Does the prosodic emphasis of sentential context cause deeper lexical-semantic processing?

Angèle Brunellière, C. Auran, L. Delrue

Title: Does the prosodic emphasis of sentential context cause deeper lexical-semantic processing?

Authors: Angèle Brunellière[1], Cyril Auran[2], Laurence Delrue[2]

Affiliation:

[1] Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France

[2] Univ. Lille, CNRS, UMR 8163 - STL - Savoirs Textes Langage, F-59000 Lille, France

Correspondence:

Angèle Brunellière
SCALab, CNRS UMR 9193, Université de Lille, Domaine Universitaire du Pont de Bois, BP 60149, 59653 Villeneuve d'Acsq, France
Tel: (+33) 3 20 41 72 04
angele.brunelliere@univ-lille.fr

Abstract

Speakers use prosodic emphasis to express the content of their message in order to help listeners to infer meaning. By measuring event-related potentials (ERPs) to semantically congruent and incongruent final words embedded in a sentential context that was emphasized or de-emphasized, we investigated whether prosodic emphasis conveyed by a sentential context leads listeners to a finer semantic analysis. The negative shift (N400) triggered by the difficulty to combine the incongruent word with the sentence representation was increased by prosodic emphasis at an early stage. Over the later stages, the amplitude of the N400 wave was increased by prosodic emphasis of the sentential context, whatever the semantic congruency of final words. As shown by the N400 wave, emphasizing a sentential context affected the lexical-semantic processing of the following word. This study provides clear evidence that prosodic emphasis plays a role in the semantic analysis of sentences by inducing a deeper analysis.

Count: 150 words

Introduction

Understanding the message of spoken sentences requires the processing of sounds into meaning. In spoken language comprehension, the accurate content of a message is highlighted by speakers accentuating words or segments in an utterance. The prosodic structure of speech formed by changes in timing, amplitude and frequency can convey the semantic information embedded in an utterance (for a review, Cutler, Dahan, & van Donselaar, 1997). However, the ability of listeners to exploit accented cues for the semantic analysis of discourse remains elusive. To gain better understanding of the impact of accentuation cues for the semantic analysis of discourse, we investigated whether emphasizing a sentential context induces consequences of semantic processing in discourse that affect the processing of the following words.

The advantage of accented speech in the semantic analysis of discourse has been well documented in the literature on accentuation and information structure. For instance, by monitoring eye fixations, Dahan, Tanenhaus and Chambers (2002) showed that the interpretation of referent status (e.g., whether it is a given or new entity) occurs very quickly and depends on accentuation. Listeners interpreted accented words from their first syllable as referring to a new entity and deaccented words as referring to a given entity. More recently, electrophysiological studies investigated how accentuation influences the analysis of information structure in spoken language comprehension (Dimitrova, Stowe, & Hoeks, 2015; Dimitrova, Stowe, Redeker, & Hoeks, 2012; Magne, Astésano, Lacheret-Dujour, Morel, Alter & Besson, 2005; Li, Hagoort, & Yang, 2008; Li & Lu, 2011; Li & Ren, 2012; Li & Yang, 2013; Wang, Bastiaansen, Yang & Hagoort, 2011). In particular, some electrophysiological studies in Dutch and French (Bögels, Schriefers, Vonk, & Chwilla, 2011; Dimitrova, Stowe, Redeker, & Hoeks, 2012; Magne, Astésano, Lacheret-Dujour, Morel, Alter & Besson, 2005)

focused on the analysis of information structure deriving from the distinction between focus and background information within discourse. After listening to a question context, listeners can expect the position at which a word should be marked prosodically as a focus within a sentence. Using such an experimental design, Magne et al. (2005) showed that incongruous prosodic patterns in French elicited positive-going potentials beginning as soon as 300 ms in comparison with congruous prosodic patterns when target words associated with these prosodic patterns occurred in a medial sentence position. The positive-going potentials observed resembled the classical P300, which is known to be triggered by unexpected stimuli around 300 ms when participants are instructed to pay attention to them (Picton, 1992; Donchin, & Coles, 1988). No such positive effects were found when the target words occurred at the final position of sentences (Magne et al., 2005). Instead, a negativity was found after incongruous prosodic patterns. Similarly, increased negativities for incongruous prosodic patterns were observed when participants were not instructed to perform a prosodic judgment task in Dutch but instead to perform a comprehension task or to identify referred objects (Bögels, Schriefers, Vonk, & Chwilla, 2011; Dimitrova, Stowe, Redeker, & Hoeks, 2012). Beyond the precise electrophysiological signatures of prosodic processing depending on processing and task demands, these studies above all highlight the fact that when listening to speech, humans are sensitive to the appropriateness of its prosody. By so doing, they are able to construct coherent representations of the informational structure of the speech. In this vein, a recent electrophysiological study (Dimitrova, Stowe, & Hoeks, 2015) investigated the contribution of different types of accents (new information accents and corrective accents) for the analysis of informational structure.

Even though accented speech appears to play a role in the semantic analysis of discourse, an intriguing question is whether emphasizing words induces a finer semantic

analysis. The electrophysiological correlate of semantic processing is the N400, a negative-going potential peaking around 400 ms after stimulus onset with a centro-parietal maximum. It is mostly associated with lexical access in both written and spoken modalities (Kutas & Federmeier, 2000). In sentence contexts, the amplitude of the N400 wave is more pronounced after words that are semantically incongruent or which fit less with contextually based constraints from the meaning of the utterance (Brunellière & Soto-Faraco, 2013, 2015; Connolly & Phillips, 1994; Connolly, Phillips, Stewart, & Brake, 1992; Connolly, Stewart, & Phillips, 1990; Hagoort, & Brown, 2000). In spoken language comprehension, the N400 reflects the lexical-semantic processing between the activation of lexical candidates from the speech input and the sentence representation built by the semantic analysis of the preceding words. Regarding the influence of accented speech in spoken language comprehension, it has been shown that the amplitude of the N400 wave is larger for accented than for deaccented words (Li, Hagoort, & Yang, 2008). This supports an attention allocation account in which listeners pay attention to parts of the utterance where the accent falls (Cutler, 1976; Sanford, Sanford, Molle, & Emmott, 2006). In line with this view, although a larger N400 amplitude was found for accent-inappropriate (accented) than accent-appropriate (deaccented) words in old information, the reverse pattern was observed in new information with a larger N400 for accent-appropriate (accented) than accent-inappropriate (deaccented) words (Li, Hagoort, & Yang, 2008).

Moreover, some authors consider that the allocation of attention to accented words produces more detailed semantic processing (Birch and Garnsey, 1995; Blutner and Sommer, 1988; Fraundorf, Watson, & Benjamin, 2010; Li & Lu, 2011). For instance, better retention in memory was found for contrastive accent words embedded in sentential context (Birch and Garnsey, 1995; Blutner and Sommer, 1988; Fraundorf, Watson, & Benjamin, 2010).

Accentuation can even facilitate a semantic priming effect between two words under adverse conditions of high-load sentences (Li & Lu, 2011). A larger N400 amplitude was found for low semantic relatedness than for high semantic relatedness when only the prime word was accented. Nonetheless, this advantage of accented speech was not observed with low-load sentences, a condition in which a semantic priming effect was observed whether the prime word was accented or not (Li & Lu, 2011). Other electrophysiological studies suggest that the easiness to combine the processing of an incoming word with the sentence representation might be facilitated or interrupted thanks to accented speech (Li & Ren, 2012; Li & Yang, 2013; Wang, Bastiaansen, Yang & Hagoort, 2011). For example, Li & Ren (2012) found that the larger amplitude of the N400 wave after semantically incongruous words from the preceding sentential context in comparison with congruous words occurred earlier when the new information was produced with emphasis than when weakly accented. Additionally, when the new information was deaccented, the N400 was not sensitive to combining the processing of the incoming word with the sentence representation because of the inconsistent accent.

Nonetheless, studies to date have mixed the perceptual processing of accented speech and the consequences of accented speech in semantic processing (e.g., using critical words where accent and semantic information are manipulated together). When speech is accented, word duration and spectral clarity are increased (Dahan and Bernard, 1996; Eefting, 1991; Klatt, 1976; van Santen and Olive, 1990), thereby providing more salient cues to be processed. On the contrary, one of the consequences of semantic processing in discourse is that the specificity of semantic representations at sentential level is increased, thereby facilitating the activation of relevant semantic information. Therefore, as suggested by Norris, Cutler, McQueen & Butterfield (2006), more than having a deeper semantic analysis of

accented words, the presence of salient cues anywhere in a spoken sentence could encourage listeners to increase the semantic analysis of sentence representation.

