	0.44						
Fixed effects:			·						
	Estimate	SE	df	t	Pr(> t)				
					τ ι)				
Intercept	0.98	0.07	32.67	14.20	<.001				
Target fixations	-0.17	0.05	35.00	-3.13	<.001				
ROI	0.11	0.01	3524.00	13.96	<.001				
Target fixations \times ROI	0.04	0.03	3524.00	1.13	.260				

For abbreviations, see legend of Table 2.

Table 11									
Predicting alpha	power	at noun	onset	from	fixations	prior	to	noun	onset.

Random effects:								
Groups	Name	Variance	SD	Corr				
Sentence	Intercept	0.01	0.07					
	Target fixations	0.06	0.24	-0.28				
Subject	(Intercept)	0.29	0.54					
	Target fixations	0.04	0.19	-0.22				
Residual		0.20	0.45					
Fixed effects:								
	Estimate	SE	df	t	Pr(> t)			
(Intercept)	1.30	0.10	32.27	13.45	<.001			
Target fixations	-0.20	0.06	38.16	-3.42	<.001			
ROI	-0.04	0.01	3534.00	-4.88	<.001			
Target fixations \times ROI	0.01	0.04	3534.00	0.18	0.85			

For abbreviations, see legend of Table 2.

congruent with the prediction, but may have interfered with processing when the noun was incongruent with the prediction. A relationship between pre-word alpha power and N400 amplitude has previously been found by Lago et al. (2023), where lower pre-word alpha power was associated with larger N400 amplitude in response to incongruent words. These recent findings are difficult to compare with the findings we present here, as we did not find the typical alpha decrease prior to predictable words.

Our findings of an effect of noun predictability on frequency power contrast with two previous studies that have found no effects of predictability on frequency power at the critical word onset (Rommers and Federmeier, 2018; Terporten et al., 2019). An increase in theta power for unpredictable nouns is consistent with an increase in processing resources being required to retrieve the lexical and semantic information of an unpredictable word, as well as integrate it with the prior sentence context. The experiment presented here was different from previous reports in that the (un)predictable referent was presented visually in parallel to the sentence. Going beyond a flat visual presentation, the referent was part of the virtual environment that the participant was immersed in. This visual and virtual presence meant that the participant could activate the semantic features of the object to a richer degree than if the participant predicted the object while reading a sentence. This may have led to a greater and more robust differentiation between the processing of predictable and unpredictable nouns in the present results. If

Model summary for predicting fixations from anterior and posterior theta power in response to verbs.

	Anterior				Posterior			
Parametric coefficients	В	SE	Z	Р	В	SE	Z	Р
Intercept	3.05	0.05	61.83	<.001	3.04	0.05	60.33	<.001
Log theta power	-0.03	0.02	-1.04	.298	-0.08	0.03	-2.61	<.001
Intercept	-1.84	0.10	-18.74	<.001	-1.85	0.10	-18.90	<.001
Log theta power	0.04	0.02	1.84	.066	0.02	0.04	0.53	.594
Smooth terms	edf	Ref. df	χ^2	Р	edf	Ref. df	χ^2	Р
S: Log theta power	6.28	8.00	246.90	<.001	5.29	8.00	158.25	.002
RS for subject	28.45	31.00	1416.51	<.001	28.61	31.00	1469.60	<.001
RS for sentence	109.61	123.00	1737.56	<.001	109.80	123.00	1890.27	<.001
S: Log theta power	0.78	8.00	1.27	.205	2.48	8.00	7.82	.044
RS for subject	23.44	31.00	89.64	<.001	23.21	31.00	86.81	<.001
RS for sentence	69.86	127.00	166.23	<.001	70.19	127.00	167.89	<.001

Table 13

Model summary for predicting fixations from anterior and posterior alpha power in response to verbs.

	Anterior				Posterior			
Parametric coefficients	В	SE	z	Р	В	SE	z	Р
Intercept	3.04	0.05	62.45	<.001	3.04	0.05	62.21	<.001
Log alpha power	< 0.01	< 0.01	-1.01	.315	< 0.01	< 0.01	-0.84	.402
Intercept	-1.84	0.10	-17.97	<.001	-1.84	0.10	-17.79	<.001
Log alpha power	< 0.01	0.01	0.23	.819	<0.01	< 0.01	-0.56	.577
Smooth terms	edf	Ref. df	χ^2	Р	edf	Ref. df	χ^2	Р
S: Log alpha power	2.67	8.00	14.94	.399	4.78	8.00	242.63	.001
RS for subject	28.36	31.00	1319.00	<.001	28.27	31.00	1412.49	<.001
RS for sentence	109.71	123.00	1791.03	<.001	110.00	123.00	1749.90	<.001
S: Log alpha power	0.03	8.00	0.03	.434	< 0.01	8.00	< 0.01	.756
RS for subject	24.18	31.00	102.91	<.001	24.35	31.00	104.69	<.001
RS for sentence	70.33	127.00	168.63	<.001	70.42	127.00	169.13	<.001

For abbreviations, see legend of Table 5.

this tentative conclusion is correct, this would confirm the added value of testing hypotheses in environments that are as rich as the settings in which listeners typically process speech in their everyday lives.

