

**FOREIGN-ACCENTED SPEECH MODULATES LINGUISTIC  
ANTICIPATORY PROCESSES**

**Carlos Romero-Rivas<sup>a,\*</sup>, Clara D. Martin<sup>b,c</sup>, Albert Costa<sup>a,d</sup>**

<sup>a</sup> Center for Brain and Cognition, Universitat Pompeu Fabra, Barcelona, Spain

<sup>b</sup> BCBL. Basque Center on Cognition, Brain and Language, San Sebastian, Spain

<sup>c</sup> IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

<sup>d</sup> Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

\* Corresponding author at: Center for Brain and Cognition, Universitat Pompeu Fabra.

Carrer de Tànger, 122, 08018 Barcelona, Spain. Phone number: +34 93 542 29 63

*E-mail address:* [romeriv@gmail.com](mailto:romeriv@gmail.com) (Carlos Romero-Rivas)

## **ABSTRACT**

Listeners are able to anticipate upcoming words during sentence comprehension, and, as a result, they also pre-activate semantically related words. In the present study, we aim at exploring whether these anticipatory processes are modulated by indexical properties of the speakers, such as a speaker's accent. Event-related brain potentials were obtained while native speakers of Spanish listened to native (Experiment 1) or foreign-accented speakers (Experiment 2) of Spanish producing highly constrained sentences. The sentences ended in: (1) the highest cloze probability completion, (2) a word semantically related to the expected ending, or (3) a word with no semantic overlap with the expected ending. In Experiment 1, we observed smaller N400 mean amplitudes for the semantically related words as compared to the words with no semantic overlap, replicating previous findings. In Experiment 2, we observed no difference in integrating semantically related and unrelated words when listening to accented speech. These results suggest that linguistic anticipatory processes are affected by indexical properties of the speakers, such as the speaker's accent.

## 1. Introduction

Comprehenders actively anticipate upcoming linguistic information when reading or listening to a sentence, thus predicting upcoming words (or discourse topics) that are likely to appear in the message (e.g., DeLong, Urbach, & Kutas, 2005). Crucially, comprehenders do not only anticipate the most likely lexical item, but also words with overlapping semantic characteristics (Ferdermeier & Kutas, 1999). This predictive mechanism reduces processing load and helps interlocutors to free resources in order to plan their utterances during a conversation, hence smoothing communication (for further discussion, see Pickering & Garrod, 2007). Nevertheless, we know little about how indexical properties of speakers interact with anticipation processes. In this study, we aim at exploring whether such anticipatory processes are modulated by contextual factors, such as indexical properties of the speakers, in our case the speaker's accent. This question is important at the theoretical level, because it will reveal whether anticipation during sentence comprehension is modulated by indexical properties of speakers, or whether it is so strongly anchored in the speech comprehension system that it is not modulated by such indexical properties. This study will also add important information to the field of speech comprehension, by showing how listeners deal with foreign accents during lexical and semantic integration in sentence comprehension.

Interactions with foreign-accented speakers are becoming frequent due to global mobility phenomena and the increasing interest in foreign language learning. Importantly, foreign-accented speech deviates in several ways from native speech. For instance, foreign-accented speech often involves non canonical and variant pronunciations (Nissen, Dromey, & Wheeler, 2007; Wade, Jongman, & Sereno, 2007; Wester, Gilbers, & Lowie, 2007; Hanulíková & Weber, 2012), as well as unusual prosody patterns (Gut, 2012). Since one of the purposes of predicting upcoming information during speech comprehension might be to facilitate communication with interlocutors (Pickering & Garrod, 2007), it is reasonable to assume that such processes would be at play when interacting with foreign-accent speakers. However, it is also possible that difficulties in processing accented speech at the acoustic/phonological level might hamper listeners' anticipatory processes.

Thus, the present study will explore anticipation processes during sentence comprehension when listening to accented speech. As far as we know, this is the first attempt to study the permeability of such anticipatory processes to indexical properties of speakers. Nevertheless, it is important to note that there is already some evidence suggesting that lexical-semantic processing differs during native and foreign-accented speech comprehension. Goslin, Duffy, and Floccia (2012) explored lexical processing during native and foreign-accented speech comprehension, by presenting listeners with low-cloze probability<sup>1</sup> sentences (“His wife managed to win a bag of goldfish”). They observed that the final words of sentences produced by foreign-accented speakers elicited smaller N400 amplitudes compared to the native speech condition. Goslin et al. (2012) concluded that foreign-accented speech could not be fully normalized during pre-lexical processing levels, therefore influencing later steps of language processing such as lexical access and integration. In other words, listeners would rely on top-down contextual cues in order to normalize the signal during foreign-accented lexical processing, narrowing the possible lexical candidates (Goslin et al., 2012).

Romero-Rivas, Martin, and Costa (2015) also explored sentence comprehension during foreign-accented speech comprehension. We observed that, during native speech comprehension, semantic violations in the critical words (“My favorite breakfast is a toast with marmalade and a coffee/hospital with a lot of milk”) elicited an N400 effect followed by a P600 effect. We took the P600 effect as suggesting that listeners carried out some sort of semantic reanalysis processing after the semantic violations (see e.g., Van Petten & Luka, 2012). However, during foreign-accented speech comprehension, semantic violations only elicited an N400 effect, more widely distributed than during native speech comprehension<sup>2</sup>. The wider N400 effect and the lack of a P600 effect after semantic violations were interpreted as suggesting that listeners were unable to carry out later semantic re-analysis processes (trying to make sense of a semantic violation), probably because of the higher cognitive demands on lexical processing

---

<sup>1</sup> The cloze probability of a given word is the proportion of individuals who provide that particular word as the most likely completion for that sentence fragment in a paper and pencil test (Taylor, 1953). It has been repeatedly shown that the lower cloze probability of a word, the higher N400 amplitude (e.g., N400 mean amplitude larger for “There was nothing wrong with the car” as compared to “He mailed the letter without a stamp”; Kutas & Hillyard, 1984).

<sup>2</sup> Similar results can be observed in Hanulíková, van Alphen, van Goch, and Weber’s (2012) Figure 3, although the authors did not perform any analysis on the P600 time window for semantic violations processing.

associated to foreign-accented speech comprehension (in order to adapt to foreign-accented speech).

Somewhat relatedly, other types of indexical properties of speakers, such as age or gender (e.g., “I always drink a glass of wine during dinner”, uttered by a children), have also been observed to interact with semantic congruity processing. For instance, Van Berkum, van den Brink, Tesink, Kos, and Hagoort (2008) observed an N400 effect for speaker inconsistencies of similar latency and topography to that observed for semantic incongruities. This result suggests that social and semantic information interact during the construction of meaning (Van Berkum et al., 2008).

In sum, a few ERP studies have revealed that indexical properties of speech, such as a foreign accent, modulate lexical and semantic processing. However, an important question remains unanswered: do listeners anticipate upcoming words when listening to foreign-accented speakers, in order to facilitate lexical processing? In the present study, we will explore this issue by presenting listeners with highly constrained sentences uttered by native or foreign-accented speakers. If anticipation serves to improve communication (Pickering & Garrod, 2007), then we would expect that listeners anticipate upcoming words when listening to both native and foreign-accented speakers. This would reveal that sentences’ best completions are anticipated independently of indexical properties of speakers, such as a native/foreign accent. On the other hand, since foreign-accented speech modulates the availability of lexical and semantic information during sentence comprehension (Goslin et al., 2012; Romero-Rivas et al., 2015), it could be that anticipatory processes differ between native and foreign-accented speech comprehension. If listeners rely to a higher extent on top-down cues in order to normalize the speech signal during foreign-accented lexical processing (Goslin et al., 2012), we would expect narrowed anticipation effects during foreign-accented compared to native speech comprehension.