Using the ecological phenomenon of prosodic emphasis in spoken language, we explored whether the prosodic emphasis of sentential context can induce a finer semantic analysis by gauging event-related potentials (ERPs) after final words embedded in a sentential context produced with or without prosodic emphasis and introducing a semantic anomaly or not. In particular, we examined the easiness to process an incoming word according to emphasizing the preceding sentential context although the incoming word was not emphasized. Therefore, we focused on the influence of prosodic emphasis of sentential context on the lexical-semantic processing of final words in spoken language comprehension. Prosodic emphasis can arise in ecological conversations when the speaker's intention is to highlight the content of message and to convince the interlocutor. For example, an increased fundamental frequency, intensity and duration were found in French stressed words during emphasis discourse (Jun & Fougeron, 2002, Lacheret-Dujour & Beaugendre 1999; Touati 1987, Pasdeloup 1990). Studying the influence of natural prosodic emphasis in semantic processing of sentence representation provides a better understanding of the expressive function provided by prosody.

To disentangle the perceptual processing of accented speech and the consequences of semantic processing, we used parenthetical structure and grammatical cataphora in French spoken sentences. Grammatical cataphora is the use of a pronoun to refer ahead to another word in a sentence (called, a postcedent). For example, in the sentence « Créer des bonbons : **ils** ont un beau métier, les **confiseurs** *Creating sweets, they have a great job, confectioners*», the pronoun "ils" appears earlier than "les confiseurs" that refers to it. This grammatical cataphora places the postcedent "les confiseurs" in the position of a parenthetical element.

The postcedent is thus isolated from the prosodic structure of the preceding clause, making it less sensitive to acoustic variations in the expressivity of speakers due to a sentential context (e.g., "Créer des bonbons : **ils** ont un beau métier). The acoustic cues of the postcedent based on fundamental frequency, intensity and duration did not differ in the expressivity occurring earlier in the sentence (i.e. emphasis compared to without emphasis). We thus explored the impact of prosodic emphasis of sentential context on semantics in right-dislocated structures in French and manipulated two factors, the prosodic emphasis of sentential context (without/with emphasis) and the semantic congruency of the final word (congruous/incongruous), thereby creating four experimental conditions (see, Table 1). The final word was semantically incongruous or expected from the semantic constraints given by the sentential context. The experimental design did not include any incongruous prosodic patterns and participants were instructed to listen to the spoken sentences to understand their meaning, instead of performing an explicit prosodic judgment task.

< Insert Table 1 about here >

We hypothesized that the prosodic emphasis of sentential context would produce more detailed semantic processing of the meaning of the sentence provided by the sentential context. A finer semantic analysis of sentential context should affect the processing of an incoming word embedded in a sentential context. If the prosodic emphasis of sentential context causes deeper lexical-semantic processing, the amplitude of the N400 wave after semantically incongruous words from the preceding sentential context in comparison with congruous words should be stronger when the sentential context is produced with emphasis than when it is not. More exactly, a finer semantic analysis of sentential context could increase both the sensitivity to detect the semantic anomaly and the ease of processing

8

semantically congruous words over the N400 wave. Moreover, increasing the specificity of sentence representation from the sentential context could induce better memory retention of congruous words with respect to the semantic constraints provided by the sentential context when the sentential context is produced with emphasis. Hence, an interactive effect between the prosodic emphasis of sentential context and the semantic congruency of final words was expected on the N400 amplitude and in terms of memory retention performance.

Methods

Participants

Twenty-six right-handed students (17 females, 9 males) from the University of Lille participated in the experiment. All were French native speakers and aged between 18 and 30 (mean age: 22.3, standard deviation: 2.5) and none reported any hearing or language impairments. The handedness of participants was assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Participants gave their written consent before beginning the experiment. The study was approved by the Research Ethics Committee of the University of Lille.

Materials

The experimental stimuli consisted of a set of 392 semantically strongly constraining sentence frames that were ended either by the most expected word or by a semantically incongruous word. The set of sentence frames was selected from a classical cloze procedure. Ninety French speakers with the same characteristics as the participants in the ERP experiment completed sentence fragments with the first word that came to mind. During this classical cloze procedure, each participant was exposed to 100 sentence frames and each list of 100 sentence frames randomly presented was completed by 15 participants. The 392 selected sentence frames had a mean cloze probability of 0.82 (range: 0.53-1) and a mean length of

15.8 words (range: 8-34). All sentence frames were semantically plausible spoken sentences and were composed of a grammatical cataphora. This created the openness of a parenthesis marked by a prosodic boundary between the final word with its article and the sentence frame. In the congruous condition, the sentence frame was ended by the expected word. In the semantically incongruous condition, the final word differed from the expected word and produced the detection of a semantic anomaly (i.e., their word's cloze probabilities were always 0). The semantically incongruous word shared gender and number with the expected word. The final words selected from the French database Lexique (New, Pallier, Ferrand, & Matos, 2001) were nouns beginning with a plosive (/p/, /t/, /k/). The plosive provided a clear physical marker to align the ERP recordings to word onset. The expected and semantically incongruous words were matched for lexical frequency, number of phonemes, number of syllables and number of phonological neighborhoods (see, Table 2).

< Insert Table 2 about here >

A French-speaking male was asked to produce each sentence several times with natural prosody at normal speaking rate. All sentences sharing the same sentence frame were successively spoken and the order of the four versions within one sentence frame varied for each sentence frame to prevent the effect of first reading in a particular condition. The speaker was asked to pronounce sentences either with emphasis or without emphasis as in neutral prosody. Stimuli were digitized at a sampling rate of 44 kHz with 16-bit. The stimuli with the best pronunciation, natural intonation and speaking rate were selected by two native speakers. Ten native French speakers judged the sentences in the condition with emphasis as being more expressive than those without emphasis. As in prior literature describing increased fundamental frequency, intensity and duration during emphasized discourse (Jun & Fougeron, 2002, Lacheret-Dujour & Beaugendre 1999; Touati 1987, Pasdeloup 1990), acoustic

measurements (duration, F0, intensity) confirmed that the sentential contexts were spoken with an intended emphasis (see, Figure 1 and Table 3).

< Insert Figure 1 and Table 3 about here >

Acoustic values (duration, F0, intensity) were extracted from the auditory stimuli with the speech-editing software Praat (version 5.3; Boersma & Weenink, 2011). The total duration of sentence context up to the onset of the article of the final word was longer in the emphasis condition ($F(1,391)=136.74$, $MSE=0.5$, $p<.001$, duration difference: 144.6 ms). Statistical analysis of the mean and standard deviation fundamental frequency (F0) of sentence context revealed higher fundamental frequency values with prosodic emphasis than without emphasis (mean F0, $F(1,391)= 3142.2$, $MSE=208$, $p<.001$, difference: 40.8 Hz; standard deviation F0, $F(1,391)=2537.8$, $MSE=63$, $p<.001$, difference: 20.2 Hz). Additionally, the mean intensity of sentence context was stronger with prosodic emphasis than without emphasis ($F(1,391)=3056.8$, $MSE=3$, $p<.001$, difference: 4.6 dB). Such acoustic differences were not observed between congruous and incongruous sentences. No interactive effects between the factors, expressive prosody and semantic congruency were observed. Crucially, the acoustic differences observed between the with emphasis and without emphasis conditions in sentential contexts were greater than the sensitivity threshold typically described in ERP studies. For example, previous ERP studies found that the sensitivity to changes in intensity ranged from 2 to 3 dB at 500 Hz and 3000 Hz when the auditory N1 response or the auditory Mismatch Negativity were measured (e.g., Harris, Mills, & Dubno, 2007; Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007). Regarding fundamental frequency and duration, changes in these two acoustic parameters elicited differential ERP responses when the difference was equal to or greater than 5 Hz and 8 ms, respectively (Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007). On the contrary, acoustic measurements based on the

11

final noun indicated differences between the with emphasis and without emphasis conditions that were below the sensitivity threshold typically described in ERP studies (mean F0, difference: 2.5 Hz; standard deviation F0, difference: 0.9 Hz; mean intensity, difference: 1.1 dB; duration, difference: 3.7 ms). It appeared that final words were never produced with emphasis. Acoustic measurements based on the final noun did not reveal any significant difference between the conditions.

