Neuropsychologia 155 (2021) 107797

Contents lists available at ScienceDirect

Neuropsychologia

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

The role of crosslinguistic differences in second language anticipatory processing: An event-related potentials study

José Alemán Bañón [a,*], Clara Martin [b,c]

[a] Centre for Research on Bilingualism, Department of Swedish and Multilingualism, Stockholm University, Universitetsvägen 10 (D355), 10691, Stockholm, Sweden
[b] Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 20009, Donostia-San Sebastián, Spain
[c] Ikerbasque, Basque Foundation for Science, María Díaz de Haro 3, 48013, Bilbao, Spain

ARTICLE INFO

Keywords:
Prediction
Crosslinguistic differences
Possessive pronouns
ERP
N400
P600

ABSTRACT

The present study uses event-related potentials to investigate how crosslinguistic (dis)similarities modulate anticipatory processing in the second language (L2). Participants read predictive stories in English that made a genitive construction consisting of a third-person singular possessive pronoun and a kinship noun (e.g., *his mother*) likely in an upcoming continuation. The possessive pronoun's form depended on the antecedent's natural gender, which had been previously established in the stories. The continuation included either the expected genitive construction or an unexpected one with a possessive pronoun of the opposite gender. We manipulated crosslinguistic (dis)similarity by comparing advanced English learners with either Swedish or Spanish as their L1. While Swedish has equivalent possessive pronouns that mark the antecedent's natural gender (i.e., *hans/hennes* "his/her"), Spanish does not. In fact, Spanish possessive pronouns mark the syntactic features (number, gender) of the possessed noun (e.g., *nosotros queremos a nuestra madre* "we-$_{\text{MASC}}$ love our-$_{\text{FEM}}$ mother-$_{\text{FEM}}$). Twenty-four native speakers of English elicited an N400 effect for prenominal possessives that were unexpected based on the possessor noun's natural gender, consistent with the possibility that they activated the pronoun's form or its semantic features (natural gender). Thirty-two Swedish-speaking learners yielded a qualitatively and quantitatively native-like N400 for unexpected prenominal possessives. In contrast, twenty-five Spanish-speaking learners showed a P600 effect for unexpected possessives, consistent with the possibility that they experienced difficulty integrating a pronoun that mismatched the expected gender. Results suggest that differences with respect to the features encoded in the activated representation result in different predictive mechanisms among adult L2 learners.

1. Introduction

Research on language comprehension has amassed extensive evidence that native speakers of a language do not just passively integrate words after they become available in the bottom-up input. Instead, they use a variety of linguistic and nonlinguistic cues to anticipate likely continuations at all levels of linguistic representation (e.g., Altmann and Mirković, 2009; Federmeier, 2007; Huettig, 2015; Kuperberg and Jaeger, 2016; Kutas et al., 2011; Pickering and Gambi, 2018; Pickering and Garrod, 2013). In light of this evidence, most theoretical models of language comprehension agree that language processing is, at least to some extent, predictive, although there remains disagreement on fundamental issues, including the ubiquity of predictive processing in language comprehension (e.g., Clark, 2013; Huettig and Mani, 2016; Kuperberg and Jaeger, 2016; Nieuwland and Kazanina, 2020; Pickering and Gambi, 2018).

In contrast, the involvement of anticipatory mechanisms in second language "L2" processing is less clear-cut. This is, in part, because research on this topic is less abundant, although both Kaan (2014) and Phillips and Ehrenhofer (2015) already noted a surge in interest. A number of studies have claimed that adult L2 learners do not generate predictions to the same extent as native speakers (e.g., Covey et al., Unpublished results; Dijkgraaf et al., 2019; Grüter et al., 2012, 2014; Ito et al., 2017a; 2017b; Lew-Williams and Fernald, 2010; Martin et al., 2013; Mitsugi and MacWhinney, 2016). Common explanations for this divergence include general differences between L1 and L2 acquisition (e.g., Grüter et al., 2012), weaker lexicosemantic representations in the L2 (e.g., Dijkgraaf et al., 2019; Grüter et al., 2012; Hopp, 2013), reduced

* Corresponding author.
E-mail address: jose.aleman.banon@biling.su.se (J. Alemán Bañón).

https://doi.org/10.1016/j.neuropsychologia.2021.107797
Received 3 November 2020; Received in revised form 24 December 2020; Accepted 7 February 2021
Available online 19 February 2021

experience with the L2 (e.g., van Bergen and Flecken, 2017), and a lack of the necessary processing/executive resources to generate predictions (e.g., Ito et al., 2017a; 2017b; Kaan et al., 2010; Martin et al., 2013; Mitsugi and MacWhinney, 2016; but see Ryskin et al., 2020).

This position aligns well with Grüter et al.'s RAGE hypothesis (2014, 2017), according to which adult L2 learners generally have Reduced Ability to Generate Expectations across all domains of grammar. The proponents of the RAGE hypothesis argue that, even in cases where predictive processing does not require extensive experience with words or morphosyntactic computations; or when the target linguistic property is similar in the L1 and the L2; or even when there are no time constraints for prediction generation (as in Grüter et al., 2014, 2017), advanced learners are less likely to rely on anticipatory mechanisms to the same degree as native speakers (Grüter et al., 2017, p. 26). Although factors such as proficiency and crosslinguistic differences did not play a significant role in their studies (e.g., Grüter et al., 2014, 2017), the proponents of the RAGE hypothesis contemplate that such factors might modulate learners' predictive abilities.

Other investigations have shown that predictive processing in the L2 is possible but contingent upon factors that are known to impact L2 processing in general. This includes factors inherent to the L2 learner, such as global proficiency (e.g., Dussias et al., 2013; Hopp, 2013; Hopp and Lemmerth, 2018), robust and stable lexical representations in the L2 (e.g., Hopp, 2013), or individual differences in processing speed and working memory (e.g., Hopp, 2013; Ito et al., 2018), as well as factors related to the L2 grammar itself. For example, Hopp (2015) found that learners across the proficiency spectrum were more likely to use lexicosemantic rather than morphosyntactic cues predictively (see also Mitsugi and MacWhninney, 2016), and work by Kaan et al. (2016) and Covey et al. (Unpublished results) suggests that L2 learners are less likely to predict abstract syntax. Learners also seem better able to generate predictions for L2 properties that exist or are realized similarly in the L1, although very few studies have addressed this question directly (e.g., Dussias et al., 2013; Hopp and Lemmerth, 2018; van Bergen and Flecken, 2017).

The latter evidence resonates more with a proposal by Kaan (2014) that L2 learners rely on the same predictive mechanisms as native speakers, although this ability can be hindered by a variety of factors. Under Kaan's proposal, conflicts between how the L1 and the L2 encode linguistic information can obscure an otherwise native-like ability to predict in the L2. Likewise, individual differences in cognitive factors that are known to modulate prediction in native speakers, such as working memory and processing speed (Huettig, 2005; Huettig and Janse, 2016; McDonald, 2006), can also impact anticipatory processing in the L2 (Hopp, 2013; Ito et al., 2018; see Kaan, 2014 for a more extensive list of factors).

In sum, both the RAGE hypothesis and Kaan's proposal assume that adult L2 learners are generally less likely to predict than native speakers, although for different reasons. While the proponents of the RAGE hypothesis argue for general limitations on the ability to generate predictions, Kaan (2014, p. 257, abstract) explicitly posits that predictive mechanisms are qualitatively (and potentially quantitatively) the same in the L1 and the L2. Crucially, both proposals contemplate the possibility that crosslinguistic differences impact predictive processing in the L2, which is the question we address here. Both proposals also highlight that the role of prediction in the L2 has important implications for L2 acquisition research. For example, Chang et al. (2006) established a clear link between the ability to generate predictions and language development (see also Jaeger and Snider, 2013; Phillips and Ehrenhofer, 2015; cf. Kaan, 2015). Importantly, this question also has implications outside the L2 literature. For example, some researchers have used the evidence that adult L2 learners have intact comprehension but do not predict (e.g., Martin et al., 2013) as an argument to advance proposals undermining the ubiquity of prediction in language comprehension (e.g., Huettig and Mani, 2016; Pickering and Gambi, 2018; Pickering and Garrod, 2013).

Here, we investigate how L1-L2 (dis)similarity affects learners' predictions. We examine a case in which the activated representation (i.e., what is predicted) encodes feature information that is either similar or different in the L1s of two groups of L2 learners. In the L2 acquisition/processing literature, some prominent theories conceptualize negative transfer effects in a similar manner (e.g., Lardiere, 2009; Tokowicz and MacWhinney, 2005; see also Alemán Bañón et al., 2014; 2017; 2018; Foucart and Frenck-Mestre, 2011; 2012; Jackson and Dussias, 2009; Sabourin and Haverkort, 2003; Shimanskaya and Slabakova, 2017). For example, within the generative framework, the Feature Reassembly Hypothesis (e.g., Lardiere, 2009) posits that differences between how the L1 and the L2 assemble features in lexical items may present obstacles for L2 development. Likewise, the Competition Model (e.g., Tokowicz and MacWhinney, 2005) argues that L2 learners are less likely to show native-like processing for properties that are shared between the L1 and the L2, but realized differently. We will examine whether L1-L2 (dis)similarities also impact learners' predictions.

Our study uses event-related potentials "ERPs", which are brain responses that are time-locked to relevant stimuli and are differentially impacted by different linguistic manipulations. We focus on two recurring ERP components in the prediction literature, the *N400* (e.g., DeLong et al., 2005; Federmeier et al., 2007; Kutas and Federmeier, 2011; Kutas and Hillyard, 1984; Lau et al., 2008; Mantegna et al., 2019) and a post-N400 positivity referred to as the *Late Anterior Positivity* or the *Late Frontal Positivity* (e.g., Brothers et al., 2020; DeLong et al., 2011; 2014; Federmeier et al., 2007; Ito et al., 2020; Martin et al., 2013; Van Petten and Luka, 2012). The N400 is a negative-going wave that tends to occur between ~300 and 500 ms in central-posterior electrodes, and that is sensitive to both semantic integration and different aspects of lexical access/retrieval, including predictability (Kutas and Hillyard, 1984; see Kutas and Federmeier, 2011 and Lau et al., 2008 for reviews). For example, DeLong et al. (2005) presented English native speakers with high-constraint sentences such as *The day was breezy so the boy went outside to fly*... and compared continuations that were expected, such as *a kite*, to those that were unexpected but plausible, such as *an airplane* (based on independent offline cloze probability ratings). Their results showed a larger N400 for unexpected relative to expected nouns, in line with accounts of the N400 as an index of lexical retrieval difficulty. This effect, however, cannot unambiguously be attributed to lexical predictability independently from integration, since less predictable words are also harder to integrate (e.g., Bornkessel-Schlesewsky and Schlesewsky, 2019; Van Petten and Luka, 2012; but see Mantegna et al., 2019). Crucially, DeLong et al. (2005) also manipulated the phonological onsets of the expected and unexpected nouns (i.e., vowel vs. consonant) such that they would be preceded by the different allomorphs of the English indefinite article *a/an*, which have approximately the same meaning and should not differ with respect to integration. The authors found a larger N400 effect for articles that were unexpected based on the phonological properties of the predicted yet unencountered nouns, providing some of the strongest evidence linking the N400 to lexical predictability. Other studies manipulating prenominal material as a function of other rules, such as gender agreement, have made similar claims (e.g., Foucart et al., 2014; 2016; Van Berkum et al., 2005; Wicha et al., 2004), although the nature of these prediction effects differs across studies (see Fleur et al., 2020).[1]

The Late Anterior Positivity is a positive deflection captured by frontal electrodes between ~500 and 900 ms (i.e., after the N400) that is

[1] Recent failures to replicate these findings have lead researchers to question whether predictive processing is as detailed or ubiquitous as traditionally assumed (e.g., Ito et al., 2017b; Kochari and Flecken, 2018; Nieuwland et al., 2018; see DeLong et al., 2017 for counterarguments to Ito et al., 2017 and Ito et al.'s rebuttal, 2017c; see Yan et al., 2017 for a commentary on Nieuwland et al.'s methodology; and see Urbach et al., 2020 for evidence in support of DeLong et al.'s original claims).

argued to signal the cost of an unmet lexical prediction (e.g., Brothers et al., 2015; 2020; DeLong et al., 2011; 2014; Federmeier et al., 2007; Van Petten and Luka, 2012). For example, in DeLong et al.'s study (2005, 2011), the N400 effect for unexpected (relative to expected) nouns was followed by a Late Anterior Positivity (see also Federmeier et al., 2007; Fleur et al., 2020; Foucart et al., 2014; Ito et al., 2020; Martin et al., 2013; cf. Lau et al., 2013). Recent reports have shown that the Late Anterior Positivity is more likely to emerge when the preceding context is rich and globally constraining (e.g., Brothers et al., 2020; Kuperberg et al., 2020), suggesting that this component indexes the reevaluation of the discourse situation built by comprehenders.

Here, we will examine these two components in a study using a "prenominal material" paradigm to probe anticipatory processing in short discourse contexts in English. In this paragraph, we only provide a succinct description of the study's rationale, which we will elaborate on in Sections 3 and 4. The contexts made a genitive construction consisting of a third-person singular possessive pronoun and a kinship noun of the opposite semantic gender (e.g., *his mother*) highly expected in an upcoming sentence. The continuation included either the expected genitive construction or an unexpected one with a possessive pronoun of the opposite gender (e.g., *his mother* vs. *her aunt*). This gender alternation in the possessive pronoun depended on whether the pronoun's antecedent was a male or a female (e.g., ***Tom**'s mother* = ***his** mother; **Ruth**'s mother* = ***her** mother*), which had been previously established in the contexts. Similar to previous studies manipulating prenominal articles/adjectives (e.g., DeLong et al., 2005; 2011; Fleur et al., 2020; Foucart et al., 2014; Ito et al., 2017a; 2020; Martin et al., 2013; Otten and van Berkum, 2009; van Berkum et al., 2005; Wicha et al., 2004), we examined the N400 for unexpected prenominal possessives, and both the N400 and the Late Anterior Positivity for unexpected nouns.

The role of L1-L2 similarity was examined by testing advanced L2 learners of English with either Swedish or Spanish as their L1. Crucially, while Swedish has third-person singular possessive pronouns that, similar to English, mark the *possessor noun*'s natural gender (i.e., *hans/hennes* "his/her"), Spanish does not. In Spanish, the possessor noun's natural gender has no impact on the form of third-person possessive pronouns. Thus, a genitive construction such as *su madre* "$_{\text{POSS}}$ mother" can refer either to Tom's mother (his mother) or to Ruth's (her mother). In fact, the Spanish possessive pronoun system provides evidence for syntactic agreement with the *possessed noun*, in contrast to English (e.g., *nosotros queremos a nuestra madre* "we-$_{\text{MASC}}$ love our-$_{\text{FEM}}$ mother-$_{\text{FEM}}$"). Given that, in the current study, the expected "possessive + noun" combination mismatched in natural gender (e.g., *his mother, her brother*), we were able to examine whether crosslinguistic differences with respect to the realization of possessive pronouns between English/Swedish and Spanish would impact the learners' predictions (e.g., Lardiere, 2009; Tokowicz and MacWhinney, 2005). We will start by reviewing the extant literature on L2 predictive processing, with a focus on studies that have directly or indirectly examined the contribution of L1-L2 (dis)similarity.