Contrary to our hypothesis and the previous literature, we did not observe an effect of noun predictability on the N400 amplitude. Although from Fig. 4 it appears as though the N400 was on average slightly larger in the unpredictable compared to predictable condition, this difference was not significant. These findings contrast with the effect of noun predictability on the pattern of anticipatory fixations in the current study, which show that participants could predict the noun in the predictable but not the unpredictable condition. Thus, together our findings provide concrete evidence supporting that the N400 reflects the integration of linguistic information, rather than directly reflecting predictive processes (Baggio and Hagoort, 2011; Mantegna et al., 2019). As previously mentioned, our paradigm differs from previous paradigms in that sentences were presented within a relevant 3D visual context, where the referent was visible in the scene. The N400 can be interpreted as reflecting the amount of new semantic information to be integrated or retrieved (Brothers and Kuperberg, 2021; Hagoort et al., 2009; Hodapp and Rabovsky, 2021; Kutas and Federmeier, 2011; Nieuwland et al., 2020). It is likely that the visual context in our experiment reduced the amount of new semantic information to be processed upon hearing unpredictable nouns, thereby resulting in a smaller N400 amplitude relative to if the referent had not been visible. Indeed, a post hoc analysis revealed a smaller N400 amplitude in response to unpredictable referents that were visible in the scene compared to those absent from the scene (see Table SM5 in the supplementary material). This finding seems to confirm the importance of the degree of richness of the environment in which stimuli are presented to the language user.

Our model estimated the post-N400 response to be more positive in response to predictable relative to unpredictable nouns. This is in the opposite direction to our expectations and the previous literature (DeLong et al., 2014; Federmeier et al., 2007; Quante et al., 2018; Thornhill and Van Petten, 2012; Van Petten and Luka, 2012). However, our results could be explained in reference to the work by Brothers et al. (2020), where plausible but unexpected words elicited a larger late anterior positivity compared to expected words only when the preceding context was globally rather than locally constraining. Similar to Experiment 2a in Brothers et al. (2020), here the overall sentence context was unconstraining and sentences were only constrained by the verb. Moreover, the verb was always positioned locally to the noun, often only separated with a single preposition and a determiner. The authors propose that post-N400 positivities are elicited only when a substantial situation model has been formed, when the prior linguistic context is rich and constraining. With locally predictive constraints, the predicted event had not yet been built into the situation model, resulting in no requirement for the situation model to be updated. Our findings are also in line with Lau et al. (2013), who found a late negativity, rather than a positivity, in response to unpredicted targets in prime-target word pairs. Again, with word pairs there was no opportunity to build up a rich and meaningful situation model. In relation to the current paradigm, this explanation raises the question of how much time is needed for the listener's situation model to be updated with the highly constraining information. Our findings also raise the question of to what extent the visual context affects post-N400 effects. If they are indeed reflecting the updating of the situation model then it might be expected for them to also be influenced by visual context constraints. Indeed, situation models should integrate information that is concurrently and consecutively perceived via the different (visual, auditory) modalities and senses.

Alternatively, it may be that the post-N400 effects are task-specific. Previous studies investigating the effect of prediction on post-N400 components have been conducted in the laboratory with (un)constraining sentence stimuli (Brothers et al., 2020; DeLong et al., 2014; Federmeier et al., 2007; Kuperberg et al., 2020; Quante et al., 2018; Rommers and Federmeier, 2018; Thornhill and Van Petten, 2012; Van Petten and Luka, 2012). Nieuwland et al. (2020) highlighted that components observed over hundreds of milliseconds, like the post-N400, are likely a result of the combined activity of multiple sources, as is evident from source localisation of the magnetoencephalography signal in response to single words (Arana et al., 2020; Huizeling et al., 2021; Pylkkänen and Marantz, 2003). The combination of components contributing to the signal can therefore be expected to be different when the linguistic stimuli are integrated with a rich visual context, as was the case in the present study. In addition to post-N400 effects being task-specific, later components, which are further from the baseline, are more vulnerable to error when traditional baseline correction techniques are used instead of including the baseline as a regressor in the statistical model, as we did here. Indeed, there is evidence to suggest that some post-N400 effects in the literature may be artefactual and disappear when improved baseline correction techniques are applied (Alday, 2019).

4.2. Noun congruence

Consistent with our hypothesis and the previous literature, in the predictable condition the N400 amplitude was significantly smaller when the referent was congruent with the context provided by the visual scene, confirming the listener's prediction, compared to when the referent was incongruent with the visual context, disconfirming their prediction. These findings are consistent with the wider N400 literature and with multimodal N400 effects to incongruent visual and linguistic information (Sitnikova et al., 2008; Willems et al., 2008a,b). A larger N400 amplitude for incongruent nouns likely reflects an increased amount of semantic information to be retrieved and integrated with the prior context at the time of word onset (Brothers and Kuperberg, 2021; Hagoort et al., 2009; Hodapp and Rabovsky, 2021; Kutas and Federmeier, 2011), in comparison to the congruent condition where the referent was visible and thus some of the semantic features may have already been retrieved prior to word onset.

Effects of noun congruence were found on both theta and alpha power in response to the noun onset, an effect that interacted with ROI for theta frequency. Consistent with the previous literature, higher theta was found in response to incongruent relative to congruent nouns and in the anterior relative to posterior ROI (Bastiaansen and Hagoort, 2015; Li et al., 2017; Rommers et al., 2017; Willems et al., 2008a), an effect that was sustained throughout the 1000ms analysis window after noun onset. An increase in theta power has been observed for both unexpected yet plausible nouns and to semantically anomalous nouns. Higher theta could reflect the retrieval of semantic and lexical information from memory or increased cognitive control to suppress the predicted word and to update the sentence interpretation (Cavanagh and Cohen, 2022; Cavanagh and Frank, 2014; Demiralp and Başar, 1992; Klimesch et al., 1994). It is likely that both information retrieval and cognitive control contributed to increased theta here.

