Using ERPs to capture inferential processes guided by prosodic cues

Coralie Chevallier, Mathilde Bonnefond, Jean-Baptiste Van der Henst & Ira A. Noveck

Imagine Mary asks Jack to buy her a rose or a book and he in fact gets her both. In doing so, he indicates that he interpreted *or* inclusively (i.e., a rose, or a book, or both). However, had he been asked to buy a rose OR a book (with contrastive accent), Jack might have inferred that Mary intended him to buy one gift or the other, but *not both*; in this case he interpreted the utterance exclusively. In this paper, we test 32 participants and investigate the electrophysiological correlates of contrastive accent and of the 'not both' inference which it can generate. Following Relevance Theory, and contra Levinson's theory of Generalized Conversational Implicatures, we argue that pragmatic inferencing based on such a contextual cue will elicit a distinct electrophysiological component, marking extra processing effort (i.e., the P600). Our results demonstrate the presence of a P3a associated with contrastive accent and a P600-like component as well as a Left Anterior Negativity when contrastive accent leads to pragmatic inferencing. We discuss our results in the context of the linguistic-pragmatic literature and argue that our data are consistent with the idea that the exclusive interpretation requires a deeper analysis than the inclusive interpretation.

Keywords: scalar inference, prosody, pragmatics, P3a, P600, LAN

1. Introduction

Consider the following scenario in which Mary is telling Jack what she would like for her birthday:

(1) I'd like a rose or a book.

One can assume that Mary would be more than happy if Jack came home with both a rose *and* a book and one would consider her wish was fulfilled. In this scenario, *or* is understood inclusively ('a rose or a book or both'): its reading is compatible with the rules of standard logic (e.g., Copi 1941) and it corresponds to its semantics (e.g., Grice 1989). Now imagine that Mary knew that Jack was short of money. In that case, Mary could very well insist that he buys her a

 (Received October 2009)

rose or a book, *but not both*. One way to get that message across is to stress the word *or* (Chevallier et al. 2008):

(2) I'd like a rose OR a book.[1]

This makes it clear that she intends *or* to be understood exclusively as 'one or the other, but not both'. This gloss corresponds with a pragmatically enriched reading, which is appropriate given the context of the conversation, but incompatible with standard logic. This exclusive interpretation is thought to be the result of a *scalar inference* (Horn 1972; Levinson 2000). At the root of such an inference is a weakly informative term (such as *or*) that leads the hearer to infer that the speaker has deliberately chosen not to be more informative by using a stronger term from the same scale (for instance *and*).[2]

Scalar inferences have prompted much experimental work over the last several years with respect to their development (Chierchia 2004; Gualmini et al. 2001; Noveck 2001; Papafragou & Musolino 2003; Pouscoulous et al. 2007), their time course (Bott & Noveck 2004; Breheny et al. 2006; Noveck et al. 2004), the effort they entail (De Neys & Schaeken 2007) and, most recently, with respect to prosody (Chevallier et al. 2008) or autism (Chevallier et al. *in press,* Pijnacker et al. 2009). Although one of these studies (Noveck & Posada 2003) employed ERP to investigate scalar inferences, it did not satisfactorily tackle the question of the electrophysiological underpinnings of scalar inference (as we describe later) nor was there a focus on the prosody-pragmatics interface. In the present work, we use event-related potentials to examine the influence of prosody as it affects the interpretation of *or*.

The remainder of the Introduction is broken down into the following two sections. The first focuses on the influence of contrastive accent on the interpretation of *or*. This section describes how the two possible readings of *or* are expected to differ psycholinguistically by drawing on work carried out with other weak terms such as *some*. We then provide a short review of the literature on the electrophysiology of prosody and pragmatics. This is where we spell out our hypotheses and our experimental paradigm, which benefits from previous work showing that contrastive accent favours the exclusive interpretation of *or*.

1.1. Prosody and pragmatic enrichment

We focus on two phenomena which are at stake in the production of the pragmatic interpretation associated with the scenario in (2):

one is the pragmatic enrichment of *or* and the other is the influence of prosody in triggering that enrichment. We begin by pointing out that the meaning of *or,* i.e., the component which constitutes its semantics (the meaning that it is assigned by the grammar), is inclusive (i.e., semantically, *A or B* means 'A or B or both'). However, that meaning is not rigidly fixed and can be pragmatically enriched into a more informative – exclusive – interpretation ('A or B but not both'). It is important to note though, that in (2) the semantics of *or* is unchanged ('A or B or both') but the meaning the speaker intends to convey is different. To access that *intended* meaning, the hearer will need to enrich the semantics of the disjunctive utterance via pragmatic mechanisms.

The example in (2) is an illustration of the now uncontroversial idea according to which there is a considerable gap between the meaning assigned by the grammar and the meaning intended by the speaker (Grice 1989). In fact, virtually all the utterances we hear require some degree of pragmatic enrichment (Wilson 2003). This necessity for constant pragmatic enrichment has led to theoretical accounts of mechanisms that enable us to bridge the gap between *sentence* meaning and *speaker* meaning (e.g., Sperber & Wilson 1986/1995). Among those mechanisms is the indispensable ability to take contextual cues into account when trying to uncover the deep meaning of an utterance. Prosodic information plays a prominent role in that respect since it can guide the hearer's interpretation and help him build an appropriate context for interpretation (House 2006). In other words, prosody provides insight on context interpretation. As far as our example is concerned, the presence of a contrastive accent on *or* guides the hearer towards a new interpretative context where it cannot be relevant that Mary wants a rose *and* a book (e.g., she has already received many presents from Jack and knows that he is short of money). Accent on *or* thus attracts the hearer's attention to the disjunction thereby leading him to apply deeper processing to the utterance (Chevallier et al. 2008). In other words, added emphasis leads the hearer to enrich the semantics into a more informative interpretation, namely an exclusive one. Subtle variations in the speaker's intonation are therefore crucial cues in verbal interactions insofar as they help the hearer identify the speaker's communicative intentions. To sum up, the examples in (1) and (2) demonstrate that the interpretation of *or* can be either 'logical' (inclusive) or 'pragmatic' (exclusive) and that prosody can influence the listener's interpretation.

Two modern pragmatic theories have put forward possible cognitive mechanisms underlying such inferences: Levinson's theory

of Generalized Conversational Implicatures (**GCI**)[3] and Relevance Theory. Both theories agree that the inclusive meaning of *or* corresponds to its literal meaning and that the exclusive interpretation results from an inference but the two theories differ with respect to the role that context plays in drawing these inferences. For the former theory, such inferences are, in most cases, pragmatically appropriate, and are thus made by default. That is, they are automatically triggered by the presence of the word *or* in the utterance, independently of the context (e.g., Levinson 2000: 104). Those automatic inferences speed up verbal communication by allowing fast default inferences that are pragmatically justified in most cases. Of course, in those cases (considered to be rare) where the default inference must be cancelled (leading to an inclusive interpretation), the cancellation process requires additional resources. For Relevance Theory, there are no GCIs, i.e., no default inferences in comprehension. In principle, listeners do not draw an inference unless it is contextually justified (Carston 2002: 111). In particular, Relevance theorists expect listeners to adopt an enriched exclusive interpretation of *or* only when they have contextual reasons to do so. Otherwise, the inclusive interpretation can suffice. In this way, interpreting *or* exclusively involves deeper processing than interpreting it inclusively. Of course, this is just the opposite of Levinson's GCI theory.

Prior empirical work has confirmed predictions from Relevance Theory showing that pragmatic enrichments arising out of other weak scalar terms such as *some* come at an cost in terms of time (Bott & Noveck 2004; Breheny et al. 2006) and effort (Chevallier et al. 2008; De Neys & Schaeken 2007). For instance, answers that indicate logical interpretations of *some* (i.e., to respond 'true' to an underinformative statement such as *Some cows are mammals* indicates that one has interpreted *some* as 'some and possibly all') are significantly faster than answers that indicate pragmatic interpretations (i.e., to say 'false' to such underinformative statements indicates that one has interpreted *some* as 'some but not all'); reaction times slow down from roughly 2700 ms for the logical responses to 3300 ms for pragmatic responses (see Bott & Noveck 2004: Fig. 3). This sort of result has been replicated using various methods, including highly ecologically valid text comprehension paradigms. Breheny et al. (2006) measured reading times of short scenarios involving the word *or* and showed that, when the context is neutral, participants treat the disjunction most naturally as inclusive. However, they also demonstrated that *or* can be treated exclusively, with a small cost in reading time, in contexts which make it clear that the inclusive interpretation is imprac-

tical. Collectively, these results favour a cognitive model like that of Relevance Theory under which, all things being equal, the pragmatic interpretation takes more time and more effort than the logical interpretation. In the next section, we consider how these two approaches make different predictions with respect to the neural mechanisms underlying the production of scalar inferences.