2. Literature review

An early study by Martin et al. (2013) provides indirect evidence that adult L2 learners might not preactivate rules that are unique to the L2. In their conceptual replication of DeLong et al.'s study (2005), both native speakers of English and advanced L1-Spanish L2-English learners showed an N400 effect for nouns that were unexpected based on sentence constraint. However, only the L1-English speakers showed an N400 effect for indefinite articles that were unexpected based on whether the expected nouns began with a consonant or a vowel (i.e., *a/an*) (see also Ito et al., 2017a), a rule that Spanish lacks. This was the case, even though Martin et al. showed that a comparable group of learners were sensitive to online violations of the *a/an* rule. Foucart et al. (2014) followed up on this question in a study where the activated representation, gender agreement, was shared by the learners' L1 and L2. The authors found that, upon reading high-constraint sentences, both Spanish native speakers and L1-French L2-Spanish learners anticipated specific nouns and their gender, as marked on prenominal articles (e.g., Wicha et al., 2004). In particular, they found an N400 effect for unexpected prenominal articles, and both an N400 effect and a Late Anterior Positivity for unexpected nouns.

Only a few studies have directly examined the role of crosslinguistic differences in L2 anticipatory processing (e.g., Dussias et al., 2013; Hopp and Lemmerth, 2018; van Bergen and Flecken, 2017), and all of them used the visual world paradigm "VWP". In these studies, participants are instructed to locate one of several items on a visual display, and the carrier sentence either does or does not include an informative cue regarding the upcoming item. For example, in languages like Italian and Spanish, all nouns are classified as masculine or feminine, and this is marked on the preceding article (e.g., *la casa* "the-$_{\text{FEM}}$ house-$_{\text{FEM}}$"). Thus, when Italian and Spanish speakers look at a visual display including only one item with feminine gender, they make anticipatory looks to it upon hearing a feminine article, before they hear the noun. Dussias et al. (2013) found that both English- and Italian-speaking learners of L2 Spanish used gender information on articles predictively (cf. Grüter et al., 2012), although this ability was modulated by proficiency (e.g., Hopp, 2013) and, in the case of Italian-speaking learners, differences in how definite articles are realized in Italian and Spanish. In particular, the L1-Italian learners made anticipatory looks after hearing the feminine article *la*, which is similar in Italian and Spanish, but not after the masculine article *el*, for which Italian and Spanish show a number of differences. Also in the domain of gender agreement, Hopp and Lemmerth (2018) found that differences in terms of how the L1 and the L2 assign nouns to gender classes and in terms of how gender is marked across syntactic contexts impacts predictive processing among intermediate but not proficient learners.

Finally, van Bergen & Flecken (2017) found that both native speakers of Dutch and L1-German L2-Dutch learners used object placement information encoded in Dutch verbs (e.g., *zetten* "put-$_{\text{STAND}}$" vs. *leggen* "put-$_{\text{LIE}}$") to anticipate the position of an upcoming object. In contrast, L1-English and L1-French learners of Dutch did not. Since only German has similar object placement verbs, these results provide direct evidence that L1-L2 similarity facilitates anticipatory processing in the L2.

As discussed by Hopp and Lemmerth (2018), studies investigating the role of transfer in L2 anticipatory processing have typically focused on whether the learners' L1 does or does not realize the target L2 property (e.g., Dussias et al., 2013; Foucart et al., 2014; Grüter et al., 2012; Hopp, 2013; Lew-Williams and Fernald, 2010; van Bergen and Flecken, 2017). Their own study is the only one to have systematically manipulated L1-L2 (dis)similarity with respect to how a shared property is realized in the L1 and L2, informed by transfer theories such as the Feature Reassembly Hypothesis (e.g., Lardiere, 2009) or the Competition Model (e.g., Tokowicz and MacWhinney, 2005). Hopp and Lemmerth (2018) addressed this question with a within-subjects design, but they point out that studies investigating this issue by probing different L1-L2 combinations are lacking, and they explicitly advocate for them. Our study fills this gap.

2.1. Previous ERP studies on L2 anticipatory processing

The study we report here is part of a larger project investigating anticipatory processing at the level of the discourse in both native and nonnative speakers of English (see also Grüter et al., 2018; Rohde et al., 2011; Rohde and Horton, 2014). To our knowledge, only Grüter et al. (2017) have investigated L2 anticipatory mechanisms at the level of the discourse (see also Contemori and Dussias, 2019). Their study examined the predictive value of aspect in transfer-of-possession events. Using a story continuation task, the authors found that, after reading preambles such as *Emily brought/was bringing a drink to Melissa. (She) …*, L1-English speakers provided more continuations about the source (Emily) when the verb showed imperfective relative to perfective aspect. In contrast, intermediate to advanced L1-Japanese and L1-Korean learners of English showed no co-reference preferences as a function of aspect, regardless of proficiency, even though they showed target knowledge of the relevant aspectual distinction in an independent task. This is interesting, given that the L1 and L2 speakers' continuations were quite similar with respect to other subtle properties investigated in the study. Thus, Grüter et al.'s results (2017) suggest that certain cues that L1 speakers use predictively in discourse comprehension (as confirmed by Grüter et al., 2018) are not exploited to the same extent by adult L2ers.

In previous studies, we examined *information structure*, a discourse property that deals with how information is packaged in a sentence to build a felicitous discourse. We built on the distinction between two information structure categories, Topic and Focus. Topic corresponds to what a sentence is about (e.g., Lambrecht, 1994; Reinhart, 1981). In turn, Focus corresponds to new or discourse-relevant information and can be operationalized as the answer to a *wh*-question (e.g., Halliday, 1967; Lambrecht, 1994; López, 2009; Reinhart, 1981). Here, we provide a detailed description of this study because it involves the same participants as the current study (with a few exceptions) and because it is the only ERP study that has investigated L2 anticipatory processing at the level of the discourse.

In Alemán Bañón and Martin (2019), English native speakers read question-answer pairs like (1). In (1), the *wh*-question establishes that *banker* is the Topic and *adviser* and *agent*, the candidates for Focus assignment. In the response, we manipulated the presence of an *it*-cleft construction (1a-b vs. 1c-d), which acts as a cue for Focus assignment, and the information structure category of the target noun: Focus (1a, 1c) vs. Topic (1b, 1d).

(1) Either an adviser or an agent can be helpful to a banker. In your opinion, which of the two should a banker hire?
   a. *It*-cleft + Focus NP: In my opinion, it is an agent that a banker should hire.
   b. *It*-cleft + Topic NP: In my opinion, it is a banker that should hire an agent.
   c. No cleft + Focus NP: In my opinion, an agent should be hired.
   d. No cleft + Topic NP: In my opinion, a banker should hire an agent.

Importantly, since the Focus and Topic nouns were always preceded by different allomorphs of the English indefinite article (i.e., *an agent, an adviser* vs. *a banker*), we were able to measure effects of prediction on the article, before the target noun was integrated in the sentence (e.g., DeLong et al., 2005).

Our results revealed an N400 effect for articles that were unexpected after the *it*-cleft (i.e., 1b vs. 1a), which we interpreted as evidence that, upon encountering the *it*-cleft, English native speakers could better predict the location of the Focus NP in the response. In addition, nouns that were unexpected following the *it*-cleft (i.e., clefted Topics) showed a central-posterior P600 (600–900 ms) relative to felicitously clefted nouns (1b vs. 1a), providing additional evidence that the *it*-cleft constrained how information was expected to be organized in the response.

In Alemán Bañón & Martin (Unpublished results) we extended this study to 32 Swedish-speaking and 25 Spanish-speaking learners of English, and we found that, overall, their processing was qualitatively different.[2] Both learner groups showed a larger N400 for incorrectly clefted nouns relative to felicitously clefted ones (1b vs. 1a), suggesting that they processed the information structure violations in (1b) as lexical-semantic inadequacies. Most importantly, although the learners showed sensitivity to the prenominal article manipulation (i.e., *a/an*) in the conditions with the *it*-cleft, this sensitivity was nonnative-like and differed as a function of their L1. The Swedish-speaking learners showed a centrally-distributed Late Negativity (600–850 ms) for unexpected relative to expected articles (1b vs. 1a), which might be indicative of a delayed N400 effect (e.g., Dijkgraaf et al., 2019; Frenck-Mestre and Pynte, 1997; Martin et al., 2013; Mitsugi and MacWhinney, 2016; van Bergen and Flecken, 2017). In contrast, the Spanish-speaking learners showed a Late Anterior Positivity (500–800 ms), which we tentatively interpreted as evidence that they disconfirmed their predictions immediately after encountering an unexpected article. This is because, in the learners' L1 Spanish, articles provide reliable cues to nouns (e.g., nouns' grammatical gender), and most adjectives are postnominal (unlike English and Swedish, where adjectives are prenominal) (e.g., DeLong et al., 2011; Ito et al., 2017a; Ito et al., 2017b; Rabovsky, 2020). Importantly, the fact that neither Spanish nor Swedish realizes the specific *a/an* alternation could explain why both groups were quantitatively or qualitatively different from the L1 controls, as in Martin et al.'s study (2013). We follow up on this possibility in the present study.

3. The present study

Here, we use a design comparable to the one in Alemán Bañón and Martin (2019) including question-answer pairs and the *it*-cleft, in order to more directly investigate how L1-L2 similarity modulates L2 learners' predictions. Participants read short contexts like (2), which are about two characters (e.g., *Julia, Albert*) who have recently taken up some activity. For each character, one relative is mentioned (e.g., *Julia's niece, Albert's sister*) and the context motivates why only one of them is likely to join the event. A *wh*-question then asks participants which of those two relatives is most likely to be invited to the event. The response to the *wh*-question includes an *it*-cleft construction, which cues comprehenders that the upcoming phrase must be focusable. The candidates for Focus assignment are the possessed Noun Phrases "NPs" (i.e., *Julia's niece, Albert's sister*), since only they can value the *wh*-word and answer the question.

(2) Julia and Albert have joined a meditation group. Julia's niece dislikes being quiet and hates meditation. However, Albert's sister really enjoys silence and loves meditation. In your opinion, which of the two will they invite to meditation, Julia's niece or Albert's sister?
   2a: In my opinion, it is his sister that they will invite.
   2b: In my opinion, it is her niece that they will invite.

We recorded participants' EEG while they read one of two responses to the question (counterbalanced). In (2a), *his sister* is expected relative to *her niece* in (2b), since the context biases comprehenders towards expecting that Albert's sister, and not Julia's niece, will join them in meditation. As is the case in English, the form of the possessive pronoun in the response depended on the *possessor noun*'s natural gender (e.g., Cardinaletti, 1998), not the gender of the *possessed noun*. An example of this gender alternation is shown in (3) and (4).

(3) Julia$_{i\text{-FEM}}$ invited her$_{i\text{-FEM}}$ sister and her$_{i\text{-FEM}}$ brother.
(4) Albert$_{i\text{-MASC}}$ invited his$_{i\text{-MASC}}$ sister and his$_{i\text{-MASC}}$ brother.

By using possessor nouns of different gender within each context (e.g., *Julia, Albert*), we were able to measure effects of prediction on the possessive pronoun, before the target kinship noun became available in the input. In turn, the use of different nouns in the expected vs. unexpected conditions (e.g., *sister, niece*) allowed us to use participants' brain responses to the nouns as a check that they had generated predictions regarding which relative would participate in the event. We point out that, since *his* and *her* carry semantic features (i.e., natural gender), our study does not address the question of whether comprehenders activate the possessive pronoun's form or its semantic features.

Unlike the study we reported in Alemán Bañón and Martin (2019), we did not manipulate the information structure category of the target

[2] A full report of this study is currently in preparation.

NPs (*sister, niece*). Both of them could be clefted and resolve the *wh*-word. Thus, unexpected NPs violated an expectation based on contextual bias, not one based on information structure constraints. We also did not manipulate the availability of the *it*-cleft, which was always provided in the responses. Importantly, the results from our previous studies suggest that both L1 (Alemán Bañón and Martin, 2019) and L2 speakers (Alemán Bañón & Martin, Unpublished results) only activate properties of the Focus phrases upon encountering the *it*-cleft.

As we mentioned above, our learners had either Spanish or Swedish as their L1. Although both languages use cleft constructions to assign Focus, they differ with respect to the features that they encode in third-person possessive pronouns. Swedish has both anaphoric (i.e., reflexive) and nonanaphoric third-person possessive pronouns. Those that are nonanaphoric with the sentential subject mark the *possessor noun*'s human status and natural gender, similar to English (*hans* "his", *hennes* "her", *dess* "its") (e.g., Cardinaletti, 1998; Tingsell, 2007). This can be seen in (5). Crucially, these are the only pronouns that would be licensed in Swedish equivalent translations of the responses in (2a/2b), where there is no anaphoric relation between the subject (i.e., *they*) and the possessive pronoun (i.e., *his/her*).[3]

(5) De bjöd in …
they invited …
a. Hennes_i syster (Julias_i syster).
Her sister (julia's sister)
b. Hennes_i bror (Julias_i bror).
Her brother (julia's brother)
c. Hans_i syster (Alberts_i syster)
his sister (albert's sister)
d. Hans_i bror (Alberts_i bror)
his brother (albert's brother)

In contrast to English, Spanish possessive pronouns show syntactic agreement with the *possessed noun*. Although this dependency does not surface morphologically for gender in third-person singular possessive pronouns (e.g., compare 7 to 8) (e.g., Cardinaletti, 1998; Picallo, 1994), it does for number in all pronouns (e.g., compare 7 to 9). Furthermore, first- and second-person plural possessive pronouns also show syntactic gender agreement with the possessed noun. The example in (10) illustrates how Spanish possessive pronouns agree in syntactic gender with the *possessed noun*, not with the *possessor noun*. Thus, the Spanish possessive system as a whole provides evidence for syntactic agreement between possessive pronouns and possessed nouns, in contrast to English. As we mentioned earlier, an important feature of our design is that the expected "possessive + noun" combination always showed a mismatch in natural gender (e.g., *his sister*). This design allowed us to examine whether L1-Spanish learners of English would experience interference when processing the responses in (2), especially because syntactic and natural gender overlap for all kinship nouns in Spanish. If Spanish speakers incorrectly establish agreement between the possessive pronoun and the possessed noun, then the masculine possessive pronoun *his* in (2a) is unexpected relative to the feminine possessive pronoun *her* in (2b), since it is incompatible with the gender of the expected noun *sister*. In turn, the feminine possessive pronoun *her* in (2b) becomes expected, since it matches the gender of the noun *sister*.

(7) Julia quiere a su hermana/hermano.
julia loves $_{\text{CASE}}$ $_{\text{POSS-3RD-SG}}$ sister-$_{\text{FEM}}$/brother-$_{\text{MASC}}$
(8) Alberto quiere a su hermana/hermano.
alberto loves $_{\text{CASE}}$ $_{\text{POSS-3RD-SG}}$ sister-$_{\text{FEM}}$/brother-$_{\text{MASC}}$
(9) Julia quiere a sus hermanas/hermanos.
julia loves $_{\text{CASE}}$ $_{\text{POSS-3RD-PL}}$ sister-$_{\text{FEM-PL}}$/brother-$_{\text{MASC-PL}}$
(10) Nosotros vimos a nuestra hermana.
we-$_{\text{MASC}}$ saw $_{\text{CASE}}$ our-$_{\text{FEM-SG}}$ sister-$_{\text{FEM-SG}}$

A study by Antón-Méndez (2011) provides evidence that the properties of Spanish interfere with the production of English third-person singular possessive pronouns. Antón-Méndez examined this question by comparing advanced L2 learners of English whose L1 was Dutch, Italian, or Spanish. While Dutch is similar to English/Swedish in that third-person singular possessives mark the possessor noun's natural gender, Italian third-person singular possessive pronouns show overt syntactic gender agreement with the possessed noun. In an analysis of error rates, Antón-Méndez found that L1-Spanish learners were indistinguishable from L1-Italian learners, but both patterned differently (i.e., made more gender errors) from the L1-Dutch learners, which she takes as evidence that both the Italian and Spanish agreement rules cause interference in the production of English *his/her*, despite the lack of overt gender morphology in Spanish *su* (see also Lago et al., 2019; Pozzan and Antón-Méndez, 2017).

