ARTICLE IN PRESS

Journal of Memory and Language xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Bilinguals reading in their second language do not predict upcoming words as native readers do

Clara D. Martin [a,b,c,*], Guillaume Thierry [d,e], Jan-Rouke Kuipers [g], Bastien Boutonnet [d], Alice Foucart [a], Albert Costa [a,f]

[a] University Pompeu Fabra, carrer Roc Boronat, 138, 08018 Barcelona, Spain
[b] Basque Center on Cognition, Brain and Language, Paseo Mikeletegi, 69, 20009 Donostia-San Sebastian, Spain
[c] IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain
[d] School of Psychology, Bangor University, Adeilad Brigantia, Bangor, Gwynedd LL57 2AS, UK
[e] Economic and Social Research Council Centre for Research on Bilingualism in Theory and Practice, Bangor University, Bangor, UK
[f] ICREA, Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain
[g] Department of Psychology, Stirling University, Stirling, Scotland FK9 4LA, UK

A R T I C L E   I N F O

Article history:
Received 3 August 2012
revision received 30 July 2013
Available online xxxx

Keywords:
Anticipation
Bilingualism
Second language processing
ERP
N400 effect

A B S T R A C T

During reading, monolingual readers actively predict upcoming words from sentence context. Here we investigated whether bilingual readers predict sentence final words when they read in their second language. We recorded event-related potentials while English monolinguals (L1 comprehenders) and late Spanish–English bilinguals (L2 comprehenders) read sentences ending in an expected or unexpected noun. Lexical prediction was indexed by the amplitude of the N400 effect elicited by the article preceding the final noun, such that the more negative the N400, the less prediction as regards the final word. Contrary to L1 comprehenders, L2 comprehenders failed to show an N400 amplitude increase for unexpected articles. We interpret these results as evidence that L2 comprehenders do not actively predict upcoming words during sentence comprehension to the same extent as L1 comprehenders. This weaker capacity of lexical prediction in L2 might be one of the consequences of overall slower and less accurate linguistic processing stages in L2 relative to L1.

Introduction

Reading or listening to sentences in a second language (L2) usually proceeds slower and less accurately than in a first language (L1) (Cook, 1997; Green, 1998). Studies aiming at unravelling the cause of this difference have identified differences in L1 and L2 lexical (Soares & Grosjean, 1984), semantic (Hahne, 2001; Weber-Fox, Davis, & Cuadrado, 2003), and syntactic (Frenck-Mestre, 2002; Hahne & Friederici, 2001; Sanders & Neville, 2003; Weber-Fox & Neville, 1996) representations. Despite these observations, the consequences of these differences between L1 and L2 processing remain to be explored. Based on the assumption that linguistic processing stages are overall slower and less accurate in L2 (Frenck-Mestre & Pynte, 1997), we will investigate here some potential implications for sentence comprehension. To do so, we will focus on semantic processing during highly constrained sentence comprehension, and more specifically on lexical prediction. We will explore the capacities of L2 comprehenders to process linguistic representations quickly enough during sentence comprehension to be able to form a message-level representation that influences lexical predictions of upcoming words in high-constraint sentences.

* Corresponding author. Address: Basque Center on Cognition, Brain and Language, Parque Tecnológico de Miramón, Paseo Mikeletegi, 69, 2nd, 20009 Donostia-San Sebastian, Spain. Fax: +34 943 309 052.
E-mail address: claramartin3@gmail.com (C.D. Martin).

0749-596X/$ - see front matter
http://dx.doi.org/10.1016/j.jml.2013.08.001

Semantic processing during sentence reading

Semantic processing during sentence reading can be studied using event-related potentials (ERPs) by monitoring the amplitude of the N400, an ERP component more negative for semantically incorrect sentence endings as compared to semantically correct ones (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). Lower L2 proficiency in bilinguals has been shown to delay the N400 effect[1] (Elston-Güttler & Friederici, 2007; Hahne, 2001; Moreno & Kutas, 2005; Weber-Fox et al., 2003), suggesting that semantic processing is slower for reading in L2 than in L1 (for reviews, see Hernandez & Li, 2007; Kotz, 2009; Kotz & Elston-Guettler, 2007; Kutas, Moreno, & Wicha, 2009). However, recent accounts have suggested that this N400 component reflecting semantic processing is sensitive to lexical prediction (DeLong, Urbach, & Kutas, 2005; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004), orthographic/phonological analysis (Deacon, Dynowska, Ritter, & Grose-Fifer, 2004), semantic memory access (Kutas & Federmeier, 2000) and semantic/conceptual unification (Hagoort, Baggio, & Willems, 2009; for a review see Kutas & Federmeier, 2011). Until now, ERP experiments investigating semantic processing in L2 have not taken into account potential quantitative and/or qualitative differences between L1 and L2 regarding the way semantic processing – the N400 effect – is modulated by factors such as lexical prediction. Here, we set out to study lexical prediction during L2 sentence comprehension and how such expectation effects modulate semantic processing.

Influence of lexical prediction on semantic processing during sentence reading

The influence of sentence context on word processing has traditionally been studied using lexical decision tasks (LDT: word/non-word decision), classically showing faster responses when a word fits the sentence context as compared to when it does not (Fischler & Bloom, 1979; Kleiman, 1980; Stanovich & West, 1979). For instance, when reading "She has a nice voice and always wanted to be a...", a lexical decision on "singer" is made faster than on "lawyer". Using ERPs, it has been shown that the N400 component induced by a word is reduced when the word is embedded in a supportive context (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kutas & Federmeier, 2000, 2011; Kutas & Hillyard, 1980). Importantly, N400 mean amplitude is also reduced for the most highly expected noun in the sentence, compared to any unexpected noun (even if the unexpected noun is semantically congruent with the sentence context; e.g., "She has a nice voice and always wanted to be a *singer*" *versus* "She has a nice voice and always wanted to be an *artist*"; DeLong et al., 2005; Federmeier & Kutas, 1999; Federmeier, McLennan, De Ochoa, & Kutas, 2002).

Unfortunately, N400 amplitude modulations elicited by the critical noun of a sentence do not distinguish between active lexical prediction mechanisms and passive integration. In fact, semantic processing of the critical noun is assumed to be modulated by these two kinds of processes: (a) On the one hand, words are processed and integrated when they are read. Comprehenders incrementally build up message-level representations of meaning as the sentence unfolds, and words embedded in the sentence are integrated based on such representations. The more the meaning of a critical word fits message-level representations, the easier its semantic integration (as reflected by N400 amplitude reduction). Alternatively, according to the passive resonance hypothesis, words embedded in a sentence context activate a semantic relatedness network so that semantic processing of a critical word is facilitated when it is part of this semantic network (see Gerrig & McKoon, 1998; Myers & O'Brien, 1998; Paczynski & Kuperberg, 2012). According to these accounts, context effects appear when the critical word is presented and integrated based on the message-level representation and/or passive resonance with the pre-activated semantic relatedness network; see Kuperberg et al., 2011). (b) On the other hand, comprehenders can use sentence context information to generate predictions regarding upcoming words (active lexical prediction mechanisms; DeLong et al., 2005; Lau, Holcomb, & Kuperberg, 2013; Neely, 1977). According to this assumption, context effects can appear before the critical word is actually presented (see Lau et al., 2013, for extensive discussion on passive resonance *versus* active lexical prediction).

Predictions from the two accounts (passive lexical integration *versus* active lexical prediction) have been tested by studying ERPs elicited by the article preceding the final noun of a sentence. For instance, taking advantage of the grammatical properties of Spanish, which requires the article preceding the target noun to be marked and to agree with the gender of the following noun ("un" for masculine nouns *versus* "una" for feminine nouns), Wicha et al. (2004) observed ERP modulations between articles with expected and unexpected gender, based on prior sentence context. The authors concluded that readers generate predictions for specific nouns and their articles (Wicha, Bates, Moreno, & Kutas, 2003; Wicha, Moreno, & Kutas, 2003; Wicha et al., 2004). Using the phonological properties of English, where the indefinite article 'a' changes to 'an' if the following noun begins with a vowel, DeLong et al. (2005) also showed that expectation effects are already observable on the article, with the N400 more negative for the article 'an' when the best completion noun starts with a consonant, and inversely for the article 'a' when the best completion noun starts with a vowel. For instance, reading "She has a nice voice and always wanted to be..." leads to lexical prediction of the final noun "singer". Because of this lexical prediction, the N400 response to the article is greater in amplitude when the article encountered before the noun is "an" as compared to "a", since "an" is incompatible with "singer". Thus, readers actively predict words when processing a sentence with a constraining context. This active role of the comprehender during sentence processing probably speeds up language comprehension (Lau et al., 2013; van Berkum et al., 2005; Wicha et al., 2004).