Four experimental lists of 98 trials per condition were constructed so that each sentence frame and each final word were presented only once to each participant. Participants were tested in all experimental conditions. To reduce the development of strategies on the semantically incongruous conditions, 196 filler congruous sentences were added to the experimental design (total ratio of incongruous sentences: 33%).

Experimental procedure

A fixation cross appeared at the center of the screen 500 ms before the onset of the auditory sentence and remained until 1,000 ms after the end of it (see Figure 2). Auditory sentences were presented binaurally at a comfortable sound pressure level via headphones. After a 2,000-ms intertrial interval, the next fixation cross was presented. To reduce motor artifacts, participants were asked to avoid making any movements when the fixation cross appeared. Participants listened to 24 training sentences, then a set of seven blocks of 84 trials. Each block lasted around 10 min and was composed of sentences from all experimental conditions and fillers randomly presented.

< Insert Figure 2 about here >

During auditory stimulation, participants were instructed to listen to the auditory sentences attentively for comprehension (for similar approaches, see Brunellière & Soto-Faraco, 2013, 2015; Hagoort & Brown, 2000; van den Brink, Brown, & Hagoort, 2001; van den Brink & Hagoort, 2004). To maintain attention when listening to spoken sentences,

participants were informed that they should perform a lexical recognition task after having heard the spoken sentences. They performed a lexical recognition task on a set of 168 words presented in the center of the screen one by one. They were asked to indicate as quickly and accurately as possible whether they had already listened to them during the auditory stimulation by pushing one of two buttons on a response box. The response buttons were counterbalanced across all participants. After a 2,000-ms intertrial interval, another fixation cross was presented. The visual presentation of words made it possible to probe the memory trace of words while avoiding the familiarity effects associated with acoustic or phonetic properties. Eighty-four of the words were never heard as sentence-final words during auditory stimulation while the others had been presented during it, i.e. 84 old words. Among the words listened to, half of them were expected from the sentence frames that were pronounced either with emphasis or without emphasis, while the other half comprised semantically incongruous words. The words never listened to as sentence-final words during auditory stimulation were divided into two groups: half of them were completely new without being linked to auditory stimulation (48) and the other half (48) were expected words from the sentence context, although not presented at the end of the sentence, i.e. in the incongruous version. In this way, words were matched for lexical frequency, number of phonemes, number of syllables, number of phonological and orthographical neighbors, and uniqueness point.

EEG recording

The electrical signal was recorded from the scalp using a 128-channel Biosemi Active Two AD box with a sampling rate of 1024 Hz. Two electrodes measured the eye movements from the right eye and two additional electrodes were placed over the right and left mastoids. Individual electrodes were adjusted to a stable offset lower than 20 mV. The EEG epochs started 100 ms before and lasted 900 ms after the onset of the final word. Each epoch was

filtered offline with a 0.1–30 Hz band-pass filter and a 50 Hz notch filter. This was corrected to a 100-ms baseline. Epochs were rejected under a rejection criterion of ± 70 µV on any channel. ERP waveforms were calculated for each participant, experimental condition and electrode and were composed of at least 80 trials for each participant and condition. The total number of accepted epochs was equal across the experimental conditions (NP-Congruous: 84, NP-Incongruous: 84, EP-Congruous: 83, EP-Incongruous: 83). Bad channels were interpolated for each participant (Perrin, Pernier, Bertrand, Giard, & Echallier, 1987) and the EEG signal was re-referenced offline to an average mastoid reference (left and right).

ERP analyses

By visual inspection of the ERP waves, we extracted the mean amplitude of the N400 component in two different time windows that we hereafter refer to as the early and late N400 time windows (260 ms-360 ms, 400-500 ms). The selection of these two time windows was mainly based on the ERP waves triggered by the semantically incongruous words when the sentential context was emphasized. To probe effects earlier or later than the N400 component, we explored whether the N2/P2 component (150-250 ms) and the late positivity (700-900 ms) were influenced by the prosodic emphasis and the semantic congruency of the final word. A three-way repeated analysis of variance (ANOVA) was conducted on the mean amplitude over each time window with independent variables: prosodic emphasis (2: without vs. with emphasis) and semantic congruency (semantically congruous vs. incongruous) and site (7: Left Anterior, Right Anterior, Frontocentral, Centroparietal, Left parietal, Right parietal, and Occipito-parietal). Each scalp site contained 9 channels and the scalp sites were chosen to provide appropriate scalp topography for the components of interest (Kutas & Federmeier, 2000; Van den Brink & Hagoort, 2004). To adjust for violations of sphericity (Greenhouse & Geisser, 1959), the Greenhouse-Geisser correction was applied when there was more than one degree of freedom in the numerator. The corrected *p* values are reported. When a significant

14

interaction was found, post-hoc Tukey tests were performed to interpret the significance of the effects. Only the significant effects are reported in the text. The main effects of site are presented only in Table 4. Additionally, a three-way repeated measures ANOVA (same factors as previously described) over the 100-ms prestimulus baseline or the time window between 50 and 150 ms after the onset of the final words did not reveal any main effects or interactions, suggesting that the pre-stimulus and the period before 150 ms of word onset did not affect the pattern of results between experimental conditions.

Results

Behavioral results

Participants performed the lexical recognition task accurately as assessed by the mean *d*-prime of 0.7 above chance ($t(25)=10.1$, $p<.001$), suggesting that they paid attention to the words embedded in the spoken sentences during auditory stimulation. *d*-prime values were obtained from hit responses (i.e., words presented during auditory stimulation and for which participants pressed the button corresponding to 'heard words') and false alarms (i.e., words that had not been presented during auditory stimulation but for which participants pressed the button corresponding to 'heard words'). To investigate whether prosodic emphasis could induce a facilitated retention in memory due to a deeper sentence representation, a two-way repeated-measures ANOVA was performed on hit rates with the factors prosodic emphasis (2: without vs. with emphasis) and semantic congruency (2: congruous vs. incongruous words). As seen in Figure 3, the hit rates were higher for congruous words than for incongruous words ($F(1,25)=28.6$, $MSE=0.03$, $p<.001$). However, neither a main effect of prosodic emphasis ($F(1,25)=2.9$, $MSE=0.008$) nor any interactive effects ($F(1,25)=0.22$, $MSE=0.01$) were found. To test whether the false alarms could be affected by prosodic emphasis, we computed a one-way ANOVA from false alarms divided into three experimental groups: completely new

15

items (24), words not heard but expected from a sentential context spoken with emphasis (24) and words not heard but expected from a sentential context spoken without emphasis (24). The analysis revealed a main effect of experimental group ($F(2,50)=13.98$, $MSE=136.11$, $p<.001$) with higher false alarms for words not heard but expected than completely new items ($p<.001$), irrespective of the expressivity of the sentential context. No significant advantage of prosodic emphasis was found with false alarms ($p>.2$)

< Insert Figure 3 about here >

EEG results

Grand-average waveforms corresponding to semantically congruous and incongruous final words are shown for each level of expressivity (see, Figure 4). The final words induced a clear positivity followed by a negativity around 200 ms in all experimental conditions. Nonetheless, the typical auditory N100 response was not observed, possibly because we measured ERPs to words embedded in continuous speech. The lack of a sufficiently long pause interval between final words and the preceding words can cause a reduced N100 amplitude or even disrupt the N100 due to its refractory period (e.g., Connolly et al., 1992; Hagoort & Brown, 2000; Näätänen & Picton, 1987). Interestingly, the semantic incongruency seemed to elicit a larger negative wave than congruous words, with a maximum peak at 330 ms after word onset in the emphasis condition. Moreover, this negative shift (N400) triggered by the semantic incongruency seemed to be more enhanced in the emphasis condition between 400 and 500 ms. Thereafter, a positivity wave appeared particularly over posterior sites and occurring around 700 ms.

< Insert Figure 4 and Table 4 about here >

In the N2/P2 time window, the analysis revealed a significant interaction between semantic congruency and site ($F(6,150)=4.79$, $MSE=1.64$, $p<.01$). The amplitude of the N2/P2

16

component was greater for incongruous words than for congruous words only over left and right anterior sites ($p<.001$) where a stronger amplitude of P2 was found for incongruous words than for congruous words. No other effects were shown (see Table 4).

