

Native and second-language processing of contrastive pitch accent: An ERP study

Second Language Research

2020, Vol. 36(4) 503–527

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0267658319838300

journals.sagepub.com/home/slr

Aleuna Lee, Michelle Perdomo 
and Edith Kaan 

University of Florida, Gainesville, FL, USA

Abstract

Prosody signals important aspects of meaning, and hence, is crucial for language comprehension and learning, yet remains under-investigated in second-language (L2) processing. The present electrophysiology study investigates the use of prosody to cue information structure, in particular, the use of contrastive pitch accent (L+H*) to define the set of elements that are contrasted. For instance, in *We ate Angela's cake, but saved BENjamin's cake*, the pitch accent on *Benjamin's* is a cue that two cakes are contrasted; *BENjamin's ice cream* is not plausible in this context. Native English speakers showed a large negativity on the target noun (*cake*) when the preceding possessive was inappropriately accented. Event-related brain potential (ERP) results from Mandarin-Chinese L2 learners of English suggest they did not use contrastive pitch accent to cue the contrast set in the way native English speakers did, even though Mandarin is similar to English in the use of prosodic cues to express contrast. Our results are in line with previous studies suggesting that L2 speakers have difficulty integrating information across domains and building information structure, especially in demanding task situations like in the present study.

Keywords

contrastive pitch accent, ERP, N400, prosody, second language learning

1 Introduction

1 Overall introduction

Prosody, or the modulation of pitch, duration, intensity and pauses across a sentence, is an important means of expressing information status in discourse. To give an example, in *We ate Angela's cake, but saved BENjamin's ...*, the contrastive pitch accent on *Benjamin's* is a cue that the contrast will be between two cakes (*Angela's* and *Benjamin's*). On the

Corresponding author:

Edith Kaan, Department of Linguistics, University of Florida, Box 115454, Gainesville, FL 32611, USA.

Email: kaan@ufl.edu

other hand, a contrastive pitch accent on the noun, e.g. *cake* in *We ate Angela's CAKE, but saved ...* signals that the contrast is between two different objects that belong to Angela (e.g. *Angela's cake* and *Angela's ice cream*). Prosody therefore is an important cue for information structure, that is, a representation of what is given, new, and focused information in a sentence (Halliday, 1967). Despite the importance of prosody for language understanding, there is a surprising lack of studies on how L2 listeners use prosody. According to various proposals regarding L2 processing, L2 learners are expected to have problems using prosody to signal information structure because it involves either integration of information across various representational domains (Interface Hypothesis, Sorace, 2011) or complex mapping of information across domains (Patterson et al., 2017). In addition, the production and perception of L2 prosody is notoriously hard for learners to fully master (Huang and Jun, 2011; Mennen and de Leeuw, 2014), especially when the correspondence between L1 and L2 prosodic patterns and their functions is not straightforward (e.g. Chen and Lai, 2011; Zubizarreta and Nava, 2011). The present study used event-related brain potentials (ERPs) to investigate to what extent native listeners and L2 learners use prosody to inform information structure. In particular, we tested to what extent listeners are sensitive to contrastive pitch accent, and to what extent they can use contrastive pitch accent as a cue to construct the appropriate focus set, that is, the set of contrasting elements. Below we will first discuss why processing and information structure is expected to be difficult in L2. We then discuss contrastive pitch accent and ERP components shown to be sensitive to prosodic manipulations before discussing the details of the current study.

2 Processing information structure in L2

Various accounts of second-language processing predict that L2 listeners will experience difficulty using prosodic information to build information structure when processing their L2. According to the Interface Hypothesis (Sorace, 2011), the integration of information across domains is a major bottle neck in L2 acquisition and processing. Building and using information structure in speech processing critically involves the combination of prosodic, syntactic, lexical and contextual information. To illustrate this, consider the question: *Did Bill eat an apple or an orange?* This question requires an answer with a particular prosody and syntax. For instance, *Bill ate an ORange* is an appropriate answer, as is a cleft construction (*It was an orange that Bill ate*); however, using a cleft in which the given information (*Bill*) is in syntactic focus, as in *It was Bill who ate an orange*, is an inappropriate answer, regardless of whether *orange* has a contrastive pitch accent. The Interface Hypothesis predicts that L2 learners will have difficulties combining information across different domains (e.g. prosody and syntax) and will not perform like native speakers in their use of prosody to build information structure, even though they may be sensitive to and familiar with the prosodic information, the lexical items, as well as syntactic structures used in the auditory input. In contrast, Patterson et al. (2017) propose that L2 learners may be able to combine information across domains but will have difficulty if this involves complex mapping. For instance, Patterson et al. show that L2 learners use simple cues such as focus particles and sentence-initial positions to determine which referent is most accessible and which can serve as an antecedent of a

pronoun; however, L2 speakers have problems using information structure signaled by syntactic means to find a pronoun antecedent. This is especially the case if the syntactic positions do not straightforwardly map onto prominence relations used to find antecedents for pronouns. For instance, in the cleft *It was the driving instructor who told the examiner on the telephone that he ...*, the noun phrase *the driving instructor* is in focus and in a prominent position at the beginning of the sentence. When reading German equivalents of these sentences, native speakers of German prefer the same-sentence pronoun *he* to refer to (part of) the topic *the examiner*, whereas L2 speakers of German do not show such preference. This suggests that L2 speakers are like native speakers in combining some forms of information across linguistic domains but not in combining others. This is problematic for the Interface Hypothesis, according to which all cross-domain mappings will be difficult for L2 speakers.

3 L2 processing of prosody

Both the Interface Hypothesis and the complex mapping account predict that L2 speakers have difficulty using prosody to build information structure to some extent. Thus far, not many studies have investigated L2 listeners' use of prosodic information to inform information structure during listening, but there is some evidence that L2 learners have difficulty with this aspect of auditory sentence comprehension, especially when the mapping between prosody and function is different between L1 and L2 and/or if the mapping is rather complex.

First, there is some evidence that L2 speakers can use L2 prosody to determine information structure but that this depends on the differences between L1 and L2. Ortega-Llebaria and Colantoni (2014) presented listeners with questions such as *Did Bobby fall out of the tree?* and had them select the correct answer out of three options which differed in their prosody (correct: *TOBY fell out of the tree*; incorrect: *Toby FELL out of the tree*; *Toby fell out of the tree*). Mandarin L2 speakers of English did not differ from native English speakers in their responses. Spanish L2 speakers, however, performed worse than Mandarin L2 speakers in this study. Ortega-Llebaria and Colantoni attribute this to the difference between Spanish on the one hand, and Mandarin and English on the other. Both Mandarin and English often use prosodic means to express an aspect of information structure, whereas Spanish often uses syntactic means.

Second, even if the L1 and L2 both use prosodic cues for information structure, L2 speakers may experience difficulty if the L2 makes a distinction the L1 does not. For instance, Dutch uses both H*L (falling) and L*H (rising) contours to indicate that information is new, whereas English only uses the former. In addition, Dutch does not have a contrastive pitch accent (Krahmer and Swerts, 2001). In a visual world study (Chen and Lai, 2011), Dutch speakers interpreted both contour types to refer to new information in their L2 English. Furthermore, the L2 speakers in the study did not appear to use the intonational information to predict the upcoming referent like native speakers do (Ito and Speer, 2008). Similarly, Namjoshi (2015) found that English learners of French were sensitive to the prosodic differences between contrastive and nuclear accent in French (which are realized differently in English), but did not map these onto information structure like the native French speakers did. Experimental results thus far therefore suggest that L2

speakers may experience problems using prosody to inform information structure due to differences in how their native and second language express information structure.

Third, in support of the complex mapping approach, mapping prosody and information structure may become even more difficult for L2 speakers if this mapping involves many steps. Shafer et al. (2015) showed that native English speakers as well as Korean and Chinese L2 learners of English preferred to have the pronoun refer to a stressed antecedent. For instance the pronoun *he* was taken to refer to *David* in contexts such as *DAVID served Paul a pint of beer. He ...*, but the pronoun *he* was taken to refer to *Paul* in *David served PAUL a pint of beer. He ...* In a second experiment, Shafer et al. stressed the pronoun and used contexts such as *David served Paul a pint of beer. HE/he obviously ...* Native English speakers more often interpreted the pronoun as referring to the second noun in the first sentence (*Paul*) when the pronoun was stressed than when it was unstressed. L2 speakers, on the other hand, preferred the pronoun to refer to the first noun, regardless of whether the pronoun was stressed. This suggests that L2 speakers can use prosody to find appropriate antecedents when the mapping is simple (a stressed noun is salient and a likely antecedent of a pronoun), whereas they have difficulty when the mapping is complex (a stressed pronoun leads to the identification of a set of potential referents, out of which the one needs to be selected that is NOT salient). This supports the complex mapping approach.

In summary, the evidence thus far suggests that L2 speakers are sensitive to modulations of L2 prosody but may have difficulty combining prosodic, syntactic and other information to build and evaluate information structure especially when the L2 makes distinctions not present in the L1 and/or when the mapping between prosody, information structure and other information is complex. Although not much investigated, proficiency may factor in as well, with more advanced learners showing more native-like patterns (Chen and Lai, 2011; Reichle, 2010; Reichle and Birdsong, 2014).

The present study further explored to what extent native and L2 listeners are sensitive to prosodic differences and can use these to build information structure, and the role of proficiency therein. In particular, we tested to what extent L2 speakers are sensitive to contrastive pitch accents and to what extent they use this pitch accent as a cue to construct a specific focus set.