Alpha power was higher in the incongruent relative to congruent condition and was stable across time. The direction of this effect is inconsistent with our hypotheses, where we expected increased attention and processing, and therefore reduced alpha, in the incongruent compared to congruent condition (Rommers et al., 2017; Wang et al., 2012; Willems et al., 2008a). For example, Willems et al. (2008a) found lower alpha power when the linguistic and visual information mismatched the sentence context. In the current paradigm, both the auditorily presented noun and the visual input mismatched the participant's expectation but were plausible within the sentence context, rather than anomalous. High alpha is often seen during sustained attention and when inhibiting irrelevant distracting information (Dockree et al., 2007; Rihs et al., 2007, 2009). It could be that a stronger inhibitory response is

needed to suppress the predicted item in the current paradigm due to the object's presence in the virtual environment with the participant. Furthermore, increased alpha may reflect the effects of sustained attention, where attention towards the virtual speaker is sustained for longer when the linguistic information is more challenging to process. This could particularly be the case in the rich visual environment in VR where there is additional irrelevant information to inhibit, including the ongoing presence of the predicted object. Indeed, alpha synchronisation has been shown to be associated with supressing irrelevant information in listening tasks (Dimitrijevic et al., 2019; Strauß et al., 2014b). Such speculative hypotheses require further investigation in relation to the current paradigm.

Consistent with our hypotheses and the previous literature, the model estimated the post-N400 response to be more positive in response to incongruent relative to congruent nouns, particularly in the posterior ROI. Typically, a greater anterior positivity has been associated with lexical items that are plausible but unexpected in highly constraining contexts (Brothers et al., 2020; DeLong et al., 2014; Federmeier et al., 2007; Kuperberg et al., 2020; Quante et al., 2018; Thornhill and Van Petten, 2012; Van Petten and Luka, 2012), but here the model estimated the positivity to be greater in incongruent relative to congruent nouns in the posterior ROI. The posterior positivity is typically seen in response to anomalous sentence endings and has been argued to reflect an inability to integrate the word with the prior context and a reanalysis of the sentence (Brothers et al., 2020; Kuperberg et al., 2020; Rommers and Federmeier, 2018; Van Petten and Luka, 2012). Here we saw a discrepancy between the data plotted in Fig. 5 and the outcome of the model, making our findings difficult to interpret in relation to previous results. Moreover, the differences between the current paradigm and previous reports make it difficult to directly compare our results with the results of others. For example, later effects are further from the baseline and thus more subject to slow drifts in the data. This could be exacerbated in a VR environment, where the participant is making small movements with their head and the data is more vulnerable to environmental noise. Further research is needed to test whether the current post-N400 findings are robust.

4.3. Verb's predictiveness

Both theta and alpha power were higher in response to predictive relative to unpredictive verbs and changed over time in response to the predictive verbs in an anterior ROI. Theta additionally changed over time in an anterior ROI in response to unpredictive verbs. Higher alpha and theta power in response to predictive relative to unpredictive verbs contrasts with the previous literature. Li et al. (2017) found lower alpha in response to predictive compared to unpredictive verbs and no effect of verb predictivity on theta power. Moreover, numerous studies have found lower alpha power in response to predictive context directly preceding predictable words (Gastaldon et al., 2020; Leon-Cabrera et al., 2022; Li et al., 2017; Molinaro et al., 2017; Piai et al., 2014; Rommers et al., 2017; Roos and Piai, 2020; Terporten et al., 2019; Wang et al., 2018). On inspection of Fig. 8, it appears as though the observed effects are a result of an anterior power increase in a broader theta-alpha frequency band in the predictable condition, ranging from 4 to 12 Hz. This could reflect the same process across frequencies. However, similar broad effects have been found in response to spoken auditory stimuli in past reports and were found to be dissociated across alpha and theta frequency bands at the source level (Strau β et al., 2014a). Conducting source level analysis went beyond the scope of the current paper and so we cannot make any such conclusions here.

In the current data, theta but not alpha power in a posterior ROI predicted more time spent fixating the object, providing the object had been fixated to begin with. Neither theta nor alpha power, however, predicted the binary outcome of whether or not the object was fixated. One reason for this could be that a prediction can take place without an eye movement towards the object. A lack of eye gaze towards the

referent could occur either when the referent has been predicted but not fixated or when the referent has not been predicted. In contrast, eye gaze towards the referent is more strongly indicative of the item having been predicted and the time spent gazing at an object may therefore be predicted more robustly by the associated neural correlates.

4.4. Application to new research directions

The results presented here lay the foundations for future work that leverages the complementary advantages of eye-tracking and EEG data to answer new theoretical questions. Combining these methods will be particularly useful to distinguish between theories about the extent that listeners use disfluencies to inform their predictions about upcoming spoken language. Huizeling et al. (2022) found that anticipatory referent fixations no longer increased if the speaker hesitated prior to naming a predictable referent, contrasting with sentences spoken fluently. Instead, looks towards the virtual speaker increased. However, it was not clear whether a change to the pattern of fixations reflected a change to the listener's prediction or a shift in their visual attention towards the speaker. It could be that the hesitation caused listeners to lose confidence in their prediction and wait for the sentence to be disambiguated as they placed less weight on their prediction. Alternatively, the salient interruption to speech may have captured the listener's attention to the speaker without a change to their prediction. Recording EEG while participants listen to (un)predictable sentences under fluent and disfluent conditions could help to disentangle these different theoretical accounts. In the current results we found higher theta power in response to unpredictable compared to predictable nouns, possibly reflecting a greater amount of resources required to retrieve the semantic and lexical information associated with unpredictable relative to predictable words. If listeners indeed abandon or place less weight on their prediction after hearing a hesitation, it might be expected that the theta response to predictable nouns would be larger after hearing a hesitation compared to when hearing a fluent sentence and would instead reflect a similar theta response to that of unpredictable nouns. Our future work aims to answer these questions.

In addition to opening new theoretical research avenues, the current work demonstrates the need to determine which effects from the literature can be replicated under conditions of more naturalistic language processing. Some of our findings were surprising in the context of previous reports that were conducted in computer-based laboratory settings (e.g. there was no effect of predictability on the N400 and alpha power was higher instead of lower in response to increased contextual constraints). Overall, embedding language within a rich context seems to introduce some changes to language processing compared to in a traditional laboratory setting. However, further research is needed to replicate the findings that we present here to determine which of our findings are robust. This research trajectory will yield many important insights for contextually rich and naturalistic language processing.