In the early N400 time window, the ANOVA indicated a main effect of semantic congruency ($F(1,25)=39.25$, $MSE=95.77$, $p<.001$). The amplitude of the N400 was stronger for incongruous words than congruous ones. In line with prior literature, the effect of semantic congruency differed over particular topographical sites as shown by the significant interaction between semantic congruency and site ($F(6,150)=8.24$, $MSE=2.84$, $p<.001$). An effect of semantic congruency was found over all sites ($p<.001$), but ANOVA based on the ERP difference between incongruous and congruous words with site as factor showed that the strongest effect of semantic congruency occurred over the centroparietal sites compared to the other sites ($p<.05$). Moreover, the effect of semantic congruency was higher over the left and right parietal sites and occipito-parietal sites than over the left and right anterior sites and frontocentral sites ($p<.05$). Importantly, a significant interaction between semantic congruency and prosodic emphasis and site was shown ($F(6,150)=2.96$, $MSE=1.01$, $p<.05$). To examine on which sites prosodic emphasis affected the effect of semantic congruency, we extracted the ERP difference between incongruous and congruous words at each level of prosodic emphasis. As seen in Figure 5, the effect of semantic congruency was stronger in the emphasis condition than in the without-emphasis condition over some topographical sites. An ANOVA based on the ERP difference between incongruous and congruous words at each level of prosodic emphasis confirmed a significant interaction between prosodic emphasis and site ($F(6,150)=2.96$, $MSE=1.82$, $p<.05$). Prosodic emphasis elicited a stronger effect of semantic congruency over the left anterior and frontocentral sites ($p<.05$).

< Insert Figure 5 about here >

17

Similar to the early N400 time window, a main effect of semantic congruency ($F(1,25)$=154.18, MSE=61.95, p<.001) was found in the late N400 time window. The amplitude of the N400 was higher for incongruous words than congruous ones. The effect of semantic congruency was topographically distributed as shown by the significant interaction between semantic congruency and site ($F(6,150)$=4.13, MSE=3.77, p<.05). An effect of semantic congruency was found over all sites (p<.001), but ANOVA based on the ERP difference between incongruous and congruous words with site as factor showed that the strongest effect of semantic congruency was found over the centroparietal sites compared to the other sites (p<.05). Moreover, a significant interaction between prosodic emphasis and site was shown ($F(6,150)$=2.84, MSE=2.05, p<.05). The amplitude of the N400 was higher in the emphasis condition than in the without-emphasis condition only over right parietal and occipito-parietal sites (p<.05). In sum, the N400 presented differential patterns over the two selected time windows. To better highlight these differential patterns, we computed mean amplitudes after normalization to the global field power (standard deviation of electrical activity across all channels) and compared the ERP response in the early N400 time window to that in the late N400 time window. A four-way repeated measures ANOVA using time window (early versus late N400), semantic congruency, prosodic emphasis and site as factors confirmed the differential patterns previously observed by showing a main effect of time window $F(1,25)$=15.14, MSE=60.65, p<.001) and interactive effects between time window and the other factors of interest (time window $\times$ semantic congruency, $F(1,25)$=7.41, MSE=45.11, p<.05; time window $\times$ site, $F(6,150)$=19.53, MSE=1.29, p<.001; time window $\times$ semantic congruency $\times$ prosodic emphasis $\times$ site, $F(6,150)$=3.12, MSE=0.58, p<.05).

In the P600 time window, statistical analysis revealed a significant interaction between semantic congruency and site ($F(6,150)$=25.59, MSE=2.15, p<.001). Unlike the N400, the

effect of semantic congruency was not observed across all topographical sites but only over left and right anterior, frontocentral and centroparietal and occipito-parietal sites ($p<.05$). Whereas more negative values for incongruous words than congruous words were found over left and right anterior, frontocentral and centroparietal sites, the pattern was reversed over occipito-parietal sites with more positive values for incongruous words. Similar to the late N400 time window, the effect of semantic congruency elicited a negative shift over left and right anterior, frontocentral and centroparietal sites during the P600 time window. The effect of semantic congruency observed over the late N400 time window somewhat persisted over the P600 time window, although another form of semantic congruency effects, marked by a positive shift for incongruous words, emerged over occipito-parietal sites in the same time window.

To sum up, semantic congruency affected the processing of final words across all time windows, unlike prosodic emphasis. An impact of prosodic emphasis was observed only over the stages of the N400 component. Over the early N400 time window, the effect of semantic congruency was increased by the prosodic emphasis over the left anterior and frontocentral sites. Incongruous words elicited a negative shift relative to congruous words. However, over the late N400 time window, this interactive effect between the semantic congruency and prosodic emphasis was not shown. Instead, a main effect of prosodic emphasis was found with a higher N400 amplitude in the emphasis condition than in the without-emphasis condition over the right parietal and occipito-parietal sites.

Discussion

The present study investigated the contribution of prosodic emphasis to the semantic analysis of sentence comprehension. Prosodic emphasis arises in ecological conversations when the speaker's intention is to highlight the content of a message, to help the listener

understand the message and to convince one's interlocutor. As previously reported, discourse pronounced with prosodic emphasis elicits increased fundamental frequency, intensity and duration, providing more salient cues for the interlocuter to process (Jun & Fougeron, 2002; Lacheret-Dujour & Beaugendre 1999; Touati 1987; Pasdeloup 1990). According to the theoretical hypothesis posited in the introduction, the presence of salient cues anywhere in a spoken sentence could encourage listeners to increase the semantic analysis of sentence representation. In other words, a finer semantic analysis of a sentential context containing emphasis should affect the ease of processing an incoming word embedded in the sentence. If the prosodic emphasis expressed by the sentential context causes deeper lexical-semantic processing, the amplitude of the N400 wave after semantically incongruous words in comparison with congruous words should be greater when the sentential context is produced with emphasis than when it is not. Moreover, increasing the specificity of sentence representation in emphasized discourse could induce better memory retention of semantically congruous words, as expected from the semantic constraints provided by a sentential context.

Our results show that the difficulty of combining the incoming word with sentence representation produced larger amplitudes after semantically incongruous words than congruous words beginning as soon as 150 ms after word onset and persisting over the following time windows of the N400 and P600 waves. In line with the hypothesis of this study, the larger amplitude observed after incongruous words than after congruous words was stronger in the early N400 time window[1] (260-360 ms) when the sentential context was produced with emphasis than when it was produced without emphasis. This effect was

---

[1] The early N400 resembles in timing the Phonological Mapping Negativity (PMN) traditionally occurring between 270 and 310 ms (see Steinhauer & Connolly, 2008). It has been suggested that this component indexes a phonological stage of word processing that serves to map phonological code in the activation of a lexical cohort of candidates. However, the latter component can be also found during the processing of pseudowords and words. We cannot rule out the emergence of PMN or an overlapping between the N400 and the PMN between 260 and 360 ms. We called the component observed over this time window the early N400, because the strongest semantic congruency effects were observed over centroparietal sites, as usually described in the N400 literature.

observed over the left anterior and frontocentral sites. Interestingly, prosodic emphasis induced a larger amplitude of the N400 in a later time window (400-500 ms) over the right parietal and occipito-parietal sites, irrespective of the semantic congruency of final words. Thereafter, in a lexical recognition task, congruous words were better recognized than incongruous words, independently of prosodic emphasis.

