CORTEX 75 (2016) 44–55

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

A role for left inferior frontal and posterior superior temporal cortex in extracting a syntactic tree from a sentence

Chotiga Pattamadilok [a,b,*], *Stanislas Dehaene* [c,d] *and Christophe Pallier* [c,e]

[a] *Laboratoire Parole et Langage (LPL), CNRS: UMR 7309, Aix-Marseille Université, Aix-en-Provence, France*
[b] *Institut National de la Santé et de la Recherche Médicale, Institut du Cerveau et de la Moelle Epinière, UMRS 975, Paris, France*
[c] *Institut National de la Santé de la Recherche Médicale, U992, Cognitive Neuroimaging Unit, Gif/Yvette, France*
[d] *Collège de France, Paris, France*
[e] *Centre National de la Recherche Scientifique, Paris, France*

ARTICLE INFO

Article history:
Received 24 November 2013
Reviewed 7 April 2014
Revised 7 July 2014
Accepted 16 November 2015
Action editor Jean Francois Demonét
Published online 26 November 2015

Keywords:
Syntactic complexity
Syntax network
Sentence processing
Verification process

ABSTRACT

On reading the sentence "*the kids who exhausted their parents slept*", how do we decide that it is the kids who slept and not the parents? The present behavioral and functional magnetic resonance imaging (fMRI) study explored the processes underlying the extraction of syntactically organized information from sentences. Participants were presented with sentences whose syntactic complexity was manipulated using either a center-embedded or an adjunct structure. The goal was to vary separately the sentence syntactic structure and the linear distance between the main verb and its subject. Each sentence was followed by a short subject + verb probe, and the participants had to check whether or not it matched a proposition expressed in the sentence. Behavioral and fMRI data showed a significant cost and enhanced activity within left inferior frontal and posterior superior temporal cortex whenever participants processed center-embedded sentences, which required extracting a nontrivial subtree formed by nonadjacent words. This syntactic complexity effect was not observed during online sentence processing but rather during the processing of the probe and only when the verification could not rely on a superficial lexical analysis. Moreover, the manipulation of linear distance affected performance and brain activity mainly when the sentences did not have a center-embedded structure. We did not find evidence suggesting that tree-extraction, a fundamental operation of a core syntax network, takes place during sentence comprehension. The present finding showed that the syntactic complexity effect, which is an outcome of this operation, became detectable later on, whenever we need to extract structural information not obvious in the superficial sequence of words.

© 2015 Elsevier Ltd. All rights reserved.

* *Corresponding author*. Laboratoire Parole et Langage (LPL), CNRS: UMR 7309, Aix-Marseille Université, Aix-en-Provence, France.
E-mail address: chotiga.pattamadilok@lpl-aix.fr (C. Pattamadilok).
http://dx.doi.org/10.1016/j.cortex.2015.11.012
0010-9452/© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Arguments from linguistics suggest that sentences are represented using a specific data structure in the form of a tree of constituents (Sportiche, Koopman, & Stabler, 2013). This theoretical proposal is motivated by the fact that sentences can be embedded within sentences and that sentences often encode nonlocal dependencies between items. For instance, on hearing or reading the sentence "*This morning the kids who exhausted their parents slept*" (cf. Embedded structure in Fig. 1), how do we know that it is the kids who slept and not the parents? Encoding the sentence as a mere temporal sequence does not suffice. Rather, the linear sequence must be transformed into an internal representation of the dependencies between words. To answer simple questions such as "Who slept?", the human brain must be equipped with a mental

