






RESEARCH ARTICLE 

Prosodic processing in sentences with ‘only’ in L1 and L2 English

Rachida Ganga¹ , Haoyan Ge² , Marijn E. Struiksma¹ , Virginia Yip³  and Aoju Chen¹ 

¹Institute for Language Sciences, Utrecht University, The Netherlands; ²School of Education and Languages, Hong Kong Metropolitan University, China; ³Department of Linguistics and Modern Languages, Chinese University of Hong Kong, China

Corresponding author: Aoju Chen; Email: aoju.chen@uu.nl

(Received 23 July 2021; Revised 24 January 2023; Accepted 06 March 2023)

Abstract

It has been proposed that second language (L2) learners differ from native speakers in processing due to either influence from their native language or an inability to integrate information from multiple linguistic domains in a second language. To shed new light on the underlying mechanism of L2 processing, we used an event-related potentials (ERP) paradigm to examine the processing of sentences with *only* in English by native speakers of English and advanced Dutch learners of English. Successful processing of sentences with *only* requires rapid integration of prosodic information with semantic and syntactic information. We found that L2 listeners showed native-like processing of the acoustics of contrastive pitch accents when adjacent to *only*. However, they needed more cues than L1 listeners to perform native-like in forming expectations for focus placement. Our results thus provide first ERP-based evidence for difficulty in the integration of information for focus expectation in difficult L2 constructs.

Introduction

There has been much discussion in second language acquisition literature on why processing in the native language (L1) and in a second language (L2) differs from each other. One circulating hypothesis is that constructions that are similar in the L1 and L2 are processed similarly, resulting in faster or more accurate processing, also known as (positive) L1 transfer. However, if the constructions differ, L2 learners have to employ different or less commonly used processing strategies, leading to slower or inaccurate processing (see reviews Clahsen & Felser, 2006; Frenck-Mestre, 2002; seen in, e.g., *wh*-fronting Juffs, 2005). Another hypothesis is that L2 learners may experience difficulties in the interface between some linguistic domains in which the integration of information is required (Sorace, 2011; Sorace & Filiaci,

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

2006; White, 2011). This view suggests that L2 learners are able to easily acquire simple syntactic structures, but they may find it more difficult to combine information from multiple constructs to adjust their interpretation of a sentence as it unfolds (Sorace, 2011), as reported for the processing of contrastive focus (Lee et al., 2020), *wh*-fronting (Marinis et al., 2005), and anaphora resolution (Sorace & Filiaci, 2006). This may be due to less (detailed) knowledge that is automatically available or simply fewer cognitive resources available for the integration in the L2 (for a review see Sorace 2011). Although both of these hypotheses assume that L1 and L2 processing differs, this does not automatically mean that L2 learners cannot achieve L1-like proficiency. For example, highly proficient Dutch learners of English use different processing strategies than native speakers of English (Akker & Cutler, 2003; Chen & Lai, 2012; Ge et al. 2021a), but they can be as quick and accurate as native speakers as measured by reaction time at the end of a target sentence (Ge et al. 2021a). This suggests that, instead of a 'wrong' approach, L2 listeners may simply use a different online processing strategy than L1 listeners (Bley-Vroman, 1983; Purdy, 2001). Taken together, L2 constructs that also occur in the L1 are processed with an L1-like strategy, but simpler ones that do not exist in the L1 can be acquired and processed with strategies similar to the ones employed by native speakers of the L2. If L2 constructs are complex and do not occur in the L1, or they require the integration of domains, L2 listeners tend to employ a strategy that differs from native speakers of either language, that is, an interlanguage processing strategy.

But what processing strategy do highly proficient L2 listeners employ when they are confronted by a challenging interface between individual constructs that occur in both their L1 and L2? The present study examined the processing strategies used during the integration of linguistic domains in the L2 in the face of complexity and L1-L2 similarities of the individual constructs, through the lens of prosodic processing in sentences with the focus particle *only*. We examined advanced listeners whose L1 and L2 are typologically highly similar in domains involved in the processing of sentences with *only*, i.e., prosody and semantics, but differ in the interface between those domains. Briefly, Dutch and English are similar in their use of prosody in focus marking (Ladd, 1980; Gussenhoven, 1983; Trommelen & Zonneveld, 1999) and the use of focus particles, such as *only* (Bouma et al., 2007; Foolen et al., 2009). However, Dutch and English differ when it comes to the interface between *only* and placement of accentuation caused by an unrelated syntactical difference concerning the positioning of the finite verb in a sentence (Bouma et al., 2007; Foolen et al., 2009; Ge et al. 2021b). Moreover, Dutch high school graduates often achieve advanced proficiency due to their English education starting from as early as the age of four years, with exposure sometimes starting earlier via (social) media, for example, undubbed English-spoken television shows (Michel et al., 2021). At the age of 16 to 18, Dutch high school graduates have achieved English A2-B2 level, with 90% reported to be able to engage in conversation (Michel et al., 2021).

Considering the high proficiency of Dutch learners of English and the dissimilarity between L1 and L2 only laying in the interface between linguistic domains, we examined whether existing similarities are sufficient for L1-like or native-like processing or whether a similarity in the interface is required as well. Specifically, we examined the anticipation and processing of prosodically realized focus in real time using an event-related potential (ERP) paradigm via electroencephalogram (EEG).

The prosodic-semantic interface in sentences with only

Constructs with *only* exist in both Dutch and English and require complex processing of the interface between prosody and semantics (Bouma et al., 2007; Gussenhoven, 1983; Filik et al., 2009; Foolen et al., 2009; Ladd, 1980; Zimmermann & Onea, 2011). Semantically, *only* signals an upcoming contrastive focus (hereafter focus) in the utterance. It flags characteristics of the contrasted element as true and activates alternatives that are related to the contrast but have characteristics that are false (Filik et al., 2009; Rooth, 1992; Zimmermann & Onea, 2011). This contrast is signaled by acoustic prominence, achieved via accentuation (i.e., placement of a pitch accent) on the word carrying the contrastive information in both English and Dutch (Gussenhoven, 1983; 2004; Ladd, 1980; 1996; 2008; Zimmermann & Onea, 2011). In example (1a), the accented element is *ball*, eliciting an object-focus reading. *Only* highlights that the ball is the one thing that John is throwing and that he is not throwing any other objects. Similarly, in (1b) containing a verb focus, throwing the ball is the one action that John does.

- (1) a. John is only throwing the BALL, not throwing the pen. (object-focus)
 b. John is only THROWING the ball, not kicking the ball. (verb-focus)

The Dutch equivalent *alleen* is similar to *only* semantically and is always coupled with a pitch accent elsewhere in the utterance. However, there is a critical difference in the importance of prosodic features in encoding the locus of focus, which is triggered by an independent difference between Dutch and English syntax. In English, verbs are preferably placed adjacent to their direct objects (Bouma et al., 2007), resulting in *only* being more frequently placed in the preverbal position (1a and 1b) regardless of the locus of focus. Pitch accents are then used as the primary cue to locate focus. In contrast, Dutch has OV word order with the verb appearing before the object in main clauses and has in general a preference for *alleen* to be placed adjacent to the focal element (2a and 2b) (Bouma et al., 2007; Foolen et al., 2009). Thus, the position of *alleen* serves as an important cue to the locus of focus in tandem with accentuation.

- (2) a. *Ik heb een appel alleen GEGETEN.*
 I have an apple only eaten.
 I have only EATEN an apple.
 b. *Ik heb alleen een APPEL gegeten.*
 I have only an apple eaten.
 I have only eaten an APPLE.

The preferred positioning of *only* during production also affects where listeners expect focus to be (Ge et al., 2021b; Mulders & Szendroi, 2016). English listeners listening to English tend to wait for accentuation to determine the contrast (Ge et al., 2021b). In contrast, Dutch listeners listening to Dutch have a strong expectation for focus adjacent to *alleen* (Mulders & Szendroi, 2016), although they can also use prosodic cues to quickly reinterpret sentences with nonadjacent focus (Dimitrova, 2012; Mulders & Szendroi, 2016).

ERP research on focus processing

In the present study, we adopted an event-related potential (ERP) paradigm using electroencephalogram (EEG) to examine prosodic focus processing in English

sentences with *only* in L1 and L2 listeners. By aligning the onset of a neural response with the acoustic onset of a pitch accent, a representation (ERP component) of how the brain processes and integrates accentuation with other information in real time can be formed (Luck, 2014). Focus particles and context affect the accessibility of certain elements (Chafe, 1994), resulting in expectations for where the focus will be and eliciting specific ERP components (Dimitrova et al., 2012). For example, in question-answer pairs, a missing or incongruent pitch accent in the answer based on the focus position implied by the preceding question often elicits a negativity peaking at 400 ms (N400) (e.g., Dimitrova et al., 2012; Hruska et al., 2000; Johnson et al., 2003; Schumacher & Baumann, 2010; Toepel et al., 2009). The N400 reflects registrations of violations in morphosyntax, phrase structures, verb tenses, grammatical gender, and so on (e.g., Baumann & Schumacher, 2011; Osterhout & Holcomb, 1992; 1993; Osterhout et al., 1997; Schumacher & Baumann, 2010; Toepel et al., 2009; Wicha et al., 2004). This ERP component is often followed by a centroparietal positivity at approximately 600 ms (P600), associated with reevaluation of sentence meaning or difficulty in integrating information (Burkhardt, 2007; Friederici, 2011; Hagoort, 2003; Kaan et al., 2000).

Prosodic focus processing elicits two additional ERP components. First, focus expectancy elicits a frontal negativity between 100–200 ms, i.e., the expectancy negativity (EN) (e.g., Heim & Alter, 2006; Hruska & Alter, 2004; Toepel et al., 2009). This is followed by ERP effects specific for processing the pitch accents, which have been found for both implicit prosody during reading (Cowles et al., 2007; Reichle & Birdsong, 2014; Stolterfoht et al., 2007) and explicit prosody during auditory stimulation. Auditory pitch accents elicit a positive deflection around 200–500 ms measured broadly over the scalp (Baumann & Schumacher, 2011; Dimitrova, 2012; Dimitrova et al., 2012; Heim and Alter, 2006; Hruska et al., 2001; Lee et al., 2020), interpreted as either frontocentral P300, depicting surprise (Baumann & Schumacher, 2011; Magne et al., 2005), or broadly distributed P200(-like) ‘accent positivity’ reflecting the processing of the acoustic features of the pitch accent (Dimitrova 2012, Dimitrova et al., 2012; Heim and Alter, 2006; Lee et al., 2020). Previously, it has been shown that this ‘accent positivity’ may be absent or weakened in L2 listeners when listening to contrastive focus (Lee et al., 2020).