Above all, our electrophysiological findings are consistent with the role of prosody in lexical-semantic processing during spoken sentence comprehension. This is evidenced by the fact that prosodic emphasis affected only the eliciting of the N400 wave. It is well known that the N400 wave reflects lexical-semantic processing between the activation of lexical candidates from speech input and the sentence representation built by the semantic analysis of preceding words (Brunellière & Soto-Faraco, 2013, 2015; Connolly & Phillips, 1994; Connolly, Phillips, Stewart, & Brake, 1992; Connolly, Stewart, & Phillips, 1990; Hagoort, & Brown, 2000). The N400 is sensitive to semantic processing and the N400 effect of semantic congruency indexes the integration of word meaning in discourse. Furthermore, N2/P2 waves have been found to reflect phonological processing (Chen, Zhang, Xu, Scheepers, Yang, & Tanenhaus, 2016; Connolly & Phillips, 1994; Connolly, Phillips, Stewart, & Brake, 1992; Connolly, Stewart, & Phillips, 1990). In line with our findings, P200 amplitude has already been observed to increase after semantic anomalies (Bohan, Leuthold, Hijikata, & Sanford, 2012). N2/P2 waves could reflect the correspondence between the phonological processing of the incoming word and the contextually based constraints arising from the meaning of the utterance. As expected, the amplitude of N2/P2 waves was not influenced by the prosodic emphasis given by the sentential context because these electrophysiological waves are associated with phonological processing and not with processing at a lexical-semantic level. In addition, prosodic emphasis expressed in the sentential context does not play a role during

repair analysis, as suggested by the P600. The amplitude of the P600 wave, which is known to reflect repair analysis after linguistic anomalies (e.g., Münte, Heinze, Matzke, Wieringa, & Johannes, 1998), did not vary as function of prosodic emphasis. It thus seems that the repair analysis that is necessary to process sentence closure takes place independently of the prosodic emphasis contained in the sentential context. Nonetheless, this may be mainly explained by the strong semantic constraints of sentential contexts, such that the integration of an incongruous word is completely discordant with the representation of the sentence during the repair analysis. It appeared that the P600 reflected general processes of semantic repair to integrate word meaning and sentence representation. Previous studies have already described a sensitivity of the P600 to prosodic cues (e.g., Dimitrova, Stowe, Redeker, & Hoeks, 2012). In those cases, the P600 was related to reanalysis of prosodic incongruity between the information structure provided by the context and word accentuation. In addition, an earlier positivity, called P300, has been observed after prosodic incongruities when the violation of focus accent occurred in the middle of a sentence. The authors suggested that this component was triggered by unexpected stimuli (Magne et al., 2005). In the present study, however, the comparison in terms of prosodic emphasis did not reveal any prosodic incongruity since the two different patterns of prosody in the sentential context were natural pronunciations.

In line with the theoretical hypothesis proposed in the introduction, an influence of prosodic emphasis during sentence comprehension was found in this study, yet the task did not require the listeners to make explicit judgments about the prosodic cues. Moreover, the potential advantage of prosodic emphasis in the semantic analysis of sentences was examined by disentangling the perceptual processing of accented speech and the consequences of semantic processing. Up to now, most studies have used either off-line measurements with explicit judgments or have gauged the processing of words from which the perceptual processing of accented speech was mixed with the consequences of semantic processing

(Birch and Garnsey, 1995; Blutner and Sommer, 1988; Fraundorf, Watson, Benjamin, 2010; Li, Hagoort, & Yang, 2008; Li & Lu, 2011; Li & Ren, 2012; Li & Yang, 2013; Wang, Bastiaansen, Yang & Hagoort, 2011). In this study, whereas the sentential context could provide more salient cues (e.g. when it was produced with emphasis), the postcedent from which we measured ERPs was not sensitive to acoustic variations in the expressivity of speakers conveyed across the sentential context, as assessed by acoustic measurements.

Under such an experimental design, it appears that the findings based on the early N400 time window are in favor of a deeper semantic analysis of sentential context thanks to the emphasis occurring early in the sentence. Over this time window, the semantic congruency effect was stronger over the left anterior and frontocentral sites when the sentential context was produced with emphasis than when it was produced without emphasis. This can be interpreted as ease or difficulty in combining the incoming word with sentence representation depending on the speech emphasized in the sentential context. The more specific semantic representations encoded from the meaning of the sentential context when delivered with emphasis could increase the sensitivity of the linguistic system to detect a possible semantic compatibility between the incoming word and representations of the sentence. In line with this view, supplementary post-hoc Tukey tests showed that the increased semantic congruency effect elicited by prosodic emphasis was due to a larger N400 amplitude for incongruous words after emphasized speech (emphasis versus without emphasis, $p<.05$). Prosodic emphasis increases the sensitivity to detect the semantic anomaly but does not lead to better processing of semantically congruous words. Moreover, the interactive effect observed over the early N400 between the prosodic emphasis of sentential context and the semantic congruency of final words can be explained by both a deeper semantic analysis of sentential context, leading to a finer sentence representation, and by a

23

deeper semantic analysis of final words induced by the prosodic emphasis occurring early in the sentence.

Surprisingly, this pattern did not last over the whole N400 time course. The second part of the N400 wave[2] was more consistent with the attention allocation account (Cutler, 1976; Sanford, Sanford, Molle, & Emmott, 2006), since N400 amplitudes were larger for any final word preceded by a sentential context produced with emphasis. The prosodic emphasis expressed by the sentential context seemed to encourage listeners to focus attention on the upcoming word over the later stages of processing that word. The influence of prosodic emphasis thus emerged independently of the semantic congruency of final words and was shown topographically over different topographical sites from the early N400 time window, i.e. right parietal and occipito-parietal sites in the case of the late N400 time window. The traditional stages proposed during the processing of spoken words (Marslen-Wilson & Welsh, 1978; Frauenfelder & Tyler, 1987) could explain the two stages of the N400 wave. During the first stage, lexical candidates from the speech input are activated until a unique lexical candidate is selected. This stage of processing could interact with the semantic details of sentence representation while the activated word with its properties remains a potential candidate. Thereafter, when the word is completely recognized from the speech input, the cognitive system could rely on the acceptability of combining the semantic properties of the recognized word with sentence representation. The constraints arising from sentence representation during word processing would be stronger when the incoming word is not yet recognized, such as during the early N400 time window. Since emphasized speech seems to trigger a deeper semantic analysis of discourse, further studies should examine in more detail the specificity of sentence representations such as manipulating shared semantic features between an incoming word and sentence representation.

_____

[2] Supplementary analyses based on frontocentral or left anterior sites separately did not reveal any significant interaction between prosodic emphasis and semantic congruency.

Beyond the influence of prosodic emphasis during the on-line processing of spoken sentences, numerous studies have demonstrated the beneficial effect of accented speech on memory retention (Birch and Garnsey, 1995; Blutner and Sommer, 1988; Fraundorf, Watson, Benjamin, 2010) and on learning of the meaning of new words (Lynott & Connell, 2010; Shintel, Anderson, & Fenn, 2014). It may be postulated that such an advantage in memory is due to the attentional allocation given to the accented words. Nonetheless, given the consequences of semantic processing driven by emphasized speech, one may hypothesize that more specific semantic representations developed when the sentential contexts are produced with emphasis could reinforce the trace memory of words expected from them. In line with this view, Neville, Kutas, Chesney, & Schmidt (1986) showed that old words that occurred previously in sentences during a listening task were better recognized when they fitted semantically with the sentence context than those that did not. Although we replicate the findings of Neville, Kutas, Chesney, & Schmidt (1986), we did not find any impact of prosodic emphasis on memory retention. This suggests that the consequences of semantic processing driven by emphasized sentential context do not lead to better memory retention.

In conclusion, listeners can exploit prosodic cues given by a sentential context in the semantic analysis of discourse. More than simply analyzing emphasized words in detail, humans listening to emphasized speech given by a sentential context may construct more specific sentence representations during the processing of an utterance.

Footnote

Footnote 1. The early N400 resembles in timing the Phonological Mapping Negativity (PMN) traditionally occurring between 270 and 310 ms (see Steinhauer & Connolly, 2008). It has been suggested that this component indexes a phonological stage of word processing that serves to map phonological code in the activation of a lexical cohort of candidates. However, the latter component can be also found during the processing of pseudowords and words. We cannot rule out the emergence of PMN or an overlapping between the N400 and the PMN between 260 and 360 ms. We called the component observed over this time window the early N400, because the strongest semantic congruency effects were observed over centroparietal sites, as usually described in the N400 literature.

Footnote 2. Supplementary analyses based on frontocentral or left anterior sites separately did not reveal any significant interaction between prosodic emphasis and semantic congruency.