In addition to the classical N400 effect elicited by unexpected lexical items, another important ERP

[1] The N400 effect is classically defined as the magnitude of the difference in amplitude between the N400 elicited by a semantic violation and that elicited by a semantically correct word in the same position.

modulation – the Anterior Positivity effect – has recently been described in studies on highly-constraint sentence processing. When comprehenders are presented with an unexpected word during sentence reading, the N400 component is followed by a positive ERP component – the Anterior Positivity – that reflects the cost of overriding or suppressing strong lexical prediction (DeLong, Groppe, Urbach, & Kutas, 2012; DeLong, Urbach, Groppe, & Kutas, 2011; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Otten & van Berkum, 2008; Van Petten & Luka, 2012). Unexpected words presented in a highly constrained sentence context elicit an enhancement of the Anterior Positivity, which reflects reprocessing cost following a failed lexical prediction. Thus, Anterior Positivity enhancement provides further evidence for active lexical prediction mechanisms and its analysis is critical to explore if readers manage or fail to actively predict upcoming lexical items. In fact, an absence of lexical prediction is accompanied by an absence of Anterior Positivity modulation (DeLong et al., 2012). Conversely, lexical prediction elicits an Anterior Positivity enhancement whenever the lexical item presented is a plausible word that violates lexical prediction (DeLong et al., 2011; Federmeier et al., 2007; Otten & Van Berkum, 2008; Van Petten & Luka, 2012).

To summarize, semantic processing of a critical word is influenced by the sentence context, (1) through passive resonance: A lexical representation is easier to integrate when it matches with the semantic network of the sentence context, and (2) through message-level build up: A lexical representation is easier to integrate when it matches message-level representations of sentence meaning. However, semantic processing of the critical word can also be facilitated by active lexical prediction based on contextual information. In that case, the pre-activated semantic relatedness network is updated and a specific item can be predicted in advance of the actual input. Thus, the N400 amplitude in response to a critical noun could be influenced by both passive integration and active lexical prediction (DeLong et al., 2005; Lau et al., 2013). On the other hand, any modulation of the N400 amplitude in response to the article preceding the critical noun can only be explained by an influence of active lexical prediction mechanisms. The Anterior Positivity effect on the critical noun reflects processing taking place when lexical prediction is highly pre-activated but not actually encountered. Thus, the modulation of the later component can only be explained by active lexical prediction mechanisms (DeLong et al., 2005, 2012). In the present study, we explored whether semantic processing in L2 is sensitive to active lexical prediction mechanisms, by assessing the N400 elicited by articles preceding expected or unexpected nouns. In order to gain further evidence about lexical prediction mechanisms in L2 sentence reading, we also investigated the Anterior Positivity elicited by the critical expected/unexpected nouns.

It is important to note that several previous studies have been conducted on the influence of sentence context on semantic processing in L2. For instance, in an eye-tracking study, Libben and Titone (2009) showed that reading in L2 is facilitated when words are embedded in highly constrained sentence contexts (see also Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Van Hell & De Groot, 2008). Schwartz and Kroll (2006) showed that L2 comprehenders (even those with low proficiency) can use highly constrained sentence context to restrict lexical competition (cross-language lexical competition being specifically explored in this study; see also Altarriba, Kroll, Sholl, & Rayner, 1996). Thus, it was repeatedly shown that semantic processing of a critical word is influenced by sentence context when reading in L2 (and when listening to speech in L2, see Bradlow & Alexander, 2007). Our study builds on these previous observations and aims at assessing whether active lexical prediction processes may be contributing to the observed effects. In fact, none of the previous studies specifically investigated whether the facilitation context effects operate through passive integration or active lexical prediction. Since those studies looked at reading times, fixation and gaze durations on the critical words of the sentences, they remain silent regarding the role of lexical prediction in L2 processing. As previously stated, the context facilitation effects observed during critical word processing do not allow differentiation between active lexical prediction and passive integration. The present study was specifically designed to explore the potential contribution of lexical prediction to semantic processing facilitation during L2 reading. We focused on lexical prediction effects defined as "electrophysiological evidence for the pre-activation of linguistic items or their features at a time preceding receipt of the actual input" (DeLong et al., 2012; p. 13).

Present study

We investigated whether reading in L2, like reading in L1, involves active lexical prediction of upcoming words on the basis of incrementally build up message-level representation. We used a similar ERP paradigm to DeLong et al. (2005) to investigate lexical prediction in late high proficient Spanish–English bilinguals (L2 comprehenders) and English monolinguals (L1 comprehenders) reading English sentences. Sentences ended with (a) an expected noun starting with a vowel (e.g., "Since it is raining, it is better to go out with an *umbrella*"); (b) an unexpected noun starting with a consonant (e.g., "Since it is raining, it is better to go out with a *raincoat*"); (c) an expected noun starting with a consonant (e.g., "She has a nice voice and always wanted to be a *singer*"); (d) an unexpected noun starting with a vowel (e.g., "She has a nice voice and always wanted to be an *artist*"). We were thus able to compare ERPs produced by the final noun of the sentence and the preceding article in the expected and unexpected conditions. We expected L1 comprehenders to show significant N400 ERP modulations by expectation for both the final noun and the preceding article, as previously observed by DeLong et al. (2005). In addition to the classical N400 effect elicited by unexpected lexical items, we also had clear hypotheses about the modulation of the Anterior Positivity elicited by the presentation of the final noun. Based on previous studies using similar paradigms, we anticipated an Anterior Positivity enhancement when the critical noun was unexpected. If L1 comprehenders actively predict upcoming words while reading, the presentation of an

Please cite this article in press as: Martin, C. D., et al. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.08.001

unexpected final noun should induce a re-processing cost reflected by an Anterior Positivity (DeLong et al., 2011, 2012; Federmeier et al., 2007; Otten & Van Berkum, 2008; Van Petten & Luka, 2012).

Regarding L2 comprehenders, we hypothesized that they would not predict upcoming words as efficiently as L1 readers. Two alternative possibilities would explain this lack of prediction efficiency: (1) First, L2 comprehenders may predict upcoming words more slowly than L1 comprehenders, because language processing is slower and more difficult in L2 relative to L1 (Frenck-Mestre & Pynte, 1997). Slower sentence comprehension might be detrimental for active lexical prediction mechanisms: online context-building, integration and comprehension may not occur dynamically enough for upcoming words to be effectively pre-activated prior to their occurrence (cf. DeLong et al., 2012 for a similar hypothesis regarding older adults). In other words, L2 comprehenders might be able to use sentence constraints to build up a message-level representation of meaning of the context, but more slowly than L1 comprehenders. In that case, the pre-activated semantic relatedness network would be updated and a specific lexical item would be predicted, but too late to influence sentence processing before the critical word is encountered. According to this alternative, L2 comprehenders would not be able to predict upcoming words quickly enough. Hence, N400 modulation by expectation would not be observed for the article preceding the critical noun (similarly to older adults, see DeLong et al., 2012). However, N400 and Anterior Positivity modulations by expectation would be observed on the noun, similarly to L1 readers. (2) The second alternative is that L2 readers rely exclusively on passive integration mechanisms during sentence processing. It may be that L2 readers incrementally build up message-level representations of meaning, but that such representations affect critical word integration only when the word is read. Alternatively, L2 readers may not incrementally build up message-level representations at all, and rely exclusively on passive resonance. In these two cases, critical words would be integrated based on contextual information stored in working memory, with no active lexical prediction affecting the integration of the critical word and the preceding article. According to this alternative, no expectation effects should be observed either on the N400 elicited by the article preceding the critical noun or on the Anterior Positivity elicited by the critical noun itself (suggesting insensitivity to failed lexical prediction; DeLong et al., 2012). But N400 modulation by expectation should still be observed on the noun. In fact, even without active lexical prediction (i.e., without expectation effect reflected in the N400 on the article and the Anterior Positivity on the noun), expected nouns should still be easier to process than unexpected nouns as a result of message-level representation influences and/or passive resonance processes.

Material and methods

Participants

Nineteen native speakers of English (12 females; 24 years ± 3) and 19 late Spanish–English bilinguals (8 females; mean age = 27 years ± 4) took part in the experiment. Bilingual participants learned English after the age of 8 (mean age of acquisition of English = 10 years ± 2) and did not speak any other language fluently. At the time of the experiment, they had been living in the UK for at least 2 years (mean immersion time = 4.7 years ± 2.3) and used both English and Spanish on an everyday basis, at home (with a dominance of Spanish) and/or at university (with a dominance of English). Participants were asked to self-rate their language proficiency on a 10 point scale in which "1" represents a low level of proficiency and "10" represents a native speaker level. The self-assessed index (averaged for speech comprehension, speech production, reading, and writing) was 9.4 (.7) in L1 (Spanish), 7.6 (1.0) in L2 (English) and 2.6 (1.3) in L3 (for 14 of the 19 participants who knew a third language which was German or French for most). None of the monolingual participants spoke any other language fluently. Their self-assessed index was 9.4 (.8) in L1 (English) and 2.4 (1.4) in L2 (for 11 of the 19 participants who knew a second language which was Spanish, French or German for most). All participants gave written consent to take part in the study that was approved by the ethics committee of Bangor University, Wales, UK.

