Structural Prediction During Language Comprehension Revealed by Electrophysiology: Evidence from Italian Auxiliaries

Giulia Bovolenta and E. Matthew Husband

Faculty of Linguistics, Philology and Phonetics, University of Oxford

Author Note

Giulia Bovolenta https://orcid.org/0000-0003-4139-6446

E. Matthew Husband https://orcid.org/0000-0002-6446-5582

Giulia Bovolenta is now at the Department of Education, University of York.

Conflict of interest: E. Matthew Husband is a Consulting Editor for *Journal of Experimental Psychology: Learning, Memory and Cognition.*

This work was supported by the Economic and Social Research Council [Doctoral award to Giulia Bovolenta].

Correspondence concerning this article should be addressed to E. Matthew Husband, St. Hugh's College, St. Margaret's Rd., Oxford OX2 6LE, United Kingdom

Email: matthew.husband@ling-phil.ox.ac.uk

Abstract

Prediction in language comprehension has become a key mechanism in recent psycholinguistic theory, with evidence from lexical prediction as a primary source. Less work has focused on whether comprehenders also make structural predictions above the lexical level. Previous research shows that processing is facilitated for syntactic structures which are predictable based on context; however, there is so far no direct evidence that speakers formulate structural predictions ahead of encountering input. We investigated whether subject noun animacy cues comprehenders to predict different verb phrase (VP) structures, with the incompatibility between a low animacy subject and an Agent interpretation of transitives/unergative VPs predicting a derived (passive/unaccusative) VP structure, using Italian auxiliaries. Native Italian speakers read sentences with subject nouns varying from high to low animacy followed by the auxiliary *avere*$_{\text{HAVE}}$, which is compatible with underived VPs, or *essere*$_{\text{BE}}$, which is compatible with derived VPs. The auxiliary *avere*$_{\text{HAVE}}$ elicited greater frontal negativity when preceded by a subject noun with lower animacy. The auxiliary *essere*$_{\text{BE}}$ elicited no differential ERP given subject animacy. We propose that this frontal negativity reflects violation of a structural prediction, with amplitude reflecting the strength of initial commitment or difficulty in revising a predicted structure. Differences between auxiliaries are proposed to follow from the more specific distribution of *avere*$_{\text{HAVE}}$. We argue that this evidence unambiguously supports a predictive mechanism for phrase-level structure.

Keywords: animacy, auxiliary selection, ERPs, Italian, structural prediction

Structural Prediction During Language Comprehension Revealed by Electrophysiology: Evidence from Italian Auxiliaries

The 21st century has seen a remarkable resurgence of interest in the role played by predictive mechanisms in language comprehension (Thornhill & Van Petten, 2012; Kutas, DeLong, & Smith, 2011; DeLong, Troyer, & Kutas, 2014; Kuperberg & Jaeger, 2016; Ferreira & Chantavarin, 2018; Pickering & Gambi, 2018). Much of this research has focused on lexical prediction, proposing that prior sentential and discourse context not only aids in accessing individual words but that comprehenders can also pre-activate upcoming words given prior context. The most convincing evidence for lexical prediction has come from studies demonstrating effects of prediction failures on articles and adjectives whose morphological form depended on features of an upcoming word and mismatches comprehenders' expectations for that word's features. Far less research has focused on comprehenders' use of context to pre-compute upcoming structure, with demonstrations of predictions for phrase-level structure being "less forthcoming" in the literature (DeLong, Troyer, & Kutas, 2014: 639). The lack of such research focusing on structural prediction is surprising, as there are good reasons to think that structural predictions might be more robust and deployed more broadly than lexical predictions during everyday language comprehension. At a given point within a sentence there are often fewer structural options than there are lexical options, and there are also tight relationships between structure and interpretation, for instance, in the distribution of thematic roles. This all suggests that structural prediction deserves further examination.

While there is evidence that syntactic processing is affected by predictability, no studies so far have investigated the effects of predictability ahead of encountering the expected syntactic structure. Evidence from such studies would provide decisive evidence for structural prediction as it has for lexical prediction. Taking a cue from the lexical prediction literature, we investigated whether comprehenders predict syntactic structure using Italian auxiliaries. The form of an Italian auxiliary depends on whether the upcoming verb phrase is underived (e.g. transitive/unergative) or not (Burzio, 1986). An example of underived verb phrase is a standard active sentence, such as *Il pittore ha completato la tela* 'The painter has finished the painting', which takes *avere*$_{\mathrm{HAVE}}$ forms for the perfect tense auxiliary. An example of derived verb phrase, by contrast, is a passive sentence such as *Il progetto è stato approvato* 'The plan has been approved', which takes *essere*$_{\mathrm{BE}}$ forms for the perfect tense auxiliary. As the two forms of the perfect tense auxiliary in Italian are diagnostic of upcoming verb phrase structure, they allow us to probe for structural predictions of different verb phrase structures prior to the appearance of the verb phrases themselves.

To set the stage, in the next section we review the recent literature on prediction in language comprehension, consisting of EEG studies that have demonstrated lexical prediction by leveraging the morphosyntactic and phonological dependencies between nouns and their preceding elements. We then review evidence suggesting that lexical predictions may not be robust or reliable outside laboratory settings. Structural prediction, we hypothesize, may be potentially more useful as a mechanism, due to the higher predictability of syntactic categories relative to lexical items. We then survey the evidence from sentence processing showing that comprehenders can form expectations about structural continuations based on both structural and semantic cues, before turning attention to our study.

Semantic and Lexical Prediction in Language Comprehension

During language comprehension, semantic information from sentential and discourse context can be used to anticipate upcoming lexical content. Words that are predictable from context are processed more easily: they elicit shorter fixations during reading (Balota, Pollatsek, & Rayner, 1985; McDonald and Shillcock, 2003; Ehrlich and Rayner, 1981; Demberg and Keller, 2008), and faster reaction times in lexical decision (Schwanenflugel and LaCount, 1988; Schwanenflugel and White, 1991; Stanovich and West, 1983) and naming tasks (Stanovich and West, 1981, 1983; Forster, 1981; Traxler and Foss, 2000). In a series of ERP studies manipulating sentential constraint and predictability, Kutas and Hillyard (1980) first showed evidence that semantically predictable words elicit a reduced N400 response (a negative shift in the centroparietal region peaking 400ms post-stimulus onset), relative to unpredictable ones (Kutas & Hillyard, 1980; Kutas & Hillyard, 1984; Kutas & Federmeier, 2011). This effect is modulated by both the cloze probability of the final word, and by sentential constraint (Federmeier and Kutas, 1999; DeLong, Urbach and Kutas, 2005; Federmeier, Wlotko, De Ochoa-Dewald and Kutas, 2007; Wlotko and Federmeier, 2012). It also driven by semantic similarity: words that are unpredictable but semantically related to the target also cause a reduction in N400 (albeit a smaller one), relative to unpredictable, unrelated ones (Federmeier and Kutas, 1999; Federmeier, McLennan, Ochoa, & Kutas, 2002; Thornhill and VanPetten, 2012; Wlotko and Federmeier, 2015). While many theories propose that these effects are driven by prediction, they are also consistent with theories of ease of integration. This is because such effects occur at the position of the expected word.

Evidence from the processing of forward-looking morphological dependencies has been more decisive in favor of predictive mechanisms by demonstrating effects of pre-activation of expected lexical items prior to their occurrence. These studies showed that comprehenders were not just anticipating certain semantic features, but that they were predicting specific lexical features, including morphosyntactic information such as gender (Wicha, Moreno and Kutas, 2004; Van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort, 2005; Otten and Van Berkum, 2008, 2009; Szewczyk and Schriefers, 2013) and grammatical animacy marking (Szewczyk and Schriefers, 2013), as well as their phonological form (DeLong et al., 2005; 2009). These studies manipulated morphosyntactic and phonological dependencies between the expected completion and prior elements in the sentence (e.g. gender-marked determiners), showing that participants were already predicting a specific lexical completion by the time they encountered the prior element. Otten and VanBerkum (2009) investigated the prediction of individual lexical items following highly constraining sentence contexts in Dutch (e.g. *The actress wore a beautiful dress, but she thought* *her neck was a little plain. She picked up... the$_{com}$ delicate yet striking necklace that had been selected by her stylist*) When sentence contexts were followed by a determiner that did not match the gender of the high-cloze completion (e.g. *...the$_{neut}$ delicate yet striking collar*), this elicited a widespread negativity in the 200-600ms time region, relative to gender-congruent determiners. In English, work by DeLong and colleagues (DeLong, et al, 2005; DeLong, et al, 2009) showed how the same kind of effect can be elicited by morphophonological incongruence. Participants read sentences such as *The day was breezy so the boy went to fly…,* which is highly constraining for the completion *(a) kite.* At this point, encountering the *an* form of the determiner (e.g. *an airplane*) elicited a significantly larger N400 compared to the form *a,* the only one compatible with the expected noun *kite*. These studies show that subjects can formulate specific predictions on upcoming lexical items, and that they are sensitive to early cues disconfirming these predictions.