Hence, going a step further on anticipation capacities during sentence comprehension, a second question of interest is whether listeners are able to pre-activate

other words semantically related to the expected ones when listening to foreign-accented speech, as they do during native speech comprehension (Federmeier, McLennan, De Ochoa, & Kutas, 2002). As we mentioned above, if listeners narrow the possible lexical candidates during foreign-accented speech comprehension (Goslin et al., 2012), then we would expect that only sentences' best completions are anticipated when listening to foreign-accented speakers. We will address this issue following the rationale put forward in Federmeier and Kutas' (1999) study. These authors observed that unexpected words that are semantically related to an expected word are processed more easily than unexpected words that do not have overlapping semantic features with the expected word. More concretely, in Federmeier and Kutas' (1999) study, participants were presented with highly constrained sentences ("They wanted to make the hotel look more like a tropical resort. So along the driveway they planted rows of ..."). There were three possible types of sentence endings: a) the most expected word ("palms"); b) an unexpected within-category word ("pines"); or c) an unexpected between-category word ("tulips"). As predicted, expected words elicited smaller N400 amplitudes than the other two conditions, revealing that expected words are more easily integrated than unexpected words. Also, and crucially, within-category words elicited smaller N400 amplitudes than between-category words. Federmeier and Kutas (1999) interpreted this gradation as suggesting that the ease with which a word (even if it is not expected) is accessed depends on its semantic similarity to the word expected in a given context (see also Federmeier et al., 2002, for an extension in the auditory modality; Thornhill & Van Petten, 2012).

Thus, based on the assumption that lexical items are grouped in long-term memory on the basis of shared perceptual and functional attributes (Kay, 1971; Rosch, 1973, 1975; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1979), the pre-activation of expected words would facilitate the processing of words that are semantically related, even if they are unexpected. In the present study, we will explore whether listeners also pre-activate semantically related words in the context of foreign-accented speech. If, as we advanced before, listeners narrow the possible lexical candidates during foreign-accented speech comprehension in order to normalize the speech signal (Goslin et al., 2012), then only the sentences' best completions might benefit from anticipatory processes. Therefore, semantically related and unrelated words should be processed similarly, being more difficult to integrate than expected words.

### *The present study*

This study contains two experiments. Experiment 1 serves as a baseline, where we aim at validating our procedure and materials. In this experiment, we want to replicate Federmeier et al.'s (2002) observations: a graded N400 amplitude for words depending on their expectancy and semantic overlap. We will do so by testing Spanish natives, listening to Spanish sentences uttered by native speakers of Spanish.

In Experiment 2, we will explore lexical and semantic anticipatory processes during foreign-accented speech comprehension. Spanish native participants will be presented with the same sentences used in Experiment 1, but uttered by foreign-accented speakers of Spanish. Note that the two studies are considered and analyzed independently because they were run on two different groups of participants, due to the large number of observations per condition needed in an ERP experiment.

In both experiments, participants will be presented with highly constrained sentences ending in three different words: Best Completion (word with the highest cloze probability in the sentence context); Related (low cloze probability word, plausible option, semantically related to the expected word); or Unrelated (low cloze probability word, plausible option, with no semantic relationship to the expected word) [e.g., “Tengo que ir a la biblioteca para devolver un libro/diccionario/portátil” (“*I have to go to the library in order to return a book/dictionary/laptop*”)]; for more examples, please refer to **Table 1**]. Comparing Best Completion and Unrelated words will reveal the effect of word anticipation during sentence comprehension. On the other hand, comparing Related and Unrelated words will characterize differences in processing unexpected words that have (or not) semantic overlap with the expected words.

PLEASE INSERT TABLE 1 ABOUT HERE

Electroencephalographic (EEG) responses will be recorded and ERPs time-locked to the critical word of each sentence will be explored. We will focus our analysis on the N400 and late negativity components. First, we will establish an early time

window for the auditory N400 (240-450 ms) based on previous literature (e.g., FitzPatrick & Indefrey, 2014). Second, a later time window will be established to explore the late negativity (500-900 ms; see e.g., Hahne & Friederici, 2001). Importantly, previous literature indicates that semantic incongruity effects start earlier and last longer in the auditory modality as opposed to the visual modality (Holcomb & Neville, 1990). For instance, the N400 semantic incongruity effect can last up to 700 ms after the onset of the critical word (Hendrickson, Walenski, Friend, & Love, 2015) or even longer (Van Petten, Coulson, Rubin, Plante, & Parks, 1999). This late negativity for semantic incongruities has been interpreted by some authors as an effect indicating that comprehenders are engaged in trying to integrate words for a longer period of time (Kutas & Kluender, 1991), and more concretely, integrating the semantic properties of words after lexical access (Hahne & Friederici, 2001; Romero-Rivas, Corey, Garcia, Thierry, Martin, & Costa, 2016).

In Experiment 1, we expect to find a graded effect of semantic appropriateness on the N400 ERP amplitudes. The more appropriate the word for a given context, the smaller the amplitude, replicating Federmeier et al. (2002). In other words, we expect the N400 amplitudes to be smaller for Best Completions than for Related words, and smaller for Related than for Unrelated words. Regarding the late negativity, this component is supposed to index continued semantic integration into the previous context, after lexical access (Hahne & Friederici, 2001; Romero-Rivas et al. 2016). Therefore, we would expect that by this late time window Best Completions and Related words were similarly integrated into the previous context, because of their semantic overlap. However, Unrelated words might still need more time to be integrated into the previous context, thus eliciting a larger late negativity than the other two conditions.

Experiment 2 is the crucial one, in which the sentences were presented in foreign-accented speech. As pointed out above, if anticipation serves to improve communication (Pickering & Garrod, 2007), then we would expect that listeners anticipate upcoming words when listening to foreign-accented speakers. This should be reflected in larger N400 amplitudes for unexpected than expected words during foreign-accented speech comprehension. Even more, if listeners narrow the possible lexical



candidates during foreign-accented speech comprehension in order to normalize the speech signal (Goslin et al., 2012), then we would expect that Related words would not have any processing advantage over Unrelated words, because of the over-anticipation of the Best Completions. This might be indexed by similar N400 and late negativity amplitudes for Related and Unrelated words.

## **2. Experiment 1: native speech comprehension**

The objective of Experiment 1 was to replicate Federmeier et al.'s (1999, 2002) observations on the ease of accessing words that are semantically related to the expected ones, using Spanish materials. This first step will establish a benchmark for interpreting the results regarding the comprehension of foreign-accented Spanish.

### *2.1 Methods*

#### *2.1.1 Participants*

Seventeen participants (8 women, all right handed, mean age = 22.8 years, range = 18-27 years) took part in Experiment 1. All participants were native speakers of Spanish. None of them reported any hearing or neurological impairments. Participation in the study was remunerated (10€/h). Before the beginning of the experiment, participants gave their written informed consent.

#### *2.1.2 Materials*

The experimental stimuli consisted of a set of 165 sentences. Each sentence was recorded three times: a version ending with the best-completion word in the cloze probability test, a version ending with a semantically related and plausible word, and a version ending with a semantically unrelated but plausible word (resulting in 495 sentences). Thus, critical words always appeared at the end of sentences, and were preceded by the exact same sentence context. Critical words were always in the final

position of the sentences in order to explore late ERP modulations without contamination of next word presentation. Additionally, since the cloze probability effect (indexed by the N400) is building up across sentence listening, effects should be maximized in the final position (Halgren et al., 2002). The three possible endings for each sentence context began with a different phoneme. All critical words were matched for phonological length ( $F(2,328) = 0.62, p = .34$ ), and word frequency ( $F(2,328) = 0.44, p = .65$ ; see **Table 4**; values extracted from BuscaPalabras (Davis & Perea, 2005), a software based on LEXESP (Sebastián-Gallés, Martí, Cuetos, & Carreiras, 2000), which is a frequency database based on approximately 5 million Spanish words).