To the best of our knowledge, there has been one published study on *only* and prosodic focus processing in L1 Dutch (Dimitrova, 2012). Dimitrova (2012) found that *alleen* elicited an anticipation for adjacent focus, that is, rapid processing of the word, seen as a positivity at 100–200 ms, followed by the ‘accent positivity’. Nonadjacent focus was processed differently. Instead of an early positivity, accentuation elicited the EN, followed by the ‘accent positivity’. The author argued that the early processes both reflect expectancies but that the expectancies are built differently, thus affecting their polarity. That is, expectancy for the adjacent element is induced by *alleen*, whereas expectancy for the nonadjacent element depends on the lack of preceding prosodic focus marking (Dimitrova, 2012). Finally, nonadjacent focus also evoked a left anterior P600, suggesting the need to reevaluate the sentence. Together, these findings show that L1 Dutch listeners expect focus adjacent to *alleen*, although they are flexible and quick to adjust their interpretation when processing nonadjacent focus (Dimitrova, 2012).

The current study

The present study examined anticipation and processing of focus in the interface between linguistic domains by L1 English listeners and highly proficient L2 English-L1 Dutch

listeners, in order to shed light on how language similarities within domains affect the processing between domains in L2 listeners in comparison to L1 listeners. We presented sentences with *only* with prosodic focus marking either adjacent (on the verb) or nonadjacent (on the object noun) to the focus particle (e.g., in 'The dinosaur is only throwing the bucket'). The locus of focus could only be determined by processing and integrating the pitch accent correctly. Additionally, these sentences were presented either in isolation, that is, single sentences, or preceded by context. The purpose of the context was not to impose a contrast. Rather, it provided possible alternatives to the adjacent and nonadjacent words that could be activated by *only* (Filik et al., 2009; Rooth, 1992; Zimmermann & Onea, 2011), contextualising the use of *only*. This may facilitate processing of focus by alleviating cognitive resources (Dimitrova et al., 2012). To explore L2 strategies and examine them against L1 strategies, we formulate three research questions, accompanied by hypotheses and predictions.

Research question 1

Do L1 and L2 English listeners differ in their processing of prosodic focus on the element adjacent to *only*, and to what extent do they differ from L1 Dutch listeners?

Hypothesis 1

As *only* can imply prosodic focus to its adjacent element in both languages (Bouma et al., 2007; Foolen et al., 2009), we hypothesize that L1 English listeners will show similar processing patterns as L1 Dutch listeners (Dimitrova, 2012) (**Hypothesis 1a**), and that L2 English listeners will do so too regardless of the presence or absence of the preceding context, as a result of positive L1 transfer (Clahsen & Felser, 2006; Frenck-Mestre, 2002) (**Hypothesis 1b**). Our prediction is that both L1 and L2 listeners will show an early broadly distributed positive peak (100–200 ms), reflecting expectancy in the form of quick processing of the adjacent element (Dimitrova 2012). This early positive peak will be followed by a broadly distributed 'accent positivity' (200–390 ms), reflecting the processing of the acoustic features of the pitch accent (Dimitrova et al., 2012; Heim and Alter, 2006; Lee et al., 2020).

Research question 2

Do L1 and L2 English listeners differ in their processing of prosodic focus when the focused element is not adjacent to *only*, and do they differ from L1 Dutch listeners?

Hypothesis 2

Because *only* is more likely to be accompanied by nonadjacent focus in English than in Dutch (Bouma et al., 2007; Foolen et al., 2009), we hypothesize that L1 English listeners will show slightly different processing patterns than those found in L1 Dutch listeners (Dimitrova, 2012) (**Hypothesis 2a**). As advanced Dutch learners of English have been shown to exhibit slow or absent integration of prosodic and focus information in English (Akker & Cutler, 2003; Ge et al., 2021b), we hypothesize that L2 listeners will experience difficulty in integrating different strands of linguistic information in L2 (Sorace, 2011; Sorace & Filiaci, 2006) and employ their own strategy to process nonadjacent focus, different from L1 English listeners' strategy and L1 Dutch listeners' strategy (Dimitrova, 2012) (**Hypothesis 2b**). Specifically, our prediction is that L1 English listeners will expect focus on the nonadjacent position and show an early EN

(100–200 ms), as this focus position is cued by the lack of prosodic marking in the preceding verb (Dimitrova, 2012; Heim & Alter, 2006; Hruska & Alter, 2004; Toepel et al., 2009), which will then be followed by the ‘accent positivity’ (200–390 ms). Different from L1 Dutch processing (Dimitrova, 2012), we do not expect to see broadly distributed P600 (500–900 ms) in L1 English processing. As both focus positions are possible in English, reevaluation of the sentence meaning should not be needed (Friederici, 2011; Ge et al., 2021b; Hagoort, 2003; Kaan et al., 2000). Our prediction for L2 listeners is that they will not expect the nonadjacent focus and therefore not show an EN, but they will be able to react to the pitch accent itself (‘accent positivity’ at 200–390 ms), after which sentence reevaluation will occur as shown by a broadly distributed P600 (500–900 ms).

Research question 3

Will the presence of context adjust L2 listeners’ processing of prosodic focus?

Hypothesis 3

We hypothesize that the presence of a preceding context will alleviate the cognitive load in L2 listeners (Chafe, 1994; Filik et al., 2009; Zimmermann & Onea, 2011). Our prediction is that L2 listeners will show more L1 English-like processing patterns of contrastive focus in sentences preceded by a context than in isolated sentences.

Method

Participants

Twenty native speakers of English (eight male, 18–34 years old, $M = 24$ years, $SD = 4.5$ years, hereafter the L1 group) and thirty-seven native speakers of Dutch (fourteen male, 19–26 years old, $M = 22$ years, $SD = 1.8$ years, hereafter the L2 group) participated in the ERP experiment.¹ Two female L1 participants were excluded due to inattentiveness as measured with a semantic relatedness task (see Procedure). Four L2 participants (two male) were excluded due to technical failure ($n = 2$), being left-handed ($n = 1$), or inattentiveness ($n = 1$). The remaining participants were right-handed, had normal or corrected to normal vision, and reported not having any neurological, psychiatric, hearing, or language impairments.

All L2 participants were university students and frequently exposed to English in their daily lives at the time of testing, for example, through (social) media (Michel et al., 2021) ($n = 5$) or doing a degree program that was taught in English ($n = 28$). The test LexTALE (Lemhöfer & Broersma, 2012) was used to assess the L2 participants’ English proficiency. LexTALE is a short validated vocabulary test in which test takers see a

¹Due to the difficulty in finding L1 English speakers interested in participating within the timeframe of this project in the Netherlands, we included native speakers of various English varieties to increase the total number of L1 listeners. The L1 participants originated from the United States ($n = 8$), England ($n = 7$), Ireland ($n = 1$), Australia ($n = 2$), Canada ($n = 1$), and a British community in Hong Kong who did not have knowledge of Mandarin Chinese or Cantonese ($n = 1$). We acknowledge that speakers of different English varieties can differ in their use (Fletcher et al., 2005) and processing (Arnhold et al., 2020) of prosody, which might have implications for the generalizability of our findings. However, we lacked the statistical power to test the effect of English variety on prosodic processing. Instead, we added participant as a random factor in the analyses to factor possible influence from individual characteristics in.

written word and rate whether it is an English word or a nonword. Although it only tests vocabulary knowledge, LexTALE scores correlate with general English proficiency as tested in the Test of English for International Communication (TOEIC) and the Quick Placement Test (QPT). On average, a LexTALE score below 59% corresponds to B1 level or lower, 60–80% to B2, and 80–100% to C1/C2 (Lemhöfer & Broersma, 2012). In our study, the participants achieved a mean LexTALE score of 87.88 ($n = 30$; three participants had missing data; score range: 65 to 100; $SD = 9.97$; 95% $CI = [86.70; 89.05]$), implying that the L2 group was highly proficient. Considering the high exposure to English and high LexTALE scores, we decided to not include proficiency as a factor in the analyses.

Stimuli

Sixty unique pairs of experimental stimuli (Table 1) and 118 fillers (Table 2) were derived from stimuli used in Ge et al. (2022).² The stimuli were recorded by a male native speaker of British English (age 25–27 years, from York, the United Kingdom) at 44.1kHz sampling frequency with 16 bits resolution in a shielded recording booth. The speaker was instructed to naturally produce context and target sentences with half of the 120 stimuli containing accentuation on the verb (Table 1.A), and the other half on the object (Table 1.B). The context sentences presented two possible objects and actions, and hinted at a correction. The correction occurred in the target sentences, which could only be correctly interpreted by accurate processing of the accentuation. To test the effect of context, we created an additional 120 stimuli by removing the context from the above-mentioned stimuli (Table 1.C and D). In total, we used 240 experimental items.

Two sets of fillers lacking the focus particle *only* were used. The first set consisted of target sentences containing an explicit correction, either preceded by context sentences with three objects and an action ($n = 26$, Table 2.A) or lacked these context sentences ($n = 52$, Table 2.B). The target sentences were semantically and prosodically congruent with the context sentences. The second set contained context sentences with an additional sentence, followed by either a target sentence with a congruent pitch accent ($n = 20$, Table 2.C), or incongruent pitch accent ($n = 20$, Table 2.D). The

Table 1. Examples of experimental materials in British-English

Context sentences	Target sentences	
	Verb accentuation	Object accentuation
The dinosaur has a bucket and a suitcase. He was going to throw them and kick them. Then he changed his mind.	(A) The dinosaur is only CARRYING the bucket.	(B) The dinosaur is only carrying the BUCKET.
-*	(C) The dinosaur is only CARRYING the bucket.	(D) The dinosaur is only carrying the BUCKET.