While empirically interesting, lexical predictions may, however, be limited in their scope and fragile and unreliable in online processing or outside of laboratory settings. In a recent extensive study of lexical predictability that elicited cloze continuations for every word in 55 natural texts, Luke and Christianson (2016) found that highly predictable content words (>0.67 cloze probability) were rare in naturally occurring texts, comprising only about 5% of all content words. Furthermore, for any given target content word, the actual word given in the text was the most frequent cloze response only 21% of the time, meaning that, for about four fifths of the content words, some word other than the actual word was more expected by participants. Such low rates for lexical predictability could present problems for a highly active predictive comprehender and make predictive mechanisms unreliable during everyday language processing.

Examples of such unreliability may be responsible for recent failed attempts to replicate prior experiments on lexical prediction. In a large-scale replication study across nine labs, Nieuwland et al. (2018) did not replicate the N400 effect of cloze on *a/an* articles reported by DeLong et al. (2005), while reproducing the N400 effect on the noun. In their discussion, Nieuwland et al. note there is only a 33% probability for indefinite articles *a/an* to be directly followed by a noun in English corpora (Nieuwland et al., 2018, p. 14), suggesting that a cue from the indefinite article's form as *a/an* may not be robust to signal prediction failure. Other evidence suggests that prediction of phonological form based on sentential context is possible, but is not as robust as prediction of other lexical features. Ito et al. (2016) found that, in high-cloze sentences, words that were semantically related to the target completion reliably caused a reduction in N400 response. By contrast, words that where form-related to the target completion (e.g. *hook* for *book*) did not cause an N400 reduction; instead, they elicited a late sustained positivity in the parietal region, though only in very high-cloze sentences and at a slower presentation rate (700ms). Similar evidence was reported in Ito et al. (2020). In their study, ERPs to articles whose gender mismatched the expected noun were elicited earlier and lasted longer than ERPs to articles whose phonological form mismatched the expected noun. Even evidence from gender has been less reliable than one would hope, replicating only partially in Kochari and Flecken (2019). Taken together, this suggests that lexical prediction may not be a robust domain for investigations into predictive mechanisms.

The Case for Structural Prediction

While lexical prediction itself may not be robust, predictive mechanisms could be quite useful at other levels of linguistic representation. Prediction for syntactic structure, in particular, is likely to be more useful and reliable, given that the alternatives are fewer at any given point in a sentence. Empirically, Luke and Christianson (2016) found that syntactic category was predictable for content words about 50% of the time, with the predictability of nouns and verbs at a higher 72%. Morphosyntactic information concerning content words was also highly predictable. The plurality of a noun was predictable 72% of the time and the tense of a verb was predictable 78% of the time. This suggests that comprehenders may benefit from predicting structural information even when lexical information is unpredictable. Because of these conditions, the comprehender may use information that favours a particular sentence structure to pre-compute that structure even without knowing what particular lexical items will instantiate it.

There is evidence to suggest that comprehenders formulate expectations about upcoming structure, as shown by ERP and behavioural responses to structural completions which are dispreferred based on context. Syntactic category violations (e.g. the ungrammatical word *about* in the sentence "*I heard Max's about story"*) are detected rapidly and robustly, eliciting brain responses as early as 130ms post-stimulus onset (Neville, Nicol, Barss, Forster, and Garrett, 1991; Friederici, 2002; Lau, Stroud, Plesch, and Phillips, 2006; Dikker, Rabagliati, and Pylkkänen, 2009; Dikker, Rabagliati, Farmer, and Pylkkänen, 2010). Evidence further suggests that comprehenders can make predictions about larger structural units, using structural elements to anticipate disjunction, filler-gap constructions, and ellipsis (Staub and Clifton, 2006; Phillips, 2006; Wagers and Phillips, 2009; Lau, Stroud, Plesch, and Phillips, 2006; Yoshida, Dickey, and Sturt, 2013). Staub and Clifton (2006) reported evidence consistent with prediction in the processing of disjunction. They manipulated the presence or absence of *either* which comprehenders could use to predict an upcoming disjunct *or* as well as the size of the phrases being joined. They found faster reading times on the phrase after *or* when it was preceded by *either*. Additionally, comprehenders were able to use the presence of *either* to avoid an NP/S coordination ambiguity in a sentence like "(Either) Louise punished the children or the parents decided to let it slide". Comprehenders can also anticipate different types of predicate structure, based on previous elements in the sentence. In a word detection task using French, Brusini, Brun, Brunet, and Christophe (2015) reported that hearing a pre-verbal object clitic made listeners more likely to expect a transitive verb, relative to sentences without an object clitic which were equally compatible with transitive and intransitive verbs.

Thematic information can also serve as a cue to guide lexical expectations, specifically concerning upcoming verbs (Chow, Smith, Lau & Phillips, 2016; Chow, Lau, Wang & Phillips,

2018) and verb arguments (Altmann & Kamide, 1999; Kamide, Altmann & Haywood, 2003). Verb selectional restrictions and the semantics of the Agentive noun subjects have been shown to jointly guide comprehenders to anticipate suitable nouns to fill the direct object Theme in the sentence, as evidenced by preferential looking in a visual world eye-tracking paradigm (Altmann & Kamide, 1999; Kamide, Altmann & Haywood, 2003). For instance, upon hearing "The man/girl will ride the...", participants launched anticipatory eye movements to suitable Themes for the verb ride (e.g. a motorbike/carousel) than to other objects in the scene. Furthermore, the specific Theme object they were most likely to look at depended on the noun filling the Agent position: if the subject was *The girl,* participants were most likely to look at the carousel, while *The man* elicited preferential looking towards the motorbike (Kamide et al., 2003). In a series of ERP studies, Chow and colleagues (Chow et al., 2016; Chow et al., 2018) showed that semantic information from preverbal arguments could similarly shape expectations for upcoming verbs: substituting an argument with a different one which rendered the verb low-cloze (e.g., "The superintendent overheard which *tenant* / *realtor* the landlord had <u>evicted</u> the end of May") elicited a greater N400 on the verb (Chow et al., 2016).

The studies just reviewed show that semantic and thematic cues can inform expectations about upcoming lexical items, whether they be verbs or argument nouns. However, there is also evidence that the processing of phrase-level syntactic structure is sensitive to semantic information provided earlier in the sentence (McRae et al., 2005; Hare et al., 2009; Kim & Osterhout, 2005; Kuperberg et al., 2003). In particular, a noun's *animacy* (the extent to which a noun refers to an animate or causal entity) is a good predictor of its capacity to fill different thematic roles, such as Agent (Causer) or Patient (Theme). An inanimate noun such as *table* is a suitable Patient but not a suitable Agent because it lacks mental states, the ability to initiate motion, and the ability to cause actions (Dowty 1991). This makes inanimate nouns unlikely to occupy the subject position of a transitive or unergative verb phrase, which take Agent subjects, and more likely to be the subject of a passive or unaccusative structure, which take Patient subjects. Comprehenders are sensitive to these regularities, which is reflected in their processing of predicate structure. The thematic fit of grammatical subjects can shape expectations for upcoming verbs. Nouns commonly associated with different types of thematic roles (agents, patients, instruments, and locations) prime verbs depicting events compatible with those thematic roles (McRae et al. 2005). Furthermore, when a verb is compatible with more than one type of syntactic structure, subject animacy can guide comprehenders' expectations for the specific syntactic structure that the verb is likely to be used with. In a self-paced reading study, Hare et al. (2009) constructed sentences from nouns rated as 'good causes' or 'good themes' and verbs that could be used transitively or intransitively (e.g. *shatter*). They found that 'good themes' led participants to expect intransitive structures (e.g. *The glass shattered into tiny pieces when it hit the floor*), while 'good causes' led them to expect transitive usage of the verb. When these expectations were not met, reading times for post-verbal regions increased significantly (while there was no difference in reading times for the verb itself). Similar findings have been reported from ERP studies of sentence processing. Research on the "semantic" P600 component has shown that verb-argument combinations which are implausible due to thematic mismatch elicit a P600 relative to plausible sentences (Hoeks, Stowe & Doedens, 2004; Kim & Osterhout, 2005; Kuperberg et al., 2003). For instance, in the sentence "The hearty meal was...devouring / devoured", the active form *devouring* elicits a P600 effect relative to *devoured*, because *The hearty meal* does not constitute a suitable Agent (Kim & Osterhout, 2005).

To summarise, evidence shows that the processing of syntactic structure is affected by its compatibility with structural and thematic features previously encountered in the input. However, these effects have so far only been observed at the point when comprehenders encounter a given syntactic structure, not before. Therefore, we do not know whether these effects result from comprehenders' prediction of structure ahead of encountering them, or whether they result as instances of access and integration, where structures which better fit with the preceding context are easier to process. In this study, we investigate this question by using features of Italian auxiliaries which allow us to probe for participants' expectations concerning syntactic structures early on.

