Journal of Neurolinguistics 55 (2020) 100908

Contents lists available at ScienceDirect

Journal of Neurolinguistics

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

Bilinguals processing noun morphology: Evidence for the Language Distance Hypothesis from event-related potentials

Adam Zawiszewski*, Itziar Laka

Department of Linguistics and Basque Studies, University of the Basque Country (UPV/EHU), Spain

A R T I C L E I N F O

Keywords:
Bilingualism
Non-native language processing
Language distance
ERPs
Ultimate attainment

A B S T R A C T

Evidence shows that second language (L2) processing depends on the Age of Acquisition (AoA), proficiency and differences between L1 and L2 grammar. Here we focus on the influence of the latter factor on L2 processing. To this end, we tested early (AoA = 3 years) and highly proficient Spanish-Basque and Basque-Spanish bilinguals by means of Event-Related Potentials (ERPs) while processing noun morphology in Basque (Experiments 1 and 2) and Spanish (Experiments 3 and 4). Both behavioral and electrophysiological results revealed significant differences between L1 and L2 speakers: non-natives made more errors and elicited a smaller P600 for violations than natives when processing ergative and allative morphology in Basque and accusative, dative and allative in Spanish. These findings reveal that, even for early and highly proficient bilinguals, (a) L2 processing is modulated by L1 grammar and (b) native vs. non-native differences obtain only when L1 and L2 morphological categories differ but not otherwise.

1. Introduction

Many psycholinguistic studies have asked how second language (L2) processing is modulated by the first language knowledge (L1) and the corresponding findings have been interpreted within several theoretical models (i.e. The Failed Functional Features Hypothesis: Hawkins & Chan, 1997; The Full Transfer/Full Access Hypothesis: Schwartz & Sprouse, 1996; Declarative/Procedural Model: Ullman, 2001; The Separate/Shared Syntax Account: Hartsuiker, Pickering, & Veltkamp, 2004; The Shallow Structure Hypothesis: Clahsen & Felser, 2006, among others). The aim of the current study is to shed more light on the role of this issue by investigating to what extent L2 processing is influenced by *grammatical similarities or differences* between the native and the non-native languages. We put forth the Language Distance Hypothesis (LDH): even at early onset of exposure and high proficiency in L2, native vs. non-native differences emerge in the processing of grammatical properties of L2 absent in L1, whereas no such differences are expected when processing traits of L2 that are present in L1. The concept of linguistic distance is not new and has been previously explored by other researchers, also when accounting for a third language (L3) learnability and processing (i.e. Bassetti, 2008, Cenoz, 2001, Gleitman, 1985, Kim, Qi, Feng, Ding, Liu and Cao, 2016; Schepens, van der Slik, & van Hout, 2016, among others). More specifically, our hypothesis is focused on (but not limited to) the specific typological features of L1 and L2 rather than orthographic or historical relatedness between L1 and L2. Our hypothesis is also compatible with the "shared syntax" approach (Hartsuiker et al., 2004) that suggests that grammatical traits shared by the two languages are represented once, reducing redundancy and increasing the efficiency when code-switching between languages.

* Corresponding author. University of the Basque Country (UPV/EHU), Michaela Portilla Research Center, Justo Vélez de Elorriaga, 1, 01006, Vitoria-Gasteiz, 01006, Vitoria-Gasteiz, Spain.

E-mail addresses: adam.zawiszewski@ehu.eus (A. Zawiszewski), itziar.laka@ehu.eus (I. Laka).

https://doi.org/10.1016/j.jneuroling.2020.100908
Received 5 September 2019; Accepted 14 March 2020
Available online 10 April 2020

In order to test our hypothesis we have resorted to the Event-Related Potentials' method, which has been successfully used to investigate how monolingual and bilingual speakers process language (i.e. Bornkessel-Schlesewsky & Schlesewsky, 2009; for an overview). Three components have been usually claimed to reflect different aspects of language processing: the Left Anterior Negativity (LAN), the N400 and the P600. LAN is a negative-going wave distributed over the left anterior regions of the scalp which occurs between 300 and 500 ms after stimulus onset and is interpreted as a response to processing morphosyntactic manipulations, most frequently to agreement anomalies (see Molinaro, Barber, & Carreiras, 2011; Molinaro, Barber, Caffara and Carreiras, 2015; Tanner, 2015 for an overview and discussion). The N400 component is a centro-parietally distributed negative deflection of the wave occurring between 300 and 500 ms after stimulus onset usually elicited as a response towards semantic violations (Kutas & Federmeier, 2011, for an overview), atypical thematic hierarchy (Frisch & Schlesewsky, 2001) or lexical-semantic expectations (Bornkessel-Schlesewsky & Schlesewsky, 2019). Finally, the P600 is a positive-going wave distributed mostly over centro-parietal sites starting approximately 500 ms after stimulus onset and lasting about 300 ms. It has been generally assumed to reflect revision, reanalysis or integration processes taking place when syntactically ungrammatical, ambiguous or complex information is being parsed (i.e. Regel, Meyer, & Gunter, 2014; for the detailed description and discussion or Brouwer, Crocker, Venhuizen, & Hoeks, 2017; for a neurocomputational model indexing N400 and P600 components). P600 effects have been also reported for sentences containing semantic anomalies (i.e. Kim & Osterhout, 2005) and this component has been also related to language monitoring processes such as checking upon the veracity of an (unexpected linguistic) event (Kolk, Chwilla, van Herten, & Oor, 2003). In addition, many studies have reported biphasic patterns as response to morphosyntactic anomalies: either LAN followed by a P600 (i.e. Molinaro et al., 2011, see Tanner & van Hell, 2014 for an overview) or a N400 followed by a P600 (i.e. Mueller, Hirotani, & Friederici, 2007; Zawiszewski, Gutiérrez, Fernández, & Laka, 2011). Lastly, it is important to highlight that all components mentioned above (LAN, N400 and P600) have also been reported as indexing non-linguistic processing (LANs as result of working memory load, i.e. Kluender & Kutas, 1993; N400 to musical stimuli, i.e. Daltrozzo & Schön, 2009; P600-like responses observed in the violation of abstract rules during cognitive sequencing tasks, i.e. Lelekov, Dominey, & Garcia-Larrea, 2000).

1.1. L1 effects on L2 representation and processing: ERP evidence

Here, we offer an overview of recent studies that have used ERPs to examine the impact of L1 on L2 processing seeking to determine to what extent the pattern of results may be accounted for given the LDH and, more precisely, in terms of L1 vs. L2 morphosyntactic similarities vs. differences.

Tokowicz and MacWhinney (2005) tested English natives at initial stages of learning Spanish while processing linguistic aspects that differ in the two languages (i.e. determiner number agreement) vs others that are similar (tense-marking). Results revealed a P600 component for shared linguistic properties while no effect was found for the one particular to Spanish and absent in English. However, these results might have been obscured by the poor performance of the participants (chance level), which precludes disentangling the contribution of L1 vs. L2 differences and proficiency to L2 processing.

Gillon Dowens, Vergara, Barber, and Carreiras (2010), investigated L2 morphosyntactic processing by proficient English native late learners of Spanish (AoA $>$ 20 years) with a relatively long exposure to L2 (at least 12 years). Participants were tested while processing grammatical and ungrammatical number (shared by English and Spanish) and gender agreement (absent in English, present in Spanish). Native speakers showed a LAN-P600 pattern for all violations in all conditions. A similar pattern was found in the non-native group in the within-phrase condition, but in the between-phrases agreement condition only a P600 was reported. Also, non-natives performed worse when the agreeing feature was absent in their L1 (gender) than when it was shared by L1 and L2 (number). Foucart and Frenck-Mestre (2012) also showed that gender agreement violations (a feature present in French, absent in English) elicited a P600 among native speakers, while English late advanced learners of French (AoA = 13.4 years) showed either no response, a N400 or a smaller P600. This suggests that different neural substrates are used by non-natives in comparison to natives when processing L2 features absent in L1 (but see Alemán Bañon, Fiorentino, & Gabriele (2014) who found a similar P600 pattern across all participants for both types of violations suggesting that native-like processing is attainable in adult L2 acquisition).

In another subset of studies on verb agreement, Ojima, Nakata, and Kakigi (2005) found that while subject-verb agreement violations in English elicit a LAN-P600 pattern for English natives, they only elicit a LAN in high proficiency and no effect for low proficiency native Japanese L2 speakers of English (AoA = 12 years). Chen, Shu, Liu, Zhao, and Li (2007) obtained similar results with L1Chinese-L2English speakers. These differences between native and non-native speakers can be explained by the fact that unlike English, Japanese and Chinese totally lack verb agreement.

Zawiszewski, Gutiérrez, Fernández, and Laka (2011) also investigated how the grammatical traits of L1 present and absent in L2 influence L2 processing by testing how native speakers of Basque and highly proficient early Spanish-Basque bilinguals (AoA = 3 years) process verb-agreement violations (Basque and Spanish both have subject agreement) and case morphology (Basque is ergative and Spanish is nominative). Both groups showed a similar N400–P600 response to verb-agreement violations. Case violations led to a broadly distributed negativity in both groups, followed by a P600 only in the native group. These differences strongly correlate with language distance indicating that features absent in the L1 are harder to learn (see also Díaz et al. (2016) for a similar pattern of results in early high proficiency (AoA = 4 years, C1 level) and late intermediate proficiency (AoA = 18 years, B2 level) learners of Basque).

Finally, some researchers focused on canonical and non-canonical sentence word order processing. Erdocia, Zawiszewski, and Laka (2014) for example investigated the impact of language distance factor on L2 processing by testing how highly proficient early L1Spanish-L2Basque bilinguals (AoA = 3 years) deal with canonical SOV and non-canonical OSV word orders in their L2 (Basque is SOV, Spanish is SVO). ERP measures revealed a different pattern in non-natives as compared to the P600 component reported for natives at sentence's second position, revealing that even early and highly proficient non-natives use different strategies when processing features of L2 absent L1.

In sum, the experimental evidence discussed above indicates that non-native language processing strongly correlates with the presence or absence of a given trait in the bilingual's native language, even at high proficiency and early AoA, as put forth by the LDH.

2. The present study

The goal of the present investigation is to further test the LDH by examining how L1 and L2 speakers of Basque and Spanish process morphological features present/absent in their L1/L2. To this end, we designed a study with two main conditions: (a) morphological category present in L1 and absent in L2 (b) morphological category present in L1 and L2. In the first series of experiments we tested L1 and L2 speakers of Basque while in the second series L1 and L2 speakers of Spanish were targeted.

2.1. Noun phrase morphology: differences and similarities between Basque and Spanish

Basque (spoken in Northern Spain and Southwestern France) and Spanish differ in their argument alignment/marking system: Basque (examples 1–2) is ergative, whereas Spanish (examples 3–5) is accusative (1–5):

(1)	Emakume-a-k	seme-a-ri	liburu-a	ekarri dio	parke-ra.	
	woman-the$_{SG}$-ERG	son-the$_{SG}$-DAT	book-the$_{SG}$-ABS	brought has	park-ALL	
	The woman brought the book to the son to the park.					
(2)	Emakume-a	parke-ra	etorri da.			
	woman-the$_{SG}$-ABS	park-ALL	come has.			
	The woman came to the park.					
(3)	La mujer	ha traído	al hijo	al parque.		
	The$_{SGfem}$ woman	has brought	DOM + the$_{SGmsc}$ son	ALL + the$_{SGmsc}$ park		
	The woman brought the son to the park.					
(4)	La mujer	le	ha traído	el libro	al hijo	al parque.
	The$_{SGfem}$ woman 3sing-DAT			has brought	the$_{SGmsc}$ book DAT + the$_{SGmsc}$ son ALL + the$_{SGmsc}$ park	
	The woman brought the son the book to the park.					
(5)	La mujer	ha venido	al parque.			
	The$_{SGfem}$ woman	has come	ALLthe$_{SGmsc}$- park			
	The woman came to the park.					