1.2. Scalars, prosody and electrophysiology

Event Related Potentials (ERPs) are obtained by recording and averaging brain potentials associated with the presentation of a specific event (e.g., listening to a specific word in the context of an utterance). After averaging, it is possible to compare the description of the evoked potential to existing descriptions in the literature thus putting one in a position to gain insight concerning the nature of the specific event which elicited the potential (e.g., if word A elicits a larger N400 than word B – a component which has traditionally been identified in response to semantic anomalies –, it will be possible to infer that there are some similarities between the processing of word A and the detection of semantic anomalies). The experiment to be presented later is designed to record ERPs among listeners who produce scalar inferences after hearing a prosodic cue. This method allows one to test pragmatic theories in a novel way since it can shed light on electrophysiological underpinning of the pragmatic processes that are resorted to during the course of interpretation. In what follows, we summarise previous findings that concern a) the (involuntary) detection of prosodic accent; b) pragmatic processes involved in the interpretation of utterances. In both cases, we will make predictions about components that are likely to emerge in the current study.

Numerous ERP studies have been carried out on various topics related to prosody, such as the understanding of emotions in the voice (e.g., Besson et al. 2002; Brosch et al. 2009; Wambacq & Jerger 2004), the interaction between syntax and prosody (e.g., Eckstein & Friederici 2005; 2006; Steinhauer et al. 1999), or the perception of word stress (e.g., Böcker et al. 1999; Friedrich et al. 2001; Friedrich et al. 2004). However, to the best of our knowledge, no ERP study has studied contrastive accent and the way it can affect pragmatic interpretations. The closest topic to the present one is the processing of *word* stress.[4] In that domain, the left frontal N325 has been identified as a possible marker of stress (Böcker et al. 1999) but those results were contradicted more recently (Friedrich et al. 2001). Even if these results had been consistent, contrastive accent is a different topic of

investigation (see Note 4). On the one hand, word stress is obligatory, involves pure decoding and is used to constrain lexical activation from a very early stage (Friedrich et al. 2004). On the other hand, contrastive accent is optional and gives access to pragmatic inferencing, presumably at a later stage in sentence processing. In this paper, we investigate contrastive accent and compare the electrophysiological profiles associated with the word *or* when it remains non-accented to when it is accented.

In the context of our experiment, we predict that contrastive accent will be associated with a P300 (more precisely a P3a), a component which is generated automatically when one perceives an abnormal, rare or deviant stimulus (Ranganath & Rainer 2003). Indeed, function words and words placed in a median position in an utterance rarely carry a contrastive accent in French and *ou* (the French word for 'or'), being a function word that necessarily appears in a median position, is thus rarely accented. Moreover, given that the disjunction is the sole word to be accented at all in our experiment and in a fraction of the trials only, it further qualifies as rare in the context of this paradigm. Once *or* is accented, we also compare those participants whose responses indicate that they interpret the utterance logically to those who interpret it pragmatically. This presence or absence of the P3a will enable us to find out whether both groups of participants perceive accent in a similar fashion thus constituting an important control. We predict that the P3a should appear both for logical and pragmatic participants upon hearing a contrastive accent, unlike other electrophysiological markers that are specific to pragmatic inference making.

What are the potential markers of pragmatic inference making? We consider two well known components – the N400 and the P600 – and explain why we do not expect effects related to the former and do predict effects with respect to the latter. Although the N400 component has traditionally been associated with semantic anomalies (e.g., *He spread the warm bread with* <u>*socks*</u>, see Kutas & Hillyard 1984), recent developments have pointed out that semantic and pragmatic factors are often confounded in standard paradigms. Once semantic and pragmatic factors are disentangled, it is arguable, in many cases, that the N400 is elicited by pragmatic and not semantic anomalies (Filik & Leuthold 2008; Van Berkum et al. 1999; Van Berkum et al. 2003). For instance, when presented a story wherein Jane's brother has done something quickly, the word *slow* in *Jane told her brother that he was exceptionally slow* elicits a larger N400 than the (same sentence but with the) word *quick*. Note that there is no semantic

anomaly here but rather a pragmatic *incongruency* between the target word and the wider discourse context (Filik & Leuthold 2008; Van Berkum et al. 1999). Unlike in such cases, the scalar inference that we investigate here is generated when the context warrants a richer pragmatically-*justifiable* interpretation of the utterance; it is not the result of a strong mismatch between a word (e.g., *slow*) and its wider discourse (e.g., a story describing a quick person). Thus, this line of work does not appear to address the kind of pragmatic inference-making we aim to investigate here.

Another notable link between the N400 and pragmatic interpretation occurs with respect to metaphor (Bonnaud et al. 2002; Coulson & Van Petten 2002; 2007; Pynte et al. 1996). It has been reported that common French metaphors (e.g., *Those soldiers are lions*, Pynte et al. 1996) and novel metaphors (e.g., *Those apprentices are lions*) trigger a greater N400 than literal sentences. However, most experiments on the topic are confounded by the fact that they resort to novel metaphors which are intrinsically more anomalous and uncommon than literal controls. In paradigms controlling for this bias, it has been shown that N400s produced by metaphorical sentences are not distinguishable from those prompted by literal sentences (Lai et al. 2009; Tartter et al. 2002). Taken together, these data indicate that the N400 is unlikely to capture the process of pragmatic enrichment that is at stake in the production of scalar inferences.

Furthermore, initial indications from the literature confirm that N400s are not directly associated with scalar inferences. Although Noveck & Posada (2003) highlighted what has become a robust reaction time difference between logical and pragmatic answers (in reaction to sentences such as *Some elephants have trunks*), the authors did not identify an ERP profile linked to those who provided pragmatic interpretations (even while focusing specifically on the N400). It remains a possibility that Noveck and Posada's paradigm was not ideal for uncovering N400s;[5] nonetheless, the previous lack of N400 effects (with respect to pragmatic inferencing) leaves us sceptical that one will find such effects in response to the pragmatic enrichments of *or* in our experiment.

As we indicated above, we do anticipate effects related to the P600. Although the P600 has long been thought of as the component associated with syntactic anomalies, its amplitude being first described as a "function of the perceived syntactic well-formedness of the sentence" (for the first report, see Osterhout & Holcomb 1992; Osterhout et al. 1994), recent studies have now highlighted that it denotes a general mechanism of reanalysis of an utterance in order

to make sense of it, in relationship to its linguistic context (Friederici 2002). Indeed, it has been identified after semantic anomalies in the context of sentences containing no syntactic problems at all. For instance, Van Herten et al. (2005) found it in sentences such as: *De vos die op de stropers joeg sloop door het bos* / 'The fox that hunted the poachers stalked through the woods'.

This led the authors to argue that the P600 reflects a reprocessing of the utterance to check whether or not the initial processing was correct. According to a more general view, the P600 occurs in various situations including cases where a specific effort is made to understand an utterance or to integrate a word in a context (Burkhardt 2006).

This leads to our predictions. Given that current research favours Relevance Theory, showing that scalar inferences are not produced automatically, the production of a pragmatic inference (in this case, prompted by a prosodic cue) ought to elicit an ERP component related to processing effort. We thus predict that the P600 component will be enhanced among the pragmatic participants who take the prosodic cue into account in order to draw an inference and seemingly put in more effort to get the 'not both' reading of *or*. It also follows that the logical participants, who do not enrich the meaning of *or* based on the prosodic cue, will not exhibit such a profile. Such an outcome would not be predicted in Levinson's framework since it postulates that it is the *cancellation* of the scalar inference, and not its *production*, that requires extra processing. According to a default view, scalar inferences are generated automatically, regardless of the context, and can be subsequently cancelled following extra costs. This implies, of course, the opposite prediction, namely that logical (but not pragmatic) answers should be associated with a P600.