*Target sentences C and D are not preceded by context sentences.

²The sixty pairs of stimuli were edited from sentences like 'The dinosaur is only CARRYING the bucket, not throwing the bucket' to sentences like 'The dinosaur is only CARRYING the bucket' by removing the 'not' part. Our informal perception test confirmed that the adapted stimuli sounded final after the object noun.

Table 2. Examples of filler materials

Filler type	Context sentences	Target sentences
<i>Correction of action</i>		
A (n = 26)	The dog has a broom, a paper, and a tomato. He was going to wash the tomato. Then he changed his mind.*	The dog is GETTING the tomato, not washing the tomato.*
B (n = 52)	-**	The dog is GETTING the tomato, not washing the tomato.*
<i>Prosodic (in)congruity</i>		
C (n = 20)	The fox has a banana and a pear. She was going to peel the banana. Then she changed her mind. I wonder what the fox is peeling.*	The fox is peeling the BANANA.*
D (n = 20)	The fox has a banana and a pear. She was going to peel the banana. Then she changed her mind. I wonder what the fox is peeling.*	The fox is PEELING the banana.*

*Accented words are in capitals.

**This type of filler did not have context sentences.

target sentences were semantically incongruent with the context sentences. Despite the limited variety of filler types, we believed they were sufficiently different from the experimental stimuli to distract the participant from our research purpose, confirmed by answers that the participants provided in an informal questionnaire after the EEG recording.

Twelve lists were created with all experimental and filler items (n = 358); participants were randomly assigned to one of the lists. Six of these lists were uniquely pseudo-randomized, meaning that (1) items with the same target sentence were never presented consecutively, (2) no more than three stimuli with the same accentuation pattern were presented consecutively, and (3) no more than three stimuli with the same context were presented consecutively. These six lists were reversed for the remaining six lists such that the first trial of the original list became the last trial. This procedure counter-balanced the trial order across participants, minimizing effect of fatigue in the later trials. Each list was preceded by the same practice round consisting of four items. Two of which were similar to the experimental items and two to the filler items. See [Supplementary Materials A](#) for the full list of stimuli³.

Acoustic analysis

To establish that the stimuli in the two accentuation conditions were acoustically different and were therefore appropriate for the ERP experiment, we conducted an acoustic analysis on the critical words (i.e., verbs and object nouns). Mean (f0), minimum (f0min), and maximum pitch (f0max), pitch range (f0range), and duration were extracted from the critical words using Praat (Boersma & Weenink, 2022), and were analyzed with Wilcoxon signed-rank test given that the acoustic data were skewed and contained missing values due to differences in voice quality.

We expected that accentuation would result in longer duration, higher f0 max, lower f0 min, and a larger f0 range. Indeed, the results (Table 3) showed that on average the accented verbs were 36 ms longer, had a 74Hz higher mean pitch, a 1Hz lower

³Audio files, lists, aggregated data, scripts, and analyses can be found on our archive: <https://doi.org/10.24416/UU01-O8Q739>. Raw EEG files can be requested by contacting the corresponding authors directly.

Table 3. Mean acoustic measurements of verbs and objects and standard deviations (in brackets) in the accented and unaccented conditions

	Unaccented verbs	Accented verbs	p^*	Unaccented objects	Accented objects	p^*
Duration (ms)	374 (45)	410 (48)	.000	345 (82)	388 (74)	.000
F0 (Hz)	84 (11)	158 (19)	.000	70 (38)	189 (26)	.000
F0 min (Hz)	73 (15)	72 (34)	.008	63 (39)	132 (63)	.000
F0 max (Hz)	96 (8)	209 (17)	.000	77 (39)	217 (24)	.000
F0 range (Hz)	22 (16)	134 (41)	.000	14 (18)	85 (65)	.000

*Significance at $p < 0.05$.

minimum pitch, a 203Hz higher maximum pitch, and a 112Hz larger pitch range (i.e. difference between the maximum and minimum pitch) than their unaccented counterparts. Similarly, in comparison to the unaccented object nouns, the accented object nouns were on average 43 ms longer, had a 119Hz higher mean pitch, a 140Hz higher maximum pitch, and 71Hz larger pitch range. Thus, the accentuation resulted in the expected acoustic changes and the stimuli could be used without further manipulation.

Procedure

This study was conducted in accordance with the Ethics Assessment Committee Linguistics (ETCL) of the Faculty of Humanities at Utrecht University. All participants provided informed consent and their visit (approximately 2.2 hours) was compensated with 23 euros.

The participants were seated in front of a computer screen in a quiet, unshielded room in the EEG lab of the Institute for Language Sciences. After electrode application and instructions, auditory stimuli were presented with Presentation software (v.20.0, Neurobehavioral Systems, 2019) via loudspeakers, positioned on either side of the screen. The experiment was preceded by four practice trials, after which the participant had the opportunity to ask questions to the experimenter. The remaining trials were divided into six blocks of fifty-one trials and one block of fifty-six trials. Each block was approximately 11 minutes long and was followed by a 2-minute obligatory break, allowing the participants to blink and stretch. In addition, optional breaks were provided approximately every 4 minutes and were as long as the participants needed. During auditory stimulation, to minimize eye movement the participants were instructed to fixate on a white cross against a black background on the computer screen, which appeared 100 ms prior to stimulus presentation and remained until 1000 ms after stimulus presentation. Each auditory stimulus was followed by a blink trial, in which the cross was replaced by three dashes for 2000 ms to indicate that participants could blink between trials.

To ensure that the participants were attentive to the stimuli, a semantic relatedness task, similar to the one used in Dimitrova et al.'s (2012) Dutch prosodic processing study, was implemented in 25% of the trials. By only implementing the task in a quarter of the trials, we minimized task-related ERPs (Dimitrova et al., 2012). In these ninety-four trials, a word was presented visually at the end of a trial. The participants were instructed to decide whether the presented word was related to the fragment they last heard by pressing the left shift key on a keyboard for YES or the right shift key for NO. These were visually marked by color to represent the different answers. These words were either unrelated or related to the subjects, verbs, or objects of the

preceding audio stimulus. All words were always presented after the same stimulus and were balanced over the four word categories (for exact word content and related stimuli see [Supplementary Materials A](#)). After the participants' response, the blink trial was presented before the next trial was initiated.

EEG was recorded in sixty-four channels at 2048 Hz using a sixty-four-channel BioSemi cap with Ag/AgCl electrodes in the 10–20 configuration. Additional electrodes on both mastoids, next to both eyes, and above and below the left eye were used for offline referencing, correction of horizontal eye movements (hEOG), and correction of vertical eye movements (vEOG) or blinks respectively. Impedances were kept under $\pm 20 \mu\text{V}$ and in rare cases $\pm 30 \mu\text{V}$. EEG markers were time-locked to the acoustic onset of the verbs and objects.

EEG preprocessing

The raw EEG signal was preprocessed using BrainVision Analyzer (v2.1.2.327, Brain Products, 2019). We performed the following steps in a chronological order. The data were first rereferenced to the average of both mastoids, after which they were bandpass-filtered at 0.1–35 Hz (24 dB slope), and downsampled to 500 Hz for further processing. Using the EEG markers, epochs were created from 100 ms before to 1000 ms after the onset of the verbs and objects. Next, mean amplitude in the 100 ms before the onset (-100 to 0 ms) was subtracted from every epoch as means of baseline correction. Finally, artefact rejection was done in two rounds. In the first round, trials that were contaminated by blinks or other eye movements were rejected from all channels. To this end, two new channels were created using the two hEOG and two vEOG electrodes. If, within an epoch, the signals in these two channels exceeded an amplitude of $\pm 75 \mu\text{V}$, displayed a voltage step of $50 \mu\text{V}$ or more between two adjacent sampling points or the difference in signal activity was lower than $0.5 \mu\text{V}$ in an interval of 100 ms, the whole epoch was rejected in all other channels. The second round served to reject trials contaminated by other artefacts, for example, muscle movement, on a channel-specific level. The same rejection criteria in the first round were applied to the signals of all sixty-four scalp electrodes. If the criteria were violated, an epoch was marked for rejection on an individual electrode and trial basis. In the end, 87%, 88%, 91%, and 91% of the trials in the L1 group, and 85%, 85%, 88%, and 89% in the L2 group in conditions A, B, C, and D, respectively, were included in the analyses.

To create the datasets for statistical analysis, the following steps were done in Matlab (R2019a, 2019). The trials marked for rejection were excluded on channel and electrode levels. ERPs, as area-under-the-curve data, per word in each condition were then calculated. To do so, epochs were averaged per channel into three time windows: an early time window (100–200 ms) to test for the frontal EN or the broadly distributed positivity, a middle time window (200–390 ms) for the global 'accent positivity', and a late time window (500–900 ms) for the anterior P600. These time windows were predefined based on previous research (Dimitrova et al., 2015; 2010a; 2010b; 2012; Luck & Gaspelin, 2017; Regel et al., 2014) and, in the case of the middle time window, also the average duration of the verbs in our study. Finally, the epochs per scalp electrode were assigned to one of nine scalp regions, which were based on Dimitrova (2012) in order to compare our results with L1 Dutch: left anterior (Fp1, AF3, AF7, F3, F5, and F7), Left Central (FC3, FC5, C3, C5, CP3, and CP5), left posterior (P3, P5, P7, PO3, PO7, and O1), Right Anterior (Fp2, AF4, AF8, F4, F6, and F8), right central (FC4, FC6, C4, C6, CP4, and CP6), right posterior (P4, P6, P8, PO4, PO8, and O2), center

anterior (FPz, AFz, and Fz), center central (FCz, Cz, and CPz), and center posterior (Pz, POz, and Oz).

In the end, we obtained two ERP datasets for the verbs and three datasets for the objects,⁴ each in the separate time windows. Each dataset contained trial data⁵ per participant in four conditions per scalp region.