Thus, our experimental set-up allows us to examine whether L2 learners' predictions are impacted by crosslinguistic differences, by comparing learners whose L1 might play a facilitatory role (Swedish), with learners whose L1 might cause interference (Spanish) (e.g., Hopp and Lemmerth, 2018).

3.1. Research question

Our research question "RQ" is formulated below, followed by our predictions:

RQ. *To what extent do the properties of the learners' L1 impact the predictions they generate in the L2?*

We address this question by comparing native speakers of English to advanced L2 learners of English with either Spanish or Swedish as their L1. Although no previous study manipulating prenominal words has examined the *his/her* rule, we hypothesize that native speakers of English will show a larger N400 for unexpected relative to expected possessive pronouns (2b vs. 2a), since this is the most common response in studies manipulating prenominal articles marked for syntactic gender (e.g., DeLong et al., 2005; Foucart et al., 2014; Martin et al., 2013). If Spanish-speaking learners overextend the syntactic agreement rule from their L1 to English, or if the rule interferes, it is possible that the learners will elicit a "reverse N400 effect", with more negative waveforms for (2a) relative to (2b). In other words, their N400 would look like a positivity. It is also possible that the Spanish-speaking group will show no sensitivity to the possessive pronoun manipulation. This would be consistent both with proposals which argue that learners have reduced ability to predict in general, such as the RAGE hypothesis (Grüter et al., 2017; Martin et al., 2013), and with proposals which explicitly claim that predictive routines are the same in the L1 and L2, but can be obscured by L1-L2 differences, such as Kaan (2014).

In Swedish, equivalent third-person singular possessive pronouns mark the possessor noun's natural gender, just as in English (*hans* "his", *hennes* "her"). Thus, if L1-L2 similarity facilitates predictive processing, it is possible that the Swedish-speaking learners will show a qualitatively (i.e., N400) and quantitatively native-like effect of prediction for unexpected possessive pronouns. This would be more consistent with Kaan's proposal, which assumes that predictive mechanisms are fundamentally the same in the L1 and L2.

All three groups are predicted to elicit an N400 effect for unexpected

[3] Swedish also has reflexive possessive pronouns that agree in number and syntactic gender (common vs. neuter) with the *possessed noun*, as shown in (6). These reflexive pronouns are unlicensed in the responses' Swedish equivalent translations, since they are not bound in their binding domain (Chomsky, 1981, 1982, 1995). Importantly, it has been reported that Swedish speakers overuse nonanaphoric third-person possessive pronouns in contexts where the reflexive pronouns belong, but they do not do the reverse (see Tingsell, 2007, 2011). (6) a. Klara såg sin mamma, sitt barn och sina vänner. klarai saw $\text{her}_{\text{i-COM}}$ mum $_{\text{-COM}}$, $\text{her}_{\text{i-NEU}}$ child-$_{\text{NEU}}$ and $\text{her}_{\text{i-PL}}$ friend $_{\text{-PL}}$.

relative to expected nouns (e.g., DeLong et al., 2005; Foucart et al., 2014; Ito et al., 2017a; Martin et al., 2013). It is possible that unexpected nouns will also yield a Late Anterior Positivity, although this component is more likely to emerge in the L1-English group (e.g., Ito et al., 2017a; Martin et al., 2013).

4. Methods and materials

All experimental procedures were vetted and approved by the regional ethics committee in Stockholm (project number: 2018/611–31) and by the ethics committee at the Basque Center on Cognition, Brain and Language (no project number was assigned).

4.1. Participants

The participants for the study included 24 native speakers of English (15 females), 25 Spanish-speaking learners of English (15 females), and 32 Swedish-speaking learners of English (20 females). They all provided their informed written consent to participate in the study. Data from two additional participants (one L1-English speaker and one L1-Swedish learner) were excluded from analysis, in both cases due to a drifty reference in the EEG recording.

English proficiency was assessed via two standardized 60-item tests, the LexTALE (Lemhöfer and Broersma, 2012) and the Quick Oxford Placement Test "QOPT" (2001). The former assesses vocabulary skills and the latter assesses grammar, vocabulary, and reading skills. Table 1 summarizes the three groups' performance in both tests, alongside other relevant demographic characteristics. As Table 1 shows, both learner groups performed in the advanced range in both proficiency measures (over 80% accuracy), but the L1-Swedish group outperformed the L1-Spanish group in both tests (LexTALE: $t(55) = 2.013, p < .05$; QOPT: $t(30.54) = 3.855, p = .001$, equal variances not assumed).

All 82 participants had normal or corrected-to-normal vision and no record of learning disabilities or neurological impairments. All but two participants (two Spanish-speaking learners) were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The majority of the learners spoke other foreign languages than English, to varying degrees of proficiency. Most L1-English controls had knowledge of at least one foreign language, also to varying levels of proficiency. The L1-English and L1-Swedish participants were tested in Sweden and the L1-Spanish learners in Spain. All participants received compensation for their time.

4.2. Materials

The materials for the study comprise 64 short stories following the structure in (2). An additional example is provided in (11). To facilitate the description of the sample story, we present it segmented into sentences:

(11)

Sentence #1: Tom and Ruth want to throw a party.

Sentence #2: Tom's mother is very social and loves parties.

Sentence #3: However, Ruth's aunt is very shy and hates parties.

Wh-question: In your opinion, which of the two will they invite to the party, Tom's mother or Ruth's aunt?

Response 11a: In my opinion, it is his mother that they will invite.

Response 11b: In my opinion, it is her aunt that they will invite.

Structure of the contexts. Each context begins with a short sentence (sentence #1) introducing two characters who are planning an event (e.g., throwing a party). Importantly, the two characters are of the opposite sex (*Tom, Ruth*). Sentences #2 and #3 introduce two genitive constructions indicating a family relation between the possessor nouns and the possessed ones (*Tom's mother, Ruth's aunt*). Sentences #2 and #3 also clearly establish which of the two relatives should most reasonably participate in the event being organized. For example, in (11), it makes more sense that Tom's mother be invited to the party, since she loves parties, while Ruth's aunt hates them.

The wh-question. Next comes a *wh*-question asking participants which of the two relatives was likely to participate in the event. The two genitive constructions provided as possible answers to the question (*Tom's mother, Ruth's aunt*) are the candidates for Focus assignment (hereinafter, the Focus NPs), since only they can fill the slot opened by the *wh*-question and value the *wh*-word. In turn, the characters whom the story is about (*Tom, Ruth*) are the Topic NPs, and the *wh*-question requests additional information about them (e.g., Reiner, 1981).

The *wh*-question begins with the phrase *in your opinion*, to ensure that participants would engage with the contexts. In addition, we used the *wh*-expression *which of the two* to present the two possible alternatives as mutually exclusive. This contrast was reinforced by presenting the two Focus NPs in contrastive focus at the end of the question (*Tom's mother or Ruth's aunt?*) (e.g., Cowles et al., 2007). In half of the contexts, the *wh*-question was negated, to ensure that participants did not ignore the question and respond exclusively based on the bias created by sentences #1–3. An example of a set-up context with a negated question is provided in (12).

(12) Fred and Cathy want to go rowing. Fred's uncle is very sporty and loves rowing. However, Cathy's father is out of shape and hates rowing. In your opinion, which of the two will they not invite to go rowing, Fred's uncle or Cathy's father?

Response 12a: In my opinion, it is her father that they will not invite.

Response 12b: In my opinion, it is his uncle that they will not invite.

The response. The response to the *wh*-question involves an *it*-cleft

Table 1
Participants' age at testing, age of L2 acquisition, and English proficiency.

	L1-English (*N* = 24)			L1-Spanish (*N* = 25)			L1-Swedish (*N* = 32)		
	Mean	*SD*	*Range*	*Mean*	*SD*	*Range*	*Mean*	*SD*	*Range*
Age at testing	30	7	19–44	28	5	19–40	30	5	22–39
AoA of English		–		9	2	6–14	8	1	6–11
LexTALE score	95	6	78–100	82	9	66–100	87	10	65–100
QOPT score	98	1	95–100	83	13	50–100	94	5	77–100

AoA: Age of acquisition.
QOPT: Quick Oxford Placement Test.
LexTALE and QOPT scores are provided as percentages.

construction, which acts as a cue for Focus assignment (e.g., Alemán Bañón and Martin, 2019; Cowles et al., 2007). That is, the cleft indicates that the upcoming phrase is the Focus NP. We then systematically manipulated the expectedness of the clefted NP (i.e., the possessed noun). For example, the response in (11a) is expected relative to (11b), since the previous context biases comprehenders towards expecting that Tom's mother, and not Ruth's aunt, will be invited to the party.

Crucially, the clefted NPs were presented as genitive constructions with prenominal possessives (e.g., *his mother* vs. *her aunt* in 11). Since the two possessor nouns (*Tom, Ruth*) were always of the opposite sex, the prenominal possessives in the response showed different gender depending on whether the clefted noun was expected (*Tom's mother* = *his mother*) or unexpected (*Ruth's aunt* = *her aunt*). This allowed us to measure effects of prediction at the prenominal possessive, at a point when the target noun was yet to appear (e.g., DeLong et al., 2005; Foucart et al., 2014; Martin et al., 2013; Wicha et al., 2004). Notice that no possessive pronouns were used in the preceding set-up context, to avoid cueing participants.

Another crucial feature of the response is that the expected "possessive + noun" sequence showed a natural gender mismatch (context in 11: *his mother*). In contrast, the unexpected possessive pronoun (context in 11: *her*) was of the same gender as the expected noun (context in 11: *mother*). This allowed us to examine the extent to which Spanish-speaking participants experienced interference from the agreement rule that is instantiated in their L1.

Item controls. For the possessor nouns, we selected 32 unambiguously masculine (e.g., *Tom, Peter, Arthur*) and 32 unambiguously feminine (e. g., *Ruth, Ann, Silvia*) proper names. Each proper name was used twice, once with a male relative and once with a female one (e.g., *Tom's brother, Tom's mother*). No pair of names was repeated across contexts (e. g., *Tom and Ruth* vs. *Silvia and Tom*). The male character was the first to be mentioned in half of the contexts, and this was counterbalanced across the two uses of a given name.

The selection of the Focus NPs (i.e., the possessed nouns) was more constraining, since the number of English noun pairs showing a sex distinction that can be used as heads of genitive constructions (e.g., *his mother*) is limited. We therefore selected four feminine kinship terms (*aunt, mother, niece, sister*) and their four masculine counterparts (*uncle, father, nephew, brother*), using each kinship term 16 times. The Focus nouns were fully counterbalanced across the expected and unexpected conditions. Thus, we could compare brain responses to the nouns *mother, father, sister, brother,* etc. when they were expected, relative to when they were unexpected. The target Focus nouns corresponded to the first or the second noun in the context and in the *wh*-question an equal number of times, and this was also counterbalanced across the expected and unexpected conditions. The use of negated questions was fully counterbalanced across the expected and unexpected conditions. In addition, each kinship term was used an equal number of times with negated and non-negated questions.

These materials were assigned to two experimental lists following a Latin Square design. The set-up contexts were identical across the two lists, while the responses were fully counterbalanced. Across the two lists, every context was followed by both response types in (11) or (12), but no participant saw more than one response for a given context. Each of the two lists was used an equal number of times within the Swedish-speaking group (16 times each). Across the English-speaking participants, one list was used 11 times and the other one, 13 times (due to the loss of one recording). Across the Spanish-speaking participants, one list was used 12 times and the other one, 13 times. All materials are provided in Appendix 1.

Control Tasks. Our design encompasses two control tasks aimed at assessing the learners' knowledge of the *his/her* rule. Both tasks were administered after the EEG recording to avoid cueing the participants. The first task was a modified version of the Truth Value Judgment Task (TVJT). The materials for this task consisted of 16 statements including genitive constructions such as *Harold's niece* or *Gloria's nephew*, where the possessor and possessed nouns had different natural gender. Each statement was paired with a sentence including a genitive construction with a possessive pronoun whose correctness participants were to evaluate (against the statement). With this set-up, we manipulated the congruency between the two genitive constructions (*Harold's niece* vs. *his/her niece*). Examples are shown in (13–14) below:

(13) I think that Linda's father is a gardener.
 a. True: I think that her father is a gardener.
 b. False: I think that his father is a gardener.

(14) We went swimming with Harold's niece.
 a. True: We went swimming with his niece.
 b. False: We went swimming with her niece.

The second task was a Production Task. The materials for this task consisted of 16 statements similar to the ones in the TVJT. Each statement was followed by an identical counterpart where the possessor noun had been replaced by a series of underscores. Examples are provided in (15) and (16) below:

(15) Mike's mother has blue eyes.

____ mother has blue eyes

(16) A shark attacked Claire's nephew.

A shark attacked ____ nephew.

To create these tasks, we selected the same eight kinship terms that we used in the EEG experiment, in addition to 16 proper names (eight masculine), also from the EEG experiment. In both tasks, the target genitive constructions were located in initial, middle, or final position approximately an equal number of times. Both tasks included six filler items not targeting possessive pronouns. In the production task, all participants saw all items. For the TVJT, we created two experimental lists following a Latin Square design. The first statement was held constant across lists, and the second one was counterbalanced across the two lists. Across lists/participants, all target sentences from the TVJT were used in their congruent and incongruent versions, but no participant saw the same item twice. All participants saw all fillers.

4.3. Procedure

The testing for the present study involved one 3-h session. We used the software *PsychoPy* (Peirce, 2007, 2008; Peirce and MacAskill, 2018) for stimuli presentation. Participants sat comfortably in front of a computer monitor and received instructions that they would read a series of stories, each with a question. Their task was to read each story attentively, including the question. They learned that, upon a button press, they would read one potential response to the question. They were asked to avoid blinks and body movements while reading the responses, and to rest their eyes while reading the contexts. Participants also learned that, occasionally, a *yes/no* question would appear after the response. Thirty-two trials (16 per condition) were followed by a comprehension question (50% in total). The questions targeted either one of the Focus NPs 16 times (the target NP was targeted 8 times) and the event in the story 16 times.

In preparation for the experiment, participants completed a 13-trial practice set. Feedback was provided for the first three trials, to ensure that participants understood the task. The feedback also clarified how some responses were more reasonable than others, based on the bias provided by the context. The practice items (see an example in 17) had a similar structure to the experimental items, although they were not identical. For example, approximately half of the questions did not include a *wh*-word and approximately half of the responses did not involve the *it*-cleft. Recall that, in the experimental items, no possessive pronouns were used before the response (… *which of the two will they* *invite to the party, Tom's mother* or *Ruth's aunt?*). Thus, one of the purposes of the practice was to increase the predictability of possessive pronouns in the responses (see the response in 17). Since the purpose of our study was not to test how pervasive predictive processing is in L2 comprehension, but whether L1-L2 (dis)similarities modulate the learners' predictions, this allowed us to better address our research question. To avoid cueing participants, in all practice items where the response included a third-person singular possessive pronoun, the possessor and possessed nouns were of the same sex (e.g., *Sharon's mum*, *Zach's dad*). Thus, no matter which agreement rule participants applied (English/Swedish or Spanish), the output would be the same.