To sum up, we hypothesize that the prosodic cue will be associated with two effects. One, the P3a should appear as a generalised response to contrastive accent because it beckons the listener to notice it (regardless of the way the listener eventually responds). In contrast, the P600 will not be generalised to all the participants and will only appear in association with responses that are indicative of a pragmatic interpretation.

2. Experiment

To test our hypotheses, we employ a paradigm which we have been using in previous experiments and which confirmed the influ-

ence of prosody on the interpretation of *or* (Chevallier et al. 2008). In this paradigm, the participants reads a word, for example *TABLE*, and then hears a sentence describing two letters in that word, for example (3):

(3) There is a T or a U.

In Chevallier et al.'s experiment, the participants then had to decide whether or not the sentence correctly describes the word on the screen. In (3), the correct answer is 'true' because there is indeed a T in *TABLE*. The main experimental condition is the one wherein two letters – which are both indeed in the word – are mentioned and joined by the connective *or* (i.e., True-True condition, henceforth **Or TT**). This is the only condition that can distinguish between inclusive and exclusive readings (for further discussion, see Pelletier 1977). Indeed, this is the condition that typically prompts two kinds of responses. For example, consider *TABLE* again, but followed by (4). When participants accept (4) as true, the response is considered **logical** and when they reject it as false, it is considered **pragmatic**.

(4) There is an A or a B.

The Or TT condition can be presented with two different prosodic contours – either in a neutral intonation or else carrying a contrastive accent (OR). Prior results from Chevallier et al.'s (2008) between-subjects experiment indicate that, under neutral circumstances, participants make the logical (i.e., inclusive) interpretation of *or* 77% of the time in the Or TT condition. On the other hand, the contrastive accent significantly decreases rates of logical interpretations to 27% (and thus by definition increases the percentage of pragmatic interpretations, to 73%). Henceforth, **OR TT** will refer to the contrastive condition.

The effect was confirmed in a within-subject version of the experiment in which participants hear *or* in both – the neutral and contrastive – conditions (yielding logical answers at rates of 68% and 48%, respectively). Based on their reaction to the two conditions (Or TT and OR TT), the participants in the within-subject experiment can be classified into three categories. Overall, the modal response (which comprised roughly half the participants) is to provide the same logical response to the two conditions (whether contrastive accent is added or not). The second ranking modal response is to process the (Or TT and OR TT) utterances pragmatically. The third type of response is

to provide a logical answer when *or* is neutral (Or TT) and a pragmatic response when *or* is contrastive (OR TT). Note that no one did the opposite (i.e., provide a logical answer when *or* is accented and a pragmatic one when it is non-accented).

This sort of distribution makes the paradigm more than apt for ERP analyses with respect to logical and pragmatic treatments of accented *or*; that is, it ultimately allows us to compare a group that provides logical response when presented stimuli from the critical OR TT contrastive condition to a group that provides pragmatic responses in this same condition (i.e., we anticipate comparing those in the first group above to those in the two latter groups). We therefore adapt the within-subject version of the paradigm just described and apply these to an EEG paradigm (essentially by adding more stimuli). This paradigm puts us in a position to investigate the electrophysiological correlates of 1) the perception of contrastive accent and 2) of the cognitive processing of pragmatic inferences that result from such a prosodic cue.

2.1. Method

Participants. Thirty-two volunteer students (11 men and 21 women) were tested in this study. All were right handed and native speakers of French. Seven participants (two men and five women) were eliminated either because of poor performance rates (for two participants who gave less than 5% of correct answers in the TF and FT Or conditions, thus treating *or* as *and*) or because the recordings were confounded with artefacts (for five participants for whom we detected artefacts in at least one of the conditions in over 20 percent of trials). The remaining 25 participants were between 18 and 28 years old (M = 22.25; SE = 2.74). All the participants received a gift worth approximately 10 euros.

Stimuli. The 500 most frequent French words were selected from Lexique and Brulex databases (Content & Radeau 1988; New et al. 2001). Words including an accented letters (e.g., *é*, *è*, *ü*, *à*, etc.) and words including two identical letters were excluded from that list and the remaining 270 monomorphemic words were used to form our experimental list. The words were presented at the centre of a black screen in white letters using Capital letters, size 16 and Font Times New Roman. The sentences all had the same form: 'There is a *letter 1* and / or a *letter 2*' with a mean length of 2013 ms. The connective (mean length = 166 ms) appeared 1110 ms on average after the beginning of the sentence. They were read by a native French female

speaker and recorded in a quiet room (microphone: Sony ECM 719) using the software WaveStudio. The speech signal was digitised at a 22 kHz / 16 bits sampling rate (see Fig. 1B).

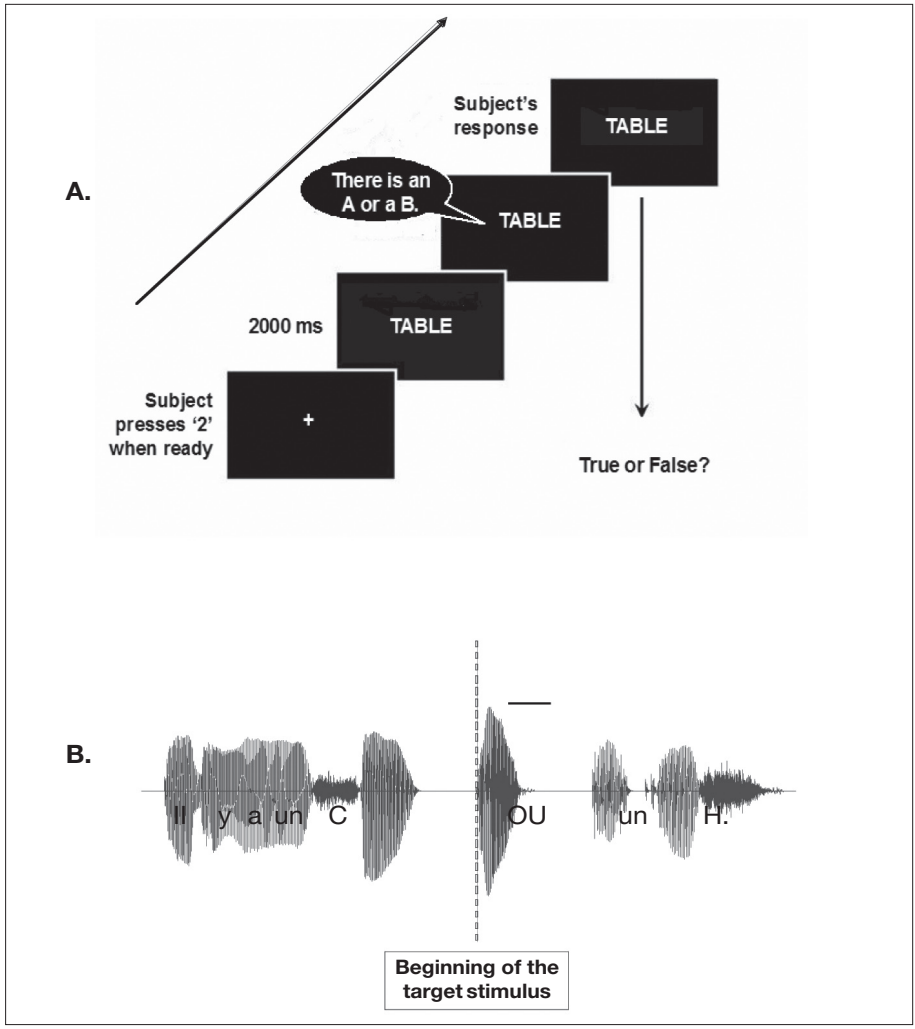

Figure 1. Experimental procedure and Spectrogram of one of the sentences. **A.** Timing of a sample trial. **B.** Waveform of one of the critical Or TT sentences.

