Brain and grammar: revealing electrophysiological basic structures with competing statistical models

Andrea Cometa[+,1,2,3], Chiara Battaglini[+,4], Fiorenzo Artoni[5], Matteo Greco[3], Robert Frank[6], Claudia Repetto[7], Franco Bottoni[8], Stefano F Cappa[3,9], Silvestro Micera[2,10], Emiliano Ricciardi[*,1], Andrea Moro[*,3]

+ These authors contributed equally to the work

* These authors share senior author responsibility

Corresponding author: Andrea Cometa, andrea.cometa@imtlucca.it

1 MoMiLab, IMT School for Advanced Studies Lucca, Piazza S.Francesco, 19, 55100 Lucca, Italy

2 The BioRobotics Institute and Department of Excellence in Robotics and AI, Scuola Superiore Sant'Anna, Viale Rinaldo Piaggio 34, Pontedera 56025, Italy

3 Cognitive Neuroscience (ICoN) Center, University School for Advanced Studies IUSS, Piazza Vittoria 15, Pavia 27100, Italy

4 Neurolinguistics and Experimental Pragmatics (NEP) Lab, University School for Advanced Studies IUSS Pavia, Piazza della Vittoria 15, 27100 Pavia, Italy

5 Department of Clinical Neurosciences, Faculty of Medicine, University of Geneva, 1, rue Michel-Servet, 1211 Genéve, Switzerland

6 Department of Linguistics, Yale University, 370 Temple St, New Haven, CT 06511, Unites States

7 Department of Psychology, Università Cattolica del Sacro Cuore, Largo A. Gemelli 1, Milan 20123, Italy

8 Istituto Clinico Humanitas, IRCCS, Via Alessandro Manzoni 56, Rozzano 20089, Italy

9 IRCCS Mondino Foundation National Institute of Neurology, Via Mondino 2, Pavia 27100, Italy

10 Bertarelli Foundation Chair in Translational NeuroEngineering, Center for Neuroprosthetics and School of Engineering, Ecole Polytechnique Federale de Lausanne, Campus Biotech, Chemin des Mines 9, Geneva, GE CH 1202, Switzerland

Abstract: Acoustic, lexical and syntactic information are simultaneously processed in the brain requiring complex strategies to distinguish their electrophysiological activity. Capitalizing on previous works which factor out acoustic information, we could concentrate on the lexical and syntactic contribution to language processing by testing competing statistical models. We exploited EEG recordings and compared different surprisal models selectively involving lexical information, part of speech or syntactic structures in various combinations. EEG responses were recorded in 32 participants during listening to affirmative active declarative sentences. We compared the activation corresponding to basic syntactic structures, such as noun phrases vs verb phrases. Lexical and syntactic processing activates different frequency bands, partially different time windows and different networks. Moreover, surprisal models based on part of speech inventory only do not explain well the electrophysiological data, while those including syntactic information do. By disentangling acoustic, lexical, and syntactic information, we demonstrated differential brain sensitivity to syntactic information. These results confirm and extend previous measures obtained with intracranial recordings, supporting our hypothesis that syntactic structures are crucial in neural language processing. This study provides a detailed understanding of how the brain processes syntactic information, highlighting the importance of syntactic surprisal in shaping neural responses during language comprehension.

Introduction

Different methods have been applied to the investigation of the brain basis of syntax. The classic anatomo-clinical approach, based on the study of lesion effects, has contributed to the definition of a set of brain areas required for sentence production and comprehension (Cappa, 2012; Fahey et al., 2024). Most of the lesion studies are based on the comparison of the performance in understanding non-canonical vs. canonical sentences, a task which is confounded by differences in lexical access, semantic processing, and working memory between the two sentence types. The results of these studies point to a crucial role of left temporo-parietal areas, involved in the computation of hierarchical structure for both production and comprehension, while the frontal component plays a major role in the production of sequential morphosyntactic structures (Matchin & Hickok, 2020). Functional imaging methods have provided additional information towards the localization of the syntactic component during sentence processing. Many functional magnetic resonance studies are based on the comparison between the brain response to word lists and to sentences, showing activations in the left inferior frontal gyrus, anterior and posterior temporal lobe, angular gyrus and temporo-parietal junction. (Matchin et al., 2017) compared three levels of structure: unstructured lists, two-word phrases, and simple, short sentences; and two levels of content: natural stimuli with real words and stimuli with open-class items replaced with pseudowords (jabberwocky). Inferior frontal gyrus and posterior superior temporal sulcus activation increase for both natural and jabberwocky sentences compared to unstructured word lists and two-word phrases, but not to word lists. More recently, tasks have been developed with the aim to define in additional detail the contribution of these different areas to syntactic processing during sentence comprehension, based on surprisal effects. Surprisal is defined as the negative log probability of a word in a context, which yields an inverse relationship between a word's probability and its surprisal value (Roark, 2001): the rarer a word is in a given context, the higher the surprisal. Surprisal is known to be positively correlated with brain activity (Henderson et al., 2016). Brennan et al. used the metric of surprisal to quantify complexity, correlating it with the blood oxygenation level dependent (BOLD) signal measured while subjects listened to stories (J. R. Brennan et al., 2016). The main result of the study is that hierarchical grammars predict time courses from left anterior and posterior temporal lobe, while Markov models are predictive across a broader network, including the inferior frontal gyrus. A recent work compared 5-gram and a hierarchical probabilistic context-free grammar (PCFG) surprisal (Shain et al., 2020). The effects were found only in language areas, for both types of surprisal, with small differences in specific regions. Another work tested if combinatory categorial grammars (CCGs) provide a better fit to MR results (story listening) than context-free grammars (CFG) (estimates of next-word predictability from a transformer neural network language model) (Stanojević et al., 2023). The main result is the unique contributions of CCG structure-building, predominant in the left posterior temporal lobe. In general CCG-derived measures offered a superior fit to neural signals compared to those derived from a CFG. These effects were spatially distinct from bilateral superior temporal effects uniquely related to predictability.

While lesion studies and functional imaging provide complementary evidence about the localization of linguistic processing, the dynamic analysis of speech perception requires the temporal resolution of neurophysiological methods. The electrophysiological correlates of sentence comprehension have been investigated with multiple methodologies. Oscillations in auditory cortex entrain (phase-lock) to temporally-regular acoustic cues (Lakatos et al., 2005). (Luo & Poeppel, 2007) showed that phase tracking in the theta range (4-8 Hz) is associated with sentences and correlated with comprehension. These results suggest that a 200ms temporal window is segmenting the input through resetting and sliding at a frequency corresponding to the syllable range, and corresponds to the temporal envelope. Additional oscillatory activity is present in the fine structure at high frequency, corresponding to the phonemic level, and at lower frequencies, in the delta range, corresponding to lexical and phrasal integration (Poeppel, 2003). According to (Giraud & Poeppel, 2012) "the close correspondences between (sub)phonemic, syllabic and phrasal processing, on the one side, and gamma, theta and delta oscillations, on the other, suggest potential mechanisms for how the brain deals with the 'temporal administrivia' that underpin speech perception" "speech temporally organizes (resets) oscillatory activity that is already visible at rest, but only in specific frequency domains corresponding to the sampling rates optimal for phonemic and syllabic sampling". Intracortical recording confirmed that during listening to connected speech, cortical activity of different timescales tracks the time course of hierarchically organized linguistic structures, with low frequency activity in the theta and delta range reflecting the phrasal and sentential levels. Electrocorticography provided evidence that gamma activity in human auditory cortex robustly tracks the speech envelope (Kubanek et al., 2013).

Another landmark study (Peelle et al., 2013) manipulated sentences with gradual vocoding, reducing intelligibility but maintaining the envelope: entrainment was maximal in the theta range, and was related to intelligibility. The authors propose that this multiscale pattern of co-occurring oscillatory activities is a mechanism transforming the hierarchical embedding of linguistic structures into hierarchical embedding of neural dynamics, responsible of the integration of information in time (Ding et al., 2015).

During sentence comprehension, syntactic information is crucially intertwined with acoustic and lexical-semantic information (Ding et al., 2015; Kayne, 2019; Magrassi et al., 2015). This makes it difficult to isolate the neural correlates of syntactic processing. To address the possible confound of differences in acoustic information we designed a set of stimuli composed of sentences containing homophonous phrases (Artoni et al., 2020). These homophonous phrases have the same acoustic content but different syntactic structures, i.e., they could be either noun phrases (NPs – article + noun) or verb phrases (VPs – clitic (pronoun) + verb). This was made possible by exploiting three characteristics of the Italian language: (i) some definite articles are pronounced exactly like some object clitic pronouns (such as [la] written as *la*; it can be both "the - fem.sing." or "her - fem.sing."); (ii) the syntax of articles and clitic pronouns is very different: they both precede the noun/verb, but usually complements follow verbs. The operation of placing a pronoun before the verb is called *cliticization* (Moro, 2016). The placement of the clitic pronoun before the verb has been taken to implicate a complex syntactic operation, which is absent in NPs. And (iii) the Italian lexicon contains several homophonous pairs of nouns and verbs, such as [ˈpɔrta] (written *porta*), which can either mean "door" or "brings". Pairs of words such as [laˈpɔrta] (written as *la porta*) can thus be interpreted either as a noun phrase ("the door") or a verb phrase ("brings her") depending on the syntactic context (homophonous phrases - HPs).

In our previous works (Artoni et al., 2020; Cometa et al., 2023), we compared the brain activity elicited by the processing of ambiguous HPs that can be interpreted as noun phrases or verb phrases depending on. the syntactic context, using stereo-electroencephalographic (SEEG) recordings. We found that the frequencies in the high-gamma band (150-300 Hz) were the main neural correlate of syntactic processing. We also observed a higher number of responsive contacts for VPs than for NPs, with the neural network supporting the processing of VPs being wider than the network processing NPs, and involving more cortical and subcortical areas, especially in the right (non-dominant) hemisphere. On that basis, we concluded that an abstract process of syntactic structure building is reflected in and therefore implicated by brain activity. However, there is an alternative interpretation of these results that does not tie them to the details of syntactic structure, deriving them instead from a quantitative measure called *surprisal* — the statistical expectation of linguistic material in a particular context, as quantified by its negative log probability (Attneave, 1959). A range of work has demonstrated that surprisal measures can explain behavioral measures in on-line sentence processing (Hale, 2001; Levy, 2011), surprisal has also been found to be positively correlated with brain activity (Henderson et al., 2016). VP experimental items have generally higher surprisal than the same HPs in NP items. For example, given the context 'Pulisce __' (He/She cleans __), the next linguistic element is very likely to be (the start of) a noun phrase, (i.e., it favors the NP interpretation of the HP 'la porta'). Since a determiner is a very likely start of an NP and 'la' is a highly probably determiner, its surprisal will be low in this context. In contrast, in a VP-favoring context like 'La donna __' (the woman __), the following linguistic element is likely to be (the start of) a verb phrase. Verb phrases begin with a range of words, from auxiliary and main verbs to (most relevant to our case) clitic pronouns. However, the latter option is much less likely in this VP context as compared to a determiner are in the NP context, and 'la' is less likely as a clitic than it is as a determiner than it is a clitic. This yields a higher surprisal in VP than NP contexts. Thus it would seem that the greater activity in the VP items found in (Artoni et al., 2020; Cometa et al., 2023) can be explained without appeal to explicit syntactic structure building (Henderson et al., 2016; Roark, 2001). It is worth noting however, that such surprisal-based explanations, while not directly referencing linguistic structure, are not necessarily in conflict. The determination of surprisal depends crucially on the way in which a linguistic unit's probability is determined, and the probability model that does this may or may not make reference to abstract linguistic representations. If the probabilities that give rise to surprisal are assigned on the basis of linguistic structure, then the viability of surprisal would in fact provide support for the role of linguistic structure, even if its effects are exerted through a quantitative bottleneck.

There are (at least) two orthogonal dimensions of linguistic representation a language probability model might incorporate (or not): (i) a model can represent the context using hierarchical structure (e.g., probabilistic context-free grammars) (Hale, 2001; Roark et al., 2009), or as an unstructured linear sequence (e.g., n-gram models) (Shannon, 1948); and (ii) a model can represent the units that assigned probability as instances of an abstract linguistic category like its part-of-speech (POS) or as unrelated items, i.e., separate words. The use of an abstract category like POS in a probability model allows generalizations across related words and functions as a strategy to mitigate the estimation problem arising from the difficulty of obtaining accurate estimates of the probability of infrequent words in infrequent contexts. (Greco et al., 2023) showed that only a probability model that incorporates hierarchical structure and makes predictions of abstract POS categories is able to account for the linguistic distinctions between VP and NP contexts in the stimuli from (Artoni et al., 2020) and (Cometa et al., 2023). For this reason (Greco et al., 2023) concluded that surprisal models must therefore incorporate syntactic structure to mirror human listeners' linguistic competence. This aligns with previous works that have compared the relative impact of subsets of the aforementioned linguistic dimensions of probability models on the utility of surprisal measures to predict both behavior (Fossum & Levy, 2012; Roark et al., 2009; Van Schijndel & Schuler, 2015) and patterns of brain response in fMRI and to a lesser extent EEG (J. R. Brennan & Hale, 2019; N. P. Brennan et al., 2016; Henderson et al., 2016; Lopopolo et al., 2017; Shain et al., 2020). These studies have generally (though not universally – see (Frank et al., 2015)) found consistent effects of hierarchical structure and abstract categorization (alongside linear and lexical effects). However, perhaps for reasons of statistical power and general interest in "naturalizing" psychological experimentation (Hamilton & Huth, 2020), these previous brain studies have eschewed controlled experiments in favor of large datasets derived from naturalistic reading or listening. This means that other factors, ranging from phonological to semantic to information structure, that are not controlled may be playing a role in determining processing difficulty. Thus, there is a need to solidify these findings through comparisons of identical stimuli which differ only along dimensions of linguistic structure.