Cross-linguistically, the difference between different types of predicates may be marked overtly ahead of the predicate. In Italian, for example, perfect tense auxiliaries come in two morphological forms: *avere* ‘have’ or *essere* ‘be’. These two forms are used to mark perfect tense constructions but do not otherwise differ in lexical meaning. Instead, the difference between them lies in the type of verb phrase they can introduce (Perlmutter, 1978; Burzio 1986). *Avere*$_{\text{HAVE}}$ forms are used with transitive and unergative verb phrases, where the grammatical subject plays an active role: e.g. *Il bambino ha*$_{\text{HAVE}}$ *mangiato un biscotto* ‘The child has eaten a biscuit’, or *La ballerina ha*$_{\text{HAVE}}$ *danzato* ‘The ballerina has danced’. *Essere*$_{\text{BE}}$ forms are used with passive and unaccusative verb phrases, where the grammatical subject plays a more passive role: e.g. *Il cameriere è*$_{\text{BE}}$ *stato licenziato* ‘The waiter has been fired’, or *La foglia è*$_{\text{BE}}$ caduta ‘The leaf has fallen’.

In a generative syntax framework, a variety of evidence has been brought forth to demonstrate that the grammatical subject of a transitive or unergative sentence has a different status from the subject of a passive or unaccusative sentence (Baker, Johnson, and Roberts, 1989; Perlmutter, 1978). Transitive and unergative clauses (Figure 1A) assign the thematic role of Agent directly to the subject, and, if transitive, they also assign the role of Theme to the grammatical object within the verb phrase. In passive and unaccusative clauses (Figure 1B), the grammatical subject underlyingly occupies the same position as the object of a transitive: it is part of the VP, where it is assigned the role of Theme. In the surface form of the sentence, the subject is derived (moved) from its position within the VP to appear before the predicate, retaining its Theme role. Grammatically, it is this derived status that is structurally selected by *essere* (Burzio, 1986; Kayne, 2010). Therefore, structures where the grammatical subject is interpreted as an Agent are introduced by *avere*$_{\text{HAVE}}$, while structures where the grammatical subject is intepreted as a Theme use *essere*$_{\text{BE}}$. Table 1 summarizes the links between auxiliary form, structure, and the subject’s thematic role.

We can make use of this relationship between auxiliary form, structure, and the subject's thematic role to probe for predictions of upcoming structure by manipulating the animacy of subject nouns. The animacy of a subject noun can be a strong predictor for the kind of thematic role it is likely to take on and therefore the structure of the upcoming predicate. Figure 2 schematizes the relationship between animate and inanimate subject nouns, their likely thematic roles, and the different verb phrase structures required to assign those roles to those subject nouns. Animate nouns (Figure 2B) can fill both the role of Agent and that of Theme, and is therefore a likely subject for both transitive/unergative and passive/unaccusative sentences. By contrast, inanimate nouns (Figure 2C) are a good fit for the Theme role but not for Agent, meaning that they are more likely to occur as the subject of a passive/unaccusative sentence. Figure 2 also shows that these different VPs select for different auxiliary forms in Italian: *avere* for transitive/unergative structures and *essere* for passive/unaccusative structures. Since these auxiliaries comes after the subject but before the VP, we can use their different forms to probe whether comprehenders are predicting the structure of the upcoming predicate. Evidence of disruption due to a mismatch between the predicted predicate's structure and the auxiliary form required for that structure would constitute strong evidence for structural prediction. Crucially, it would constitute evidence that participants are formulating predictions about syntactic structures before encountering them, as demonstrated for lexical prediction by ERP studies exploiting morphosyntactic and phonological dependencies (Wicha et al., 2004; Van Berkum et al., 2005; Otten and Van Berkum, 2008, 2009; Szewczyk and Schriefers, 2013; DeLong et al., 2005; 2009).

The following study manipulated subject animacy and the morphological form of auxiliaries in Italian to investigate whether comprehenders predict syntactic structure using the animacy of a subject noun phrase. We predicted a difference in the ERP response to *avere*$_{\text{HAVE}}$ when preceded by inanimate subjects versus animate subjects because inanimate subjects generate a strong expectation for a passive or unaccusative structure given that they are incompatible with an agent thematic role of transitive or unergative structures which is signalled by *avere*$_{\text{HAVE}}$. In contrast, we made two different predictions for ERPs elicited by *essere*$_{\text{BE}}$. On the one hand, ERP responses to *essere*$_{\text{BE}}$ could different when preceded by animate versus inanimate subjects since the presence of *essere*$_{\text{BE}}$ violates the expectation animate subjects may generate for an active or unergative structure, given their compatibility with the Agent thematic role of these verb phrases. On the other hand, ERP responses to *essere*$_{\text{BE}}$ may not differ when preceded by either animate or inanimate nouns, as both animate and inanimate nouns are compatible with a Patient/Theme thematic role, which is consistent with verb phrases that require *essere*$_{\text{BE}}$. Both predictions are compatible with norming data elicited from native Italian speakers (Table 2). On one hand, the likelihood of *essere*$_{\text{BE}}$ being used as auxiliary decreases as subject noun animacy increases, and at the highest level of animacy, *essere*$_{\text{BE}}$ is the dispreferred option (Total column in Table 2), which suggests incompatibility between high animacy and *essere*$_{\text{BE}}$ as auxiliary. On the other hand, *essere*$_{\text{BE}}$ is the more likely completion on average, and it is still more likely than *avere*$_{\text{HAVE}}$ following mid- to high-animacy nouns; therefore, it may be perceived as a suitable completion even at higher animacy levels.

Method

Participants

30 native Italian speakers from the University of Oxford and surrounding community (14 females, mean age 28 years) participated in the study. All participants were right-handed as assessed by the Edinburgh Handedness Test (Oldfield, 1971), and received a compensation of £20 for taking part in the study. Research ethics was approved by the University of Oxford's Social Sciences and Humanities Inter-divisional Research Ethics Committee (IDREC), Ref No: R44463/RE001, Title: "Syntactic and semantic comprehension as reflected in behaviour, eye movements, and electrophysiology".

Materials

To obtain a measure of animacy for our subject nouns, we created an online survey using Google Forms (https://www.google.com/forms/about/) and distributed it to a sample of 80 native Italian speakers. Follow Dowty's (1991) definition of Proto-Agents, participants were given a list of common Italian nouns; for each noun, they had to answer the following questions about the noun's referent: "Can it think?", "Can it move on its own?" and "Can it cause something to happen?". Questions were answered by giving the appropriate rating on a scale from 1 to 7, with 1 meaning "Not at all" and 7 "Definitely yes". The mean rating obtained by each item across the three questions was used as animacy score for the item.

To obtain cloze probabilities for *essere*$_{\text{BE}}$ and *avere*$_{\text{HAVE}}$ following the nouns, we recruited a further sample of 185 native Italian speakers who provided sentence completions for the nouns, using online survey platform Qualtrics (https://www.qualtrics.com). Participants were presented with a sentence fragment consisting of a determiner and noun, and asked to complete the sentence as follows: *"Complete each beginning with a sentence of your choice. E.g. "THE DOG… has bitten the bone", "THE STUDENT… arrived late" or "PIZZA… is my favourite dish". There are no rules to follow, just write the first sentence that comes to mind."* (Original Italian instructions: *'Completa ciascun inizio con una frase a tua scelta. Ad esempio: "IL CANE... ha morso l'osso", "LO STUDENTE... è arrivato in ritardo", oppure "LA PIZZA... è il mio piatto preferito". Non ci sono regole da seguire, semplicemente scrivi la prima frase che ti viene in mente!'*). Sentence completions were coded for overt auxiliary, when an auxiliary was provided, and for the auxiliary required by the verb phrase when an auxiliary was not overtly provided. Cloze probabilities for the auxiliaries *essere*$_{\text{BE}}$ and *avere*$_{\text{HAVE}}$ following each noun are shown in Table 2. A breakdown of these figures by animacy level showed the expected pattern: in total *essere*$_{\text{BE}}$ decreased as animacy increased from 80.83% to 39.31% (*Est.* = -5.465, t = -5.345, p = .003). The opposite pattern was found for *avere*$_{\text{HAVE}}$ which increased as animacy increased from 19.17% to 60.34% (*Est.* = 5.014, t

= 4.312, p = .008). Isolating those cases with overt elicitation of auxiliaries only showed the same global pattern (*essere*$_{\mathrm{BE}}$: *Est.* = -5.919, t = -5.475, p = .003; *avere*$_{\mathrm{HAVE}}$: *Est*. = 1.769, t = 4.228, p = .008) while analysis of main verb phrase auxiliary requirements for cases without an overt auxiliary only found an effect for *avere*$_{\mathrm{HAVE}}$ (*essere*$_{\mathrm{BE}}$: *Est.* = 0.455, t = 0.281, p = .790; *avere*$_{\mathrm{HAVE}}$: *Est.* = 3.244, t = 2.802, p = .038).

Each of 120 normed nouns (distributed evenly across the animacy scale) was used as the grammatical subject of two sentences in the present perfect tense, one using auxiliary *essere*$_{\mathrm{BE}}$ and one with *avere*$_{\mathrm{HAVE}}$, for a total of 240 experimental stimuli. A sample stimulus is shown in Table 3 and a full set of stimuli is provided in supplemental materials. Stimuli were counterbalanced across two lists, so that every subject saw 120 experimental stimuli mixed with 120 filler sentences. Each sentence was followed by a Yes / No comprehension question (the number of Yes and No correct answers was also counterbalanced across lists).