4 Contrastive pitch accent

Contrastive pitch accent in English is instantiated as a low-falling followed by high-rising F_0 contour, also denoted L+H* in the ToBI (tones and break indices) annotation system (Beckman and Ayers Elam, 1994; Beckman and Hirschberg, 1994). This prosodic pattern is used to inform the listener which elements should be contrasted. For instance, the contrastive pitch accent on *Benjamin* in *We took Angela's cake, but not BENjamin's ...* is a cue to the listener that what is contrasted is a set of cakes. No such narrow expectation is set up when *Benjamin* does not have a contrastive pitch accent and the prosody of the first clause is kept the same. For instance, the fragment *We took Angela's cake, but not Benjamin's ...*, where *Benjamin* does not receive a contrastive pitch accent, could be followed by about anything.

Visual-world (eye-tracking) studies have shown that native listeners actively use contrastive pitch accent in comprehension to anticipate the set of upcoming referents (the contrast set), or at least to speed up identification of referents (Dahan et al., 2002; Ito

et al., 2014; Ito and Speer, 2008; Sedivy et al., 1995). For instance, Ito and Speer (2008) instructed participants to decorate a holiday tree. When hearing instructions such as *Hang the green drum. Now hang the BLUE drum*, participants were quicker to move their eyes to a drum than when the last adjective was not accented. In addition, participants were slower to fixate on the intended target if the pitch accent was used infelicitously (e.g. *Hang the red angel. Now hang the BLUE drum*). This suggests that native speakers can immediately use contrastive pitch accent to restrict the set of possible referents.

Testing Mandarin-Chinese L2 English speakers, Takahashi et al. (2018) showed that these L2 English speakers are sensitive to the correct and incorrect use of a contrastive pitch accent on an adjective in English: participants in the study were faster to click on the target picture (mittens) when the pitch accent was used correctly (e.g. *Click on the purple mittens. Now click on the SCARLET mittens ...*) than when it was used incorrectly (as in, for example, *Click on the scarlet necklace. Now click on the SCARLET mittens ...*). This suggests that Mandarin L2 English speakers recognize that the accented adjective expresses a contrast with another adjective: the oddity of having a repeated adjective accented may have slowed the listeners down in identifying the target. The current study tested whether the L2 speakers could use contrastive pitch accent to restrict the set of upcoming referents.

5 Prosody and ERPs

In the current study we used Event-Related brain Potentials (ERPs) to investigate the use of prosodic information in native and L2 listeners. This technique allowed us to track processing while the sentence unfolds without having to restrict the set of potential referents by means of a visual display, as in eye-tracking studies. Studies on native listeners have identified several ERP components that may reflect different aspects of the processing and use of prosodic information. First, elements that bear a prosodic accent typically elicit a 200–500ms positivity that is widely distributed over the scalp (Dimitrova et al., 2012; Heim and Alter, 2006), as do prosodic cues marking clausal boundaries (Hruska et al., 2001; Sheppard et al., 2018; Steinhauer, 2003). Similar effects have been found for focused positions in written language (e.g. Cowles et al., 2007; Reichle and Birdsong, 2014).

Second, several ERP components have been found to reflect the processing of a mismatch between the prosody and the intended information structure. Studies typically manipulate whether an accent is missing or superfluous given the intended information structure. In the example *Did the club give a bonus to the player or to the trainer? They gave a BONUS to the player* (from Dimitrova et al., 2012), the word *bonus* is not in informational focus but is superfluously accented, whereas *player* is in focus but is missing a pitch accent, making the response presumably prosodically infelicitous. Superfluous or missing accents have been reported to elicit an increased N400 component (Bögels et al., 2011; Dimitrova et al., 2012; Hruska et al., 2001; Magne et al., 2005; Toepel and Alter, 2004), a more frontal negativity (Li et al., 2011), and/or a late positivity (Dimitrova, 2012; Li et al., 2018; Magne et al., 2005); sometimes no effects are reported for superfluous (Bögels et al., 2011; Hruska et al., 2001) or missing accents (Dimitrova, 2012; Li et al., 2018). The N400 is a negative-going potential, peaking between 300–500ms, which is typically associated with semantic integration difficulty. A more frontal

Table 1. Examples of the four experimental conditions.

	Contrastive pitch accent	Same or different nouns	Example
	Context: Angela and Benjamin came to our potluck. They brought dessert.		
A	–Accent	Different	We ate Angela's ice cream but saved Benjamin's <u>cake</u> in the fridge.
B*	+ Accent	Different	We ate Angela's ice cream but saved BENjamin's <u>cake</u> in the fridge.
C**	–Accent	Same	We ate Angela's cake but saved Benjamin's <u>cake</u> in the fridge.
D	+ Accent	Same	We ate Angela's cake but saved BENjamin's <u>cake</u> in the fridge.

Notes. Bold indicates the word that has contrastive pitch accent; critical noun is underscored; * accent is used infelicitously; ** accent is missing.

negativity has been associated with difficulties finding an unambiguous referent for the noun phrase (Nref component; Van Berkum et al., 2007). The positivity found for inappropriate accents has been interpreted as a P300, associated with the processing of unexpected or novel stimuli (Magne et al., 2005), as a Closure Positive Shift associated with prosodic phrasing (Hruska et al., 2001), or as a P600 effect associated with revision processes resulting from inappropriate prosodic phrasing (Pauker et al., 2011).

6 The current study and predictions

In the present ERP study, we manipulated the presence of contrastive pitch accent, which could either be appropriate or inappropriate given the target information structure. In our paradigm, a pitch accent did not become inappropriate in terms of information structure until the next word. Examples of our four conditions are given in Table 1.

Experimental trials consisted of two context sentences, in which two characters were introduced, and a target sentence. The target sentence always expressed a contrast between two objects, one belonging to one character, one to the other. The proper name and noun mentioned last were our critical positions. The critical proper name (*Benjamin* in Table 1) either had contrastive pitch accent (+ Accent conditions B, D) or no accent (–Accent conditions A, C). The critical noun itself (e.g. *cake*) was either different from the noun mentioned earlier (Different conditions, A, B) or not (Same conditions, C, D). In English, a contrastive pitch accent on the critical proper name is a cue to the listener that the following element refers to an object of the same kind that has been mentioned previously. The pitch accent on the proper name is therefore felicitous in condition D where the two nouns are the same (+Accent Same), but infelicitous in condition B where the two nouns are different (+Accent Different). In contrast, the absence of a contrastive pitch accent is expected in A where the two nouns are different (–Accent Different) but may be perceived as infelicitous in C where the two nouns are the same (+ Accent Same). Note that several steps are involved in arriving at the intended information structure in the +Accent Same condition. First, when encountering *but* followed by the verb, a contrast set needs to be

formed of actions that can contrast with those mentioned in the first clause. When the critical proper name is introduced, another contrast set is constructed, namely between the object belonging to the first protagonist (e.g. Angela) and objects belonging to the second (e.g. Benjamin). If the proper name is accented, the latter contrast needs to be even more restricted to objects of the same type as mentioned in the first clause (Angela's cake contrasted with Benjamin's cake).

If listeners notice the pitch accent or at least the difference in prominence at the proper names in the second clause, the ERPs should show an effect of pitch accent at the critical proper name (*BENjamin* vs. *Benjamin*). Based on the studies cited above in which prosodic prominence was manipulated (Dimitrova et al., 2012; Heim and Alter, 2006; Hruska et al., 2001; Sheppard et al., 2018; Steinhauer, 2003) we expected a widely distributed 200–500 ms positivity for the contrastively accented vs. non-accented proper name.

If listeners use the contrastive pitch accent on the proper name to build the contrast set, they will expect a noun referring to the same type of object when the proper name is accented (+Accent Same), or at least such a noun should be easier to integrate into the message. However, if the accented proper name is followed by a noun referring to a different object (+Accent Different), the expectation of the same type of object is violated, and the noun is harder to integrate. In this case we expected to see a large N400 component in the +Accent Different relative to the +Accent Same condition. In addition, a frontal negativity and/or positivity could be seen as these effects have been observed in previous studies for infelicitously accented words (Bögels et al., 2011; Dimitrova et al., 2012; Hruska et al., 2001; Li et al., 2018; Magne et al., 2005; Toepel and Alter, 2004). In the –Accent conditions, there is no strong expectation for what kind of noun will be mentioned. The –Accent Same condition may be perceived as strange at the noun since an accent is missing. However, in studies manipulating intonational boundaries, anomalies caused by the absence of a prosodic marking have been shown to have a less detrimental effect on processing than anomalies caused by a superfluous marking on a previous word (Pauker et al., 2011). We therefore expected the difference in ERPs between the two –Accent conditions to be smaller than that between the two +Accent conditions. In summary, at the critical noun, we expected the N400 to be largest for the +Accent Different (B) condition vs. the +Accent Same condition (D); the difference in the N400 effect was expected to be smaller between the –Accent conditions, with the N400 being larger in the –Accent Same condition (C) than the –Accent Different (A) condition.