Importantly for our purposes, Basque and Spanish diverge regarding argument alignment and case marking. Spanish is a nominative-accusative language, like English, while Basque is an ergative–absolutive language, like Mayan or Eskimo-Aleut languages (Dixon, 1994). Thus, in Spanish, subject Noun Phrases are always unmarked (see examples 3–5). In contrast, Basque marks agentive subjects with a dedicated marker called "ergative" case (*Emakumeak* 'the woman' in 1), while non-agentive subjects (i.e. *Emakumea* 'the woman' in 2) are unmarked, like objects (*liburua* 'the book' in 1) (De Rijk, 2007; Hualde & Ortiz de Urbina, 2003). Basque and Spanish also diverge regarding complement-head direction (OV-VO): heads follow complements in Basque (i.e. *Emakume-a* 'teacher-the', *parke-ra* 'park-to', etc.) but precede complements in Spanish (i.e. *La mujer* 'the woman', *al parque* 'to the park'). Note that in Spanish, animate objects and dative and allative phrases have homophonous markers (*a*) unlike Basque, where they are distinct (transitive subject *emakumea**k***, object *emakumea*, dative NP *emakumea**ri**,* allative *parke**ra***) (see Table 1).

In sum, the great typological distance between Basque and Spanish, manifested in their contrasting morphological marking systems make Basque-Spanish and Spanish-Basque bilingual populations optimal to investigate how L2 speakers process morphological aspects of grammar present or absent in their L1, that is, to examine to what extent L2 representation and processing is shaped by the properties of L1 as predicted by the LDH.

2.2. Hypotheses and predictions

Taking into account the impact of the linguistic distance between L1 and L2 on L2 morphosyntactic processing, we expect that, given the LDH, differences between L1 and early very proficient L2 speakers in Basque will emerge when processing ergative morphology. However, we expect the same ERP pattern in both native and non-native groups when processing the dative and allative conditions, because Spanish also marks them morphologically, same as Basque. By the same reasoning, we expect to find native vs. non-native differences in Spanish for animate differential object marking (DOM), because this is a morphological feature absent in Basque. Hence, we do not expect to observe L1/L2 differences for dative and allative morphology. More specifically, given previous

Table 1
Noun Phrase morphology: differences and similarities between Basque and Spanish.

		Subject			Object		PP	
Basque	-k		Ø		Ø		-ri	-ra
Spanish		Ø		Ø		a	a	a

Table 2
Results of relative use of language and self-proficiency ratings reported by the participants in the Experiments 1 and 2. The following seven-point scale was applied for measuring the relative use of language: 1 – I speak only Basque, 2 – I speak mostly Basque, 3 – I speak Basque 75% of the time, 4 –I speak Basque and Spanish with similar frequency, 5 – I speak Spanish 75% of the time, 6 – I speak mostly Spanish, 7 – only Spanish. Proficiency level was determined by using the following four-point scale: 7 – native-like proficiency, 6 – high proficiency, 5 – full proficiency, 4 – working proficiency, 3 – limited proficiency, 2 – low proficiency, 1 – very low proficiency. Standard deviations values are in parentheses.

	L1 speakers of Basque n = 32	L2 speakers of Basque n = 30
Age	21.62 (4.82)	21.77 (4.79)
AoA of Basque	------------	3.23 (1.07)
Sex (# males)	5	8
Relative use of language		
Before primary school (0–3yrs)	1.28 (0.52)	5.97 (1.74)
Primary school (4–12 yrs)		
Home	1.06 (1.37)	5.97 (1.53)
School	1.44 (0.72)	2.33 (1.81)
Others	1.65 (1.03)	5.20 (1.30)
Secondary school (12–18 yrs)		
Home	1.25 (1.36)	6.06 (1.40)
School	1.75 (0.77)	2.20 (1.57)
Others	2.13 (1.16)	4.90 (1.36)
At time of testing		
Home	1.53 (1.48)	5.60 (1.63)
University/Work	2.09 (0.86)	2.60 (1.85)
Others	2.75 (1.25)	4.67 (1.27)
Self-rated proficiency: Basque		
Comprehension	6.84 (0.37)	6.60 (0.56)
Speaking	6.78 (0.42)	6.10 (0.88)
Reading	6.75 (0.51)	6.57 (0.57)
Writing	6.66 (0.55)	6.07 (0.69)
Self-rated proficiency: Spanish		
Comprehension	6.16 (0.99)	6.73 (0.45)
Speaking	5.69 (0.90)	6.63 (0.49)
Reading	6.13 (0.55)	6.73 (0.45)
Writing	5.69 (1.12)	6.47 (0.68)

results reported for Basque (Zawiszewski et al., 2011) we expect ergative marking violations to elicit an N400 followed by a P600 in the native group and an N400 with a significantly reduced P600 in the non-native group. Regarding dative and allative morphology, and based on previous works on noun morphology processing in Basque (Díaz, Sebastián-Gallés, Erdocia, Mueller, & Laka, 2011; Erdocia, Laka, Mestres-Missé, & Rodríguez-Fornells, 2009; Zawiszewski et al., 2011), we expect that violations will yield a P600 component, preceded by a negativity, either a LAN or a N400. In Spanish, we expect similar ERP components for morphological violations, that is, a P600 possibly preceded by negativity. But in the animate object condition the P600 should be significantly reduced (or absent) in the non-native group as compared to natives, while we expect no substantial differences between natives and non-natives for the dative and allative morphology conditions.

2.3. Experiments 1 and 2

Participants. 66 neurologically healthy speakers of Basque (students at the University of the Basque Country) participated in the experiment: 34 Basque natives (Experiment 1, 6 men, mean age 21.47 years, $SD = 4.72$) and 32 Spanish natives L2 speakers of Basque (Experiment 2, 9 men, mean age 21.71 years, $SD = 4.70$), who started acquiring Basque when they were 3 (mean AoA = 3.25 yrs, $SD = 1.11$). According to the Edinburgh Inventory for assessment of handedness (Oldfield, 1971) they were all right-handed. Data from 2 native and 2 non-native participants were excluded from the analyses because of excessive eye movements and other artifacts; consequently the results of 62 participants were submitted to statistical analyses. All participants were paid for their participation. According to the language questionnaire (see Table 2) all participants reported themselves as very skilled users of Basque (6.74 vs. 6.38 out of 7 for natives and non-natives, respectively). The study was approved by the Ethics Committee of the University of the Basque Country (UPV/EHU).

The non-native participants were highly proficient in L2 (C1 proficiency level) and reported a similar frequency of use of Basque in their everyday life as native speakers did (see Table 2).

Materials. The experiment was carried out in standard Basque (De Rijk, 2007; Salaburu, 2016). 240 experimental sentences were

Table 3
Sample of the materials used in the experiments 1 and 2.

ERGATIVE CASE	GRAM	(1) Liburua ekarri dio neska-ri goizean irakasle-ak klase-ra. Book-ABS brought girl-DAT morning teacher-ERG classroom-ALL
	UNGR	(2) Liburua ekarri dio neska-ri goizean *irakaslea klase-ra. Book-ABS brought girl-DAT morning teacher-ABS classroom-ALL 'This morning the teacher brought a book to the classroom for (to) the girl.'
DATIVE CASE	GRAM	(3) Liburua ekarri dio irakasle-ak goizean **neska-ri** klase-ra. Book-ABS brought teacher-ERG morning girl-DAT classroom-ALL
	UNGR	(4) Liburua ekarri dio irakasle-ak goizean ***neska** klase-ra. Book-ABS brought teacher-ERG morning girl-ABS classroom-ALL 'This morning the teacher brought a book to the classroom for (to) the girl.'
ALLATIVE POSTPOSITION	GRAM	(5) Liburu berria ekarri du irakasleak **klase-ra** goizean. Book new-ABS brought teacher-ERG classroom-ALL morning
	UNGR	(6) Liburu berria ekarri du irakasleak ***klasea** goizean. Book-ABS brought girl-DAT morning teacher-ABS classroom-ALL 'This morning the teacher brought a new book to the classroom.'

created and divided into six sets of sentences (1–6; 40 per condition) resulting from crossing the three different types of sentences (ergative, dative and allative) with the Grammaticality factor (Grammatical vs. Ungrammatical) (see Table 3). In (1) the subject of the sentence bears the ergative marker (-*k*), which is missing in (2) rendering it ungrammatical. In (3), the NP carries the dative marker, but it does not in (4), which is ungrammatical. Finally, in (5), the NP is headed by the allative marker, whereas in (6) it is not, yielding ungrammaticality.

The lexical material used at critical word (CW) position in all experimental conditions was matched for length (in letters) and frequency (*Euskal Hiztegiaren Maiztasun Egitura* "Frequency Structure in the Basque Dictionary" (EHME http://www.ehu.es/ehg/ehme/): The mean length of words in the ergative, dative and allative conditions was of 6.83 ± 1.9, 6.43 ± 1.6, and 6.46 ± 2.1, respectively. The mean frequency (per million) for the nouns used as CW in ergative, dative and allative conditions was of 88.17 (SDE = 15.5), 106.25 (SDE = 15.6) and 91.03 (SDE = 26.9), respectively. All comparisons between conditions were non-significant (ps > .2).

Additionally, 120 filler sentences (of different length and word orders) were added in order to make the material as diverse as possible. All blocks were counterbalanced so that each participant read only one version of each experimental sentence.

Procedure. Personal computers (Windows XP operating system) and Presentation® software (Version 16.0; www.neurobs.com) were used to present the stimuli. Before the experiment started, participants were instructed about the EEG procedure and seated comfortably in a quiet room in front of a 17 inch monitor. All sentences were displayed word-by-word in the middle of the screen for 350 ms (ISI = 250 ms). The fixation cross (+) indicated the beginning of each trial. After each sentence the questions CORRECT or INCORRECT appeared on the screen (in their corresponding Basque versions) and subjects had to press one of two keyboard buttons (left Ctrl or right Intro) depending on whether the previously displayed sentence was grammatical or not. 50% of participants used the left hand in order to respond "CORRECT" and the other 50% used the right hand. All 360 sentences were distributed over 4 blocks and after each block, which lasted approximately 10 min; subjects were given a short break (3 in total). Before the experiment began, participants ran a short training session in order to become familiar with the procedure. They were also instructed not to blink or move when the sentences were being displayed and to make the acceptability judgment as fast as possible. Each session, including the training trial, four experimental blocks, electrode-cap application and removal lasted no more than 1 h and 30 min.

EEG recording. The electroencephalogram (EEG) was recorded from 58 Ag/AgCl electrodes secured in an elastic cap (ElectroCap International, Eaton, USA). Electrodes were placed in the following sites: Fp1/2, Fz, F3A/4A, Fz, F1/2, F3/4, F5/6, F7/8, CZA, C1A/2A, C3A/4A, C5A/6A, Cz, C1/2, C3/4, C5/6, T3/4, PZA, C1P/2P, C3P/4P, TCP1/2, T3L/4L, PZ, P1/2, P3/4, P5/6, T5/6, PZP, P1P/2P, P3P/4P, CB1/2, Oz and O1/2. All recordings were referenced to the right mastoid and re-referenced off-line to the linked mastoids. Vertical eye movements and blinks were monitored by means of an electrode positioned beneath the right eye. Horizontal eye movements were monitored by an electrode positioned to the right of the right eye. Electrode impedance was kept below 5 kOhm at all scalp and mastoid sites and below 10 kOhm at the eye electrodes. The electrical signals were digitized on-line at a rate of 250 Hz by a BrainVision amplifier system and filtered off-line within a bandpass of 0.1–35 Hz. After the EEG data were recorded, the artifact rejection procedure was applied when the amplitude (from bottom to top) of the electrooculogram (EOG) was higher than 150 μV and a procedure based on independent component analysis (ICA) was used.