The stimuli were then presented to two independent judges unaware of the goal of the study. For each item, they were asked to judge whether the connective carried an accent or not. Both correctly identi-

fied all the utterances containing contrastive accent, with an inter-ater reliability of 100%.

The participant was required to evaluate 270 sentences during the experimental phase. The words and their corresponding sentences were presented in a random order with the sole constraint being that there were 30 sentences in each of our nine conditions. There were eight conditions based on Connective (*and* vs. *or*) and the four lines of the truth-table: First letter *True*-Second letter *True* (**TT**), True-False (**TF**), False-True (**FT**) and False-False (**FF**). The ninth condition (Accented OR TT) is identical to the Or TT condition except that it stresses the *or*. One can see that the True-True cases for the disjunctive sentence (e.g., 'There is an A or a B' following *TABLE*) are the situations that are most relevant and that they are just two of nine possible conditions. The remaining seven conditions serve as controls and distractors. We chose to put an accent uniquely where the prosodic cue can affect the truth value by turning a potentially true statement into a false one (the TT condition). As far as the other conditions are concerned, a prosodic cue cannot affect the participant's applied truth value in a detectable way. In the FT, TF and FF conditions, a contrastive accent on *or* could lead the participant to enrich its semantics towards an exclusive interpretation but this would amount to redundant (and potentially confusing) information since the enrichment would have no impact on the response pattern. For instance, 'There is an A OR a B' after the word *CHILD* would lead the participant to answer 'false' regardless of whether she had an inclusive or an exclusive reading of the word. In the FT, TF and FF conditions, a contrastive accent on *or* could lead the participant to enrich the utterance towards an exclusive interpretation but we would have no way of confirming that since the enrichment would not affect the response pattern.

Procedure. The experiment was conducted in our laboratory and lasted approximately 90 minutes. The participant sat comfortably in a chair 90 cm away from the screen in a dimly lit room. The stimuli were presented in a random order using E-Prime (Schneider et al. 2002). The sound files were synchronised with the EEG recording. For each trial, the participant would see a fixation cross at the centre of the screen which indicated that he was allowed to blink and move (see Fig. 1A). When ready, the participant pressed the middle ('2') key of a three-key button-box which made the word appear. The word remained on the screen for 2000 ms before the participant would hear the sentence from the speakers placed on each side of the computer screen (150 cm away from the participant). At the end of the sentence, a 100 ms silent delay was inserted. The participant could then give his answer using his right index to press '1' or '3' (associated with the

'true' response for half the participants). The experiment was divided into two phases: a training phase including 6 trials and an experimental phase including 270 trials divided into three blocks of 90 trials.

EEG recording and pre-processing of the recordings. A 64-electrode Geodesic Sensor Net was mounted on the participant's head. We used a 500 Hz sampling frequency and a high impedance amplifier (200 MΩ, Net Amps TM, Electrical Geodesics INC.). The signal was filtered to keep only the frequencies between 0.1 and 200 Hz and impedances were kept below 5 KΩ). Before analysing the data, the raw recordings were segmented using Netstation™: Each segment consisted in a 200 ms baseline before the onset of the connective and an 800 ms segment after the onset of the connective. All further ERP analyses were conducted using ELAN-Pack software developed at INSERM U821 (Lyon, France). The signal was re-filtered with a 25 Hz low-pass filter, the baseline was corrected from the 200 ms pre-stimulus interval and the artefacts were then removed to get a cleaner signal (contaminated trials containing ocular artefacts or movement artefact amounted to 5.9% of the total number of trials). All electrodes were referenced to Cz online and rereferenced off-line using average reference.[6] The P3a component was defined as the positive wave within a latency window between 200 ms and 400 ms; the P600-like component was defined as the positive wave occurring within the 550-750 ms range. The scalp topographies correspond to difference waveforms which were obtained by subtraction.

3. Results

3.1. Behavioural results

Reaction times of more than three standard deviations from the mean were considered outliers (0.53% of the trials) and were removed from both choice proportions and reaction time data. In the control conditions, the analysis of accuracy rates revealed rates of correct answers that were as high as what we have observed in our previous behavioural studies, ranging from 92.8% to 100%. This indicates that the task was well understood by the participants and that the EEG experimental setting did not influence the participants' accuracy (for more details see Tab. 1A).

As far as the critical conditions are concerned, rates of responses indicating an inclusive (logical) interpretation are higher in the Or TT (neutral) condition (77.2%) than in the OR TT (accented)

Table 1A. Mean accuracy for all the participants.

		Sentence	Example words	Logically correct answer	Percentage correct (standard deviation)
And	TT	There is an A and a B.	TABLE	True	97.9 (0.9)
	TF	There is an A and a B.	PATIN	False	96.8 (2.2)
	FT	There is an A and a B.	BUSTE	False	99.7 (0.9)
	FF	There is an A and a B.	CHIEN	False	100.0 (0.0)
Or	TTacc	There is an A OR a B.	TABLE	True*	67.2 (43.4)
	TT	There is an A or a B.	TABLE	True*	77.2 (39.4)
	TF	There is an A or a B.	PATIN	True	93.5 (4.8)
	FT	There is an A or a B.	BUSTE	True	92.8 (7.0)
	FF	There is an A or a B.	CHIEN	False	96.0 (1.7)

* A 'true' response is logically correct because the sentence contains a true disjunct (note, though, that a 'false' response is pragmatically warranted because *or* can be taken to imply 'not both'). By convention, logical interpretations of *or* (i.e., inclusive interpretations) are coded as 'correct', and pragmatic answers (i.e., exclusive interpretations) are coded as 'incorrect'.

condition (67.2%) ($p < .05$). In other words, the accent placed on *or* gives rise to more exclusive interpretations. Also, as was found in Chevallier et al. (2008) and summarised above, most participants adopt a particular response pattern when faced with the two Or TT conditions and maintain it throughout the experiment. Given our focus on responses to the OR TT (accented) condition, participants were classified into two groups. One group is made up of those who consistently provide inclusive interpretations (at an 80% level or above) and are classified as logical; of 25 participants, 16 participants fall into this category. The other group is made up of those who consistently provide exclusive interpretations (at an 80% level or above) and are classified as pragmatic; nine participants fall into this category. Among these nine participants, six respond exclusively generally (to Or TT and OR TT) and three respond exclusively only when the *or* in the True-True condition is accented. It is important to note that rates of accuracy between the logical and the pragmatic group were otherwise consistent (i.e., among the numerous control conditions). The only difference is performance across the two critical Or TT conditions (see Tab. 1B).

Table 1B. Mean accuracy and reaction times in each group (logical vs. pragmatic).

		Percentage correct (standard deviation)		Reaction Times (ms) (standard deviation)	
		Logical	Pragmatic	Logical	Pragmatic
And	TT	97.9 (2.9)	97.8 (2.2)	482 (220)	558 (256)
	TF	96.2 (2.3)	97.4 (2.2)	550 (243)	607 (288)
	FT	99.8 (0.8)	99.6 (1.1)	524 (191)	626 (174)
	FF	100.0 (0.0)	100.0 (0.0)	574 (281)	640 (239)
Or	TTacc	98.1 (2.9)	12.2 (17.9)	479 (166)	808 (403)
	TT	98.9 (3.4)	38.5 (44.7)	479 (175)	721 (268)
	TF	98.3 (2.4)	93.7 (6.8)	546 (219)	786 (318)
	FT	95.4 (6.3)	93.3 (8.5)	496 (213)	608 (233)
	FF	96.2 (1.1)	95.5 (2.4)	584 (292)	701 (365)

Let us now analyse accuracy rates and reaction times in more detail. For the purpose of the following analyses, we only include the Or TT contrastive condition, i.e., the condition which allows us to analyse two clearly distinct groups of participants (logical or pragmatic) and which will serve as the basis for subsequent ERP analyses. A repeated measures ANOVA concentrating on the contrastive condition with the within-subject factors Truth condition (FF, FT, TF, TT), and Connective (And vs. Or) and the between subject factor Group (Logical vs. Pragmatic) revealed a main effect of Group (Logical, Pragmatic), $F(1,23) = 65.3$; $p < .0001$, of Connective (Or, And), $F(1,23) = 93.3$; $p < .0001$, and of Truth condition (FF, FT, TF, TT) $F(3,21) = 128.3$; $p < .0001$, each interacting with one another, all $Fs > 53.3$; all $ps < .0001$. *Post-hoc* Least Significant Difference (**LSD**) tests revealed that all these effects were due to the OR TT condition in the pragmatic group, in all comparisons including this condition: all $ps < .01$; all other comparisons do not reach significance.