References

Birch, S. L., & Garnsey, S. M. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language*, *34*, 232-267. doi: 10.1006/jmla.1995.1011

Blutner, R., & Sommer, R. (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language*, *27*, 359-367. doi: 10.1037/e636952013-638

Boersma, P., & Weenink, D. (2011). *Praat: doing phonetics by computer* [Computer program].Version 3.4, retrieved 2 Jan 2011 from http://www.praat.org/

Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D. (2011). Pitch accents in context: How listeners process accentuation in referential communication. *Neuropsychologia*, *49*, 2022-2036. doi: 10.1016/j.neuropsychologia.2011.03.032

Bohan, J., Leuthold, H., Hijikata, Y., & Sanford, A.J. (2012). The processing of good-fit semantic anomalies: an ERP investigation. *Neuropsychologia*, *50*, 3174-3184. doi: 10.1016/j.neuropsychologia.2012.09.008

Brunellière, A., & Soto-Faraco, S. (2013). The speakers' accent shapes the listeners' phonological predictions during speech perception. *Brain and Language*, *125*, 82-93. doi: 10.1016/j.bandl.2013.01.007

Brunellière, A., & Soto-Faraco, S. (2015). The interplay between semantic and phonological constraints during spoken-word comprehension. *Psychophysiology*, *52*, 46-58. doi: 10.1111/psyp.12285

Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M.K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition*, *154*, 11-21. doi: 10.1016/j.cognition.2016.05.007

Connolly, J.F., & Phillips, N.A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, *6*, 256-266. doi: 10.1162/jocn.1994.6.3.256

Connolly, J.F., Phillips, N.A., Stewart, S.H., & Brake, W.G. (1992). Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language, 43*, 1-18. doi: 10.1016/0093-934x(92)90018-a

Connolly, J.F., Stewart, S.H., & Phillips, N.A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and Language, 39*, 302-318. doi: 10.1016/0093-934x(90)90016-a

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics*, *20*, 55-60. doi: 10.3758/bf03198706

Cutler A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141-201. doi: 10.1177/002383099704000203

Dahan, D., & Bernard, J. M. (1996). Interspeaker variability in emphatic accent production in French. *Language and Speech*, *39*, 341-374. doi: 10.1177/002383099603900402

Dahan, D., Tanenhaus, M.K., & Chambers, C.G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*, 292-314. doi: 10.1016/s0749-596x(02)00001-3

Dimitrova, D.V., Stowe, L.A., & Hoeks, J.C.J. (2015). When Correction Turns Positive: Processing Corrective Prosody in Dutch. *PLoS One*, *10*, e0126299. doi: 10.1371/journal.pone.0126299

Dimitrova, D.V., Stowe, L.A., Redeker, G., & Hoeks, J.C.J. (2012). Less is not more: neural responses to missing and superfluous accents in context. *Journal of Cognitive Neuroscience*, *24*, 2400-2418. doi: 10.1162/jocn_a_00302

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 355-425. doi: 10.1017/s0140525x00058027

Eefting, W. (1991). The effect of "information value" and "accentuation" on the duration of Dutch words, syllables and segments. *Journal of the Acoustical Society of America*, *89*, 412-424. doi: 10.1121/1.400475

Fraundorf, S.H., Watson, D.G., & Benjamin, A.S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*. *63*, 367-386. doi: 10.1016/j.jml.2010.06.004

Frauenfelder, U.H., & Tyler, L.K. (1987). The process of spoken word recognition: an introduction. *Cognition*, *25*, 1-20. doi: 10.1016/0010-0277(87)90002-3

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95-111. doi: 10.1007/bf02289823

Hagoort, P., & Brown, C.M. (2000). ERP effects of listening to speech: semantic ERP effects. *Neuropsychologia*, *38*, 1518-1530. doi: 10.1016/s0028-3932(00)00052-x

Harris, F.C., Mills, J., & Dubno, J.R. (2007). Electrophysiological correlates of intensity discrimination in cortical evoked potentials of younger and older adults. *Hearing Research*, *228,* 58-68. doi: 10.1016/j.heares.2007.01.021

Jun, S.A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, *14*, 147-172. doi: 10.1515/prbs.2002.002

Klatt, D.H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208-1221. doi: 10.1121/1.380986

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *12*, 463-470. doi: 10.1016/s1364-6613(00)01560-6

Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français*. Paris. CNRS Editions.

Li, X., Hagoort, P., & Yang, Y. (2008). Event-related potential evidence on the influence of accentuation in spoken discourse comprehension in Chinese. *Journal of Cognitive Neuroscience*, *20*, 906-915. doi: 10.1162/jocn.2008.20512

Li, X-Q., & Lu, Y. (2011). How accentuation influences semantic short-term memory representations during on-line speech processing: an event-related potential study. *Neuroscience*, *193*, 217-228. doi: 10.1016/j.neuroscience.2011.07.010

Li, X-Q., & Ren, G-Q. (2012). How and when accentuation influences temporally selective attention and subsequent semantic processing during on-line spoken language comprehension: An ERP study. *Neuropsychologia*, *50*, 1882-1894. doi: 10.1016/j.neuropsychologia.2012.04.013

Li, X., & Yang, Y. (2013). How long term memory and accentuation interact during spoken language comprehension. *Neuropsychologia*, *51*, 967-978. doi: 10.1016/j.neuropsychologia.2012.12.016

Lynott, D., & Connell, L. (2010). The Effect of Prosody on Conceptual Combination. *Cognitive Science*, *34*, 1107-1123. doi: 10.1111/j.1551-6709.2010.01119.x

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., & Besson, M. (2005). On-line processing of "pop-out" words in spoken French dialogues. *Journal of Cognitive Neuroscience*, *17*, 740-756. doi: 10.1162/0898929053747667

Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29-63. doi: 10.1016/0010-0285(78)90018-x

Münte, T.F., Heinze, H.J., Matzke, M., Wieringa, B.M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. *Neuropsychologia*, 36, 217-226. doi: 10.1016/s0028-3932(97)00119-x

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*, 375-425. doi: 10.1111/j.1469-8986.1987.tb00311.x

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. L. (1986). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language*, *25*, 75-92. doi: 10.1016/0749-596x(86)90022-7

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique*, *101*, 447-462. doi: 10.3406/psy.2001.1341

Norris, D., Cutler, A., McQueen, J.M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, *53*, 146-193. doi: 10.1016/j.cogpsych.2006.03.001

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97-113. doi: 10.1016/0028-3932(71)90067-4

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology*, *118*, 177-185. doi: 10.1016/j.clinph.2006.09.001

Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole* (Doctoral dissertation). Retrieved from A.N.R.T.

Perrin, F., Pernier, J., Bertrand, O., Giard, M.-H., & Echallier, J.F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and Clinical Neurophysiology*, *66*, 75-81. doi: 10.1016/0013-4694(87)90141-6

Picton, T.W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*, 456-479. doi: 10.1097/00004691-199210000-00002

Sanford, A.J.S., Sanford, A.J., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, *42*, 109-130. doi: 10.1207/s15326950dp4202_2

Shintel, H., Anderson, N.L., & Fenn, K.M. (2014). Talk this way: the effect of prosodically conveyed semantic information on memory for novel words. *Journal of Experimental Psychology: General*, *143*, 1437-1442. doi: 10.1037/e633262013-968

Steinhauer, K., & Connolly, J.F. (2008). Event-Related Potentials in the Study of Language. In Stemmer, B. & Whitaker, H. (Eds.) *Handbook of the Neuroscience of Language*, Elsevier, Amsterdam, pgs. 91-104.

Taouti, P. (1987). *Structure prosodique du suédois et du français* (Doctoral dissertation). Retrieved from http://www.bibliotheque.bordeaux.fr.

van den Brink, D., Brown, C.M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, *13*, 967-985. doi: 10.1162/089892901753165872

van den Brink, D., & Hagoort, P. (2004). The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. *Journal of Cognitive Neuroscience*, *16*, 1068-1084. doi: 10.1162/0898929041502670

van Santen, J.P.H., & Olive, J.P. (1990). The analysis of contextual effects on segmental duration. *Computer Speech & Language*, *4*, 359-390. doi: 10.1016/0885-2308(90)90016-y

Wang, L., Bastiaansen, M., Yang, Y., & 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*, 813-820. doi: 10.1016/j.neuropsychologia.2010.12.035

Figure captions

Figure 1. Spectrograms of example of sentence frames in each experimental condition. Intensity shown in yellow and scale from 50 to 100dB (in green). Fundamental frequency (F0) shown in blue color with scale from 75Hz to 795Hz.

Figure 2. Trial Scheme.