The L2 speakers of English in this study were L1 speakers of Mandarin from mainland China. Like English, Mandarin uses prosody to mark contrastive focus rather than relying solely on syntactic means (Xu, 1999). However, whereas in English contrastive focus is marked by an L+H* pitch accent, in Mandarin Chinese, information structure is denoted by the duration and range of the F_0 contour to preserve lexical tones (Chen, 2006; Chen and Gussenhoven, 2008; Xu, 1999). In spite of these differences in the prosodic expression of contrastive focus, the Ortega-Llebaria and Colantoni (2014) and Takahashi et al. (2018) studies discussed above, they found that Mandarin L2 English speakers performed as well as native English speakers in tasks probing perception of sentence-level prosody and correctly recognized the relation between informational focus and prosodic prominence. We therefore expected that the L2 speakers in our study would notice the prominence of the contrastive pitch accent and, like the native English speakers, would show a positivity for

the accented vs. non-accented proper names. If the L2 speakers also recognize the L+H* accent as expressing a contrast, as in the Takahashi et al. (2018) study, and use this information to construct the intended contrast set, we expected the ERPs at the second noun to show the same pattern as in the native speakers, with the +Accent Different condition showing the largest negativity. On the other hand, given the complexity and number of steps involved in restricting the contrast set in the +Accent conditions, the L2 speakers may break down, especially if they are less proficient. In that case we expected to see less of a difference between the +Accent Same (D) and +Accent Different (B) conditions than in the native speakers, with this difference being modulated by L2 proficiency.

II Methods

I Participants

Two groups of participants were recruited for this study: 33 functionally monolingual native speakers of English and 26 native speakers of Mandarin Chinese who were second-language learners of English. Participants were undergraduate and graduate students from the University of Florida and received course credits and/or cash compensation for the participation. Nine native speakers of English were excluded from the analysis due to technical difficulty (1 participant), not completing the study (3 participants), or excessive artifacts leading to a data loss rate exceeding 70% in at least two of the experimental conditions (5 participants). Three Mandarin learners of English were excluded because of low familiarity ratings of the nouns and names used in the trials¹ (1 participant) and excessive data loss (2). The remaining participants were 24 native speakers of English (15 women and 9 men, 18–31 years, mean age 20.9) and 23 Mandarin learners of English (14 women and 9 men, 18–35 years, mean age 24.5).

All participants completed the LEAP-Q language background questionnaire (Marian et al., 2007). Mandarin speakers indicated that they started learning English since around 9 years of age (range 5–13 years) and had been in an English-speaking country approximately 2 years (range 1 month–11 years). Immediately after the ERP session, participants completed the LexTALE task, a standardized English lexical decision task (Lemhöfer and Broersma, 2012). The native English speaker group obtained a mean score of 88 (SD 9.8, range 63–100) on this task while Mandarin speakers had a mean score of 64 (SD 8.6, range 51–84). In addition, the Mandarin speakers of English completed the grammar and cloze portion of the MELICET (Michigan English Language Institute, 2001). The mean score on this task was 67 (SD 9.1, range 50–84), indicating the L2 group varied widely in proficiency scores. All participants were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) with normal or corrected-to-normal vision and no history of reading difficulties or neurological disorders. All had a minimal bilateral hearing range of 500 to 8,000 Hz measured at 25 dB HL as tested on-site.

2 Materials

One hundred and sixty quadruplets were created of the format illustrated in Table 1. Each trial consisted of three sentences. The first sentence introduced a male and a female

character using their proper names (e.g. *Angela and Benjamin came to our potluck*); the second sentence provided a situational context (e.g. *They brought dessert*). The third sentence was the target sentence. The target consisted of two clauses connected by the conjunction *but* which each contained a proper name in the possessive followed by a noun (*Angela's cake, Benjamin's cake*) and a verb that expressed a contrast with the verb in the other clause (e.g. *ate* vs. *saved*; *tried* vs. *refused*) or that was a negated synonym (e.g. *cleared out* vs. *did not clean up*). Four conditions were created such that the second proper name was either contrastively accented or not, and the nouns were either the same or different (see Table 1). The second proper name was always three syllables in length, with stress on the first syllable. In conditions in which the two nouns were different, the two nouns were semantically related. The critical second noun was followed by at least two other words to avoid effects of sentence-final intonation.

All auditory stimuli were spoken by a female native speaker of American English and recorded at a sampling rate of 41.4 kHz in a soundproof booth. To control the acoustic variations between four conditions, the speaker only recorded the two felicitous versions of the items (+Accent Same and -Accent Different). Recordings were done in blocks of 20 sentences at a time, each block containing sentences of only one condition. The amplitudes of all the audio recordings were normalized using Praat (Boersma and Weenink, 2017). Next, recordings were spliced right after the offset of the connective *but* in the target sentence. Infelicitous conditions were created by cross-splicing the first part of the +Accent Same recordings with the second part of the -Accent Different recordings, and vice versa. For the -Accent conditions, the pitch contour was realized as L* for the critical proper name (e.g. *Benjamin's*) across all trials (see Figure 1), while the pitch contour was L+H* for the +Accent conditions (Figure 2). Tables 2 and 3 summarize the acoustic analysis on F_0 and duration of the proper names and nouns in the second and first clause, respectively. In the critical second clause (Table 2), the +Accent and -Accent conditions differed significantly in all measures for the proper name. The second noun was significantly longer in the +Accent than -Accent conditions. In the first clause (Table 3), the proper name had a longer duration, higher maximum F_0 and larger F_0 range in the conditions recorded as +Accent Same compared to the -Accent Different conditions; the following noun had a lower minimum F_0 and a larger F_0 range in this condition. These first clauses were used as context sentences in the Same and Different conditions, respectively. The differences in prosodic properties of the proper noun and noun in the first clause may have created expectations for the proper name in the second clause to be accented or not, which could have led to differences in the ERP between the infelicitous vs. felicitous conditions already at the second proper name. However, as we will discuss below, we did not find such effects in the native speakers.

Materials also included 120 filler items distributed over four different sentence types: (1) containing connectives other than *but* (*after, and, if, since, because* or *until*); (2) contrastive pitch accent in a different position within the sentence, such as other nouns or verbs; (3) nouns from different semantic categories (e.g. *Lisa was not feeling well. Burt came to her home to care for her. Burt cooked some soup for Lisa and gave her some aspirin*); (4) cleft constructions (e.g. *It was Barney who graduated* or *What Sarah wants is a ...*) in which the pitch accent signals informational focus. The prosody was incorrect or unusual in 23% of the filler items.

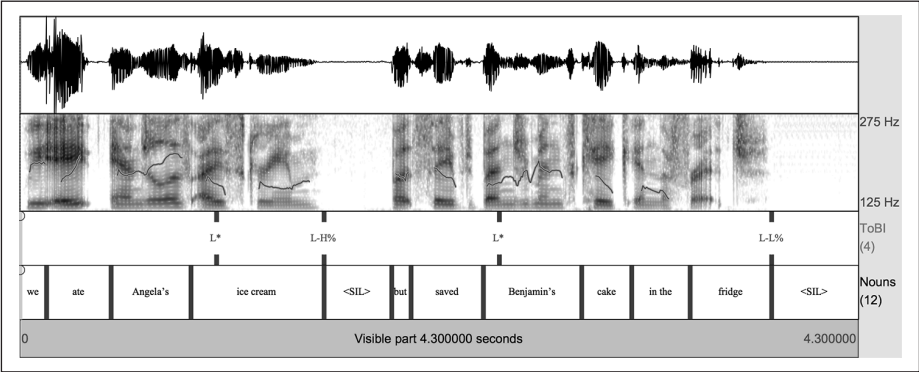


Figure 1. Time wave form and ToBI-annotated F_0 contour for an example sentence in the –Accent Different condition.
Note. <SIL> means ‘silence’.

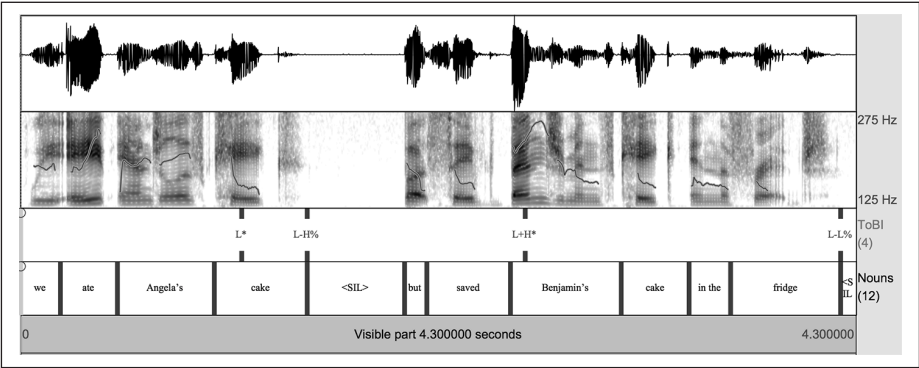


Figure 2. Time wave form and ToBI-annotated F_0 contour for an example sentence in the +Accent Same condition.
Note. <SIL> means ‘silence’.

Table 2. Mean duration (ms) and F_0 (Hz) of the critical proper name (*Benjamin’s*) and noun (*cake*) in the +Accent (L+H*) and –Accent conditions, standard deviation in parentheses.

	Proper name		Noun	
	–Accent	+ Accent (L+H*)	–Accent	+ Accent (L+H*)
Duration (ms)	476 (62)	575 (73)**	405 (115)	475 (122)**
Minimum F_0 (Hz)	149 (29)	143 (24)**	135 (28)	131 (31)
Maximum F_0 (Hz)	239 (110)	334 (65)**	207 (99)	222 (127)
Range F_0 (Hz)	89 (110)	191 (68)**	72 (98)	91 (133)

Notes. ** $p < .05$ as per pairwise t -tests between +Accent and –Accent (df 159).