Procedure

Participants read stimuli displayed on a screen using word-by-word rapid serial visual presentation, while their EEG was being recorded. Stimuli were displayed on a 32” HD LED screen (Samsung Smart TV) positioned at approximately 120cm from the participant, in black 50-point serif typeface, on a light grey background. Each trial was initiated by a fixation cross that appeared at the centre on the screen and remained for 2 sec. Sentence stimuli were then presented using rapid serial visual presentation, one word at a time. Each word remained on the screen for 200 msec and was followed by a 300 msec blank screen, yielding a stimulus onset asynchrony of 500 msec. Participants were instructed to avoid eye blinks and body movements while the sentences were presented on screen, and were encouraged to blink when the fixation cross was shown. After each sentence, a relevant comprehension question appeared on the screen. Participants had to answer it by pressing the appropriate button on a computer mouse, which they held between their hands.

Recording

EEG was recorded on a 64-channel ANT Neuro system, mounted in an elastic cap, and referenced to the Cz electrode. Blinks and eye movements were registered by placing an electrode under each eye. Electrode impedance was kept below 20 kΩ throughout the experiment. The EEG was amplified with an ANT Neuro amplifier and sampled with a frequency of 512 Hz.

Analysis

Offline preprocessing and measurement of EEG data was done in Matlab using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014). EEG data was filtered (0.1-40 Hz), segmented -200 to 1000 msec time-locked to the onset of the auxiliary, rereferenced to the average, and baseline corrected (-100 to 0 msec). Channels with local artefacts (flatlined, low correlation, line noise) were identified using EEGLAB's clean_rawdata() function and interpolated. Trials with remaining artefacts were identified using a moving window peak-to-peak function (voltage change in EEG channels > 100μV in any 200 msec window every 50ms from the beginning of each epoch) and step function (voltage change in eye channels > 25μV in any 200 msec window every 50ms from the beginning of each epoch). Four participants were eliminated due to excess artefacts leaving them with fewer than 50% of the total trials. Artefact-free trials from the remaining 26 subjects were submitted for statistical analysis and grand averaged for visualization purposes.

For initial visual inspection, trials were split into low (1 to 4.4) and high (4.5 to 7) animacy groups for both *avere* and *essere* auxiliaries. Visual inspection of the grand average waveforms and topographic plots revealed a potential ERP difference for *avere* onsetting after 300 msec (see Figure 2 and supplemental figures S1 and S2). To determine the extent of this effect, a two-tailed cluster mass permutation analysis comparing the effect of high vs. low animacy on *avere*$_{\text{HAVE}}$ and on *essere*$_{\text{BE}}$ was carried out on 9 frontal electrodes (Fp1, FPz, Fp2, AF3, AFz, AF4, F1, Fz, F2) for all time points from 300 to 1000 msec (358 time point) with a family-wise $\alpha = 0.049$ (Bonferroni $\alpha =$

0.000016) using the Mass Univariate ERP Toolbox (Groppe, Urbach, & Kutas, 2011). Electrodes within approximately 5.44 cm of one another were considered spatial neighbors.

On the basis of the cluster analysis for low vs. high animacy, we further investigated the effects of continuous animacy on mean ERP amplitudes elicited by the two auxiliaries within a temporal window from 400 to 600 msec across five ROIs (see supplemental figure S3): Anterior (AF7, Fp1, FPz, Fp2, AF8, F5, AF3, AFz, AF4, F6, F1, Fz, F2), Left (FT7, FC5, FC3, T7, C5, C3, TP7, CP5, CP3), Center (FCz, FC1, Cz, FC2, CP1, CPz, CP2), Right (FC4, FC6, C4, C6, T8, CP4, CP6, TP8), and Posterior (P1, Pz, P2, P3, PO1, PO2, P4, P7, O1, Oz, O2, P8). Linear mixed effects models were fit to the data using the *lme4* (v1.1-17) and *lmerTest* (v3.0-1) packages in R (R Development Core Team, 2010) with Animacy, Auxiliary, and ROI as fixed effects and random intercepts by subjects and by items and random slopes for Animacy, Auxiliary, and their interaction by subjects and by items. Auxiliary and ROI were sum coded and Animacy was centered to avoid collinearity effects. Model criticism was applied after an initial model fit to all data, removing data falling outside 2.5 standard deviations, resulting in 2.80% data loss (Baayen, 2008), and the model was refit to the truncated data. Since our predictions concern the effect of subject noun's animacy on auxiliary form, three-way interactions for both initial and final models were followed up with a Bonferroni corrected post-hoc trend analysis of the effects of Animacy within Auxiliary and ROIs using the emtrends() and test() functions from the *emmeans* (v1.6.2-1) package. We report the analyses on both the full data and the truncated data that passed model criticism, though the pattern of effects did not greatly differ between analyses using all data and only data that passed model criticism.

Results

Comprehension Accuracy

The mean response accuracy rate was 94%.

ERPs

Figures 3 and 4 show the ERP effects elicited by *avere* and *essere* given low and high subject animacy. The initial (all data) linear mixed effect model reported in Table 4 revealed a significant effect of ROI ($F = 47.164$, $p < .001$) and a significant two-way interaction between Animacy and ROI ($F = 2.827$, $p = .023$). These effects were qualified by a significant three-way interaction between Animacy, Auxiliary and ROI ($F = 2.826$, $p = .023$). The final (data truncated) linear mixed effect model reported in Table 5 revealed a significant effect of ROI ($F = 60.041$, $p < .001$) and a significant two-way interaction between Animacy and ROI ($F = 4.599$, $p = .001$). However, this analysis revealed that the three-way interaction between Animacy, Auxiliary, and ROI was not robust ($F = 1.949$, $p = .099$).

Avere*$_{HAVE}$ *ERPs

The initial cluster analysis revealed one significant negative cluster for *avere*$_{\mathrm{HAVE}}$ from 416 to 568 msec involving all frontal channels except for F1 ($t = -2.066$, $p = .038$; see supplemental Figure S4). Figure 5 shows the effect of structural prediction on *avere* auxiliaries when preceded by high to low animate subject nouns. Post-hoc trend analysis of the initial (all data) linear mixed effect model revealed a significant effect of Animacy in the Anterior ROI for *avere* (*Est.* = 0.154, $t = 3.300$, $p = .010$; Table 4) such that amplitude in the Anterior ROI became more negative as animacy decreased. All other effects of Animacy in other ROIs were not significant. Post-hoc trend analysis of the final (data truncated) linear mixed effect model also revealed a significant effect of Animacy in the Anterior ROI for *avere* (*Est.* = 0.139, $t = 3.414$, $p = .006$; Table 5) such that amplitude in the Anterior ROI became more negative as animacy decreased. All other effects of Animacy in other ROIs were again not significant.

Essere*$_{\mathrm{BE}}$ *ERPs

The initial cluster analysis did not discover any significant clusters for *essere*$_{\mathrm{BE}}$ (all $p > 0.446$). Figure 5 shows the effect of structural prediction on *essere* auxiliaries when preceded by high to low animate subject nouns. Post-hoc trend analysis of the initial (all data) linear mixed effect model revealed a marginal effect of Animacy in the Center ROI for *essere* (*Est.* = -0.121, *t* = -2.595, *p* = .095; Table 4). However, this marginal effect was not robust in the final (data truncated) linear mixed effect model where the post-hoc trend analysis revealed no significant effects of Animacy in any ROI (Animacy in the Center ROI for *essere*: *p* = .221; see Table 5).

Discussion

The aim of this study was to investigate whether comprehenders use semantic information to make predictions about upcoming syntactic structure. We examined ERP responses to Italian auxiliaries in present perfect constructions, manipulating subject noun animacy and auxiliary form. We hypothesised that if participants were making predictions about upcoming predicate structures based on animacy, they should be sensitive to auxiliary form given the structure predicted by the animacy of the subject, particularly in the case of the auxiliary *avere*$_{\text{HAVE}}$ which mismatches the passive/unaccusative structure predicted by highly inanimate nouns.

Based on this hypothesis, we examined the effect of Animacy within Auxiliary form (*avere* and *essere*) and ROI, which maintains the same word form in comparing across the Animacy range, and found differential ERP effects on auxiliaries given the animacy of the subject noun, suggesting that the animacy of the subject noun triggered comprehenders to predict the structure of an upcoming verb phrase, with auxiliary forms either matching or mismatching this prediction. Specifically, we found a frontal negativity spanning 400-600ms post-stimulus onset, which was elicited by *avere*$_{\text{HAVE}}$ auxiliaries when preceded by inanimate subjects, compared to animate ones. This effect is consistent with the proposal that inanimate subjects, being implausible as Agents of transitive or unergative VPs, cued the parser to predict a derived VP-structure. Such a structure, however, cannot take *avere*$_{\text{HAVE}}$ as its auxiliary, eliciting a violation of expectation reflected by the frontal negativity. Importantly, this effect was found on the auxiliary prior to the verb phrase itself, supporting the theory that comprehenders make structural predictions during real-time sentence processing. Importantly, this effect does not appear to be simply due to co-occurrence patterns between different nouns and *avere*$_{\mathrm{HAVE}}$ in Italian. As the auxiliary norming data in Table 2 shows, Overt realisation of *avere* (which includes both *avere* as auxiliary and *avere* as the verb to have meaning "to possess") is relatively rare both after Low animacy and High animacy nouns ("Overtly produced" column group in Table 2): under 10% for all animacy levels apart from 7. If participants were just relying on co-occurrence (and therefore, overt forms only) we would not expect to see the graded effect of animacy on processing of *avere* that we found. On the other hand, the effect we found is compatible with the pattern in the Total column where *avere* increases from 19% to 60% as animacy increases, and the No Overt Auxiliary column where *avere* increases from 17% to 46%, which suggests that participants were making predictions at the structural level.