Data analysis. For the data analysis the critical words of grammatical sentences were compared to their ungrammatical counterparts. For ERP measures, segments were constructed from 200 ms before to the onset of the critical words in the sentences and included 1000 ms after the critical word onset trigger. The trials associated with each sentence type were averaged for each participant. The EEG 200 ms prior to the onset was also used as a baseline for all sentence type comparisons. Based on the literature and visual inspection of the data, the following temporal windows were considered during statistical analysis: 300–500 ms, 500–700 ms and 700–900 ms for all conditions. After the stimuli were recorded and averaged, ANOVAs were carried out in the following 9 five-electrode regions of interest (henceforth ROI): left anterior (F3, F5, F7, C3A, C5A), left-central (C3, C5, T3, C3P, TCP1), left-posterior (P3, P5, T5, P3P, CB1), central anterior (F1, F2, C1A, C2A), central (C1, C2, C1P, C2P, PZA) central posterior (P1, P2, P1P, P2P, PZP),

Table 4
Experiments 1 and 2. Mean reaction times in milliseconds and percentage of correct responses for both in all experimental conditions, standard deviation errors (SDE) between parentheses.

	RESPONSE TIMES IN MS				ACCURACY IN %			
	GRAM		UNGR		GRAM		UNGR	
	NAT	NNAT	NAT	NNAT	NAT	NNAT	NAT	NNAT
Ergative	1142 (104)	1264 (121)	832 (80)	1180 (115)	90.1 (0.7)	85.8 (1.2)	95.5 (0.5)	71.8 (2.1)
Dative	1093 (95)	1203 (117)	847 (92)	1074 (121)	91.3 (0.5)	88.8 (1.0)	96.3 (0.4)	88.0 (1.1)
Allative	1073 (99)	1372 (138)	951 (96)	1202 (119)	92.2 (0.5)	87.8 (0.9)	87.8 (0.9)	79.6 (1.2)

right anterior (F4, F6, F8, C4A, C6A), right-central (C4, C6, T4, C4P, TCP2) and right-posterior (P4, P6, T6, P4P, CB2). An ANOVA was performed for each of the three experimental conditions over the between-subject factor GROUP (natives and non-natives) and the three within-subjects factors: grammaticality (grammatical, ungrammatical), hemisphere (left, right) and anteriority (anterior, central and posterior). Midline (central anterior, central and central posterior) electrodes were analyzed independently. Whenever the sphericity of variance was violated, Greenhouse and Geisser (1959) correction was applied to all the data with greater than one degree of freedom in the numerator. Finally, further statistical analyses were conducted for each particular region of interest whenever appropriate. Effects for the hemisphere or region factors are only reported when they interact with the experimental manipulation. For the behavioral results, error rates and response latencies were submitted to by subject (F_1) and by item (F_2) repeated measures ANOVAs with grammaticality condition (two levels: grammatical and ungrammatical) as a within-subjects factor and with GROUP as a between-subject factor.

2.4. Behavioral results

Acceptability judgment task. Results show that non-native speakers made significantly more errors than natives in all ungrammatical conditions (ergative: F_1 (1, 60) = 20.66; $p < .001$, F_2 (1, 318) = 195.33, $p < .001$; dative: F_1 (1, 60) = 8.61, $p = .005$, F_2 (1, 318) = 39.48, $p < .001$; allative: F_1 (1, 60) = 5.05, $p = .028$, F_2 (1, 158) = 14.77, $p < .001$) and a significant effect was found also in the grammatical allative condition (by item analysis) (F_1 (1, 60) = 2.89, $p = .094$, F_2 (1, 158) = 8.65, $p = .004$). In the grammatical ergative and dative conditions native and non-native groups did not differ in the amount of errors (ergative: F_1 (1, 60) = 1.56, $p = .216$, F_2 (1, 318) = 5.65, $p = .018$; dative: F_1 (1, 60) = 0.86, $p = .358$, F_2 (1, 318) = 2.08, $p = .150$) (see Table 4).

Response times. Response times revealed that non-natives required more time than natives to perform the acceptability judgment task in the ungrammatical ergative condition (F_1 (1, 60) = 5.83, $p = .019$; F_2 (1, 318) = 100.71, $p < .001$). No differences between groups were observed in the grammatical ergative condition: F_1 (1, 60) = 0.585, $p = .447$; F_2 (1, 318) = 7.00, $p = .009$). In dative and allative conditions (grammatical and ungrammatical) no differences were found between the native and the non-native group in the analysis by subject: (dative grammatical: F_1 (1, 60) = 0.508, $p = .479$; F_2 (1, 318) = 6.98, $p = .009$; dative ungrammatical: F_1 (1, 60) = 2.37, $p = .129$; F_2 (1, 318) = 37.91, $p < .001$; allative grammatical: F_1 (1, 60) = 3.26, $p = .076$; F_2 (1, 158) = 40.92, $p < .001$; postposition ungrammatical: F_1 (1, 60) = 2.90, $p = .095$; F_2 (1, 158) = 37.91, $p < .001$).

2.5. ERP results

After baseline correction, epochs with artifacts were rejected, which resulted in the exclusion of 14.95% (SD = 5.45) of the trials in the native group (Expe 1) and 9.57% (SD = 5.14) of the trials in the non-native group (Expe 2) (see Table 5 for the detailed statistical results of the main effects and the relevant interactions).

2.5.1. Ergative condition

Between 300 and 500 ms a main effect of grammaticality was found over the lateral and midline electrodes; the negativity elicited by ungrammatical sentences was larger in comparison to that elicited by grammatical sentences in both groups. No significant interactions involving both grammaticality and group factors were found at lateral or midline electrodes in the 500–700 ms time window. Between 700 and 900 ms a significant grammaticality × anteriority × group interaction was found at lateral electrodes. Follow-up analyses (by grammaticality) showed that ungrammatical sentences elicited a significantly larger positivity than grammatical sentences over the central (F (1, 60) = 12.77; $p = .001$) and parietal (F (1, 60) = 24.49, $p < .001$) sites in the native group and over the parietal sites only in the non-native group (F (1, 60) = 4.31, $p = .042$). This positivity in the ungrammatical condition was significantly larger for the native than for the non-native group over central (F (1, 60) = 24.49, $p < .001$) and posterior sites (F (1, 60) = 8.27, $p = .006$). No effects were found at midline electrodes (see Fig. 1).

2.5.2. Dative condition

Between 300 and 500 ms the main effect of grammaticality and the grammaticality × hemisphere interaction were significant revealing that the differences between ungrammatical and grammatical sentences were more prominent over the right hemisphere. Subsequent analyses of this interaction (by grammaticality) showed that the grammaticality effect was significant in both

Table 5
Experiments 1 and 2. Statistical results. . Notes: GRAM: Sentence Type (two levels); HEM: Hemisphere (two levels) ANT: Anterior-Posterior factor (3 levels); df: degrees of freedom. ^{a} $p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

			ERGATIVE CASE		DATIVE CASE		ALLATIVE POSTPOSITION	
			Lateral	Midline	Lateral	Midline	Lateral	Midline
		df	*F*	*F*	*F*	*F*	*F*	*F*
300-500	GRAM	1, 60	**12.63*** **	**13.93*** **	**24.92*** **	**29.70*** **	0.53	0.19
	GRAM x HEM	1, 60	0.01	-----	**6.87***	-----	*3.18*^{a}	-----
	GRAM x ANT	2,120	1.82	0.81	**5.20***	**5.91***	0.32	0.21
	GRAM x HEM x ANT	2,120	0.92	-----	0.69	-----	*4.06*^{a}	-----
	GRAM x GROUP	1, 60	0.17	0.14	1.63	0.90	0.28	0.04
	GRAM x HEM x GROUP	1, 60	0.46	-----	0.33	-----	2.33	-----
	GRAM x ANT x GROUP	2, 120	0.63	0.53	0.35	0.58	0.32	0.23
	GRAM x HEM x ANT x GROUP	2, 120	0.67	-----	1.15	-----	0.36	-----
500-700	GRAM	1, 60	2.02	1.01	0.21	0.40	0.06	0.41
	GRAM x HEM	1, 60	051	-----	0.99	-----	**8.19** **	-----
	GRAM x ANT	2, 120	1.15	*2.91*^{a}	0.50	1.78	0.24	0.08
	GRAM x HEM x ANT	2, 120	1.54	-----	0.39	-----	1.89	-----
	GRAM x GROUP	1, 60	0.32	0.43	0.77	0.83	0.62	1.27
	GRAM x HEM x GROUP	1, 60	0.04	-----	1.13	-----	1.94	-----
	GRAM x ANT x GROUP	2, 120	1.49	1.20	0.45	0.20	2.38	*3.23*^{a}
	GRAM x HEM x ANT x GROUP	2, 120	0.01	-----	0.40	-----	0.15	-----
700-900	GRAM	1, 60	**13.57*** **	**9.54*** **	*3.26*^{a}	0.75	**9.09** **	**6.02** **
	GRAM x HEM	1, 60	*3.26*^{a}	-----	2.69	-----	**9.69** **	-----
	GRAM x ANT	2, 120	**17.24*** **	**15.31*** **	**22.12*** **	**25.49*** **	**17.00*** **	**16.95*** **
	GRAM x HEM x ANT	2, 120	0.23	-----	2.47	-----	**3.92***	-----
	GRAM x GROUP	1, 60	1.55	1.56	2.77	*2.99*^{a}	0.08	0.15
	GRAM x HEM x GROUP	1, 60	0.35	-----	0.02	-----	**5.13***	-----
	GRAM x ANT x GROUP	2, 120	**4.54***	***3.11***^{a}	0.05	0.21	*3.05*^{a}	**3.52***
	GRAM x HEM x ANT x GROUP	2, 120	0.21	-----	0.33	-----	0.01	-----

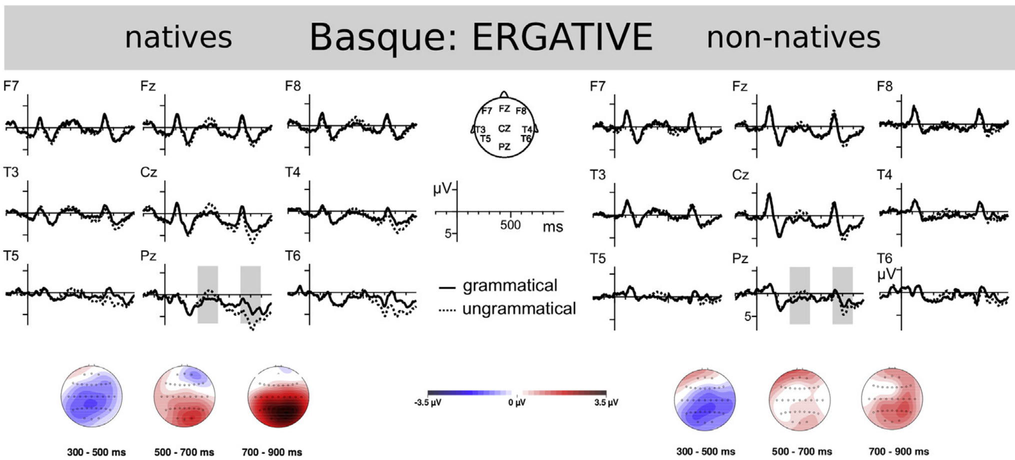

Fig. 1. ERPs elicited at the critical word position in the Ergative condition (Experiments 1–2). Red lines represent the ungrammatical stimuli and the black lines represent the grammatical stimuli. Significant differences between the grammaticality conditions are highlighted by the gray areas. Topographical amplitude difference maps for the grammaticality effect below were calculated as the average subtracting grammatical sentences from ungrammatical ones. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

hemispheres (left: $F(1, 61) = 13.01, p = .001$; right: $F(1, 61) = 27.24, p < .001$), with larger negativity for ungrammatical than grammatical sentences. The grammaticality × anteriority interaction was also significant. Follow-up analyses showed that the grammaticality effect was larger over the centro-parietal sites than over the frontal sites, but significant in all the three regions (frontal: $F(1, 61) = 8.02, p = .006$; central: $F(1, 61) = 28.05, p < .001$; parietal: $F(1, 61) = 28.80, p < .001$). The analysis of the midline electrodes revealed a main effect of grammaticality as well as a grammaticality × anteriority interaction, revealing a larger negativity for ungrammatical than grammatical dative conditions, with the grammaticality effect being larger over the centro-