A recent body of work has found that the surprisal values derived from large language models (LLMs), like those that underlie AI agents like ChatGPT, Gemini and Claude, are able to model neural activity during sentence processing (Goldstein et al., 2022; Heilbron et al., 2022; Russo et al., 2022). On the face of it, these conclusions appear to challenge the necessity for linguistic structure in explaining the details of linguistic cognitive computation. However, it is important to recall that the mechanisms governing the behaviour of such models remain largely opaque: it is not obvious from their inspection exactly what kind of representations the models are constructing. Investigations into the internal representations of some LLMs have revealed that they do in fact encode linguistic abstractions, encompassing both grammatical categories and hierarchical structures (Lin et al., 2019; Manning et al., 2020; Tenney et al., 2019). Nonetheless, given their complexity, pinpointing the precise role played by these abstractions in computing word probabilities is virtually unfeasible, leading us to set aside these models for the present discussion.

To evaluate whether syntactic surprisal modulations are similarly mirrored in brain data and provide an electrophysiological analysis of the theoretical conclusions reached in (Greco et al., 2023), in this paper, we used a new set of auditory stimuli containing homophonous sentences. In this new set of stimuli, the predictability of the syntactic content of the homophonous phrase is considered, allowing us to produce a high number of analytical contrasts among stimuli features (predictability, homophonous phrase type, and surprisal), to refine the knowledge of how syntactic information is processed in our brain, and how this neural processing is linked to the lexical and the syntactic surprisal. We presented this new set of auditory stimuli to 32 healthy participants while recording their electroencephalographic (EEG) signal. Participants listened to affirmative active declarative sentences designed to vary in syntactic predictablity, featuring noun phrases (NPs) and verb phrases (VPs). We anticipate investigating into the correlation between various surprisal models and the syntactic modulation of our stimuli. To evaluate the neural processing of syntactic structures, we used both Representational Similarity Analysis (RSA) and linear modeling on the time-frequency transformed EEG data. RSA assesses the pattern similarity of neural responses to different stimuli across the entire time-frequency space (Kriegeskorte et al., 2008). This method allows us to understand how different syntactic structures are represented in the brain by capturing complex, high-dimensional relationships between neural responses. In contrast, linear modeling directly quantifies the effects of different predictors, such as lexical and syntactic surprisal, on the EEG signal. This approach allows us to test specific hypotheses about the influence of these predictors on neural activity and to identify the precise timing and localization of these effects. Using both RSA and linear modeling permits to take advantage of the unique strengths of each method. RSA provides a richer understanding of the overall structure and organization of neural representations, while linear modeling offers precise quantification and hypothesis testing regarding the influence of syntactic and lexical surprisal. We expect that syntactic predictability influences the neural processing of syntactic structures, both in the number of EEG electrodes affected and in the timing of these effects, and that each type of surprisal model has an influence on brain activity, even though in distinct manners and locations. We hypothesize that surprisal models incorporating both syntactic and morphological information (POS), will exhibit greater accuracy in discerning the neural activity associated with syntactic processing. This anticipation arises from our observation, supported by mathematical models, that these models uniquely discriminate our stimuli based on predictability (Greco et al., 2023). More importantly, we expect the temporal dynamics of the neural activity divergence between NPs and VPs to be highly affected by the predictability of the syntactic structure. Finally, we aim to replicate our prior findings using SEEG and a simplified set of stimuli, eliminating consideration for the predictability of syntactic structure (Artoni et al., 2020; Cometa et al., 2023). This comprehensive exploration promises to deepen our understanding of the electrophysiological correlates of syntactic processing.

Methods

Stimuli

To modulate the relation between the syntactic and surprisal information we crucially relied on the paradigm introduced in (Greco et al., 2023). More specifically, three experimental conditions have been generated here by modulating the syntactic context preceding the HPs, which predicts the syntactic type of the HPs:

- **Unpredictable HPs** (Unpred.): the syntactic context preceding HPs is an adverb. Thus, the syntactic category of the HP is not predictable as the context allows both NPs and VPs. The syntactic category of the HP becomes discernible only after the HP: if it is followed by a verb, it is a NP (such as in *Forse* ***la porta*** *è aperta*, 'Maybe **the door** is open'): otherwise, it is very likely a VP (*Forse* ***la porta*** *a casa*, 'Maybe **s/he brings it** at home'). No differences will exist in the lexical surprisal values at the two HPs because the context preceding the HP is the same for both syntactic categories.
- **Early predictable HPs** (E. Pred.): the syntactic type of the HP is predictable at its onset. If the syntactic context preceding the HP is a verb the HP can only be a NP (such as in *Pulisce* ***la porta*** *con l'acqua*, 'S/he cleans **the door** with water'). If the HP is preceded by a noun, it can only be a VP (*La donna* ***la porta*** *domani, 'The woman* ***brings her*** *tomorrow'*). Our previous works (Artoni et al., 2020; Cometa et al., 2023) exploited only this type of stimuli. The different lexical context preceding NPs and VPs allows for different lexical surprisal values.
- **Late predictable HPs** (L. Pred.): this is the mixed class. The sentences are introduced by a temporal adverb requiring a past tense (e.g., *Yesterday*). Thus, the first word of the HP (*la*) could either be an article ('the') or a clitic pronoun ('her'), as in the Unpred. case; while the second word of the HP (*porta*) can only be a noun ('door'), since the verbal form of the VP would involve a present tense verb ([s/he] 'brings') that is incompatible with the temporal adverb requiring a past tense (i.e. yesterday). An example stimulus is: *Ieri* ***la porta*** *era aperta*, ('Yesterday **the door** was opened'). Using this structure, in Italian, L. Pred. VP sentences are impossible, and thus L. Pred. HPs could only be NPs.

Syntactic predictability	Syntactic class	Example			Duration (mean ± sd) [s]
Unpredictable HPs (Unpred.)	NP	Forse (maybe 'Maybe the door is open'	la porta the door	è aperta is open)	3.42 ± 0.25
	VP	Forse (maybe 'Maybe s/he brings it at home'	la porta s/he brings it	a casa at home)	3.48 ± 0.22
Early predictable HPs (E. Pred.)	NP	Pulisce (s/he cleans 'S/he cleans the door with water'	la porta the door	con l'acqua with water)	3.60 ± 0.25
	VP	La donna (the woman 'The woman brings her tomorrow'	la porta brings her	domani tomorrow')	3.59 ± 0.18
Late predictable HPs (L. Pred.)	NP	Ieri (yesterday 'Yesterday the door was opened'	la porta the door	era aperta was opened	3.63 ± 0.29

Table 1. Example stimuli. The first column is the syntactic predictability, the second column is the syntactic class of the hophonous phrase. In the third column there are some example stimuli in Italina with their English translation. The fourth column shows the mean duration of the stimuli (± the standard deviation) in seconds. HPs: homophonous phrases, NP: noun phrase, VP: verb phrase.

A total of 150 trials were prepared: 60 for Unpred. HPs, 30 NPs and 30 VPs, 60 for E. Pred. HPs, 30 NPs and 30 VPs, and 30 for L. Pred. HPs, only NPs since there cannot be VPs of this type. **Table 1** shows one example stimulus for each of the five conditions, with their mean duration.

The stimuli were selected in order to balance the mean duration and some important psycholinguistic and psychological dimensions, i.e., the frequency, familiarity, imageability, concreteness, emotional valence, arousal, and motor and sensory dimensions (Montefinese et al., 2014; Repetto et al., 2023).

The acoustic stimuli were recorded by using a Sennheiser Microphone MH40P48, Sound Card: Motu Ultralight Mk3, Connection: Firewire 400, Computer: Apple OSX. The stimuli were edited by using Audiodesk 3.02 and mastered using Peak Pro7. Files were generated in 16 bit, 44.1 kHz (Sampling Frequency); intensity was normalized to 0 Db and rendered in .wav format. All sentences were read by the same person: Italian native speaker, male, 59 years old.

Surprisal calculation

Surprisal calculation is based upon language probability models. Briefly, language probability models can be distinguished along two dimensions (Greco et al., 2023):

- Structure: (i) Linear models (which we will instantiate as *n-gram models*) view language as an unstructured sequence. In such a model, the probability of an element is determined by the linear sequence of *n* elements before it, and the probability of the entire sequence is the product of the individual elements. (ii) Hierarchical models (which we will instantiate as probabilistic context-free grammars), assume language is structured hierarchically, following (Chomsky, 1957) so that probabilities are assigned to each element on the basis of the structural configuration it occurs, which can potentially span linearly unbounded distances.
- Prediction: (i) Word models predict individual words' probabilities. (ii) Category models predict POS categories' probabilities. Both can be easily computed using n-gram models. (Roark et al., 2009) show how word and category predictions can be separated in a hierarchical model, and we use their techniques to calculate syntactic surprisal for categories and lexical surprisal for words.

For both n-gram and structural models, we estimate their parameters using a 10M word subset of the PAISÀ corpus of Italian text (Lyding et al., 2014), for which constituency parses and POS tags were computed with the Stanza package (Qi et al., 2020).

For subsequent analysis, we use only the surprisal values associated with the article (if NP) or clitic (if VP) of the HPs. As already demonstrated in (Artoni et al., 2020; Greco et al., 2023), this is the word for which the lexical surprisal difference between NPs and VPs is maximal for E. Pred. items.

Human participants and EEG recordings

In total, 32 right-handed Italian native speakers were recruited (16 males and 16 females; median age 27, range 24-54). We selected a sample size of 32 participants, which is considered robust based on standards in similar studies (Frank et al., 2013; Meyer & Gumbert, 2018; Momenian et al., 2024; Weissbart et al., 2020). All participants retained the right to withdraw from the study at any time and received a small monetary compensation for their participation. They had no history of neurological or psychiatric conditions, normal hearing, normal or corrected vision, and no reported history of drug or alcohol abuse. All participants completed all experimental sessions. EEG recordings were carried out in a sound-isolation booth using a 65-channel HydroCel Geodesic Sensor Net (Electrical Geodesies, Inc., Eugene, OR). Electrode impedance was maintained below 30 kΩ throughout the recording session. Participants were instructed to listen carefully to the sentences to be able to answer questions about them. After reading instructions on a screen, they listened to the stimuli. The stimuli were administered four times, for a total of 600 trials. More specifically, all the 150 sentences were presented once for each repetition block, in a randomized order. For each trial, the following events were annotated: the start of the sentence, the start of the HP, the start of the noun/verb of the HP, and the start of the first word following the HP. At the end of each repetition block, participants were asked two questions about two sentences (mean percentage of correct answers, across both questions: 50%) (**Figure 1**). The 8 questions (2 questions for each of the 4 repetition blocks) were the same across all participants, but different across repetition blocks. The answer was produced orally by the participants. These questions were not intended as rigorous attention checks but were included as occasional prompts to encourage participants to attend to the stimuli. Given the infrequency of the questions (every 150 sentences), we did not expect them to provide a precise measure of continuous attention. Instead, these prompts were designed to minimally interrupt the participants' engagement with the linguistic stimuli, while still providing some level of encouragement to maintain focus throughout the experiment. Participants were offered a break every 75 trials. The experiment lasted about one hour. EEG was acquired at 500 Hz. The present study received the approval of the Joint Ethics Committee of the Scuola Normale Superiore and the Scuola Superiore Sant'Anna (protocol n. 22/2022) and informed consent was obtained.