By contrast, subject animacy did not have an effect on the ERP responses elicited by the *essere*$_{\mathrm{BE}}$ auxiliary. Since animate subjects can fill both Agent and Patient thematic roles, we proposed that a high-animacy noun does not provide strongly constraining cues as to which structural continuations are more likely to follow. This suggests that comprehenders do not formulate strong predictions for upcoming structure when encountering an animate noun, which is reflected in the fact that we did not see any evidence of prediction violation for *essere*$_{\mathrm{BE}}$ following animate nouns.

In addition to the overall pattern above, we also found that the more inanimate a subject noun phrase was rated, the greater the frontal negativity elicited in the 400-600ms time window. This suggests that structural predictions may be graded in a manner potentially similar to lexical predictions (DeLong, Urbach, and Kutas, 2005). Comprehenders have a strong preference to take inanimate subjects as derived from an underlying verb phrase internal position as in a passive or unaccusative so that they can receive a non-Agent thematic role (Gennari & MacDonald 2008). This preference may override other parsing preferences, such as those that assume an Agent-first (and therefore a transitive/unergative structure) analysis for clauses (Bever, 1970). Our results suggest that comprehenders may use this non-Agent inanimacy preference in a graded fashion to predict passive/unaccusative structures. Alternatively, the gradability of structural prediction may reflect competition between subject interpretations and their resulting structural commitments, similar to competition models for structural ambiguity resolution (MacDonald, Pearlmutter, and Seidenberg, 1994; McRae, Spivey-Knowlton, and Tanenhaus, 1998). The relative incompatibility of a subject noun with an Agent thematic role may guide comprehenders' commitment to predict a non-active versus active verb phrase structure, with non-active verb phrases outcompeting active ones when the subject noun is highly inanimate.

The lack of an effect on *essere*$_{\text{BE}}$ may be explained by the strength of this preference. It appears that comprehenders do not have a similar preference to take animate subjects as necessarily being Agents which also commits them to an unergative/transitive verb phrase, even when highly animate. This suggests that the any Agent-first parsing preference or constraint is relatively weak, with subject inanimacy playing the main role in determining what structures comprehenders predict. However, as previously mentioned, the lack of an effect on *essere*$_{\text{BE}}$ is also compatible with the distribution of this auxiliary, which is quite different from that of *avere*$_{\text{HAVE}}$ (see Table 2). While the highest animacy rated nouns elicited completions with or requiring *avere*$_{\text{HAVE}}$ over half the time (0.603), *essere*$_{\text{BE}}$ completions were much more common across our items, even among the more animate of nouns. Nouns rated 5 or 6 on our 7-point animacy scale were still more frequently completed with a verb phrase requiring *essere*$_{\text{BE}}$. Besides being used in passive and derived structures, the auxiliary *essere*$_{\text{BE}}$ is also used as a copula in copula constructions (in predicates such as *Il mio vicino è*$_{\text{BE}}$ *un pittore* 'My neighbour is a painter' or *La ragazza è*$_{\text{BE}}$ *felice* 'The girl is happy'), which are compatible with high-animacy subjects. This suggests that comprehenders may use the full range of animacy to make structural predictions, but that the wider distribution of *essere*$_{\text{BE}}$ and the structures it is compatible with mean it is less likely to be incompatible with the structure comprehenders predicted. *Avere*HAVE, it seems, is the more specialized of the two auxiliaries, leading it to be more sensitive to structural predictions.

Turning to ERP components, frontal and widespread negativities in the same time region have been observed in response to unexpected continuations following highly constraining contexts, in a variety of domains. They can be elicited at the level of morphosyntactic processing, specifically by gender marking violations (Otten and Van Berkum, 2009; Otten et al., 2007; Szewczyk and Schriefers, 2013; Wicha et al., 2004). Otten and VanBerkum (2009) reported a widespread negativity in the 200-600ms time region in response to a determiner that did not match the gender of the high-cloze completion relative to gender-congruent determiners. Such negativities can also be triggered by unmet expectations in anaphora resolution (Van Berkum, Brown, and Hagoort, 1999; Van Berkum, Brown, Hagoort, and Zwitserlood, 2003; Nieuwland and Van Berkum, 2006; Nieuwland and Van Berkum, 2008) and in situations where comprehenders are met with unexpected, but plausible, sentence continuations. For example, Wlotko and Federmeier (2012) observed a fronto-temporal negativity onsetting at 400-500ms post-stimulus onset in response to plausible but low-cloze continuations following highly constraining contexts, with greater amplitude after contexts with fewer possible continuations. The variety of contexts in which this kind of ERP response has been observed suggest that it may reflect a general index of unmet expectations, particularly following strongly constraining contexts, which is not specific to grammaticality or plausibility.

In addition to a cost for violated prediction, studies also often report a reduction in N400 amplitude for expected completions (Federmeier et al., 2007; Thornhill and Van Petten, 2012; Kuperberg et al., 2020; DeLong et al., 2014; Quante et al., 2018). In our study, we might have expected a reduction in N400 when the animacy of the subject noun better fit with the structural expectations of the thematic roles which required a certain auxiliary, namely higher animacy for *avere*HAVE (related to an Agent role), and lower animacy for *essere*BE (related to a non-Agent role).

However, we did not find any such effect, which may be explained by the nature of the stimuli we used. The N400 component is sensitive to ease of lexical access as modulated by the semantic relatedness to a highly predictable completion (e.g. Kutas and Hillyard, 1984; Lau, Phillips, and Poeppel, 2008; Thornhill and Van Petten, 2012). The studies which found an N400 alongside a late positivity (Federmeier et al., 2007; Thornhill and Van Petten, 2012; Kuperberg et al., 2020; DeLong et al., 2014; Quante et al., 2018) used nouns as critical words, following a sentential context. Such contexts created expectations for a noun with certain semantic features, which the noun could match more or less closely, influencing N400 amplitude. We think that this was not the case in our stimuli: inanimate nouns created an expectation for a certain type of structure (passive/unaccusative) which requires a particular auxiliary form (e.g. *essere*$_{\text{BE}}$ for passive/unaccusative); it is not an expectation for an auxiliary specifically or for the semantic properties of a particular auxiliary.

Based on these premises, we interpret the lack of an N400 difference on the auxiliary as suggesting that animacy did not robustly facilitate lexical access or integration for either auxiliary form. Rather, it appears that when participants encountered an auxiliary, expectations for the predicate type were checked against the auxiliary's form, with a mismatch signalling prediction error and, possibly, revision. This interpretation, while speculative, is compatible with our hypothesis that comprehenders were engaging in structural prediction and that the ERP response we observed is an index of prediction violation, rather than facilitation for expected completions.

The fact that we found a frontal negativity, but no P600 effect, is also in line with the nature of the stimuli we used. Variation in animacy levels had an effect on the cloze probability for each auxiliary form (Table 2); accordingly, we saw that the amplitude of the frontal negativity was graded, being modulated by animacy level (Figure 4), which is in line with previous findings (DeLong et al., 2005; Wlotko and Federmeier, 2012). However, although plausibility was not directly normed, all the sentences in our study were both grammatical and plausible: an inanimate noun followed by *avere*$_{\text{HAVE}}$, while infrequent, is not anomalous, e.g. *La chitarra ha*$_{\text{HAVE}}$ *perso due* *corde* ‘The guitar has lost two strings’. Therefore, we think that there were no significant effects of grammaticality or plausibility in our items, explaining the lack of a P600 effect.

Overall, this pattern of findings provides strong support for our hypothesis concerning inanimate nouns. It suggests that comprehenders formulate specific predictions about upcoming VP structures, which in turn affect processing of the auxiliary forms associated with those structures. When a prediction was disconfirmed by a mismatching auxiliary form (specifically, *avere*$_{\text{HAVE}}$ following an inanimate noun), it elicited a frontal negativity, indicating a prediction violation. Furthermore, the specific ERP component we found – a frontal negativity starting at 400ms – is indicative of an unexpected structural continuation, rather than a semantic effect. This supports our thesis that comprehenders were formulating predictions about specific syntactic structures and the relevant auxiliary forms, and constitutes strong evidence for a mechanism of structural prediction.