Experimental design and materials

Stimuli consisted of 80 sentence contexts with two possible final noun-phrases (indefinite article + noun): expected or unexpected (see Table 1 for examples of sentences). In 40 sentence contexts, the expected final noun began with a vowel ("an + noun" expected noun-phrase) and the unexpected noun began with a consonant ("a + noun" unexpected noun-phrase). In the other 40 sentence contexts, the expected final noun began with a consonant and the unexpected noun began with a vowel. The 160 sentences were divided into two lists of 80 and each participant was presented with one list. Sentence contexts and final noun-phrases were used only once per list. Each list contained equal numbers of expected and unexpected nouns as well as 'a' and 'an' articles. Importantly, all sentences were semantically and grammatically correct. There were no semantic violations as final noun-phrases were always semantically correct, albeit that one was more expected than the other (e.g. "She has a nice voice and always wanted to be a singer" *versus* "She has a nice voice and always wanted to be an artist"). There were no agreement violations such as in "an singer" or "a artist". Mean written frequency of expected and unexpected words did not significantly differ so that expectancy effects observed on the noun could not be attributed to lexical frequency (mean frequency = 47 ± 60 and 59 ± 96 respectively; $p = .35$). One-quarter of the sentences were followed by yes/no comprehension questions in order to keep participants engaged in the silent reading task.

The sentence endings were submitted to a cloze probability test conducted with 21 English native speakers and 20 Spanish–English late bilinguals who did not take part in the ERP experiment. These participants were presented with 80 sentences truncated before the final noun-phrase and asked to complete the sentence with the first noun

Table 1
Examples of sentences used as experimental material.

Sentences	Conditions	
	Expected noun phrase	Unexpected noun phrase
He was very tired so he sat on	a chair	an armchair
She has a nice voice and always wanted to be	a singer	an artist
To have fresh milk you have to milk	a cow	an animal
After several weeks without working, Paul started looking for	a job	an occupation
He and his friends could play their own instruments so they decided to form	a band	an orchestra
He got it wrong but he does not want to recognize that he made	a mistake	an error
He has breathing difficulties so he is on the transplant list for	a lung	an organ
He really wants things to get better with you. You need to give him	a chance	an opportunity
As it is rainy it is better to go out with	an umbrella	a raincoat
I asked you a question because I need	an answer	a response
To throw away your cigarette butt you need	an ashtray	a bin
If you put a flame in front of a gas pipe you will cause	an explosion	a fire
He studied the whole night because in the morning he had	an exam	a test
The traffic jam on the motorway was due to	an accident	a breakdown
John's house was robbed twice. To increase the house's security, he installed	an alarm	a camera
David has a passion for constructions and edifices. He wants to become	an architect	a builder

that came to their mind. The cloze probability of a noun was defined as the percentage of times it was used. Mean cloze probability of the expected nouns was .69 ± .21 in monolinguals and .65 ± .26 in bilinguals. Mean cloze probability of unexpected nouns was .08 ± .10 in monolinguals and .09 ± .14 in bilinguals. Expected nouns had larger cloze probabilities than unexpected words in both groups ($ps < .0001$ in monolinguals and bilinguals). Importantly, cloze probability ratings did not significantly differ between monolingual and bilingual participants (expected nouns: $p = .26$; unexpected nouns: $p = .86$).

The cloze probability test suggested that bilinguals would predict the same words as monolinguals. We also controlled for phonological agreement between articles and nouns to ensure that L2 comprehenders were aware of the phonological rule in English. For instance, it could be the case that L2 comprehenders predict the upcoming noun ''singer'' at the end of the sentence ''She has a nice voice and always wanted to be...'' but that they would be less clear regarding the requirement of ''a'' or ''an'' before ''singer''. Such a situation would lead to null expectations on the article, not because of lack of lexical prediction but because of greater ''tolerance'' to article-noun phonological disagreement. This assumption is unlikely, however, given that the bilinguals recruited in the study were highly proficient in English and had been living in the UK for at least 2 years. Moreover, phonological disagreements such as ''an singer'' or ''a artist'' are not typical errors made by Spanish natives speaking English (Swan & Smith, 2001). Nevertheless, we tested this possible caveat by performing a second off-line cloze probability test with a group of 13 Spanish–English late bilinguals who did not take part in the ERP experiment or in the previous cloze probability rating. This group was presented with 80 sentences truncated before the final noun and asked to complete the sentences with the first noun that came to their mind. In this version of the test, we systematically used the least expected article, e.g., if the expected word was ''singer'' the article was ''an''. Thus, the most expected answer would create a phonological agreement violation. Under these conditions, the mean cloze probability of the most expected noun (e.g., ''singer'') became 3.5% and the mean cloze probability of the less expected noun used in the experiment (e.g., ''artist'') was 37.4%. Participants only made 4.1% phonological agreement errors. These results confirm that late Spanish–English bilinguals are highly sensitive to phonological agreement between articles and nouns when reading in English.

The off-line test revealed that L2 readers are sensitive to phonological agreement but it did not necessarily ensure that they have this information readily available during on-line sentence comprehension. Thus, in order to test on-line sensitivity to phonological agreement between articles and nouns in English, we conducted a control ERP experiment. Previous literature has shown that agreement violation between a noun and its preceding article induces a P600 effect: P600 amplitude elicited by a noun is larger when it is preceded by an article incongruent with the noun's features (e.g., gender), relative to a congruent article. This effect has been found in L1 (Barber & Carreiras, 2005; Hagoort & Brown, 1999) as well as in L2 (Foucart & Frenck-Mestre, 2011; Gillon-Dowens, Vergara, Barber, & Carreiras, 2010). In this control experiment, we presented L2 readers with sentences containing phonological agreement violations between an article and a noun in order to test sensitivity to these violations during on-line sentence comprehension. Given the results of the off-line test, we expected L2 readers to be sensitive to violations when processing phonological agreement on-line. Thus, we expected to observe a P600 effect in L2 readers, meaning larger P600 amplitudes for nouns preceded by incongruent articles relative to nouns preceded by congruent articles (Foucart & Frenck-Mestre, 2011; Gillon-Dowens et al., 2010).

Fourteen intermediate Spanish speakers of English[2] (who did not take part in the main study and the off-line tests; 11 females; 22 years ± 2) read 80 short sentences in

[2] Participants learned English after the age of 6 (mean age of acquisition of English = 9 years ± 3). At the time of the experiment, they were living in Spain and were not using English on an everyday basis. Participants' self-assessed index of English proficiency (averaged for speech comprehension, speech production, reading, and writing) was 6.8 ± 1.2.

English. Forty sentences ended with a noun preceded by the article "a" and 40 ended with a noun preceded by the article "an" (the 80 critical nouns were the ones used in the main experiment and the off-line tests). Half of the sentences were syntactically correct and half of them contained violations on the article (e.g., 'I would like to go to a/an *concert*'; 'He got a job in an/a *agency*'). Given the low constraint sentence context, the final critical noun was not highly expected. Each sentence was presented one word at a time (200 ms duration and 500 ms stimulus onset asynchrony) in the centre of a computer screen and the instruction was to read each sentence silently. Electrophysiological recording and data analyses were similar to the procedure used for the main experiment (see below "*Electrophysiological recording and data analyses*" for details). The P600 effect elicited by the critical noun was analyzed for the factors Congruency (congruent *versus* incongruent phonological agreement between the article and the noun) and Region (Frontal, Central, Parietal), in the [600–900] ms time-window.

The ANOVA performed on P600 mean amplitude of the critical noun (following a congruent/incongruent article) revealed a significant Congruency effect ($F[1,13] = 6.80$, $p < .03$), no effect of Region ($F[2,26] = .97$, $p = .39$) and no Congruency × Region interaction ($F[2,26] = .23$, $p = .79$). The P600 component was significantly larger when the phonological agreement was incongruent than when it was congruent (see Fig. 1).

Analyses of the final critical noun revealed a significant P600 effect suggesting that Spanish speakers of English show on-line sensitivity to phonological agreement violations (Barber & Carreiras, 2005; Foucart & Frenck-Mestre, 2011; Gillon-Dowens et al., 2010; Hagoort & Brown, 1999). Note that the sensitivity to phonological agreement observed in this control experiment is valid for participants with intermediate L2 proficiency, which is the profile of the participants of the main study (see above "*Participants*" section). Altogether, the results of the off-line and the on-line pre-tests show that it is very unlikely that Spanish–English bilinguals would predict the final noun of sentences without taking into account article-noun phonological agreement.