Statistical analysis

The EEG datasets were analyzed using mixed-effect linear regression modeling by means of R (R Core Team, 2015), RStudio (RStudio Team, 2019), and the package lmer4 package (v.1.1.21, Bates et al., 2015). Separate models were created for the verbs for the early and middle time windows (as the late window went beyond the duration of the verb) and objects and for each of the three time windows, resulting in five different models.

The models were constructed using the outcome variable *ERP* and random factors *Participant* and *Item*. Fixed factors included *Context* (with or without), *Accent* (on the verb or object), *Region of Interest* (henceforth *ROI*: left anterior, center anterior, right anterior, left central, center central, right central, left posterior, center posterior, right posterior), and *Group* (L1 or L2). We also added a random slope of *Accent* in *Participants* to allow some individual variability in responses to the pitch accent. We added the factors iteratively and used the -2LL chi-square test to examine whether models improved ($p < .05$).

Because we are interested in how listeners process accentuation (i.e., the effect of *Accent*) and how the presence of context affects accent processing (i.e., the interaction between *Context* and *Accent*), we only analyzed main effects of and interactions with *Accent* in detail. We also only examined the highest level of an interaction that included *Accent*. By limiting analyses to main effects and highest level of interactions with *Accent*, we could eliminate unnecessary analyses and thus lower our experiment-wise error rate (Luck & Gaspelin, 2017).

Significant interactions with *Accent* were further analyzed using pairwise comparisons using R (R Core Team 2015) and the package emmeans (v.1.3.5, Russell Lenth, 2019). As the emmeans package typically conducts all possible pairwise comparisons within an interaction, we chose to only analyze and manually adjust the p -value for simple comparisons of *Accent* and *Context*, differing only in one factor, using Benjamini–Hochberg correction. This correction method controls the expected proportion of false positives, while keeping a high power (Hemmelmann et al., 2005).

⁴It was possible that objects were processed differently across accentuation condition due to the presence of absence of a pitch accent on the verb. We checked this concern visually by computing two ERP figures of the objects: one with a baseline preceding the verbs and one with a baseline preceding the objects. Both figures looked similar at the acoustic onset of the object, suggesting that the objects were not processed differently based on the accentuation status of the verbs across the two conditions.

⁵Considering the discrepancy in sample size in the two groups, we analyzed the data on an individual trial basis, instead of averaging all trials per condition per participant. This increased the number of data points, which in turn increased power. Furthermore, trial-based analysis allowed us to include all participants whose data would have otherwise been excluded due to the loss of too many epochs during artefact rejection. We therefore did not impose a criterium stating a minimum percentage of retained trials to include a participant in the analysis.

Results

Performance on semantic relatedness task

To determine attentiveness, the participants needed to achieve a minimum accuracy of 80% in the semantic relatedness task. Three participants did not meet this requirement and were excluded from the analysis. The accuracy of the final sample ranged from 81% to 96% ($n = 18$, $M = 87\%$, $SD = 4.20\%$, 95% C.I. [89, 92]) in the L1 group, and from 81% to 96% ($n = 33$, $M = 90\%$, $SD = 3.83\%$, 95% C.I. [89, 91]) in the L2 group.

Results on the ERP datasets

Complete models, model summaries, and the emmeans output of relevant significant interactions can be found in [Supplementary Materials B](#) for the verbs and [Supplementary Materials C](#) for the objects. Only the significant findings are reported here and were taken from the best-fit models. All best-fit models contained the random effects of *Item* and *Participant*, with the random slope for *Accent* over *Participants*, and the fixed effects of *Context*, *Accent*, *ROI*, and *Group*. In addition, each model had several interactions that improved the final model (see [Table 4](#)).

Verbs: 100–200 ms time window

There were significant main effects of *Intercept*, *Accent*, *ROI*, and *Group* and significant interactions of *Accent*ROI*, *Context*Group*, *Accent*Group*, *Context*Accent*Group*, and *Context*Accent*ROI*Group*, although all effect sizes were rather small. The interaction *Context*Accent*ROI*Group* was tested further. [Table 5](#) lists the emmeans output for reported findings (see [Supplementary Materials B](#) for the complete list).

Table 4. Interactions in best-fit models

Critical word	Time window (ms)	Included interactions
Verbs	100–200	<i>Context*ROI</i> <i>Accent*ROI</i> <i>Context*Group</i> <i>Accent*Group</i> <i>Context*Accent*Group</i> <i>Context*Accent*ROI*Group</i>
	200–390	<i>Context*ROI</i> <i>Group*ROI</i> <i>Context*Group</i> <i>Context*ROI*Group</i> <i>Context*Accent*Group</i> <i>Accent*ROI*Group</i>
Objects	100–200	<i>Accent*ROI</i> <i>Context*Group</i> <i>Accent*Group</i> <i>Context*ROI*Group</i> <i>Context*Accent*ROI*Group</i>
	200–390	<i>Context*ROI</i> <i>Context*Group</i> <i>ROI*Group</i> <i>Context*Accent*ROI*Group</i>
	500–900	<i>Context*ROI</i> <i>Context*Group</i> <i>ROI*Group</i> <i>Context*Accent*ROI*Group</i>

Table 5. Significant statistics of pairwise comparisons of examined interactions found for ERPs at verbs in the early and middle time-windows

Time window (ms)	Interaction in model	Contrast	Group	ROI*	Estimated means	Standard error	df	t	p (adj)	Cohen d	
100–200	<i>Context*Accent*ROI*Group</i>	Unaccented – Accented without context	L1	LA	–1.78	0.47	955	–3.80	0.002	–0.21	
				CA	–1.44	0.47	956	–3.06	0.017	–0.17	
				RC	–1.53	0.47	939	–3.28	0.010	–0.18	
200–390	<i>Accent*ROI*Group</i>	Unaccented – Accented with context	L2	CA	–1.47	0.37	997	–3.95	0.002	–0.17	
				RA	–1.52	0.37	997	–4.10	0.002	–0.18	
		Unaccented – Accented	L1	LA	–1.76	0.39	383	–4.46	0.001	–0.19	
				CA	–1.18	0.39	382	–3.00	0.006	–0.13	
				RA	–1.14	0.31	447	–3.64	0.002	–0.12	
				CC	–1.15	0.39	383	–2.93	0.006	–0.13	
				RC	–1.26	0.39	383	–3.21	0.004	–0.14	
				L2	LA	–0.78	0.31	447	–2.54	0.015	–0.08
					CA	–1.01	0.31	446	–3.27	0.004	–0.11
					RA	–1.12	0.31	447	–3.64	0.002	–0.12
LC	–1.40	0.31	443		–4.55	0.001	–0.15				
				CC	–1.03	0.31	442	–3.36	0.003	–0.11	
				RC	–1.09	0.31	443	–3.56	0.002	–0.11	
				LP	1.01	0.37	164	2.69	0.011	–0.07	
				CP	–0.92	0.31	440	–3.00	0.006	–0.10	
				RP	–0.83	0.31	441	–2.71	0.010	–0.09	

*LA = left anterior; CA = center anterior; RA = right anterior; LC = left central; CC = center central; RC = right central; LP = left posterior; CP = center posterior; RP = right posterior.

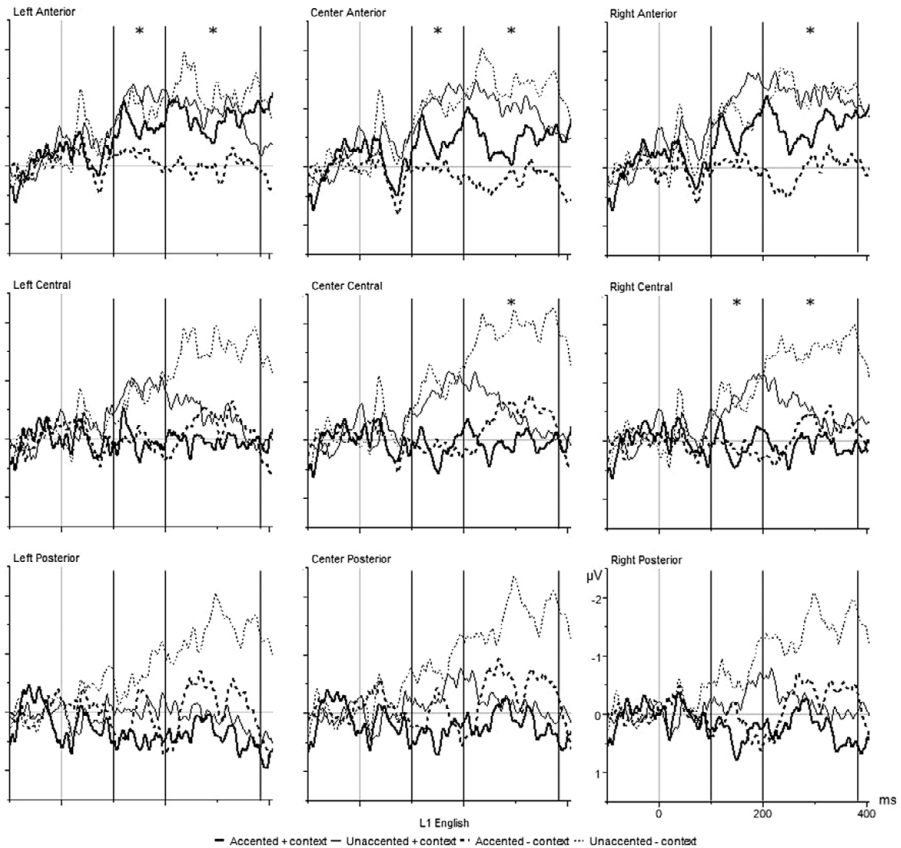


Figure 1. ERPs to the verbs for the native speakers of English (L1 English) in nine scalp regions. Thick solid ERP curves correspond to the accented verbs with context; thick dotted curves to the accented verbs without context; thin solid curves to the unaccented verbs with context; and thin dotted curves to the unaccented verbs without context. Gray vertical line corresponds to onset of stimulus and black vertical lines to the analyzed time windows. Asterisk indicates significant findings per time window.