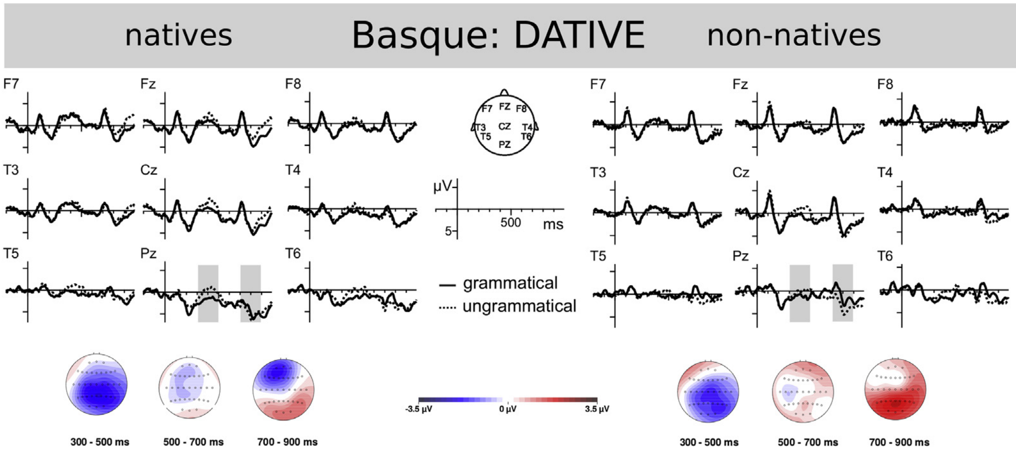

Fig. 2. ERPs elicited at the critical word position in the Dative condition (Experiments 1–2). Green lines represent the ungrammatical stimuli and the blue lines represent the grammatical stimuli. Significant differences between the grammaticality conditions are highlighted by the gray areas. Topographical amplitude difference maps for the grammaticality effect below were calculated as the average difference amplitude between the ungrammatical condition and the grammatical baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

parietal sites (frontal: F (1, 61) = 14.17, $p < .001$; central: F (1, 61) = 33.39, $p < .001$; parietal: F (1, 61) = 31.35, $p < .001$).

The analysis of the 700–900 ms time window carried out over the lateral electrodes revealed a significant grammaticality × anteriority interaction. The follow-up analyses of this interaction (by grammaticality) showed that the positivity elicited by the ungrammatical sentences in comparison to the grammatical ones was only significant over the central (marginally; F (1, 61) = 3.62, p = 0.062) and posterior electrodes; F (1, 61) = 14.41, $p < .001$). The analyses of the midline electrodes revealed a grammaticality × anteriority interaction, showing that the grammaticality effect was only significant over the posterior electrodes (F (1, 61) = 9.78, p = .003) (see Fig. 2).

2.5.3. *Allative condition*

The analyses of the 300–500 ms and 500–700 ms time windows performed over the lateral electrodes revealed a significant grammaticality × hemisphere × anteriority and a grammaticality × hemisphere interactions, respectively. Further analyses (by grammaticality) yielded no statistically significant effect and the analysis of the midline electrodes revealed no effects either. The analysis of lateral electrodes carried out within the 700–900 ms time window revealed a significant grammaticality effect as well as grammaticality × anteriority and grammaticality × hemisphere × anteriority interactions. Follow-up analyses of the grammaticality × anteriority interaction showed that the grammaticality effect was only significant over the central (F (1, 61) = 10.44, p = .002) and parietal electrodes (F (1, 61) = 22.62, $p < .001$). The analyses of the 3-way interaction revealed that the grammaticality effect was only significant over the left-posterior (F (1, 61) = 17.73, $p < .001$), right-central (F (1, 61) = 15.76, $p < .001$) and right-posterior electrodes; (F (1, 61) = 25.30, $p < .001$). Additionally, there was also a grammaticality × hemisphere × group interaction. Subsequent (by grammaticality) analyses of the this interaction revealed that the grammaticality effects were significant over the left (F (1, 60) = 4.40, p = .040) and the right hemisphere (F (1, 60) = 5.34, p = .024) in the native group and only significant over the right hemisphere in the non-native group (F (1, 60) = 8.94, p = .004). The analysis by group revealed that the positivity elicited in the grammatical condition over the right hemisphere was larger in natives than in non-natives (F (1, 60) = 4.42, p = .04). The analysis of the midline electrodes revealed a significant grammaticality effect and a grammaticality × anteriority interaction, revealing that ungrammatical sentences elicited a larger positivity than grammatical ones, and this grammatical effect was larger over the posterior (F (1, 61) = 16.82, $p < .001$) than over central sites (F (1, 61) = 5.95, p = .018. Also, there was a significant grammaticality × anteriority × group interaction. Follow-up analyses revealed that the grammaticality effect was larger over the central (F (1, 60) = 4.86, p = .031) and posterior electrodes (F (1, 60) = 4.86, p = .031) in the native group and over the posterior electrodes in the non-natives (F (1, 60) = 4.60, p = .036) (see Fig. 3).

2.6. *Summary of results from experiments 1 and 2*

Behavioral results showed native speakers to be more accurate than non-natives during the acceptability judgment task in the ungrammatical conditions. Natives were also significantly faster than non-natives to respond in the ungrammatical ergative condition.

Regarding electrophysiological results, in the ungrammatical ergative condition native speakers displayed a broadly distributed negativity, roughly corresponding to an N400, followed by a robust P600 component. L2 speakers showed a similar negativity, but a significantly smaller P600. The ungrammatical dative condition elicited similar N400–P600 ERP patterns in both L1 and L2 groups. The ungrammatical allative condition led to a P600 in both groups of speakers and this positivity was larger in the native group.

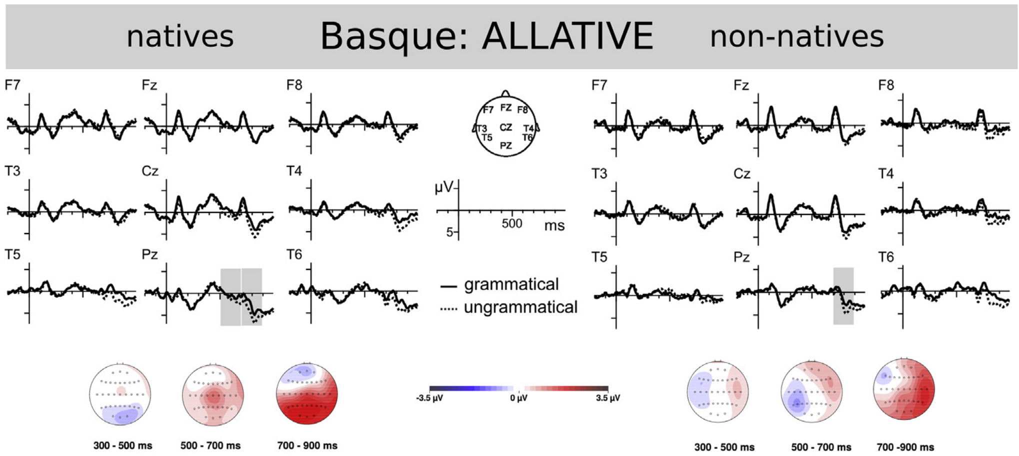

Fig. 3. ERPs elicited at the critical word position in the Allative condition (Experiments 1–2). Brown lines represent the ungrammatical stimuli and the magenta lines represent the grammatical stimuli. Significant differences between the grammaticality conditions are highlighted by the gray areas. Topographical amplitude difference maps for the grammaticality effect below were calculated as the average difference amplitude between the ungrammatical condition and the grammatical baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.7. ERP experiments 3 and 4

Participants. 56 neurologically healthy speakers of Spanish (undergraduates and graduates at the University of the Basque Country) participated in the experiment: 28 natives (Experiment 3, 5 men, mean age 19.8 years, $SD = 1.7$) and 28 L2 speakers of Spanish natives of Basque (Experiment 4, 10 men, mean age 20.7 years, $SD = 1.1$), who started acquiring Spanish at 5 (mean AoA = 5.3 yrs, $SD = 2.2$). According to Edinburgh Inventory for assessment of handedness (Oldfield, 1971) they were all right-handed. Data of one L2 participant were excluded from the analyses because of excessive eye movements and other artifacts. Consequently, the results of 55 speakers were submitted to the statistical analyses. All participants were paid for their participation. According to the language questionnaire (see Table 6) all participants reported themselves as very skilled users of Spanish (6.3 (1.2) vs. 6.0 (1.5) out of 7) for natives and non-natives, respectively). The study was approved by the Ethics Committee of the University of the Basque Country (UPV/EHU).

Materials. The experiment was carried out in standard Spanish (Bosque & Demonte, 1999). As in Experiments 1 and 2, 240 experimental sentences were created and divided in six sets (1–6, 40 per condition) (see Table 7). In (1) the animate direct object bears a differential marker *a* "to", contracted with the determiner (a + el, *al*), which is missing in (2) yielding ungrammaticality (Leonetti, 2004). In (3) the indirect object is marked (*a*), but in (4) the dative marker is missing and hence the sentence is ungrammatical. Finally, in (5) the prepositional phrase bears the corresponding allative preposition (*a*) whereas in (6) it does not, yielding ungrammaticality.

The lexical material used at critical word (CW) position in all experimental conditions was matched for length and frequency (Guasch, Boada, Ferré, & Sánchez-Casas, 2013). The mean length of words in DOM, dative and preposition condition was of 7.71 ± 1.8, 7.93 ± 1.9, and 6.8 ± 2.1, respectively. The mean frequency (per million) for the nouns used as CW in DOM, dative and allative preposition conditions was of 63.05n (SDE = 10.8), 60.07 (SDE = 10.6) and 75.05 (SDE = 19.8), respectively. Only DOM vs. allative preposition (t (79) = 2.94, $p < .01$) and dative vs. allative preposition (t (79) = 3.68, $p < .001$) length comparisons resulted significant (all other ps > .4).

Additionally, 120 filler sentences (of different length and word orders) were added in order to make the material as diversified as possible. All experimental blocks were counterbalanced so that each participant read only one version of each sentence.

Procedure. Same as in the Experiments 1 and 2.

EEG recording. The EEG was recorded from 32 active electrodes secured in an elastic cap (Acticap System, Brain Products, Gilching, Germany). Electrodes were placed on standard positions according to extended Internationals 10–20 system in the following sites: Fp1/Fp2, Fz, F3/F4, F7/F8, FC5/FC6, FC1/FC2, T7/T8, C3/C4, Cz, CP5/CP6, CP1/CP2, P7/P8, P3/P4, Pz, PO9/PO10, O1/02, and Oz. All recordings were referenced to the right mastoid and re-referenced off-line to the linked mastoids. Vertical and horizontal eye movements and blinks were monitored by means of two electrodes positioned beneath and to the right of the right eye. Electrode impedance was kept below 10 kOhm at all scalp and eye electrodes. The electrical signals were digitized online at a rate of 500 Hz by a BrainVision amplifier system and filtered offline within a bandpass of 0.1–35 Hz. After the EEG data were recorded, the artifact rejection procedure was applied (*off-line*) when the amplitude (from bottom to top) of the electrooculogram (EOG) was higher than 150 μV and a procedure based on independent component analysis (ICA) was used.

Data analysis. For the data analysis the two critical words of grammatical sentences were compared separately to their ungrammatical counterparts (see Table 7). For ERP measures, segments were constructed from 200 ms before to the onset of the critical words in the sentences and included 1000 ms after the critical word onset trigger. The trials associated with each sentence type were

Table 6
Results of relative use of language and self-proficiency ratings reported by the participants in the Experiments 3 and 4. The following seven-point scale was applied for measuring the relative use of language: 1 – I speak only Basque, 2 – I speak mostly Basque, 3 – I speak Basque 75% of the time, 4 –I speak Basque and Spanish with similar frequency, 5 – I speak Spanish 75% of the time, 6 – I speak mostly Spanish, 7 – only Spanish. Proficiency level was determined by using the following four-point scale: 7 – native-like proficiency, 6 – high proficiency, 5 – full proficiency, 4 – working proficiency, 3 – limited proficiency, 2 – low proficiency, 1 – very low proficiency. Standard deviations values are in parentheses.