Table 3. Mean duration (ms) and F₀ (Hz) of the proper name (*Angela's*) and noun (*ice cream/cake*) in the first clause used for the +/–Accent Same and +/–Accent Different conditions, standard deviation in parentheses.

	Proper name		Noun	
	–/+Accent different	+/–Accent same	–/+Accent different	+/–Accent same
Duration (ms)	437 (77)	457 (79)**	498 (114)	506 (121)
Minimum F ₀ (Hz)	164 (27)	165 (29)	126 (35)**	114 (35)
Maximum F ₀ (Hz)	240 (80)	263 (104)**	209 (89)	226 (116)
Range F ₀ (Hz)	76 (79)	98 (104)**	83 (95)	112 (122)**

Notes. ** $p < .05$ as per pairwise t-tests between Same and Different conditions (df 159).

The experimental items were Latin-Squared based on the four conditions, creating four participant lists. Within and across list, the items in the conditions were matched in terms of word frequency of the critical noun, order of the gender of the characters in the critical sentence, and position of the critical nouns within a target sentence. Each participant in the ERP study listened to a total of 280 trials (160 experimental items pseudorandomly intermixed with 120 fillers) divided into 8 blocks of 35 trials. A quarter of the trials were followed by a yes/no comprehension question to encourage participants to pay attention to the sentences. We did not have participants judge the appropriateness of the prosody or of the sequence of clauses. This was to minimize task demands and to avoid introducing decision or response-related components before the last word of the trial. The eight blocks were presented in a different random order for each participant.

3 Procedure

During the experiment, participants were seated at a distance of 1,000 mm from a computer monitor in a sound-attenuating chamber, wearing a cap with electrodes. They were asked to listen to short passages over pneumatic earphones and to respond to yes/no comprehension questions after certain trials by pressing buttons on a game controller. Each trial started with black screen. After 500 ms, a white square was presented on the screen and the sentence audio began playing. The square was replaced by a white fixation cross 3,500 ms before the onset of the second proper name. This point coincided roughly with the onset of the target sentence but was different for each trial. Participants were asked not to blink when they saw the fixation cross. After the offset of the target sentence, the screen turned black for 750 ms. This was followed in some trials by a comprehension question that was presented auditorily. Participants had unlimited time to answer the question by pressing buttons on the game controller. Trials were separated by a visual message that read *Press for next*. Participants were asked to press to continue when they felt ready to proceed to the next trial.

A practice session including five items was run before the experiment to ensure that participants were familiar with the task and could hear the sounds clearly. These practice items contained different structures than those in the critical items within the experiment.

Participants took a 2- to 3-minute break in between each of the eight blocks. Recording time was approximately one hour and a half. Participants completed the LEAP-Q questionnaires during the EEG setup, and the LexTALE proficiency task immediately after the ERP experiment. The L2 speakers also completed the MELICET proficiency task and an on-line questionnaire in which they judged their familiarity with the critical nouns and proper names used in the study.²

a EEG recording. The electroencephalogram (EEG) was recorded from 64 Ag/AgCl electrodes (Fp1/2, AF3/4/7/8, F1/2/3/4/5/6/7/8, FC1/2/3/4/5/6, FT7/8, C1/2/3/4/5/6/, T7/8, CP1/2/3/4/5/6, TP7/8, P1/2/3/4/5/6/7/8, PO3/4/7/8, O1/2, Fpz, Fz, FCz, Cz, CPz, Pz, POz, Oz) mounted in a Waveguard cap. Blinks and eye movements were monitored by electrodes on the left and right outer canthi and electrodes placed above and below the left eye. Electrodes were also placed over the left and right mastoids. Impedance was maintained below 5 k Ω . EEG was sampled at 512 Hz. An average reference was used.

b EEG Analysis. The EEG data were re-referenced offline to the mean of both mastoids and filtered between 0.01 and 30 Hz. ERPs were time-locked to the onset of the critical proper name (e.g. *Benjamin's*) and noun (e.g. *cake*). Individual averages were computed for epochs ranging from 200 ms prior to and 1,500 ms post the onset of the critical word for each condition and electrode, using a time window of 200 ms preceding word onset as a baseline. Trials contaminated by artifacts were rejected from these averages; for native speakers of English, average rejection rates ranged from 31–34% per condition; for L2 speakers, average rejection rates were 40–43% per condition for the proper name epochs, and 35–40% for the noun epochs.

Based on the timing of the effects found in the literature, statistical analyses were conducted on the mean amplitudes in the following time windows, relative to a 200 ms pre stimulus baseline: (1) between 200–500 ms after the onset of the second proper name (*Benjamin's*); this was to assess whether the contrastively accented proper names would elicit a posterior positivity vs. the non-accented proper names (Dimitrova, 2012); and (2) 300–500 ms after the onset of the second noun (*cake*). This was to capture the N400 and other negativities observed in previous studies (Bögels et al., 2011; Dimitrova et al., 2012; Hruska et al., 2001; Magne et al., 2005; Toepel and Alter, 2004). In addition, (3) we probed the 700–1,000 ms time window to test late positive effects (Dimitrova, 2012). However, since no effects were found in the latter time window, the results from this window will not be further discussed.

A repeated-measures generalized linear model (GLM) was run with Accent (+ Accent, –Accent), Type of noun (Same, Different), Laterality (Left, Midline, Right) and Anteriority (frontal: F5/F3/Fz/F4/F6; fronto-central: FC5/FC3/FCz/FC4/FC6; central: C5/C3/Cz/C4/C6; central-parietal: CP5/CP3/CPz/CP4/CP6; parietal: P5/P3/Pz/P4/P6) as within-participants factors. The Greenhouse–Geisser correction was applied for effects involving factors with more than two levels (Greenhouse and Geisser, 1959). These analyses were conducted separately for each language group. We will only report significant effects involving the factor Accent and/or Type of noun. In the final analysis, we compared two groups, including Language as a between-participants factor.

III Results

I Comprehension questions

Mean accuracy of the comprehension question was 84.95% ($SD = 5.46\%$) for the native English group and 74.15% ($SD = 6.37\%$) for the Mandarin-English L2 speakers, suggesting that participants were paying attention. Native English speakers had overall higher accuracy on the questions compared to L2 speakers ($T(43) = 6.23, p < .001$).

2 ERP results from native English speakers

a Effects of contrastive pitch accent at the proper name. ERPs to the second proper name for the English speakers are given in Figure 3. The effects of contrastive pitch accent were assessed at 200–500ms after the onset of second proper name, which could either have a contrastive pitch accent or not. Accented proper names evoked a large positive deflection compared to non-accented proper names (main effect of Accent, $F(1,23) = 8.73, p < .01$, partial $\eta^2 = .28$). Isovoltage maps are given in Figure 4. The positive effect was largest over midline and left in frontal and central regions, but was absent at posterior regions (interaction of Accent by Laterality, $F(2,46) = 6.46, p < .01$, partial $\eta^2 = .22$; Accent by Laterality by Anteriority, $F(2,46) = 4.66, p < .01$, partial $\eta^2 = .17$; Effects of Accent (posthoc, Bonferroni corrected) were significant at left and midline sites at frontal, fronto-central, central, and centro-parietal regions, $ps < .035$, differences between +Accent and –Accent ranged from 1.11 μV at midline fronto-central to 0.44 μV at left-central sites).

b Effects at the noun. We predicted that a noun referring to the same type of object when the proper name is accented (+Accent Same) would be easier to integrate into the message. However, if the accented proper name is followed by a noun referring to a different object (+Accent Different), this noun should be hard to integrate, yielding a larger N400. Results for the critical second noun (*cake*) are shown in Figures 4 and 5. The +Accent Different condition showed the largest negativity, and the +Accent Same condition elicited the smallest negativity of the four conditions; no apparent differences were seen between the –Accent Same and –Accent Different conditions. Statistical analyses on the mean amplitude in the 300–500ms time window yielded a main effect of Type of noun ($F(1,23) = 5.61, p < .05$, partial $\eta^2 = .20$, larger negativity for Different vs. Same nouns), a main effect of Accent ($F(1,23) = 7.52, p < .05$, partial $\eta^2 = .25$, larger negativity when the proper name was accented), and an interaction of Accent by Anteriority ($F(4,92) = 24.29, p < .001$, partial $\eta^2 = .51$: the negativity for Accented conditions was largest over anterior sites). The effect of Accent could be due to the pre-noun baseline covering a part of the Accent-related positivity seen at the proper name, which also had an anterior distribution. Importantly, though, a significant interaction of Accent by Type of noun ($F(1,23) = 13.96, p < .01$, partial $\eta^2 = .38$) was obtained, with the three-way interaction between Type of noun, Accent and Anteriority approaching significance ($F(4,92) = 3.47, p = .06$, partial $\eta^2 = .13$).