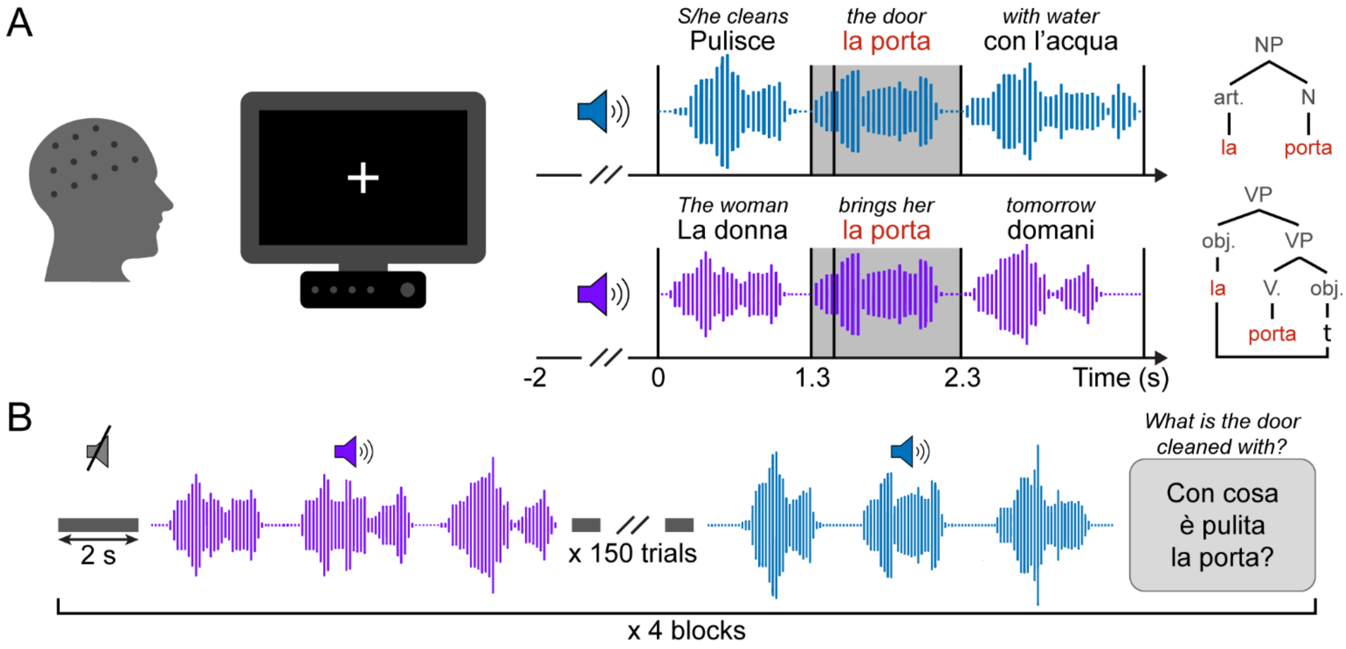

Figure 1. Recording protocol. (A) The EEG data were acquired while the participants fixated on a fixation cross displayed on a screen. After two seconds of silence, the sentence was played back through a speaker. For each sentence, its beginning, the beginning of the homophonous part, the beginning of the second word of the homophonous part, and the beginning of the word following the homophonous part were annotated. The black vertical lines on the graphs on the left represent the annotated events. The grey-shaded areas are the homophonous part. On top of each graph, there is an example stimulus sentence and its English translation. The top graph depicts an example of an early predictable noun phrase, and the bottom graph shows an early predictable verb phrase. The homophonous part (*la porta*) is highlighted in red. The syntactic trees of the example stimuli are drawn on the right of the graphs. *t* indicates the position where the pronoun is base generated in the verb phrase. **(B)** Data acquisition was divided into 4 blocks. In each block, the 150 stimulus sentences were presented in a randomized order, with 2 s of silence between one stimulus and the other. At the end of each block, the participants were asked to answer two questions regarding the content of the stimuli. The text in the grey block to the right of the panel shows one example question of the 8 administered in total.

EEG pre-processing

First, EEG data were downsampled at 250 Hz to reduce computational time. Then, 2 pre-processing steps were carried out, using a semiautomatic pipeline (Bottari et al., 2020; Martina Berto et al., 2022; Stropahl et al., 2018). The pre-processing was divided into two steps to minimize the removal of brain activity while maximizing the quality of the Independent Component Analysis (ICA) decomposition.

Pre-processing step 1

EEG data were band-pass filtered at (1 – 40 Hz) using a Hamming windowed sinc FIR filter (Widmann, 2006). Then, the signal was divided into non-overlapping windows of length equal to 1 s. The windows with a joint probability larger than three standard deviations with respect to the mean probability of occurrence of a trial were rejected. This method is used to identify and remove outliers, which are segments of data that deviate significantly from the rest of the dataset. Outliers can arise from various sources of noise or artifacts, such as muscle movements, eye blinks, or external interferences. By rejecting these windows, we aim to minimize the impact of these artifacts on our analysis, ensuring that the remaining data more accurately reflects the neural activity related to linguistic stimulation. This threshold of three standard deviations is commonly used in statistical analysis to identify extreme values that are unlikely to be part of the normal distribution of the data. Epoch rejection affected all five conditions equally (one-way ANOVA, $F(4, 155) = 0.25$, $p = 0.91$). Independent Components (ICs) were calculated using the Infomax algorithm (Bell & Sejnowski, 1995). Finally, ICs representing notable eye artifacts were rejected (Delorme & Makeig, 2004).

Pre-processing step 2

Final ICA weights resulting from Pre-processing Step 1 were applied to data more conservatively pre-processed within EEG Pre-processing Step 2. The data were epoched from 2 s before the stimulus onset to 4.5 s after and band-pass filtered (0.1 – 40 Hz) using a Hamming windowed sinc FIR filter (Widmann, 2006). Epoch length was chosen to always contain the entire stimulus. Bad channels and epochs containing high-amplitude artifacts, high-frequency noise, and other irregular artifacts were removed. Finally, bad channels were interpolated using cubic-spline interpolation (Marsden, 1974) and the EEG data were re-referenced to the average.

Event-related spectral perturbation estimation

Previous work showed that the frequency of the EEG signal activity plays an important role in syntactic processing (Artoni et al., 2020; Bastiaansen et al., 2002; Cometa et al., 2023; Grodzinsky & Friederici, 2006; Sauppe et al., 2021). Thus, we computed event-related spectral perturbations (ERSPs) to characterize the neural response to our stimuli in both frequency and time. Time-frequency transforms of each trial were normalized to the baseline (divisive baseline, ranging from -2000 ms to 0 ms before the start of the sentence), time-warped to the stimulus events, and averaged across trials for each participant, for each stimulus class to obtain the ERSPs (Grandchamp & Delorme, 2011). Time-warping refers to the process of aligning the neural responses across trials to account for temporal variability in the stimulus presentation or participant responses. We used linear time warping to adjust the timing of neural responses so that corresponding events in each trial occur at the same time point (Artoni et al., 2020; Cometa et al., 2021).

Representational similarity analysis

Representational similarity analysis (RSA) is an analysis technique used to compare the information content carried by a representation in the brain with that carried by a model (Kriegeskorte et al., 2008). This is done by comparing the representational dissimilarity matrices (RDMs) of brain activity with those computed on some features of the stimuli. The RDMs are square matrices of pairwise dissimilarity values for all pairs of stimulus-specific patterns.

More specifically, the steps of the RSA analysis are:

- ***Computation of model RDMs.*** The model RDMs are calculated on some features of the stimuli and not on the EEG data. They were calculated on several dimensions of our stimuli: the phrase type (NP or VP), the predictability (E. Pred., L. Pred., Unpred.), the lexical surprisal of the article/clitic of the HP, the syntactic surprisal of the article/clitic of the HP, the n-gram surprisal of the article/clitic of the HP, and the POS n-gram surprisal of the article/clitic of the HP (**Figure 2A** and **Figure 4A**). For the surprisal values, the RDMs were calculated using the Euclidean distance between the pairs of averages of the surprisal values for a given stimulus class. Having a total of 5 classes (Unpred. VPs and NPs, E. Pred. NPs and VPs, and L. Pred. NPs), the RDMs are 5x5 matrices. The value in row i and column j of the RDMs for the lexical and the syntactic surprisal is the difference between the mean value of the lexical (or syntactic) surprisal across all the stimuli of class i and the mean value of the lexical (or syntactic) surprisal calculated across all the stimuli of class j. The lexical surprisal and syntactic surprisal RDMs are thus matrices composed of continuous real values.

For the phrase type RDM, the Hamming distance was used, resulting in a binary RDM. The value in row *i* and column *j* of the RDM for the phrase type is 0 if the phrase type of class *i* is the same as the phrase type of class *j*, 1 otherwise.

The predictability RDM is a 3-valued matrix, with the distance between items belonging to the same class being 0, the distance between E. Pred. and Unpred. is 1, and the distance between L. Pred. and the other two classes is 0.5. The distance value of 0.5 was chosen because L. Pred. is an intermediate class between E. Pred. and Unpred.

It should be noted that the unpredictable sentences include different adverbs at the beginning, even within the same pair of homophonous sentences. As a result, the adverbs preceding the article/clitic introduce slight variations in surprisal values for the unpredictable NPs and VPs. These variations are reflected in the model RDMs, leading to non-zero Euclidean distances between Unpred. NPs and Unpred. VPs pairs. This approach allows us to capture a more refined representation of linguistic processing by accounting for the subtle differences introduced by these varying adverbs.

- ***Computation of brain* RDMs.** To calculate brain RDMs, the 5 condition-specific ERSPs were windowed in the time domain with a window length of 200 ms and an overlap of one time sample (4 ms at 250 Hz). The pairwise Euclidean distance between the ERSPs time-frequency samples of two analogous time windows was then calculated for each pair of conditions. This procedure was repeated for each time window, for each frequency, for each channel, and for each participant, resulting in channel-time-frequency-varying participant-specific RDMs.
- ***Comparison between brain RDMs and model RDMs.*** The brain RDMs were then compared to the model RDMs using the correlation coefficient as an index of similarity, resulting in a correlation value for each time-frequency point, for each model, for each channel, and for each participant.
- ***Statistical analysis on the correlation values.*** Finally, these correlation values were tested against the null hypothesis of being equal to 0 (see Statistical analysis).

Linear modeling

For each participant, we modeled the ERSPs using linear regression. The linear regression aims to model the observed neural response (time-frequency point, for each channel, for each participant) as a linear combination of different features of the stimuli. The features of the stimuli (model regressors) used were: the phrase type (NP or VP), the predictability (E. Pred., L. Pred., or Unpred.), their interaction (phrase type : predictability), the lexical surprisal of the article/clitic of the HP, and the syntactic surprisal of the article/clitic of the HP. QR decomposition was used to solve the linear model and find the regression coefficient for each regressor (Anderson et al., 1992). Specifically, the time-warped trial-by-trial time-frequency transforms of the EEG signal were used to estimate the regression coefficients for each participant, for each time-frequency point, and for each channel, resulting in 4 (one for each regressor) 4-dimensional regression coefficients. Finally, these regression coefficients were tested against the null hypothesis of being equal to 0 (see Statistical analysis). We repeated the analysis by deleting the syntactic surprisal regression term to avoid collinearity between this and the interaction between the phrase type and predictability. These two terms, by design, are highly correlated (Greco et al., 2023). This high correlation can lead to collinearity in the model, which may distort the estimation of regression coefficients and their statistical significance. By removing the syntactic surprisal term, we aim to ensure the robustness of our analysis and obtain clearer insights into the unique contributions of phrase type and predictability.

Statistical analysis

Both the correlation coefficients of the RSA, and the regression coefficient of the linear modeling have 4 dimensions: time, frequency, channels, and participants. Thus, it is possible to perform a cluster-based permutation test in the time-frequency-spatial domain (Nichols & Holmes, 2002).

The participant-specific correlation values of RSA were compared against matrices of zeros of the same size to test for the null hypothesis of zero correlation.

To increase the power of the statistical test, the regression coefficients of the linear model were first averaged in 5 frequency bands (delta: 0-4 Hz, theta: 4-8 Hz, alpha: 8-13 Hz, beta: 13-30 Hz, and gamma: 30 – 40 Hz) and 3 time windows (from the start of the sentence to the start of the HP, the HP, and from the end of the HP to the end of the sentence). The windowed regression coefficients were thus compared against the null hypothesis of their value being equal to 0, similar to the correlation values of the RSA (J. R. Brennan & Hale, 2019).

Pair-wise comparisons of NPs and VPs trials were carried out in the same way, but directly on the time-frequency transforms of the EEG data. Pair-wise comparison of E. Pred. and Unpred. items were computed using a cluster-based permutation test directly on the time-warped pre-processed EEG signal.

Results

Syntactic surprisal and syntactic class are represented by neural activity

Representational similarity analysis (RSA) was performed to investigate the effect of four dimensions of our stimuli [the lexical surprisal, the syntactic surprisal, the phrase type (NP or VP), and the predictability (E. Pred., L. Pred., and Unpred.) on the brain response, as coded by the ERSPs (i.e. time-frequency transforms).

First, we found that ERSPs power did not correlate with lexical surprisal. One cluster of significant negative correlation was found for the syntactic surprisal in the gamma band during the presentation of the homophonous parts of the stimuli (mean correlation coefficient $r = -0.17$, $p < 0.05$). This negative correlation was found between frontal-left electrodes (**Figure 2B**, top row).

For the phrase type, one cluster of significative negative correlation between the brain RDM and the model RDM was found ($r = -0.12$, $p < 0.05$). Right electrodes responded to the phrase type just after the start of the sentences, in theta band. The significance lasted up until the start of the second word of the homophonous part of the stimuli (noun or verb) (**Figure 2B**, second row).

For the predictability, two clusters of significant negative correlation were found: (i) frontal electrodes (with no evident lateralization) responded to the predictability of the stimuli after the start of the second word of the homophonous part (noun or verb), in the alpha band (r = -0.11, $p < 0.05$); (ii) frontal electrodes (with slight lateralization to the right) significantly correlated with the predictability RDM after the start of the first word following the homophonous part, with the significance being in a frequency band between alpha and beta (r = -0.12, $p < 0.05$) (**Figure 2B**, graphs three and four, from the top, one for each significant cluster).