4.5. Conclusions

In summary, we demonstrate that it is possible to combine EEG and eye-tracking in VR to study the prediction of speech through the analysis of anticipatory eye movements, ERPs and spectral power effects. We were able to use the proportion of anticipatory fixations as an estimate of predictability, independently of condition, to successfully predict EEG signal at noun onset. Furthermore, we were able to predict the proportion of anticipatory fixations from the theta power in response to the constraining verb. Verb predictivity and noun predictability modulated alpha and theta frequency power. Overall, we found that increased processing resources, as reflected in increased theta power, were applied either at the verb onset when the verb was predictive of the noun, or at noun onset if the verb was not predictive of the noun. As such, listeners flexibly allocate cognitive resources at different moments in time while processing a sentence as a function of when this allocation is most effective and required for quickly and efficiently grasping the overall message-level meaning provided through the incoming information. Spectral power effects were best observed in lower frequency bands (alpha and theta, as opposed to beta) that have a better signal to noise ratio than higher frequency bands and are less susceptible to interference from muscle artifacts. Surprisingly, alpha power was higher, rather than lower, in response to the predictive verb and to unpredictable nouns. We replicated typical effects of noun congruence but not predictability on the N400 in response to the noun. The rich visual context that accompanied speech therefore altered the findings compared to previous reports. The visual context seems to ease the processing of unpredictable nouns, possibly as part of the semantic information has already been retrieved, and may result in a stronger inhibitory response to suppress the predicted item that is present in the scene. As such, our findings suggest that predictive processing in rich and dynamic everyday situations may be subserved by mechanisms that partially differ from those proposed by experimental studies that used a computer monitor as their means of stimulus display.

Author contributions

EH, DP and PH contributed towards the conceptualization of the project. EH collected and analysed the data, managed the project and wrote the original draft. DP provided support throughout the development of the project. PA assisted with the data analysis. PH provided funding for the project. All authors contributed towards reviewing and editing the manuscript.

Declaration of competing interest

The authors have no conflict of interest to declare.

Data availability

Dat and analysis scripts can be downloaded from the Max Planck Institute for Psycholinguistics' Language Archive (https://hdl.handle. net/1839/607d53ad-9a81-4891-856c-8bfcaff99a8b)

Acknowledgements

Research was funded by the Max Planck Society. PH was supported by the NWO Grant Language in Interaction, grant number 024.001.006. The authors would like to thank research assistants Iris Schmits, Eva Poort, Marjolijn Dijkhuis and Birgit Knudsen for helping to develop the Dutch sentence stimuli and with data collection. The authors are grateful for the Max Planck Institute for Psycholinguistics' Technical Group, particularly Reiner Dirksmeyer and Albert Russel for providing technical support, and Jeroen Derks for programming the experiment and creating the visual stimuli. Many thanks to Evelien Heyselaar for sharing R scripts for pre-processing the eye-tracking data.

Appendix ASupplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2023.108730.

References

Alday, P.M., 2019. How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. Psychophysiology 56 (12), e13451. https://doi.org/10.1111/psyp.13451.

- Alday, P.M., Kretzschmar, F., 2019. Speed-accuracy tradeoffs in brain and behavior: testing the independence of P300 and N400 related processes in behavioral responses to sentence categorization. Front. Hum. Neurosci. 13 (285) https://doi. org/10.3389/fnhum.2019.00285.
- Allopenna, P.D., Magnuson, J.S., Tanenhaus, M.K., 1998. Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. J. Mem. Lang. 38 (4), 419–439. https://doi.org/10.1006/jmla.1997.2558.

Altmann, G.T.M., Kamide, Y., 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. Cognition 73 (3), 247–264. https://doi.org/ 10.1016/S0010-0277(99)00059-1.