As shown in [Figure 1](#) and the left half of [Figure 2](#), the L1 group responded to the accented verbs with more positivity than to the unaccented verbs in the absence of context in the 100–200 ms time window. This main effect of *Accent* was significant only for the left and center anterior regions, and right central region.

The main effect of *Accent* was also found for the L2 group, as shown in the right half of [Figure 2](#) and [Figure 3](#). However, it was only significant for the center and right anterior region in the presence of context.

The accented verbs therefore elicited an early anterior positive response in both groups but in different context conditions: without context for the L1 group and with context for the L2 group. This is in line with our predictions that both groups would show the early positivity and that the context would have an effect in the L2 listeners. The fact that they only showed this positivity in the presence of context suggests that they needed more contextual information during the processing of focus on the verb in the early time window.

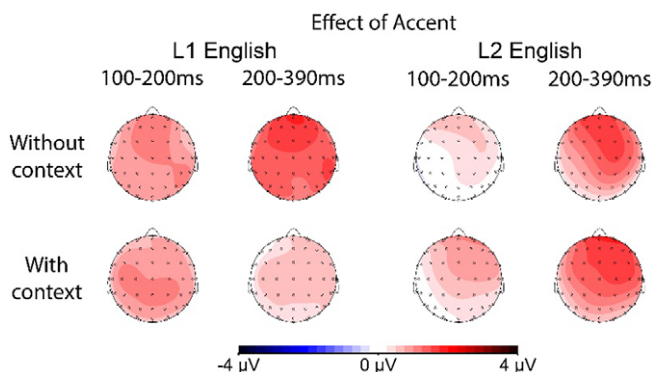


Figure 2. Topography maps for the effect of accent in the two context conditions in the two time windows 100–200 ms and 200–390 ms. The maps from the native speakers of English (L1 English) are on the left side, and the maps from the Dutch learners of English (L2 English) are on the right side. The topographies were created by subtracting the value from the unaccented verbs in each context condition from the accented verbs in the respective context condition.

Verbs: 200–390 ms time window

We found significant main effects of *Intercept*, *Context*, *Accent*, *ROI*, and *Group* and significant interactions of *Context*Group*, *ROI*Group*, *Context*ROI*Group*, and *Accent*ROI*Group*. The interaction of *Accent*ROI*Group* was further examined (see Table 5 for statistics).

Both the L1 and L2 groups showed a larger positivity to the accented than unaccented verbs that was broadly distributed (Figures 1, 2, and 3). The groups differed in involved regions: For the L1 group the positivity was present for the frontal, and center and right central regions, whereas all regions were involved in the case of the L2 group (Figure 1 and left half of Figure 2). Again, the results of the statistical analyses were in line with our prediction that the processing of the accented verbs would be reflected by the globally distributed positivity as a result of the processing of the pitch accent's acoustic features.

Objects: 100–200 ms time window

For the objects in the first time window, we found a significant main effect of *ROI* and significant interactions of *ROI*Group* and *Context*Accent*ROI*Group*. The interaction *Context*Accent*ROI*Group* was further examined. However, after adjustment, there were no significant effects of accentuation in neither group in any of the scalp regions (see Supplementary B for all findings, and Figures 4, 5, and 6 for the ERP graphs). This suggests that neither the L1 nor the L2 group showed early responses to the pitch accent, which goes against our prediction that nonadjacent accentuation would elicit an early frontal negativity in the L1 group.

Objects: 200–390 ms time window

We did not find a significant main effect of *Accent* or significant interactions that included the effect of *Accent* in the best-fit model (see Supplementary B), as can also be observed between 200 to 390 ms in Figures 4 and 6. Instead, we found significant effects of *ROI* and *ROI*Group*, which were not examined any further. The lack of significant

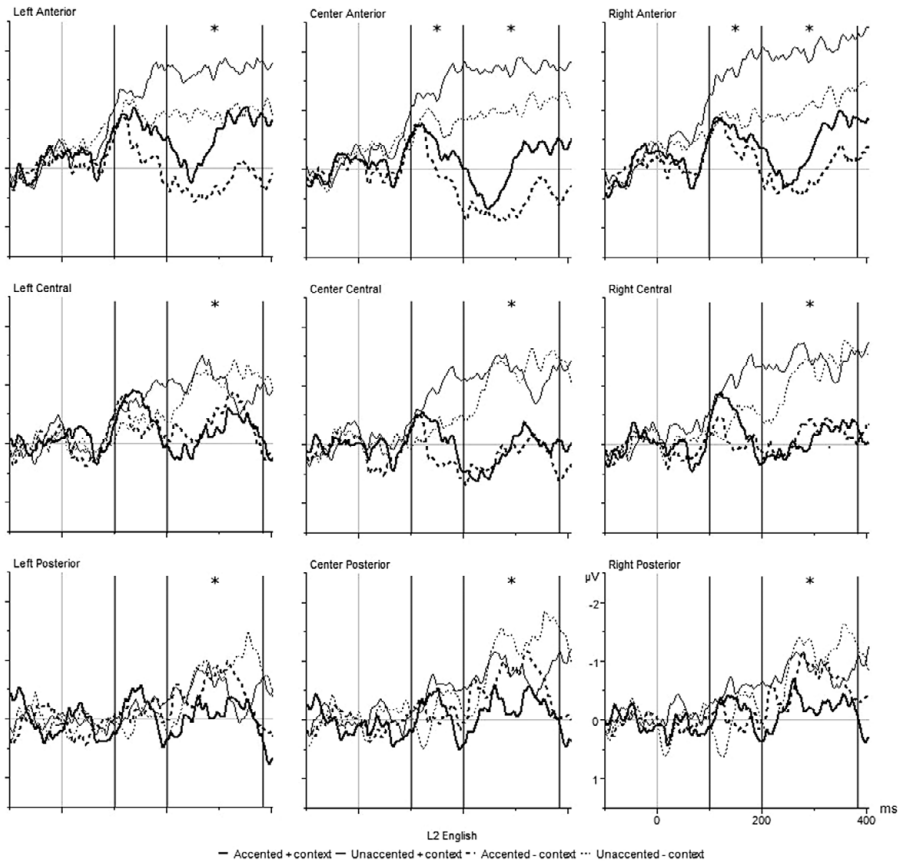


Figure 3. ERPs to the verbs for the Dutch learners of English (L2 English) in the nine scalp regions. Thick solid ERP curves correspond to the accented verbs with context; thick dotted curves to the accented verbs without context; thin solid curves to the unaccented verbs with context; and thin dotted curves to the unaccented verbs without context. The gray vertical line in each panel corresponds to onset of stimulus and black vertical lines to the analyzed time windows. Asterisk indicates significant findings per time window.

Accent effects goes against our predictions that both groups would show a positivity as a response to accentuation, suggesting that the groups did not process the pitch accent on the object noun.

Objects: 500–900 ms time window

Different from the previous two time windows, we found significant main effects *Accent* and *ROI* and significant interactions of *Context*Group*, *ROI*Group*, and *Context*Accent*ROI*Group*. The interaction of *Context*Accent*ROI*Group* was analyzed further. However, after adjustment, there were no significant effects of *Accent* in either group in any of the nine scalp regions (see [Supplementary B](#)). This finding is not in line with our predictions that the L1 group would show more positivity to the accented than unaccented objects, reflecting sentence reinterpretation, and the L2 group would not show such an effect.

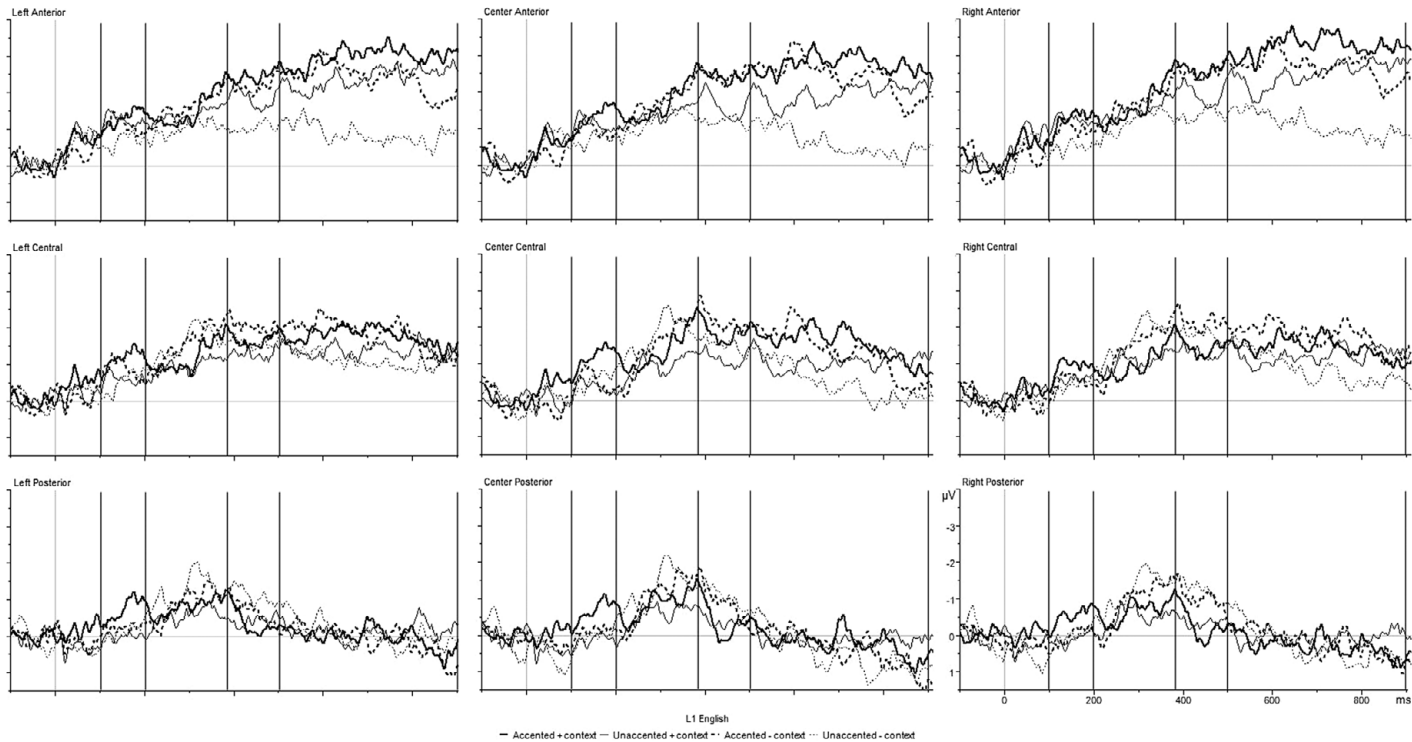


Figure 4. ERPs to the objects for the native speakers of English (L1 English) in nine scalp regions. Thick solid ERP curves correspond to the accented objects with context; thick dotted curves to the accented objects without context; thin solid curves to the unaccented objects with context; and thin dotted curves to the unaccented objects without context. Gray vertical line corresponds to onset of stimulus and black vertical lines to the analyzed time-windows.