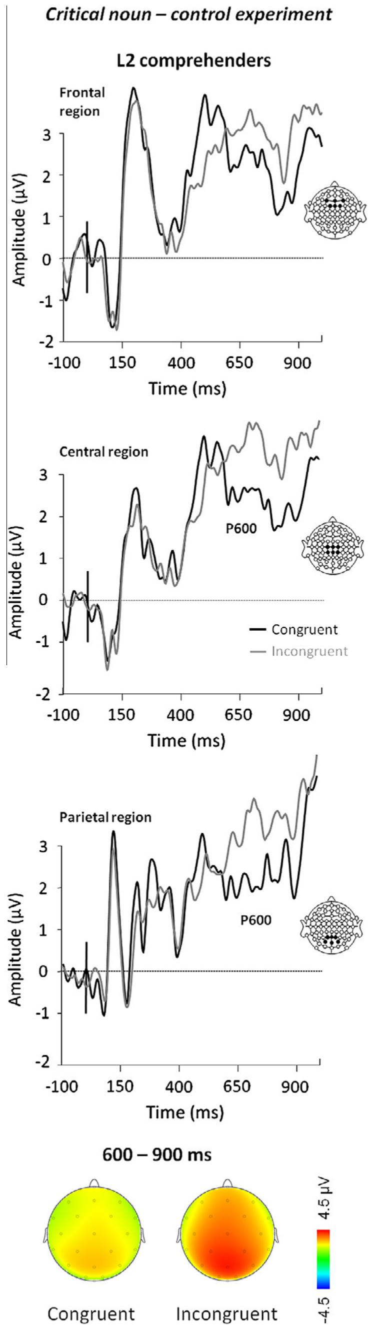

Fig. 1. Top: Event-related potential results for the critical noun of the sentence, in the control experiment, in *L2 comprehenders*. Time zero indicates the presentation of the noun. Black lines depict ERPs measured for nouns phonologically congruent with preceding articles; grey lines depict ERPs measured for nouns phonologically incongruent with preceding articles. ERPs measured over the frontal (F3, Fz, F4, FC1, FCz, FC2), central (C1, Cz, C2, CP1, CPz, CP2) and parietal (P1, Pz, P2, PO3, POz, PO4) scalp. Negativity is plotted down. Bottom: Topographic distributions of congruent and incongruent article-noun phonological agreement in the [600–900] ms time-window, in *L2 comprehenders*.

Procedure

Participants were presented with the first part of each sentence as a whole on a computer screen. They were instructed to press the space bar when ready to start reading the second part. The second part of each sentence was presented one word at a time (200 ms duration and 500 ms stimulus onset asynchrony) in the centre of a computer screen. The instruction was to read each sentence silently and to answer 'yes' or 'no' when a comprehension question appeared on the screen (¼ of the trials) by pressing a YES or NO button on a keyboard.

Electrophysiological recording and data analyses

Electrophysiological data were recorded in reference to electrode Cz at a rate of 1 kHz from 64 Ag/AgCl electrodes placed according to the 10–20 convention. Impedances were kept below 5 kOhm. EEG activity was filtered on-line band pass between 0.1 Hz and 200 Hz and re-filtered off-line with a 30 Hz low pass zero-phase shift digital filter.

Please cite this article in press as: Martin, C. D., et al. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.08.001

Eye blink artifacts were mathematically corrected based on a model artifact computed from a minimum of 50 individual artifacts in each participant using the procedure implemented in Scan 4.3 (Neuroscan, Inc., El Paso, TX, USA) and remaining artifacts were manually removed. Epochs ranged from −100 to 1000 ms after the onset of the final article and from −100 to 1000 ms after the onset of the final noun. Baseline correction was performed in reference to pre-stimulus activity and individual averages were digitally re-referenced to a global average reference. ERP components were defined based on the grand averages and analyzed in time-windows classically used to explore the N400 and Anterior Positivity in similar paradigms (see for instance DeLong et al., 2012; Federmeier et al., 2007). This allowed mean amplitude analyses in the following intervals: 250–400 ms for the N400 component following the presentation of the article (N400a), 300–500 ms for the N400 following the presentation of the final noun (N400n) and 500–900 ms for the Anterior Positivity following the presentation of the final noun. Moreover, visual inspection of the ERP signal time-locked on the presentation of the final noun revealed that Expectation effects started earlier than the classical N400 time-window. Thus, ERP mean amplitudes were also analyzed in the [200–300] ms interval following the presentation of the final noun. Mean peak amplitudes were subjected to a repeated measures analysis of variance (ANOVA) for each component with Expectation (expected *versus* unexpected) and Region (3 levels) as within-subject factors and Group (L1 *versus* L2) as between-subject factor. Analyses of the 3 peaks (N400a, N400n and Anterior Positivity) were conducted in 3 regions: Frontal (F3, Fz, F4, FC1, FCz, FC2 electrodes), Central (C1, Cz, C2, CP1, Cpz, CP2 electrodes) and Parietal (P1, Pz, P2, PO3, POz, PO4 electrodes). In addition, the time-window of differences between conditions was estimated using ms-by-ms paired *t*-tests for the contrast of interest (expected *versus* unexpected words). Unstable differences (remaining below $p = .05$ for less than 30 ms) were discarded (Rugg & Doyle, 1993).

Results

The general ANOVA performed on N400a mean amplitude (N400 elicited by the article preceding the final noun; [250–400] ms time-window) revealed a significant Expectation effect ($F[1,36] = 6.45$, $p < .02$), a significant Region effect ($F[2,72] = 14.33$, $p < .001$) and a marginal Group effect ($F[1,36] = 4.01$, $p = .06$; Fig. 2). The Expectation × Group interaction was almost significant ($F[1,36] = 3.40$, $p = .07$). Neither the Region × Group interaction ($F[2,72] = 2.41$, $p = .10$), the Expectation × Region interaction ($F[2,72] = .40$, $p = .67$) nor the triple interaction ($F[2,72] = .69$, $p = .51$) was significant. Because of the significant Region effect, we performed 3 separated ANOVAs on each of the 3 regions.

The repeated measures ANOVA performed on N400a mean amplitude over the frontal region revealed a significant Expectation effect ($F[1,36] = 3.81$, $p = .05$), a marginal Group effect ($F[1,36] = 4.52$, $p = .06$) and a significant Expectation × Group interaction ($F[1,36] = 3.17$, $p = .02$). The N400 was larger for unexpected than expected articles. *Post-hoc* analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferroni test,* $p = .05$) but not L2 ($p = .08$) comprehenders. Planned comparisons also showed that the Expectation effect was significant in L1 ($F[1,18] = 4.46$, $p = .04$) but not L2 ($F[1,18] = .21$, $p = .60$) comprehenders.

The repeated measures ANOVA performed on N400a mean amplitude over the central region revealed a significant Expectation effect ($F[1,36] = 3.95$, $p = .04$), a marginally significant Group effect ($F[1,36] = 3.32$, $p = .05$) and a significant Expectation × Group interaction ($F[1,36] = 6.01$, $p = .02$). The N400 was larger for unexpected than expected articles. *Post-hoc* analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferroni test,* $p = .03$) but not L2 ($p = 1.00$) comprehenders. The Group effect was marginally significant for unexpected articles ($p = .05$) but not for expected articles ($p = 1.00$). Planned comparisons also showed that the Expectation effect was significant in L1 ($F[1,18] = 6.39$, $p = .02$) but not L2 ($F[1,18] = .54$, $p = .69$) comprehenders, and that the Group effect was significant for unexpected articles ($F[1,18] = 8.03$, $p = .03$) but not for expected articles ($F[1,18] = .15$, $p = .81$).

The repeated measures ANOVA performed on N400a mean amplitude over the parietal region did not reveal any significant effects (Group effect: $F[1,36] = 1.60$, $p = .19$; Expectation effect: $F[1,36] = 2.92$, $p = .10$; Expectation × Group interaction $F[1,36] = .97$, $p = .45$).

In sum, the N400 Expectation effect on the article was significant in L1 but not L2 comprehenders and was significant over frontal and central regions (see Fig. 2). Moreover, the N400 mean amplitude significantly differed between the two groups for unexpected articles but not for expected articles.[3]

The ANOVA performed in the [200–300] ms time-window following the presentation of the final noun revealed a significant Expectation effect ($F[1,36] = 17.84$, $p < .001$), a significant Region effect ($F[2,72] = 12.25$, $p < .001$) and a significant Group effect ($F[1,36] = 15.02$, $p < .001$; Fig. 3). The Region × Group ($F[2,72] = 7.86$, $p < .01$), the Expectation × Region ($F[2,72] = 25.91$, $p < .01$) and the triple interaction ($F[2,72] = 4.31$, $p = .02$) were significant. The Expectation × Group interaction ($F[1,36] = 4.64$, $p = .07$) was marginally significant. Because of the significant Region effect, we performed 3 separated ANOVAs on each of the 3 regions.

The repeated measures ANOVA performed in the [200–300] ms time-window over the frontal region revealed no significant effects (Expectation effect: $F[1,36] = 3.09$, $p = .09$; Group effect: $F[1,36] = 1.78$, $p = .19$; Expectation × Group interaction ($F[1,36] = 1.66$, $p = .18$).

[3] The time-window for N400a analyses (250–400 ms) was chosen based on previous research using similar paradigms (DeLong et al., 2012; Federmeier et al., 2007). For the sake of completeness, ERP mean amplitudes were also analyzed in the time-window of maximal differences between expected and unexpected articles (440–670 ms) and the results were similar. The only difference was that the Expectancy effect was significant in L1 comprehenders over central and parietal regions instead of frontal and central regions.