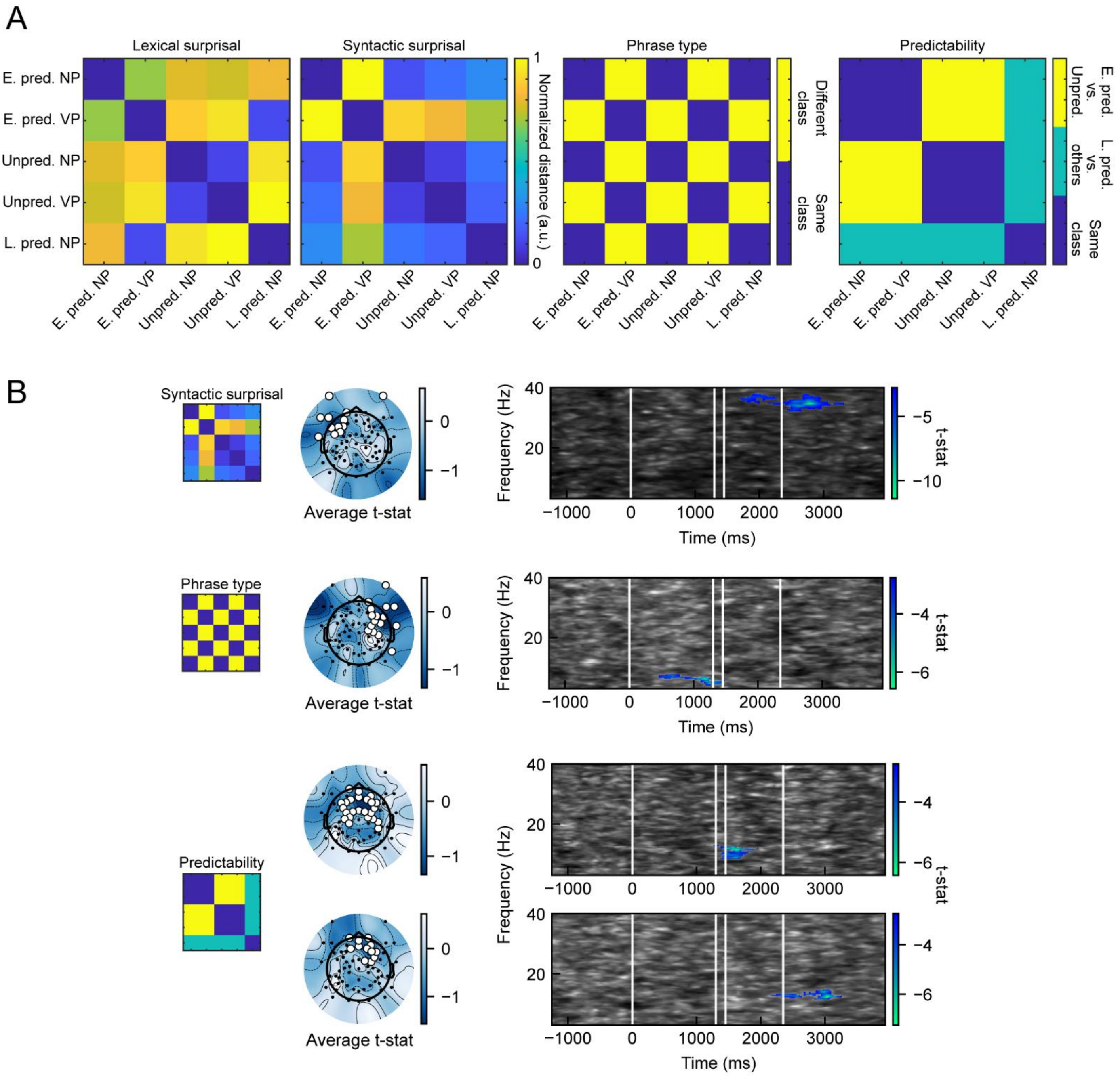

Figure 2. Representational similarity analysis.
(A) Models used in representational similarity analysis. Each of the four matrices is called Representational Dissimilarity Matrix (RDM). Each RDM is a different representation of our stimuli, along 4 dimensions: the hierarchical lexical surprisal, the syntactic surprisal, the phrase type (NP or VP), and the predictability (E. Pred, L. Pred, or Unpred.). **(B)** Significant clusters for the representational similarity analysis. The topographic plots show the average t-statistic across significant time points and frequencies. White dots represent the significant electrodes. The time-frequency graphs represent the minimum t-value across significant electrodes. Significant time-frequency points are colored in blue (top graph for the syntactic surprisal, second graph for the phrase type, and graphs three and four for the two significant clusters for the predictability). The time is adjusted according to the stimulus onset (0 ms). The four white vertical lines respectively represent: (i) stimulus onset, (ii) the start of the article/clitic, (iii) the start of the noun/verb, (iv) the start of the word that follows the noun/verb.

The neural response to the syntactic class is not dependent on lexical surprisal

RSA is not a multivariate analysis, i.e., treating all the models independently, it is not able to decouple the effect of confounding factors such as the lexical surprisal from the neural response to features of interest, such as the phrase type. Thus, we performed linear modelling, a multivariate analysis, on the ERSP power values with the syntactic surprisal, lexical surprisal, phrase type, predictability, and the interaction between phrase type and predictability as regressors. This analysis revealed only an effect of the lexical surprisal in delta band at the start of the sentences (mean regression coefficient ϱ = -0.053, $p < 0.05$) and during the presentation of HPs (ϱ = -0.043, $p < 0.05$), and in theta band at the start of the sentences (ϱ = -0.045, $p < 0.05$), on posterior electrodes (**Supplementary Figure 1**). A new linear model was treated without the syntactic surprisal under the hypothesis that the syntactic surprisal and the interaction between predictability and phrase type in the linear model are linearly correlated (Greco et al., 2023). This collinearity may cause to decrease the single effects of the syntactic surprisal and the interaction term making them (singularly) not significant. We chose to utilize the interaction between phrase type and predictability instead of syntactic surprisal because the term 'interaction' precisely describes our stimuli. Investigating the interplay between phrase type and predictability allows us to discern syntactic computation in the brain from other types of language processing. **Figure 3A** shows the results of the linear modeling without the syntactic surprisal term. This analysis revealed: (i) a significant negative regression coefficient for the phrase type in the beta band, during the homophonous part of the sentences, in central and left electrodes (ϱ = -0.098, $p < 0.05$); (ii) a significant positive regression coefficient for the interaction between phrase type and predictability in the beta band, during the homophonous phrases, in central and frontal electrodes (ϱ = 0.038, $p < 0.05$); and (iii) a significant negative regression coefficient for the lexical surprisal in the delta band, during the start of the sentences (ϱ = -0.041, $p < 0.05$) and their homophonous parts (ϱ = -0.037, $p < 0.05$), in posterior electrodes, coherently with the results obtained using the linear model which includes the syntactic surprisal. No significant effect was found for the predictability predictor alone ($p > 0.05$). Importantly, lexical surprisal correlates with brain activity in a different frequency band, a partially different time window, and different electrodes than the interaction term. To address the potential collinearity between syntactic surprisal and the interaction between predictability and phrase type, we conducted an additional analysis replacing the interaction term with syntactic surprisal. This analysis revealed that only the effect of lexical surprisal remained significant, consistent with our original findings. Additionally, an analysis with only lexical surprisal and syntactic surprisal regressors indicated a syntactic surprisal effect similar to the interaction term effect, but with a negative regression coefficient and in the gamma band, but still within high-frequency ranges (**Supplementary Figure 2**) (ϱ = -0.064, $p < 0.05$). This suggests that the collinearity between syntactic surprisal and the interaction term extends to the phrase type regressor (i.e. NP vs. VP).

Syntactic predictability modulates the response to NPs and VPs

Figure 3B displays the results of the cluster-based permutation test on ERSPs for the interaction between phrase type and predictability, broken down by predictability. For E. Pred. sentences, the contrast between NPs and VPs revealed a stronger negative response for VPs, in the beta band (beta desynchronization), during the homophonous part of the stimuli. This stronger beta desynchronization for VPs was found in the central and right electrodes. For Unpred. sentences, the difference between VPs and NPs was found in lower frequency bands, after the homophonous phrases, where syntactic interpretation is hypothesized to be carried out by the participants. The response to VPs was stronger than the response elicited by NPs, in the delta band, for almost all the EEG recording contacts. In theta band, response to Unpred. VPs was stronger than the response to Unpred. NPs on the channels over the left hemisphere.

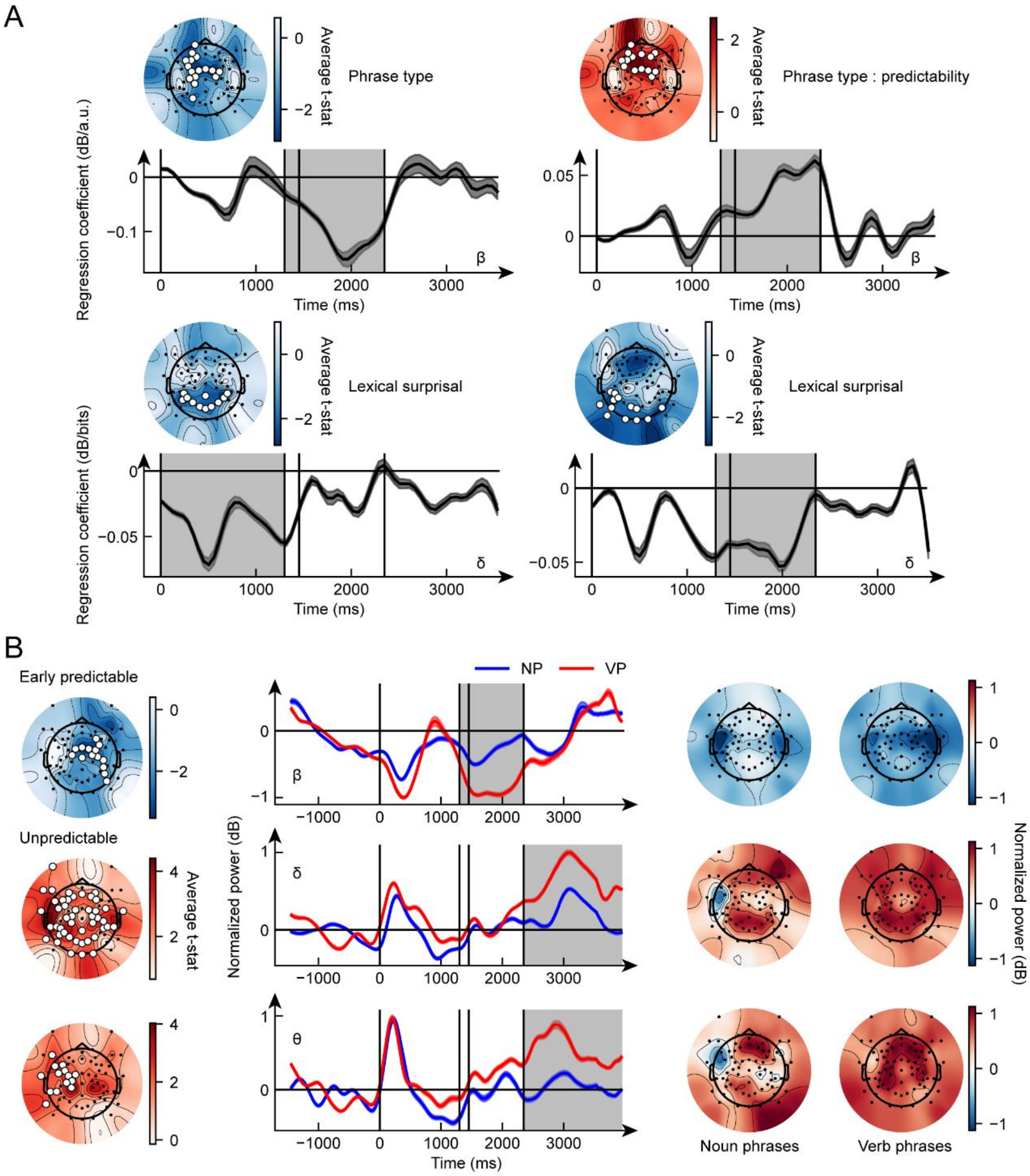

Figure 3. Noun phrases vs. verb phrases.

(A) Significant clusters ($p < 0.05$) for the linear modeling analysis. The topographic plots (head plots) show the average t-statistic across significant time points and frequencies. White dots represent the significant electrodes. The line graphs represent the temporal evolution of the coefficients of the linear model, averaged across significant electrodes, for the given frequency band. The dark grey shaded area represents the standard error across participants, light grey shaded area shows the significant time points. The time is adjusted according to the stimulus onset (0 ms). The four black vertical lines respectively represent: (i) stimulus onset, (ii) the start of the article/clitic, (iii) the start of the noun/verb, (iv) the start of the word that follows the noun/verb. The interaction between phrase type and predictability is denoted as *phrase type : predictability*. **(B)** Results of the cluster-based permutation test on the ERSPs for the contrast noun phrases vs. verb phrases, for the early predictable items (top row) and the unpredictable items (second and third rows). No contrast was done on late predictable items since there were only noun phrases. Topographic maps in the first column are the same as in (A). The line graphs represent the temporal evolution of the power of the ERPSs averaged across significant electrodes and for the given frequency bands. The last two columns represent the average power across significant time points and participants, for the given frequency band, for the sentences containing noun phrases (left) and verb phrases (right).