- Arana, S., Marquand, A., Hultén, A., Hagoort, P., Schoffelen, J.-M., 2020. Sensory modality-independent activation of the brain network for language. J. Neurosci. 40 (14), 2914–2924. https://doi.org/10.1523/jneurosci.2271-19.2020.
- Arnold, J.E., Fagnano, M., Tanenhaus, M.K., 2003. Disfluencies signal theee, um, new information. J. Psycholinguist. Res. 32 (1), 25–36. https://doi.org/10.1023/a: 1021980931292.
- Arnold, J.E., Tanenhaus, M.K., 2011. Disfluency effects in comprehension: how new information can become accessible. In: The Processing and Acquisition of Reference, pp. 197–217.
- Arnold, J.E., Tanenhaus, M.K., Altmann, R.J., Fagnano, M., 2004. The old and thee, uh, new:disfluency and reference resolution. Psychol. Sci. 15 (9), 578–582. https://doi. org/10.1111/j.0956-7976.2004.00723.x.
- Badia, S.B.i., Morgade, A.G., Samaha, H., Verschure, P.F.M.J., 2013. Using a hybrid brain computer interface and virtual reality system to monitor and promote cortical reorganization through motor activity and motor imagery training. IEEE Trans. Neural Syst. Rehabil. Eng. 21 (2), 174–181. https://doi.org/10.1109/ TNSRE.2012.2229295.
- Baggio, G., Hagoort, P., 2011. The balance between memory and unification in semantics: a dynamic account of the N400. Lang. Cognit. Process. 26 (9), 1338–1367. https://doi.org/10.1080/01690965.2010.542671.
- Bastiaansen, M., Hagoort, P., 2015. Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. J. Cognit. Neurosci. 27 (11), 2095–2107. https://doi.org/10.1162/jocn_a_00829.
- Bates, D., 2019. Complexity in fitting linear mixed models. nextjournal. https://doi.org/ 10.33016/nextjournal.100002.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67 (1), 1–48. https://doi.org/10.18637/jss.v067.i01.
- Bengtsson, H., 2017. R.matlab: Read and Write MAT Files and Call MATLAB from within R.
- Bonhage, C.E., Mueller, J.L., Friederici, A.D., Fiebach, C.J., 2015. Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. Cortex 68, 33–47. https://doi.org/10.1016/j.cortex.2015.04.011.
- Bortfeld, H., Leon, S.D., Bloom, J.E., Schober, M.F., Brennan, S.E., 2001. Disfleency rates in conversation: effects of age, relationship, topic, role, and gender. Lang. Speech 44 (2), 123–147. https://doi.org/10.1177/00238309010440020101.
- Brennan, S.E., Williams, M., 1995. The feeling of Another's knowing: prosody and filled pauses as cues to listeners about the metacognitive States of speakers. J. Mem. Lang. 34 (3), 383–398. https://doi.org/10.1006/jmla.1995.1017.
- Brothers, T., Kuperberg, G.R., 2021. Word predictability effects are linear, not logarithmic: implications for probabilistic models of sentence comprehension. J. Mem. Lang. 116, 104174 https://doi.org/10.1016/j.jml.2020.104174.
- Brothers, T., Wlotko, E.W., Warnke, L., Kuperberg, G.R., 2020. Going the extra mile: effects of discourse context on two late positivities during language comprehension. Neurobiology of Language 1 (1), 135-160. https://doi.org/10.1162/nol_a_00006.
- Cavanagh, J.F., Cohen, M.X., 2022. Frontal midline theta as a model specimen of cortical theta. In: The Oxford Handbook of EEG Frequency, vol. 178.
- Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cognit. Sci. 18 (8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012.
- Coco, M.I., Nuthmann, A., Dimigen, O., 2020. Fixation-related brain potentials during semantic integration of object–scene information. J. Cognit. Neurosci. 32 (4), 571–589. https://doi.org/10.1162/jocn_a_01504.%M.31765602.
- Coopmans, C.W., Nieuwland, M.S., 2020. Dissociating activation and integration of discourse referents: evidence from ERPs and oscillations. Cortex 126, 83–106. https://doi.org/10.1016/j.cortex.2019.12.028.
- Corley, M., MacGregor, L.J., Donaldson, D.I., 2007. It's the way that you, er, say it: hesitations in speech affect language comprehension. Cognition 105 (3), 658–668. https://doi.org/10.1016/j.cognition.2006.10.010.
- Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C., 1992. The CAVE: audio visual experience automatic virtual environment. Commun. ACM 35 (6), 64–73.
- Dambacher, M., Kliegl, R., Hofmann, M., Jacobs, A.M., 2006. Frequency and predictability effects on event-related potentials during reading. Brain Res. 1084 (1), 89–103. https://doi.org/10.1016/j.brainres.2006.02.010.
- DeLong, K.A., Quante, L., Kutas, M., 2014. Predictability, plausibility, and two late ERP positivities during written sentence comprehension. Neuropsychologia 61, 150–162. https://doi.org/10.1016/j.neuropsychologia.2014.06.016.
- Demiralp, T., Başar, E., 1992. Theta rhythmicities following expected visual and auditory targets. Int. J. Psychophysiol. 13 (2), 147–160. https://doi.org/10.1016/0167-8760 (92)90054-F.
- Dimigen, O., 2020. Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. Neuroimage 207, 116117. https://doi.org/10.1016/j. neuroimage.2019.116117.
- Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A.M., Kliegl, R., 2011. Coregistration of eye movements and EEG in natural reading: analyses and review. J. Exp. Psychol. Gen. 140 (4), 552–572. https://doi.org/10.1037/a0023885.
- Dimitrijevic, A., Smith, M.L., Kadis, D.S., Moore, D.R., 2019. Neural indices of listening effort in noisy environments. Sci. Rep. 9 (1), 11278 https://doi.org/10.1038/ s41598-019-47643-1.
- Dockree, P.M., Kelly, S.P., Foxe, J.J., Reilly, R.B., Robertson, I.H., 2007. Optimal sustained attention is linked to the spectral content of background EEG activity: greater ongoing tonic alpha (~10 Hz) power supports successful phasic goal activation. Eur. J. Neurosci. 25 (3), 900–907. https://doi.org/10.1111/j.1460-9568.2007.05324.x.

- Eichert, N., Peeters, D., Hagoort, P., 2018. Language-driven anticipatory eye movements in virtual reality. Behav. Res. Methods 50 (3), 1102–1115. https://doi.org/10.3758/ s13428-017-0929-z.
- Federmeier, K.D., Kutas, M., 1999. A rose by any other name: long-term memory structure and sentence processing. J. Mem. Lang. 41 (4), 469–495. https://doi.org/ 10.1006/jmla.1999.2660
- Federmeier, K.D., Kutas, M., 2001. Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. J. Exp. Psychol. Learn. Mem. Cognit. 27 (1), 202–224. https://doi.org/10.1037/ 0278-7393.27.1.202.
- Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., Kutas, M., 2007. Multiple effects of sentential constraint on word processing. Brain Res. 1146, 75–84. https://doi.org/ 10.1016/j.brainres.2006.06.101.
- Foxe, J.J., Simpson, G.V., Ahlfors, S.P., 1998. Parieto-occipital ~1 0Hz activity reflects anticipatory state of visual attention mechanisms. Neuroreport 9 (17), 3929–3933. Retrieved from. https://journals.lww.com/neuroreport/Fulltext/1998/12010/Pari eto_occipital_1_0Hz_activity_reflects.30.aspx.
- Fraundorf, S.H., Watson, D.G., 2014. Alice's adventures in um-derland: psycholinguistic sources of variation in disfluency production. Language, Cognition and Neuroscience 29 (9), 1083–1096. https://doi.org/10.1080/01690965.2013.832785.
- Ganis, G., Kutas, M., Sereno, M.I., 1996. The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. J. Cognit. Neurosci. 8 (2), 89–106. https://doi.org/10.1162/jocn.1996.8.2.89.
- Gastaldon, S., Arcara, G., Navarrete, E., Peressotti, F., 2020. Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. Cortex 133, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026.