	L1 speakers of Spanish n = 28	L2 speakers of Spanish n = 27
Age	19.8 (1.7)	20.6 (1.1)
AoA of Spanish	------------	4.9 (2.2)
Sex (# males)	5	9
Relative use of language		
Before primary school (0–3yrs)	1.50 (0.8)	5.7 (2.21)
Primary school (4–12 yrs)		
Home	1.3 (0.5)	6.3 (1.5)
School	4.5 (1.7)	6.2 (1.9)
Others	1.9 (0.7)	5.5 (1.9)
Secondary school (12–18 yrs)		
Home	1.3 (0.5)	5.9 (2.1)
School	4.2 (1.7)	5.7 (2.0)
Others	2.0 (1.0)	5.0 (1.9)
At time of testing		
Home	1.3 (0.5)	5.6 (1.9)
University/Work	2.0 (2.0)	5.5 (1.9)
Others	2.2 (0.9)	4.8 (1.9)
Self-rated proficiency: Spanish		
Comprehension	6.6 (0.5)	5.9 (1.9)
Speaking	6.5 (0.5)	5.7 (1.9)
Reading	6.1 (0.7)	5.9 (1.9)
Writing	5.9 (1.0)	5.6 (2.0)
Self-rated proficiency: Basque		
Comprehension	5.8 (0.7)	6.2 (1.8)
Speaking	5.1 (0.9)	6.3 (1.9)
Reading	5.8 (0.7)	6.2 (1.9)
Writing	5.1 (1.1)	6.0 (2.0)

Table 7
Sample of the materials used in the experiments 3 and 4.

ACCUSATIVE CASE	GRAM	(1) La reina criticó al ministro con dureza. The queen criticized the-ACC minister with strength
	UNGR	(2) La reina criticó *el ministro con dureza. The queen criticized the-NOM minister with strength 'The queen strongly criticized the minister.'
DATIVE CASE	GRAM	(3) Le saqué una foto **al entrenador** del Athletic. Him took (I) a picture the-DAT manager of Athletic
	UNGR	(4) Le saqué una foto ***el entrenador** del Athletic. Him took (I) a picture the-NOM manager of Athletic 'I took a picture of the manager of Athletic.'
ALLATIVE PREPOSITION	GRAM	(5) Los niños llegaron **al parque** muy cansados. The children arrived the-ALL park very tired
	UNGR	(6) Los niños llegaron ***el parque** muy cansados. The children arrived the-NOM park very tired 'The children arrived at the park very tired.'

averaged for each participant. The EEG 200 ms prior to the onset was also used as a baseline for all sentence type comparisons. Based on the literature and visual inspection of the data, the following temporal windows were considered during statistical analysis: 300–500 ms, 500–700 ms and 700–900 ms for both words in all conditions. After the stimuli were recorded and averaged, ANOVAs were carried out in the following 9 two-electrode ROIs: left anterior (F7, F3), left-central (T7, C3), left-posterior (P7, P3), right anterior (F4, F8), right-central (C4, T8) and right-posterior region (P4, P8). The same type of ANOVA analyses as in Experiments 1 and 2 were performed.

2.8. Behavioral results

Grammaticality judgment task. Results showed that non-native speakers made significantly more errors than natives in the ungrammatical DOM condition (F_1 (1, 53) = 24.47; $p < .001$, F_2 (1, 158) = 115.67, $p < .001$) and a marginally significant effect was also found in the ungrammatical allative preposition condition (F_1 (1, 53) = 3.88; $p = .054$, F_2 (1, 158) = 16.17, $p < .001$) while no significant differences were found in the grammatical DOM (F_1 (1, 53) = 0.68; $p = .415$, F_2 (1, 158) = 2.79, $p = .101$), grammatical dative (F_1 (1, 53) = 0.36; $p = .549$, F_2 (1, 158) = 0.69, $p = .407$), ungrammatical dative (F_1 (1, 53) = 1.66; $p = .203$, F_2 (1, 158) = 38.98, $p < .001$) and grammatical allative preposition conditions (F_1 (1, 53) = 0.90; $p = .347$, F_2 (1, 158) = 1.33, $p = .251$).

Response times. Results from response times (by subject analysis) showed that non-natives performed the acceptability judgment task faster than natives in the grammatical allative preposition condition (F_1 (1, 53) = 4.89, $p = .031$; F_2 (1, 158) = 2.11, $p = .148$), while no differences emerged in the grammatical DOM condition (F_1 (1, 53) = 2.80, $p = .100$; F_2 (1, 158) = 20.53, $p < .001$) and in the grammatical dative condition (F_1 (1, 53) = 3.17, $p = .081$; F_2 (1, 158) = 18.89, $p < .001$). Regarding the ungrammatical conditions no differences were found (by subject analysis) in the DOM (F_1 (1, 53) = 0.36, $p = .550$; F_2 (1, 158) = 2.11, $p = .148$), in the dative condition (by subject analysis) (F_1 (1, 53) = 1.50, $p = .226$; F_2 (1, 158) = 11.19, $p = .001$) or in the allative preposition condition (by subject analysis) (F_1 (1, 53) = 3.07, $p = .086$; F_2 (1, 158) = 24.56, $p < .001$) (see Table 8).

2.9. ERP results

For natives, 4.91% (SD = 2.80) and 4.94% (SD = 2.90) of the trials were rejected due to artifacts in the first and second critical words, respectively. For non-natives, 6.25% (SD = 3.52) and 6.17% (SD = 3.50) of the trials were rejected (see Tables 9 and 10 for the detailed statistical results of the main effects and the relevant interactions).

2.9.1. Accusative case (animate object)

At the first CW position analyses carried out over the lateral electrodes within the 300–500 ms time window revealed a significant grammaticality effect, that is, ungrammatical sentences elicited a larger negativity than their grammatical counterparts. The analysis of the midline electrodes showed a marginally significant grammaticality × anteriority interaction, revealing a larger negativity for ungrammatical sentences than for grammatical ones only over the central site of the scalp (F (1, 54) = 5.25, $p = .026$). Regarding the second CW, analyses of the lateral electrodes showed a significant grammaticality × anteriority interaction, indicating that the grammaticality effect (larger positivity for ungrammatical than grammatical sentences) was only present over frontal electrodes (F (1, 54) = 6.55, $p = .013$). No effects were found at midline electrodes. Between 500 and 700 ms no significant results obtained over the lateral electrodes at the first CW. The analysis of the midline electrodes revealed a significant grammaticality × anteriority interaction. The follow-up analysis of this interaction showed a larger positivity for the ungrammatical DOM condition as compared to the grammatical one over the posterior electrodes (F (1, 54) = 6.40, $p = .014$). At the second CW, there was a significant grammaticality effect and a grammaticality by hemisphere by anteriority interaction over the lateral electrodes, revealing that the grammaticality effect (larger positivity for ungrammatical than grammatical sentences) was only significant over the left-posterior (F (1, 54) = 4.38, $p = .041$), the right-frontal (F (1, 54) = 13.22, $p = .001$) and right-central electrodes (F (1, 54) = 12.95, $p = .001$). Importantly, there was a grammaticality × group interaction: while native speakers displayed a significant grammaticality effect (F (1, 53) = 11.51, $p = .001$), the effect was not significant for non-natives.. The analysis of the midline electrodes showed a main grammaticality effect indicating a larger positivity for ungrammatical sentences than for grammatical ones. Regarding the 700–900 ms time window, no significant effects were found at the first CW. At the second CW there was a grammaticality effect, as well as grammaticality × anteriority and grammaticality × anteriority × hemisphere interactions over the lateral electrodes. The former interaction revealed that the main grammaticality effect (larger positivity elicited by ungrammatical than grammatical sentences) mostly prominent over the posterior electrodes (F (1, 54) = 33.49, $p < .001$). The triple interaction showed that grammaticality effects were significant at all sites but the left-frontal one (p values $< .01$)). The analysis of the midline electrodes showed a main effect of grammaticality and a grammaticality × anteriority interaction, revealing that the grammaticality effect was

Table 8
Experiment 3 and 4. Mean reaction times in milliseconds and percentage of correct responses for both in all experimental conditions, standard deviation errors (SDE) between parentheses.

	RESPONSE TIMES IN MS				ACCURACY IN %			
	GRAM		UNGR		GRAM		UNGR	
	NAT	NNAT	NAT	NNAT	NAT	NNAT	NAT	NNAT
Accusative	1224 (106)	1014 (65)	1135 (92)	1067 (64)	93.7 (1.4)	95.2 (1.2)	88.8 (1.5)	65.0 (4.6)
Dative	1258 (100)	1048 (59)	1029 (103)	877 (65)	89.5 (1.7)	90.7 (1.3)	96.2 (0.9)	93.8 (1.6)
Prepositions	1220 (89)	975 (65)	1044 (85)	862 (57)	92.5 (1.3)	94.2 (1.1)	92.5 (1.1)	88.1 (2.0)

Table 9
Experiments 3 and 4. Statistical results Notes: GR: Sentence Type (two levels); H: Hemisphere (two levels) ANT: Anterior-Posterior factor (3 levels); df: degrees of freedom. $^{a}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

			ACCUSATIVE CASE		DATIVE CASE		ALLATIVE PREPOSITION	
			Lateral	Midline	Lateral	Midline	Lateral	Midline
		df	*F*	*F*	*F*	*F*	*F*	*F*
300-500	GR	1, 53	**4.88***	1.81	2.64	1.37	**4.31***	**16.33*****
	GR x H	1, 53	0.02	----	1.53	----	0.25	----
	GR x ANT	2, 106	0.06	*2.73*[a]	0.24	2.42	**3.81***	1.15
	GR x H x ANT	2, 106	0.01	----	**4.65***	----	2.33	----
	GR x GROUP	1, 53	0.06	0.01	0.95	1.51	0.43	1.18
	GR x H x GROUP	1, 53	1.20	----	0.08	----	*3.25*[a]	----
	GR x ANT x GROUP	2, 106	0.15	0.03	0.18	0.02	0.25	0.69
	GR x H x ANT x GROUP	2, 106	0.78	----	2.21	----	0.32	----
500-700	GR	1, 53	0.18	0.61	2.12	**5.43***	1.57	0.01
	GR x H	1, 53	1.63	----	0.12	----	0.02	----
	GR x ANT	2, 106	1.26	**6.53****	0.37	**4.65***	1.49	0.76
	GR x H x ANT	2, 106	0.25	----	1.99	----	**4.17***	----
	GR x GROUP	1, 53	0.03	0.01	0.01	0.11	0.02	2.50
	GR x H x GROUP	1, 53	0.39	----	0.02	----	0.40	----
	GR x ANT x GROUP	2, 106	0.58	0.66	0.97	0.53	0.50	0.12
	GR x H x ANT x GROUP	2, 106	0.60	----	1.11	----	0.21	----
700-900	GR	1, 53	0.36	0.57	0.03	1.13	0.01	0.87
	GR x H	1, 53	0.36	----	1.55	----	*3.21*[a]	----
	GR x ANT	2, 108	0.43	1.83	0.77	**11.56*****	0.28	*3.13*[a]
	GR x H x ANT	2, 108	0.63	----	1.45	----	0.25	----
	GR x GROUP	1, 53	0.71	0.11	0.07	0.13	0.02	1.34
	GR x H x GROUP	1, 53	1.61	----	0.42	----	1.54	----
	GR x ANT x GROUP	2, 108	0.13	1.17	0.01	0.31	0.60	0.02
	GR x H x ANT x GROUP	2, 108	0.57	----	0.78	----	0.03	----

most prominent over central (F (1, 54) = 25.78, $p < .001$) and parietal electrodes (F (1, 54) = 28.27, $p < .001$) (see Fig. 4).