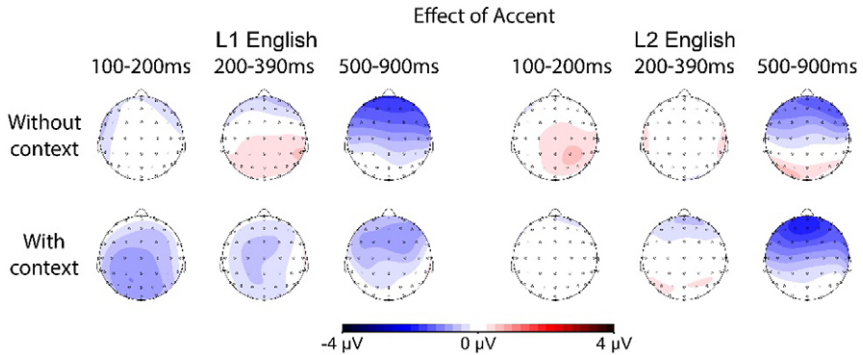


Figure 5. Topography maps for the effect of accent in the two context conditions in the three time windows 100–200 ms, 200–390 ms, and 500–900 ms. The maps from the native speakers of English (L1 English) are on the left side and the maps from the Dutch learners of English (L2 English) are on the right side. The topographies were created by subtracting the value from the unaccented objects in each context condition from the accented objects in the respective context condition.

The main effect of *Accent* in the best-fit model did suggest that the accented objects elicited more negative responses than the unaccented objects ($b = -1.208$, $t = 2.262$, $p = 0.024$), although this effect may be skewed due to the presence of interactions.

Discussion

The current ERP study examined whether native speakers of English (L1 group) and advanced Dutch learners of English (L2 group) processed English focus adjacent and nonadjacent to *only* differently and, if so, whether the L2 group employed an L1 Dutch-strategy (L1 transfer; Clahsen & Felser, 2006; Frenck-Mestre, 2002; Juffs, 2005) or an interlanguage strategy for difficult constructs (Bley-Vroman, 1983; Lee et al., 2020; Marinis et al., 2005; Purdy, 2001; Sorace, 2011; Sorace & Filiaci, 2006; White, 2011). For adjacent focus, we hypothesized that the L1 English listeners would be similar in processing to L1 Dutch listeners (Dimitrova, 2012) (Hypothesis 1a), and so will the L2 English listeners, regardless of the presence or absence of context, as a result of positive L1 transfer (Hypothesis 1b). For nonadjacent focus, we hypothesized that the L1 English listeners would show slightly different processing patterns than L1 Dutch listeners (Dimitrova, 2012), resulting in an L1-English specific strategy (Hypothesis 2a). The L2 listeners were hypothesized to experience difficulty integrating L2 information and to employ their own interlanguage strategy as a result (Hypothesis 2b), but they may show more L1 English-like processing patterns when the nonadjacent focus was preceded by the context (Hypotheses 3).

Processing of adjacent focus

Our results on the processing of adjacent focus partly support Hypothesis 1. The L1 and L2 group displayed mostly similar responses to the accented verbs. Specifically, the early anterior positivity at 100–200 ms was elicited in both groups, except that it occurred when the accented verbs were in absence of context in the L1 listeners and in presence of context in the L2 listeners. These responses were followed by the frontal-central and global positivity in the L1 and L2 listeners respectively at 200–390 ms.

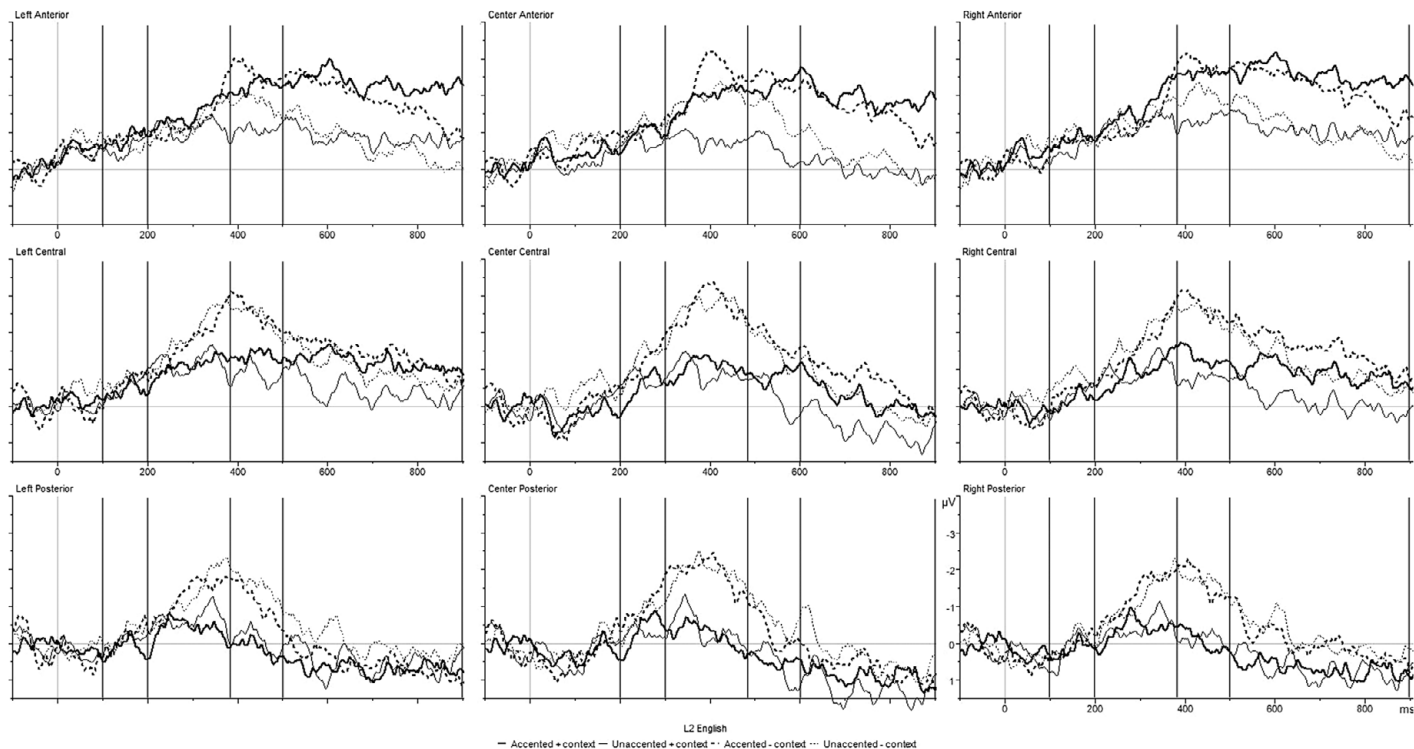


Figure 6. ERPs to the objects for the Dutch learners of English (L2 English) in the nine scalp regions. Thick solid ERP lines correspond to the accented objects with context; thick dotted curves to the accented objects without context; thin solid curves to the unaccented objects with context; and thin dotted curves to the unaccented objects without context. Gray vertical line corresponds to onset of stimulus and black vertical lines to the analyzed time-windows.

The early positivity at 100–200 ms

In accordance with our predictions, in both groups, the early positivity at 100–200 ms occurred shortly after word onset, indicating that the processing of the pitch accent was facilitated by prior expectations for accentuation, which were elicited by *only* (Bouma et al., 2007; Clahsen & Felser, 2006; Dimitrova, 2012; Foolen et al., 2009; Frenck-Mestre, 2002). However, we did not predict an effect of context. We will attempt to provide an explanation below.

First, regarding the effect of context on the L1 group, Dimitrova's (2012) study on focus expectancy in L1 Dutch has shown that native listeners of Dutch process adjacent focus earlier when presented with isolated sentences than sentences preceded by a context. Although it is unknown from previous research what effect context has on L1 English processing, we will attempt to provide an explanation for the disappearance of early processing in the L1 English group when the target sentence is embedded in context. In English, adjacent and nonadjacent focus are equally preferred (Bouma et al., 2007). When provided with alternatives to verbs and objects in the context, the L1 group may have concluded that focus can be either on adjacent or nonadjacent to *only*, resulting no expectancy effects. However, this additional information was missing in isolated sentences, in which *only* was the sole cue for upcoming contrast (Filik et al., 2009; Zimmermann & Onea, 2011). It is possible that the L1 group may have defaulted to an expectation for adjacent focus to speed up processing.

In contrast, the L2 group's early processing of the pitch accent in the presence of context may be explained as follows. They expected the pitch accent on the verb when the context provided alternatives and hinted at a contrast (Filik et al., 2009; Zimmermann & Onea, 2011) but not when *only* was the sole cue for contrast (before the verb unfolds itself) in isolated sentences. As the processing of *only* increases cognitive load (Chafe, 1994; Filik et al., 2009; Zimmermann & Onea, 2011), the L2 group may not have had the cognitive resources to expect focus, suggesting a difficulty in integrating semantic information to update the sentence meaning online (Sorace, 2011). When accented verbs were preceded by context, the possibility for contrast and alternatives were provided much earlier before the contrast was prosodically realized. It is then plausible that the L2 group used the information from the context and were then simply waiting for *only* for the exact placement of the contrast. Based on their L1 experience, that is, via positive L1 transfer, this would be the word adjacent to *only*.