PLEASE INSERT TABLE 4 ABOUT HERE

The 165 experimental sentences were selected based on a preliminary cloze probability test. Thirty native speakers of Spanish (24 women) participated in this pre-test (none of them took part later in the ERP experiment). Participants were asked to report the most likely completion for 266 sentence contexts. For stimuli selection we only chose those sentence contexts having a highly predictable best completion word (cloze probability equal or higher than .65; best completions:  $M = 86.67, SD = 10.77$ ; semantically related words:  $M = 2.58, SD = 0.06$ ; unrelated words:  $M = 0.71, SD = 0.03$ ). In contrast to Federmeier et al. (2002) we did not use a control condition with low cloze probability sentences, because we were mainly interested in exploring the processing of expected words and words semantically related to the expected ones (for similar cases in which only high cloze probability sentences were used see e.g., Wicha, Moreno, & Kutas, 2003, 2004). We did not have any constraints concerning the preceding sentence contexts and no specific matching was needed since those contexts were strictly identical in each of the 3 conditions (counterbalanced across participants).

In Experiment 1, sentences were uttered by four female and four male Spanish native speakers, being recorded and edited with Audacity (© Audacity Team), at 44.1 kHz, 32 bits and stereo sound. Each speaker received a list containing the experimental sentences (in a randomized order) in the three versions (ending with the expected, related or unrelated word). Each speaker recorded 1/8 of the sentences (each triplet was

recorded by the same speaker). They were asked to utter each sentence with neutral prosody. Three experimental lists were created, each of them containing only one version of the 165 experimental sentences (55 sentences for each condition). This way, each participant listened to all speakers and sentence contexts, and conditions were counterbalanced across participants.

We carried out three separate t-tests comparing the duration of the final word of the sentences across conditions. T-test comparisons revealed significant shorter durations for Best Completion ( $M: 445.93$  ms;  $SD: 9.74$ ) than Related ( $M: 491.53$  ms;  $SD: 9.25$ ) words ( $t(164) = 3.85$ ;  $p < .001$ ) and Unrelated ( $M: 493.50$  ms;  $SD: 10.21$ ) words ( $t(164) = 4.03$ ;  $p < .001$ ). However, the t-test comparison between Related and Unrelated words was not significant ( $t(164) = 0.87$ ;  $p = .16$ ). Importantly, there is ample evidence showing that predictability has an effect on reading times (see e.g., Smith & Levi, 2013). Regarding sentence context durations, t-test comparisons revealed no significant differences between conditions (Best Completion:  $M: 3227.16$  ms,  $SD: 445.93$ ; Related:  $M: 3231.55$  ms,  $SD: 491.53$ ; Unrelated:  $M: 3200.76$  ms,  $SD: 493.5$ ).

### 2.1.3 *Experimental procedure*

Participants were seated in front of a computer screen, in a sound-proof room. They were asked to listen carefully in order to comprehend all the sentences during a passive listening task. We did not provide any information about the speakers or their accents, and only told the participants that they would be listening to people speaking in an everyday context. The experiment was run on E-Prime 2.0. Sentences were presented binaurally at a constant sound level via headphones. Each trial started with a fixation cross, presented 500 ms before the onset of each sentence and remained there until 500 ms after sentence offset. Participants were asked to look at the fixation cross and to avoid blinking throughout the auditory sentence presentation. Between each sentence presentation, a blank screen was presented during 3000 ms, where participants were allowed to blink. Sentence order was randomized between participants. Three pauses were programmed during the experimental session, in order to let participants have a

rest if they needed. During these pauses, pressing the space bar resumed the experiment. The whole experiment lasted approximately 40 minutes.

#### 2.1.4 EEG recording

The EEG signal was recorded from 31 electrodes (impedances kept below 5 k $\Omega$ ) mounted on an elastic cap, at standard 10-20 locations. The on-line reference electrode was attached to the tip of the nose, and the signal was re-referenced off-line to the mastoid average. Lateral eye movements were recorded with an electrode beside the right eye, and eye blinks were recorded with another electrode below the right eye. EEG signal was filtered on-line with a 0.1-100 Hz bandpass filter and digitized at 500 Hz.

#### 2.1.5 ERP analyses

EEG epochs were determined for the last word of each sentence. We extracted the epochs from 200 ms before to 1200 ms after the onset of the critical (final) word of each sentence. EEG waveforms were baseline corrected to a 200 ms pre-stimulus onset, and averaged per participant and condition. Mean amplitudes in specific time windows were analyzed with repeated measures ANOVAs, analyzing three regions: frontal (F3, Fz, F4, FC1, FC2, FC5, and FC6), central (C3, Cz, C4, CP1, CP2, CP5 and CP6), and posterior (P3, Pz, P4, PO1, PO2, O1 and O2).

Statistical analyses were performed on two main time windows: 250-450 ms, and 500-900 ms. All analyses were repeated measures ANOVAs, and included the factors *topography* (frontal, central, posterior) and *semantic status* (Best Completion, Related, Unrelated). All effects and interactions were corrected for sphericity using the Greenhouse-Geisser correction (degrees of freedom are reported without the G-G correction). In addition, we used the Bonferroni correction for post-hoc analyses.

## 2.2 Results

### 2.2.1 250-450 ms

We obtained significant effects of *topography* ( $F(2,32) = 7,26; p < .05$ ), and *semantic status* ( $F(2,32) = 22,02; p < .001$ ). In addition, we also observed a significant interaction between the two factors ( $F(4,64) = 14,99; p < .001$ ). Post-hoc tests revealed that Best Completion elicited an attenuated N400 amplitude compared to Related and Unrelated, over the three topographic regions (**Figures 1 and 3**). Moreover, Related elicited a decreased N400 amplitude compared to Unrelated over central and posterior regions (statistical values are presented in **Table 2**). Finally, the mean amplitudes over the posterior region of the scalp were significantly more positive than over the frontal and central regions.

PLEASE INSERT FIGURE 1 AND 3 AND TABLE 2 ABOUT HERE

### 2.2.2 500-900 ms

We obtained a significant effect of *topography* ( $F(2,32) = 16,72; p < .001$ ). In addition, we also observed a significant interaction between *topography* and *semantic status* ( $F(4,64) = 4,53; p < .05$ ). Post-hoc tests revealed that Best Completion and Related elicited attenuated late negativity amplitudes compared to Unrelated over central and posterior regions. Moreover, the late negativity was similar in amplitude for Best Completion and Related (statistical values are presented in **Table 2**). Finally, there were significant differences between the three topographic regions (most negative mean amplitudes over the frontal region, and most positive mean amplitudes over the posterior region).

To summarize, in Experiment 1 we replicate previous findings on the ease of accessing words semantically related to expected words during speech comprehension (Federmeier et al., 2002) in a native speech context. In addition, we also contribute by showing that by the later time window (500-900 ms) the differences between Best Completion and Related words disappear, while Unrelated words keep eliciting a more negative amplitude. These results will be further discussed below.

### **3. Experiment 2: foreign-accented speech comprehension**

The objective of Experiment 2 is to explore the impact of foreign-accented speech on lexical and semantic processing of words embedded in highly constrained sentences. Two questions are important here: whether listeners are able to anticipate upcoming words during foreign-accented speech comprehension and, if they do so, whether semantically related words also benefit from these anticipatory processes, being easier to integrate than semantically unrelated words.