Distinct neural effects of sequential surprisal and syntactic structure

We repeated the RSA and the linear modelling analysis account also for the linear models of surprisal, i.e., the n-gram surprisal, and the POS n-gram surprisal (**Figure 4A**). The RSA revealed that the POS n-gram surprisal (unlike the syntactic surprisal) is not represented by the electrophysiological activity. The n-gram surprisal showed one cluster of significant negative correlation in the gamma band during the presentation of the homophonous parts of the stimuli ($r = -0.15$, $p < 0.05$) and at the end of the sentences ($r = -0.16$, $p < 0.05$). This negative correlation was found between frontal-left electrodes (**Figure 4B**) and only partially overlaps with that associated to the syntactic surprisal (**Figure 2B**).

The n-gram surprisal values and the POS n-gram surprisal values were added to the regressor for the linear modeling. Thus, the regressor for the final linear model were n-gram surprisal, POS n-gram surprisal, lexical surprisal, phrase type, predictability, and the interaction between phrase type and predictability. Both n-gram surprisal and POS n-gram surprisal significantly drive EEG activity. The POS n-gram surprisal has an effect at the start of the sentence (before the homophonous part) in delta ($\varrho = 0.096$, $p < 0.05$) and beta ($\varrho = -0.069$, $p < 0.05$) bands (**Figure 5**, top 2 graphs). The n-gram surprisal has an effect at the start of the sentence (before the homophonous part) in delta band ($\varrho = 0.049$, $p < 0.05$) (**Figure 5**, bottom-left graph). Finally, the effect of the lexical surprisal remained constant with a significant negative regression coefficient in the delta band, during the start of the sentences in posterior electrodes ($\varrho = -0.058$, $p < 0.05$) (**Figure 5**, bottom-right graph) but with no significant regression coefficient during the presentation of the HPs. In this new linear model, phrase type, predictability, and their interaction did not have a significant regression coefficient. This lack of significance is most likely due to the complexity of the linear model relative to the amount of data available. The model may be too complex to detect the weaker effects of phrase type and predictability, resulting in these effects not reaching statistical significance.

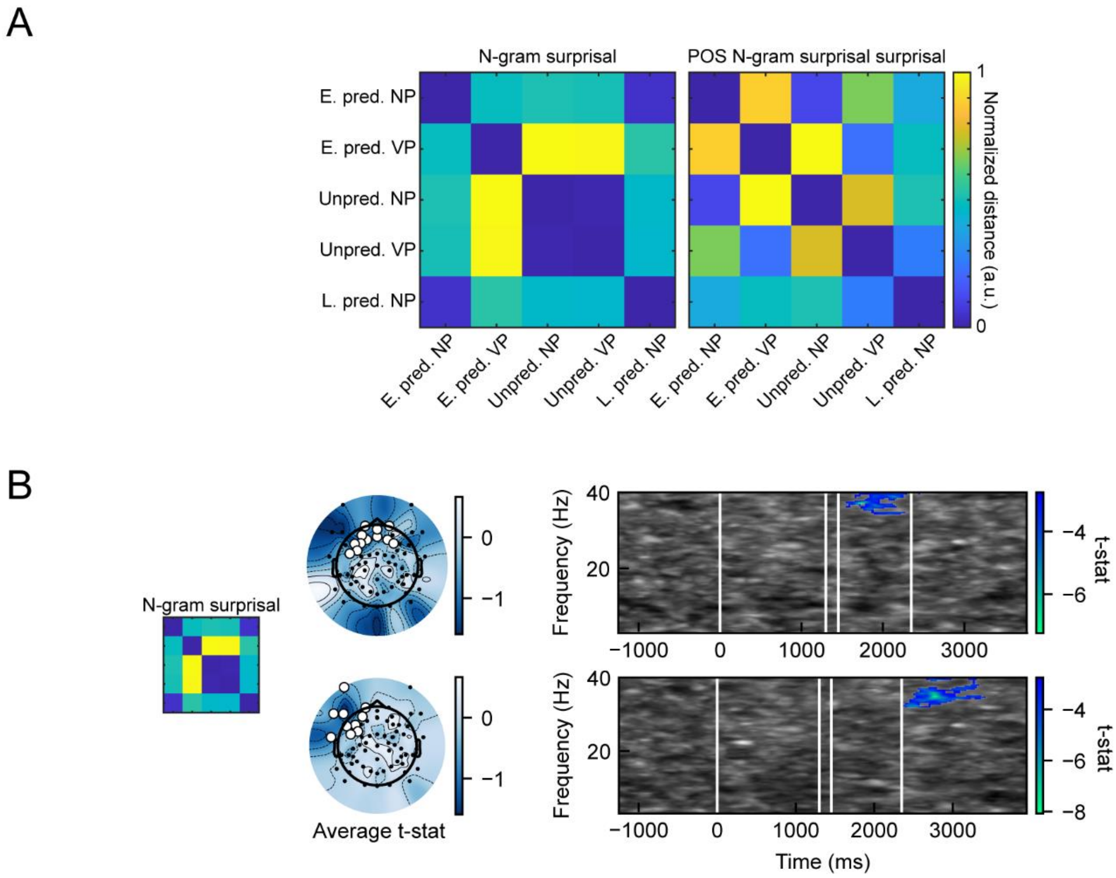

Figure 4 Representational similarity analysis of the linear surprisal.
(A) Models used in representational similarity analysis. Each of the two matrices is called Representational Dissimilarity Matrix (RDM). Each RDM is a different representation of our stimuli, along two dimensions: the n-gram surprisal, and the POS n-gram surprisal. **(B)** Significant clusters for the representational similarity analysis. The topographic plots show the average t-statistic across significant time points and frequencies. White dots represent the significant electrodes. The time-frequency graphs represent the minimum t-value across significant electrodes. Significant time-frequency points are colored in blue (both graph for the n-gram surprisal, for the two significant clusters). The time is adjusted according to the stimulus onset (0 ms). The four white vertical lines respectively represent: (i) stimulus onset, (ii) the start of the article/clitic, (iii) the start of the noun/verb, (iv) the start of the word that follows the noun/verb.

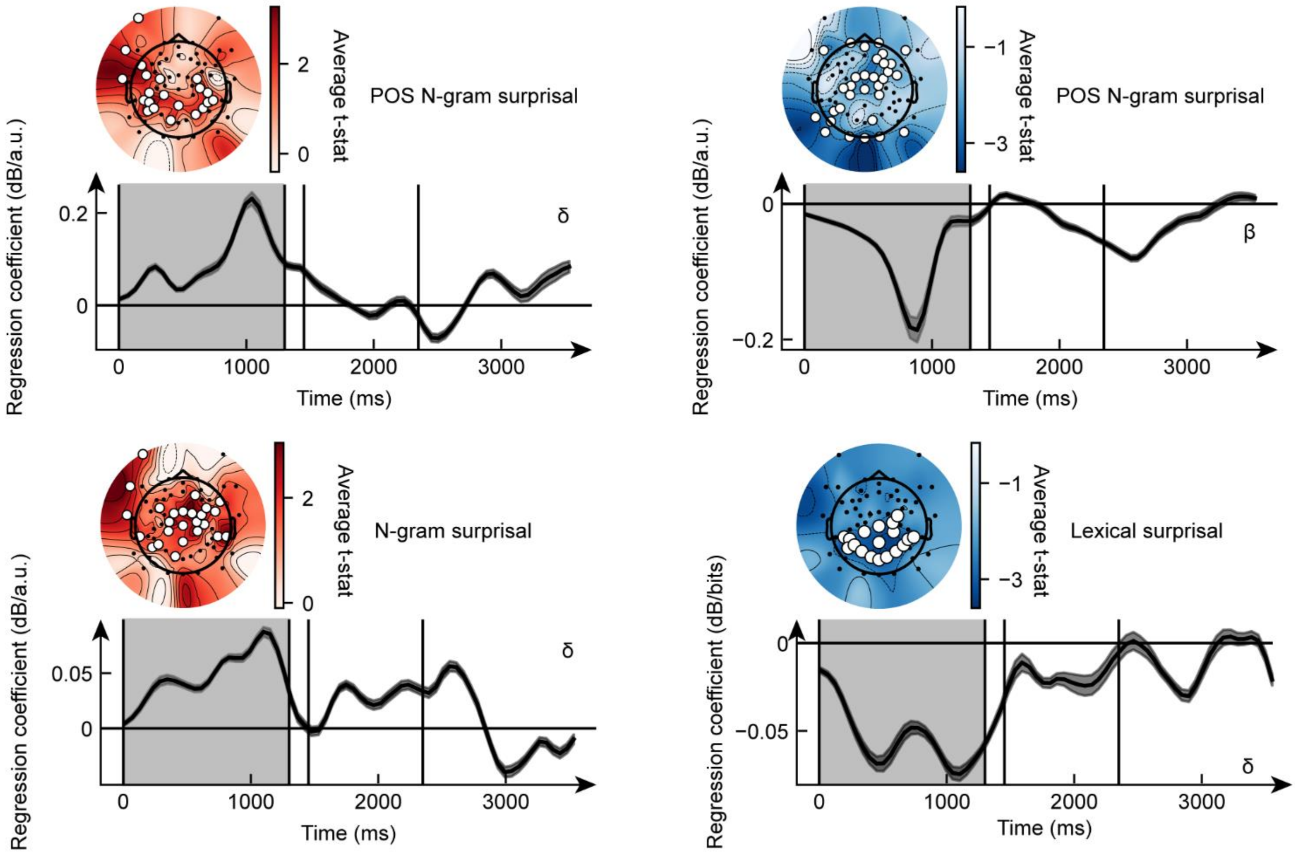

Figure 5. Linear modeling analysis with linear surprisal.
Significant clusters ($p < 0.05$) for the linear modeling analysis. The topographic plots (head plots) show the average t-statistic across significant time points and frequencies. White dots represent the significant electrodes. The line graphs represent the temporal evolution of the coefficients of the linear model, averaged across significant electrodes, for the given frequency band (Greek letter in the graph). The dark grey shaded area represents the standard error across participants, light grey shaded area shows the significant time points. The time is adjusted according to the stimulus onset (0 ms). The four black vertical lines respectively represent: (i) stimulus onset, (ii) the start of the article/clitic, (iii) the start of the noun/verb, (iv) the start of the word that follows the noun/verb.

Neural response to syntactic category does not depend on systematic confounding effects

To exclude systematic differences in potential lexico-semantic confounding factors between NPs and VPs, we used a cluster-based permutation test directly on the pre-processed EEG signal and defined two contrasts: E. Pred. NP vs. Unpred. NP, and E. Pred. VP vs. Unpred. VP. The rationale behind this analysis is that if we find the same significant effects for both contrasts, then the systematic differences between E. Pred. NPs and E. Pred. VPs that could have affected the previous results can be excluded. We do not expect systematic differences in confounding factors in Unpred. NPs and Unpred. VPs since the structure of Unpred. sentences do not allow them by design. We performed this analysis directly on the filtered data, and not on the time-frequency transformed EEG. This choice allowed us to explore the differences occurring in shorter time windows than those possible with ERSPs, due to the dilution effect of the filtering on the signal. Indeed, we expect those systematic differences to occur during short time windows.

For the E. Pred. NP versus Unpred. NP contrast, 10 clusters were found. For the E. Pred. VP vs. Unpred. VP 14 clusters were found. No differences were found during the baseline period. Six pairs of clusters were comparable across the two contrasts (**Figure 6**). Among these six pairs, two spatio-temporal significant clusters were found during the start of the sentence: one involving frontal, temporal, and posterior electrodes, denoting a higher positive potential peak for Unpred. sentences right at the start of the stimuli (**Figure 6**, first row); the other involving central electrodes and denoting a negative potential deflection for Unpred. sentences, absent in E. Pred. stimuli, right before the start of the homophonous phrase (**Figure 6**, second row). Two significant clusters, with opposite polarity, were found in the frontal and posterior regions during the homophonous phrases, indicating a stronger potential deflection for E. Pred. sentences (**Figure 6**, third and fourth rows). The last two significant clusters, again with opposite polarity, were found during the last part of the sentences, in frontal and posterior electrodes, respectively (**Figure 6**, fifth and sixth rows).

We additionally repeated the analysis using ERSPs instead of ERPs to ensure consistency with the time-frequency analysis used in our main results. Upon reanalysis with ERSPs, we did not find any significant cluster. This indicates that the original differences observed are not due to systematic lexico-semantic differences between NPs and VPs, reinforcing the robustness of our findings.