We then carried out a repeated measures ANOVA on participants' reaction times (see Tab. 1B), including the same factors as the ones used in the analysis of accuracy rates (focusing on the OR TT contrastive condition, which is the basis for classifying the two groups). A log transformation was carried out before analysis to improve the conformity of the data to the standard assumptions of ANOVA (see, e.g., Howell 1997). Outliers were removed and only correct responses were kept in the analysis (thereby removing a further 10.2% of the

responses). The ANOVA revealed no main effect of the Group (Logical, Pragmatic), $F(1,23) = 2.22$; p = n.s., no main effect of Connective (Or, And), $F(1,23) = 3.51$; p = n.s., and a main effect of Truth condition (FF, FT, TF, TT), $F(3,21) = 8.98$; $p < .001$. The only significant interaction was one between the Group and Connective, $F = 9.44$; $p < .005$. *Post-hoc* LSD tests revealed that the OR TT constrastive condition yielded slower responses in the pragmatic group than in the logical group ($p < .004$). We found one other significant difference between the two groups in the Or TF condition, $p < .008$, possibly because the pragmatic participants were invested in waiting for the second letter to be uttered so as to make their decision. All other comparisons between the two groups did not reach significance. Overall, our reaction time data provide further evidence in favour of the claim that pragmatic processing is time consuming and takes some amount of effort since pragmatic participants were slower than logical participants in the OR TT contrastive condition. It is important to note the absence of a main effect of Group and the absence of differences, through the post-hoc tests, with respect to the And condition, all *ps* = n.s., and the FF and FT Or conditions, both *ps* = n.s. Taken together, these data indicate that pragmatic participants are not simply slower overall.

Since the TF, FT and FF conditions (for *and* and *or*) and the TT condition (for *and* only) served as distractors for the participants and as a means for us to detect poor overall performances, we now turn to the Or TT conditions and the ERP analyses. The contrastive vs. neutral conditions and the distribution it produces allows for two comparisons with respect to ERP profiles. First, we can determine whether or not contrastive accent has an effect on the ERP profile; we thus compare the contrastive (OR TT) and the neutral (Or TT) conditions across all participants. Second, we can determine how an ERP profile is affected by a participant's response; we thus compare logical and pragmatic responses to stimuli in the critical OR TT (contrastive) condition only. In the following sections, we present the results from these comparisons for two time windows (200 – 400 ms and 550 – 750 ms). These analyses then lead us to focus on a wider time window (100 – 800 ms) which reveals the presence of another interesting component (the Left Anterior Negativity).

3.2. ERP analyses

200 – 400 ms time window. Visual inspection of the grand-average waveforms (Fig. 2A, Cz and Fz) and their topographies reveals the presence of a greater positivity in the contrastive condition com-

pared to the neutral condition (Fig. 2A, scalp maps). This positivity is maximal over frontal and central areas of the scalp and peaks at around 300 ms after the presentation of *or*, thus presenting distinctive features of the P3a. In this same time range, a posterior negativity appears and could correspond to the negative pole of the P3a.

As far as the comparison between logical and pragmatic groups is concerned, both the scalps and the waveforms indicate the presence of a slightly greater positivity for pragmatic answers peaking at around 300 ms after the presentation of *or* (Fig. 2B). This positivity is visible over the central area (Cz) but does not appear on Fz.

On the basis of this visual inspection, we chose two relevant electrode sites for our statistical analysis: one frontal site (Fz) and one central site (Cz). We conducted an ANOVA using repeated measures with two factors: the between subject factor Group (Logical vs. Pragmatic) and the within-subject factor Accent (Contrastive vs. Neutral). For this analysis and further *post hoc* tests the dependent measure was the average voltage amplitude at the surface of the scalp in the 200-400 ms latency range. The ANOVA on Cz revealed no main effect of Group, $F(1,23) = 3.81$; p = n.s.), no main effect of Accent (Contrastive vs. Neutral), $F(1,23) = 2.31$; p = n.s., and no Group X Accent interaction, $F(1,23) = 0.19$; p = n.s. However, the ANOVA on Fz revealed a significant effect of Accent, $F(1,23) = 26.01$; $p < .0001$, no main effect of Group, $F(1,23) = 0.57$; p = n.s. and no Group X Accent interaction, $F(1,23) = 2.32$; p = n.s. The difference in the mean amplitude of this P3a component on Fz indicates that, following our predictions, the contrastive condition is associated with a more positive waveform (Mean amplitude at Fz = 1.15 μV) than the neutral condition (Mean amplitude at Fz = -0.57 μV).

This analysis has two implications: first, the detection of a contrastive accent on *or* is followed by a greater P3a component on Fz; second, this positivity is not different across the pragmatic and logical groups (no Group X Accent interaction).[7] This positivity is thus not specific to one group of participants and appears regardless of the answer that is ultimately provided. This is evidence that all the participants perceived the contrastive accent (and that they did so in a similar fashion). Overall, the graphs, the scalp map and the statistical analyses highlight the presence of a P3a component, most evident on Fz and appearing when *or* is accented, regardless of the group.

550 – 750 ms time window. Visual inspection of the grand-average waveforms (Fig. 2A, Cz) and their topographies (Fig. 2A, scalp maps) reveals a slightly greater positivity in the contrastive condition compared to the neutral condition and could indicate the presence of