(17) Zach and Sharon need a babysitter. Zach's dad can't stand babies. However, Sharon's mum adores babies. In your opinion, will they ask Zach's dad or Sharon's mum for help?
a. In my opinion, they will ask her mum for help.

All practice trials involved kinship terms that did not appear in the experimental stimuli, such as *mum-dad, grandpa-grandma, daughter(s)-son(s)*. Five of the practice trials included a comprehension question. The experiment began immediately after the practice. The experiment encompassed four blocks of 16 items, separated by three short breaks. Within each block, items from both experimental conditions were intermixed and randomized. Words were displayed in black text (Courier New font) against a grey background. The last word of each response was marked with a period.

Trial structure. On each trial, participants first read the set-up context, which remained visible until they were ready to read the response. Upon a button press, the response began. First, participants saw a central fixation cross, displayed for 500 ms. Then, words were presented one at a time for 450 ms, followed by a 300 ms pause. We added an interval ranging from 500 to 1000 ms between the set-up context and the beginning of the response, pseudorandomly varied at 50 ms increments. After the response, participants either answered a *yes/no* comprehension question or moved on to the next context. The prompts for the comprehension question remained on the screen until the participants provided a response, which they did with their left hand: middle finger for *yes* and index finger for *no*.

Additional tasks. After the EEG recording, participants completed the two control tasks described in the *Materials* section, both of which were computerized. In the TVJT, participants saw one statement in black print and another one below in red print. They were asked to evaluate whether the latter was a correct way of expressing the information in the former, which they did by pressing one of two keys. In the Production Task, participants saw one statement in red print. Right below, they saw the same statement, but this time one word had been replaced with a series of underscores. Their task was to type in one word that would preserve the meaning of the original statement. A blinking cursor was placed right above the underscores.

4.4. EEG recording and analysis

L1-English and L1-Swedish data. The continuous EEG was recorded from 32 sintered Ag/AgCl active electrodes snapped into an elastic cap (Biosemi, Amsterdam, NL). The electrodes were positioned according to the International 10–20 System (midline: Fz, Cz, Pz, Oz; lateral: FP1/2, AF3/4, F3/4, F7/8, FC1/2, FC5/6, C3/4, T7/8, CP1/2, CP5/6, P3/4, P7/8, PO3/4, O1/2). To detect ocular artifacts, four flat electrodes were placed above and below the left eye, and on the left and right outer canthi. Two additional flat electrodes were placed on the mastoids. Electrodes CMS (Common Mode Sense, between C3 and Cz) and DRL (Driven Right Leg, between Cz and C4) were used as reference and ground, which is standard in Biosemi systems. Only active electrodes were used, which kept impedances very low. The recordings were amplified with an ActiveTwo amplifier (Biosemi, Amsterdam, NL) and digitized continuously with a sampling rate of 2048 Hz. Offline, the recordings were decimated (with an anti-aliasing filter) to 1024 Hz.

L1-Spanish data. We used Brain Vision Recorder (Brain Products, GmbH, Germany) to record the continuous EEG from 27 electrodes mounted on an elastic cap. The electrodes were placed according to the International 10–20 System (midline: Fz, Cz, Pz; lateral: FP1/2, F3/4, F7/8, FC1/2, FC5/6, C3/4, T7/8, CP1/2, CP5/6, P3/4, P7/8, O1/2). To monitor blinks and horizontal eye movements, we placed four electrodes above and below the right eye and on the left and right outer canthi. Two additional electrodes were placed on the mastoids. The recordings were referenced online to the left mastoid. Electrode impedances were kept below 5 KΩ for scalp electrodes and below 8 KΩ for eye electrodes. The recordings were amplified with an online bandpass filter of 0.016–250 Hz (with a 12dB/octave roll-off) by a BrainAmp DC amplifier, and digitized continuously at a sampling rate of 1 kHz.

EEG data analysis. The EEG data from both labs were analyzed with the Brain Vision Analyzer 2.1 software (Brain Products, GmbH, Germany). We first applied a 0.1 Hz high-pass filter to remove drift. We then re-referenced the recordings to the average of the mastoids. We then corrected ocular artifacts via Independent Component Analysis. This procedure decomposes each participant's EEG into independent components. We then removed components that are characteristic of blinks and horizontal eye movements upon visual inspection of their topographic distribution. On average, three components were removed per participant. The mean number of removed components per participant did not differ across groups, $F(2, 79) = 0.366$, $p = .694$. When necessary, bad electrodes were interpolated by using spherical spline interpolation. Interpolation of one critical electrode (i.e., an electrode that would be used for analysis) was needed for three L1-English participants and five L1-Swedish participants. Interpolation of two critical electrodes was needed for one L1-English participant. In all cases, interpolation was conducted for all trials.[4]

After ocular correction, we segmented the EEG into epochs in two intervals: between −300 and + 1000 ms relative to the onset of the possessive; between −100 and + 1000 ms relative to the onset of the noun. Remaining artifacts exceeding ±75 μV were automatically rejected. The remaining trials were included in the analysis, regardless of accuracy in the comprehension question. For the analyses on the possessive, the mean number of trials per condition was similar across conditions and groups (L1-English, expected possessive, $M = 31$, range = 18/32; unexpected, $M = 30$, range = 19/32; L1-Spanish, expected possessive, $M = 32$, range = 30/32; unexpected, $M = 31$, range = 27/32; L1-Swedish, expected possessive, $M = 31$, range = 22/32; unexpected, $M = 31$, range = 20/32). The same was true of the analyses on the noun (L1-English, expected noun, $M = 31$, range = 23/32 unexpected, $M = 30$, range = 20/32; L1-Spanish, expected noun, $M = 32$, range = 31/32; unexpected, $M = 32$, range = 28/32; L1-Swedish, expected noun, $M = 31$, range = 20/32; unexpected, $M = 30$, range = 20/32).

After artifact rejection, the epochs for the possessive were baseline-corrected relative to the 300 ms pre-stimulus interval. A shorter baseline was chosen for the noun (100 ms pre-stimulus interval) to ensure that any effects on the prenominal possessive would not contaminate the noun's baseline. Epochs were then averaged per condition and per subject. Finally, we filtered the averaged waveforms with a phase-shift free Infinite Impulse Response Butterworth filter, with a high cutoff of 30 Hz and a 12dB/octave roll-off.

Similar to Alemán Bañón and Martin (2019), we opted for a spatio-temporal approach to analyze the EEG data. Previous ERP studies manipulating prenominal material have examined article/adjective-noun syntactic gender agreement (e.g., Fleur et al., 2020; Foucart et al., 2014; Ito et al., 2020; van Berkum et al., 2005;

[4] Additional interpolation was conducted for a different set of seven participants (two L1-English speakers, five L1-Swedish learners). In these cases, one to two electrodes were interpolated, but those electrodes were not used for analysis.

Wicha et al., 2004), but no previous study on prediction has examined English possessive pronouns, which mark their antecedent's natural gender, not the syntactic gender of the head noun. In addition, there is substantial variability across previous studies with respect to the latency, topography, and polarity of prediction effects. Thus, we decided to quantify ERPs for the possessive via mean amplitudes between 250 and 400 ms, corresponding to the time window where unexpected indefinite articles yielded an N400 effect in Martin et al.'s study (2013). This is also the time window where the same L1-English group yielded more negative waveforms for unexpected indefinite articles relative to expected ones in our previous study manipulating information structure (Alemán Bañón and Martin, 2019). For the analysis on the noun, ERPs were quantified via mean amplitudes between 200-500 ms and 500–800 ms. These two intervals encompass the time window where incorrectly focused nouns yielded more negative waveforms in a previous study by Cowles et al. (2007), an effect that was most robust between 200 and 500 ms (i.e., the canonical N400 time window). In addition, the 500–800 ms approximates the time window where a Late Anterior Positivity emerged for unexpected relative to expected nouns in Martin et al.'s study (2013).

Since the data were collected with two different EEG systems, with different electrode arrays, we created regions of interest (ROI) including electrodes common to both systems. These ROIs are comparable to the ones in Alemán Bañón and Martin (2019): Frontal (F3, FC1, Fz, F4, FC2), Central (C3, CP1, Cz, C4, CP2), and Posterior (P3, O1, Pz, P4, O2). Based on Alemán Bañón and Martin (2019) and Martin et al. (2013), the analyses on the possessive were carried out in the Frontal and Central-Posterior regions, where previous studies have reported effects of prediction on prenominal material (e.g., Alemán Bañón and Martin, 2019; DeLong et al., 2005; Foucart et al., 2014; Martin et al., 2013). The analyses on the noun were carried out in the Central-Posterior region, where N400 effects have been reported for both unexpected or incorrectly focused nouns (e.g., Cowles et al., 2007; DeLong et al., 2005; Foucart et al., 2014). We also analyzed brain responses to the noun in the Frontal region, where previous studies have reported a Late Anterior Positivity for unexpected nouns (Delong et al., 2011; Martin et al., 2013).

For each analysis, mean amplitudes were entered into a mixed-factors ANOVA with Expectedness (Expected, Unexpected) as the repeated measure and Group (L1-English, L1-Spanish, L1-Swedish) as the between-subjects factor. Since the L1-Spanish data were collected with a different EEG system from the English/L1-Swedish data, we did not interpret main effects of Group. Notice, however, that the main effect of Group does not directly address our RQ. In total, we planned two analyses on the possessive (250–400 ms: Frontal, Central-Posterior). Then, upon visual inspection of the data, we ran an additional analysis in the 500–800 ms time window. On the noun, we planned a total of three analyses (200–500 ms: Central-Posterior; 500–800 ms: Frontal, Central-Posterior). We used a false discovery rate correction (Benjamini and Hochberg, 1995) in all follow-up tests. We report all effects regardless of significance.

5. Results

Mean accuracy in the comprehension questions approached ceiling performance in all three groups (L1-English, $M = 98\%$, $SD = 2\%$; L1-Spanish, $M = 96\%$, $SD = 4\%$; L1-Swedish, $M = 96\%$, $SD = 4\%$), with every participant scoring above 75% accuracy. This suggests that, across groups, participants were attentive to the stories and engaged in the task. A one-way ANOVA with Group (L1-English, L1-Spanish, L1-Swedish) as the independent variable revealed no differences in accuracy across groups, $F(2, 78) = 0.017$, $p = .838$; $\eta_p^2 = 0.005$.

To assess participants' knowledge of the *his/her* rule, we created a composite score by averaging across the scores for the two controls tasks. Mean accuracy across these tasks also approached ceiling performance within each group (L1-English, $M = 99\%$, $SD = 2\%$; L1-Spanish, $M = 96\%$, $SD = 6\%$; L1-Swedish, $M = 99\%$, $SD = 2\%$), with every participant scoring above 75% accuracy. This suggests that, when presented with genitive constructions such as *Harold's niece* or *Gloria's nephew*, both learner groups selected/produced a possessive pronoun agreeing with the possessor noun's natural gender (*Tom, Gloria*), not with the gender of the possessed noun (*niece, nephew*). Here, the effect of Group on accuracy was significant, $F(2, 78) = 6.545$, $p = .002$; $\eta_p^2 = 0.144$, driven by the fact that the L1-Spanish group was slightly less accurate than both the L1-English group, $F(1, 47) = 5.151$, $p = .028$; $q^* = 0.033$; $\eta_p^2 = 0.099$, and the L1-Swedish group, $F(1, 55) = 8.975$, $p = .004$; $q^* = 0.017$; $\eta_p^2 = 0.140$. Despite this difference, which is expected under accounts of transfer that capitalize on differences in feature realization (e.g., Lardiere, 2009; Tokowicz and MacWhinney, 2005), what is important is that the L1-Spanish group had robust knowledge of the *his/her* rule, as indicated by their mean accuracy score of 96%.

Figs. 1–3 show the waveforms for expected and unexpected possessive pronouns separately for the L1-English, L1-Swedish, and L1-Spanish groups. Visual inspection of the waveforms suggests that, in the L1-English and L1-Swedish groups, unexpected possessive pronouns yielded more negative waveforms than expected ones between ~250 and 400 ms in central-posterior electrodes, the effect emerging slightly later in the L1-English group (Figs. 1 and 2). In this time window, no negativity is visible for unexpected relative to expected possessive pronouns in the L1-Spanish group (Fig. 3). Unexpected possessive pronouns also yielded more positive waveforms than expected ones between 500 and 800 ms in central posterior electrodes, mainly in the L1-Spanish group (Fig. 3). A less robust positivity in this time window is also apparent in the L1-Swedish group (Fig. 2) and, to a lesser extent, in the L1-English group (Fig. 1). These differences can be seen in Fig. 4, which provides topographic maps of the relevant effects. Fig. 5 illustrates the effect size in the 250–400 ms time window for each group, together with a measure of uncertainty (i.e., within-subject standard error of the mean).

Figs. 6–8 provide the waveforms for expected and unexpected nouns for the L1-English, L1-Swedish, and L1-Spanish groups separately. Visual examination of the ERPs reveals more negative waveforms for unexpected relative to expected nouns between ~200 and 500 ms in all three groups. This effect shows a central-posterior distribution and is consistent with the N400. Although the negativity appears to approach baseline around ~500 ms, it remains sustained until ~800 ms in all three groups, mainly in the L1-English group (Fig. 6). These effects can be seen in Fig. 4, which also provides topographic maps of the relevant effects on the noun. No evidence of a Late Anterior Positivity is apparent for unexpected relative to expected nouns in any of the groups, although an earlier frontal positivity is apparent between 400 and 600 ms in the L1-Swedish and L1-Spanish groups (see Figs. 7 and 8).