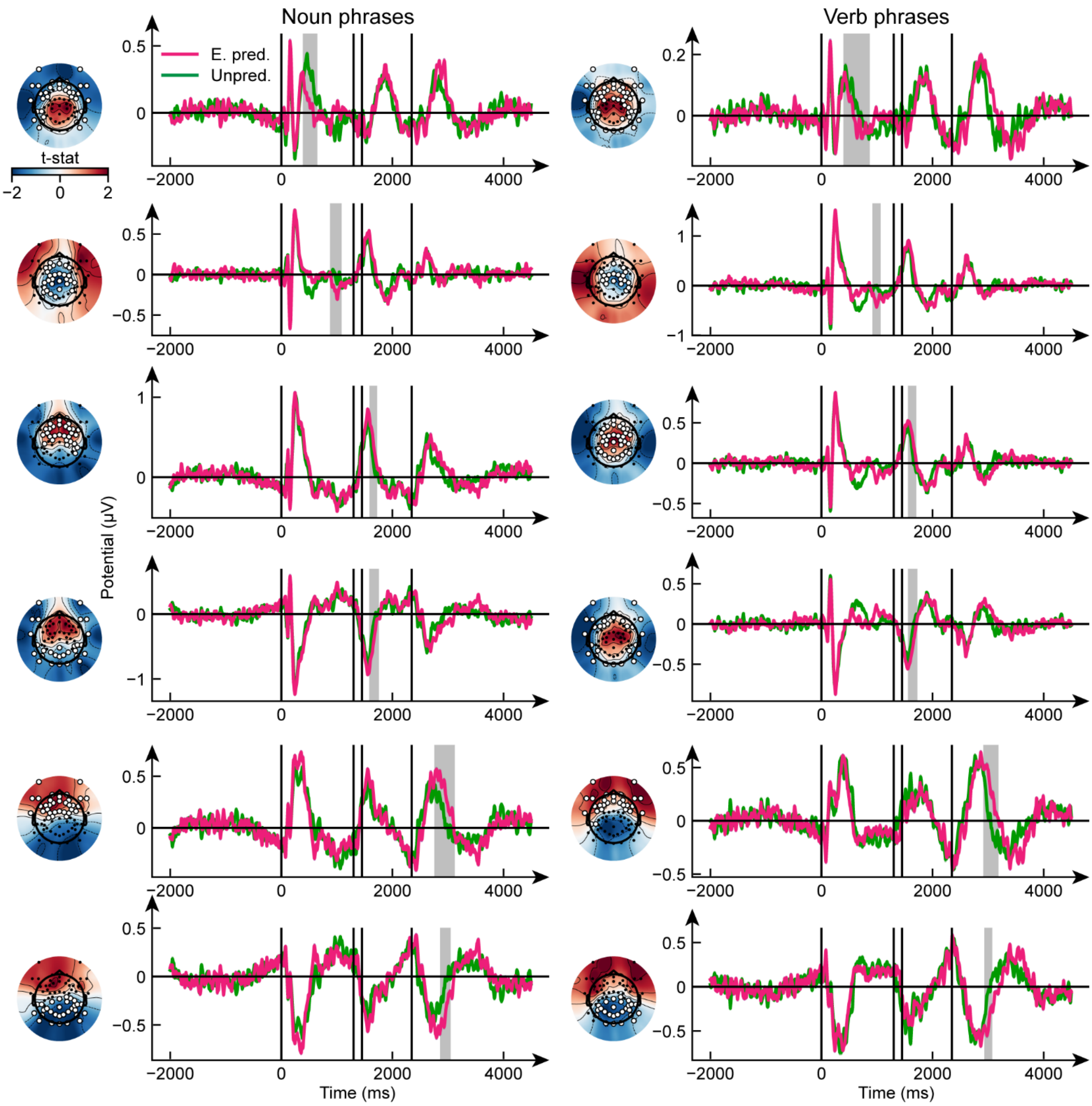

Figure 6. Early predictable vs. Unpredictable sentences

Results of the cluster-based permutation test on ERPs for the contrasts E. Pred. NP vs. Unpred. NP (left) and E. Pred. VP vs. Unpred. VP (right). The topographic plots show the average t-statistic across significant time points. The line graphs represent the temporal evolution of the EEG signal, averaged across participants and significant electrodes. The light grey shaded area shows the significant time points (cluster-based permutation test). The time is adjusted according to the stimulus onset (0 ms). The four black vertical lines respectively represent: (i) stimulus onset, (ii) the start of the article/clitic, (iii) the start of the noun/verb, (iv) the start of the word that follows the noun/verb.

Discussion

In our earlier research (Artoni et al., 2020; Cometa et al., 2023), SEEG recordings were used to compare the brain activity elicited by the processing of (homophonous) noun phrases or verb phrases. The primary neural correlate of syntactic processing was found in the increase in power of frequencies in the high-gamma band (150-300 Hz). We discovered that there were more responsive contacts for VPs than for NPs, with the neural network supporting VP processing being wider and involving more cortical and subcortical areas than the network processing NPs. Unfortunately, the concept of n-gram surprisal represented a confounding factor for these findings. The higher the rarity of a sequence of words, the higher the surprisal, which is defined as the negative log probability of a given word following another in a sentence given a corpus (Roark, 2001). It is well known that surprisal and brain activity are positively correlated (Henderson et al., 2016), and VPs were associated to a generally higher n-gram surprisal than NPs. Here, we modulated both the syntactic and lexical surprisal values of the HPs, showing that the difference in the neural processing of NPs or VPs is better mirrored by the syntactic surprisal.

Lexical predictability does not affect syntactic processing

Here we used two models of surprisal calculated using hierarchical structures: the lexical surprisal (word prediction), and the syntactic surprisal (hierarchical POS prediction). In line with (Greco et al., 2023), our analysis shows that both syntactic surprisal and lexical surprisal significantly affect the electrophysiological signal. However, it appears that syntactic surprisal provides a better explanation for the neural responses observed in our study. On the other hand, the neural correlates of lexical surprisal may be related to other aspects of language processing, such as processing difficulty during language comprehension (Brouwer et al., 2021; Frank et al., 2013), and semantic information retrieval (Russo et al., 2020). Moreover, the RSA showed that lexical surprisal failed at eliciting a significant brain response thus that our stimuli were able to induce syntactic computation while controlling for other types of language computation (**Figure 2**). However, all significant correlations between model RDMs and brain RDMs were negative. A negative correlation means that when the model predicts higher dissimilarity between two stimuli, the brain shows lower dissimilarity, and vice versa. Further investigation is needed to understand the underlying reasons for this negative correlation. It could reflect differences in how the model and brain organize and process the information, or it might indicate that the model captures some aspects of the data that are not directly aligned with the neural representations observed. The linear model showed an effect of the lexical surprisal in posterior electrodes, in the delta band, from the start of the sentence to the end of the homophonous part (**Figure 3A**). This result aligns with the large body of evidence demonstrating a crucial role of temporo-parietal areas for language comprehension (Binder et al., 2009; Heilbron et al., 2022; Liuzzi et al., 2020). Moreover, the prominence of delta oscillations in this context could be attributed to their recognized role in orchestrating long-range communication and coordinating cognitive processes, such as memory retrieval and attention allocation (Aktürk et al., 2022; Harmony, 2013). There is extensive neurophysiological evidence for the role of delta oscillations in sentence level processing, Using EEG while subjects listened to naturally spoken sentences, jabberwocky stimuli with morphemes and sentential prosody, word lists and backward acoustically matched stimuli, (Kaufeld et al., 2020) were able to show that the activity in the delta band is not prosody-dependent, but driven to the lexically-driven combinatorial process. Comparing phrases and sentences with similar meaning and high phonological similarity (De rode vaas (The red vase) and De vaas is rood (The vase is red) (Bai et al., 2022) found that in theta and delta band sentences were associated with more power, more connectivity, and more phase synchronization than phrases. The sentential response appears to reflect hierarchical organization rather that lexical properties of the words, as indicated by the study of (Lo et al., 2022), which compared syntactically correct sentences with their reversed form ("white sheep eat grass" and "sheep white grass eat"). Robust delta synchronization was found only for the correct sentences. It is noteworthy that in another EEG study (Weissbart et al., 2020) found tracking of lexical surprisal in the delta, beta and gamma bands.

This convergence of findings substantiates the notion that lexical surprisal exerts a distinct and lasting impact on neural processing. Importantly, this impact does not completely overlap with the neural activity related to syntactic processing, as the latter is prominent in different electrodes, different frequency bands, and partially different time windows.

Our findings demonstrate that while lexical surprisal was not represented in the EEG signal according to RSA (**Figure 2**), it had a significant effect in the linear modeling analysis (**Figure 3A**). This discrepancy can be attributed to the different underlying assumptions of the two methods, underscoring the importance of using multiple analytical approaches to gain a more comprehensive understanding of the neural responses to linguistic stimuli.

The electrophysiological correlates of NPs and VPs

Phrase type was represented by the observed differences in neural patterns during the start of the sentences, and the homophonous part. The significant correlation observed before the onset of the noun/verb in the phrase type RDM (**Figure 2**) can be attributed to the E. Pred. stimuli. The E. Pred. stimuli may exert a stronger effect that masks the statistical significance of the effect of Unpredictable stimuli after the noun/verb phase. This interpretation is supported by the significant electrodes being predominantly on the right side, as observed in both our current linear model and in our previous work with SEEG data. The pre-noun/verb significance suggests that the brain is already processing predictive information before the critical noun/verb appears, particularly in E. Pred. conditions where the context provides strong predictive cues early on. Verbs and nouns in isolation are known to elicit different brain responses (Lukic et al., 2021; Pulvermüller et al., 1996; Vigliocco et al., 2011). However, we introduced the concept of predictability for this purpose: eliminating confounding factors thanks to the number of contrasts that we can obtain by comparing our stimuli along different dimensions.

To this aim, we used the linear modeling analysis to isolate the effects of the phrase type, the predictability, their interaction, and the lexical surprisal (**Figure 3A**). We found an effect of the phrase type in the beta band, during the homophonous part, on central/left electrodes. The effect of the interaction between the phrase type and the predictability was in central/frontal electrodes in the beta band during the homophonous part. We broke down the effect of the phrase type by the predictability (**Figure 3B**) and we found that: (i) for the E. Pred. items, VP elicited higher beta desynchronization during the homophonous part, over central/right electrodes; and (ii) in the Unpred. case, VPs caused a delta (over most electrodes) and theta (on right electrodes) synchronization after the homophonous part of the sentence. The lack of significance of the interaction effect between the phrase type and the predictability modelled with linear regression at the end of the sentence, in delta/theta bands may be attributed to the lower statistical power of the linear regression analysis compared to the post-hoc comparison. The fact that in the E. Pred. case VPs elicited higher brain activity than NPs, during the homophonous part confirmed our previous findings using SEEG (Artoni et al., 2020). In this earlier study, only E. Pred. stimuli were used. Furthermore, SEEG and EEG signals have different characteristics. For instance, the high-gamma band is not recordable using EEG due to the low-pass filtering effect of the scalp and the skull (Ramon et al., 2009), and future work should further establish whether there is a direct link between the beta band desynchronization in EEG and high-gamma increase in SEEG. In summary, although expanding on our previous work, we have also confirmed with EEG data what we previously observed using SEEG recordings, opening new research and treatment possibilities as EEG is much simpler to record than SEEG, and comes with fewer limitations. However, if spatial resolution is of utter importance for the investigation of the neural correlates of syntactic processing, SEEG still has an advantage (Cometa et al., 2022), and future work is needed to precisely identify the cortical areas whose activity correlates with the syntactic predictability.

The latency of the difference between NPs and VPs in the Unpred. scenario (i.e., after the homophonous part) is consistent with the hypothesis on the timing in which the participant realizes what type of phrase has been heard (NP or VP). Thus, the detected activity in delta and theta bands may be an index of late disambiguation. It is not surprising that we found different bands for the two predictability scenarios. In the case of E. Pred. sentences, evidence of disambiguation is absent. This is because participants know if they are listening to a NP or a VP prior to the onset of the homophonous part. On the other hand, in the Unpred. case, participants must remain uncertain about the structure of the sentence following the homophonous phrase and should continue to show evidence for disambiguation. Thus, the late response in delta-theta bands probably reflects this process. The shift of electrophysiological activity in lower frequencies, together with the higher number of selective electrodes, may be a correlate of the higher syntactic processing effort required from the listener in the Unpred. scenario (Clarke et al., 2019; Edward M. Bernat et al., 2008; Harmony, 2013).

Of note, the unpredictable sentences shift the significant activity in the portion that follows the homophonous phrase, thus the attribution of the class of the HP (NP or VP) takes place after the phrase has been listened to.

The significant clusters for the contrasts E. Pred. NP vs. Unpred. NP, and E. Pred. VP vs. Unpred. VP in the time domain are highly superimposable (**Figure 6**). This shows that there are no systematic differences in potential confounding factors between NP and VP stimuli, in both the E. Pred. sentences and in the Unpred. sentences. Upon reanalysis with ERSPs, we did not find any significant clusters. This outcome indicates that the contrasts E. Pred. NP vs. Unpred. NP and E. Pred. VP vs. Unpred. VP are not different, which strengthens the conclusion that there are no systematic lexico-semantic differences between E. Pred. NPs and E. Pred. VPs. Additionally, the linear modeling with ERSPs did not find an effect of predictability. If such an effect were present, it would manifest as equal significant clusters in the two comparisons. However, the absence of clusters, either equal or different, is consistent with the lack of a predictability effect in our linear modeling. Thus, the results shown in **Figure 3B** and section **Syntactic predictability modulates the response to NPs and VPs** are only due to the effect of the different neural activities underlying the processing of different syntactic structures.

We did not perform any direct contrast involving L. Pred. sentences because L. Pred. VP sentences are impossible in Italian and thus the responses to NPs vs. VPs could not be compared. However, their inclusion in the experiment was fundamental for the finer modulation of syntactic surprisal values that we achieved. Without L. Pred. sentences we would have less variety in the values of surprisal and therefore a more limited ability to infer what is happening in the brain due to the values of surprisal. Nevertheless, L. Pred. sentences open many possibilities for future studies.

Syntactic processing vs. surprisal

Syntactic surprisal has a significant effect on the neural response, especially in beta and gamma bands (**Figure 2**). The electrophysiological response to the syntactic class depends on the predictability of the syntactic class, and thus on syntactic surprisal, but not on lexical surprisal, indicating that the observed response is necessarily syntactic.