In order to break down the interaction, we conducted follow-up analyses for the +Accented and –Accented conditions separately. By comparing conditions within the same Accent manipulation, we minimized effects due to Accent-related differences

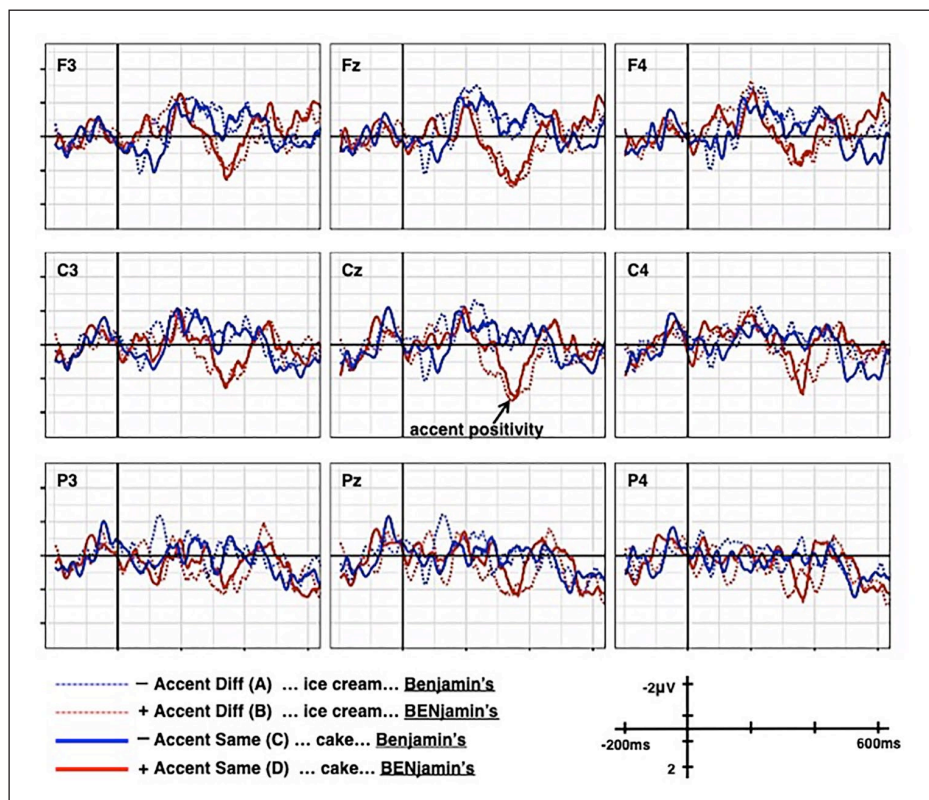


Figure 3. ERPs to the proper name (e.g. *Benjamin's*) for native English speakers for nine electrodes.

Notes. For examples of the conditions, see Table 1. Red lines denote the +Accent conditions; blue lines: -Accent conditions; dotted lines: Different noun conditions; solid lines: Same noun conditions.

in the ERPs preceding the noun. Recall that the contrastive pitch accent was correctly used in the +Accent Same condition but was incorrectly used in the +Accent Different condition. The +Accent Different condition showed a larger negativity than the +Accent Same condition ($F(1, 23) = 12.56, p < .01$, partial $\eta^2 = .35$) at all regions except frontal (Interaction of Type by Anteriority, $F(4, 92) = 11.41, p < .01$, partial $\eta^2 = .33$; pairwise comparisons of Type for each region: fronto-central through parietal: $p < .05$ to $< .001$; frontal: $p = .81$, Bonferroni corrected).

In the -Accent conditions, the pitch accent was inappropriately missing in the -Accent Same condition and correctly absent in the -Accent Different condition. No significant differences were found between the two -Accent conditions. This suggests, first, that a missing accent had not much effect on the ERP signal, and second, that the repetition of the noun in the Same condition had no detectable effects in the ERP relative to the presentation of the new noun in the Different condition (Van Petten et al., 1991). The effects observed in the +Accent conditions are therefore not likely to be due to noun repetition in the Same condition.

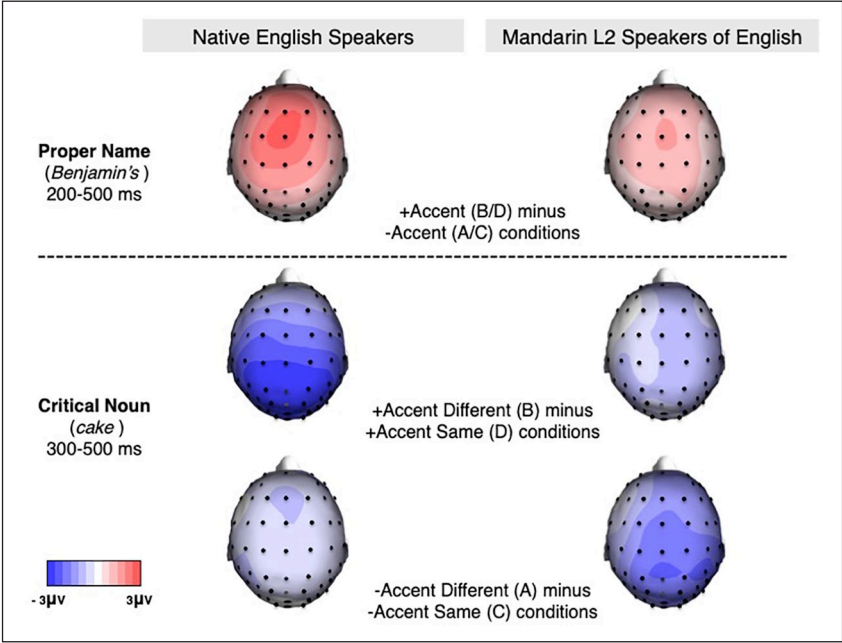


Figure 4. Isovoltage maps for the Native English speakers (left column) and L2 English speakers (right column).

Notes. Top row: Mean difference in voltage for 200–500 ms after onset of the proper name, for the + Accent minus –Accent conditions (collapsed over Type of noun); Center row: Mean difference in voltage between for 300–500 ms after onset of the critical noun for the +Accent Different minus +Accent Same conditions; Bottom row: Mean difference in voltage between 300–500 ms after onset of the critical noun for the –Accent Different minus –Accent Same conditions.

c Summary of results for the Native English speakers. Overall, consistent with previous studies, native speakers of English showed sensitivity to the presence of contrastive pitch accent: a broad positivity in a 200–500 ms time window was seen for the proper name when it was accented vs. when it was not. More importantly, when the proper name was accented, the following noun showed large negativity when it referred to a different object compared to when it referred to an object that had been previously mentioned. This suggests that native speakers were able to use contrastive pitch accent as a cue to integrate information as shown by semantic facilitation of a noun with an appropriate pitch accent and integration difficulty with an inappropriate pitch accent.

3 ERP results from Mandarin L2 learners of English

a Effects of contrastive pitch accent at the proper name. ERPs for the proper name in the L2 speakers are given in Figures 4 and 6. The +Accented conditions elicited a numerically larger positivity in the 200–500 ms window ($M = 0.02 \mu V$, $SD = 0.19 \mu V$) than the –Accented conditions ($M = -0.27 \mu V$, $SD = 0.15 \mu V$), especially at frontal and central

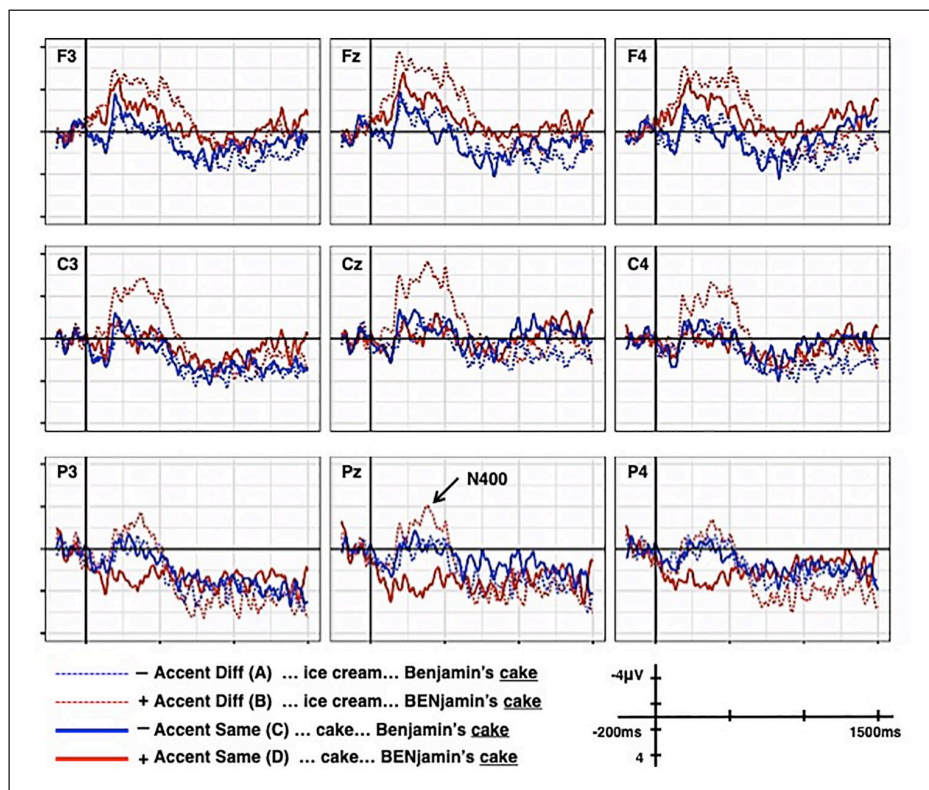


Figure 5. ERPs to the critical noun (e.g. *cake*) for native English speakers.