5.1. Effects on the possessive

5.1.1. Time window between 250 and 400 ms

In the Central-Posterior region, the omnibus ANOVA revealed a main effect of Expectedness, $F(1, 78) = 5.488$, $p = .022$; $\eta_p^2 = 0.066$, driven by the fact that, across groups, unexpected possessives ($M = 2.079$; $SD = 2.33$) yielded more negative waveforms than expected ones ($M = 2.652$; $SD = 2.99$). Crucially, the main effect of Expectedness was modified by an interaction with Group, $F(2, 78) = 3.666$, $p = .030$; $\eta_p^2 = 0.086$. Follow-up tests to this interaction revealed that the main effect of Expectedness was marginal in the L1-English group, $F(1, 23) = 3.643$, $p = .069$, $q^* = 0.033$; $\eta_p^2 = 0.137$, and significant in the L1-Swedish group, $F(1, 31) = 7.410$, $p = .011$, $q^* = 0.017$; $\eta_p^2 = 0.193$. In the L1-Spanish group, the main effect of Expectedness was not significant, $F(1, 24) = 0.868$, $p = .361$, $q^* = 0.05$; $\eta_p^2 = 0.035$. As we pointed out above, visual inspection of the L1-English data suggested that the waveforms for expected and unexpected possessives started to diverge slightly later than our predetermined time window (i.e., 250–400 ms). As an exploratory analysis, we ran an additional follow-up in the 275–425 ms time window in the L1-English group, which showed that the main effect of

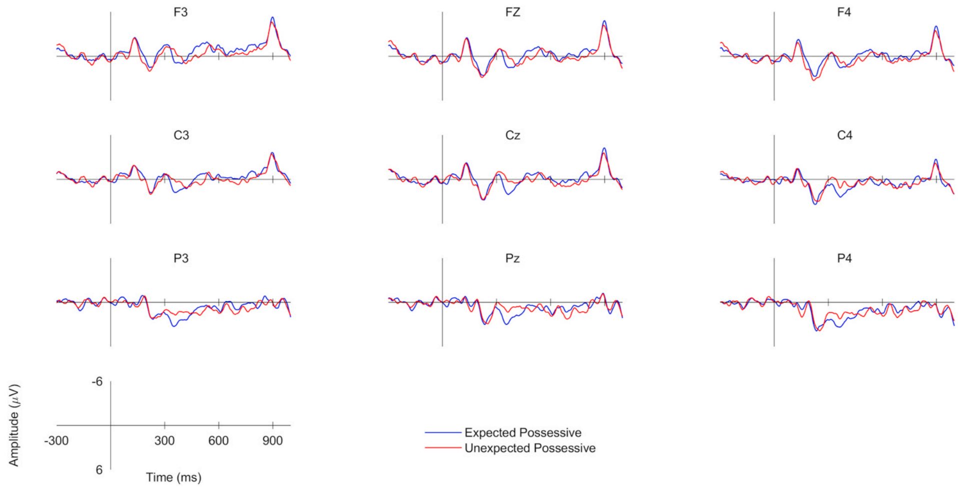

Fig. 1. Grand average ERP waveforms for expected and unexpected possessive pronouns in the L1-English group. ERPs are plotted for nine equidistant representative electrodes within each region of interest.

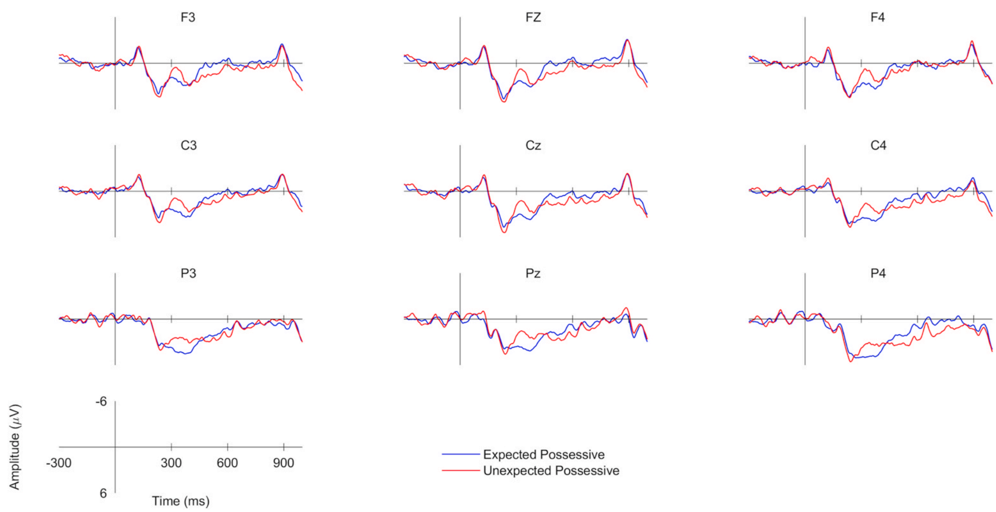

Fig. 2. Grand average ERP waveforms for expected and unexpected possessive pronouns in the L1-Swedish group. ERPs are plotted for nine equidistant representative electrodes within each region of interest.

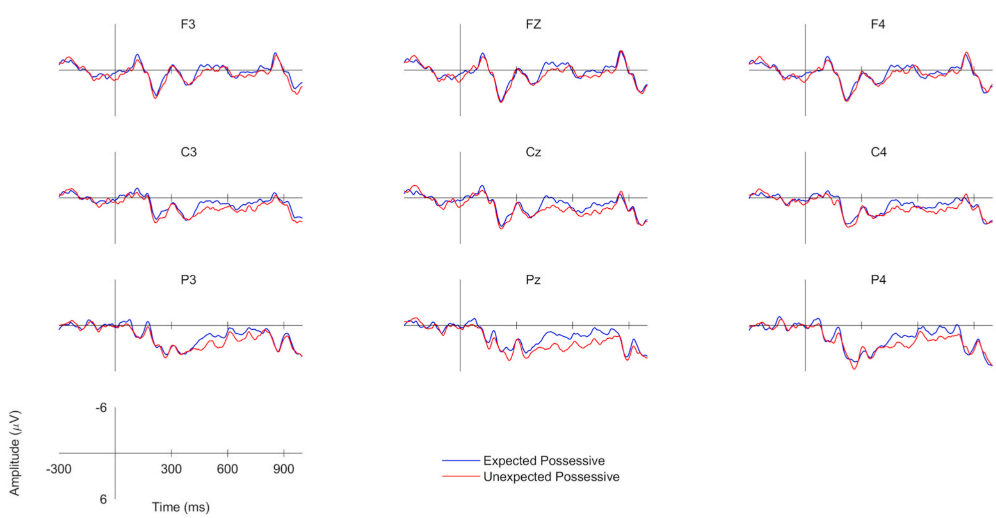

Fig. 3. Grand average ERP waveforms for expected and unexpected possessive pronouns in the L1-Spanish group. ERPs are plotted for nine equidistant representative electrodes within each region of interest.

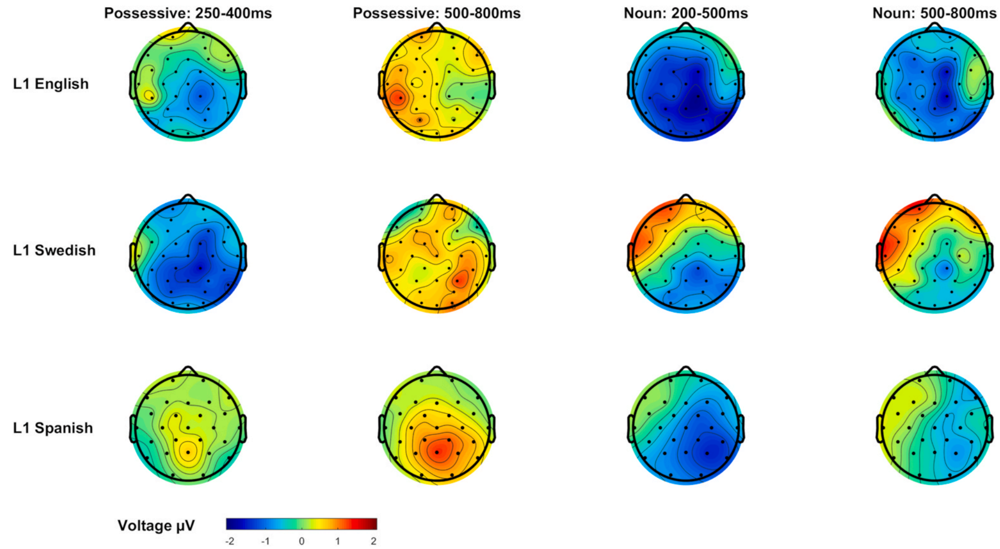

Fig. 4. Topographic plots for the effects at the possessive pronoun in the 250–400 ms and 500–800 ms time windows, and at the noun in the 200–500 ms and 500–800 ms time windows, for all three groups (upper row: L1-English; middle row: L1-Swedish; lower row: L1-Spanish). Plots were computed by subtracting the expected from the unexpected condition.

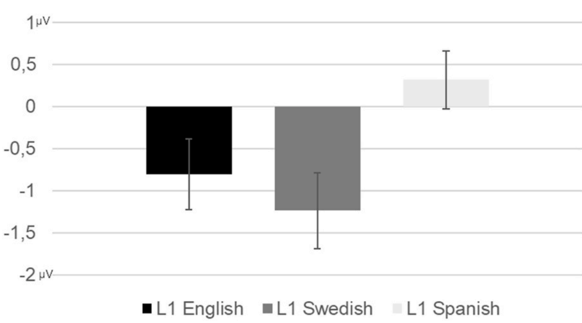

Fig. 5. Effect size for unexpected possessives in the 250–400 ms time window for all three groups (left bar: L1-English; middle bar: L1-Swedish; right bar: L1-Spanish). Effect size was computed by subtracting the expected from the unexpected condition in the Central-Posterior region. Error bars provide the within-subject standard error of the mean.

Expectedness was significant (at α = .05), $F(1, 23) = 4.751$, $p = .040$; $\eta_p^2 = 0.171$.[5]

Finally, the omnibus ANOVA also revealed a significant main effect of Group, $F(2, 78) = 3.814$, $p = .026$; $\eta_p^2 = 0.089$. As indicated above, this effect is difficult to interpret, given that we used two different EEG systems to collect the data. Moreover, it is not directly relevant to our discussion. Thus, no further comparisons were run.

In the *Frontal region*, the omnibus ANOVA only revealed a significant main effect of Group, $F(2, 78) = 4.442$, $p = .015$; $\eta_p^2 = 0.102$. Neither the main effect of Expectedness, $F(1, 78) = 1.888$, $p = .173$; $\eta_p^2 = 0.024$, nor the Group by Expectedness interaction reached significance, $F(2, 78) = 1.844$, $p = .165$; $\eta_p^2 = 0.045$).

5.1.2. Follow-up with a Bayesian approach

Since the absence of a central-posterior negativity for unexpected (relative to expected) possessive pronouns between 250 and 400 ms in the L1-Spanish group is theoretically relevant, we complemented this analysis with a Bayesian approach to determine the probability of the null hypothesis, relative to the alternative hypothesis. We used the statistical package JASP to conduct this analysis (JASP Team, 2017; Version 0.12.2). We selected a Cauchy prior centered on zero with a scale of 0.707 (i.e., the default) for the directional alternative hypothesis (Unexpected < Expected). A Bayesian paired samples *t*-test revealed moderate evidence in favor of the null hypothesis ($BF_{01} = 8.401$), relative to the alternative hypothesis. This suggests that the null hypothesis (i.e., that the waveforms for expected and unexpected possessive pronouns would not differ) was ~8.4 times more likely than the alternative hypothesis.

5.1.3. Time window between 500 and 800 ms (central-posterior region)

In this time window, the omnibus ANOVA revealed a main effect of Expectedness, $F(1, 78) = 8.375$, $p = .005$; $\eta_p^2 = 0.097$, driven by the fact that, across groups, unexpected possessives ($M = 1.181$; $SD = 2.20$) yielded more positive waveforms than expected ones ($M = 0.535$; $SD = 2.16$). No other effects reached significance (Group: $F(2, 78) = 0.771$, $p = .466$; $\eta_p^2 = 0.019$; Group by Expectedness interaction: $F(2, 78) = 0.524$, $p = .594$; $\eta_p^2 = 0.013$).

Despite the lack of an Expectedness by Group interaction, the positivity seems more robust in the L1-Spanish group (compare Fig. 3 to Figs. 1 and 2). Thus, we conducted pairwise comparisons, which showed that the positivity was only significant in the L1-Spanish group, $F(1, 24) = 6.279$, $p = .019$, $q^* = 0.017$ $\eta_p^2 = 0.207$ (L1-Swedish: $F(1, 31) = 2.763$, $p = .107$, $q^* = 0.033$ $\eta_p^2 = 0.082$; L1-English: $F(1, 23) = 0.939$, $p = .343$, $q^* = 0.05$ $\eta_p^2 = 0.039$). Although the effect in the L1-Spanish group became marginal after adjusting the original *p* value to .017, we calculated Cohen's D by dividing the mean difference (mean amplitude for unexpected possessive pronouns, $M = 1.665$, minus mean amplitude for expected ones, $M = 0.734$) by the standard deviation difference ($SD = 1.86$), which revealed that the effect was of medium size (0.5).

To sum up, these analyses revealed an N400-like effect between 250 and 400 ms for unexpected relative to expected possessive pronouns in

[5] Since our study is mainly concerned with between-group differences as a function of L1 similarity to English, we also compared the magnitude of the N400 effect across groups. N400 effect magnitude was calculated by subtracting the expected from the unexpected condition, separately for each group. These analyses showed that the N400 effect was significantly larger in the L1-English group ($M = -0.801$; $SD = 2.06$) compared to the L1-Spanish group ($M = 0.319$; $SD = 1.71$), but only before correcting for Type I error, $F(1, 47) = 4.310$, $p = .043$, $q^* = 0.033$; $\eta_p^2 = 0.084$. It was significantly larger in the L1-Swedish group ($M = -1.238$; $SD = 2.57$) relative to the L1-Spanish group, $F(1, 55) = 6.793$, $p = .012$, $q^* = 0.017$; $\eta_p^2 = 0.110$. In contrast, the magnitude of the N400 effect did not differ across the L1-English and L1-Swedish groups, $F(2, 54) = 0.468$, $p = .497$, $q^* = 0.05$; $\eta_p^2 = 0.009$.

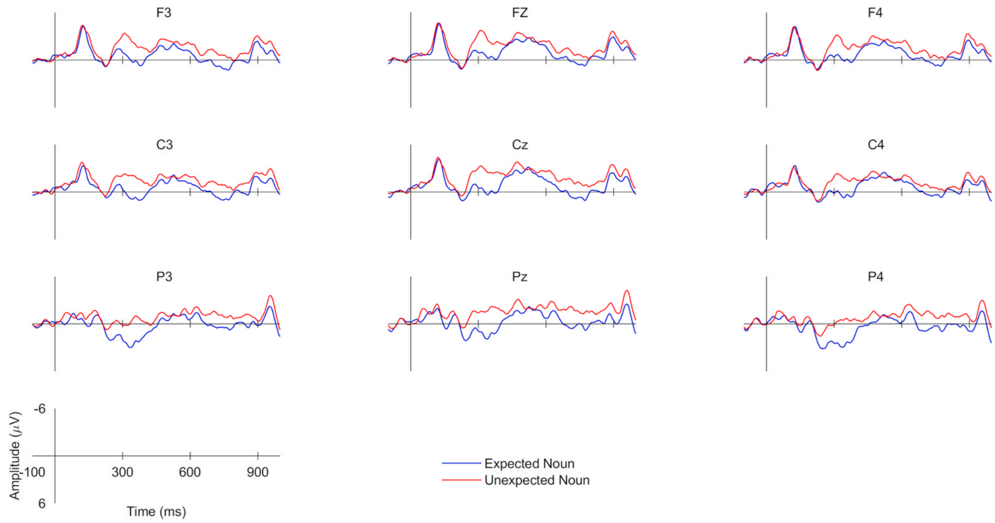

Fig. 6. Grand average ERP waveforms for expected and unexpected nouns in the L1-English group. ERPs are plotted for nine equidistant representative electrodes within each region of interest.