2.9.2. Dative condition

Between 300 and 500 ms, at the first CW (lateral electrodes) there was a grammaticality x × hemisphere × anteriority interaction. Further analyses showed that the negativity elicited by the ungrammatical dative condition was larger than that elicited by the grammatical dative condition and this difference was most prominent over the left posterior electrodes (F (1, 54) = 9.12; p = .004). No significant effects were found over the midline electrodes. Regarding the second CW (lateral electrodes), there were grammaticality × hemisphere and grammaticality × hemisphere × anteriority interactions. The former interaction revealed that grammaticality effects (larger positivity for ungrammatical than grammatical sentences) were only significant over the right hemisphere (F (1, 54) = 5.82, p = .019). The triple interaction showed that grammaticality effects were only significant over the left-posterior (F (1, 54) = 6.93, p = .011) and over the right-frontal (marginally; F (1, 54) = 3.51, p = .066) and right-central electrodes (F (1, 54) = 9.38, p = .003). Importantly, there was a grammaticality × group interaction, revealing that grammaticality effects were significant for natives (F (1, 53) = 10.20, p = .002), but not for non-natives. At midline electrodes, a grammaticality effect and a grammaticality × anteriority interaction, revealed that ungrammatical sentences elicited a larger positivity than grammatical ones only over central (marginally, F (1, 54) = 3.97, p = .057) and parietal electrodes (F (1, 54) = 8.24, p = .006). Additionally, the grammaticality × group interaction showed that grammaticality effects were significant for natives (F (1, 53) = 16.46, $p < .001$) but not for non-natives. Regarding the 500–700 ms time window, no significant results were found over the lateral electrodes at the first CW. The analyses of the midline electrodes revealed a main effect of grammaticality, that is, the ungrammatical sentences elicited larger positivity than the grammatical ones, and a grammaticality × anteriority interaction. Subsequent analyses of the latter interaction demonstrated that ungrammatical sentences yielded larger positivity than grammatical ones over the central (F (1, 54) = 6.45, p = .014) and posterior electrodes (F (1, 54) = 10.46, p = .002). At the second CW (lateral electrodes), there was a main effect of grammaticality, as well as grammaticality × hemisphere, grammaticality × anteriority and grammaticality × hemisphere × anteriority interactions. Follow-up analysis showed that grammaticality effects (larger positivity for ungrammatical than grammatical sentences) were only significant over the right hemisphere (F (1, 54) = 14.22, $p < .001$) and over posterior electrodes (F (1, 54) = 34.12, $p < .001$). Additionally, the triple interaction demonstrated that the grammaticality effect was most prominent over the left-posterior (F (1, 54) = 29.44, $p < .001$), right-central (F (1, 54) = 29.44, p < .001) and right-posterior electrodes (F (1, 54) = 27.39, $p < .001$). At midline electrodes, there was a main effect of grammaticality and a grammaticality × anteriority interaction, revealing that grammaticality effects were only significant over the central (F (1, 54) = 9.70, p = .003) and parietal electrodes (F (1, 54) = 24.85, $p < .001$). No significant effects were found at the first CW (lateral electrodes) between 700 and

Table 10
Experiments 3 and 4. Statistical results, second word. Notes: GRAM: Sentence Type (two levels); HEM: Hemisphere (two levels) ANT: Anterior-Posterior factor (3 levels); df: degrees of freedom. [a] $p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

			ACCUSATIVE CASE		DATIVE CASE		ALLATIVE PREPOSITION	
			Lateral	Midline	Lateral	Midline	Lateral	Midline
		df	*F*	*F*	*F*	*F*	*F*	*F*
300-500	GRAM	1, 53	*3.55*[a]	0.95	2.15	**4.79***	0.04	0.58
	GRAM x HEM	1, 53	0.05	----	**4.42***	----	*3.10*[a]	----
	GRAM x ANT	2, 106	**4.48***	*1.95*	0.79	**3.92***	0.56	0.67
	GRAM x HEM x ANT	2, 106	1.56	----	**14.83*****	----	**10.77*****	----
	GRAM x GROUP	1, 53	0.07	0.02	**9.01****	**12.23****	*3.43*[a]	2.19
	GRAM x HEM x GROUP	1, 53	0.07	----	0.11	----	1.22	----
	GRAM x ANT x GROUP	2, 106	0.36	0.09	0.22	0.51	*3.62*[a]	0.87
	GRAM x HEM x ANT x GROUP	2, 106	1.69	----	0.78	----	0.03	----
500-700	GRAM	1, 53	**7.48****	**7.40****	**5.11***	**9.31****	**10.09****	**16.53*****
	GRAM x HEM	1, 53	*3.31*[a]	----	**12.34****	----	**6.52***	----
	GRAM x ANT	2, 106	0.93	0.92	**26.48*****	**19.59*****	**9.54****	**7.51****
	GRAM x HEM x ANT	2, 106	**5.35***	----	**18.34*****	----	**12.32*****	----
	GRAM x GROUP	1, 53	**4.08***	*3.12*[a]	*3.37*[a]	2.55	**6.91***	2.65
	GRAM x HEM x GROUP	1, 53	1.94	----	0.01	----	0.27	----
	GRAM x ANT x GROUP	2, 106	0.77	0.08	0.16	0.10	1.02	0.28
	GRAM x HEM x ANT x GROUP	2, 106	0.52	----	0.33	----	1.05	----
700-900	GRAM	1, 53	**21.10*****	**23.73*****	0.09	0.59	**7.32****	**10.45****
	GRAM x HEM	1, 53	1.98	----	2.55	----	**4.32***	----
	GRAM x ANT	2, 108	**7.21****	**8.69****	**38.64*****	**20.81*****	**40.74*****	**21.34*****
	GRAM x HEM x ANT	2, 108	**9.93****	----	**15.01*****	----	18.69	----
	GRAM x GROUP	1, 53	2.25	1.29	2.36	1.76	*3.59*[a]	1.24
	GRAM x HEM x GROUP	1, 53	0.32	----	0.10	----	0.01	----
	GRAM x ANT x GROUP	2, 108	0.41	0.02	0.05	0.09	0.34	0.28
	GRAM x HEM x ANT x GROUP	2, 108	0.21	----	0.30	----	1.31	----

900 ms. The analysis of the midline electrodes showed a significant grammaticality × anteriority interaction. Subsequent comparisons indicated that ungrammatical sentences elicited larger positivity than grammatical ones in both groups of speakers over the posterior sites of the scalp (F (1, 54) = 13.39, p = .001). The analyses conducted at the second CW over the lateral electrodes revealed the grammaticality × anteriority and the grammaticality × hemisphere × anteriority interactions, showing a significantly larger negativity for ungrammatical sentences in comparison to grammatical ones over frontal locations (F (1, 54) = 8.11, p = .006) and a significantly larger positivity over posterior electrodes (F (1, 54) = 22.96, p < .001). The triple interaction indicated that ungrammatical sentences yielded a larger negativity than grammatical ones over left-frontal electrodes (F (1, 54) = 13.49, p = .001), and they also yielded a larger positivity over the left-posterior (left-posterior: F (1, 54) = 31.92, p < .001) and right-posterior electrodes (F (1, 54) = 10.28, p = .002) (see Fig. 5).

2.9.3. Allative preposition condition

In the 300–500 ms time window at the first CW (lateral electrodes), the analyses revealed a significant grammaticality effect indicating that ungrammatical sentences elicited a larger negativity than grammatical counterparts. Also, there was a marginally significant grammaticality × hemisphere × group interaction and a significant grammaticality × anteriority interaction. Further analyses of the former (by grammaticality) revealed that ungrammatical sentences elicited a larger negativity over the left hemisphere in the native group (F (1, 53) = 4.76, p = .034) while no significant differences were found in the non-natives. Finally, further analyses of the grammaticality × anteriority interaction showed that the ungrammatical sentences yielded larger negativity than the grammatical ones over the parietal electrodes (F (1, 54) = 15.07, p < .001). The analysis of the midline electrodes revealed a main grammaticality effect, that is, ungrammatical sentences elicited larger negativity than grammatical ones over the anterior, central and posterior sites. Regarding the second CW (lateral electrodes), there was a grammaticality × hemisphere × anteriority interaction. The follow-up analyses of this interaction (by grammaticality) showed no statistically significant results (all p's > .12). No effects were found at midline electrodes either. Within the 500–700 ms time window, at the first CW (lateral electrodes) a significant grammaticality × hemisphere × anteriority interaction emerged; however, further comparisons (by grammaticality) yielded no significant results (all p's > 0.10). The analysis of midline electrodes yielded no significant results either At the second CW (lateral electrodes), there was a main grammaticality effect, as well as grammaticality × hemisphere, grammaticality × anteriority and grammaticality × hemisphere × anteriority interactions. These 2-way interactions revealed that the grammaticality effect (larger positivity for ungrammatical than grammatical conditions) was only significant over the right hemisphere (F (1, 54) = 17.77, p < .001) and over the central (F (1, 54) = 7.84, p = .007) and posterior electrodes (F (1, 54) = 46.47, p < .001). The three-way interaction indicated that the grammaticality effect was only significant over the left-posterior (F (1, 54) = 32.65, p < .001), right-

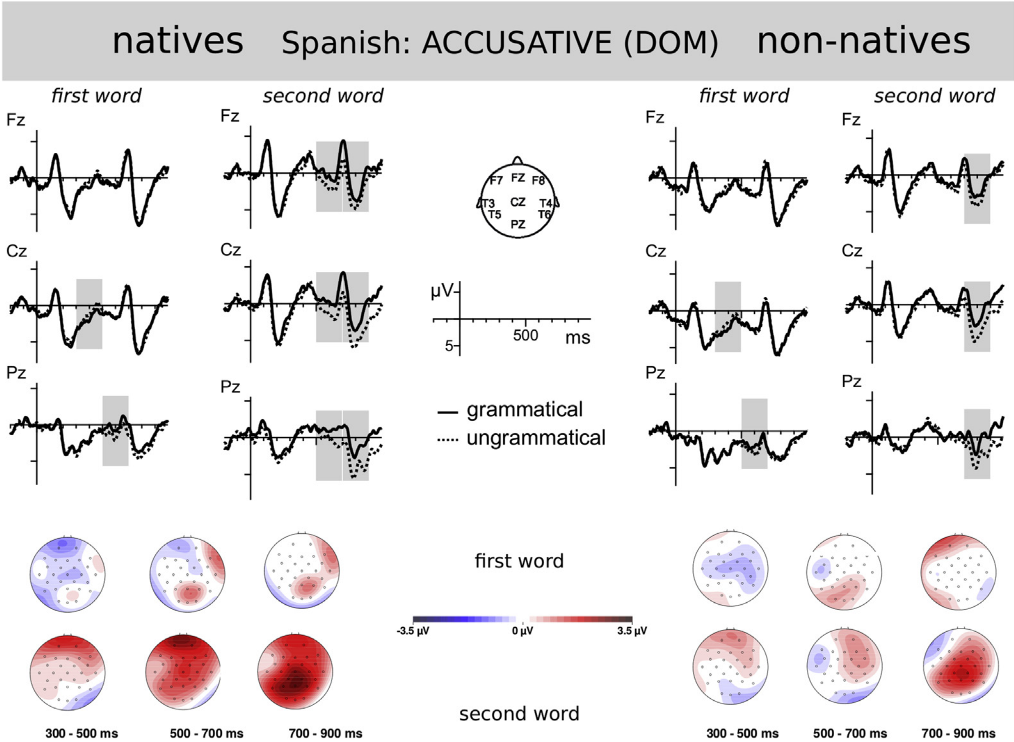

Fig. 4. ERPs elicited at the critical word positions in the Accusative condition (Experiments 3–4). Red lines represent the ungrammatical stimuli and the black lines represent the grammatical stimuli. Significant differences between the grammaticality conditions are highlighted by the gray areas. Topographical amplitude difference maps for the grammaticality effect below were calculated as the average difference amplitude between the ungrammatical condition and the grammatical baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