Thus, partly in line with our Hypothesis 1, the L1 English listeners showed similar processing patterns to L1 Dutch listeners (Dimitrova, 2012). However, the L2 English listeners showed a processing strategy that is different from that of the L1 English listeners and L1 Dutch listeners, suggesting an interlanguage strategy that requires more information to facilitate expectations for prosodic focus in the semantic-syntactic-prosodic interface on the locus adjacent to *only*.

The positivity at 200–300 ms

As predicted, the positivity at 200–390 ms occurred over a large portion of the scalp in both groups, representing the 'accent positivity'. Both the L1 and L2 listeners, therefore, processed the prosodic information in the verbs (Dimitrova, 2012; Dimitrova et al., 2012; Heim and Alter, 2006; Lee et al., 2020). As the 'accent positivity' was comparable between the two groups, we interpreted the L2 prosodic processing to be native-like, supporting Hypothesis 1b.

The native-like processing found in our study is in contrast with what Lee et al. (2020) found, that is, a diminished 'accent positivity' in the L2 group. This discrepancy

may be due to differences in typological similarity between the L1-L2 pairings of the two studies, which has found to be of importance to L2 processing (e.g., Dussias et al., 2013; Ortega-Llebaria & Colantoni, 2014; Steinhauer, 2014). In our study, the L1 and L2 (Dutch versus English) are intonation languages with a lexical stress system and use the same prosodic parameters for focus marking. In Lee et al.'s (2020) study, the L1 and L2 (English versus Mandarin Chinese) are similar in the prosodic parameters involved but different in how they are deployed due to differences in the prosodic grammar, such as English being an intonation language and Mandarin Chinese being a tone language. The L1-L2 similarity is thus bigger in our study than in Lee et al.'s (2020) study.

Underlying mechanisms for adjacent focus processing

Our data show that, at least for adjacent focus, the L1 and L2 English listeners were very similar in prosodic focus processing. Both groups were capable of expecting prosodic cues to focus, but they differed significantly in the amount of available information they needed from the preceding context. The L1 group in our study showed similar early processing patterns to L1 Dutch listeners when presented with isolated sentences (Dimitrova, 2012), indicating that expected focus in the native language was processed similarly by the L1 English and L1 Dutch listeners when adjacent to *only*. In contrast, the L2 group were unable to expect adjacent focus upon hearing *only* in isolated sentences and required the aid in the form of alternatives to verbs and objects provided in the context. This suggests that the L2 listeners may have experienced difficulty with quickly integrating information from *only* to rapidly form focus expectations (Sorace, 2011).

We also found evidence for native-like processing of pitch accents, that is, similar processing in L1 Dutch, L1 English, and L2 English. This native-like processing can be either a result of successful learning or positive L1 transfer (Clahsen & Felser, 2006; Frenck-Mestre, 2002), which is difficult to disentangle at this location, as pitch accents adjacent to *only* are similar in acoustics and function in both languages (Filik et al., 2009; Gussenhoven, 1983; Ladd, 1980; Zimmermann & Onea, 2011).

Taken together, our findings on focus adjacent to *only* suggest that theories of L1 transfer and interlanguage processing strategies are not mutually exclusive. Rather, it appears that L2 listeners show almost simultaneous occurrence of difficulty in information integration in some aspects and native-like processing in other aspects.

Processing of the nonadjacent focus

Whether the L2 listeners were indeed successful in learning English constructs with *only* or benefited from positive L1 transfer was expected to be elucidated by data on nonadjacent accentuation, as English and Dutch differ in the preferential positioning of *only* at this location (Bouma et al., 2007; Foolen et al., 2009). However, rather unexpectedly, we did not find any effects related to the accentuation on the objects. This suggests that the L1 and L2 listeners did not expect accentuation, respond to its acoustic features, and feel the need to reinterpret the sentence. Furthermore, the presence of context did not affect prosodic processing.

A possible explanation for the lack of expectation and prosodic processing could be that the listeners were simply not sensitive to the pitch accent on the object. However, this is not likely as both English and Dutch listeners have shown to react to and interpret object-focus correctly in offline and online methods in their L1 (Dimitrova, 2012; Ge

et al., 2021b; Mulders & Szendroi, 2016). An alternative explanation could be that the listeners used information prior to the object to interpret the sentence and did not need additional prosodic information provided at the object. That is, an accented verb that preceded the object implied that the object was unaccented, and vice versa. Information on accentuation in the object may become superfluous. The listeners may thus have ignored the pitch accent and focused on other information given in the object position. That listeners can actively ignore superfluous pitch accents has previously been shown (Llanos et al., 2021). Figures 4 and 6 support this speculation: Objects appear not to be processed based on accentuation status, but rather on presence of context. This is especially evident in Figure 6 as a frontocentral negative deflection peaking around 400 ms in the L2 group for the objects in sentences without context relative to with context. We did not statistically analyze this negativity as we had predetermined our time windows, but the polarity and peak latency suggest a possible N400 effect. This component has previously been associated with detection of anomalies (Baumann & Schumacher, 2011; Osterhout & Holcomb, 1992; 1993; Osterhout et al., 1997; Schumacher & Baumann, 2010; Toepel et al., 2009; Wicha et al., 2004) and with difficulty integrating semantic information into sentence meaning based on its context (van Berkum, 2009; Brown & Hagoort, 1993; Kutas & Hillyard, 1980). As the potential N400 only occurred in the absence of context, this finding suggests that the L2 listeners may have had difficulties accepting the objects in isolated sentences. We therefore propose that, instead of prosodic processing, the listeners diverted to lexical processing instead (similar to Lee et al., 2020). However, we are unable to test this theory using the current data.

Taken together, our results do not support our hypotheses on nonadjacent prosodic focus processing and we are unable to test our proposal on possible underlying mechanisms at this point. More research is needed to explore why prosodic processing does not occur for accentuation information in the object nouns. One way to do this is to study the processing of accentuation in non-sentence-final object nouns in more complex constructions where a constituent following the object can also be the focus (e.g., *The dinosaur is only throwing the BUCKET at the tiger* versus *The dinosaur is only throwing the bucket at the TIGER*) to only to determine whether the positioning within the sentence is of influence.

Conclusions

In conclusion, the present ERP study has examined the differences and similarities in L1 and L2 processing of the prosodic-semantic interface in English sentences with *only* by native and nonnative listeners. Our results have provided evidence for native-like processing with regard to the processing of prosodically-realized focus when adjacent to *only*. However, in line with findings from previous studies (Akker & Cutler, 2003; Chen & Lai, 2012; Ge et al., 2021a; 2021b; Lee et al., 2020), L2 listeners perform nonnative-like in forming expectations for focus placement when presented with the same amount of information as L1 listeners. That is, *only* as the singular cue results in expectations for adjacent contrastive focus in native listeners, whereas L2 listeners need additional information to do. This extends the view that L2 listeners experience difficulty when integrating information from different domains in difficult and non-L1-like constructs to build expectations for the rest of the sentence (Sorace, 2011) to interfaces that involve L1-like and relatively simple constructs. Our study has thus provided first ERP evidence for interface-based processing in which similarities in domains on their own are not sufficient for native-like prosodic processing in L2.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S0272263124000019>.

Acknowledgments. We thank all the participants for their participation, and Zenghui Liu, Aelish Hart, and Nelleke Jansen for their help with testing the participants. We also thank Piet van Tuijl for his input on the statistical analyses and Chris van Run and his colleagues at the Institute for Language Sciences Labs (previously UiL-OTS Linguistics Labs) for their technical support. We also thank Karsten Steinhauer for his feedback on the data preprocessing and the anonymous reviewers for their useful comments. This research was funded by a Westerdijk Talent grant awarded to Aojun Chen by the Dutch Research Council. Portions of the (preliminary) results were reported at the International Workshop on the L1 and L2 acquisition of Information Structure (2019), the Ninth Annual Society for the Neurobiology of Language Meeting (2017), the Second Hanyang International Symposium on Phonetics and Cognitive Sciences of Language (2019), the Third Phonetics and Phonology in Europe Conference (2019), and the 35th Annual Conference on Human Sentence Processing (2021).

Competing interest. The authors declare that there were no competing interests during this research.

References

- Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition*, 6, 81–96.
- Arnhold, A., Porretta, V., Chen, A., Verstegen, S., Mok, I., & Järvikivi, J. (2020). (Mis) understanding your native language: Regional accent impedes processing of information status. *Psychon Bull Rev.*, 27, 801–808. <https://doi.org/10.3758/s13423-020-01731-w>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Baumann, S., & Schumacher, P. B. (2011). (De-)Accentuation and the processing of information status: Evidence from event-related brain potentials. *Language and Speech*, 55, 361–381. <https://doi.org/10.1177/0023830911422184>
- Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning*, 33, 1–17. <https://doi.org/10.1111/j.1467-1770.1983.tb00983.x>
- Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer [Computer program]. <http://www.praat.org/>
- Bouma, G., Hendriks, P., & Hoeksema, J. (2007). Focus particles inside prepositional phrases: A comparison of Dutch, English, and German. *Journal of Comparative German Linguistics*, 10, 1–24.
- BrainVision Analyzer (Version 2.1.2.327) [Software]. (2019). Gilching, Germany: Brain Products GmbH.
- Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5, 34–44. <https://doi.org/10.1162/jocn.1993.5.1.34>
- Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Cognitive Neuroscience and Neuropsychology*, 18, 1851–1854. <https://doi.org/10.1097/WNR.0b013e3282f1a999>
- Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago/London: University of Chicago Press.
- Chen, A., & Lai, V. (2012). Comb or coat: The role of intonation in online reference resolution in a second language. *Sound and Sounds Utrecht*, 57–68.
- Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, 10(12), 564–570. <https://doi.org/10.1016/j.tics.2006.10.002>
- Cowles, H. W., Kluender, R., Kutas, M., & Polinsky, M. (2007). Violations of information structure: An electrophysiological study of answers to wh-questions. *Brain and Language*, 102, 228–242.
- Dimitrova, D., Stowe, L., & Hoeks, J. (2015). When correction turns positive: Processing corrective prosody in Dutch. *PLoS One*, 10, e0126299. <https://doi.org/10.1371/journal.pone.0126299>
- Dimitrova, D. V. (2012). *Neural correlates of prosody and information structure* (University of Groningen). Retrieved from http://www.rug.nl/research/portal/files/2434008/DVDimitrova_thesis_20120507.pdf.
- Dimitrova, D. V., Stowe, L. A., Redeker, G., & Hoeks, J. C. J. (2010a). Focus Particles and Prosody Processing in Dutch: Evidence from ERPs. *Speech Prosody*, 979.
- Dimitrova, D. V., Stowe, L. A., Redeker, G., & Hoeks, J. C. J. (2010b). ERP Correlates of Focus Accentuation in Dutch. *Speech Prosody 2010*, 100873, 1–4.