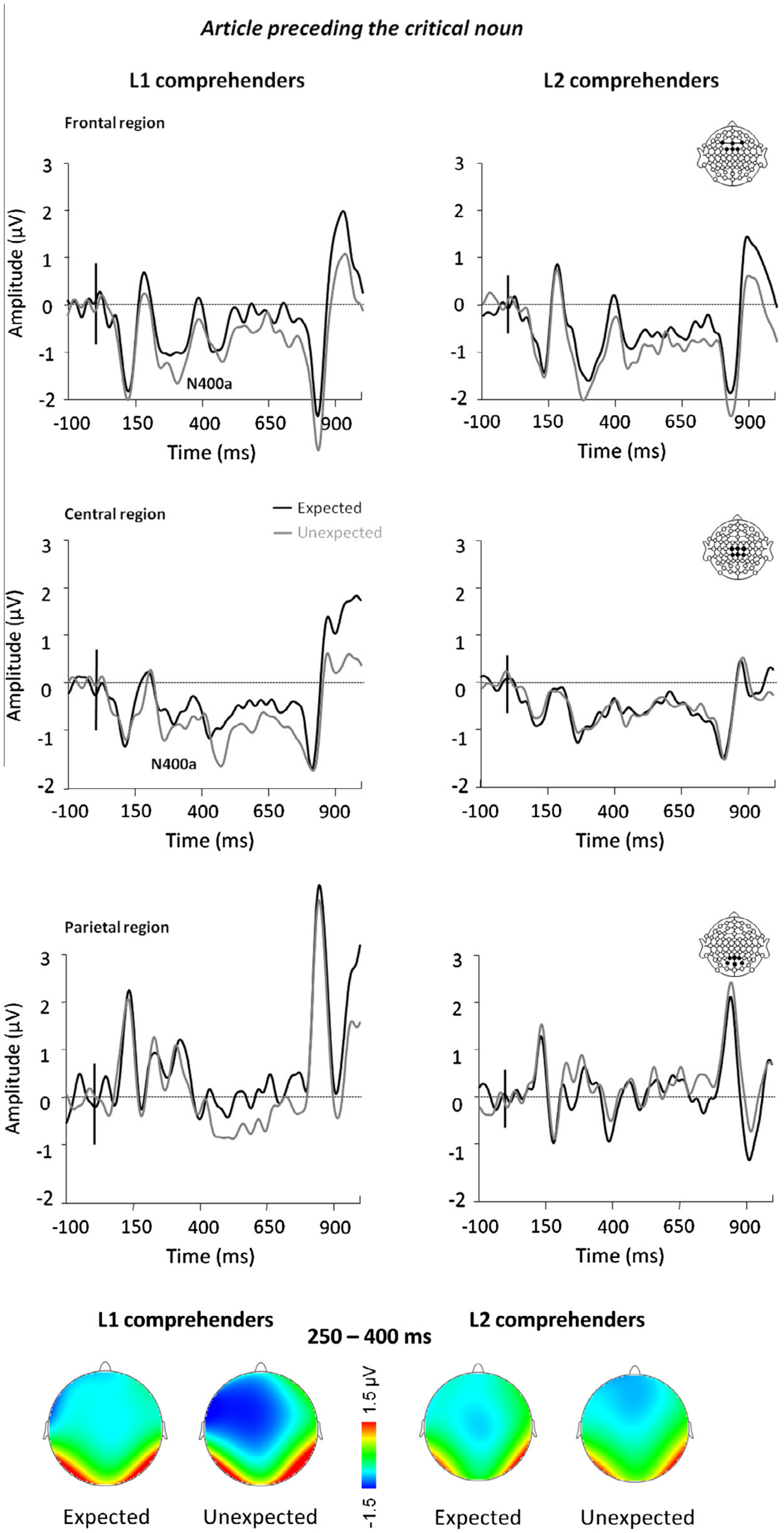

Fig. 2. Top: Event-related potential results for the article preceding the critical noun of the sentence, in *L1 comprehenders* (left panel) and *L2 comprehenders* (right panel). Time zero indicates the presentation of the article. Black lines depict ERPs measured for expected articles; grey lines depict ERPs measured for unexpected articles. ERPs measured over the frontal (F3, Fz, F4, FC1, FCz, FC2), central (C1, Cz, C2, CP1, CPz, CP2) and parietal (P1, Pz, P2, PO3, POz, PO4) scalp. N400a is the N400 wave elicited by the article. Negativity is plotted down. Bottom: Topographic distributions of expected and unexpected articles in the [250–400] ms time-window, in *L1 comprehenders* (left panel) and *L2 comprehenders* (right panel).

Please cite this article in press as: Martin, C. D., et al. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.08.001

The repeated measures ANOVA performed in the [200–300] ms time-window over the central region revealed a significant Expectation effect ($F[1,36] = 42.40$, $p < .001$), a significant Group effect ($F[1,36] = 5.79$, $p = .02$) and a marginally significant Expectation × Group interaction ($F[1,36] = 4.03$, $p = .05$). The ERP mean amplitude was more negative for unexpected than expected nouns and the overall ERP mean amplitude in this region was more positive in L1 than L2 comprehenders. *Post-hoc* analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferonni test*, $p < .001$) but not in L2 comprehenders ($p = .08$). Planned comparisons showed that the Expectation effect was significant in L1 ($F[1,18] = 18.85$, $p < .001$) but not in L2 ($F[1,18] = 1.95$, $p = .09$) comprehenders, and that the Group effect was significant for expected nouns ($F[1,18] = 11.22$, $p < .01$) and for unexpected nouns ($F[1,18] = 9.97$, $p < .01$).

The repeated measures ANOVA performed in the [200–300] ms time-window over the parietal region revealed a significant Group effect ($F[1,36] = 10.32$, $p < .01$), a significant Expectation effect ($F[1,36] = 7.52$, $p < .01$) and a significant Expectation × Group interaction ($F[1,36] = 10.25$, $p < .01$). The ERP mean amplitude was more negative for unexpected than expected nouns and the overall ERP mean amplitude in this region was more positive in L1 than L2 comprehenders. Post-hoc analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferonni test*, $p = .01$) but not in L2 comprehenders ($p = .81$). Planned comparisons showed that the Expectation effect was significant in L1 ($F[1,18] = 6.88$, $p < .01$) and not in L2 ($F[1,18] = .29$, $p = .60$) comprehenders, and that the Group effect was significant for expected nouns ($F[1,18] = 25.11$, $p < .001$) and for unexpected nouns ($F[1,18] = 11.23$, $p < .01$).

In sum, there was a significant modulation by Expectancy of the ERP mean amplitudes in the [200–300] ms time-window following the presentation of the final noun: Unexpected nouns elicited more negative ERP waves than expected nouns over central and parietal regions in L1 comprehenders only (see Fig. 3). Overall, the ERP mean amplitude was more positive in L1 than L2 comprehenders for both expected and unexpected nouns.

The ANOVA performed on N400n mean amplitude (N400 elicited by the final noun; [300–500] ms time-window) revealed a significant Expectation effect ($F[1,36] = 33.65$, $p < .001$), a significant Region effect ($F[2,72] = 13.46$, $p < .001$) and a significant Group effect ($F[1,36] = 17.25$, $p < .001$; Fig. 3). The Expectation × Group interaction ($F[1,36] = 4.71$, $p = .03$), the Region × Group ($F[2,72] = 3.34$, $p = .04$), the Expectation × Region ($F[2,72] = 6.09$, $p < .01$) and the triple interaction ($F[2,72] = 4.83$, $p < .01$) were significant. Because of the significant Region effect, we performed 3 separated ANOVAs on each of the 3 regions.

The repeated measures ANOVA performed on N400n mean amplitude over the frontal region revealed a significant Expectation effect ($F[1,36] = 15.11$, $p < .001$), a significant Group effect ($F[1,36] = 5.88$, $p < .01$) and a significant Expectation × Group interaction ($F[1,36] = 5.29$, $p = .01$). The ERP amplitude was more positive over this region in L1 than L2 comprehenders and for unexpected than expected nouns. *Post-hoc* analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferonni test*, $p < .001$) and not in L2 comprehenders ($p = .08$). Planned comparisons confirmed that the Expectation effect was significant in L1 ($F[1,18] = 19.23$, $p < .001$) and not in L2 ($F[1,18] = .15$, $p = .70$) comprehenders. The Group effect was significant for unexpected nouns ($F[1,18] = 17.77$, $p < .01$) but not for expected nouns ($F[1,18] = .17$, $p = .84$).

The repeated measures ANOVA performed on N400n mean amplitude over the central region revealed a significant Expectation effect ($F[1,36] = 35.34$, $p < .001$), a significant Group effect ($F[1,36] = 10.21$, $p < .01$) and a significant Expectation × Group interaction ($F[1,36] = 5.14$, $p = .03$). The N400 component was larger for unexpected than expected nouns and the overall ERP amplitude in this region was more positive in L1 than L2 comprehenders. *Post-hoc* analysis of the interaction revealed that the Expectation effect was significantly larger in L1 (*Bonferonni test*, $p < .001$) than L2 comprehenders ($p = .04$). Planned comparisons showed that the Expectation effect was significant in both L1 ($F[1,18] = 9.19$, $p = .001$) and L2 ($F[1,18] = 7.78$, $p < .01$) comprehenders, and that the Group effect was significant for expected nouns ($F[1,18] = 14.22$, $p < .01$) and for unexpected nouns ($F[1,18] = 12.98$, $p < .01$).