central (F (1, 54) = 17.31, p < .001) and right-posterior regions (F (1, 54) = 35.88, p < .001). Importantly, there was also a grammaticality × group interaction, revealing that the grammaticality effect was significant only for native speakers (F (1, 53) = 17.16, p < .001). At midline electrodes, there was a grammaticality effect, and a grammaticality × anteriority interaction, revealing that grammaticality effects were only significant over central (F (1, 54) = 17.31, p < .001) and posterior electrodes (F (1, 54) = 35.57, p < .001). Between 700 and 900 ms, at the first CW (lateral electrodes) analyses revealed a marginal grammaticality × hemisphere interaction. Further analyses (by grammaticality) showed no differences between grammatical and ungrammatical conditions (all p's > 0.39). The statistical tests conducted over the midline electrodes revealed a marginally significant grammaticality × anteriority interaction; but subsequent analyses yielded no grammaticality effect over any site (all p's > . 18). Regarding the second CW (lateral electrodes), there was a grammaticality effect, as well as grammaticality × hemisphere and grammaticality × hemisphere × anteriority interactions. Follow-up analyses revealed that the grammaticality effects (larger positivity for grammatical than ungrammatical conditions) were only significant over the right hemisphere (F (1, 54) = 12.13, p = .001). The 3-way interaction revealed that while the ungrammatical sentences elicited larger negativity than grammatical sentences over the left anterior region (F (1, 54) = 8.89, p = .004), they elicited a larger positivity over the left-posterior (F (1, 54) = 52.59, p < .001) and right-central (F (1, 54) = 11.51, p = .001) and right-posterior regions (F (1, 54) = 40.68, p < .001). At midline electrodes, there was a grammaticality effect and a grammaticality × anteriority interaction, revealing that ungrammatical sentences elicited larger positivity than grammatical ones only over the central (F (1, 54) = 12.32, p = .001) and posterior regions (F (1, 54) = 28.23, p < .001) (see Fig. 6).

2.10. Summary of results of experiments 3 and 4

Behavioral results showed that native speakers were more accurate than non-natives detecting the ungrammaticality in the DOM and allative preposition conditions. Natives were also slower than non-natives when judging acceptability in the grammatical allative condition.

Regarding electrophysiological results, the pattern found at the first CW (*al* vs. **el*) revealed no differences between natives and non-natives: A similar N400 – P600 pattern was found for violations in the DOM and dative conditions and an N400 for violations in the allative condition. At the second critical word (Noun), overall, DOM violations elicited a larger positivity than grammatical sentences across all time windows. This positivity was significantly larger (500–700 ms) in the native group. A similar picture emerged in the dative condition: violations yielded a long-lasting positive component, larger in natives than in non-natives

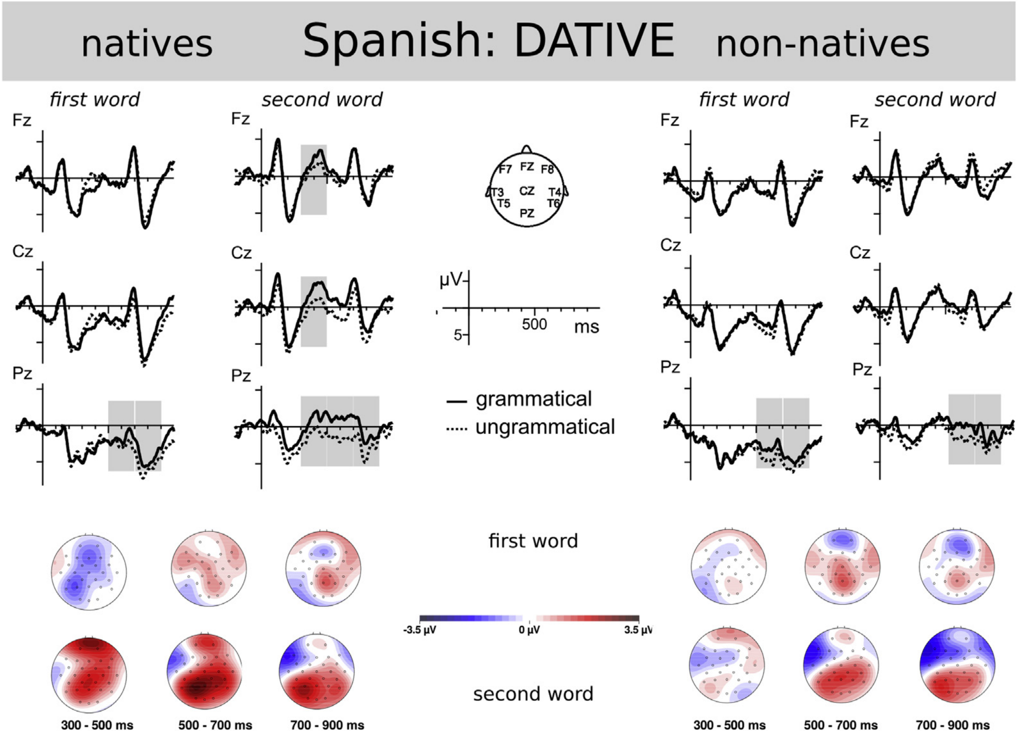

Fig. 5. ERPs elicited at the critical word positions in the Dative condition (Experiments 3–4). Green lines represent the ungrammatical stimuli and the blue lines represent the grammatical stimuli. Significant differences between the grammaticality conditions are highlighted by the gray areas. Topographical amplitude difference maps for the grammaticality effect below were calculated as the average difference amplitude between the ungrammatical condition and the grammatical baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(300–500 ms). In addition, this positivity was accompanied by negativity over the fronto-lateral sites at late stages of processing (700–900 ms). Finally, native vs. non-native differences also obtained for the allative condition: In the ungrammatical condition, the positivity was larger for the native group than for the non-native (500 and 700 ms). Additionally, at late stages of processing (700–900 ms) the positivity was accompanied by a negative component over the left-frontal electrodes in all participants.

3. General discussion

The aim of the present study was to investigate whether highly proficient and early (AoA = 3) non-native speakers differ in their representation and processing of L2 grammatical traits absent in their native language. We studied NP morphology in Basque and Spanish. We first compared how native and non-native speakers process (a) ergative marked NPs, (b) dative marked NPs and (c) allative PPs in Basque (Exp. 1 and 2). Importantly, Basque and Spanish differ regarding condition (a) but do not regarding (b) and (c). Then we turned to native vs. non-native speakers of Spanish (Exp. 3 and 4), and explored the processing of (d) Differential Object Marking (DOM), (e) dative NP marking and (f) allative PPs. These conditions are comparable to those studied in Basque: Basque does not have (d), but has (e) and (f).

Based on previous research (Díaz et al., 2016; Zawiszewski et al., 2011), we put forth the Language Distance Hypothesis, and we hypothesized that the linguistic distance between L1 and L2 (that is, what traits are present or absent in the languages) plays a significant role in L2 representation and processing. These predictions are based on the experimental evidence gathered from many studies, where L1 vs. L2 differences have been found to significantly modulate L2 processing. Weber-Fox and Neville (1996), for instance, reported differences between in Chinese speakers of English and L1 speakers of English when testing them on subjacency effects in English wh-questions (Chinese lacks overt wh-movement while English is an overt wh-movement language). Similar findings were described in Mueller, Hahne, Fujii, and Friederici (2005), who tested L1 German speakers of (an artifical mini-) Japanese on classifier morphology, a feature absent in German. Finally, when L1 speakers of Japanese (Ojima et al., 2005) or Chinese (Chen et al., 2007) were tested processing English verb agreement (a feature absent in their L1), significant differences in the ERP pattern obtained as well.

Given this evidence, we hypothesized that native/non-native differences would emerge for grammatical traits absent in L1 grammar even at high proficiency and early AoA. Conversely, non-native speakers should display native-like processing signatures for grammatical traits that are shared by L1 and L2.

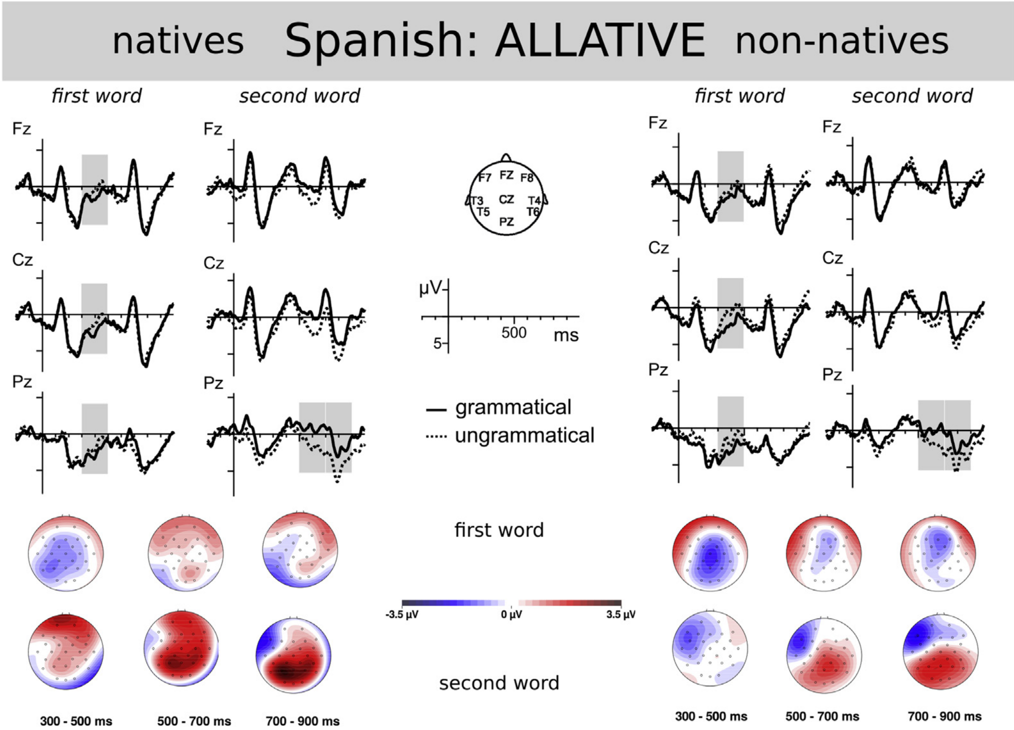

Fig. 6. ERPs elicited at the critical word positions in the Allative condition (Experiments 3–4). Brown lines represent the ungrammatical stimuli and the magenta lines represent the grammatical stimuli. Significant differences between the grammaticality conditions are highlighted by the gray areas. Topographical amplitude difference maps for the grammaticality effect below were calculated as the average difference amplitude between the ungrammatical condition and the grammatical baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.1. Testing L2 grammatical traits absent L1: ergative and nominative alignments

Results from the experiments 1 and 2 carried in Basque revealed significant differences between native and non-native speakers, consistent with previous findings (Díaz et al., 2016; Zawiszewski et al., 2011). Both groups showed a similar a N400 – P600 pattern for ergative violations, but non-natives displayed a smaller P600. Non-natives were also less accurate in judging sentence acceptability and slower performing the task. A mirror-like picture emerged in Spanish when processing differential object marking (DOM): although violations elicited a broadly distributed negativity in both groups, the following P600 was smaller for non-natives. Non-natives were also less accurate judging ungrammatical sentences.

These findings further support the hypothesis that native grammars modulate the representation and processing of later learned languages and provide more evidence that native/non-native differences are observed for traits absent in the native grammar, even for languages learned at early ages, frequently used throughout adulthood and having attained a very high proficiency.