Hagoort, P., Baggio, G., Willems, R.M., 2009. Semantic unification. In: The Cognitive Neurosciences, fourth ed. MIT press, pp. 819–836.

- Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. Science 304 (5669), 438–441. https://doi.org/10.1126/science.1095455.
- Hasson, U., Egidi, G., Marelli, M., Willems, R.M., 2018. Grounding the neurobiology of language in first principles: the necessity of non-language-centric explanations for language comprehension. Cognition 180, 135–157. https://doi.org/10.1016/j. cognition.2018.06.018.
- Heine, A., Tamm, S., Hofmann, M., Bosel, R.M., Jacobs, A.M., 2006. Event-related theta activity reflects memory processes in pronoun resolution. Neuroreport 17 (18), 1835–1839. https://doi.org/10.1097/WNR.0b013e328010a096.
- Hendrix, P., Bolger, P., Baayen, H., 2017. Distinct ERP signatures of word frequency, phrase frequency, and prototypicality in speech production. J. Exp. Psychol. Learn. Mem. Cognit. 43 (1), 128–149. https://doi.org/10.1037/a0040332.
- Heyselaar, E., Peeters, D., Hagoort, P., 2020. Do we predict upcoming speech content in naturalistic environments? Language, Cognition and Neuroscience 36 (4), 1–22. https://doi.org/10.1080/23273798.2020.1859568.
- Hodapp, A., Rabovsky, M., 2021. The N400 ERP component reflects an error-based implicit learning signal during language comprehension. Eur. J. Neurosci. 1–16. https://doi.org/10.1111/ejn.15462.

Holler, J., Levinson, S.C., 2019. Multimodal language processing in human communication. Trends Cognit. Sci. 23 (8), 639–652.

- Huettig, F., Rommers, J., Meyer, A.S., 2011. Using the visual world paradigm to study language processing: a review and critical evaluation. Acta Psychol. 137 (2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003.
- Huizeling, E., Arana, S., Hagoort, P., Schöffelen, J.-M., 2021. Lexical frequency and sentence context influence the brain's response to single words. Neurobiology of Language 1–31. https://doi.org/10.1162/nol_a_00054.
- Huizeling, E., Peeters, D., Hagoort, P., 2022. Prediction of upcoming speech under fluent and disfluent conditions: eye tracking evidence from immersive virtual reality. Language, Cognition and Neuroscience 37 (4), 481–508. https://doi.org/10.1080/ 23273798.2021.1994621.
- Jensen, O., Mazaheri, A., 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front. Hum. Neurosci. 4, 186. Retrieved from. https: //www.ncbi.nlm.nih.gov/pmc/articles/PMC2990626/pdf/fnhum-04-00186.pdf.
- Keuleers, E., Brysbaert, M., New, B., 2010. SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. Behav. Res. Methods 42 (3), 643–650. https://doi. org/10.3758/BRM.42.3.643.
- Klimesch, W., Sauseng, P., Hanslmayr, S., 2007. EEG alpha oscillations: the inhibition–timing hypothesis. Brain Res. Rev. 53 (1), 63–88. https://doi.org/ 10.1016/j.brainresrev.2006.06.003.
- Klimesch, W., Schimke, H., Schwaiger, J., 1994. Episodic and semantic memory: an analysis in the EEG theta and alpha band. Electroencephalogr. Clin. Neurophysiol. 91 (6), 428–441. https://doi.org/10.1016/0013-4694(94)90164-3.
- Knoeferle, P., 2015. Language Comprehension in Rich Non-linguistic Contexts: Combining Eye-Tracking and Event-Related Brain Potentials.
- Knoeferle, P., Urbach, T.P., Kutas, M., 2011. Comprehending how visual context influences incremental sentence processing: insights from ERPs and picture-sentence verification. Psychophysiology 48 (4), 495–506. https://doi.org/10.1111/j.1469-8986.2010.01080.x.
- Kretzschmar, F., Schlesewsky, M., Staub, A., 2015. Dissociating word frequency and predictability effects in reading: evidence from coregistration of eye movements and EEG. J. Exp. Psychol. Learn. Mem. Cognit. 41 (6), 1648–1662. https://doi.org/ 10.1037/xlm0000128.
- Kuperberg, G.R., Brothers, T., Wlotko, E.W., 2020. A tale of two positivities and the N400: distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. J. Cognit. Neurosci. 32 (1), 12–35. https://doi.org/ 10.1162/jocn.a.01465.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62 (1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Kutas, M., Hillyard, S., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427), 203–205. https://doi.org/10.1126/ science.7350657.

Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307 (5947), 161–163. https://doi.org/10.1038/ 307161a0.

Kuznetsova, A.B., B, P., Christensen, R.H.B., 2017. ImerTest package: tests in linear mixed effects models. J. Stat. Software 82 (13), 1–26. https://doi.org/10.18637/jss. v082.i13.

Lago, S., Pezzetta, R., Gastaldon, S., Peressotti, F., Arcara, G., 2023. Trial-by-trial fluctuations of pre-stimulus alpha power predict language ERPs. Psychophysiology n/a (n/a), e14388. https://doi.org/10.1111/psyp.14388.