Conclusion

After some initial scepticism on the role of prediction mechanisms in language processing, there is now a growing body of research showing that sentential context can cause comprehenders to formulate predictions for specific categories or lexical items. This research adds to this body of research by demonstrating unambiguous evidence for predictions at the level of syntactic structure. Comprehenders draw on semantic information to formulate expectations for the upcoming syntactic structure of the input. In this particular case, comprehenders used the inanimacy of a subject noun to predict different upcoming VP structures. In naturalistic settings, structural predictions like these may even be more useful than lexical predictions given the different statistical properties of everyday language as compared to more typical experimental paradigms. Therefore, understanding the scope and limits of structural prediction will form an important element of research as we continue to explore the role of prediction in language comprehension.

References

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264. https://doi.org/10.1016/s0010-0277(99)00059-1

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, M., Johnson, K., & Roberts, I. (1989). Passive Arguments Raised. *Linguistic Inquiry, 20*(2), 219–251. http://www.jstor.org/stable/4178625

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390.

Bever, T. G., & Others. (1970). The cognitive basis for linguistic structures. *Cognition and the Development of Language, 279*(362), 1–61. https://pdfs.semanticscholar.org/72b8/cdab0bf33bc812615b9597ac446636e1c1ab.pdf

Brusini, P., Brun, M., Brunet, I., & Christophe, A. (2015). Listeners Exploit Syntactic Structure On-Line to Restrict Their Lexical Search to a Subclass of Verbs. *Frontiers in Psychology*, *6*, 1841.

Burzio, L. (1986). *Italian Syntax: A Government-Binding Approach.* Dordrecht; Boston: D. Reidel Publishing Company.

Chow, W.-Y., Lau, E., Wang, S., & Phillips, C. (2018). Wait a second! delayed impact of argument roles on on-line verb prediction. *Language, Cognition and Neuroscience*, *33*(7), 803–828. https://doi.org/10.1080/23273798.2018.1427878

Chow, W.-Y., Smith, C., Lau, E., & Phillips, C. (2016). A “bag-of-arguments” mechanism for initial verb predictions. *Language, Cognition and Neuroscience*, *31*(5), 577–596. https://doi.org/10.1080/23273798.2015.1066832

DeLong, K.A. (2009). *Electrophysiological explorations of linguistic pre-activation and its consequences during online sentence processing.* Doctoral dissertation. San Diego: University of California.

DeLong, K. A., Troyer, M., & Kutas, M. (2014). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure: Pre-processing in sentence comprehension. *Language and Linguistics Compass, 8*(12), 631–645. https://doi.org/10.1111/lnc3.12093

DeLong, K.A., Urbach, T. P., and Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117–1121.

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

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

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, *21*(5), 629–634.

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*, 293–321.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*(6), 641–655.

Federmeier, K. D., and Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, *41*(4), 469–495.

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133-146.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75–84.

Ferreira, F., & Chantavarin, S. (2018). Integration and Prediction in Language Processing: A Synthesis of Old and New. *Current Directions in Psychological Science, 27*(6), 443–448. https://doi.org/10.1177/0963721418794491

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 33*(4), 465–495.

Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.

Gennari, S. P., & Macdonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language, 58*(4), 161–187. https://doi.org/10.1016/j.jml.2007.07.004

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711-1725.

Hare, M., Elman, J. L., Tabaczynski, T., & McRae, K. (2009). The wind chilled the spectators, but the wine just chilled: Sense, structure, and sentence comprehension. *Cognitive Science*, *33*(4), 610–628.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Brain Research. Cognitive Brain Research*, *19*(1), 59–73. https://doi.org/10.1016/j.cogbrainres.2003.10.022

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, *86*, 157–171.

Ito, A., Gambi, C., Pickering, M. J., Fuellenbach, K., & Husband, E. M. (2020). Prediction of phonological and gender information: An event-related potential study in Italian. Neuropsychologia, 136, 107291. https://doi.org/10.1016/j.neuropsychologia.2019.107291

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kayne, R.S. (2010). A note on auxiliary alternations and silent causation. In R.S. Kayne, *Comparisons and Contrasts*, pp. 146–164.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*(2), 205–225. https://doi.org/10.1016/j.jml.2004.10.002

Kochari, A. R., & Flecken, M. (2019). Lexical prediction in language comprehension: a replication study of grammatical gender effects in Dutch. *Language, Cognition and Neuroscience*, *34*(2), 239–253.

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A Tale of Two Positivities and the N400: Distinct Neural Signatures Are Evoked by Confirmed and Violated Predictions at Different Levels of Representation. *Journal of Cognitive Neuroscience, 32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Brain Research. Cognitive Brain Research*, *17*(1), 117–129. https://doi.org/10.1016/s0926-6410(03)00086-7

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future*, pp. 190–207. Oxford University Press.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62,* 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M. and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Kutas, M. and Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. Nature Reviews. Neuroscience, 9(12), 920–933. https://doi.org/10.1038/nrn2532

Lau, E., Stroud, C., Plesch, S., and Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language, 98,* 74–88.

Lopez-Calderon, J. and Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8, 213*.

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

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676–703.

McDonald, S. A., and Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14,* 648–652.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition, 33*(7), 1174–1184.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the Influence of Thematic Fit (and Other Constraints) in On-line Sentence Comprehension. *Journal of Memory and Language, 38*(3), 283–312. https://doi.org/10.1006/jmla.1997.2543

Neville, H., Nicol, J., Barss, A., Forster, K., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*, 151–165.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife, 7.* https://doi.org/10.7554/eLife.33468

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: evidence from ERPs. *Brain Research, 1118*(1), 155–167. https://doi.org/10.1016/j.brainres.2006.08.022

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language, 106*(2), 119–131. https://doi.org/10.1016/j.bandl.2008.05.001

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Otten, M., and Van Berkum, J.J.A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes, 45*, 464–498.

Otten, M., and Van Berkum, J.J.A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain research*, *1291*, 92-101.

Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. In *annual meeting of the Berkeley Linguistics Society* (Vol. 4, pp. 157-190).

Phillips, C. (2006). The real-time status of island phenomena. *Language, 82*, 795–823.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin, 144*(10), 1002–1044. https://doi.org/10.1037/bul0000158

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

R Development Core Team (2010) *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna. http://www.R-project.org

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(2), 344–354.

Schwanenflugel, P. J., and White, C. R. (1991). The influence of paragraph information on the processing of upcoming words. *Reading Research Quarterly, 26,* 160–177.

Stanovich, K. E., and West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 658.

Stanovich, K. E., & West, R. F. (1983). The generalizability of context effects on word recognition: A reconsideration of the roles of parafoveal priming and sentence context. *Memory & Cognition, 11*(1), 49–58.

Staub, A. and Clifton Jr, C. (2006). Syntactic prediction in language comprehension: evidence from *either… or. Journal of Experimental Psychology: Learning, Memory and Cognition, 32,* 425–436.

Szewczyk, J.M., and Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language, 68,* 297–314.

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology, 83*(3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007

Traxler, M. J., and Foss, D. J. (2000). Effects of sentence constraint on priming in natural language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1266–1282.

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Brain Potentials. *Journal of Memory and Language, 41*(2), 147–182. https://doi.org/10.1006/jmla.1999.2641

van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology, 40*(2), 235–248. https://doi.org/10.1111/1469-8986.00025

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, memory, and Cognition, 31*, 443–467.

Wagers, M., and Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics, 45,* 395–433.

Wicha, N.Y., Moreno, E.M., and Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16,* 1272–1288.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage, 62*(1), 356–366. https://doi.org/10.1016/j.neuroimage.2012.04.054

Wlotko, E. W., & Federmeier, K. D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex, 68,* 20-32.

Yoshida, M., Walsh Dickey, M., and Sturt, P. (2013). Predictive processing of syntactic structure: sluicing and ellipsis in real-time sentence processing. *Language and Cognitive Processes, 28,* 272–302.

Table 1

Function of avere$_{HAVE}$ *and* essere$_{BE}$ *auxiliaries in Italian.*

	Italian Form	Required structure	Preferred subject interpretation	Example
HAVE	*avere*	Underived (transitive/unergative)	Agent/Causer	*Il bambino ha$_{HAVE}$ mangiato un biscotto* (The child has eaten a biscuit)
BE	*essere*	Derived (passive/unaccusative)	Non-Agent/Non-Causer	*Il cameriere è$_{BE}$* stato licenziato (The waiter has been fired)

Table 2
Normed cloze probabilities for avere$_{\text{HAVE}}$ *and* essere$_{\text{BE}}$ *by animacy level. Animacy scores are binned (by rounding down to the nearest integer) for presentation purposes.*

	Overtly produced		No overt auxiliary produced – auxiliary required by verb			Total	
Animacy	*essere*$_{\text{BE}}$	*avere*$_{\text{HAVE}}$	Either	*essere*$_{\text{BE}}$	*avere*$_{\text{HAVE}}$	*essere*$_{\text{BE}}$	*avere*$_{\text{HAVE}}$
1	0.598	0.019	0.000	0.211	0.173	0.808	0.192
2	0.517	0.006	0.003	0.107	0.366	0.625	0.372
3	0.479	0.058	0.054	0.090	0.320	0.569	0.377
4	0.320	0.081	0.021	0.284	0.293	0.605	0.374
5	0.253	0.063	0.080	0.280	0.323	0.533	0.387
6	0.310	0.072	0.048	0.190	0.379	0.500	0.452
7	0.259	0.138	0.003	0.135	0.466	0.393	0.603
Average	**0.389**	**0.063**	**0.030**	**0.185**	**0.333**	**0.574**	**0.395**

Table 3

Sample stimuli demonstrating the use of *avere*HAVE and *essere*BE with a low animacy noun (*impronte* ‘footprints’) and a high animacy noun (*boscaiolo* ‘lumberjack’). The auxiliary form tracks the structure of the upcoming verb phrase, with *avere* selected by transitive (a,c) and unergative verb phrases and *essere* selected by unaccusative (b) and passive (d) verb phrases.