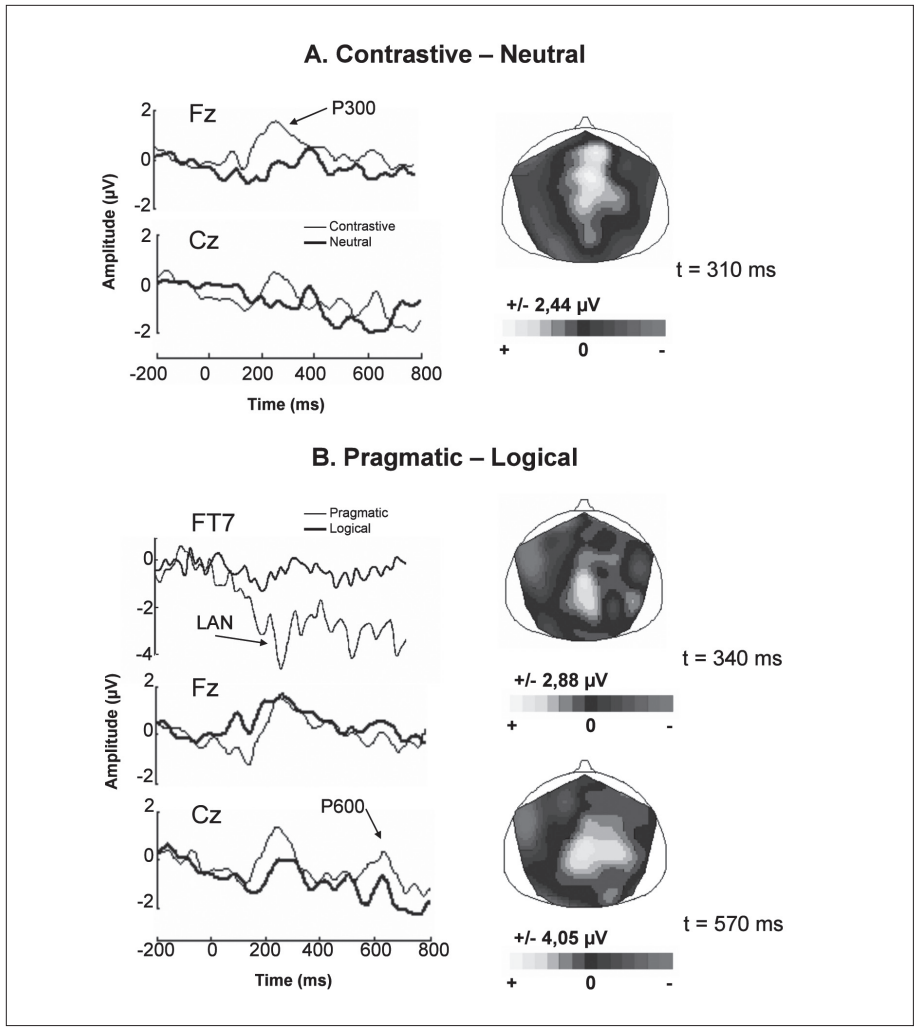

Figure 2. Effects of accent and pragmatic enrichment in the Or TT condition. **A.** On the left, stimulus locked grand-average waveforms evoked by the appearance of the target item in the contrastive (light gray line) and in the neutral condition (dark gray line) at Fz and Cz. On the right, topographic distribution of the difference waves created by subtracting the mean voltage of ERPs elicited by contrastive items (in the OR TT condition) from that elicited by neutral items (in the Or TT condition). **B.** On the left, stimulus locked grand-average waveforms evoked by the appearance of the target item for the pragmatic (light gray line) and for the logical participants (dark gray line) at FT7, Fz and Cz. On the right, topographic distribution of the difference waves created by subtracting the mean voltage of ERPs elicited by pragmatic participants from that elicited by logical participants in the OR TT contrastive condition.

a P600-like component. Furthermore, the scalp maps and the wave forms in the contrastive condition both indicate the presence of a greater positivity for the pragmatic answers when compared to the logical answers. In both types of comparisons (i.e., Contrastive vs. Neutral and Logical vs. Pragmatic), this positivity is maximal over the central area (Cz) and peaks at around 600 ms after the presentation of *or*, thus resembling an enhanced P600-like component. In the frontal area (Fig. 2A, Fz), the curves are very similar overall.

On the basis of this visual inspection, we chose to focus our statistical analysis on Cz. We conducted an ANOVA using repeated measures with two factors: the within-subject factor Accent (Contrastive vs. Neutral), and the between subject factor Group (Logical vs. Pragmatic). For this analysis and further *post hoc* tests the dependent measure was the average voltage amplitude at the surface of the scalp in the 550-750 ms latency range. The ANOVA revealed a significant effect of Accent on Cz, $F(1,23) = 8.21$; $p < .001$, with a Group X Accent interaction, $F(1,23) = 11.11$; $p < .05$. The Group X Accent interaction indicates that contrastive accent has a different influence on the pragmatic and logical response types. The comparison of the mean amplitudes across groups in the contrastive condition confirms that the effect goes in the direction predicted by our hypothesis with a higher P600-like component for pragmatic answers (Mean amplitude at Cz = 1.08 μV) than for logical answers (Mean amplitude at Cz = -1.05 μV). Furthermore, LSD *post hoc* tests indicate that the main effect of Accent is driven mainly by the pragmatic group. When the logical participants (N = 16) are considered separately, the waveform in the contrastive condition turns out to be comparable to the one in the neutral condition, p = n.s. This implies that the mere presence of a contrastive accent is not enough to trigger a P600. However, when the remaining participants are analysed, the contrastive condition is associated with a greater positivity when compared to the neutral condition, $p < .01$. Overall, the enhanced positivity evidenced in the 550 – 750 ms time window presents the classical topography of the P600 (i.e., maximal over Cz) and is arguably linked with pragmatic inferencing generally or with pragmatic inferencing linked with the presence of a contextual cue (i.e., the contrastive accent).

100 – 800 ms time window. Visual inspection of the grand-average waveforms (Fig. 2B, FT7) and their topographies (Fig. 2B, scalp maps) reveals the presence of an early and long lasting negativity over the left frontal areas of the scalp (especially, F7, FT7 and T7) for pragmatic answers when compared to logical answers. This negativity, both in terms of latency and topography, closely resembles the profile of the Left Anterior Negativity (**LAN**).

On the basis of this visual inspection, we chose to focus our analysis on FT7 (a standard site for the LAN). We conducted an ANOVA using repeated measures with two factors: the within-subject factor Accent (Contrastive vs. Neutral), and the between subject factor Group (Logical vs. Pragmatic). For this analysis and further *post hoc* tests the dependent measure was the average voltage amplitude at the surface of the scalp in the 100-800 ms latency range. The ANOVA on FT7 revealed a main effect of Group, $F(1,23) = 6.71$; $p < .01$, a main effect of Accent, $F(1,23) = 5.55$; $p < .05$, and a Group X Accent interaction, $F(1,23) = 6.29$; $p < .05$.

As with the P600-like component, LSD *post hoc* tests indicate that the main effect of the condition is driven mainly by the inclusion of the pragmatic group. Indeed, logical participants have the same profile in the contrastive and in the neutral conditions, p = n.s., conversely, a strong LAN is revealed when the remaining participants are analysed, $p < .01$. Again, the mere presence of a contrastive accent is not enough to elicit a LAN, p = n.s. The comparison of the mean amplitudes in the contrastive condition confirmed that the pragmatic answers are associated with a more negative profile (Mean amplitude at FT7 = -1.60 μV) than the logical answers (Mean amplitude at FT7 = -0.40 μV). It appears then that for pragmatic participants, the making of a pragmatic inference or the presence of a contextual cue is linked with a greater negativity. Overall, these data resemble the pattern described for the P600-like component.

4. General discussion

This experiment aimed to explore the ERP correlates of the pragmatic interpretation of the word *or* ('A or B but not both') when it is prompted by a prosodic cue. We predicted that the presence of a contrastive accent would induce a P3a component, i.e., a specific component that is elicited following unexpected stimuli. Furthermore, we predicted that pragmatic answers in the contrastive condition would be associated with a different waveform profile than logical ones. More precisely, we hypothesised that we would observe a P600-like component in association with pragmatic (but not logical) answers because the P600 is associated with increased processing effort in language understanding and it has been established in the cognitive literature that scalar inference-making is effort-demanding. Our exploratory data go in the direction of both these hypotheses. First, we found a fronto-central P3a component in the condition when the *or*

was accented but not when it was non-accented. This first component is evidence that both logical and pragmatic participants detected the accent on *or*. Second, we also demonstrated that when considering answers associated with a prosodic cue, pragmatic interpretations – aside from being associated with longer response times – elicited a larger central P600-like component and LAN compared to the logical answers. This suggests that the pragmatic processing of *or* relies on different mechanisms than what we call logical processing.

The presence of the P3a can be easily explained by two facts. First, accented items were quite rare, compared to non-accented items. Among 270 sentences, only 30 carried a contrastive accent. Second, contrastive accent is characterised by a number of specific acoustic properties, among which is the presence of a pause before the accented item as well as increased pitch and intensity; all of these features render the accented item physically different from its non-accented counterpart. Those two conditions (rarity and physical 'deviance') are favourable to the elicitation of a P3a (Osterhout & Hagoort 1999). Indeed, the amplitude of the P3a depends on the probability that the item appears and on its salience; the most salient and rarest stimuli being those associated with the amplest P3a. The presence of this precocious component reflects the activity of a general network which enables participants to rapidly orient their attention towards a new and unexpected stimulus (Ranganath & Rainer 2003). Disentangling the two factors (rarity vs. acoustic properties) would require additional tests where the probability to hear accented and non-accented utterances would be equal. In the present paradigm, we chose to put accent only in the relevant condition (OR TT) so that the accent would be construed as carrying specific information. This design allowed us to easily elicit the P3a, which then served as a control ensuring that differences between logical and pragmatic participants observed in later time windows did not merely reflect perceptual differences between the two groups.