The repeated measures ANOVA performed on N400n mean amplitude over the parietal region revealed a significant Group effect ($F[1,36] = 11.09$, $p < .001$), a significant Expectation effect ($F[1,36] = 8.50$, $p < .01$) and a significant Expectation × Group interaction ($F[1,36] = 8.61$, $p < .01$). The N400 component was larger for unexpected than expected nouns and the overall ERP amplitude over this region was more positive in L1 than L2 comprehenders. Post-hoc analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferonni test*, $p = .01$) and marginally significant in L2 comprehenders ($p = .07$). Planned comparisons showed that the Expectation effect was significant in L1 ($F[1,18] = 4.71$, $p = .01$) and marginally significant in L2 ($F[1,18] = 3.20$, $p = .05$) comprehenders, and that the Group effect was significant for expected nouns ($F[1,18] = 31.01$, $p < .001$) and for unexpected nouns ($F[1,18] = 10.99$, $p < .01$).

In sum, there was a significant N400 modulation by expectancy for the final noun in both L1 and L2 comprehenders over central and parietal regions, with a larger effect in L1 comprehenders (see Fig. 3). The expectancy effect was significant in L1 comprehenders over frontal region. Overall, in the [300–500] ms time-window, the ERP amplitude was more positive in L1 than L2 comprehenders for both expected and unexpected nouns.[4]

The ANOVA performed on the Anterior Positivity mean amplitude ([500–900] ms time-window) revealed a significant Region effect ($F[2,72] = 36.07$, $p < .001$) and a significant Expectation × Group interaction ($F[2,72] = 6.43$, $p = .02$;

[4] The time-window for N400n analyses (300–500 ms) was chosen based on previous research using similar paradigms (DeLong et al., 2012; Federmeier et al., 2007). For the sake of completeness, ERP mean amplitudes were also analyzed in the time-window of maximal differences between expected and unexpected articles (220–420 ms for L1 and 320–520 ms for L2 comprehenders) and the results were similar. The only difference was that the Expectancy effect was significant in both groups over the central region instead of central and parietal regions.

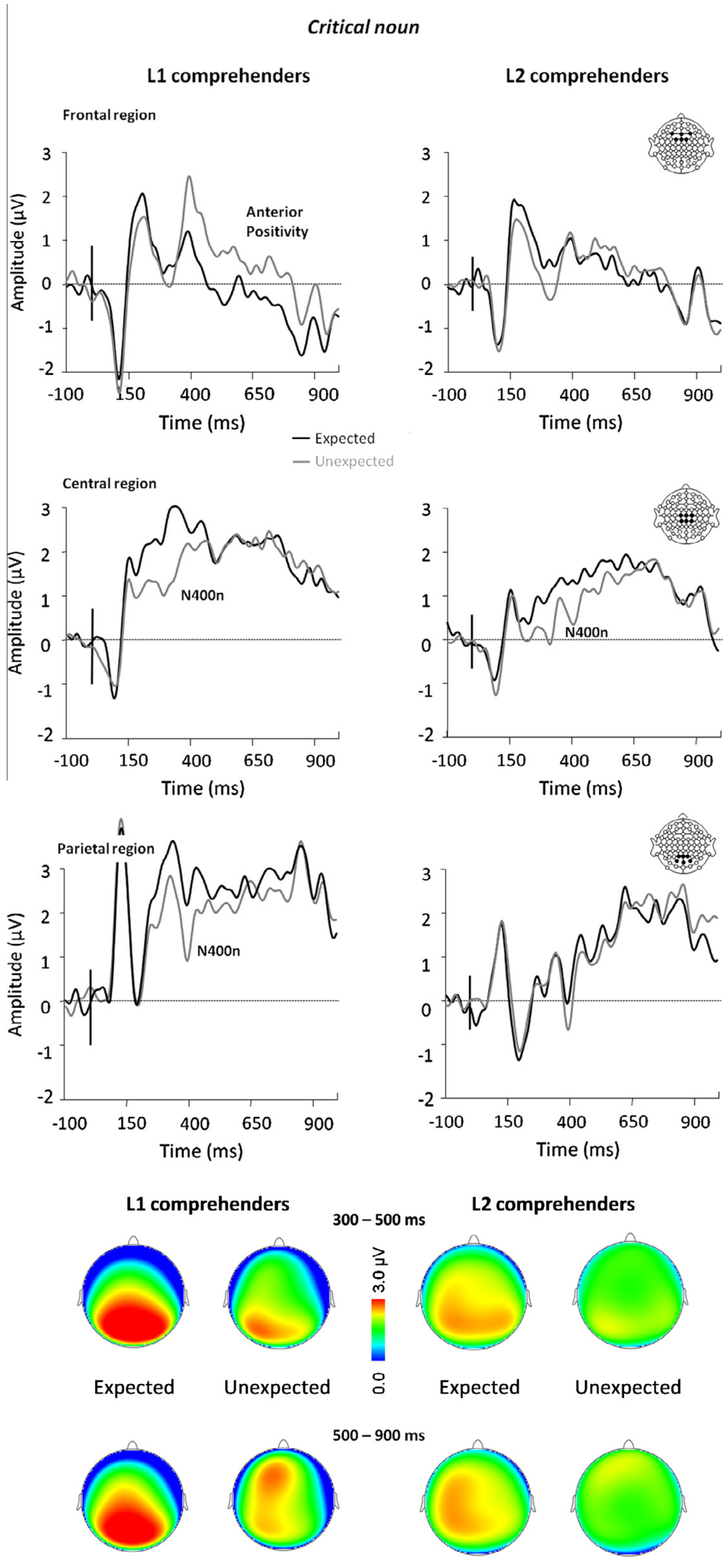

Fig. 3. Top: Event-related potential results for the critical noun of the sentence, in *L1 comprehenders* (left panel) and *L2 comprehenders* (right panel). Time zero indicates the presentation of the final noun. Black lines depict ERPs measured for expected nouns; grey lines depict ERPs measured for unexpected nouns. ERPs measured over the frontal (F3, Fz, F4, FC1, FCz, FC2), central (C1, Cz, C2, CP1, CPz, CP2) and parietal (P1, Pz, P2, PO3, POz, PO4) scalp. N400n is the N400 wave elicited by the noun. Anterior Positivity is the Post-N400 Positivity elicited by the noun. Negativity is plotted down. Bottom: Topographic distributions of expected and unexpected nouns in the [300–500] ms and [500–900] ms time-windows, in *L1 comprehenders* (left panel) and *L2 comprehenders* (right panel).

Fig. 3). No other effect or interaction was significant (all *ps* > .11). Because of the significant Region effect, we performed 3 separated ANOVAs on each of the 3 regions.

The subsequent ANOVA performed on the Anterior Positivity mean amplitude over the frontal region revealed a significant Expectation effect ($F[1,36] = 9.09$, $p < .01$), a significant Expectation × Group interaction ($F[1,36] = 7.75$, $p < .01$) but no effect of Group ($F[1,36] = 1.45$, $p = .24$). The Anterior Positivity was larger for unexpected than expected nouns. Post-hoc analysis of the interaction revealed that the Expectation effect was significant in L1 (*Bonferonni test*, $p < .01$) but not L2 ($p = 1.00$) comprehenders. Planned comparisons also showed that the Expectation effect was significant in L1 ($F[1,18] = 5.75$, $p < .01$) but not L2 ($F[1,18] = .85$, $p = .25$) comprehenders.

The repeated measures ANOVA performed on the Anterior Negativity mean amplitude over the central region did not reveal any significant effects (Group effect: $F[1,36] = 1.88$, $p = .17$; Expectation effect: $F[1,36] = .53$, $p = .47$; Expectation × Group interaction: $F[1,36] = .13$, $p = .73$).

The repeated measures ANOVA performed on the Anterior Negativity mean amplitude over the parietal region did not reveal any significant effects (Group effect: $F[1,36] = 1.09$, $p = .35$; Expectation effect: $F[1,36] = .00$, $p = .98$; Expectation × Group interaction: $F[1,36] = .88$, $p = .36$).

In sum, the Expectancy effect on the Anterior Negativity elicited by the noun was observed in L1 but not L2 comprehenders over frontal region (see Fig. 3).

In the ms-by-ms paired *t*-test analysis, the first sustained significant difference between expected and unexpected articles was found at 450 ms in L1 comprehenders. The two conditions did not differ significantly at any time of the epoch in L2 comprehenders. The first sustained significant difference between expected and unexpected nouns was found at 195 ms in L1 and 300 ms in L2 comprehenders.

Discussion

The goal of the present study was to investigate whether L2 comprehenders predict upcoming words during sentence reading in the same way as L1 comprehenders. To assess expectation effects, we measured the N400 elicited by the article preceding the final noun of sentences featuring a highly predictive context. We also explored the N400 effect elicited by the final word and the subsequent Anterior Positivity effect over frontal regions of the scalp.