3.2. Testing grammatical traits in L2 present in bilinguals' L1: the dative condition

The results obtained from Basque-Spanish and Spanish-Basque early and proficient bilinguals processing dative marked NPs were consistent with the LDH: since dative marking is found both in Basque and in Spanish grammars, no differences between natives and non-natives were expected. Indeed, when processing Basque dative all participants showed a N400–P600 pattern towards the violations, although native speakers were faster and more accurate judging ungrammatical sentences. Dative violations in Spanish also elicited a similar N400–P600 at both critical words. However, against predictions, this positivity was significantly larger for the native group. We think these results might be due to the dative manipulations used in the experiments. In the present study all nouns marked by dative case and the corresponding ungrammatical versions were obligatorily accompanied by a pronominal clitic *le* (third person singular).This clitic-NP dependency is required in contemporary peninsular Spanish but it is missing in Basque where the dative NP agrees with the auxiliary verb. Consequently, the clitic pronoun significantly increased the predictability of the upcoming dative NP. Violations of this expectation elicited a larger positivity among the natives than the non-natives. This suggests that in comparison to L1 speakers, L2 speakers are less sensitive to the contrasts absent in their L1, even at high proficiency and low AoA.

3.3. Testing grammatical traits in L2 present in bilinguals' L1: allative P morphology

In line with the LDH, we expected to find no differences between native and non-native speakers when processing allative morphology because Basque and Spanish are similar regarding this particular trait of grammar. However, against our predictions, significant differences between L1 and L2 speakers were encountered. L1 speakers showed a larger positivity than L2 speakers when processing allative P violations in both Basque and Spanish. Although these findings need to be interpreted with caution, the type of constructions used in the present study may be a possible source of the results obtained here. More precisely, Basque is a head final language where the head of the phrase follows the complement while Spanish is head initial where the heads precede the complements (cf. *parke-ra* 'park-the-to' in Basque vs. *al parque* 'to the park'). This means that the non-native speakers of Basque and Spanish had to deal with allative constructions in the L2 diverging from their L1 specifications. These L1 vs. L2 differences would be manifested by a larger processing cost for the non-natives than for the natives. A possible caveat to this explanation lies on the fact that this argumentation would be valid only for the structures containing pre-/postpositional phrases excluding those involving case. However, all the experimental manipulations in Basque contained head final phrases whereas all the materials tested in Spanish involved head-initial structures. Therefore, if native vs. non-native differences were due to the head-directionality parameter, differences would obtain for all the conditions tested in both Basque and Spanish. However, similar ERP pattern was found for dative case manipulations in Basque, casting doubt on the plausibility of such an interpretation.

Summing up, the results from allative morphology conditions can be accounted for by head directionality only if we assume that different processing mechanisms operate for case and pre-/postpositional constructions. This view coincides with Chomsky's (1995) proposal, which claims that pre-/postpositions (P) differ fundamentally from Case in that they have semantically interpretable features while Case does not. Unlike (uninterpretable) case features which are deleted during the derivation, interpretable features cannot be deleted and they remain accessible to the computation and are visible at Logical Form (LF). To recapitulate, although our results need to be interpreted with caution, they indicate that probably more factors than AoA, proficiency or linguistic distance between L1 and L2 have to be considered in order to fully account for native vs. non-native differences.

3.4. Impact of linguistic distance on the processing mechanisms in the L1 and the L2

Given the differences between Basque and Spanish regarding case morphology and head directionality, comparing the time course of electrophysiological processing can shed more light on the mechanisms used by native and non-native speakers when dealing with structures in their L2 converging or diverging with their L1 grammars. The results from ergative and dative case manipulations in Basque and those from accusative and dative case manipulations in Spanish suggest that information about the thematic hierarchy of the arguments is processed and established during the early stage of processing (300–500 ms) as indicated by the N400 component (cf. Frisch & Schlesewsky, 2001; Frisch & Schlesewsky, 2005), whereas the errors are reanalyzed and repaired later on (500–900 ms). Interestingly, data from Spanish show that predictions concerning the thematic roles of the arguments are made in absence of any semantic information: a N400 is present at the first critical word position (*al* vs. **el*) when no lexical information about the upcoming word has been provided yet.

As concerns the allative condition, the violations elicited a different ERP pattern in both languages: a P600 for Basque and a N400 (first word) followed by a P600 (second word) in Spanish. This suggests that allative constructions are processed differently in head-final and head initial languages. It seems that in the head-final language where the lexical information precedes the morphosyntactic one in the postpositional phrase (*klase-ra* 'classroom-the-to') no difficulty arises when identifying the thematic role of the argument after the semantics of the word has been presented and the structure of the whole phrase can be anticipated and interpreted (but bear in mind that this approach is build upon the null effect and consequently must be interpreted with caution). In a head-initial language like Spanish, thematic/semantic information about the prepositional phrase is accessed before the lexical one (*al parque* 'to the park'). Therefore, a violation at this point leads to wrong expectations and predictions at thematic/semantic level, as indicated by the N400 component. Alternatively, the divergent ERP patterns in Basque and Spanish may have come about because in Basque the ungrammaticality appeared at one word while it appeared in two words in Spanish. (Footnote 1: We thank an anonymous reviewer for suggesting this possibility.) However, the overall ERP pattern observed in both Basque and Spanish is consistent with previous literature on the topic (i.e. Barber & Carreiras, 2005; Guajardo & Wicha, 2014; for gender violations in Spanish; Molinaro, Vespignani, & Job, 2008, for gender violations in Italian; Mancini, Molinaro, Rizzi, & Carreiras, 2011 for subject-verb agreement violations in Spanish; Zawiszewski, Santesteban, & Laka, 2016; Chow, Nevins, & Carreiras, 2018; Martínez de la Hidalga, Zawiszewski, & Laka, 2019 for subject-verb agreement violations in Basque).

4. Conclusion

The aim of the present work was to investigate the impact of the linguistic distance on morphological processing in native and non-native speakers. Experimental evidence from Basque and Spanish suggests that the characteristics of the L1 have a deep impact on the way the L2 is processed. Although a more detailed research is needed to fully account for the influence of the native language over the non-native one during the L2 processing, our data indicate that L2 processing is to a large extent shaped by the characteristics of the L1 grammar, that is, native vs. non-native differences obtain for those traits in which L1 and L2 differ (ergative and accusative alignment, clitic doubling, head-directionality) while a similar pattern is found for the characteristics present in both the L1 and the L2 (dative case morphology).

Acknowledgments

Research leading to these results received funding from (1) European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 613465, (2) the Spanish Ministerio de Economía y Competitividad (FFI2015-64183-P) and (3) the Basque Department of Education, Universities and Research (IT1169-19).

References

Alemán Bañón, J. A., Fiorentino, R., & Gabriele, A. (2014). The processing of number and gender agreement in Spanish: An event-related potential investigation of the effects of structural distance. *Second Language Research, 30*, 275–306.
Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.
Bassetti, B. (2008). Orthographic input and second language phonology. In P. Thorsten, & M. Young-Scholten (Eds.). *Input matters in SLA. Multilingual matters* (pp. 191–206). Clevedon, UK.
Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). *Processing syntax and morphology: A neurocognitive perspective.* Oxford: Oxford University Press.
Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Towards a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology, 10*(298), 1–17.
Bosque, I., & Demonte, V. (1999). *Gramática descriptiva de la lengua española.* Madrid: Espasa Calpe S.A.
Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science, 41*, 1318–1352.
Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on crosslinguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.). *Cross-linguistic influence in third language acquisition, psychological perspectives, 8-20*. Clevedon: Multilingual Matters LTD.
Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition, 10*(2), 161.
Chomsky, N. (1995). *The minimalist program.* Cambridge MA: MIT Press.
Chow, W., Nevins, A., & Carreiras, M. (2018). Effects of subject-case marking on agreement processing: ERP evidence from Basque. *Cortex, 99*, 319–329.
Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied PsychoLinguistics, 27*, 3–42.
Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience, 21*, 1882–1892.
De Rijk, R. (2007). *Standard Basque: A progressive grammar.* Cambridge MA: MIT Press.
Díaz, B., Erdocia, K., de Menezes, R. F., Mueller, J. L., Sebastián-Gallés, N., & Laka, I. (2016). Electrophysiological correlates of second-language syntactic processes are related to native and second language distance regardless of age of acquisition. *Frontiers in Psychology, 7*(133), 1–13.
Díaz, B., Sebastián-Gallés, N., Erdocia, K., Mueller, J., & Laka, I. (2011). On the cross-linguistic validity of electrophysiological correlated of morphosyntactic processing: A study of case and agreement violations in Basque. *Journal of Neurolinguistics, 24*(3), 357–373.
Dixon, R. M. (1994). *Ergativity.* Cambridge: Cambridge University Press.
Dowens, M. G., Vergara, M., Barber, H., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience, 22*(8), 1870–1887.
Erdocia, K., Laka, I., Mestres-Missé, A., & Rodríguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain and Language, 109*, 1–17.
Erdocia, K., Zawiszewski, A., & Laka, I. (2014). Word order processing in a second language: From VO to OV. *Journal of Psycholinguistic Research, 43*(6), 815–837.
Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language, 66*(1), 226–248.
Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *Basic and Clinical Neurophysiology, 12*(15), 3391–3394.
Frisch, S., & Schlesewsky, M. (2005). The resolution of case conflicts from a neurophysiological perspective. *Cognitive Brain Research, 25*(2), 484–498.
Gleitman, L. R. (1985). Orthographic resources affect reading acquisition — if they are used. *Remedial and Special Education, 6*, 24–36.
Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.
Guajardo, L., & Wicha, N. (2014). Morphosyntax can modulate the N400 component: Event related potentials to gender-marked post-nominal adjectives. *NeuroImage, 91*, 262–272.
Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods, 45*, 765–771.
Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? *Psychological Science, 15*(6), 409–414.
Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research, 13*, 187–226.
Hualde, J., & Ortiz de Urbina, J. (2003). *A grammar of Basque.* Berlin: Mouton de Gruyter.
Kim, A., & Osterhout, L. (2005). The independence of combinatory semanticprocessing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205–225.
Kim, S., Qi, T., Feng, X., Ding, G., Li, L., & Cao, F. (2016). How does language distance between L1 and L2 affect the L2 brain network? An fMRI study of Korean–Chinese–English trilinguals. *NeuroImage, 129*, 25–39.
Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*, 196–214.
Kolk, H., Chwilla, D., van Herten, M., & Oor, P. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language, 85*, 1–36.
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.
Lelekov, T., Dominey, P., & Garcia-Larrea, L. (2000). Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. *NeuroReport, 11*, 1129–1132.
Leonetti, M. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics, 3*, 25–114.
Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011). A person is not a number: Discourse involvement in subject–verb agreement computation. *Brain Research, 1410*, 64–76.
Martínez de la Hidalga, G., Zawiszewski, A., & Laka, I. (2019). Eppur non si muove: Experimental evidence for the Unaccusative Hypothesis and distinct ϕ-feature processing in Basque. *Glossa: A Journal of General Linguistics, 4*(1), 120.
Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex, 66*, 156–159.
Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 47*, 908–930.
Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research, 1228*, 161–176.
Mueller, J., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience, 17*(8), 1229–1244.
Mueller, J., Hirotani, M., & Friederici, A. (2007). ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC Neuroscience, 8*, 18.
Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience, 17*(8), 1212–1228.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.
Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing neurocognitive processes reflected by P600 effects: Evidence from ERPs and neural oscillations. *PloS One, 9*(5), 1–11 e96840.
Salaburu, P. (2016). *Writing words: The unique case of the standardization of Basque.* Nevada: The Center for Basque Studies, University of Reno.
Schepens, J., van der Slik, F., & van Hout, R. (2016). L1 and L2 distance effects in learning L3 Dutch. *Language Learning, 66*, 224–256.
Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research, 12*, 40–72.
Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement. *Cortex, 66*, 149–155.
Tanner, D., & van Hell, J. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289–301.
Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An Event-Related Potential Investigation. *Studies in Second Language Acquisition, 27*(2), 173–204.
Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*, 37–69.
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*(3), 231–256.
Zawiszewski, A., Gutiérrez, E., Fernández, B., & Laka, I. (2011). Language distance and non-native syntactic processing: Evidence from event-related potentials. *Bilingualism: Language and Cognition, 14*(3), 400–411.
Zawiszewski, A., Santesteban, M., & Laka, I. (2016). Phi-features reloaded: An ERP study on person and number agreement processing. *Applied PsychoLinguistics, 37*, 601–626.