Notes. Red lines denote the +Accent conditions; blue lines: -Accent conditions; dotted lines: Different noun conditions; solid lines: Same noun conditions.

sites along the midline. The interaction between Accent, Anteriority and Laterality just failed to reach significance ($F(8, 176) = 2.61, p = .06$, partial $\eta^2 = .11$). No other effects of Type of noun were found.

b Effects at the second noun. ERPs to the second noun for the L2 speakers are given in Figures 4 and 7. In contrast to L1 speakers, the -Accent Different condition elicited a larger central-parietal negativity than the +Accent Different condition, even though the pitch accent is inappropriately used in the latter condition. Analysis of ERPs at the second noun (*cake*) to assess the N400 effect (300–500 ms) for L2 speakers yielded a main effect of Type of noun ($F(1,22) = 5.86, p < .05$, partial $\eta^2 = .21$), indicating that the Different noun conditions elicited a larger negativity than the Same noun conditions; an interaction of Accent and Anteriority ($F(4,88) = 15.99, p < .001$, partial $\eta^2 = .42$) indicating that ERPs in the +Accent conditions were more negative than in the -Accent conditions at frontal sites, but less negative at parietal sites; and a nearly significant interaction of Accent by Type of noun ($F(1,22) = 4.03, p = .06$, partial $\eta^2 = .16$).

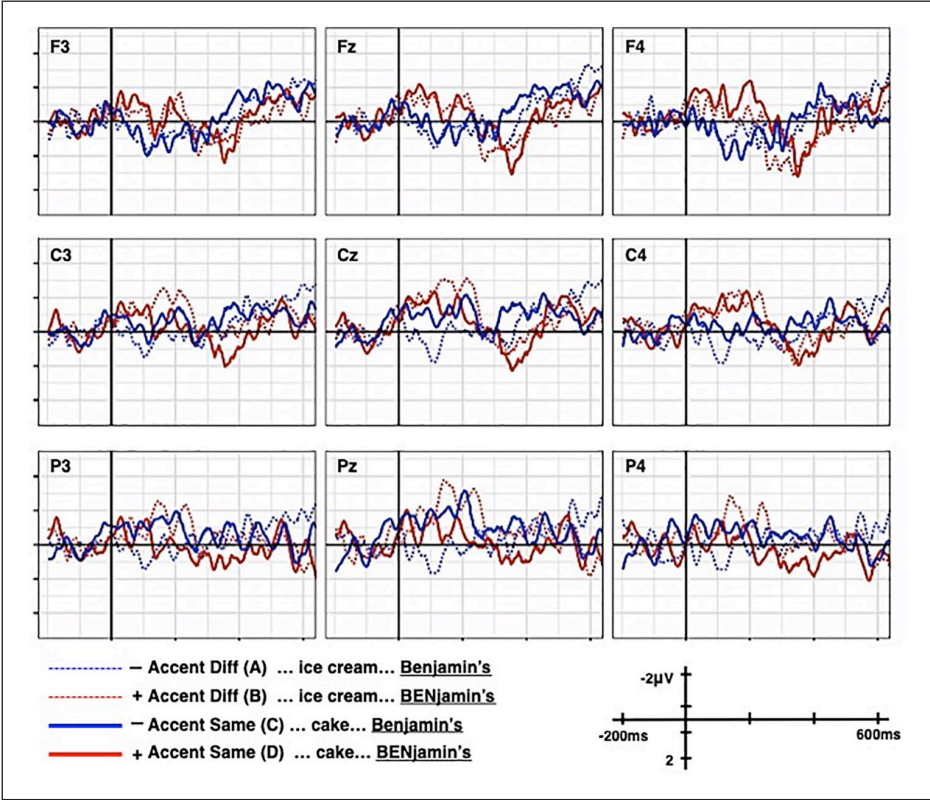


Figure 6. ERPs to the critical proper name (e.g. *Benjamin's*) for Mandarin L2 speakers of English.

Notes. Red lines denote the +Accent conditions; blue lines: –Accent conditions; dotted lines: Different noun conditions; solid lines: Same noun conditions.

Follow-up comparisons for the +Accent and –Accent conditions separately showed that the ERPs for the Different condition were more negative than those for the Same condition when the preceding proper name was not accented (–Accent conditions, $F(1,22) = 8.70, p < .01$, partial $\eta^2 = .28$); but ERPs to the Same and Different conditions did not differ when the proper name was accented (+Accent conditions, $F < 1$). This is in contrast to the native English speakers, who showed the largest difference between Same and Different noun in the +Accented conditions.

c Effects of proficiency. We had expected that more proficient L2 speakers would show more native English like responses. To test the relation between English proficiency and the Accent-related positivity seen at the proper name, we calculated the difference in the 200–500 ms window between the +Accent and –Accent conditions for the Cz electrode, collapsed over type of noun. Performance on the LexTALE proficiency task did not correlate with that on the MELICET, so we tested the effects of these proficiency

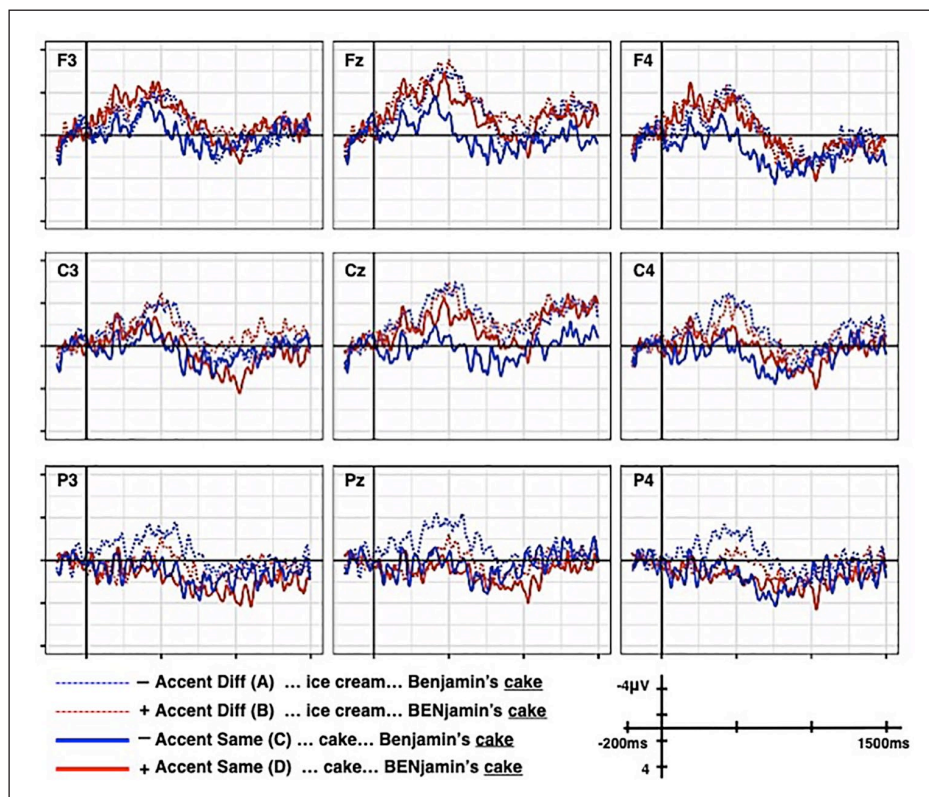


Figure 7. ERPs to the critical noun (e.g. *cake*) for Mandarin L2 speakers of English.

Notes. Red lines denote the +Accent conditions; blue lines: -Accent conditions; dotted lines: Different noun conditions; solid lines: Same noun conditions.

tasks separately. The difference in amplitude for the +Accent vs. -Accent condition negatively correlated with MELICET scores ($r = -.45$, $df = 21$, $p = .03$). In contrast to our expectation, the L2 speakers with the lower proficiency demonstrated enhanced positivity for accented conditions than non-accented conditions.

Next, we tested whether, as proficiency increased, the ERPs at the noun would show (1) a larger negativity for the +Accent Different vs. the +Accent Same conditions; and (2) less of a negativity for the -Accent Different vs. the -Accent Same conditions. For each comparison, we calculated the mean difference in ERPs in the 300–500 ms time window at Cz and tested the correlation with LexTALE and MELICET scores. Only the second comparison showed a small correlation in the expected direction with the MELICET score: the higher the scores on the MELICET task, the smaller the difference between the -Accent Same and -Accent Different conditions at the noun ($r = -.45$, $df = 21$, $p = .03$). In other words, the negativity seen for the -Accent Different conditions in the L2 speakers was less pronounced in those learners with higher English proficiency in grammar.

d Summary of the L2 results. In sum, the L2 speakers showed a numeric effect of the contrastive pitch accent on the proper name. The L2 speakers who were less proficient in English showed a stronger sensitivity to the pitch accent on the proper name. At the following noun, the ERPs were more negative for the condition in which there was no pitch accent and the noun referred to a different object. This effect was reduced in more proficient learners.

4 Comparing the groups directly

To compare the groups directly, we conducted an analysis in which we included Language Group as a between-participants factor. The analysis on the proper name data did not yield any significant interactions involving Language Group. The analyses on the critical noun revealed a marginal interaction of Accent by Language ($F(1,45) = 3.83, p = .06$, partial $\eta^2 = .08$) and a significant three-way interaction between Type of Noun by Accent by Language Group ($F(1,45) = 15.23, p < .001$, partial $\eta^2 = .25$), supporting the observation above that the native and L2 speakers differ with respect to their ERP effects at the critical noun.