#### *3.1 Method*

##### *3.1.1 Participants*

Seventeen participants (9 women, all right handed, mean age = 22.8 years, range = 20-27 years) took part in experiment 2. All participants were native speakers of Spanish. None of them took part in Experiment 1, nor reported any hearing or neurological impairments. Participation in the study was remunerated (10€/h). Before the beginning of the experiment, subjects gave their written informed consent.

##### *3.1.2 Materials*

The experimental stimuli consisted of the same set of 165 sentences used in Experiment 1. In Experiment 2, sentences were recorded by eight foreign accented speakers of Spanish (4 Italians (2 women), and 4 French (2 women)). The decision to use these speakers was rooted in the aim to test the main effect of foreign-accented speech, independently of the native language of the foreign speakers and the similarities between Spanish and those other languages. As we did for Experiment 1, three experimental lists were created, each of them containing only one version of the 165 experimental sentences. Foreign-accented speakers were asked to utter each sentence with a neutral prosody. They were presented with native accented versions of the sentences before their recordings, in order to minimize possible differences in speech

rate and prosody. Foreign-accented recordings did not contain mispronunciations that could lead to word/sentence misinterpretation. Thus, the main difference between native and foreign-accented sentences was on canonical versus non canonical phonetic information (that listeners may treat as allophonic variations, that is, context-sensitive phonetic variants of phonemes; see e.g., Samuel & Larraza, 2015, for further discussion on some circumstances under which listeners may treat accented non-canonical phonetic information as allophonic variants for the purpose of lexical access).

We carried out three separate t-tests comparing the duration of the final word of the sentences across conditions. T-test comparisons revealed significantly shorter durations for Best Completion ( $M$ : 479.76 ms;  $SD$ : 11.25) compared to Related ( $M$ : 515.65 ms;  $SD$ : 11.72) words ( $t(164) = 2.66$ ;  $p < .05$ ) or compared to Unrelated ( $M$ : 540.13 ms;  $SD$ : 11.04) words ( $t(164) = 4.37$ ;  $p < .001$ ). However, the t-test comparison between Related and Unrelated words was not significant ( $t(164) = 1.83$ ;  $p = .07$ ). Importantly, the directions of these comparisons were along the same lines as in Experiment 1. Regarding sentence context durations, t-test comparisons revealed significant differences between the three conditions: Best Completion ( $M$ : 4194.64 ms,  $SD$ : 479.76) vs. Related ( $M$ : 4080.39 ms,  $SD$ : 515.65) ( $t(164) = 4.02$ ;  $p < .001$ ); Best Completion vs. Unrelated ( $M$ : 3998.8 ms,  $SD$ : 540.13) ( $t(164) = 6.78$ ;  $p < .001$ ); and Related vs Unrelated ( $t(164) = 3.55$ ;  $p < .001$ ).

In addition, we carried out two separate t-tests comparing sentence and final-word duration across Experiments. Both t-test comparisons revealed significant differences across native and foreign-accented speech (Sentence duration:  $M$  for native speech = 3219 ms,  $SD$ : 117.32;  $M$  for foreign-accented speech = 4091 ms,  $SD$ : 157.12;  $t(494) = 17.15$ ,  $p < .001$ ; Final words:  $M$  for native speech = 477 ms,  $SD$ : 11.95;  $M$  for foreign-accented speech = 512 ms,  $SD$ : 14.21;  $t(494) = 6.83$ ,  $p < .001$ ). Importantly, Van den Brink, Brown, and Hagoort (2006) showed that early or late isolation points of words do not affect the onset of the N400 ERP. Also, following Goslin et al. (2012), no attempt was made to control or adjust the temporal features of the stimuli, since longer productions are an inherent part of foreign-accented speech.

Accent strength and intelligibility of the native (Experiment 1) and foreign-accented (Experiment 2) speakers were rated by an independent sample of 18 native

speakers of Spanish (12 women, mean age = 23.4 years, range = 21-27 years). These pre-tests were run in order to ensure that native and foreign-accented speakers were perceived differently, and that, beyond this difference, they were all comprehensible. Participants carried out two tasks. During the first task, they had to listen to the experimental sentences of Experiments 1 and 2 and rate them from 1 (native accent) to 5 (the speaker has a very strong foreign accent). For the second task, subjects had to write down the final word of each sentence (comprehension task). Regarding the first task, because the scoring of some native speakers had no variance, we carried out a general repeated measures ANOVA including the within subject factors Accent (native, foreign) and Condition (best completion, related, unrelated). A significant effect of Accent was obtained ( $F(1,17) = 765,81; p < .001$ ), revealing that foreign speakers' accents ( $M = 3.63, SE = 0.18$ ) were evaluated as stronger than native speakers' accent ( $M = 1.12, SE = 0.08$ ). Regarding the second task, participants recognized the last word of the sentences one hundred per cent of the times both for the native and for the foreign accented speakers, and did not report any difficulties in understanding the sentences. Based on this pre-test, we were confident that native and foreign-accented speakers were perceived differently, although all of them were intelligible<sup>3</sup>.

The experimental procedure, EEG recordings, and ERP analyses were carried in the same way as during Experiment 1.

### 3.2 Results

#### 3.2.1 250-450 ms

We obtained a significant effect of *semantic status* ( $F(2,32) = 8,85; p < .01$ ), and a significant interaction between *topography* and *semantic status* ( $F(4,64) = 7,79; p < .01$ ). Post-hoc tests revealed that Best Completion elicited a decreased N400 amplitude

---

<sup>3</sup> As a secondary objective, we explored whether accent type and strength correlated with the ERP modulations. However, due to the limitations of the current design and the typical constraints of EEG experiments (small amount of epochs per condition considering this sub-division), we did not obtain reliable results. In addition, a preliminary inspection of the effect of first language (Italian vs. French) showed very similar ERP responses during the comprehension of both accents.



compared to Related and Unrelated over central and posterior regions. However, Related and Unrelated words elicited similar N400 amplitudes over the entire scalp (statistical values are presented in **Table 3**, see also **Figure 1**). Finally, the mean amplitudes over the posterior region of the scalp were significantly more positive than over the frontal and central regions.

PLEASE INSERT TABLE 3 ABOUT HERE

### 3.2.2 500-900 ms

We obtained significant effects of *topography* ( $F(2,32) = 41,81; p < .001$ ), and *semantic status* ( $F(2,32) = 6,68; p < .01$ ). In addition, we also observed a significant interaction between the two factors ( $F(4,64) = 6,64; p < .01$ ). Post-hoc tests revealed that Best Completion elicited attenuated N400 late negativity amplitudes compared to Related and Unrelated over central and posterior regions. However, Related and Unrelated words elicited similar late negativity amplitudes over the entire scalp (statistical values are presented in **Table 3**). The post-hoc analyses for Topography revealed that the mean amplitudes over the posterior region of the scalp were significantly more positive than over the frontal and central regions.

In brief, our first observation in Experiment 2 is that during foreign-accented speech comprehension, sentences' Best Completions elicited attenuated N400 and late negativity amplitudes as compared to Related and Unrelated words. Second, we did not observe significant differences in the processing of Related and Unrelated words in any time window. Results will be further discussed below.