Figure 3. Hit rates in % for old words per experimental condition (in black, congruous words, in white, incongruous words, at the left, sentential context spoken with emphasis and at the right, sentential context spoken without emphasis) and false alarms in % for unheard words (expected from a sentential context spoken with emphasis, expected from a sentential context spoken without emphasis, new). ***$p<.001$

Figure 4. Grand average of ERPs time-locked to onset of final word across four experimental conditions. NP: Neutral prosody, EP: Expressive prosody. *Significant effects due to prosodic emphasis, $p<.05$

Figure 5. Mean ERP difference between incongruous and congruous words in early N400 time window across all topographical sites (LA: Left Anterior, RA: Right Anterior, F: Frontocentral, C: Centroparietal, LP: Left parietal, RP: Right parietal, OP: Occipito-parietal). NP: Neutral prosody, EP: Expressive prosody. *$p<.05$
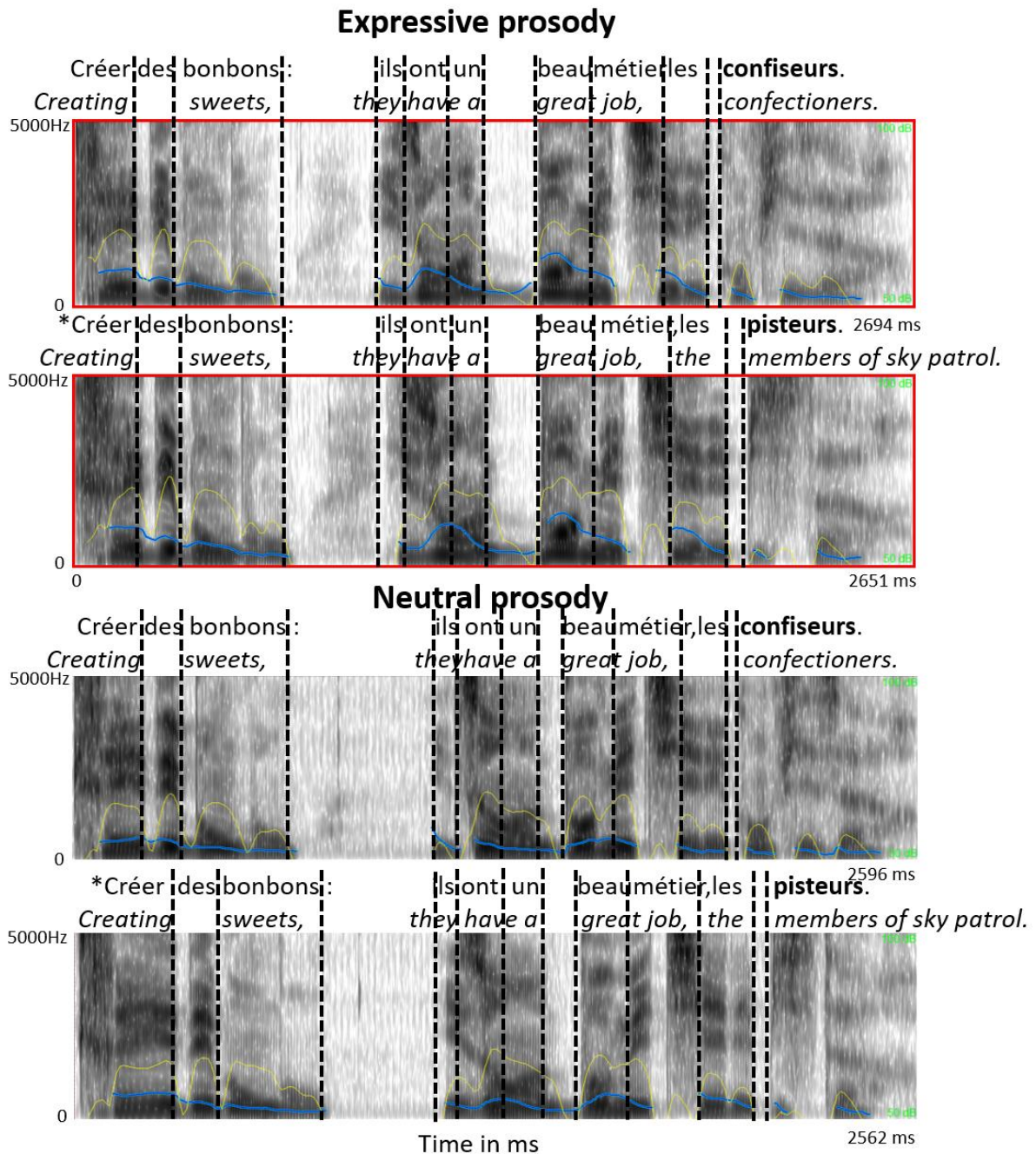
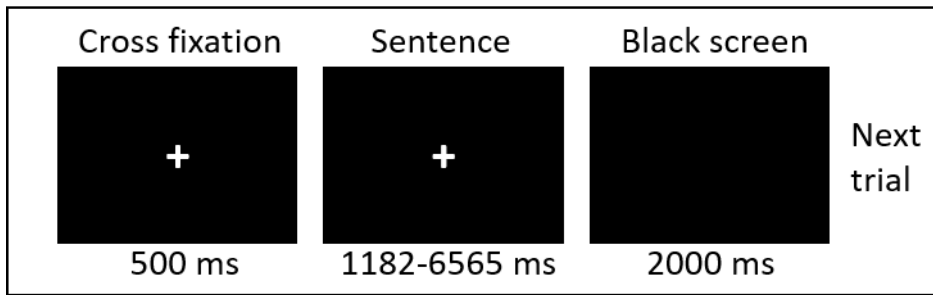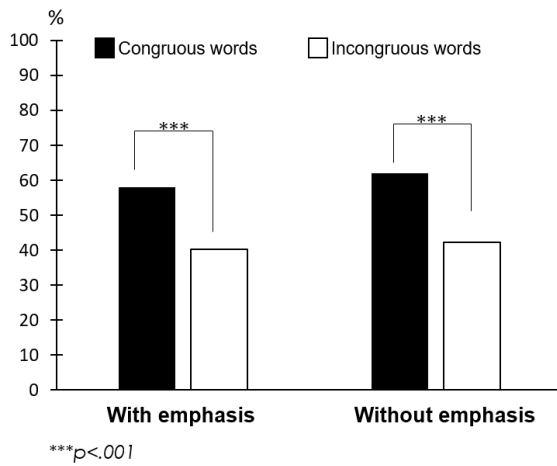# Figures