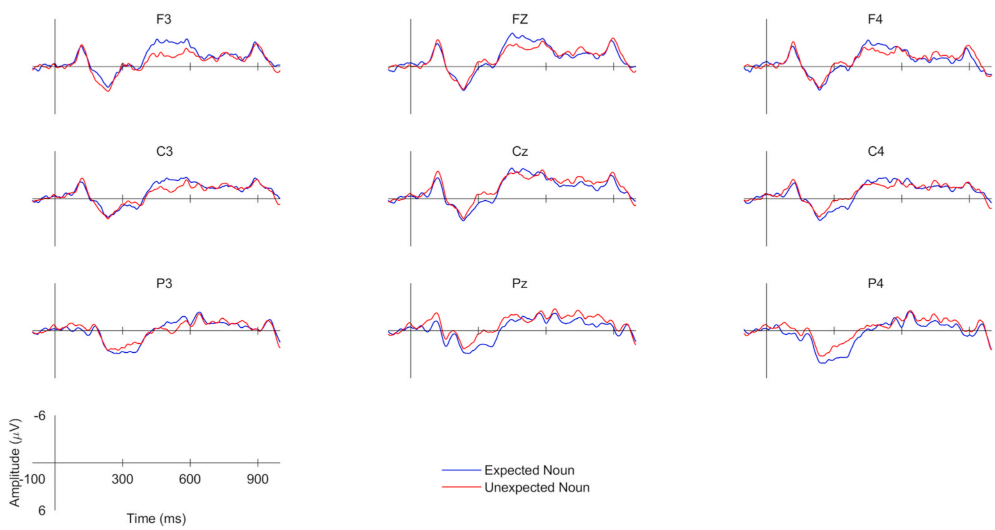

Fig. 7. Grand average ERP waveforms for expected and unexpected nouns in the L1-Swedish group. ERPs are plotted for nine equidistant representative electrodes within each region of interest.

the L1-English group (where it was marginal) and in the L1-Swedish group.[6] Analyses also revealed a positivity between 500 and 800 ms for unexpected relative to expected possessives across all three groups.

5.2. Effects on the noun

5.2.1. Time window between 200 and 500 ms (central-posterior)

The omnibus ANOVA revealed a main effect of Expectedness, $F(1, 78) = 18.379$, $p < .0001$; $\eta_p^2 = 0.191$, driven by the fact that, across groups, unexpected nouns ($M = 0.157$; $SD = 2.49$) yielded more negative waveforms than expected ones ($M = 1.181$; $SD = 2.51$). The main effect of Group was also significant, $F(2, 78) = 3.844$, $p = .026$; $\eta_p^2 = 0.090$, but the Group by Expectedness interaction was not, $F(2, 78) = 1.238$, $p = .296$; $\eta_p^2 = 0.045$.

5.2.2. Time window between 500 and 800 ms (central-posterior region)

The omnibus ANOVA revealed a main effect of Expectedness, $F(1, 78) = 5.856$, $p = .018$; $\eta_p^2 = 0.070$, driven by the fact that unexpected nouns ($M = -1.015$; $SD = 1.98$) yielded more negative waves than expected ones ($M = -0.432$; $SD = 2.10$) across groups. The main effect of Group was also significant, $F(2, 78) = 7.984$, $p = .001$; $\eta_p^2 = 0.170$. As in the previous time window, the Group by Expectedness interaction did not reach significance, $F(2, 78) = 1.024$, $p = .364$; $\eta_p^2 = 0.026$).

5.2.3. Time window between 500 and 800 ms (frontal region)

The omnibus ANOVA only revealed a main effect of Group, $F(2, 78) = 4.066$, $p = .021$; $\eta_p^2 = 0.094$. Neither the main effect of Expectedness, $F(1, 78) = 2.464$, $p = .121$; $\eta_p^2 = 0.031$, nor the Group by Expectedness interaction was significant, $F(2, 78) = 2.047$, $p = .136$; $\eta_p^2 = 0.050$).

To sum up, the analyses on the noun revealed an N400 effect between 200 and 500 ms for unexpected relative to expected nouns across

[6] A reviewer wondered whether the same pattern of results would emerge in a longer time window. Therefore, we analyzed the 250–500 ms time window, which revealed a marginal Group by Expectedness interaction, $F(2, 78) = 3.093$, $p = .051$; $\eta_p^2 = 0.073$. Follow-ups to this interaction revealed a pattern of results that was qualitatively similar to those from our predetermined time window, although the effects of Expectedness did not reach significance. This analysis in provided in full in Appendix 2.

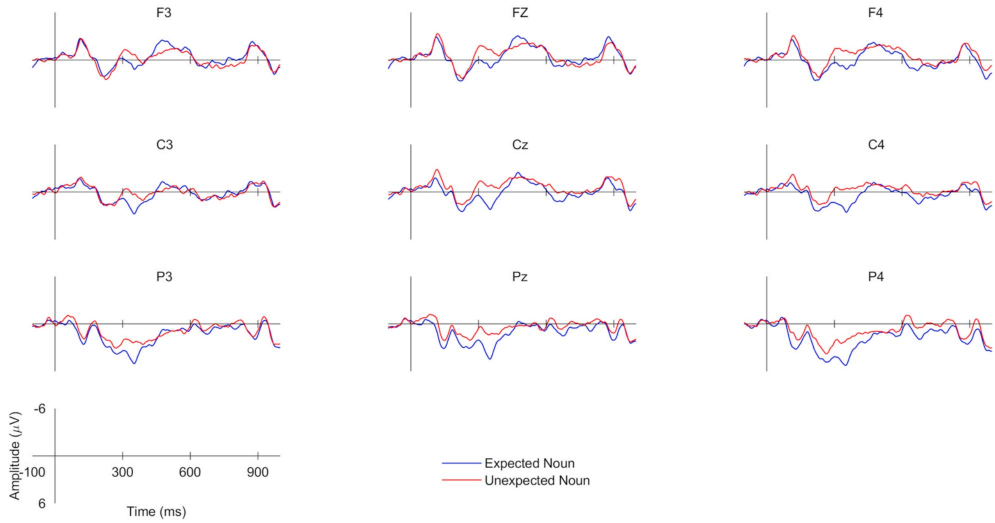

Fig. 8. Grand average ERP waveforms for expected and unexpected nouns in the L1-Spanish group. ERPs are plotted for nine equidistant representative electrodes within each region of interest.

all three groups. The negativity remained sustained until 800 ms, also across groups. Unexpected nouns did not yield a Late Anterior Positivity.[7]

5.3. Additional analyses on the possessive: multiple regression analysis

The analyses on the possessive pronouns revealed an N400 effect for unexpected relative to expected possessives in the L1-Swedish group, but not in the L1-Spanish group, even though some L1-Spanish learners showed negativities in this time window. Recall, however, that, although both groups were of advanced proficiency, the Swedish-speaking group scored higher than the Spanish-speaking group in the two standardized English proficiency tests that were administered in the study (i.e., the LEXtale and the QOPT). Thus, we used multiple regression to tease apart the individual contribution of both L1 background and English proficiency towards explaining variability in N400 effect magnitude for unexpected possessives (i.e., the prediction effect). The outcome variable was N400_Size, corresponding to the mean amplitude between 250 and 400 ms in the unexpected minus expected possessive condition, calculated for the Central-Posterior region. The predictor variables were Group, which was dummy-coded as zeros (L1-Spanish) and ones (L1-Swedish), and Proficiency, corresponding to the percentage score in the LEXtale. We selected the LEXtale as a measure of proficiency because, even if the L1-Swedish group outperformed the L1-Spanish group, both groups showed comparable variance in this measure (Levene statistic: $F(1, 55) = 0.208$, $p = .650$). This was not the case with the QOPT, where the L1-Swedish group showed very little variability (Levene statistic: $F(1, 55) = 18.500$, $p = .00007$), which is unsurprising given that our study does not manipulate proficiency.

The assumptions of multiple regression were met, as revealed by an analysis of residuals. First, no residual was smaller than −3.29 or larger than 3.29 (standardized residual minimum: −2.32; maximum: 2.38), suggesting that there were no outliers in the solution. In addition, the standardized residuals looked approximately normally distributed when plotted in a histogram (skewness = −0.29). The P–P plot of standardized residuals also showed a normal distribution (i.e., the points were close to the regression line). The data also met the assumptions of homogeneity of variance and linearity, as indicated by the scatterplot of standardized predicted values. In addition, there was no perfect multicollinearity between the predictors (both *VIF* values = 1.074). Finally, the data met the assumption of no zero variance (N400_Size, Variance = 94.641; Group, Variance = .251; Lextale_Score, Variance = 5.527).

The model including both Group and Proficiency as predictors accounted for a significant amount of the variance in N400_Size, $F(2, 54) = 3.382$, $p = .041$, $R^2 = 0.111$, $R^2_{Adjusted} = 0.078$. Crucially, the individual contribution of Group towards explaining variability in N400_Size was significant after controlling for Proficiency, $B = -1.509$, $t(54) = -2.418$, $p = .019$, $q^* = 0.025$. The unstandardized regression coefficient suggests that, on average, the size of the N400 effect for unexpected possessive pronouns was 1.509 μV larger in the L1-Swedish group, relative to the L1-Spanish group. In contrast, LEXtale_Score did not significantly predict N400_Size, $B = -0.009$, $t(54) = -0.290$, $p = .773$, $q^* = 0.05$. Thus, these results suggest that the qualitatively different ERP patterns shown by the L1-Swedish and L1-Spanish groups are more likely to stem from their having different L1s, as opposed to differences in proficiency.

6. Discussion

In the present study, native and nonnative speakers of English read short discourse contexts about a male and a female character planning some event. The contexts were followed by a *wh*-question asking which of two relatives were or were not likely to participate in the event. This set-up context made a genitive construction consisting of a third-person singular possessive pronoun and a kinship term of the opposite gender (e.g., *his mother*) likely in the upcoming response to the question. The response included an *it*-cleft, a syntactic device that cues comprehenders that the upcoming phrase is the response to the *wh*-question (e.g., Alemán Bañón and Martin, 2019; Cowles et al., 2007). After the cleft came either the expected genitive construction or an unexpected one with a possessive pronoun of the opposite gender (e.g., *his mother* vs. *her aunt*). This gender alternation in the possessive pronoun depended on whether the pronoun's antecedent was the male or the female character in the story. Importantly, the responses were always semantically plausible and syntactically correct (i.e., no gender agreement rule was violated). The participants' brain responses to the expected and unexpected possessive pronouns and kinship nouns were recorded with EEG.

The main purpose of the study was to investigate whether similarities

[7] Figs. 7 and 8 show that unexpected nouns yielded more positive waveforms than expected ones in anterior electrodes between ~400 and 600 ms, in the L1-Swedish group and, to a lesser extent, in the L1-Spanish group. We submitted mean amplitudes in this time window to an additional ANOVA, which showed that neither the main effect of Expectedness, $F(1, 78) = 0.060$, $p = .808$; $\eta_p^2 = 0.001$, nor the Group by Expectedness interaction reached significance, $F(1, 78) = 2.475$, $p = .091$; $\eta_p^2 = 0.060$.

and differences with respect to the features that English and the L1s of two learner groups (Spanish, Swedish) encode in third-person singular possessive pronouns would impact the learners' predictions (e.g., Dussias et al., 2013; Hopp and Lemmerth, 2018; Kaan, 2014; Lardiere, 2009; Tokowicz and MacWhinney, 2005; van Bergen and Flecken, 2017). Notably, English third-person singular possessive pronouns mark the *possessor noun*'s human status (*his/her* vs. *its*) and natural gender (*his* vs. *her*), which is the feature that we manipulated. In Spanish, in contrast, third-person singular possessive pronouns encode the syntactic features of the *possessed noun* (i.e., number and syntactic gender). Although this dependency is not morphologically obvious for syntactic gender, the Spanish possessive pronoun system as a whole provides evidence for it (see also experimental evidence from Antón-Méndez, 2011 and Lago et al., 2019). In addition, none of the Spanish possessive pronouns marks the possessor noun's natural gender. Finally, Swedish third-person singular possessive pronouns encode the same features as English, i.e., the possessor's human status (*hans/hennes* vs. *dess* "his/her vs. its") and natural gender (*hans* "his"; *hennes* "her") (e.g., Cardinaletti, 1998; Tingsell, 2007; see footnote #4).

We start by discussing the N400 effects on the kinship nouns, which we use as a confirmation check that participants generated predictions in the expected direction. Across all three groups, we found a central-posterior N400 effect for kinship nouns that were unexpected based on contextual bias, relative to expected nouns between 200 and 500 ms, a time window that was predetermined based on the study by Cowles et al. (2007). This effect indeed has similar latency and topography to the N400 effect reported by Cowles et al. (2007) for nouns that were unexpected following an *it*-cleft construction (although unexpected nouns in their study also violated information structure constraints). These findings are consistent with previous studies that have compared plausible nouns that differed with respect to predictability based on offline cloze probability (e.g., DeLong et al., 2005; Fleur et al., 2020; Foucart et al., 2014; Ito et al., 2017b; 2020; Kutas and Hillyard, 1984; Martin et al., 2013; Nieuwland et al., 2018). Our design, however, is quite different, since the contextual bias depended on which relative participants thought was or was not likely to participate in the event described in the story. These results suggest (1) that the stories successfully biased comprehenders towards one of the two possible genitive constructions; and (2) that this bias did not differ across the three groups. Thus, we are in a good position to compare how the three groups processed the prenominal possessive pronouns.

Crucially, all three groups showed some type of brain sensitivity to prenominal possessives that were unexpected based on their antecedent's natural gender, compared to expected possessives. This is not surprising, since (a) participants had enough time for prediction generation between the *wh*-question and the response; (b) the task encouraged participants to predict by asking their opinion in a question (e.g., Kuperberg and Jaeger, 2016); and (c) the contextual bias did not depend on a subtle linguistic property (e.g., Ito et al., 2018), as was the case in Grüter et al.'s study (2017), where the predictive cue was a morphologically encoded aspectual distinction. In addition, in all three languages, Focus can be assigned syntactically via cleft constructions (e.g., Lambrecht, 1994, 2001). Remember, however, that the purpose of our study was not to examine the ubiquity of predictive processing in L2 comprehension, but the impact of L1-L2 (dis)similarity on learners' predictions. Thus, although we erred on the side of encouraging predictive processing (e.g., Huettig and Mani, 2016), this allowed us to better address our research question (e.g., Ito et al., 2018). Crucially, participants' brain responses to unexpected possessive pronouns differed qualitatively as a function of their L1. As expected, the L1-English controls elicited a central-posterior N400 effect between ~250 and 400 ms (i.e., a time window that was predetermined based on the study by Martin et al., 2013; see also Alemán Bañón and Martin, 2019 and Ito et al., 2020), which might reflect activation of the pronoun's form or its semantic features (i.e., natural gender). The morphology of this effect is consistent with previously reported N400 effects for unexpected prenominal function words (e.g., Delong et al., 2005; Foucart et al., 2014). A qualitatively and quantitatively similar N400 effect emerged in the L1-Swedish group, but not in the L1-Spanish group. The latter showed a posteriorly-distributed positivity between 500 and 800 ms, which could be reminiscent of previously reported P600-like effects for function words showing unexpected gender (e.g., van Berkum et al., 2005; Wicha et al., 2004). Thus, the use of a design that facilitated predictive processing did not translate into a native-like pattern for the L1-Spanish learners, despite their advanced proficiency in English and despite their robust offline knowledge of the *his/her* rule, which suggests that they knew that English possessive pronouns mark the *possessor noun*'s gender, not the gender of the *possessed noun*. Only the L1-Swedish group, whose L1 encodes the same linguistic features as English in third-person singular possessive pronouns (i.e., the activated representation), showed a native-like pattern, as confirmed by a Group by Expectedness interaction in the N400 time window. This is in line with theories of transfer which argue that L1-L2 differences with respect to the realization of shared features impact L2 processing (e.g., Tokowicz and MacWhinney, 2005).