## 4. Comparisons of Experiments 1 and 2

Although we are aware that the between subjects design may create some problems when comparing the results of Experiments 1 and 2, we carried out an exploratory analysis in which we compared the two datasets. More concretely, we explored whether the lexical anticipation effect (that is, the difference waveform after subtracting Best Completion from Unrelated words) and the semantic relatedness effect (the difference waveform after subtracting Related from Unrelated words) differed across accent conditions. Comparing Best Completion and Unrelated words will reveal the effect of word anticipation during sentence comprehension. On the other hand, comparing Related and Unrelated words will characterize differences in processing unexpected words that have (or not) semantic overlap with the expected words (Federmeier & Kutas, 1999; Federmeier et al., 2002).

We conducted two repeated measures ANOVAs (one for each time window) for each effect (lexical anticipation effect and semantic relatedness effect), including the factors *topography* (frontal, central, posterior), and *accent* (Native – Experiment 1, Foreign – Experiment 2). In addition, we compared each difference wave against zero, in order to explore whether each effect was independently significant. For this purpose, we carried out one-sample t-tests for each accent and time window, over the posterior topographic region, which was the one showing the strongest effects.

#### *4.1 Lexical anticipation effect (Unrelated – Best Completion)*

In the 250-450 ms time window, we obtained a significant effect of *topography* ( $F(2,64) = 31.4; p < .001$ ). The magnitude of the N400 effect was largest over central and posterior regions. The effect of *accent* ( $F(1,32) = 2.53; p = .12$ ), and the interaction between the two factors ( $F(2,64) = 0.38; p = .55$ ) were not significant.

In addition, the lexical anticipation effect was significantly different from zero both for native ( $t(1,16) = -7.03; p < .001$ ) and foreign-accented speech comprehension ( $t(1,16) = -4.37; p < .001$ ).

In the 500-900 ms time window, we also obtained a significant effect of *topography* ( $F(2,64) = 16.13; p < .001$ ). The magnitude of the late negativity effect was

largest over central and posterior regions. The effect of *accent* ( $F(1,32) = 0.71; p = .41$ ), and the interaction between the two factors ( $F(2,64) = 0.66; p = .44$ ) were not significant.

Also, the lexical anticipation effect was significantly different from zero both for native ( $t(1,16) = -3.03; p < .01$ ) and foreign-accented speech comprehension ( $t(1,16) = -5.05; p < .001$ ).

In sum, the lexical anticipation effect was significantly different from zero, in the two time windows, for both native and foreign-accented speech comprehension. Critically, the effect size did not differ between native and foreign-accented speech comprehension.

#### 4.2 Semantic relatedness effect (*Unrelated – Related*)

In the 250-450 ms time window, we observed a significant effect of *topography* ( $F(2,64) = 9.78; p < .01$ ), showing that the magnitude of the semantic relatedness N400 effect was largest over central and posterior regions. Importantly, we also observed a significant interaction between *topography* and *accent* ( $F(2,64) = 4.29; p < .05$ ). Post-hoc tests showed that the magnitude of the semantic relatedness N400 effect was larger in native than foreign-accented speech comprehension over the posterior region of the scalp ( $t(64) = 2.03; p < .05$ ; see **Figures 2 and 3**), but not over the frontal ( $t(64) = 0.16; p = .8$ ) or central ( $t(64) = 1.13; p = .26$ ) regions.

PLEASE INSERT FIGURE 2 ABOUT HERE

Also, the semantic relatedness effect was significantly different from zero in this time window for native speech comprehension ( $t(1,16) = -4.88; p < .001$ ). Nevertheless, the effect was not significantly different from zero for foreign-accented speech comprehension ( $t(1,16) = -1.47; p = .16$ ).

In the 500-900 ms time window, we observed a significant effect of *topography* ( $F(2,64) = 7,29; p < .01$ ), showing that the magnitude of the semantic relatedness late negativity effect was largest over central and posterior regions. The effect of *accent* ( $F(1,32) = 1,10; p = .30$ ) and the interaction between the two factors ( $F(2,64) = 1,11; p = .34$ ) were not significant.

In addition, as in the previous time window, the semantic relatedness effect was significantly different from zero in this time window for native speech comprehension ( $t(1,16) = -3.62; p < .01$ ). However, the effect was not significantly different from zero for foreign-accented speech comprehension ( $t(1,16) = -1.62; p = .13$ ).

In sum, the semantic relatedness effect was significantly different from zero in both time-windows, for native speech comprehension. Crucially, it was not significantly different from zero for foreign-accented speech comprehension. In addition, the size of the semantic relatedness effect was different between the two accent contexts (even if the difference did not reach significance in the late negativity time window).

These results are further discussed below.

## 5. Discussion

In the present study, we explored whether listeners are able to anticipate upcoming words when listening to foreign-accented speakers. In addition, we also explored whether listeners are able to anticipate not only the expected words, but also other words with overlapping semantic characteristics. These issues were assessed by studying the modulation of the N400 and late negativity ERPs in different conditions. Participants were presented with highly constrained sentences, in which final words were of different gradations in terms of semantic congruity (the best completion, an unexpected and semantically related word, and an unexpected and semantically

unrelated word). In Experiment 1, participants were presented with sentences uttered by native speakers of Spanish. In Experiment 2, participants were presented with the same sentences, uttered by foreign-accented speakers of Spanish.

Three main observations were found:

- First, the results on native speech comprehension replicated previous findings (Federmeier et al., 2002), thus validating our materials and design.
- Second, during foreign-accented speech comprehension, sentences' best completions elicited decreased N400 and late negativity amplitudes compared to semantically related and unrelated words.
- Third, during foreign-accented speech comprehension, semantically related and unrelated words elicited similar N400 and late negativity amplitudes.

Thus, regarding native speech comprehension, we were able to replicate Federmeier et al.'s (2002) observations, extending previous results to native Spanish comprehension. We observed an attenuated N400 amplitude for expected compared to semantically related words, and also an attenuated N400 amplitude for semantically related words compared to semantically unrelated words during the 250-450 ms time window. That shows, as previously reported during English sentence comprehension (Federmeier & Kutas, 1999; Federmeier et al., 2002; Thornhill & Van Petten, 2012), that listeners are not only able to predict upcoming words, but also their semantic features, which also allows them to pre-activate semantically related words.

In addition to this replication, we extend previous results by showing that in the late negativity time window (500-900 ms), only semantically unrelated words elicited more negative amplitudes than sentences' best completions (best completions and semantically related words elicited similar amplitudes). That is, soon after the onset of the critical word ( $\approx 500$  ms), words that are semantically related to the best completion are easily integrated into the previous context, independently of their expectancy (or cloze probability). These results support previous literature suggesting that different processing mechanisms are involved during the classical N400 time window and the late negativity (Hahne & Friederici, 2001; Romero-Rivas et al., 2016). Hahne and Friederici (2001) suggested that late negative effects might indicate that listeners are

still engaged in trying to integrate the semantic properties of words into the previous context, after lexical access (see also Romero-Rivas et al., 2016, for late negativity effects during world knowledge processing). Therefore, in our case, listeners would have managed to integrate semantically related words 500 ms after the onset of the word, while they would be trying to integrate the semantic properties of unrelated words into the previous context for a longer period of time.

Regarding foreign-accented speech comprehension, our first observation was that listeners were able to anticipate the sentence's best completion when listening to foreign-accented speakers. In fact, we did not observe significant differences in the lexical anticipation effect (that is, the difference between Unrelated and Best Completion words) between native and foreign-accented speech comprehension. These results are congruent with previous studies showing that listeners adapt very rapidly to foreign-accented speech, using lexical information (Clarke & Garrett, 2004; Romero-Rivas et al., 2015). In addition, our results suggest that lexical anticipation is a very efficient process, since it takes place even when listening to foreign-accented speakers, who usually produce non canonical phonetic and prosodic information (Nisse et al., 2007; Wade et al., 2007; Wester et al., 2007; Gut, 2012; Hanulíková & Weber, 2012). Some authors (see e.g., Pickering & Garrod, 2007) have suggested that comprehenders predict upcoming words and discourse topics, through the involvement of the production system, in order to smooth communication. In this context, anticipating upcoming words when listening to foreign-accented speech would allow listeners to improve their comprehension of foreign-accented speakers.