IV Discussion

1 Aims and results

The aim of the current study was to investigate to what extent native English speakers and L2 speakers are sensitive to prosodic differences in the L2 and can use prosodic cues to build information structure, and the role of proficiency therein. Native speakers, and to some extent L2 speakers, showed a 200–500 ms positive-going component in the ERPs for the pitch accented words vs. non-accented words. This positivity was larger for less proficient L2 speakers, suggesting that less-proficient L2 speakers may pay more attention to modulations of the acoustic signal. In addition, both L2 and native groups responded differently at the critical noun to conditions with and without an accent. This suggests both groups noticed the difference in prosody at the proper name.

More critical for our research question were the effects observed at the noun since these could inform us how the prosodic information was used. Native English speakers showed the N400 pattern we expected. When the proper name was infelicitously accented, a large negativity was seen for the following noun (*We took Angela's ice cream, but not BENjamin's cake*) relative to a noun following a proper name with a felicitous contrastive pitch accent (*We took Angela's cake, but not BENjamin's cake*); there was no difference between the –Accent Same and –Accent Different conditions.

On the other hand, the L2 speakers showed a very different pattern: the largest negativity occurred at the noun for the –Accent Different condition (*We took Angela's ice cream, but not Benjamin's cake*), which, according to native English standards, is felicitous given the absence of an accent on the preceding proper name. The negativity for the –Accent Different conditions was smaller in more proficient L2 speakers.

These differences in ERP responses suggest that the two groups used the pitch accent in different ways during sentence comprehension. Below we will discuss the ERP effects

observed and what these can tell us about native and L2 processing of contrastive pitch accent. We are aware that the ERP results observed at the critical noun may have been partially affected by the use of a baseline that covered the effects of Accent seen at the preceding proper name. However, the effects at the proper name were not different between the Same and Different conditions; the ERPs at the noun, on the other hand, did show a significant difference between the Same and Different conditions in the +Accented conditions for the native speakers and the –Accented conditions in the L2 speakers. This suggests that not all of our findings can be attributed to baseline differences. One, of course, cannot exclude that effects observed at the noun may have already started before the onset of the noun, so any interpretation should be treated with caution.

2 Native speakers: Negativity for noun after infelicitous accent

We interpret the effects in the native speakers as a modulation of the N400 component, which may index the ease of integration of the noun (e.g. Dimitrova et al., 2012; Magne et al., 2005). For instance, in *We took Angela's cake but not BENjamin's ...*, the pitch accent on the proper name suggests that Angela's cake will be contrasted with Benjamin's cake; that is, the focus set will contain elements that have been mentioned before (*cakes*). If the next word is *cake*, this expectation is fulfilled, whereas if the next word is different, the integration of this different noun is harder. This accounts for the difference observed between the +Accent Different and Same conditions. In the non-accented conditions, there is no clear expectation of what is to come next. This could account for the N400 falling in between the two accented conditions.

The absence of a difference between the two –Accent conditions was not unexpected given previous reports of no effects for missing accents (Dimitrova, 2012; Li et al., 2018). In studies manipulating intonational phrasal boundaries (Pauker et al., 2011), smaller effects for missing accents than for superfluous accents have been accounted for by assuming that the superfluous accent needs to be mentally erased in order to obtain the correct interpretation of the sentence. This is a more elaborate revision process than mentally inserting a missing accent (Pauker et al., 2011).

The larger differences in the N400 in the +Accent compared to –Accent condition may also be due to the accent on the proper name drawing attention to the processing of the subsequent noun (see Wang et al., 2011; but see also below). The native listeners may then have noticed more easily that the noun following the accented proper noun is new (Different) or old (Same). This may also have contributed to larger N400 effects for Different vs. Same nouns in the +Accent than –Accent conditions.³

3 ERP effects in L2 speakers

In contrast to the native English speakers, the Mandarin learners of English in our study showed a negativity for the –Accent Different condition. This is the condition in which the prosody was intended to be plausible by native English standards. The negativity was smaller for more proficient L2 speakers. We are, however, careful in interpreting these results since the negativity starts right at the onset of the noun and may therefore be triggered by the proper name rather than the noun. In addition, the effect

of proficiency seen at the noun may in part have been modulated by the proficiency effects observed on the preceding proper name.

The ERP results in the L2 speakers were rather unexpected but can be accounted for in the following way. The critical noun in the Accent Different is a new noun in the context, whereas the nouns in the Same conditions are repeated from the first clause. In terms of lexical access, the nouns in the Same conditions are therefore easier to access than those in the Different conditions. An accent on the preceding word may have drawn attention toward the noun, which may have made lexical access easier for L2 speakers. This may have led to a reduced negativity at the noun in the +Accented Different condition, on par with the Same conditions in which lexical access is easier because of noun repetition. This leaves the –Accent Different as the condition for which lexical access is hardest. In our less proficient L2 speakers, lexical access of new nouns (Different conditions) may have been harder to start with and may have been additionally affected by the lack of accent at the preceding proper name. This can account for the observation that the size of the negativity for the –Accent Different vs. Same condition is smaller for more proficient L2 speakers. If this explanation is correct, L2 speakers use the accent in a different way from the native speakers. In native speakers, the accent may draw attention to the given or new status of the noun, leading to an increased negativity for the +Accent Different condition and a decrease in the +Accent Same condition; in the L2 speakers, the effect of the accent is to ease lexical access, resulting in a larger N400 in the –Accent Different conditions vs. the other conditions.

In sum, our data suggest that both L2 and native speakers are sensitive to differences in prosodic prominence but use these cues differently. We suggest that native speakers use these cues to inform information structure, whereas in L2 speakers, the presence of these cues mainly affects lexical processing. These differences between L2 and native speakers can be accounted for in several ways. First, despite superficial similarities between Mandarin and English in the prosodic expression of contrast, Mandarin speakers signal pitch accent differently since the identity of the lexical tone needs to be retained (Chen and Braun, 2006; Chen and Gussenhoven, 2008; Xu, 1999). L2 Mandarin speakers of English may therefore not interpret the English L+H* pitch accent as contrastive or may not systematically do so, depending on task demands.

In addition, or instead, the mapping of prosodic cues onto information structure may have been too complex for the L2 speakers in the current study. As we explained earlier, several steps are involved in arriving at the intended information structure in our materials. First, when processing *but*, a contrast is set up between the first and second clause. At the proper name in the second clause, another contrast set is constructed, namely between objects belonging to the first and the second protagonist. If the proper name is accented, the latter contrast needs to be even more restricted to objects of the same type as mentioned in the first clause. This may have been too complex for the L2 speakers in our study.

Suggestive evidence that complexity may have played a role comes from an eye-tracking study conducted in our lab (Perdomo & Kaan, submitted) using participants from the same L2 population. This study used a subset of the materials from the current study in a visual world paradigm, in which the participants looked at a display with pictures while listening. For instance, in *We took Mary's cake but not BENjamin's*

... the pictures were Mary's ice cream, Mary's cake, Benjamin's ice cream and Benjamin's cake. The display therefore greatly restricted the potential contrast sets. In addition, the pictures may have primed the lexical items to be used in the sentences, making lexical access easier. Unlike in the current ERP study, the eye-tracking results for Mandarin L2 speakers of English were very similar to that of native English speakers: before the critical target noun occurred in the spoken input, both groups initially showed a preference to look at objects that had not been mentioned, but this preference was reduced when the proper noun was contrastively accented. Both L2 and native English participants preferred to look at the target over the competitor earlier in the +Accent Same than -Accent Same condition, and earlier in the -Accent Different than +Accent Different conditions. This suggests that by reducing the options for contrast sets, the task can be made easier for L2 speakers, and they can use contrastive pitch accent in the same way as native English speakers do. This is in line with other work suggesting that information structure is easier to process for L2 speakers if the mapping between cues is simple (one-to-one mapping) rather than complex (Patterson et al., 2017; Shafer et al., 2015). A follow-up to the current study in which simpler scenarios are used (e.g. *This is Mary's ice cream/cake, and this is Benjamin's/BENjamin's ice cream*) combining auditory and visual stimuli, could shed more light on the use of contrastive pitch accent and the role of complexity. Such a follow-up could also include more highly advanced learners of English to test whether native-like use of prosodic cues to build information structure is achievable for highly-proficient speakers or whether the integration of information across domains remains a bottle neck (Hopp, 2009).

V Conclusions

In summary, our data confirm findings from previous studies showing that L2 speakers perform non-native-like when using prosodic cues during listening (Akker and Cutler, 2003; Chen and Lai, 2011; Namjoshi, 2015) and support the view that L2 speakers have difficulty integrating information from various domains (in this case, prosody and information structure), as proposed by the Interface Hypothesis (Sorace, 2011), or that L2 speakers do not build detailed information structure (Patterson et al. 2017), especially not when the mapping between syntax, prosody and information structure involves many steps (Shafer et al., 2015).

Prosodic structure is essential for communication since differences in prosodic prominence and phrasing can change the information structure and message of the utterance. Yet, the processing and acquisition of prosodic information has remained under-investigated, in particular in L2 learners. Our study shows that L2 speakers can detect differences in prosodic prominence in the L2, but may not necessarily use them in the way native speakers do to infer what is old and new information, and what is contrasted with what, especially not under demanding circumstances. Future research should investigate ways to make it easier for L2 speakers to use and interpret prosodic information. This, in turn, could inform teaching methods and make it easier for L2 learners to learn which prosodic cues are critical and how these relate to information structure.