Subject	Animacy	Aux	Sentence
impronte “footprints”	LOW	AVERE	a. Le impronte hanno condotto la polizia al colpevole. *The footprints have.*3SG *led the police to.the culprit* “The footprints have led police to the culprit.”
		ESSERE	b. Le impronte sono scomparse sotto la neve. *The footprints be.*3SG *disappeared under the snow* “The footprints have disappeared under the snow.”
boscaiolo “lumberjack”	HIGH	AVERE	c. Un boscaiolo ha trovato la volpe ferita. *A lumberjack have.*3SG *found the fox wounded* “A lumberjack has found the wounded fox.”
		ESSERE	d. Un boscaiolo è stato colpito da un fulmine. *A lumberjack be.*3SG *been struck by a lightning* “A lumberjack has been struck by lightning.”

Table 4

Type III Analysis of Variance Table (with Satterthwaite's method) and the post-hoc trend analysis of Animacy within Auxiliary and ROI (Bonferroni corrected) for the initial (all data) linear mixed effects model for evoked potentials to *avere*$_{\text{HAVE}}$ and *essere*$_{\text{BE}}$ from 400-600 msec. Model: Voltage ~ Animacy*Auxiliary*ROI + (1 + Animacy*Auxiliary || Subjects) + (1 + Animacy*Auxiliary || Items)

	Sum Sq	Mean Sq	df	Den df	F	p
Animacy	7.78	7.78	1	15554	0.670	.413
Auxiliary	19.05	19.05	1	4517	1.641	.200
ROI	2190.48	547.62	4	15555	47.164	<.001 ***
Animacy: Auxiliary	9.23	9.23	1	15555	0.795	.373
Animacy: ROI	131.30	32.83	4	15555	2.827	.023 *
Auxiliary: ROI	47.42	11.86	4	15555	1.021	.395
Animacy: Auxiliary: ROI	131.24	32.81	4	15555	2.826	.023 *

		Estimate	Std. Err.	*t* value	*p* value
Avere	Anterior	0.154	0.047	3.300	.010 **
	Center	-0.002	0.047	-0.033	1.00
	Left	-0.021	0.047	-0.457	1.00
	Posterior	-0.113	0.047	-2.428	.152
	Right	-0.012	0.047	-0.265	1.00
Essere	Anterior	-0.000	0.047	-0.011	1.00

	Center	-0.121	0.047	-2.595	.095 .
	Left	-0.022	0.047	-0.011	1.00
	Posterior	0.005	0.047	0.105	1.00
	Right	0.013	0.047	0.279	1.00

Table 5

Type III Analysis of Variance Table (with Satterthwaite's method) and the post-hoc trend analysis of Animacy within Auxiliary and ROI (Bonferroni corrected) for the final (data-truncated via model criticism) linear mixed effects model for evoked potentials to *avere*$_{\text{HAVE}}$ and *essere*$_{\text{BE}}$ from 400-600 msec. Model: Voltage ~ Animacy*Auxiliary*ROI + (1 + Animacy*Auxiliary || Subjects) + (1 + Animacy*Auxiliary || Items)

	Sum Sq	Mean Sq	df	Den df	F	p
Animacy	1.42	1.42	1	95.7	0.169	.682
Auxiliary	6.47	6.47	1	2354.6	0.767	.381
ROI	2025.58	506.40	4	15116.8	60.041	<.001 ***
Animacy: Auxiliary	7.49	7.49	1	13564.4	0.888	.346
Animacy: ROI	155.16	38.79	4	15116.4	4.599	.001 **
Auxiliary: ROI	49.81	12.45	4	15115.9	1.477	.206
Animacy: Auxiliary: ROI	65.75	16.44	4	15116.1	1.949	.099 .

		Estimate	Std. Err.	*t* value	*p* value
Avere	Anterior	0.139	0.041	3.414	.006 **
	Center	0.014	0.040	0.332	1.00
	Left	-0.007	0.040	-0.169	1.00
	Posterior	-0.097	0.041	-2.399	.165
	Right	-0.018	0.040	-0.440	1.00
Essere	Anterior	0.053	0.041	1.315	1.00

	Center	-0.092	0.040	-2.290	.221
	Left	-0.054	0.040	-1.353	1.00
	Posterior	-0.023	0.041	-0.561	1.00
	Right	0.028	0.040	0.692	1.00

Figure 1. Structures for *avere*$_{HAVE}$ and *essere*$_{BE}$. A) Structures that select *avere* are transitive and unergative, assigning Agent directly to their subject position. B) Structures that select *essere* are passive and unaccusative, assigning Theme to their derived subject from an underlying direct object position.

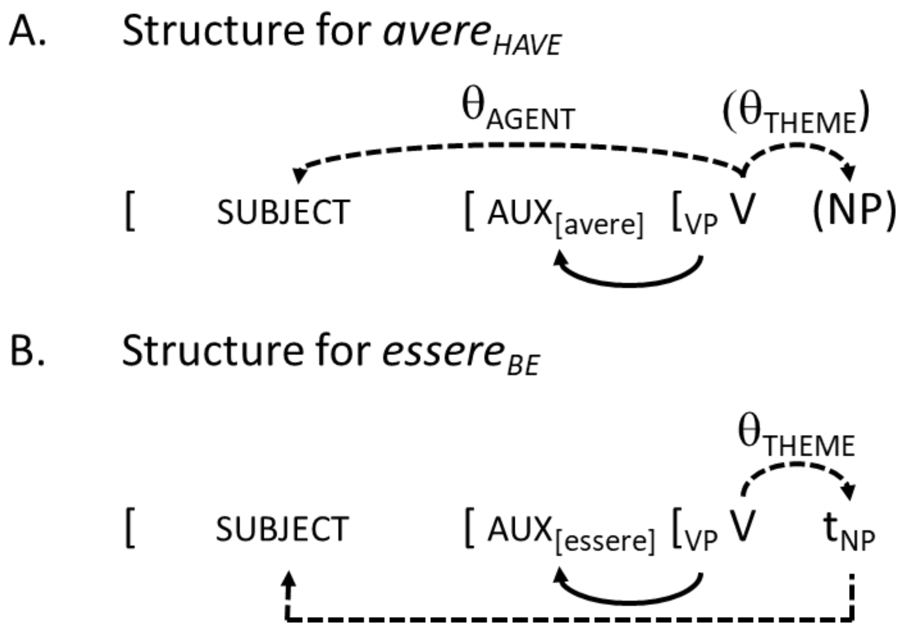

Figure 2. The role of animacy as a semantic cue for verb phrase structure and the structural relationships between a grammatical subject, auxiliary, and verb phrase. A) The animacy of the subject acts as a semantic cue to the comprehender to expect an underived or derived verb phrase. The form of auxiliary *avere*$_{\text{HAVE}}$ or *essere*$_{\text{BE}}$ is then structurally dependent on whether the upcoming verb phrase is underived or derived respectively. B) More animate subjects are likely to be initially interpreted as either Agents with underived verb phrases (transitive/unergative) that select for *avere* or Themes with derived verb phrases (passive/unaccusative) that select for *essere*. C) More inanimate subjects are likely to be initially interpreted only as Themes which require derived verb phrases that select for *essere*.