Our previous paper (Greco et al., 2023) established that the predictability associated with three distinct classes of stimuli is more accurately represented in a model that incorporates syntactic information. This assertion is further corroborated by our recent findings. However, the RSA analysis conducted on the linear surprisal values revealed no significant role of surprisal models based solely on POS. This suggests that the morphological information encoded in POS alone may be insufficient, and syntax is indeed necessary to elucidate brain activation patterns. If validated, this would serve as additional evidence supporting our hypothesis.

Conversely, our analysis identified a significant role of surprisal based on n-gram models (**Figure 4**). This aligns with expectations, given the numerous studies demonstrating the impact of n-gram-based surprisal on cognitive tasks. However, our previous paper on surprisal indicated that n-gram models are not optimal for distinguishing between different classes of stimuli in terms of predictability. Therefore, it is not concerning that we observed activations correlating with both n-gram surprisal models and syntactic models. Importantly, the electrodes that were significant in the two conditions only partially overlapped, suggesting that they represent different facets of linguistic stimuli processing.

These results confirm the pivotal role of the computation of syntactic structures in human languages (Berwick & Chomsky, 2016). We showed different roles for the various EEG frequency bands (Mai et al., 2016), showing an immediate response that is syntactic specific in the beta band in the contrast between E. Pred. sentences and a late response in delta and theta bands in the contrast between unpredictable items (**Figure 3B**). This delta-theta response is an index of disambiguation of the syntactic type of the homophonous phrase after that the phrase type becomes discernible, coherently with the higher cognitive load associated with syntactic processing for the Unpred. items. Overall, our findings suggest that the processing of noun phrases and verb phrases is modulated by the syntactic surprisal as encoded by the predictability of the HPs and that there are distinct neural representations for early syntactically predictable and syntactically unpredictable stimuli (**Figure 7**).

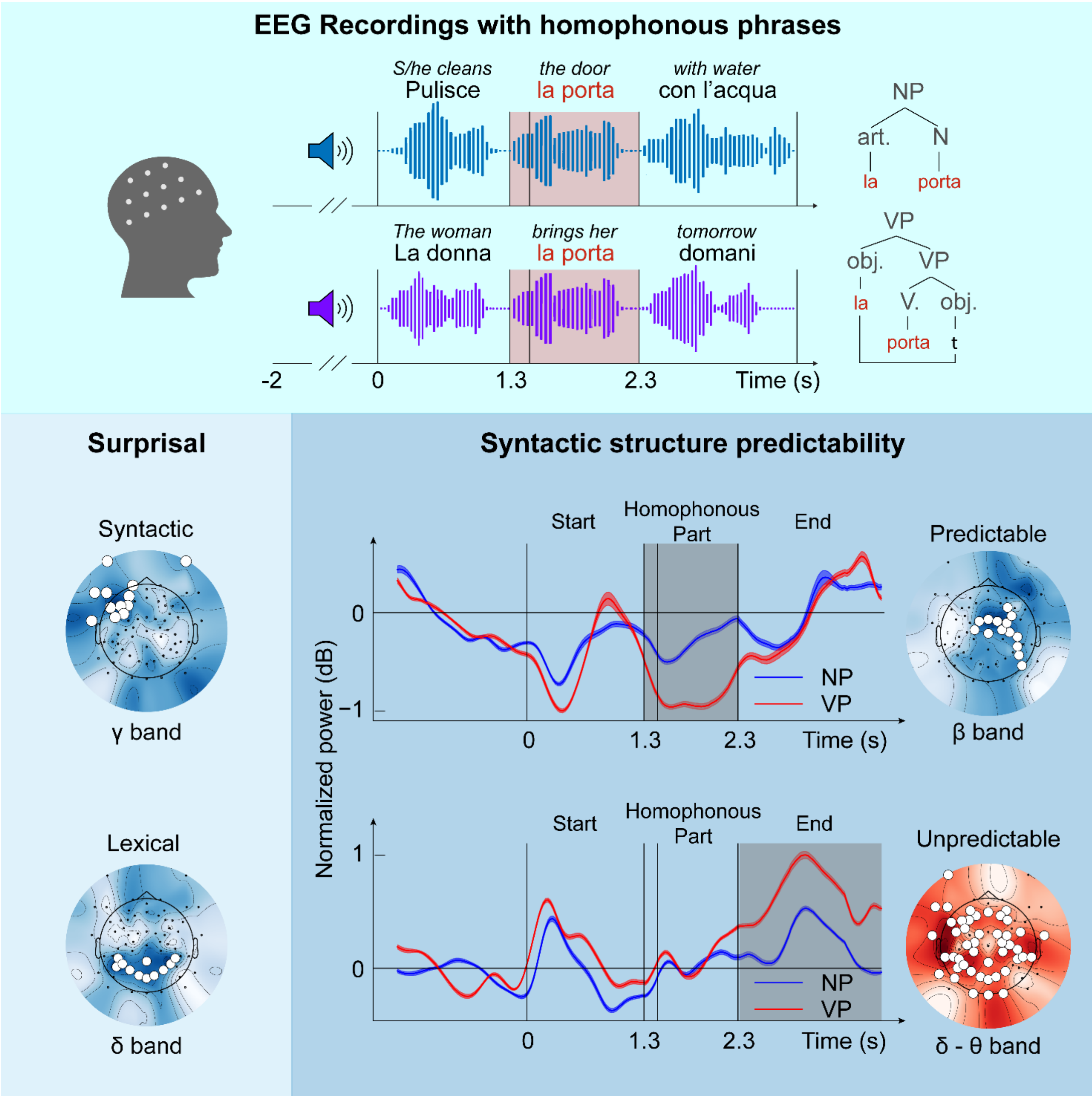

Figure 7. Main Results.
Graphical summary of the results presented in this paper. Refer to the other figures for the panel legends.

Conclusions

This paper showed that the class of predictability correlates with brain activity as predicted by mathematical models (Greco et al., 2023). The observed responses are inherently syntactic because of the distinctions in NPs and VPs contrasts in E. Pred. and Unpred. cases, the alignment of their temporal dynamics with the expected syntactic processing timings, and the exclusion of acoustic factors, coupled with the integration of surprisal and syntactic processing concepts.

In this sense, these results constitute a fundamental factor of a broader picture toward the cracking of the syntactic code of human languages. More specifically, the present results strongly correlate and provide novel evidence with two previous ones, namely: (i) the distinct electrophysiological correlates of NPs vs. VPs (Artoni et al., 2020); (ii) the distinct cortical connectivity related to NPs and VPs (Cometa et al., 2023). The whole picture provides a first electrophysiological fine-grained contrast of two basic syntactic units, namely NPs and VPs, having excluded the confounding factor of phonological information. We found activations correlating with syntactic and lexical processing in different electrodes, different frequency bands and partially different time windows. These results showed that both syntactic and lexical information are important for language processing but rely on distinct computations. Moreover, the present study showed that surprisal models based only on morphological information (the POS) do not play a significant role. Syntactic information is needed to explain brain activations.

What are the implications of these results for current neural models of sentence comprehension? The model proposed by Friederici proposes a crucial role for the posterior part of Broca's area for the core combinatorial merge operation, while the process of semantic-syntactic integration is based on the left temporo-parietal cortex (Friederici, 2017; Maran et al., 2022). The two regions are connected via the arcuate and superior longitudinal fasciculus (dorsal stream). The dynamic operation of the system during comprehension is based on a syntax-first, rule-based automatic syntactic analysis, while integration, revision processes are considered to take place on a longer scale time frame. The model proposed by Matchin and Hickok takes a different perspective, based on a distinction between comprehension and production processes (Matchin & Hickok, 2020). A hierarchical lexical-syntactic function is assigned to the posterior middle temporal gyrus, integrating speech perception and semantic systems in the temporal and inferior parietal areas, which are required for both sentence production and comprehension. In contrast, activity in the left posterior inferior frontal gyrus is required for the transformation of the relational hierarchies into morpho-syntactic sequences, a process crucial for production. Finally, Hagoort model is proposing a central role of Broca's area in the unification of information, which however is not language specific (Bickerton & Szathmáry, 2009). In addition, syntax is not considered to have a privileged role in sentence processing, which is based on spontaneous brain prediction of upcoming input at multiple levels of abstraction (Heilbron et al., 2022). Our findings, showing distinct neural effects of lexical and syntactic surprisal in different frequency bands and regions, provide support for these models. The dissociation between the effects of lexical and syntactic surprisal aligns with the idea of different neural processes and regions being involved in handling these aspects of language. However, to better integrate these findings into the existing neural models of sentence comprehension, we acknowledge the need for higher localization power than what EEG can provide. This limitation highlights the necessity for intracranial recordings to achieve more precise localization of these neural processes.

These findings contribute to the broader field of cognitive neuroscience by elucidating the temporal dynamics of language processing and highlighting the importance of predictive coding in syntactic computation., revolutionizing how we interact with linguistic computation in the brain.

Data and code availability

The data and custom *Python* and *MATLAB* code supporting these findings are available from the corresponding author upon reasonable request.

Acknowledgements

This work was supported by the Italian Ministry for Universities and Research (INSPECT PRIN 2017JPMW4F) and by the Bertarelli foundation.

Author contributions

Conceptualization: S F C, A M, and S M; Methodology: A C, C B, F A, M G, F B, C R, S F C, S M, A M, and E R; Experimental setup: C B, F B; Data Collection: A C and C B, Software: A C and C B; Validation: A C, C B, F A, and E R; Formal Analysis: A C and C B; Theoretical syntax contribution: A M; Computational linguistic contribution: R F; Lexical analysis: M G; Resources: S M and E R; Data Curation: A C and C B; Writing - Original draft: A C, C B, and M G; Writing - Review and Editing: all authors; Visualization: A C; Supervision: A M, and E R; Project Administration: S F C, A M, and S M; Funding Acquisition: S M, A M, and E R.

C B contributed equally as first author and has the right to put her name first in her CV.

Competing interests

The authors declare no competing interests.

References

Aktürk, T., de Graaf, T. A., Erdal, F., Sack, A. T., & Güntekin, B. (2022). Oscillatory delta and theta frequencies differentially support multiple items encoding to optimize memory performance during the digit span task. *NeuroImage*, *263*, 119650. https://doi.org/10.1016/j.neuroimage.2022.119650

Anderson, E., Bai, Z., & Dongarra, J. (1992). Generalized QR factorization and its applications. *Linear Algebra and Its Applications*, *162*, 243–271.

Artoni, F., d'Orio, P., Catricalà, E., Conca, F., Bottoni, F., Pelliccia, V., Sartori, I., Russo, G. L., Cappa, S. F., Micera, S., & Moro, A. (2020). High gamma response tracks different syntactic structures in homophonous phrases. *Scientific Reports*, *10*(1), 7537. https://doi.org/10.1038/s41598-020-64375-9

Attneave, F. (1959). *Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods, and Results*. Holt. https://books.google.it/books?id=VnB9AAAAMAAJ

Bai, F., Meyer, A. S., & Martin, A. E. (2022). Neural dynamics differentially encode phrases and sentences during spoken language comprehension. *PLOS Biology*, *20*(7), e3001713. https://doi.org/10.1371/journal.pbio.3001713

Bastiaansen, M. C. M., van Berkum, J. J. A., & Hagoort, P. (2002). Syntactic Processing Modulates the θ Rhythm of the Human EEG. *NeuroImage*, *17*(3), 1479–1492. https://doi.org/10.1006/nimg.2002.1275

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159.

Berwick, R. C., & Chomsky, N. (2016). *Why Only Us: Language and Evolution*. MIT Press. https://books.google.it/books?id=8eBRCwAAQBAJ

Bickerton, D., & Szathmáry, E. (Eds.). (2009). *Biological Foundations and Origin of Syntax*. The MIT Press. https://doi.org/10.7551/mitpress/9780262013567.001.0001

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelyova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., & Röder, B. (2020). EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. *NeuroImage*, *223*, 117315. https://doi.org/10.1016/j.neuroimage.2020.117315

Brennan, J. R., & Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLOS ONE*, *14*(1), e0207741. https://doi.org/10.1371/journal.pone.0207741

Brennan, J. R., Stabler, E. P., Van Wagenen, S. E., Luh, W.-M., & Hale, J. T. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*, *157*, 81–94.

Brennan, N. P., Peck, K. K., & Holodny, A. (2016). Language mapping using fMRI and direct cortical stimulation for brain tumor surgery: The good, the bad, and the questionable. *Topics in Magnetic Resonance Imaging*, *25*(1), 1–10.

Brouwer, H., Delogu, F., Venhuizen, N. J., & Crocker, M. W. (2021). Neurobehavioral Correlates of Surprisal in Language Comprehension: A Neurocomputational Model. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.615538

Cappa, S. F. (2012). Neurological accounts of agrammatism. In *Perspectives on agrammatism* (pp. 49–59). Psychology Press.