The other main contribution of the study is to show that accessing and comprehending words that are semantically related to the expected ones is more demanding when listening to foreign-accented speakers, compared to native speakers (even if intelligibility<sup>4</sup> of sentences was at ceiling across native and foreign-accented speech, as indexed by the pre-test data). More concretely, we observed that semantically related and unrelated words elicited similar N400 and late negativity amplitudes when

---

<sup>4</sup> Comprehensibility refers to bottom-up activation processes from phonetic representations up to the lexicon to retrieve a possible candidate. In contrast, intelligibility refers to top-down decision processes involving lexical and pragmatic knowledge, arising from the computation of these lexical candidates. For further details on this comparison, please refer to Lahiri & Marslen-Wilson (1991) and Pallier et al. (2001).

listening to foreign-accented speakers. That is, when listening to foreign-accented speech, listeners are able to anticipate the sentence's best completion; however, if the expectation is not met, words sharing semantic features with the sentence's best completion do not benefit from anticipatory processes.

This latter result is consistent with those showing that adverse listening conditions (such as foreign-accented or degraded speech) narrow lexical expectations (Goslin et al., 2012; Strauß, Kotz, & Obleser, 2013). Interestingly, even if getting to the same conclusion that adverse listening conditions narrow lexical expectations, Strauß et al. (2013) and Goslin et al. (2012) did not interpret this effect in the same way. Strauß et al. (2013) observed that degraded speech narrowed the expectancies about the upcoming final word of spoken sentences. More concretely, Strauß et al. (2013) observed similar N400 amplitudes for “typical” and “atypical” word endings (e.g., “He peels a lot of potatoes” vs. “He peels a lot of bananas”) in clear speech. However, in degraded speech (noise-vocoding), “atypical” words elicited larger N400 amplitudes than “typical” words. The authors concluded that the limited perceptual evidence during degraded speech comprehension forced listeners to rely on perceptual (phonetic/acoustic) expectancies, rather than more abstract (semantic) expectancies. Conversely, Goslin et al. (2012) proposed that during foreign-accented speech comprehension (as compared to native speech comprehension), listeners are not able to normalize the speech signal during pre-lexical processing levels. Therefore, listeners rely on top-down contextual cues to a greater degree in order to normalize the signal, narrowing the possible lexical candidates. Our results are compatible with these two explanations. Hence, consistent with Strauß et al.'s (2013) conclusions, it might be that perceptual expectancies would make only the expected words to benefit from anticipatory processes during foreign-accented speech comprehension. However, we cannot discard the possibility that lexical expectations influence comprehension of foreign-accented speech. Thus, sentences' best completions (highest cloze-probability endings) would be so strongly anticipated (due to narrowing of lexical candidates) that even semantically related words would not be good enough candidates to complete the sentences. This latter explanation would be more consistent with Goslin et al.'s (2012) proposal. Since our study was not designed with the purpose of clarifying these two opposite explanations, further studies are needed to clarify this issue.

As an alternative explanation, our results might be interpreted according to the typicality and exemplar based models' framework. That is, those exemplars sharing the most perceptual features with the instances already stored in memory are considered typical and lead to a quicker categorization of the incoming input signal (Smith & Medin, 1999). In other words, less typical items (for instance, in our case, foreign-accented words, as compared to native spoken words) would require a more extensive search through memory, in order to find a matching exemplar already stored (Reisberg, 2013). Drawing a parallel, since most listeners probably have more familiarity with words spoken by native than foreign-accented speakers, accessing the lexical representations stored in semantic memory should be easier when listening to native than foreign-accented speakers. Therefore, it would be easier to access semantically related words when listening to native speech compared to foreign-accented speech. Somewhat relatedly, Hendrickson et al. (2015) showed graded N400 amplitudes for words based on semantic similarity [e.g., watching the picture of a dog while listening to the words “dog” (match), “cat” (near violation), or “lion” (far violation)], whereas environmental sounds did not elicit graded N400 amplitudes when comparing their semantic similarity [e.g., watching the picture of a dog while listening to *barkings* (match), *meows* (near violation), or *roars* (far violation)]. They concluded that listeners have more time to semantically organize words (vs. environmental sounds) in long-term memory, because of differences in the levels of familiarity and exposure that listeners have with words vs. environmental sounds. Thus, due to the low typicality of the exemplars, accessing the semantic properties of unexpected words that are semantically related to the expected ones would be more difficult when listening to foreign-accented speakers (compared to native speakers).

Therefore, our results suggest that linguistic anticipatory processes are modulated by indexical properties, such as the speaker's accent. These results are in line with previous studies showing that indexical properties of the speakers (such as age or gender of the speaker, or a foreign accent) interact with semantic congruity processing during auditory comprehension (Van Berkum et al., 2008; Romero-Rivas et al., 2015). Still, our result is novel in the sense that we show that not only semantic congruency processing is influenced by indexical properties of speakers, but also lexical and semantic anticipation processes. Altogether, these results indicate that the retrieval of



information from semantic memory during speech comprehension is not an encapsulated process, but rather depends on the indexical properties of the speaker.

Regarding topographic distribution of the effects, it is important to note that N400 effects were largest over central and posterior topographic regions, which is in accordance with the classic literature on the N400 (see e.g., Kutas & Hillyard, 1980, 1984). As for the late negativity (500-900 ms), effects were also largest over central and posterior topographic regions. Hahne and Friederici (2001) observed a right anterior-central distribution for the late negativity, while Romero-Rivas et al. (2016) observed a centro-posterior distribution for this effect. The distribution of our late negativity effect is clearly closer to Romero-Rivas et al.'s (2016) observation. This divergence could be explained by the fact that Hahne and Friederici (2001) used both syntactic and semantic manipulations (as well as the combination of the two previous) in their study, while Romero Rivas et al. (2016) only used semantic manipulations, as we did in this study.

A possible caveat of this study is that longer sentence and critical word durations during foreign-accented speech, compared to native speech, might have affected language comprehension. Previous literature on hesitations in speech (which, by definition, make sentences last longer) has shown that when target words are preceded by a hesitation, the N400 effect for unpredictable vs. predictable words is reduced (compared to fluent utterances; Corley, MacGregor, & Donaldson, 2007). In our study, the duration of sentences and critical words was longer in the foreign than in the native condition, and this might have led to a reduction in the magnitude of the semantic gradient for the foreign accents. However, during native speech comprehension, the three experimental conditions had similar sentence durations and, nevertheless, elicited very dissimilar N400 amplitudes (which therefore could not be explained by differences in length). Similarly, during foreign-accented speech comprehension, the three experimental conditions had different sentence durations but, nevertheless, Related and Unrelated conditions elicited similar N400 amplitudes. Similar conclusions can be drawn from word duration analyses. Therefore, the lengths of sentences and words do not seem to be a critical factor for anticipating upcoming words and accessing their semantic features. In addition, longer sentence and word durations (compared to native

speech) is a natural feature of foreign-accented speech that should be taken into account when studying this phenomenon (see e.g., Goslin et al., 2012).