- Dimitrova, D. V., Stowe, L. A., Redeker, G., & Hoeks, J. C. J. (2012). Less is not more: neural responses to missing and superfluous accents in context. *Journal of Cognitive Neuroscience*, 24, 2400–2418.
- Dussias, P., Valdés Kroff, J., Guzzardo Tamargo, R., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387. <https://doi.org/10.1017/S0272263112000915>
- Filik, R., Paterson, K. B., & Liversedge, S. P. (2009). The influence of only and even on online semantic interpretation. *Psychonomic Bulletin and Review*, 16, 678–683. <https://doi.org/10.3758/PBR.16.4.678>
- Fletcher, J., Grabe, E., & Warren, P. (2005). Intonational variation in four dialects of English: The high rising tune. In S.-A. Jun (Ed.), *Prosodic typology. The phonology of intonation and phrasing* (pp. 390–409). Oxford: OUP.
- Foolen, A., Gerrevink, R. Van, Hogeweg, L., & Prawiro-Atmodjo, P. (2009). The placement of focus particles in Dutch. *Linguistics in the Netherlands*, 26(April 2016), 51–63. <https://doi.org/10.1075/avt.26.06foo>
- Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In J. Altarriba & R. Herridia (Eds.), *Bilingual sentence processing*, 217–236. Elsevier.
- Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91, 1357–1392. <https://doi.org/10.1152/physrev.00006.2011>
- Ge, H., Chen, A., & Yip, V. (2021a). Comprehension of focus-to-accentuation mapping in sentence with ONLY by advanced Cantonese learners and Dutch learners of English. *Studies in Second Language Acquisition*, 43 (2021), 25–49. <https://doi.org/10.1017/S0272263120000248>.
- Ge, H., Mulders, I., Kang, X., Chen, A. & Yip, V. (2021b). Processing focus in native and non-native speakers of English: An eye-tracking study in the visual world paradigm. *Applied Psycholinguistics*, 42(4), 1057–1088. <https://doi.org/10.1017/S0142716421000230>
- Ge, H., Mulders, I., Kang, X., Chen, A., & Yip, V. (2022, September 27). Experimental materials for the study entitled “Processing focus in native and non-native speakers of English: An eye-tracking study in the visual world paradigm”. Retrieved from osf.io/k23b4
- Gussenhoven, C. (1983). *On the grammar and semantics of sentence accents*. Dordrecht: Floris.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Hagoort, P. (2003). How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *NeuroImage*, 20, S18–S29. <https://doi.org/10.1016/j.neuroimage.2003.09.013>
- Heim, S., & Alter, K. (2006) Prosodic pitch accents in language comprehension and production: ERP data and acoustic analysis. *Acta Neurobiologiae Experimentalis*, 66, 55–68.
- Hemmelmann, C., Horn, M., Süsse, T., Vollandt, R., & Weiss, S. (2005). New concepts of multiple tests and their use for evaluating high-dimensional EEG data. *Journal of Neuroscience Methods*, 142, 209–217. <https://doi.org/10.1016/j.jneumeth.2004.08.008>
- Hruska, C., Alter, K., Steinhauer, K., & Steube, A. (2000). Can wrong prosodic information be mistaken by the brain? *Journal of Cognitive Neuroscience, Suppl.* 122, E82
- Hruska, C., Alter, K., Steinhauer, K., & Steube, A. (2001). Misleading dialogues: Human’s brain reaction to prosodic information. *Paper presented at ORAGE 2001, Aix en Provence, France*
- Hruska, C., & Alter, K. (2004). Prosody in dialogues and single sentences: How prosody can influence speech perception. In Steube, A. (Ed.), *Information structure: Theoretical and empirical aspects* (pp. 211–226). Berlin: De Gruyter.
- Johnson, S. M., Clifton, C., Breen, M., & Morris J. (2003). An ERP investigation of prosodic and semantic focus. *Journal of Cognitive Neuroscience, Suppl.* 174, E157.
- Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, 21(2), 121–151. <https://doi.org/10.1191/0267658305sr2550a>
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201. <https://doi.org/10.1080/016909600386084>
- Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.
- Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.
- Lee, A., Perdomo, M., & Kaan, E. (2020). Native and second-language processing of contrastive pitch accent: An ERP study. *Second Language Research*, 36(4), 503–527. <https://doi.org/10.1177/0267658319838300>

- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Llanos, F., German, J. S., Nike Gnanateja, G., & Chandrasekaran, B. (2021). The neural processing of pitch accents in continuous speech. *Neuropsychologia*, 158, 107883. <https://doi.org/10.1016/j.neuropsychologia.2021.107883>
- Luck, S. J. (2014). A closer Look at ERPs and ERP components. In *An introduction to the event-related potential technique* (2nd ed., pp. 35–70). Bradford Books.
- Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, 54, 146–157. <https://doi.org/10.1111/psyp.12639>
- Magne, C., Astesano, C., Lacheret-Dujour, A., Morel, M., Alter, K., & Besson, M. (2005). On-line processing of “Pop-Out” words in spoken French dialogues. *Journal of Cognitive Neuroscience*, 17, 740–756
- Marinis, T., Roberts, L., Felsler, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27, 53–78. <https://doi.org/10.1017/S0272263105050035>
- Matlab (Version R2019a, 9.6.0.1072779) [Software]. (2019). Natick, Massachusetts, USA: The MathWorks, Inc.
- Michel, M., Vidon, C., Graaff, R., & Lowie, W. (2021). Language learning beyond English in the Netherlands: A fragile future? *European Journal of Applied Linguistics*, 9(1), 159–182. <https://doi.org/10.1515/eujal-2020-0020>
- Mulders, I., & Szendroi, K. (2016). Early association of prosodic focus with alien “only”: Evidence from eye movements in the visual-world paradigm. *Frontiers in Psychology*, 7, 150. <https://doi.org/10.3389/fpsyg.2016.00150>
- Presentation (Version 20.0) [Software]. (2019). Berkeley, CA, USA: Neurobehavioral Systems, Inc.
- Ortega-Llebaria, M., & Colantoni, L. (2014). L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition*, 36, 331–353. <https://doi.org/10.1017/S0272263114000011>
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806. [https://doi.org/10.1016/0749-596X\(92\)90039-Z](https://doi.org/10.1016/0749-596X(92)90039-Z)
- Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, 8, 413–437. <https://doi.org/10.1080/01690969308407584>
- Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition*, 25, 273–285. <https://doi.org/10.3758/BF03211283>
- Purdy, J. D. (2001). The comparative fallacy in studies of the L2 acquisition of unaccusative verbs. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 4, 1. <http://doi.org/10.7916/salt.v4i1.1610>
- R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- RStudio Team (2019). RStudio: Integrated development for R. Boston, MA, USA: RStudio.
- Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing neurocognitive processes reflected by P600 effects: Evidence from ERPs and neural oscillations. *PLoS ONE*, 9, e96840. <https://doi.org/10.1371/journal.pone.0096840>
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75–116.
- Reichle, R. V., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French. *Studies in Second Language Acquisition*, 36, 535–564
- Russell Lenth. (2019). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.3.5. Retrieved from <https://cran.r-project.org/package=emmeans>.
- Schumacher, P. B., & Baumann, S. (2010). Pitch accent type affects the N400 during referential processing. *Neuroreport*, 21(9), 618–622. <https://doi.org/10.1097/WNR.0b013e328339874a>
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1, 1–33. <https://doi.org/10.1075/lab.1.1.01sor>
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3), 339–368. <https://doi.org/10.1191/0267658306sr271oa>
- Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: A brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Applied Linguistics*, 35(4), 393–417. <https://doi.org/10.1093/applin/amu028>

- Stolterfoht, B., Friederici, A. D., Alter, K., & Steube, A. (2007). Processing focus structure and implicit prosody during reading: Differential ERP effects. *Cognition*, *104*, 565–590. <https://doi.org/10.1016/j.cognition.2006.08.001>
- Toepel, U., Pannekamp, A., & van der Meer, E. (2009). Fishing for information: The interpretation of focus in dialogs. In Alter, K., Horne, M., Lindgren, M., Roll, M., Torkildsen, J. (Eds.), *Papers from BrainTalk* (pp. 176–198). Lund: Lund University.
- Trommelen, M., & Zonneveld, W. (1999). Word stress in Western Germanic languages. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 478–515). Berlin: Mouton de Gruyter.
- van Berkum, J. J. A. (2009). The neuropragmatics of “simple” utterance comprehension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory* (pp. 276–316). London: Palgrave.
- White, L. (2011). Second language acquisition at the interfaces. *Lingua*, *121*, 577–590. <https://doi.org/10.1016/j.lingua.2010.05.005>
- Wicha, N. Y. Y., Morena, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*, 1272–1288. <http://doi.org/10.1162/0898929041920487>
- Zimmermann, M., & Onea, E. (2011). Focus marking and focus interpretation. *Lingua*, *121*, 1651–1670. <https://doi.org/10.1016/j.lingua.2011.06.002>

Cite this article: Ganga, R., Ge, H., Struiksmā, M. E., Yip, V., & Chen, A. (2024). Prosodic processing in sentences with ‘only’ in L1 and L2 English. *Studies in Second Language Acquisition*, 1–26. <https://doi.org/10.1017/S0272263124000019>