Chomsky, N. (1957). *Syntactic Structures*. Mouton. https://books.google.it/books?id=55YaAAAAIAAJ

Clarke, A. R., Barry, R. J., Karamacoska, D., & Johnstone, S. J. (2019). The EEG Theta/Beta Ratio: A marker of Arousal or Cognitive Processing Capacity? *Applied Psychophysiology and Biofeedback*, *44*(2), 123–129. https://doi.org/10.1007/s10484-018-09428-6

Cometa, A., d'Orio, P., Revay, M., Bottoni, F., Repetto, C., Russo, G. L., Cappa, S. F., Moro, A., Micera, S., & Artoni, F. (2023). Event-related causality in stereo-EEG discriminates syntactic processing of noun phrases and verb phrases. *Journal of Neural Engineering*, *20*(2), 026042. https://doi.org/10.1088/1741-2552/accaa8

Cometa, A., D'Orio, P., Revay, M., Micera, S., & Artoni, F. (2021). Stimulus evoked causality estimation in stereo-EEG. *Journal of Neural Engineering*, *18*(5), 056041. https://doi.org/10.1088/1741-2552/ac27fb

Cometa, A., Falasconi, A., Biasizzo, M., Carpaneto, J., Horn, A., Mazzoni, A., & Micera, S. (2022). Clinical neuroscience and neurotechnology: An amazing symbiosis. *iScience*, *25*(10), 105124. https://doi.org/10.1016/j.isci.2022.105124

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*. https://doi.org/10.1038/nn.4186

Edward M. Bernat, Lindsay D. Nelson, Clay B. Holroyd, William J. Gehring, & Christopher J. Patrick. (2008). *Separating cognitive processes with principal components analysis of EEG time-frequency distributions*. *7074*, 70740S. https://doi.org/10.1117/12.801362

Fahey, D., Fridriksson, J., Hickok, G., & Matchin, W. (2024). Lesion-symptom Mapping of Acceptability Judgments in Chronic Poststroke Aphasia Reveals the Neurobiological Underpinnings of Receptive Syntax. *Journal of Cognitive Neuroscience*, *36*(6), 1141–1155.

Fossum, V., & Levy, R. (2012). *Sequential vs. Hierarchical syntactic models of human incremental sentence processing*. 61–69.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2013). *Word surprisal predicts N400 amplitude during reading*.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, *140*, 1–11. https://doi.org/10.1016/j.bandl.2014.10.006

Friederici, A. D. (2017). Neurobiology of Syntax as the Core of Human Language. *Biolinguistics*, *11*, 325–338. https://doi.org/10.5964/bioling.9093

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., Nastase, S. A., Feder, A., Emanuel, D., Cohen, A., Jansen, A., Gazula, H., Choe, G., Rao, A., Kim, C., Casto, C., Fanda, L., Doyle, W., Friedman, D., … Hasson, U. (2022). Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience*, *25*(3), 369–380. https://doi.org/10.1038/s41593-022-01026-4

Grandchamp, R., & Delorme, A. (2011). Single-Trial Normalization for Event-Related Spectral Decomposition Reduces Sensitivity to Noisy Trials. *Frontiers in Psychology*, *2*, 236. https://doi.org/10.3389/fpsyg.2011.00236

Greco, M., Cometa, A., Artoni, F., Frank, R., & Moro, A. (2023). False perspectives on human language: Why statistics needs linguistics. *Frontiers in Language Sciences*, *2*. https://www.frontiersin.org/articles/10.3389/flang.2023.1178932

Grodzinsky, Y., & Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology*, *16*(2), 240–246. https://doi.org/10.1016/j.conb.2006.03.007

Hale, J. (2001). *A probabilistic Earley parser as a psycholinguistic model*. Second meeting of the north american chapter of the association for computational linguistics.

Hamilton, L. S., & Huth, A. G. (2020). The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, *35*(5), 573–582.

Harmony, T. (2013). The functional significance of delta oscillations in cognitive processing. *Frontiers in Integrative Neuroscience*, *7*(December), 83. https://doi.org/10.3389/fnint.2013.00083

Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P., & de Lange, F. P. (2022). A hierarchy of linguistic predictions during natural language comprehension. *Proceedings of the National Academy of Sciences*, *119*(32), e2201968119. https://doi.org/10.1073/pnas.2201968119

Henderson, J. M., Choi, W., Lowder, M. W., & Ferreira, F. (2016). Language structure in the brain: A fixation-related fMRI study of syntactic surprisal in reading. *NeuroImage*, *132*, 293–300. Scopus. https://doi.org/10.1016/j.neuroimage.2016.02.050

Kaufeld, Hans Rutger Bosker, Sanne ten Oever, Phillip M. Alday, Antje S. Meyer, & Andrea E. Martin. (2020). Linguistic Structure and Meaning Organize Neural Oscillations into a Content-Specific Hierarchy. *The Journal of Neuroscience*, *40*(49), 9467. https://doi.org/10.1523/JNEUROSCI.0302-20.2020

Kayne, R. (2019). What Is Suppletion? On *Goed and on Went in Modern English. *Transactions of the Philological Society*, *117*. https://doi.org/10.1111/1467-968x.12173

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*. https://www.frontiersin.org/articles/10.3389/neuro.06.004.2008

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., & Schalk, G. (2013). The Tracking of Speech Envelope in the Human Cortex. *PLOS ONE*, *8*(1), e53398. https://doi.org/10.1371/journal.pone.0053398

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. https://doi.org/10.1152/jn.00263.2005

Levy, R. (2011). *Integrating surprisal and uncertain-input models in online sentence comprehension: Formal techniques and empirical results*. 1055–1065.

Lin, Y., Tan, Y. C., & Frank, R. (2019). Open Sesame: Getting inside BERT's Linguistic Knowledge. *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, 241–253. https://doi.org/10.18653/v1/W19-4825

Liuzzi, A. G., Aglinskas, A., & Fairhall, S. L. (2020). General and feature-based semantic representations in the semantic network. *Scientific Reports*, *10*(1), 8931. https://doi.org/10.1038/s41598-020-65906-0

Lo, C.-W., Tung, T.-Y., Ke, A. H., & Brennan, J. R. (2022). Hierarchy, Not Lexical Regularity, Modulates Low-Frequency Neural Synchrony During Language Comprehension. *Neurobiology of Language*, *3*(4), 538–555. https://doi.org/10.1162/nol_a_00077

Lopopolo, A., Frank, S. L., Van den Bosch, A., & Willems, R. M. (2017). Using stochastic language models (SLM) to map lexical, syntactic, and phonological information processing in the brain. *PloS One*, *12*(5), e0177794.

Lukic, S., Borghesani, V., Weis, E., Welch, A., Bogley, R., Neuhaus, J., Deleon, J., Miller, Z. A., Kramer, J. H., Miller, B. L., Dronkers, N. F., & Gorno-Tempini, M. L. (2021). Dissociating nouns and verbs in temporal and perisylvian networks: Evidence from neurodegenerative diseases. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *142*, 47–61. https://doi.org/10.1016/j.cortex.2021.05.006

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Lyding, V., Stemle, E., Borghetti, C., Brunello, M., Castagnoli, S., Dell'Orletta, F., Dittmann, H., Lenci, A., & Pirrelli, V. (2014). *The paisa'corpus of italian web texts*. 36–43.

Magrassi, L., Aromataris, G., Cabrini, A., Annovazzi-Lodi, V., & Moro, A. (2015). Sound representation in higher language areas during language generation. *Proceedings of the National Academy of Sciences*, *112*(6), 1868–1873. https://doi.org/10.1073/pnas.1418162112

Mai, G., Minett, J. W., & Wang, W. S.-Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, *133*, 516–528. https://doi.org/10.1016/j.neuroimage.2016.02.064

Manning, C. D., Clark, K., Hewitt, J., Khandelwal, U., & Levy, O. (2020). Emergent linguistic structure in artificial neural networks trained by self-supervision. *Proceedings of the National Academy of Sciences*, *117*(48), 30046–30054. https://doi.org/10.1073/pnas.1907367117

Maran, M., Friederici, A. D., & Zaccarella, E. (2022). Syntax through the looking glass: A review on two-word linguistic processing across behavioral, neuroimaging and neurostimulation studies. *Neuroscience & Biobehavioral Reviews*, 104881.

Marsden, M. (1974). Cubic spline interpolation of continuous functions. *Journal of Approximation Theory*, *10*(2), 103–111. https://doi.org/10.1016/0021-9045(74)90109-9

Martina Berto, Emiliano Ricciardi, Pietro Pietrini, Nathan Weisz, & Davide Bottari. (2022). Distinguishing fine structure and summary representation of sound textures from neural activity. *bioRxiv*, 2022.03.17.484757. https://doi.org/10.1101/2022.03.17.484757

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex*, *88*, 106–123.

Matchin, W., & Hickok, G. (2020). The Cortical Organization of Syntax. *Cerebral Cortex (New York, N.Y. : 1991)*, *30*(3), 1481–1498. https://doi.org/10.1093/cercor/bhz180

Meyer, L., & Gumbert, M. (2018). Synchronization of Electrophysiological Responses with Speech Benefits Syntactic Information Processing. *Journal of Cognitive Neuroscience*, *30*(8), 1066–1074. https://doi.org/10.1162/jocn_a_01236

Momenian, M., Vaghefi, M., Sadeghi, H., Momtazi, S., & Meyer, L. (2024). Language prediction in monolingual and bilingual speakers: An EEG study. *Scientific Reports*, *14*(1), 6818. https://doi.org/10.1038/s41598-024-57426-y

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, *46*(3), 887–903. https://doi.org/10.3758/s13428-013-0405-3

Moro, A. (2016). *Impossible languages*. MIT Press. https://doi.org/10.7551/mitpress/9780262034890.001.0001

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*. https://doi.org/10.1002/hbm.1058

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as "asymmetric sampling in time." *Speech Communication*, *41*(1), 245–255. https://doi.org/10.1016/S0167-6393(02)00107-3

Pulvermüller, F., Preissl, H., Lutzenberger, W., & Birbaumer, N. (1996). Brain Rhythms of Language: Nouns Versus Verbs. *European Journal of Neuroscience*, *8*(5), 937–941. https://doi.org/10.1111/j.1460-9568.1996.tb01580.x

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A Python natural language processing toolkit for many human languages. *arXiv Preprint arXiv:2003.07082*.

Ramon, C., Freeman, W. J., Holmes, M., Ishimaru, A., Haueisen, J., Schimpf, P. H., & Rezvanian, E. (2009). Similarities between simulated spatial spectra of scalp EEG, MEG and structural MRI. *Brain Topography*, *22*(3), 191–196. https://doi.org/10.1007/s10548-009-0104-7

Repetto, C., Rodella, C., Conca, F., Santi, G. C., & Catricalà, E. (2023). The Italian Sensorimotor Norms: Perception and action strength measures for 959 words. *Behavior Research Methods*, *55*(8), 4035–4047. https://doi.org/10.3758/s13428-022-02004-1

Roark, B. (2001). Probabilistic top-down parsing and language modeling. *Computational Linguistics*. https://doi.org/10.1162/089120101750300526

Roark, B., Bachrach, A., Cardenas, C., & Pallier, C. (2009). Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 324–333. https://aclanthology.org/D09-1034

Russo, A. G., Ciarlo, A., Ponticorvo, S., Di Salle, F., Tedeschi, G., & Esposito, F. (2022). Explaining neural activity in human listeners with deep learning via natural language processing of narrative text. *Scientific Reports*, *12*(1), 17838. https://doi.org/10.1038/s41598-022-21782-4

Russo, A. G., De Martino, M., Mancuso, A., Iaconetta, G., Manara, R., Elia, A., Laudanna, A., Di Salle, F., & Esposito, F. (2020). Semantics-weighted lexical surprisal modeling of naturalistic functional MRI time-series during spoken narrative listening. *NeuroImage*, *222*, 117281. https://doi.org/10.1016/j.neuroimage.2020.117281

Sauppe, S., Choudhary, K. K., Giroud, N., Blasi, D. E., Norcliffe, E., Bhattamishra, S., Gulati, M., Egurtzegi, A., Bornkessel-Schlesewsky, I., Meyer, M., & Bickel, B. (2021). Neural signatures of syntactic variation in speech planning. *PLOS Biology*, *19*(1), e3001038. https://doi.org/10.1371/journal.pbio.3001038

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, *138*, 107307. https://doi.org/10.1016/j.neuropsychologia.2019.107307

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Stanojević, M., Brennan, J. R., Dunagan, D., Steedman, M., & Hale, J. T. (2023). Modeling Structure-Building in the Brain With CCG Parsing and Large Language Models. *Cognitive Science*, *47*(7), e13312. https://doi.org/10.1111/cogs.13312

Stropahl, M., Bauer, A.-K. R., Debener, S., & Bleichner, M. G. (2018). Source-Modeling Auditory Processes of EEG Data Using EEGLAB and Brainstorm. *Frontiers in Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnins.2018.00309

Tenney, I., Xia, P., Chen, B., Wang, A., Poliak, A., McCoy, R. T., Kim, N., Van Durme, B., Bowman, S. R., & Das, D. (2019). What do you learn from context? Probing for sentence structure in contextualized word representations. *arXiv Preprint arXiv:1905.06316*.

Van Schijndel, M., & Schuler, W. (2015). *Hierarchic syntax improves reading time prediction*. 1597–1605.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and Biobehavioral Reviews*, *35*(3), 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007

Weissbart, H., Kandylaki, K. D., & Reichenbach, T. (2020). Cortical Tracking of Surprisal during Continuous Speech Comprehension. *Journal of Cognitive Neuroscience*, *32*(1), 155–166. https://doi.org/10.1162/jocn_a_01467

Widmann, A. (2006). Firfilt EEGLAB plugin, version 1.5. 1. *Leipzig: University of Leipzig*.