To conclude, we provide preliminary evidence that linguistic anticipatory processes are modulated by indexical properties, such as the speaker's accent. More concretely, we observed that listeners were able to predict upcoming words during foreign-accented speech comprehension, but this prediction did not facilitate the integration of semantically related words. However, when listening to native speakers, listeners were not only able to anticipate upcoming words, but also other words with overlapping semantic features. These observations could be explained by narrowed lexical expectations during foreign-accented speech comprehension (that is, because a lexical candidate becomes much more salient than the rest). Alternatively, it could be that, during foreign-accented speech comprehension, listeners are not able to retrieve the semantic properties of words in an effective way during lexical integration. Irrespective of the mechanism behind this effect, what is important for our purposes is the observation of differences in the anticipatory processes associated with native and foreign-accented speech comprehension.

## **Acknowledgements**

We thank Xavier Mayoral and Silvia Blanch for their technical support, Ana Gutiérrez and Vania de la Fuente for assistance in testing participants, and Joanna D. Corey for revising the document. This research was funded by an FPI grant (BES-2012-056668) and three project grants (PSI2014-54500, PSI2011-23033 and Consolider INGENIO CSD2007-00012) awarded by the Spanish Government; by one grant from the Catalan Government (SGR 2014-1210); by one grant from the Basque Government (PI\_2015\_1\_25); and by one grant from the European Research Council under the European Community's Seventh Framework (FP7/2007-2013 Cooperation grant agreement 613465-AThEME). C.D.M. is supported by the IKERBASQUE institution, the Basque Center on Cognition, Brain and Language, and by the Severo Ochoa

program grant SEV-2015-049. A.C. is supported by the ICREA institution and the Center for Brain and Cognition.

## References

- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116*, 3647-3658.
- Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way you, er, say it: hesitations in speech affect language comprehension. *Cognition, 105*, 658-668.
- Davis, C. J., & Perea, M. (2005). BuscaPalabras: a program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37*, 665-671.
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117-1121.
- Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469-495.
- Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*, 133-146.
- FitzPatrick, I., & Indefrey, P. (2014). Head start for target language in bilingual listening. *Brain Research, 1542*, 111-130.
- Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain & Language, 122*, 92-102.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related potentials. *Bilingualism: Language and Cognition, 4*, 123-141.

Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*, *17*, 1101-1116.

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: the effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, *24*, 878-887.

Hanulíková, A., & Weber, A. (2012). Sink positive: Linguistic experience with the substitutions influences nonnative word recognition. *Attention, Perception, & Psychophysics*, *74*, 613-629.

Hendrickson, K., Walenski, M., Friend, M., & Love, T. (2015). The organization of words and environmental sounds in memory. *Neuropsychologia*, *69*, 67-76.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*, 281-312.

Kay, P. (1971). Taxonomy and semantic contrast. *Language*, *47*, 866-887.

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

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

Kutas, M., & Kluender, R. (1991). What is who violating? A reconsideration of event-related brain potentials. In H. J. Heinze, T. F. Münte, & G. R. Mangun (eds.), *Cognitive electrophysiology*, pp. 183-210. Boston: Birkhäuser.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: a phonological approach to the recognition of the lexicon. *Cognition*, 38, 245-294.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: concrete exemplar-based vs. abstract lexical entries. *Psychological Science*, 12, 445-449.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Science*, 11, 105-110.

Reisberg, D. (2013). *Cognition: Exploring the Science of the Mind*. 5<sup>th</sup> ed. W. W. Norton & Co. New York.

Romero-Rivas, C., Corey, J. D., Garcia, X., Thierry, G., Martin, C. D., & Costa, A. (2016). World knowledge and novel information integration during L2 speech comprehension. *Bilingualism: Language and Cognition*, available on CJO2016. doi:10.1017/S1366728915000905.

Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech. *Frontiers in Human Neuroscience*, 9:167.

Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (Vol. xii, p. 308). New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.

Samuel, A. G., & Larraza, S. (2015). Does listening to non-native speech impair speech perception? *Journal of Memory and Language*, *81*, 51-71.

Sebastián-Gallés, N., Martí, M. A., Cuetos, F., & Carreiras, M. (2000). *LEXESP: Léxico informatizado del español*. Barcelona: Edicions de la Universitat de Barcelona.

Smith, N. J., & Levi, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*, 302-319.

Smith, E., & Medin, D. (1999). The Exemplar View. *Concepts: Core Reading*, 207-209.

Strauß, A., Kotz, S. A., & Obleser, J. (2013). Narrowed expectancies under degraded speech: revisiting the N400. *Journal of Cognitive Neuroscience*, *25*, 1383-1395.

Taylor, W. L. (1953). 'Cloze procedure': a new tool for measuring readability. *Journalism Quarterly*, *30*, 415-433.

Thornhill, D. E., & van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, *83*, 383-392.

Van Berkum, J. J. A., Van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*, 580-591.

Van den Brink, D., Brown, C. M., & Hagoort, P. (2006). The cascaded nature of lexical selection and integration in auditory sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 364-372.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 394-417.

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

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica*, *64*, 122-144.

Wester, F., Gilbers, D., & Lowie, W. (2007). Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*, 477-491.

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

Wicha, N. Y., Moreno, E., & 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.



## Figure Captions

### **Figure 1.** Grand average ERPs for Experiments 1 and 2.

Grand average ERPs from critical words in native (Experiment 1; Left panel) and foreign-accented speech (Experiment 2; Right panel) from Pz electrode. Time zero indicates the onset of the critical word being the Best Completion (full lines), a Related word (dotted lines) or an Unrelated word (dashed lines). Negativity plotted up. The colored blue rectangle marks the 250-450 ms time window. The empty blue rectangle marks the 500-900 ms time window. Below, topographic distribution of the ERPs for each of the three conditions, in the two critical time-windows.

### **Figure 2.** Lexical anticipation and semantic relatedness effects.

Grand average ERP effects (Left panel: Unrelated – Best Completion; Right panel: Unrelated - Related) from critical words in native (full lines) and foreign-accented (dashed lines) speech from Pz electrode. The colored blue rectangles mark the 250-450 ms time window. The empty blue rectangles mark the 500-900 ms time window. Below, topographic distributions of voltage difference between conditions in native and foreign-accented speech and in the two critical time-windows.

### **Figure 3.**

Mean ERP amplitudes in the N400 (250-450 ms) and late negativity (500-900 ms) time-windows, for each of the three conditions, over the posterior topographic region (where the effects were strongest). Left panel: Mean ERP amplitudes in native speech (Experiment 1). Right panel: Mean ERP amplitudes in foreign-accented speech (Experiment 2). \* =  $p < .001$ . As can be observed, Related and Unrelated conditions were significantly different from each other during native speech comprehension, in both time windows, but not during foreign-accented speech comprehension.

## Table 1

### *Examples of sentences with English translations*

---

Nunca mete gol, es un pésimo jugador de fútbol/balonmano/videojuegos.

“He never scores, he is an awful football/handball/video games player.”

En el mapa de los piratas estaba marcada con una X la ubicación del tesoro/cofre/enemigo.

“In the pirates’ map there was an X showing the location of the treasure/chest/enemy.”

Lo tenía en la punta de la lengua, pero no conseguía recordar aquella palabra/expresión/fecha.

“He had it on the tip of his tongue, but was unable to remember that word/expression/date.”

---

Critical words are in italics and underlined. The first critical word is the best completion, expected word. The second critical word is a word semantically related to the most expected candidate, with a very low cloze probability. The third critical word is a word with no semantic relationship to the most expected candidate, but still plausible in the sentence context, and also with a very low cloze probability.