Figure 1

Figure 2

Figure 3

A. Hit rates during the lexical recognition task

***p<.001

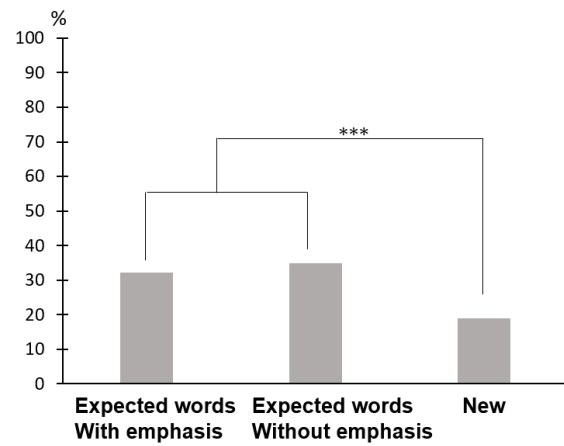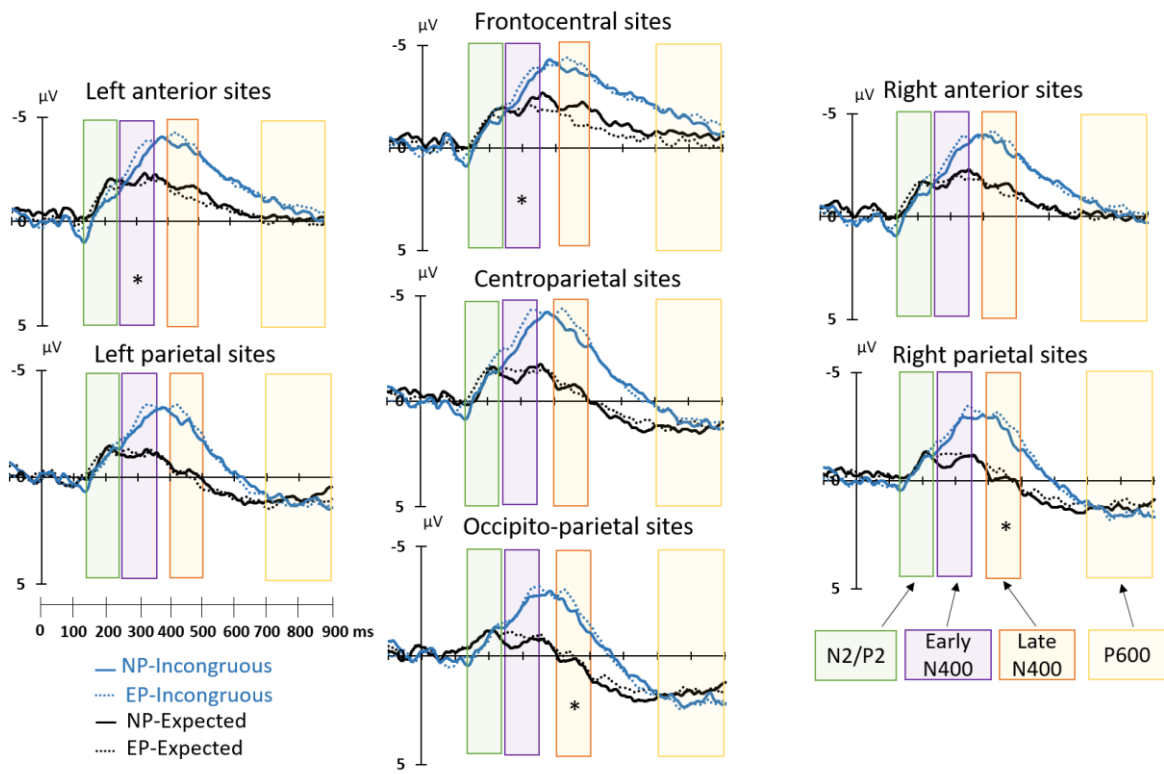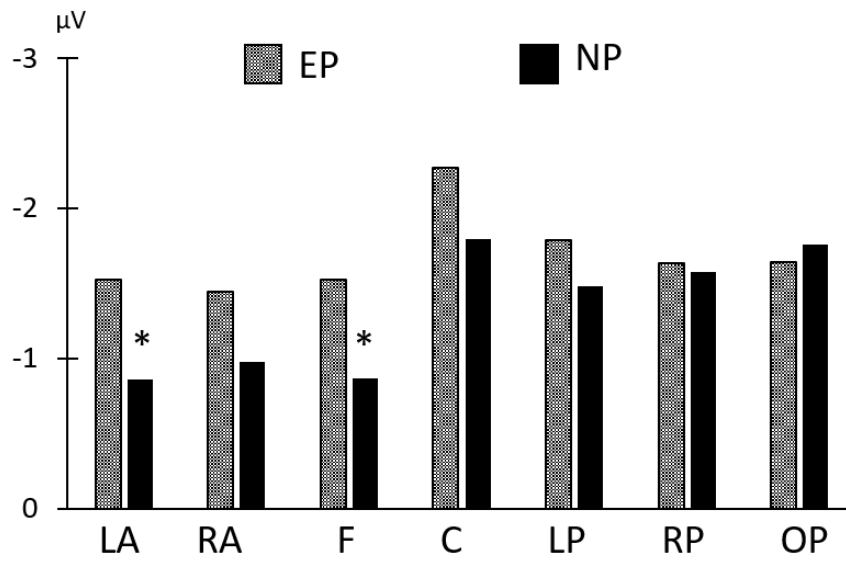B. False alarms during the lexical recognition task

Figure 4

Figure 5

Tables

Table 1. Examples of experimental conditions during auditory stimulation

Conditions	Prosodic emphasis-Sentential context	Semantic congruency-Final word	Examples
NP-Congruous	Neutral prosody (without emphasis)	Congruous	Créer des bonbons : ils ont un beau métier, les **confiseurs**. *Creating sweets, they have a great job, confectioners.*
NP-Incongruous	Neutral prosody (without emphasis)	Incongruous	Créer des bonbons : ils ont un beau métier, les **pisteurs**. *Creating sweets, they have a great job, the members of sky patrol.*
EP-Congruous	Expressive prosody (with emphasis)	Congruous	Créer des bonbons : ils ont un beau métier, les **confiseurs**. *Creating sweets, they have a great job, confectioners.*
EP-Incongruous	Expressive prosody (with emphasis)	Incongruous	Créer des bonbons : ils ont un beau métier, les **pisteurs**. *Creating sweets, they have a great job, the members of sky patrol.*

NP: Neutral prosody, EP: Expressive prosody

Table 2. Properties of final words presented during auditory stimulation

Final words	Lexical frequency	Number of phonemes	Number of syllables	Number of phonological neighborhoods
Congruous	7.3	5.7	2.3	4
Incongruous	6.9	5.8	2.3	3.7

Lexical frequency (number of occurrences per million extracted from corpus of film subtitles), **Number of phonological neighbors** (number of words by changing one phoneme without affecting the others)

Table 3. Acoustic properties of experimental conditions during auditory stimulation

Conditions	Context **Mean F0** (in Hz)	Context **SD F0** (in Hz)	Context **Mean Intensity** (in dB)	Context **Duration** (in ms)	Final word **Mean F0** (in Hz)	Final word **SD F0** (in Hz)	Final word **Mean Intensity** (in dB)	Final word **Duration** (in ms)
NP-Congruous	136.6	23.2	61.1	3014	107	8.9	52.2	395
NP-Incongruous	137.6	23.8	61.2	3027	107.2	9	52.6	398
EP-Congruous	177.8	43.6	65.7	3177	109.8	10.2	53.2	391
EP-Incongruous	178.2	43.7	65.8	3152	109.5	9.6	53.5	394

F0, Fundamental frequency**; SD,** Standard deviation.

Table 4. Statistical results of EEG analyses

N2/P2 Time window		**Early N400 Time window**	
Semantic congruency	*F*(1,25)=0.56, *p*>.2	Semantic congruency	*F*(1,25)=39.25, *p*<.001
Prosodic Emphasis	*F*(1,25)=1.22, *p*>.2	Prosodic Emphasis	*F*(1,25)=1.40, *p*>.2
Site	*F*(6,150)=3.38, *p*<.05	Site	*F*(6,150)=14.41, *p*<.001
Semantic congruency × Prosodic Emphasis	*F*(1,25)=0.005, *p*>.2	Semantic congruency × Prosodic Emphasis	*F*(1,25)=0.75, *p*>.2
Semantic congruency × Site	*F*(6,150)=4.79, *p*<.01	Semantic congruency × Site	*F*(6,150)=8.24, *p*<.001
Prosody Emphasis × Site	*F*(6,150)=1.37, *p*>.2	Prosody Emphasis × Site	*F*(6,150)=1.05, *p*>.2
Semantic congruency × Prosodic Emphasis × Site	*F*(6,150)=1.15, *p*>.2	Semantic congruency × Prosodic Emphasis × Site	*F*(6,150)=2.96, *p*<.001
Late N400 Time window		**P600 Time window**	
Semantic congruency	*F*(1,25)=154.18, *p*<.001	Semantic congruency	*F*(1,25)=1.39, *p*>.2
Prosodic Emphasis	*F*(1,25)=0.59, *p*>.2	Prosodic Emphasis	*F*(1,25)=0.04, *p*>.2
Site	*F*(6,150)=50.58, *p*<.001	Site	*F*(6,150)=87.95, *p*<.001
Semantic congruency × Prosodic Emphasis	*F*(1,25)=1.49, *p*>.2	Semantic congruency × Prosodic Emphasis	*F*(1,25)=0.09, *p*>.2
Semantic congruency × Site	*F*(6,150)=4.13, *p*<.05	Semantic congruency × Site	*F*(6,150)=25.59, *p*<.001
Prosody Emphasis × Site	*F*(6,150)=2.84, *p*<.05	Prosody Emphasis × Site	*F*(6,150)=1.52, *p*>.2
Semantic congruency × Prosodic Emphasis × Site	*F*(6,150)=1.46, *p*>.2	Semantic congruency × Prosodic Emphasis × Site	*F*(6,150)=1.25, *p*>.2