Lau, E.F., Holcomb, P.J., Kuperberg, G.R., 2013. Dissociating N400 effects of prediction from association in single-word contexts. J. Cognit. Neurosci. 25 (3), 484–502. https://doi.org/10.1162/jocn_a_00328.%M.23163410.

Leon-Cabrera, P., Piai, V., Moris, J., Rodriguez-Fornells, A., 2022. Alpha power decreases associated with prediction in written and spoken sentence comprehension. Neuropsychologia 173. https://doi.org/10.1016/j.neuropsychologia.2022.108286.

Li, X., Zhang, Y., Xia, J., Swaab, T.Y., 2017. Internal mechanisms underlying anticipatory language processing: evidence from event-related-potentials and neural oscillations. Neuropsychologia 102, 70–81. https://doi.org/10.1016/j. neuropsychologia.2017.05.017.

Luck, S.J., Gaspelin, N., 2017. How to get statistically significant effects in any ERP experiment (and why you shouldn't). Psychophysiology 54 (1), 146–157. https:// doi.org/10.1111/psyp.12639.

MacGregor, L.J., Corley, M., Donaldson, D.I., 2010. Listening to the sound of silence: disfluent silent pauses in speech have consequences for listeners. Neuropsychologia 48 (14), 3982–3992. https://doi.org/10.1016/j.neuropsychologia.2010.09.024.

Maess, B., Mamashli, F., Obleser, J., Helle, L., Friederici, A.D., 2016. Prediction signatures in the brain: semantic pre-activation during language comprehension. Front. Hum. Neurosci. 10 (591) https://doi.org/10.3389/fnhum.2016.00591.

Mantegna, F., Hintz, F., Ostarek, M., Alday, P.M., Huettig, F., 2019. Distinguishing integration and prediction accounts of ERP N400 modulations in language processing through experimental design. Neuropsychologia 134, 107199. https:// doi.org/10.1016/j.neuropsychologia.2019.107199.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., Bates, D., 2017. Balancing Type I error and power in linear mixed models. J. Mem. Lang. 94, 305–315. https://doi.org/ 10.1016/j.jml.2017.01.001.

Meyer, L., Grigutsch, M., Schmuck, N., Gaston, P., Friederici, A.D., 2015. Frontalposterior theta oscillations reflect memory retrieval during sentence comprehension. Cortex 71, 205–218. https://doi.org/10.1016/j.cortex.2015.06.027.

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. https://doi.org/10.1016/j. cognition.2017.03.012.

Nieuwland, M.S., Barr, D.J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D.I., Von Grebmer Zu Wolfsthurn, S., 2020. Dissociable effects of prediction and integration during language comprehension: evidence from a large-scale study using brain potentials. Phil. Trans. Biol. Sci. 375 (1791), 20180522 https://doi.org/ 10.1098/rstb.2018.0522.

Nieuwland, M.S., van Berkum, J., 2006. When peanuts fall in love: N400 evidence for the power of discourse. J. Cognit. Neurosci. 18, 1098–1111. https://doi.org/10.1162/ jocn.2006.18.7.1098.

Nikolaev, A.R., Meghanathan, R.N., van Leeuwen, C., 2016. Combining EEG and eye movement recording in free viewing: pitfalls and possibilities. Brain Cognit. 107, 55–83. https://doi.org/10.1016/j.bandc.2016.06.004.

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 9. https://doi.org/10.1155/ 2011/156869.

Özyürek, A., Willems, R.M., Kita, S., Hagoort, P., 2007. On-line integration of semantic information from speech and gesture: insights from event-related brain potentials. J. Cognit. Neurosci. 19 (4), 605–616. https://doi.org/10.1162/jocn.2007.19.4.605.

Peeters, D., 2019. Virtual reality: a game-changing method for the language sciences. Psychon. Bull. Rev. 26 (3), 894–900. https://doi.org/10.3758/s13423-019-01571-3.

Peeters, D., 2020. Bilingual switching between languages and listeners: insights from immersive virtual reality. Cognition 195, 104107. https://doi.org/10.1016/j. cognition.2019.104107.

Perniss, P., 2018. Why we should study multimodal language. Front. Psychol. 9 https:// doi.org/10.3389/fpsyg.2018.01109.

Piai, V., Roelofs, A., Maris, E., 2014. Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. Neuropsychologia 53, 146–156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014.

Porretta, V., Kyröläinen, A.-J., van Rij, J., Järvikivi, J., 2018. Visual World Paradigm Data: from Preprocessing to Nonlinear Time-Course Analysis (Cham).

Prystauka, Y., Lewis, A.G., 2019. The power of neural oscillations to inform sentence comprehension: a linguistic perspective. Language and Linguistics Compass 13 (9), e12347. https://doi.org/10.1111/lnc3.12347.

Pylkkänen, L., Marantz, A., 2003. Tracking the time course of word recognition with MEG. Trends Cognit. Sci. 7 (5), 187–189. https://doi.org/10.1016/S1364-6613(03) 00092-5.

Quante, L., Bölte, J., Zwitserlood, P., 2018. Dissociating predictability, plausibility and possibility of sentence continuations in reading: evidence from late-positivity ERPs. PeerJ 6, e5717. https://doi.org/10.7717/peerj.5717.

Raghavan, R., Raviv, L., Peeters, D., 2023. What's your point? Insights from virtual reality on the relation between intention and action in the production of pointing gestures. Cognition 240, 105581. https://doi.org/10.1016/j. cognition.2023.105581.

Rayner, K., Slowiaczek, M.L., Clifton, C., Bertera, J.H., 1983. Latency of sequential eye movements: implications for reading. J. Exp. Psychol. Hum. Percept. Perform. 9 (6), 912–922. https://doi.org/10.1037/0096-1523.9.6.912.

Rihs, T.A., Michel, C.M., Thut, G., 2007. Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. Eur. J. Neurosci. 25 (2), 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x.