Acknowledgements

The authors would especially like to thank Souad Kheder, Ethan Kutlu and Lauren Stookey for their help with data collection, and Julia Barrow, Naomi Braun, Torrey Cranston, Ivette De Aguiar, Rosemary Murray, and Lisha Yang for their help preparing and checking the stimulus materials.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. The familiarity rating of the rejected participant (31%) was two standard deviations below L2 participants' overall rating mean ($M = 69\%$, $SD = 18\%$).
2. We also collected forward and backward digit span scores. These measures did not correlate with the ERP measures, and hence will not be further discussed.
3. A reviewer pointed out to us that the +Accent Different (B) condition can be rendered felicitous by stressing the noun: *We ate Angela's ice cream, but saved BENjamin's CAKE in the fridge*. The effects we find at the noun can therefore also be interpreted as a missing accent. However, since we did not explicitly manipulate accent on the noun, we cannot further test this interpretation. Regardless of the interpretation of the effects, our results support the view that native English speakers use prosody to inform information structure and notice when the prosody and information structure do not match.

ORCID iDs

Michelle Perdomo  <https://orcid.org/0000-0002-8845-7384>

Edith Kaan  <https://orcid.org/0000-0002-0862-2993>

References

- Akker E and Cutler A (2003) Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition* 6: 81–96.
- Beckman ME and Ayers Elam G (1994) *Guidelines for ToBI labelling*. Online MS and accompanying files. Available at: http://www.ling.ohio-state.edu/phonetics/E_ToBI
- Beckman ME and Hirschberg J (1994) *The ToBI annotation conventions*. Online MS. Available at: http://www.ling.ohio-state.edu/~tobi/ame_tobi/annotation_conventions.html
- Boersma P and Weenink D (2017) *Praat: Doing phonetics by computer: Version 6.0.36*. Available at: <http://www.praat.org/> (accessed March 2019).
- Bögels S, Schriefers H, Vonk W, and Chwilla DJ (2011) Pitch accents in context: How listeners process accentuation in referential communication. *Neuropsychologia* 49: 2022–36.
- Chen A and Lai V (2011) *Comb or coat*: The role of intonation in online reference resolution in a second language. In: Zonneveld W and Quené H (eds) *Sound and sounds: Studies presented to M.E.H. (Bert) Schouten on the occasion of his 65th birthday*. Utrecht: UiL-OTS, pp. 57–68.
- Chen Y (2006) Durational adjustment under corrective focus in Standard Chinese. *Journal of Phonetics* 34: 176–201.

- Chen Y and Braun B (2006) Prosodic realization in information structure categories in standard Chinese. In: Hoffmann R and Mixdorff H (eds) *Proceedings of Speech prosody*. Dresden: TUD Press.
- Chen Y and Gussenhoven C (2008) Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36: 724–46.
- Cowles HW, Kluender R, Kutas M, and Polinsky M (2007) Violations of information structure: An electrophysiological study of answers to wh-questions. *Brain and Language* 102: 228–42.
- Dahan D, Tanenhaus MK, and Chambers CG (2002) Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language* 47: 292–314.
- Dimitrova DV (2012) *Neural correlates of prosody and information structure*. PhD dissertation, Rijksuniversiteit Groningen, The Netherlands.
- Dimitrova DV, Stowe LA, Redeker G, and Hoeks JCJ (2012) Less is not more: Neural responses to missing and superfluous accents in context. *Journal of Cognitive Neuroscience* 24: 2400–18.
- Greenhouse SW and Geisser S (1959) On methods in the analysis of profile data. *Psychometrika*, 24: 95–112.
- Halliday MAK (1967) *Intonation and grammar in British English*. The Hague: Mouton.
- Heim S and Alter K (2006) Prosodic pitch accents in language comprehension and production: ERP data and acoustic analysis *Acta Neurobiologiae Experimentalis* 66: 55–68.
- Hopp H (2009) The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition* 12: 463–83.
- Hruska C, Alter K, Steinhauer K, and Steube A (2001) Misleading dialogues: Human’s brain reaction to prosodic information. *Paper presented at ORAGE 2001*, Aix en Provence, France.
- Huang BH and Jun S-A (2011) The effect of age on the acquisition of second language prosody. *Language and Speech* 54: 387–414.
- Ito K and Speer S (2008) Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language* 58: 541–73.
- Ito K, Bibyk SA, Wagner L, and Speer SR (2014) Interpretation of contrastive pitch accent in six- to eleven-year-old English-speaking children (and adults). *Journal of Child Language* 41: 84–110.
- Krahmer E and Swerts M (2001) On the alleged existence of contrastive accents. *Speech Communication* 34: 391–405.
- Lemhöfer K and Broersma M (2012) Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavioral Research* 44: 325–43.
- Li W, Deng N, Yang Y, and Wang L (2018) Process focus and accentuation at different positions in dialogues: An ERP study. *Language, Cognition and Neuroscience* 33: 255–74.
- Li X, Chen Y, and Yang Y (2011) Immediate integration of different types of prosodic information during on-line spoken language comprehension: An ERP study. *Brain Research* 1386: 139–52.
- Magne C, Astésano C, Lacheret-Dujour A, Morel M, Alter K, and Besson M (2005) On-line processing of ‘pop-out’ words in spoken French dialogues. *Journal of Cognitive Neuroscience* 17: 740–56.
- Marian V, Blumenfeld HK, and Kaushanskaya M (2007) The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multi-linguals. *Journal of Speech and Hearing Research* 50: 940–67.
- Mennen I and de Leeuw E (2014) Beyond segments. *Studies in Second Language Acquisition* 36: 183–94.
- Michigan English Language Institute (2001) *Michigan English Language Institute College English Test*. Ann Arbor, MI: Michigan English Language Institute.
- Namjoshi J (2015) *The processing and production of prosodic focus in French by native and non-native speakers*. PhD dissertation, University of Illinois at Urbana-Champaign, IL, USA. ERIC ED589671.

- Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9: 97–113.
- Ortega-Llebaria M and Colantoni L (2014) L2 English intonation: Relations between form–meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition* 36: 331–53.
- Patterson C, Esaulova Y, and Felser C (2017) The impact of focus on pronoun resolution in native and non-native sentence comprehension. *Second Language Research* 33: 403–29.
- Pauker E, Itzhak I, Baum SR, and Steinhauer K (2011) Effects of cooperating and conflicting prosody in spoken English garden path sentences: ERP evidence for the boundary deletion hypothesis. *Journal of Cognitive Neuroscience* 23: 2731–51.
- Perdomo M and Kaan E (submitted) Integrating prosodic cues in second-language speech processing: A visual world eye-tracking study.
- Reichle RV (2010) Near-nativelike processing of contrastive focus in L2 French. In: VanPatten B and Jegerski J (eds) *Research on second language processing and parsing*. Amsterdam: John Benjamins, pp. 321–44.
- Reichle RV and Birdsong D (2014) Processing focus structure in L1 and L2 French. *Studies in Second Language Acquisition* 36: 535–64.
- Sedivy JC, Tanenhaus MK, Spivey-Knowlton MJ, Eberhard K, and Carlson GN (1995) Using intonationally marked presuppositional information in on-line language processing: Evidence from eye movements to a visual model. In: Moore JD and Fain Lehman J (eds) *Proceedings of the 17th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, pp. 375–80.
- Shafer AJ, Takeda A, Rohde H, and Grüter T (2015) Mapping prosody to reference in L2. *Poster at BUCLD 40*. Boston, MA.
- Sheppard SM, Midgley KJ, Love T, Shapiro LP, and Holcomb PJ (2018) Electrophysiological evidence for the interaction of prosody and thematic fit during sentence comprehension. *Language, Cognition and Neuroscience* 33: 547–62.
- Sorace A (2011) Pinning down the concept of ‘interface’ in bilingualism. *Linguistic Approaches to Bilingualism* 1: 1–33.
- Steinhauer K (2003) Electrophysiological correlates of prosody and punctuation. *Brain and Language* 86: 142–64.
- Takahashi C, Kao S, Baek H, Yeung AHL, Hwang J, and Broselow E (2018) Native and non-native speaker processing and production of contrastive focus prosody. *Proceedings of the Linguistics Society of America* 3: 1–13.
- Toepel U and Alter K (2004) On the independence of information structure processing from prosody. In: Steube A (ed.) *Information structure: Theoretical and empirical evidence*. Berlin: Walter de Gruyter, pp. 227–40.
- Van Berkum JJA, Koornneef AW, Otten M, and Nieuwland MS (2007) Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research* 1146: 158–71.
- Van Petten C, Kutas M, Kluender R, Mitchiner M, and McIsaac H (1991) Fractionating the word repetition effect with Event-Related Potentials. *Journal of Cognitive Neuroscience* 3: 131–50.
- Wang L, Bastiaansen M, Yang Y, and Hagoort P (2011) The influence of information structure on the depth of semantic processing: How focus and pitch accent determine the size of the N400 effect. *Neuropsychologia* 49: 813–20.
- Xu Y (1999) Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27: 55–105.
- Zubizarreta ML and Nava E (2011) Encoding discourse-based meaning: Prosody vs. syntax. Implications for second language acquisition. *Lingua* 121: 652–69.