The presence of a P600-like component in the context of our study is interesting with respect to the ongoing debate about the functional significance of this marker. At first, the P600 was thought of as a component marking the processing of all sorts of syntactic violations: phrase and argument structure (Friederici & Meyer 2004), morphosyntax (Friederici & Meyer 2004), garden-path sentences (Frisch et al. 2003; Osterhout et al. 1994). In brief, the P600 has been considered a good measure of syntactic anomaly and parsing difficulty (Friederici et al. 1998; Hagoort et al. 1993; Osterhout & Holcomb 1992). Similarly, the LAN has traditionally been associated with morphosyntactic violations (e.g., number-person agreement,

Friederici et al. 1993). For instance, Morris & Holcomb (2005) demonstrated that incorrect verb forms, when compared to their correct counterpart, elicit a greater LAN. In Friederici's model (2002), the LAN is thus described as a component which reflects the integration of semantic and morphosyntactic information. However, these views have been challenged towards a more general conception of the P600 and of the LAN. As far as the P600 is concerned, it seems that it is not a "monolithic component but rather seems to comprise a mixture of subcomponents, possibly reflecting very different processes such as diagnosis, syntactic and prosodic reanalysis, and integration" (Friederici et al. 2001: 321). For instance, as indicated in the Introduction to this paper, the P600 has recently been identified following semantic anomalies (Kolk et al. 2003; Van Herten et al. 2005) which could be taken to indicate that the P600 is caused by the gap between the expected word and the actual word participants see (Van Herten et al. 2005). In line with this idea, manipulations at the pragmatic level (e.g., in metaphor, jokes, and irony) have previously been observed to modulate the amplitude of this component (Coulson & Kutas 2001; Coulson & Van Petten 2002; Regel et al. 2009). Other authors have proposed a slightly different interpretation according to which the P600 would reflect the cost of reprocessing (Osterhout et al. 1994). Concomitantly, it has been argued that the P600 reflects the amount of necessary resources to integrate words in the context of the utterance (Kaan et al. 2000; Patel et al. 1998). Similarly, the LAN has been found in experimental paradigms involving syntactically correct items that require extra processing (e.g., the long-distance syntactic dependencies described by Kluender & Kutas 1993). This last view is especially interesting in the context of recent developments in experimental pragmatics where the manipulation of the amount of cognitive resources available to the listener to process an utterance has yielded interesting results that shed light on an ongoing theoretical debate.

In this debate, two modern theories of pragmatics have traditionally proposed contrasting views concerning the way logical terms (e.g., *or*, *and*, *some*, *if*, etc.) are processed (for a review, see Noveck 2004). According to Levinson (2000), these terms are enriched automatically, or by default. Thus, *or* would be automatically understood as 'one or the other but not both', *some* as 'some but not all', *if* as 'if and only if', and so on. On the other hand, Relevance Theory (Sperber & Wilson 1986/1995) argues that utterances are only enriched when context warrants it. Furthermore, when enrichment does occur, it involves a certain processing cost. Our results are in line with the latter view and are consistent with numerous studies showing that scalar inferences

are effortful enrichments (see Introduction). Not only are pragmatic responses associated with longer reaction times when compared to logical responses, pragmatic responses in the wake of a prosodic cue prompt a P600. Together, these findings indicate that pragmatic enrichment requires effort and results from a deeper analysis of the sentence. In this framework, our experiment can be analysed along the following lines: upon detecting a prosodic cue, the pragmatic participants enrich, or process more deeply, the disjunctive utterance into a more informative one which allows them to make better sense of it. This deeper analysis consists in considering alternative interpretations along a path of least effort, in the order of their accessibility, until an optimally relevant interpretation is found (Sperber & Wilson 1986/1995). This could explain why not all participants provide the same answer since, effort and effects being relative to each individual, the optimally relevant interpretation may differ according to the individual. If this analysis is correct, the P600 could be construed as reflecting the amount of resources necessary to integrate words in the context of the utterance, not only syntactically (Kaan et al. 2000), but also semantically (Kolk et al. 2003; Van Herten et al. 2005) and pragmatically.

Although these exploratory results are informative, more work remains and there are obvious limitations to the present study. First, the effects we observed here were obtained on a very small sample and can only be taken as pilot data. A bigger sample would have allowed us to isolate profiles of those participants who consistently make the pragmatic inference and to compare them to participants who make the inference when it is prompted by a prosodic cue only. Second, *or* was accented in only one condition (TT) and it would be interesting to run a new study where *or* is accented in all the *or* conditions for half the trials. This too, might make it possible to dissociate the acoustic manipulation from its inferential consequences. In any case, future work can benefit from the present pilot study which demonstrates how contrastive accent can influence, not only truth-evaluations but, the ERP profiles that are concomitant with them.

Addresses of the Authors

Coralie Chevallier (Corresponding Author), Mathilde Bonnefond, Jean-Baptiste Van der Henst, Ira A. Noveck: Laboratoire sur le Langage, le Cerveau et la Cognition (L2C2), Université de Lyon – CNRS, 67 Boulevard Pinel, 69765 Bron, France
coralie.chevallier@gmail.com

Notes

[1] Each time *or* is written in capital letters and underlined, it indicates that it carries contrastive accent.

[2] A stimulus X is considered to be more informative than a stimulus Y if X entails Y, but not *vice versa*. Given that *A and B* entails 'A or B', *and* is more informative than *or*. *Or* is thus considered to be at the bottom of the <or, and> scale. Scalar inferences do not only occur in relation to utterances containing the word *or*, as attested in examples (a)-(d):

(a) Sometimes, I like cooking with olive oil. (implies 'not always')
(b) This restaurant is good. (implies 'not excellent')
(c) I think Ira likes ERPs. (implies 'I don't know for sure')
(d) She has three children. (implies 'not four')

[3] *Generalized* Conversional Implicatures (as opposed to *Particularized* Conversional Implicatures) are thought to occur automatically, regardless of the context. For example, *John walked into a house yesterday and saw a tortoise* is always taken to imply that John walked into a house that was not his, regardless of the context.

[4] Contrastive accent and word stress differ in English. For instance, when a speaker utters *The uniVERsity MUST be rebuilt*, the word stress on university is obligatory whereas the contrastive accent on *must* is optional. In other words, word stress is obligatory and is encoded in the phonological representation of each word; contrastive accent is optional and allows the speaker to express a certain attitude towards the proposition she is expressing (Rossi 1999).

[5] There are two factors that could have masked N400 effects in the Noveck & Posada study. First, all test sentences employed the term under study, *some*, when it could have been useful to create a greater variety of distractors so as to obscure the goal of the task (e.g., by including sentences that begin with *all*). Moreover, the analysis concentrated on the last word of sentences like *Some elephants have trunks* and did not check for potential effects on the word *some* itself.

[6] Given that most language ERP studies rereference the signal to the mean of the mastoids, one could argue that the choice of a different reference makes it difficult to compare our data with some relevant literature. Our choice of the average reference was motivated by the fact that Bertrand et al. have long argued in favour of the use of the average reference, both on theoretical and on experimental grounds (Bertrand et al. 1985). Moreover, we ran our statistical analysis using the mean mastoids and noticed no major difference in the results.

[7] This last analysis is, however, slightly problematic since it includes participants who responded pragmatically in both the contrastive and the neutral condition and participants who responded pragmatically only in the contrastive condition. When the three participants belonging to the latter group are removed from the analysis, there is still no difference between the contrastive and the neutral condition (p = n.s.) but obviously the numbers are much too small to draw firm conclusions.

Bibliographical References

Bertrand Olivier, François Perrin & Jacques Pernier 1985. A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalography & Clinical Neurophysiology* 62. 462-464.

Besson Mireille, Cyrille Magne & Daniele Schön 2002. Emotional prosody: Sex differences in sensitivity to speech melody. *Trends in Cognitive Sciences* 6. 405-407.

Böcker Koen B.E., Marcel C.M. Bastiaansen, Jean Vroomen, Cornelis H.M. Brunia & Beatrice De Gelder 1999. An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology* 36. 706-720.

Bonnaud Véronique, Roger Gil & Pierre Ingrand 2002. Metaphorical and non-metaphorical links: a behavioral and ERP study in young and elderly adults. *Neurophysiologie Clinique/Clinical Neurophysiology* 32. 258-268.

Bott Lewis & Ira A. Noveck 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* 51. 437-457.

Breheny Richard, Napoleon Katsos & John Willliams 2006. Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100. 434-463.

Brosch Tobias, Didier Grandjean, David Sander & Klaus Scherer 2009. Cross-modal emotional attention: Emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience* 21. 1670-1679.

Burkhardt Petra 2006. Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. *Brain and Language* 98. 159-168.

Carston Robyn 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell Publishing.