**Table 2****Exp. 1**

	250-450 ms			500-900 ms		
	<i>Frontal</i>	<i>Central</i>	<i>Posterior</i>	<i>Frontal</i>	<i>Central</i>	<i>Posterior</i>
<b>BC vs R</b>	3,68**	7,23**	7,36**	< 1	< 1	< 1
<b>BC vs U</b>	6,04**	9,46**	11,64**	< 1	3,23*	4,57**
<b>R vs U</b>	< 1	5,33**	8,26**	< 1	2,90*	5,36**

T values for the interactions between Topographic Region and Semantic Status observed in Experiment 1. BC = Best Completion; R = Related; U = Unrelated; \* =  $p < .01$ ; \*\* =  $p < .001$ .

**Table 3****Exp. 2**

	250-450 ms			500-900 ms		
	<i>Frontal</i>	<i>Central</i>	<i>Posterior</i>	<i>Frontal</i>	<i>Central</i>	<i>Posterior</i>
<b>BC vs R</b>	< 1	4,25**	7,32**	< 1	3,58**	5,41**
<b>BC vs U</b>	1,12	6,44**	8,51**	< 1	5,24**	7,28**
<b>R vs U</b>	< 1	< 1	< 1	< 1	< 1	< 1

T values for the interactions between Topographic Region and Semantic Status observed in Experiment 1. BC = Best Completion; R = Related; U = Unrelated; \* =  $p < .01$ ; \*\* =  $p < .001$ .

**Table 4**

	<u>Phonological</u> <u>Length</u>		<u>Word</u> <u>Frequency</u>	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
<b>BC</b>	6.53	0.15	68.08	5.80
<b>R</b>	6.62	0.16	57.79	10.59
<b>U</b>	6.73	0.15	66.03	9.25

Mean values (*M*) and Standard Error (*SE*) for the phonological length and word frequency of the critical words. BC = Best Completion; R = Related; U = Unrelated.

**Figure 1**

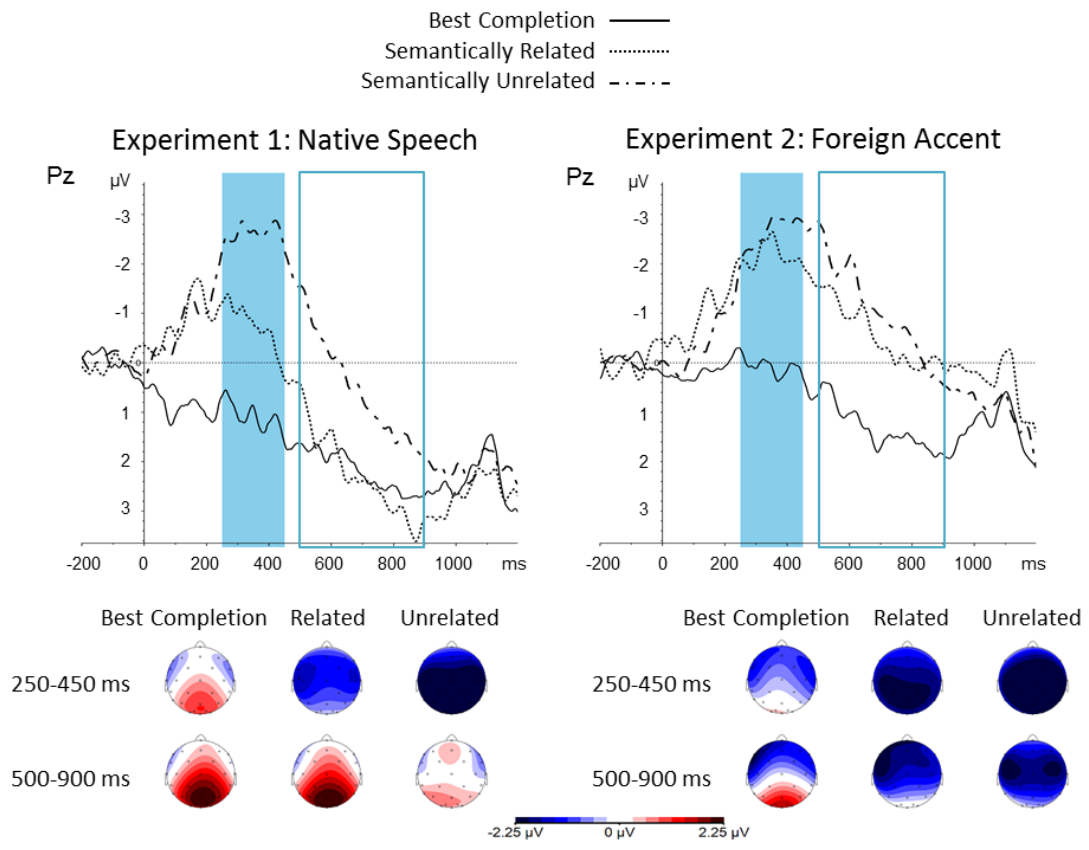
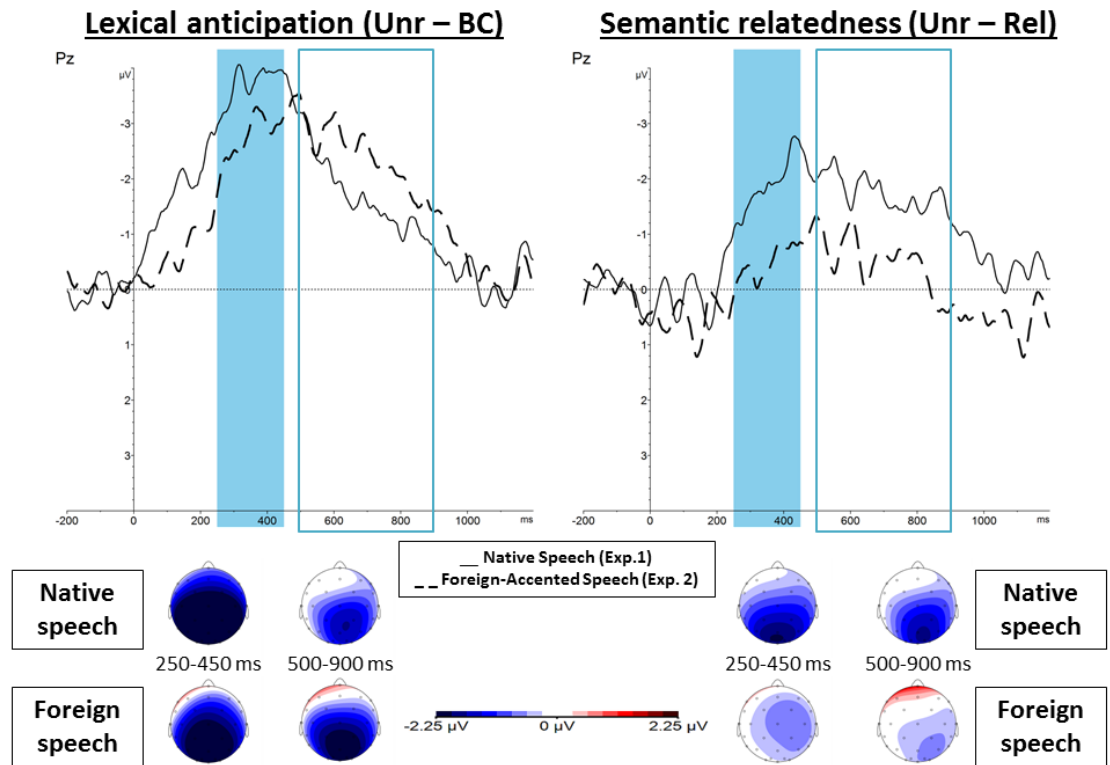


Figure 2



**Figure 3**