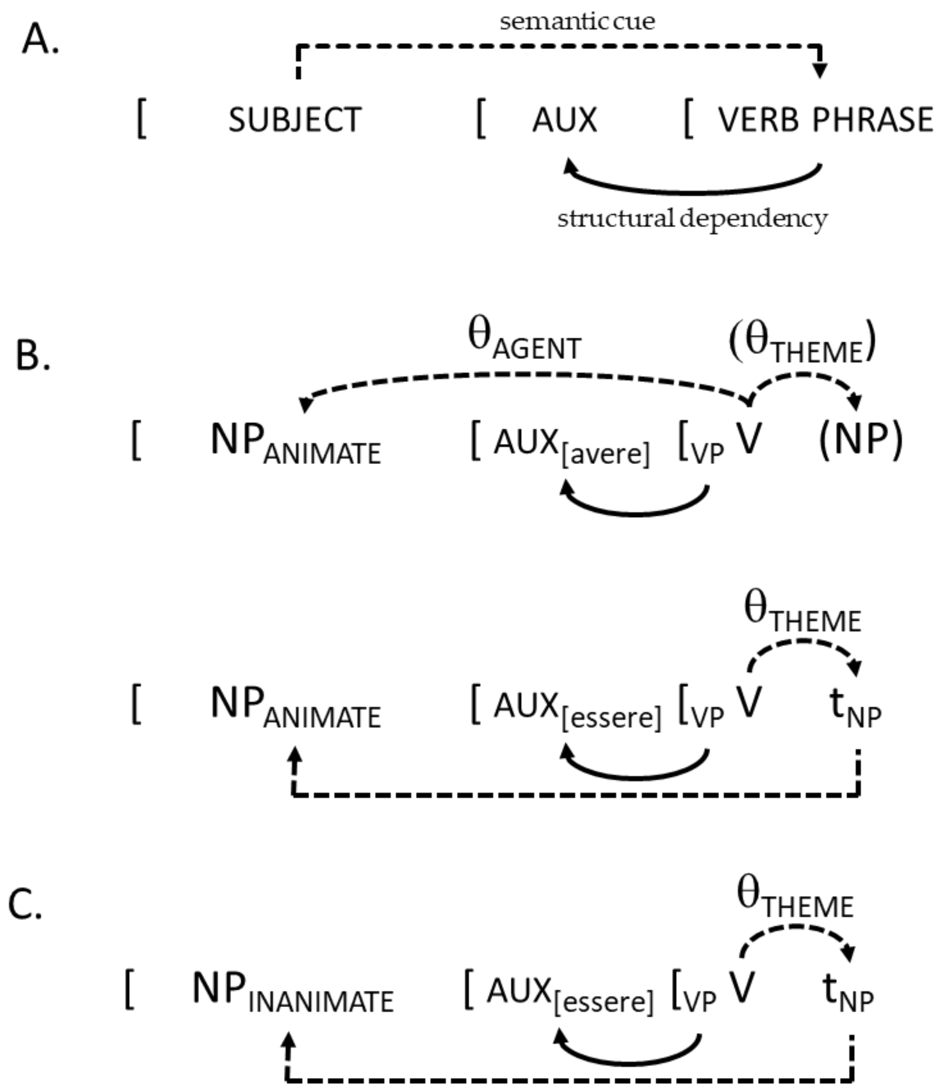

Figure 3. Voltage maps comparing *avere*$_{\mathrm{HAVE}}$ and *essere*$_{\mathrm{BE}}$ in 100 msec windows starting from 0-100 msec to 700-800 msec (Low Animacy - High Animacy).

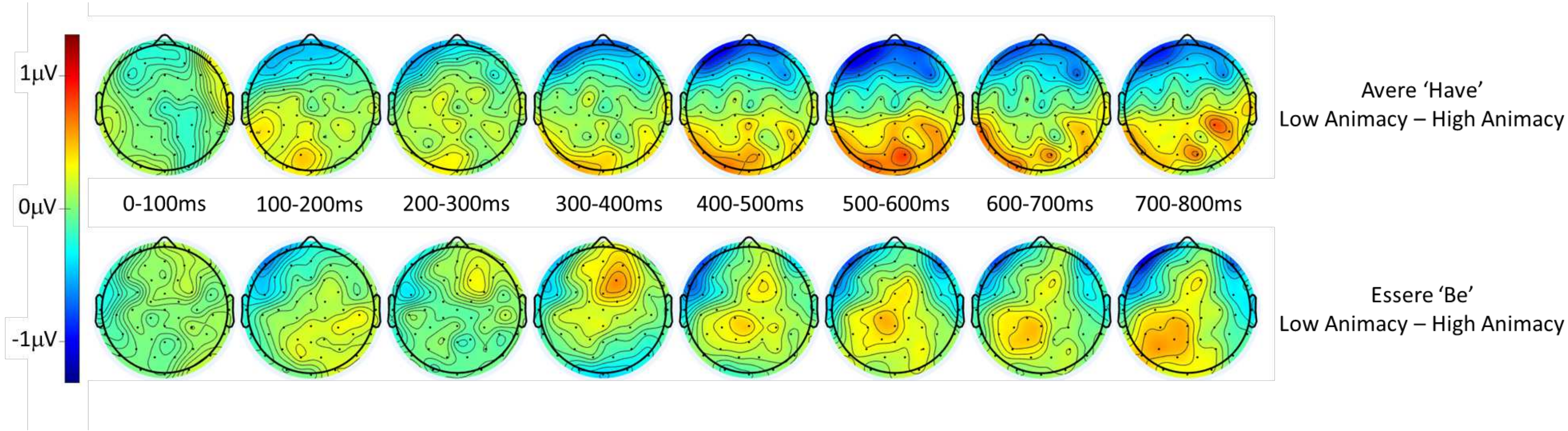

Figure 4. Voltage maps comparing *avere*HAVE and *essere*BE from 400-600 msec (Low Animacy - High Animacy). ERP plots show grand averaged waveforms to target auxiliaries *avere* ‘have’ and *essere* ‘be’ at electrode AFz. Animacy was dichotomized into high and low animacy and waves were low pass filtered at 15 Hz.

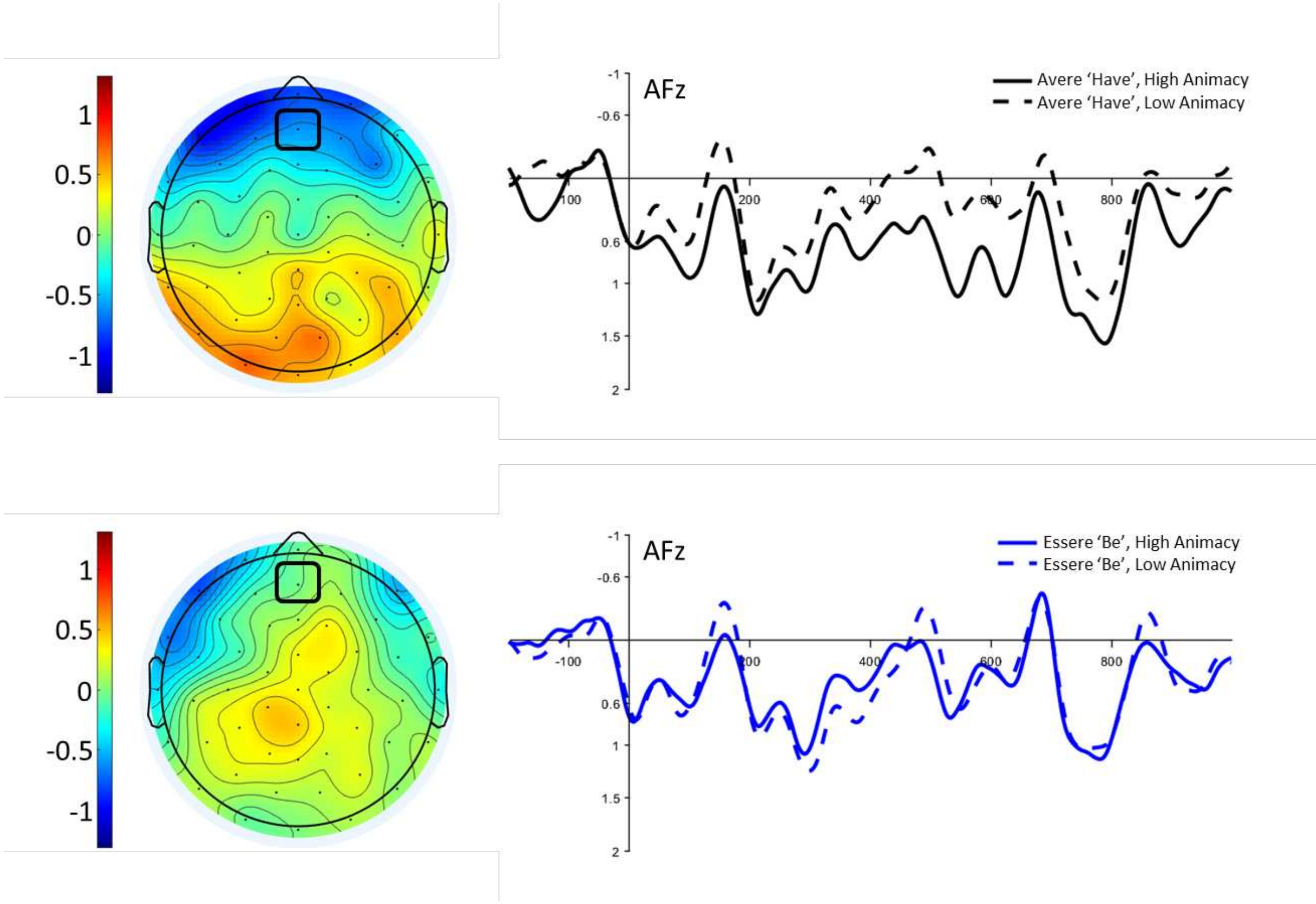

Figure 5. Voltage maps comparing *avere*HAVE and *essere*BE from 400-600 msec (Low Animacy - High Animacy). The line plots show the model regressions and confidence intervals for the effect of Animacy in the Anterior ROI (electrodes AF7, Fp1, FPz, Fp2, AF8, F5, AF3, AFz, AF4, F6, F1, Fz, and F2), for each of the two auxiliaries. Open circles are the average subject means for each item.