Chevallier Coralie, Ira A. Noveck, Tatjana Nazir, Lewis Bott, Valentina Lanzetti & Dan Sperber 2008. Making disjunctions exclusive. *The Quarterly Journal of Experimental Psychology* 61. 1741-1760.

Chevallier Coralie, Deirdre Wilson, Francesca Happé & Ira A. Noveck *in press*. Scalar inferences in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders.*

Chierchia Gennaro 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In Belletti Adriana (ed.). *Structures and Beyond*. Oxford: Oxford University Press. 39-103.

Content Anne & Marie Radeau 1988. Données statistiques sur la structure orthographique du français. *Cahiers de Psychologie Cognitive* Numéro hors séries.

Copi Irving M. 1941. *Introduction to Logic*. New-York: Macmillan.

Coulson Seana & Marta Kutas 2001. Getting it: human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters* 316. 71-74.

Coulson Seana & Cyma Van Petten 2002. Conceptual integration and metaphor: An event-related potential study. *Memory and Cognition* 30. 958-968.

Coulson Seana & Cyma Van Petten 2007. A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Research* 1146. 128-145.

De Neys Wim & Walter Schaeken 2007. When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology* 54. 128-133.

Eckstein Korinna & Angela D. Friederici 2005. Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cognitive Brain Research* 25. 130-143.

Eckstein Korinna & Angela D. Friederici 2006. It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience* 18. 1696-1711.

Filik Ruth & Hartmut Leuthold 2008. Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology* 45. 554-558.

Friederici Angela D. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6. 78-84.

Friederici Angela D., Axel Mecklinger, Kevin M. Spencer, Karsten Steinhauer & Emanuel Donchin 2001. Syntactic parsing preferences and their on-line revisions: a spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research* 11. 305-323.

Friederici Angela D. & Martin Meyer 2004. The brain knows the difference: two types of grammatical violations. *Brain Research* 1000. 72-77.

Friederici Angela D., Erdmut Pfeifer & Anja Hahne 1993. Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research* 1. 183-192.

Friederici Angela D., Karsten Steinhauer, Axel Mecklinger & Martin Meyer 1998. Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology* 47. 193-221.

Friedrich Claudia, Kai Alter & Sonja Kotz 2001. An electrophysiological response to different pitch contours in words. *Neuroreport* 12. 31-89.

Friedrich Claudia, Sonja Kotz, Angela D. Friederici & Thomas C. Gunter 2004. ERPs reflect lexical identification in word fragment priming. *Journal of Cognitive Neuroscience* 16. 541-552.

Frisch Stefan, Sonja Kotz, D. Yves Von Cramon & Angela D. Friederici 2003. Why the P600 is not just a P300: The role of the basal ganglia. *Clinical Neurophysiology* 114. 336-340.

Grice H. Paul 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Gualmini Andrea, Stephen Crain, Luisa Meroni, Gennaro Chierchia & Maria Teresa Guasti 2001. At the semantics/pragmatics interface in child language. *Proceedings of SALT XI* (2001). Ithaca, NY: Cornell University. 231-247.

Hagoort Peter, Colin Brown & Jolanda Groothusen 1993. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes* 8. 439-483.

Horn Laurence 1972. *On the semantic properties of logical operators in English*. Los Angeles: UCLA. Ph.D. dissertation.

House Jill 2006. Constructing a context with intonation. *Journal of Pragmatics* 38. 1542-1558.

HOWELL David 1997. *Statistical Methods for Psychology* (4th edition). Belmont, CA: Duxbury.

KAAN Edith, Anthony HARRIS, Edward GIBSON & Phillip HOLCOMB 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15. 159-201.

KLUENDER Robert & Marta KUTAS 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5. 196-214.

KOLK Herman, Dorothee CHWILLA, Marieke VAN HERTEN & Patrick OOR 2003. Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language* 85. 1-36.

KUTAS Marta & Steven A. HILLYARD 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307. 161-163.

LAI Vicky, Tim CURRAN & Lise MENN 2009. Comprehending conventional and novel metaphors: An ERP study. *Brain Research* 1284. 145-155.

LEVINSON Steven C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

MORRIS Joana & Phillip HOLCOMB 2005. Event-related potentials to violations of inflectional verb morphology in English. *Cognitive Brain Research* 25. 963-981.

NEW Boris, Christophe PALLIER, Ludovic FERRAND & Rafael MATOS 2001. Une base de données lexicales du Français contemporain sur Internet: LEXIQUE. *L'Année Psychologique*. 447-462.

NOVECK Ira A. 2001. When children are more logical than adults: experimental investigations of scalar implicature. *Cognition* 78. 165-188.

NOVECK Ira A. 2004. Pragmatic inferences linked to logical terms. In NOVECK Ira A. & Dan SPERBER (eds.), *Experimental Pragmatics*. Basingstoke: Palgrave Macmillan. 301-321.

NOVECK Ira A., Vinod GOEL & Kathleen SMITH 2004. The neural basis of conditional reasoning with arbitrary content. *Cortex* 40. 613-622.

NOVECK Ira A. & Andres POSADA 2003. Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language* 85. 203-210.

OSTERHOUT Lee & Peter HAGOORT 1999. A superficial resemblance doesn't necessarily mean you're part of the family: Counterarguments to Coulson, King, and Kutas (1998) in the P600/SPS debate. *Language and Cognitive Processes* 14. 1-14.

OSTERHOUT Lee & Phillip HOLCOMB 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31. 785-806.

OSTERHOUT Lee, Phillip HOLCOMB & David SWINNEY 1994. Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20. 786-803.

PAPAFRAGOU Anna & Julien MUSOLINO 2003. Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition* 86. 253-282.

PATEL Aniruddh, Edward GIBSON, Jennifer RATNER, Mireille BESSON & Phillip J. HOLCOMB 1998. Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience* 10. 717-733.

PELLETIER Francis J. 1977. Or. *Theoretical Linguistics* 4. 61-74.
PIJNACKER Judith, Peter HAGOORT, Jan BUITELAAR, Jan-Pieter TEUNISSE & Bart GEURTS 2009. Pragmatic inferences in High-Functioning adults with autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders* 39. 607-618.
POUSCOULOUS Nausicaa, Ira A. NOVECK, Guy POLITZER & Anne BASTIDE 2007. Processing costs and their impact on the development of scalar implicature. *Language Acquisition* 14. 347-375.
PYNTE Joseph, Mireille BESSON, Fabrice-Henri ROBICHON & Jézabel POLI 1996. The time-course of metaphor comprehension: An event-related potential study. *Brain and Language* 55. 293-316.
RANGANATH Charan & Gregor RAINER 2003. Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience* 4. 193-202.
REGEL Stefanie, Seana COULSON & Thomas C. GUNTER 2010. The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Research* 1311. 121-135.
ROSSI Mario 1999. *L'Intonation, le système du français: description et modélisation*. Paris: Ophrys.
SPERBER Dan & Deirdre WILSON 1986/1995. *Relevance: Communication and Cognition*. Oxford: Blackwell Publishing.
SCHNEIDER Walter, Amy ESCHMAN & Anthony ZUCCOLOTTO 2002. *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools Inc.
STEINHAUER Karsten, Kai ALTER & Angela D. FRIEDERICI 1999. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience* 2. 191-196.
TARTTER Vivien, Hilary GOMES, Boris DUBROVSKY, Sophie MOLHOLM & Rosemarie VALA STEWART 2002. Novel metaphors appear anomalous at least momentarily: Evidence from N400. *Brain and Language* 80. 488-509.
VAN BERKUM Jos J. A., Peter HAGOORT & Colin BROWN 1999. Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience* 11. 657-671.
VAN BERKUM Jos J. A., Pienie ZWITSERLOOD, Peter HAGOORT & Colin BROWN 2003. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research* 17. 701-718.
VAN HERTEN Marieke, Herman KOLK & Dorothee CHWILLA 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research* 22. 241-255.
WAMBACQ Ilse J. A. & James JERGER 2004. Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. *Cognitive Brain Research* 20. 427-437.
WILSON Deirdre 2003. Relevance and lexical pragmatics. *Italian Journal of Linguistics/Rivista di Linguistica* 15. 273-291.