These results suggest that L2 learners' predictions can be qualitatively and quantitatively native-like, at least when the form of the activated representation depends on a rule that is similar in the learners' L1, as was the case for the L1-Swedish learners (e.g., Dussias et al., 2013; Hopp and Lemmerth, 2018). They also suggest that L1-L2 dissimilarities modulate learners' predictions. As for the L1-Spanish group, we contemplated the possibility that the lack of an N400 might be due to their lower proficiency in English, rather than crosslinguistic differences. Recall that, although both groups had a mean score of "advanced" in the two standardized tests that we used, the L1-Swedish group outperformed the L1-Spanish group. Our multiple regression analysis, however, revealed that only L1 background made a significant contribution towards explaining variability in N400 magnitude for unexpected possessives after controlling for proficiency. The contribution of proficiency towards accounting for N400 size after controlling for L1 background was not significant.[8] The lack of an N400 effect in the 250–400 ms time window among Spanish-speaking learners minimally suggests that they were slower in detecting the gender mismatches than the L1-English and L1-Swedish learners (e.g., Lago et al., 2019). Although the L1-Spanish learners demonstrated robust knowledge of the *his/her* rule in two offline control tasks, it is possible that accessing this knowledge online was costlier for them. Future studies should include an additional control task testing the learners' online sensitivity to the *his/her* alternation (cf. Martin et al., 2013).

Thus, the response to our research question *To what extent do the properties of the learners' L1 impact the predictions they generate in the L2?* is that similarities and differences between how the L1 and the L2 realize the activated representation modulate the nature of the learners' predictions. These findings seem in line with Kaan's proposal (2014), since they suggest that learners can generate qualitatively and quantitatively similar predictions to L1 speakers. We point out, however, that it is difficult to adjudicate between the RAGE hypothesis and Kaan's proposal (2014), since the RAGE hypothesis was purposefully formulated in broad terms at a time when L2 reports on predictive processing were scarce, and thus it overlaps with Kaan's proposal on several aspects. Both contemplate the possibility that factors such as proficiency and L1-L2 differences might diminish or obscure L2 learners' ability to generate expectations. Importantly, only Kaan (2014, p. 257, abstract) explicitly

[8] Since the L1-Spanish group was outperformed by the L1-English and L1-Swedish groups in the control tasks testing knowledge of the *his/her* rule, we also correlated N400 size with the Spanish-speaking learners' composite score for these tasks, but the correlation was not significant, $r(23) = -0.18$, $p = .389$. The correlation was not computed for the L1-Swedish learners, since they scored at ceiling in these tasks. This is also the reason why we did not include this variable in the multiple regression.

claims that predictive mechanisms are qualitatively (and potentially quantitatively) the same in the L1 and L2, which is the pattern that we found in the L1-Swedish group.

With respect to the L1-Spanish group, we initially hypothesized that, if learners predicted a possessive pronoun and incorrectly inflected it for the gender of the possessed kinship noun (as opposed to the possessor noun's) (e.g., *Tom's mother* = **her mother*), then the N400 effect for unexpected possessives would look like a positivity. In other words, possessive pronouns that were unexpected based on the English rule would be expected based on the Spanish rule (since syntactic and natural gender overlap for kinship nouns), and vice-versa. Our results are not consistent with this possibility. Although unexpected possessive pronouns did yield a positivity relative to expected ones, this effect did not emerge in the 250–400 ms time window. Furthermore, the effect was longer-lasting than the N400 effect that we found for unexpected possessive pronouns in the L1-English and L1-Swedish groups. Thus, we are more inclined towards interpreting this positivity as a qualitatively different effect, a P600 (e.g., van Berkum et al., 2005; Wicha et al., 2004). Recall, however, that we did not find a Group by Expectedness interaction in the P600 time window. Thus, the only reliable qualitative difference between the L1-Spanish group and the L1-English and L1-Swedish groups concerns the lack of an N400 effect for unexpected possessive pronouns in the former. Crucially, the fact that the L1-Spanish group showed a P600 effect for pronouns that were inflected for the unexpected gender shows that they were sensitive to the manipulation, just not in the same way as the L1-English and L1-Swedish groups.

We are unaware of previous ERP studies on prediction manipulating the *his/her* alternation, which complicates an attempt to integrate into the previous literature the qualitative variability that we found in the L1-Spanish group (P600), relative to the L1-English and L1-Swedish groups (N400 and perhaps a P600). Interestingly, we found similar qualitative variability in our previous study with the same participants. In that study, both the L1-English (Alemán Bañón and Martin, 2019) and L1-Swedish groups (Alemán Bañón & Martin, Unpublished results) showed a negativity for indefinite articles that were unexpected based on whether the expected Focus nouns began with a consonant or a vowel (i.e., *a/an*), although the effect was delayed in the L1-Swedish group. In contrast, the L1-Spanish group showed a Late Anterior Positivity (500–800 ms), which might indicate that they disconfirmed their predictions as soon as they encountered an incompatible article (Alemán Bañón & Martin, Unpublished results). It is also interesting that similar qualitative variability has been found across L1 studies on prediction manipulating prenominal material marked for syntactic gender (e.g., Fleur et al., 2020; Foucart et al., 2014; Ito et al., 2020; Kochari and Flecken, 2018; Martin et al., 2018; Otten and van Berkum, 2009; van Berkum et al., 2005; Wicha et al., 2004). For example, Wicha et al. (2004) found that native speakers of Spanish yielded a positivity between 500 and 700 ms for articles that mismatched the syntactic gender of predictable but unencountered nouns. In contrast, the studies by Foucart et al. (2014) and Martin et al. (2018), which conceptually replicated Wicha et al.'s design, found an N400 effect for unexpected gender-marked articles in another group of Spanish native speakers. Ito et al. (2020) and Fleur et al. (2020) discuss this variability in detail and suggest that, to some extent, it might be due to crosslinguistic or methodological differences (see also Ito et al., 2017c). Here, we found variability with similar methods and the same materials. One tentative interpretation is that the N400 effect for unexpected possessive pronouns in the L1-English and L1-Swedish groups reflects difficulty with the retrieval of a form or semantic feature (i.e., semantic gender) that was unexpected based on contextual constraint. In contrast, the P600 in the L1-Spanish group might reflect the difficulty with the integration of a possessive pronoun that mismatched the gender of the expected possessor. Although this would still suggest that the L1-Spanish learners predicted, the mechanism itself might have been different. Alternatively, it is possible that the L1-Spanish learners used similar mechanisms to generate their predictions, but were sensitive to different cues when evaluating whether or not their predictions had been borne out.

Unexpected nouns in the present study did not yield a Late Anterior Positivity. This is at odds with proposals which assume that this component reflects the cost of an unmet lexical prediction (e.g., DeLong et al., 2011; 2014; Van Petten and Luka, 2012). All three groups in our study showed brain sensitivity to unexpected prenominal possessives, suggesting that they had committed themselves to one of the kinship nouns. When this expectation was violated, however, no Late Anterior Positivity emerged at the noun. Brothers et al. (2020) and Kuperberg et al. (2020) suggest that the Late Anterior Positivity is more likely to emerge when the context is rich and globally constraining and allows comprehenders to build a situation model. Under this account, the Late Anterior Positivity reflects the reevaluation of the situation model, rather than prediction disconfirmation. For example, Brothers et al. (2020) found a Late Anterior Positivity for *foil* relative to *dough* in cases where the sentence *he flattened the foil/dough* followed a story about a character baking apple pie and mixing ingredients to make pie crust. In contrast, in cases where the carrier sentence followed an uninformative context, no Late Anterior Positivity emerged for unexpected nouns. Although our design is very different from these studies, it seems to us that our contexts were sufficiently rich and globally constraining to have allowed comprehenders to build a situation model. A crucial difference between these studies and our own, however, is that both the expected and unexpected nouns in our study had already been activated in the preceding context, and participants had to adjudicate between them. Thus, despite the contextual bias against them, unexpected nouns still represented an alternative in the upcoming continuation, which could have impacted the disconfirmation processes associated with the Late Anterior Positivity.[9]

7. Conclusion

In the present study, we used ERPs to investigate how native speakers of English and advanced L2 learners with either Spanish or Swedish as their L1 processed continuations to highly predictive discourse contexts. The contexts made a genitive construction consisting of a third-person singular possessive pronoun and a kinship noun with semantic gender (e.g., *his mother*) highly likely in the response. We then examined participants' brain responses to both expected and unexpected possessive pronouns and nouns. Crucially, the learners' L1s were either similar (Swedish) or different (Spanish) from English with respect to the linguistic features encoded in third-person singular possessive pronouns, i.e., the activated representation. Our results revealed a qualitatively and quantitatively similar N400 effect between 250 and 400 ms for unexpected prenominal possessives in the L1-English and L1-Swedish groups. Although the Spanish-speaking learners were sensitive to the pronoun manipulation, they elicited a qualitatively different component for unexpected possessives, a P600. These results provide evidence that dissimilarities with respect to how the L1 and the L2 realize a given rule can result in qualitatively different predictions

[9] A reviewer noted that a Late Anterior Positivity might have emerged on the possessive pronoun, as opposed to the noun, since it is the possessive pronoun that signals a prediction violation. In fact, Figs. 1 and 2 show more positive waveforms for unexpected compared to expected possessive pronouns following the N400 effect in the L1-English and L1-Swedish groups. Thus, we submitted mean amplitudes for the possessive pronouns in the 500–650 ms time window to an additional ANOVA. This is the time window where the positivity is most apparent in the L1-English group. This analysis was done in the Frontal region and only included the L1-English and L1-Swedish, which are the two groups that also showed an N400 effect for unexpected possessives. The ANOVA revealed no main effect of Expectedness, $F(1, 54) = 1.445$, $p = .235$; $\eta_p^2 = 0.026$, and no Group by Expectedness interaction, $F(1, 54) = 0.334$, $p = .566$; $\eta_p^2 = 0.006$.

among L2 learners, even at an advanced level of proficiency.

With respect to future directions, here we have examined a scenario in which the target linguistic feature encoded in the activated representation (i.e., the possessor noun's natural gender) is not encoded in the learners' L1 Spanish. In turn, the features encoded by Spanish possessive pronouns (i.e., the possessed noun's number and syntactic gender) are not encoded in their L2 English. A more challenging scenario would involve Spanish-speaking learners of Swedish, a language where third-person possessive pronouns either encode the possessor noun's natural gender (*hans/hennes* "his/her") or the possessed noun's grammatical gender (*sin* "$_{\text{POSS-COM}}$" vs. *sitt* "$_{\text{POSS-NEU}}$"), depending on whether or not they are bound by the sentential subject. Another challenge in this learning scenario is that natural and syntactic gender do not align in Swedish the way they do in Spanish. We hypothesize that such a learning scenario, where both the L1 rule and a novel rule apply in complementary distribution, would be more prone to L1 interference.

Funding

This work was supported by a Riksbankens Jubileumsfond grant to the first author [grant number: P18–0756:1]. The first author was also supported by a postdoctoral fellowship from the Spanish Ministry of Economy and Competitiveness [grant number FPDI-2013-15813]. The second author acknowledges financial support from the Spanish Ministry of Economy and Competitiveness (SEV-2015-490; PSI2017-82941-P; RED2018-102615-T), from the Basque Government (BERC, 2018–2021 program; PIBA18_29), and from the H2020 European Research Council (ERC Consolidator Grant ERC-2018-COG-819093).

Credit author statement

José Alemán Bañón: Conceptualization, Methodology, Investigation, Formal analysis, Data interpretation, Visualization, Writing – original draft, Writing – review & editing, Funding acquisition. Clara Martin: Conceptualization, Investigation, Data interpretation, Writing – review & editing.

Acknowledgements

The authors thank Elena Fano, Pia Järnefelt, and Rita Simonis for their help with data collection at Stockholm University; Manex Lete for his assistance with data collection at the BCBL; Magda Altman for her help with the manuscript; and all participants for their time.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2021.107797.

References

Alemán Bañón, J., Fiorentino, R., Gabriele, A., 2014. Morphosyntactic processing in advanced second language (L2) learners: an event-related potential investigation of the effects of L1–L2 similarity and structural distance. Sec. Lang. Res. 30 (3), 275–306. https://doi.org/10.1177/0267658313515671.

Alemán Bañón, J., Miller, D., Rothman, J., 2017. Morphological variability in second language learners: an examination of electrophysiological and production data. J. Exp. Psychol. Learn. Mem. Cognit. 43 (10), 1509–1536. https://doi.org/10.1037/xlm0000394.

Alemán Bañón, J., Fiorentino, R., Gabriele, A., 2018. Using event-related potentials to track morphosyntactic development in second language learners: the processing of number and gender agreement in Spanish. PloS One 13 (7), 1–35. https://doi.org/10.1371/journal.pone.0200791.

Alemán Bañón, J., Martin, C., 2019. Anticipating information structure: an event-related potentials study of focus assignment via the it-cleft. Neuropsychologia 134. https://doi.org/10.1016/j.neuropsychologia.2019.107203.

Alemán Bañón, J., & Martin, C. (Unpublished results). Can second language learners predict information structure? An event-related potentials study.

Altmann, G.T.M., Mirković, J., 2009. Incrementality and prediction in human sentence processing. Cognit. Sci. 33 (4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x.

Antón-Méndez, I., 2011. Whose? L2-English speakers' possessive pronoun gender errors. Biling. Lang. Cognit. 14 (3), 318–331.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B 57 (1), 289–300.

Bornkessel-Schlesewsky, I., Schlesewsky, M., 2019. Toward a neurobiologically-plausible model of language-related, negative Event-related potentials. Front. Psychol. 10 (298). https://doi: 10.3389/fpsyg.2019.00298.

Brothers, T., Swaab, T.Y., Traxler, M.J., 2015. Effects of prediction and contextual support on lexical processing: prediction takes precedence. Cognition 136, 135–149.

Brothers, T., Wlotko, E.W., Warnke, L., Kuperberg, G.R., 2020. Going the extra mile: effects of discourse context on two late positivities during language comprehension. Neurobiology of Language (Cambridge, Mass 1 (1), 135–160. https://doi.org/10.1162/nol_a_00006.

Cardinaletti, A., 1998. On the deficient/strong opposition in possessive systems. In: Wilder, C., Alexiadou, A. (Eds.), Possessors, Predicates, and Movement in the Determiner Phrase. John Benjamins, Amsterdam, Netherlands, pp. 17–53.

Chang, F., Dell, G.S., Bock, K., 2006. Becoming syntactic. Psychol. Rev. 113, 234–272. https://doi.org/10.1037/0033-295X.113.2.234.

Chomsky, N., 1981. Lectures on Government and Binding: the Pisa Lectures. Foris, Dordrecht, Netherlands.

Chomsky, N., 1982. Some Concepts and Consequences of the Theory of Government and Binding. MIT Press, Cambridge, Mass.

Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, Mass.

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 3, 181.

Contemori, C., Dussias, P.E., 2019. Prediction at the discourse level in Spanish–English bilinguals: an eye-tracking study. Front. Psychol. 10 https://doi.org/10.3389/fpsyg.2019.00956.

Covey, L., Fiorentino, R. Pham, C., Wilson, D. and Gabriele, A. (Unpublished Results) Examining Island Sensitivity in Native and Nonnative Speakers: Evidence from Acceptability Judgments and Event-Related Potentials.

Cowles, H.W., Kluender, R., Kutas, M., Polinsky, M., 2007. Violations of information structure: an electrophysiological study of answers to wh-questions. Brain Lang. 102 (3), 228–242.

DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8 (8), 1117.