Rihs, T.A., Michel, C.M., Thut, G., 2009. A bias for posterior α-band power suppression versus enhancement during shifting versus maintenance of spatial attention. Neuroimage 44 (1), 190–199. https://doi.org/10.1016/j.neuroimage.2008.08.022.

Rommers, J., Dickson, D.S., Norton, J.J.S., Wlotko, E.W., Federmeier, K.D., 2017. Alpha and theta band dynamics related to sentential constraint and word expectancy. Language Cognition and Neuroscience 32 (5), 576–589. https://doi.org/10.1080/ 23273798.2016.1183799.

Rommers, J., Federmeier, K.D., 2018. Predictability's aftermath: downstream consequences of word predictability as revealed by repetition effects. Cortex 101, 16–30. https://doi.org/10.1016/j.cortex.2017.12.018.

Roos, N.M., Piai, V., 2020. Across-session consistency of context-driven language processing: a magnetoencephalography study. Eur. J. Neurosci. 52 (5), 3457–3469. https://doi.org/10.1111/ejn.14785.

Schachter, S., Christenfeld, N., Ravina, B., Bilous, F., 1991. Speech disfluency and the structure of knowledge. J. Pers. Soc. Psychol. 60 (3), 362–367. https://doi.org/ 10.1037/0022-3514.60.3.362.

Sitnikova, T., Holcomb, P.J., Kiyonaga, K.A., Kuperberg, G.R., 2008. Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. J. Cognit. Neurosci. 20 (11), 2037–2057. https://doi.org/10.1162/ iogn. 2008 20143

Smith, V.L., Clark, H.H., 1993. On the course of answering questions. J. Mem. Lang. 32 (1), 25–38, https://doi.org/10.1006/imla.1993.1002.

Strauß, A., Kotz, S.A., Scharinger, M., Obleser, J., 2014a. Alpha and theta brain oscillations index dissociable processes in spoken word recognition. Neuroimage 97, 387–395. https://doi.org/10.1016/j.neuroimage.2014.04.005.

Strauß, A., Wöstmann, M., Obleser, J., 2014b. Cortical alpha oscillations as a tool for auditory selective inhibition. Front. Hum. Neurosci. 8 https://doi.org/10.3389/ fnhum.2014.00350.

ter, Bekke, M, Drijvers, L., Holler, J., 2020. The predictive potential of hand gestures during conversation: an investigation of the timing of gestures in relation to speech. PsyArXiv. https://doi.org/10.31234/osf.io/b5zq7.

Terporten, R., Schoffelen, J.-M., Dai, B., Hagoort, P., Kösem, A., 2019. The relation between alpha/beta oscillations and the encoding of sentence induced contextual information. Sci. Rep. 9 (1), 20255 https://doi.org/10.1038/s41598-019-56600-x.

Thornhill, D.E., Van Petten, C., 2012. Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. Int. J.

Psychophysiol. 83 (3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007.
Tremblay, A., Newman, A.J., 2015. Modeling nonlinear relationships in ERP data using mixed-effects regression with R examples. Psychophysiology 52 (1), 124–139. https://doi.org/10.1111/psyp.12299.

Tromp, J., Peeters, D., Meyer, A.S., Hagoort, P., 2018. The combined use of virtual reality and EEG to study language processing in naturalistic environments. Behav. Res. Methods 50 (2), 862–869. https://doi.org/10.3758/s13428-017-0911-9.

Van Humbeeck, N., Meghanathan, R.N., Wagemans, J., van Leeuwen, C., Nikolaev, A.R., 2018. Presaccadic EEG activity predicts visual saliency in free-viewing contour integration. Psychophysiology 55 (12), e13267. https://doi.org/10.1111/ psyp.13267.

Van Petten, C., Luka, B.J., 2012. Prediction during language comprehension: benefits, costs, and ERP components. Int. J. Psychophysiol. 83 (2), 176–190. https://doi.org/ 10.1016/j.ijpsycho.2011.09.015.

Van Rij, J., Hendriks, P., Van Rijn, H., Baayen, R.H., Wood, S.N., 2019. Analyzing the time course of pupillometric data. Trends in Hearing 23, 2331216519832483. https://doi.org/10.1177/2331216519832483.

Van Rij, J., Wieling, M., Baayen, R.H., Van Rijn, H., 2017. Itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs, vol. 2. R package version.

Wang, L., Bastiaansen, M., Yang, Y., Hagoort, P., 2012. Information structure influences depth of syntactic processing: event-related potential evidence for the chomsky illusion. PLoS One 7 (10), e47917. https://doi.org/10.1371/journal.pone.0047917.

Wang, L., Hagoort, P., Jensen, O., 2018. Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. J. Cognit. Neurosci. 30 (3), 432–447. https://doi.org/10.1162/jocn.a_01190.%M.28949823.

E. Huizeling et al.

- Wieling, M., 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. J. Phonetics 70, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002.
 Willems, R.M., Oostenveld, R., Hagoort, P., 2008a. Early decreases in alpha and gamma
- Willems, R.M., Oostenveld, R., Hagoort, P., 2008a. Early decreases in alpha and gamma band power distinguish linguistic from visual information during spoken sentence comprehension. Brain Res. 1219, 78–90. https://doi.org/10.1016/j. brainres.2008.04.065.
- Willems, R.M., Özyürek, A., Hagoort, P., 2008b. Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. J. Cognit. Neurosci. 20 (7), 1235–1249. https://doi.org/10.1162/jocn.2008.20085.
- Winter, B., Wieling, M., 2016. How to analyze linguistic change using mixed models, growth curve analysis and generalized additive modeling. Journal of Language Evolution 1 (1), 7–18.
- Wood, S.N., 2017. Generalized Additive Models: an Introduction with R. Chapman and Hall/CRC.