In L1 comprehenders, unexpected articles elicited greater N400 amplitudes than expected articles, replicating the findings of DeLong et al. (2005). We conclude, as argued previously (DeLong et al., 2005; van Berkum et al., 2005; Wicha et al., 2004), that L1 comprehenders tend to actively predict the final nouns of sentences with a highly constrained context. In contrast, L2 comprehenders who have acquired their second language after the age of 8 do not show the same lexical prediction effect. This result is interesting because L2 comprehenders do predict the same final noun as L1 comprehenders according to cloze probability tests and they are also aware of the phonological rule which governs the selection of the appropriate article. Note that the negative ERP deflection observed here does not have the classical N400 latency and scalp distribution. Since unexpected articles have been shown to elicit negative (DeLong et al., 2005) or positive ERP deflections (van Berkum et al., 2005; Wicha et al., 2004) with various latencies, further research is needed before strong claims can be made regarding the cognitive processes underlying ERP modulations elicited by unexpected articles (see Van Petten & Luka, 2012 for an extensive review). Because of tentative nature of the interpretation of the negative component elicited by unpredicted articles, we can only speculate as regards the cognitive processes involved, and refer to the elicited negativity as the N400a. Whatever the cognitive interpretation of this effect, differences of N400a amplitude indicate differential processing of expected and unexpected articles in L1, but not in L2 readers.

Based on the pattern of results observed for the processing of the article, we can conclude that L2 comprehenders do not actively predict final nouns when reading highly constrained sentences to the same extent as L1 comprehenders do. Furthermore, the fact that N400a mean amplitude was larger for unexpected articles in L1 readers compared to the other three conditions suggests that processing of unexpected articles was more effortful for participants who actively predicted final nouns, whereas it was not so for participants who did not predict. Since the conclusion is based on a null result, we cannot claim that L2 comprehenders do not predict at all. Nevertheless, the lack of a reliable ERP modulation by expectation on the article indicates, at the very least, that L2 comprehenders predict to a weaker extent than L1 comprehenders. Therefore, the two alternative possibilities presented in the introduction provide alternative explanations of our pattern of results: (1) L2 comprehenders may predict upcoming words more slowly than L1 comprehenders, and thus not fast enough to influence linguistic processing prior to the onset of the critical noun. If this were the case, N400 and Anterior Positivity modulations by expectation would be observed on the noun, similarly to L1 readers; (2) the second alternative is that L2 readers rely exclusively on passive integration during sentence processing by building up message-level representations of context meaning that affect word processing when the word is encountered. Alternatively, L2 readers might not construct such representations but rather exclusively rely on passive resonance. According to this alternative (passive integration), no expectation effect should be observed on the Anterior Positivity elicited by the critical noun itself (suggesting insensitivity to failed lexical prediction; DeLong et al., 2012). But N400 modulation by expectation should still be observed on the noun. Since these two alternatives do not lead to the same hypotheses regarding the final noun, we will now discuss expectation effects on the final noun to provide further support for weaker (or absent) lexical prediction in L2 comprehension.

In L1 comprehenders, unexpected final nouns elicited greater N400 amplitudes than expected final nouns, replicating previous findings (see, for instance, DeLong et al., 2005, 2012; van Berkum et al., 2005; Wicha et al., 2004).

Please cite this article in press as: Martin, C. D., et al. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.08.001

Furthermore, the expectation effect started earlier than the N400 component (195 ms after the onset of the noun according to the ms-by-ms paired *t*-test analysis) and was already significant in the [200–300] ms time-window. This observation is important evidence of active lexical prediction in L1 comprehenders, based on the assumption that lexical predictions involve pre-activation of sub-lexical representations (conceptual, phonological and orthographic) of the predicted words (Lau et al., 2013). In fact, if a lexical item and its orthographic representation are actively predicted, orthographic overlap (or non-overlap) between the predicted lexical item and the actual word presented in the sentence should modulate orthographic processing stages ([200–300] ms time-window; see Grainger & Holcomb, 2009). Given the topographical distribution (central region) and temporality of the expectation effect in the [200–300] ms time-window, the most likely interpretation of this effect is a modulation of the N250 component, related to orthographic processing (Grainger & Holcomb, 2009). Classically observed in visual masked priming studies, the N250 is sensitive to the degree of prime-target orthographic overlap (Grainger & Holcomb, 2009; Holcomb & Grainger, 2006; Kiyonaga, Midgley, Holcomb, & Grainger, 2007). As previously argued by Lau et al. (2013), the reduced N250 mean amplitude for expected words in the present study likely arises from orthographic overlap between the predicted and the actual lexical item (the opposite to a lack of orthographic overlap between the predicted and the actual lexical item in the unexpected word condition, which is reflected in a greater N250 mean amplitude). Note that several other studies have provided significant evidence that lexical predictions may affect orthographic and phonological processing stages, earlier than the classical [300–500] ms time-window of the N400 component (see for instance Federmeier, Mai, & Kutas, 2005; Molinaro & Carreiras, 2010; Roehm, Bornkessel-Schlewiesky, Roesler, & Schlesewsky, 2007; Vespignani, Canal, Molinaro, Fonda, & Cacciari, 2010).

Finally, the Anterior Positivity was significantly enhanced for unexpected relative to expected final nouns in L1 readers.[5] This effect confirms that L1 readers actively predicted lexical items, and incurred a reprocessing cost following a failed lexical prediction (DeLong et al., 2011, 2012; Federmeier et al., 2007; Otten & Van Berkum, 2008; Van Petten & Luka, 2012).

In L2 comprehenders, unexpected final nouns also elicited greater N400 amplitudes than expected final nouns. However, this expectation effect was significantly smaller than in L1 comprehenders. Secondly, the expectation effect did not start earlier than the classical N400 time-window (300 ms after the onset of the noun according to the ms-by-ms paired *t*-test analysis). Finally, the Anterior Positivity mean amplitude did not differ for expected and unexpected final nouns. In the following, we will discuss these three important differences between final noun processing in L1 and L2 comprehenders.

[5] Note that the 'reversed' N400 expectation effect over frontal region in L1 comprehenders (cf. 'Result' section and Fig. 3) is probably the trace of an early Anterior Positivity (see DeLong et al., 2011; Otten & Van Berkum, 2008 for similar early anterior effects of expectation).

The presence of an N400 expectancy effect in L2 comprehenders at least guarantees that sentences were processed for meaning and that L2 comprehenders experienced semantic processing difficulties for the unexpected nouns. In other words, the lack of N400 modulation elicited by the article cannot be explained in terms of relative semantic processing incapacity on the part of L2 comprehenders. In both groups, semantic processing of best completion nouns was facilitated, which suggests message-level representation build up and/or activation of a semantic relatedness network through passive resonance during sentence context processing (Gerrig & McKoon, 1998; Myers & O'Brien, 1998; Paczynski & Kuperberg, 2012). Note that this observation of sentence context influence on semantic processing in L2 is consistent with previous literature on L2 sentence processing (Altarriba et al., 1996; Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Assche et al., 2011; Van Hell & De Groot, 2008).

More importantly, while it was observed in both participant groups, the N400 Expectation effect was significantly larger in L1 than in L2 comprehenders. This interaction is another important argument in concluding that L2 comprehenders did not predict upcoming words to the same extent as L1 comprehenders did. The N400 modulation during semantic processing of the critical noun in L1 comprehenders is probably influenced by both passive integration processes and active lexical prediction mechanisms (DeLong et al., 2005; Lau et al., 2013). In other words, the size of the N400 effect in L1 comprehenders would reflect semantic processing of the 'expected' noun facilitated by message-level context and/or passive resonance but also active lexical prediction. The relatively smaller N400 effect in L2 comprehenders would then reflect semantic processing of the final nouns facilitated by message-level context and/or passive resonance but to a lesser extent by lexical prediction. In this group, the sentence context stored in working memory made expected nouns easier to integrate than unexpected nouns but expected nouns were not actively predicted before their presentation, or at least were predicted to a weaker extent.

Another important indication that L2 readers were not actively predicting as L1 readers were is the observation that expectancy effects appeared later in L2 than in L1 comprehenders (300 ms *versus* 195 ms). As previously discussed, the modulation of the ERP signal in the [200–300] ms time-window by expectation likely arises from orthographic overlap (or non-overlap) between the predicted and the actual lexical item (Lau et al., 2013). Thus, the lack of such modulation of early ERP signal is likely due to a lack of active lexical prediction in L2 comprehenders. Nevertheless, we acknowledge that the later onset of expectation effects in L2 relative to L1 comprehenders could be only the consequence of a smaller size of the N400 mean amplitude modulation in L2 comprehenders.