DeLong, K.A., Urbach, T.P., Groppe, D.M., Kutas, M., 2011. Overlapping dual ERP responses to low cloze probability sentence continuations. Psychophysiology 48 (9), 1203–1207.

DeLong, K.A., Troyer, M., Kutas, M., 2014. Pre-processing in sentence comprehension: sensitivity to likely upcoming meaning and structure. Language and Linguistics Compass 12, 631.

DeLong, K.A., Urbach, T.P., Kutas, M., 2017. Is there a replication crisis? Perhaps. Is this an example? No: a commentary on Ito, Martin, and Nieuwland (2016). Language, Cognition and Neuroscience 32 (8), 966–973.

Dijkgraaf, A., Hartsuiker, R.J., Duyck, W., 2019. Prediction and integration of semantics during L2 and L1 listening. Language, Cognition and Neuroscience 7, 881.

Dussias, P.E., Kroff Valdés, J.R., Tamargo Guzzardo, R.E., Gerfen, C., 2013. When gender and looking go hand in hand: grammatical gender processing in L2 Spanish. Stud. Sec. Lang. Acquis. 35 (2), 353.

Federmeier, K.D., 2007. Thinking ahead: the role and roots of prediction in language comprehension. Psychophysiology 4, 491.

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

Fleur, D.S., Flecken, M., Rommers, J., Nieuwland, M.S., 2020. Definitely saw it coming? The dual nature of the pre-nominal prediction effect. Cognition 204. https://doi-org.ezp.sub.su.se/10.1016/j.cognition.2020.104335.

Foucart, A., Frenck-Mestre, C., 2011. Grammatical gender processing in L2: electrophysiological evidence of the effect of L1-L2 syntactic similarity. Biling. Lang. Cognit. 3, 379.

Foucart, A., Frenck-Mestre, C., 2012. Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. J. Mem. Lang. 66 (1), 226–248. https://doi-org.ezp.sub.su.se/10.1016/j.jml.2011.07.007.

Foucart, A., Martin, C.D., Moreno, E.M., Costa, A., 2014. Can bilinguals see it coming? Word anticipation in L2 sentence reading. J. Exp. Psychol. Learn. Mem. Cognit. 40 (5), 1461.

Foucart, A., Ruiz-Tada, E., Costa, A., 2016. Anticipation processes in L2 speech comprehension: evidence from ERPs and lexical recognition task. Biling. Lang. Cognit. 19 (1), 213–219. https://doi.org/10.1017/S1366728915000486.

Frenck-Mestre, C., Pynte, J., 1997. Syntactic ambiguity resolution while reading in second and native languages. Q. J. Exp. Psychol. Sect. A Hum. Exp. Psychol. 1, 119.

Grüter, T., Lew-Williams, C., Fernald, A., 2012. Grammatical gender in L2: a production or a real-time processing problem? Sec. Lang. Res. 28 (2), 191–215. https://doi.org/10.1177/0267658312437990.

Grüter, T., Rohde, H., Schafer, A.J., 2014. The role of discourse-level expectations in non-native speakers' referential choices. In: Proceedings of the Annual Boston University Conference on Language Development, p. 179.

Grüter, T., Schafer, A.J., Rohde, H., 2017. Coreference and discourse coherence in L2: the roles of grammatical aspect and referential form. Linguistic Approaches to Bilingualism 7 (2), 199–229. https://doi.org/10.1075/lab.15011.gru.

Grüter, T., Takeda, A., Schafer, A.J., Rohde, H., 2018. Intersentential coreference expectations reflect mental models of events. Cognition 177, 172–176. https://doi.org/10.1016/j.cognition.2018.04.015.
Halliday, M.A.K., 1967. Notes on transitivity and theme in English, part II. J. Linguist. 3, 199–244.
Hopp, H., 2013. Grammatical gender in adult L2 acquisition: relations between lexical and syntactic variability. Sec. Lang. Res. 29 (1), 33–56. https://doi.org/10.1177/0267658312461803.
Hopp, H., 2015. Semantics and morphosyntax in predictive L2 sentence processing. IRAL - International Review of Applied Linguistics in Language Teaching 53 (3), 277–306. https://doi.org/10.1515/iral-2015-0014.
Hopp, H., Lemmerth, N., 2018. Lexical and syntactic congruency in L2 predictive gender processing. Stud. Sec. Lang. Acquis. 40 (1), 171–199.
Huettig, F., 2015. Four central questions about prediction in language processing. Brain Res. 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014.
Huettig, F., Janse, E., 2016. Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. Language, Cognition and Neuroscience 31 (1), 80–93.
Huettig, F., Mani, N., 2016. Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience* 1, 19.
Ito, A., Martin, A.E., Nieuwland, M.S., 2017a. How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. Language, Cognition and Neuroscience 32 (8), 954–965.
Ito, A., Martin, A.E., Nieuwland, M.S., 2017b. Why the A/AN prediction effect may be hard to replicate: a rebuttal to Delong, Urbach, and Kutas (2017). Language, Cognition and Neuroscience 32 (8), 974–983.
Ito, A., Martin, A.E., Nieuwland, M.S., 2017c. On predicting form and meaning in a second language. J. Exp. Psychol. Learn. Mem. Cognit. 43 (4), 635–652.
Ito, A., Corley, M., Pickering, M.J., 2018. A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. Bilingualism 21 (2), 251–264. https://doi.org/10.1017/S1366728917000050.
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, vol. 136. Neuropsychologia. https://doi-org.ezp.sub.su.se/10.1016/j.neuropsychologia.2019.107291.
Jackson, C.N., Dussias, P.E., 2009. Cross-linguistic differences and their impact on L2 sentence processing. Bilingualism 12 (1), 65–82. https://doi.org/10.1017/S1366728908003908.
Jaeger, T.F., Snider, N.E., 2013. Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience. Cognition 127 (1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013.
Kaan, E., Dallas, A., Wijnen, F., 2010. Syntactic predictions in second-language sentence processing. In: Zwart, C.J., Vries, M.D., Koster, J. (Eds.), Structure Preserved: Studies in Syntax for Jan Koster. John Benjamins, Amsterdam, Netherlands, pp. 206–213.
Kaan, E., 2014. Predictive sentence processing in L2 and L1. Linguistic Approaches to Bilingualism 4 (2), 257–282.
Kaan, E., 2015. Knowing without predicting, predicting without learning. Linguistic Approaches to Bilingualism 4, 482.
Kaan, E., Kirkham, J., Wijnen, F., 2016. Prediction and integration in native and second-language processing of elliptical structures. Biling. Lang. Cognit. 19 (1), 1–18. https://doi.org/10.1017/S1366728914000844.
Kochari, A.R., Flecken, M., 2018. Lexical prediction in language comprehension: a replication study of grammatical gender effects in Dutch. Language, Cognition and Neuroscience 2, 239.
Kuperberg, G.R., Jaeger, T.F., 2016. What do we mean by prediction in language comprehension? Language, Cognition and Neuroscience 31 (1), 32–59.
Kuperberg, G.R., Brothers, T., Wlotko, E.W., 2020. A tale of two positivities and the N400: distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. J. Cognit. Neurosci. 32 (1), 12–35. https://doi-org.ezp.sub.su.se/10.1162/jocn_a_01465.
Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307 (5947), 161.
Kutas, M., DeLong, K.A., Smith, N.J., 2011. A look around at what lies ahead: prediction and predictability in language processing. Predictions in the brain: Using our past to generate a future 190–207.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). In: Annual Review of Psychology, vol. 62. Oxford University Press, New York, pp. 621–647.
Lago, S., Stutter Garcia, A., Felser, C., 2019. The role of native and non-native grammars in the comprehension of possessive pronouns. Sec. Lang. Res. 35 (3), 319–349. https://doi-org.ezp.sub.su.se/10.1177/0267658318770491.
Lambrecht, K., 1994. Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents. Cambridge University Press, Cambridge.
Lambrecht, K., 2001. A framework for the analysis of cleft constructions. Linguistics 39 (3), 463–516. ISSU 373).
Lardiere, D., 2009. Some thoughts on the contrastive analysis of features in second language acquisition. Sec. Lang. Res. 25 (2), 173–227. https://doi.org/10.1177/0267658308100283.
Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics:(de) constructing the N400. Nat. Rev. Neurosci. 9 (12), 920–933.
Lau, E.F., Holcomb, P., Kuperberg, G., 2013. Dissociating N400 effect of prediction from association in single-word contexts. J. Cognit. Neurosci. 25, 484–502.
Lemhöfer, K., Broersma, M., 2012. Introducing LexTALE: a quick and valid lexical test for advanced learners of English. Behav. Res. Methods 44 (2), 325. https://doi-org.ezp.sub.su.se/10.3758/s13428-011-0146-0.
Lew-Williams, C., Fernald, A., 2010. Real-time processing of gender-marked articles by native and non-native Spanish speakers. J. Mem. Lang. 63 (4), 447–464. https://doi-org.ezp.sub.su.se/10.1016/j.jml.2010.07.003.
López, L., 2009. A Derivational Syntax for Information Structure. Oxford University Press, Oxford, UK.
Mantegna, F., Hintz, F., Ostarek, M., Alday, P.M., Huettig, F., 2019. Distinguishing integration and prediction accounts of ERP N400 modulations in language processing through experimental design. Neuropsychologia 134, 2–9.
Martin, C.D., Thierry, G., Kuipers, J.R., Boutonnet, B., Foucart, A., Costa, A., 2013. Bilinguals reading in their second language do not predict upcoming words as native readers do. J. Mem. Lang. 69 (4), 574–588.
Martin, C.D., Branzi, F.M., Bar, M., 2018. Prediction is Production: the missing link between language production and comprehension. Sci. Rep. 8 (1), 1–9. https://doi-org.ezp.sub.su.se/10.1038/s41598-018-19499-4.
McDonald, J.L., 2006. Beyond the critical period: processing-based explanations for poor grammaticality judgment performance by late second language learners. J. Mem. Lang. 55 (3), 381–401. https://doi-org.ezp.sub.su.se/10.1016/j.jml.2006.06.006.
Mitsugi, S., MacWhinney, B., 2016. The use of case marking for predictive processing in second language Japanese. Biling. Lang. Cognit. 19 (1), 19–35. https://doi.org/10.1017/S1366728914000881.
Nieuwland, M.S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Mézière, D., 2018. Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. ELife 7, e33468.
Nieuwland, M.S., Kazanina, N., 2020. The neural basis of linguistic prediction: introduction to the special issue. Neuropsychologia 146, 107532. https://doi-org.ezp.sub.su.se/10.1016/j.neuropsychologia.2020.107532.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.
Otten, M., Van Berkum, J.J.A., 2009. Does working memory capacity affect the ability to predict upcoming words in discourse? Brain Res. 1291, 92–101. https://doi-org.ezp.sub.su.se/10.1016/j.brainres.2009.07.042.
Peirce, J.W., 2007. PsychoPy - psychophysics software in Python. J. Neurosci. Methods 162 (1–2), 8–13.
Peirce, J.W., 2008. Generating stimuli for neuroscience using PsychoPy. Front. Neuroinf. 2 (10), 1–8.
Peirce, J.W., MacAskill, M.R., 2018. Building Experiments in PsychoPy. Sage, London.
Phillips, C., Ehrenhofer, L., 2015. The role of language processing in language acquisition. Linguistic Approaches to Bilingualism 5 (4), 409–453. https://doi-org.ezp.sub.su.se/10.1075/lab.5.4.01phi.
Picallo, M.C., 1994. Catalan possessive pronouns: the avoid pronoun principle revisited. Nat. Lang. Ling. Theor. 12 (2), 259–299. https://doi-org.ezp.sub.su.se/10.1007/BF00993146.
Pickering, M.J., Gambi, C., 2018. Predicting while comprehending language: a theory and review. Psychol. Bull. 144 (10), 1002–1044. https://doi.org/10.1037/bul0000158.
Pickering, M.J., Garrod, S., 2013. An integrated theory of language production and comprehension. Behav. Brain Sci. 36 (4), 329–347. https://doi-org.ezp.sub.su.se/10.1017/S0140525X12001495.
Pozzan, L., Antón-Méndez, I., 2017. English possessive gender agreement in production and comprehension: similarities and differences between young monolingual English learners and adult Mandarin–English second language learners. Appl. Psycholinguist. 38 (4), 985–1017. https://doi-org.ezp.sub.su.se/10.1017/S0142716417000017.
Rabovsky, M., 2020. Change in a probabilistic representation of meaning can account for N400 effects on articles: a neural network model. Neuropsychologia 143. https://doi.org/10.1016/j.neuropsychologia.2020.107466.
Reinhart, T., 1981. Pragmatics and linguistics: an analysis of sentence topics. Philosophica 27 (1), 53–94.
Rohde, H., Levy, R., Kehler, A., 2011. Anticipating explanations in relative clause processing. Cognition 118 (3), 339–358.
Rohde, H., Horton, W.S., 2014. Anticipatory looks reveal expectations about discourse relations. Cognition 133 (3), 667–691.
Ryskin, R., Levy, R.P., Fedorenko, E., 2020. Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. Neuropsychologia 136. https://doi.org/10.1016/j.neuropsychologia.2019.107258.
Sabourin, L., Haverkort, M., 2003. Neural substrates of representation and processing of a second language. In: Hout, R.V., Hulk, A., Kuiken, F., Towell, R. (Eds.), Lexicon–syntax Interface in Second Language Acquisition. John Benjamins, Amsterdam, Netherlands, pp. 175–195.
Shimanskaya, E., Slabakova, R., 2017. Re-assembling objects: a new look at the L2 acquisition of pronominal clitics. Bilingualism 20 (3), 512–529. https://doi.org/10.1017/S1366728915000784.
Tingsell, S., 2007. Reflexivt Och Personligt Pronomen: Anaforisk Syftning Hos Ungdomar I Flerspråkiga Storstadsmiljöer. (Doctoral Dissertation). University of Gothenburg, Gothenburg, Sweden.
Tingsell, S., 2011. Tracking language change. Anaphoric binding in multilingual urban settings in Sweden. In: Källström, R., Lindberg, I. (Eds.), Young Urban Swedish. Variation and Change in Multilingual Settings. Department of Swedish Language, University of Gothenburg, Gothenburg, pp. 237–250.
Tokowicz, N., MacWhinney, B., 2005. Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. Stud. Sec. Lang. Acquis. 27 (2), 173–204.

Urbach, T.P., DeLong, K.A., Chan, W.-H., Kutas, M., 2020. An exploratory data analysis of word form prediction during word-by-word reading. Proc. Natl. Acad. Sci. U. S. A 117 (34), 20483–20494. https://doi.org/10.1073/pnas.1922028117.

Van Petten, C., Luka, B.J., 2012. Prediction during language comprehension: benefits, costs, and ERP components. Int. J. Psychophysiol. 83 (2), 176–190.

van Bergen, G., Flecken, M., 2017. Putting things in new places: linguistic experience modulates the predictive power of placement verb semantics. J. Mem. Lang. 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003.

Van Berkum, J.J., Brown, C.M., Zwitserlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. J. Exp. Psychol. Learn. Mem. Cognit. 31 (3), 443.

Wicha, N.Y., Moreno, E.M., 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. J. Cognit. Neurosci. 16 (7), 1272–1288.

Yan, S., Kuperberg, G.R., Jaeger, T.F., 2017. Prediction (or not) during language processing. A commentary on Nieuwland et al. (*2017*) and DeLong et al. (*2005)*. bioRxiv. https://doi.org/10.1101/143750.