Finally, further evidence in favour of a lack of lexical prediction in L2 comprehenders comes from the investigation of the Anterior Positivity component (DeLong et al., 2011, 2012; Federmeier et al., 2007; Otten & Van Berkum, 2008; Van Petten & Luka, 2012). As stated earlier in the

Please cite this article in press as: Martin, C. D., et al. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.08.001

discussion, results on the article revealed that L2 comprehenders were not predicting upcoming words as L1 comprehenders were. If L2 comprehenders were predicting upcoming words, but not fast enough to influence linguistic processing prior to the onset of the critical noun, then the Anterior Positivity elicited by the final noun would have been modulated by expectation. In fact, DeLong et al. (2012) used a similar experimental design to investigate lexical prediction in older adults and observed an Anterior Positivity modulation on the critical noun in the absence of a significant N400 effect for the article. They proposed that lexical prediction is slower in older than younger adults, leading to a pre-activation of the expected noun somewhere between the presentation of the article and that of the noun, and thus resulting in a modulation of the ERP elicited by the noun but not of that elicited by the article. Alternatively, if L2 comprehenders were relying mainly on passive integration mechanisms without lexical prediction, no expectation effect should be observed on the Anterior Positivity elicited by the critical noun itself. As no Anterior Positivity effect was observed in L2 comprehenders, we can conclude that they were not sensitive to violations of sentence-based lexical predictions. This result strongly suggests that L2 comprehenders were not actively predicting lexical items, or at least not to the same extent as L1 comprehenders.

Note that this lack of late ERP modulation does not entirely ensure that L2 comprehenders failed to actively predict critical words. It might be that they predicted upcoming words but did not suffer from the cost of overriding lexical predictions in the same manner as L1 comprehenders. Nevertheless, considering the lack of N400 modulation by expectation on the preceding article, as well as the lack of early expectation effects on the noun, and the smaller N400 modulation on the noun relative to L1 comprehenders, the absence of Anterior Positivity modulation is consistent with a lack of lexical prediction during L2 sentence reading. Recall also that the reading rate was rather slow (one word each 700 ms) which gave ample opportunity to predict upcoming words. In any case, the main interest in studying lexical prediction in L2 comprehension is to know whether L2 comprehenders can actively predict when they listen to speech rather than whether they do so when they read. As the speech perception rate is much faster than the reading rate of one word every 700 ms used here, it is likely that L2 comprehenders do not predict upcoming words when they listen to speech either.

The present results suggest that L2 comprehenders do not predict upcoming words in highly constrained sentences as L1 comprehenders do. L2 comprehenders appear to mainly rely on passive integration. Nevertheless, we cannot distinguish between two 'passive integration' alternatives. It remains to be determined whether (1) L2 comprehenders use all constraints to build up message-level representations of context meaning but do it slower, and therefore experience effects of these representations on word integration only when the word is read, or whether (2) they fail to incrementally build up message-level representations of context meaning and rely exclusively on passive resonance.

A potential limitation of this study relates to article-noun phonological agreement in English. In pre-tests, we observed that late Spanish–English bilinguals are highly sensitive to phonological agreement between articles and nouns when reading in English, both off-line and on-line. Moreover, the fact that L2 comprehenders showed a reduced N400 effect and no Anterior Positivity effect on the noun in the main task suggests that they predicted words to a lesser extent and not that they predicted similarly to L1 comprehenders but without taking into account article-noun phonological agreement. Nevertheless, further investigation of this is needed, manipulating gender agreement for example (see Wicha, Bates, et al., 2003; Wicha, Moreno, et al., 2003; Wicha et al., 2004). It would also be informative to test critical words in non-final position within the sentences. Finally, it must be noted that our conclusions refer to late, mid-proficient Spanish–English bilinguals. Further investigation is needed to establish whether lexical prediction in L2 depends on language proficiency and/or age of acquisition, which is likely.

It also remains to be established whether the pattern of results found here is due to the bilingual status of the participants, or to task difficulty (L2 readers having less cognitive resources available for lexical prediction relative to L1 readers). Indeed, it is possible that weaker lexical prediction is an inherent side effect of processing a non-native language. Alternatively, weaker lexical prediction might be due to comprehension mechanisms that are not dynamic enough, as has been previously observed in older adults (DeLong et al., 2012).

To conclude, we show that semantic processing in L1 and L2 differ at least in one way: L2 comprehenders do not actively predict upcoming words during language comprehension to the same extent as L1 comprehenders. The lack of lexical prediction of the critical noun (Anterior Positivity) and its preceding article in L2 comprehenders suggests that active lexical prediction in this group is more limited or at least less efficient than in L1 comprehenders. Such weaker lexical prediction in L2 might be one of the consequences of overall slower and less accurate linguistic processing stages in L2 relative to L1.

Acknowledgments

The authors would like to thank the editor, the anonymous reviewers and Margaret Gillon-Dowens for their valuable comments and suggestions to improve the manuscript. This work was supported by grants from the Spanish Government (PSI2008-01191, PSI2011-23033, Consolider Ingenio 2010 CSD2007-00012) and the Catalan Government (Consolidado SGR 2009-1521). C.M. was supported by the Spanish Government (Grant Juan de la Cierva) and is now supported by the Basque Foundation for Science (IKERBASQUE) and the BCBL institution. G.T. is supported by the Economic and Social Research Council (ES/E024556/1) and the European Research Council (ERC-SG-209704). A.F. is supported by the Catalan Government (Grant Beatriu de Pino).

References

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*, 477–492.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America, 121*, 2339–2349.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 279–300). Lawrence Erlbaum.

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology, 41*, 60–74.

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

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

DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language, 121*, 226–239.

Elston-Güttler, K. E., & Friederici, A. D. (2007). Ambiguous words in sentences: Brain indices for native and non-native disambiguation. *Neuroscience Letters, 414*, 85–89.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495.

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

Federmeier, K. D., Mai, H., & Kutas, M. (2005). Both sides get the point: Hemispheric sensitivities to sentential constraint. *Memory & Cognition, 33*, 871–886.

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

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior, 18*, 1–20.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition, 14*, 379–399.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology, 50*, 119–148.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). North Holland.

Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes, 26*, 67–86.

Gillon-Dowens, M., Vergara, M., Barber, H., & Carreiras, M. (2010). Morpho-syntactic processing in late L2 learners. *Journal of Cognitive Neuroscience, 22*, 1870–1887.

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word. *Language and Linguistics Compass, 3*, 128–156.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81.

Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research, 28*, 715–728.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 819–836). Boston, MA: MIT Press.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438–441.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 123–142.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30*, 251–266.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*, 638–650.

Holcomb, P. J., & Grainger, J. (2006). The time-course of masked repetition priming: An event-related brain potential investigation. *Journal of Cognitive Neuroscience, 18*, 1631–1643.

Kiyonaga, K. K., Midgley, J., Holcomb, P. J., & Grainger, J. (2007). Masked cross-modal repetition priming: An ERP investigation. *Language and Cognitive Processes, 22*, 337–376.

Kleiman, G. M. (1980). Sentence frame contexts and lexical decisions: Sentence-acceptability and word-relatedness effects. *Memory & Cognition, 8*, 336–344.

Kotz, S. A., & Elston-Guettler, K. E. (2007). Bilingual semantic memory revisited: ERP and fMRI evidence. In J. Hart & M. Kraut (Eds.), *The neural basis of semantic memory* (pp. 105–132). Cambridge: Cambridge University Press.

Kotz, S. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language, 109*, 68–74.

Kuperberg, G. R., Paczynski, M., & Ditman, T. (2011). Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience, 23*(5), 1230–1246.

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

Kutas, M., & Federmeier, KD. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 4*, 463–470.

Kutas, M., Moreno, EM., & Wicha, NY. (2009). Code-switching and the brain. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 289–306). London: Cambridge Univ. Press.

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

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single word contexts. *Journal of Cognitive Neuroscience, 25*, 484–502.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence From eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 381–390.

Molinaro, N., & Carreiras, M. (2010). Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. *Biological Psychology, 83*, 176–190.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research, 22*, 205–220.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*, 131–157.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology General, 106*, 226–254.

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes, 45*, 464–496.

Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language, 67*, 426–448.

Roehm, D., Bornkessel-Schlewiesky, I., Roesler, F., & Schlesewsky, M. (2007). To predict or not to predict: Influences of task and strategy on the processing of semantic relations. *Journal of Cognitive Neuroscience, 19*, 1259–1274.

Rugg, M. D., & Doyle, M. C. (1993). An ERP study of the effects of within and across-modality word repetition. *Language and Cognitive Processes, 8*, 337–640.

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing II: Segmentation, semantics and syntax in non-native speakers. *Cognitive Brain Research, 15*, 214–227.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197–212.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode – The effect on lexical access. *Memory & Cognition, 12*, 380–386.

Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition, 7*, 77–85.

Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*, 88–107.

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

Van Hell, J. G., & De Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*, 431–451.

Van Petten, C., & Luka, B. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*, 176–190.

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience, 22*, 1682–1700.

Weber-Fox, C., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*, 231–256.

Weber-Fox, C., Davis, L. J., & Cuadrado, E. (2003). Event-related brain potential markers of high-language proficiency in adults. *Brain and Language, 85*, 231–244.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39*, 483–508.

Wicha, N. Y. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not Pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters, 346*, 165–168.

Wicha, N. Y. 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. *Journal of Cognitive Neuroscience, 16*, 1272–1288.

Please cite this article in press as: Martin, C. D., et al. Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.08.001