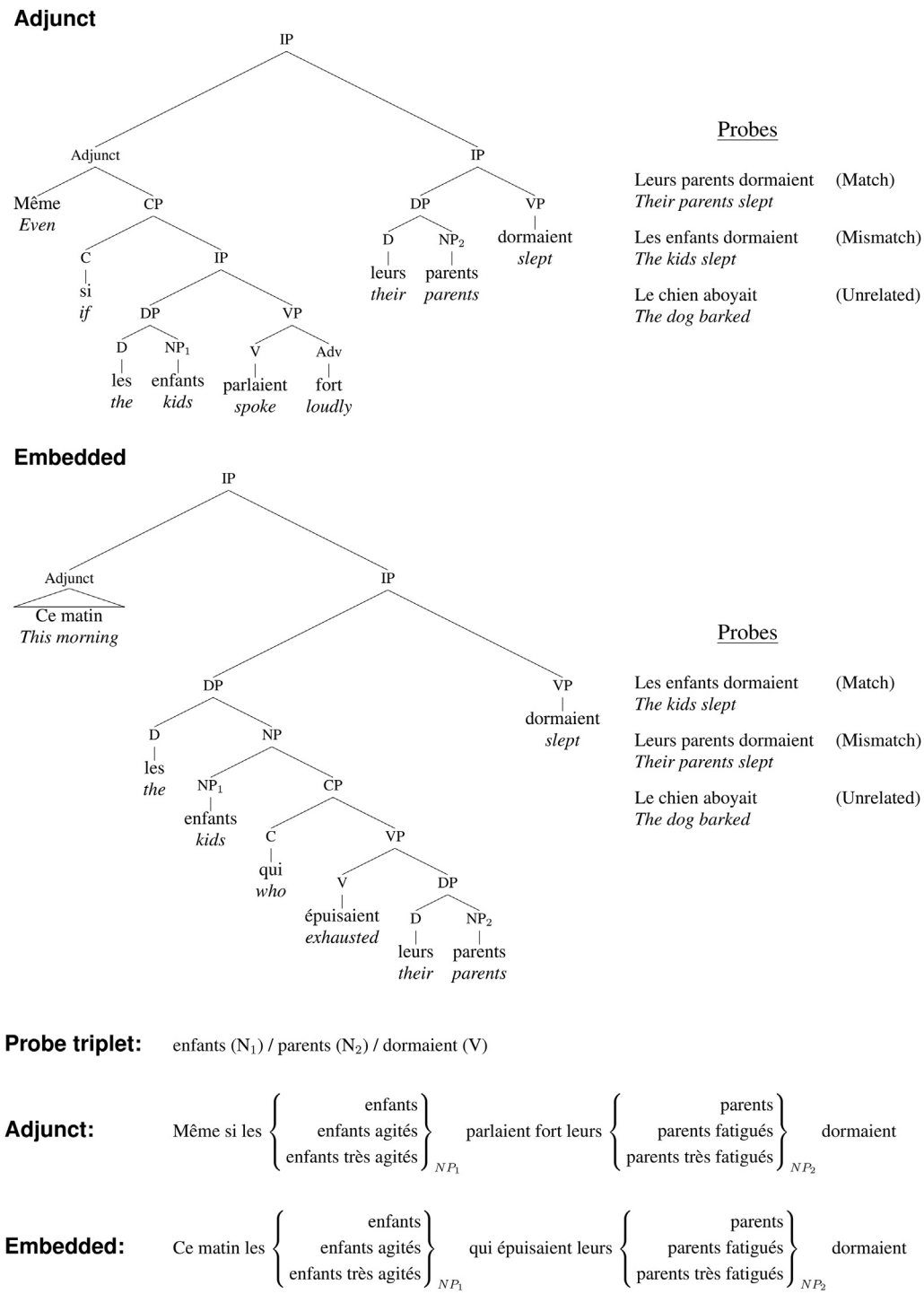

Fig. 1 – Syntactic trees representing the adjunct and embedded structure of the main sentence and their corresponding probes (upper part). The experimental design illustrating the 18 sentences created from each triplet of probes (lower part).

operation for extracting information from the tree representing the sentence. This tree-extraction operation recovers the relation between elements internal to the sentence (e.g., the kids slept), stripped of any additional irrelevant information (e.g., the kids exhausted their parents). If the linguistic analysis is correct, representing trees and manipulating them (including merging, stripping and extracting subtrees) must form a natural toolkit of operations available to the human language system.

Here, we investigate the postulated tree-extraction operation both behaviorally and with brain imaging. Participants performed a verification task based on the semantic content of main sentences and probes. Since a simple reliance on lexical content would not be sufficient, they were required to extract subparts of syntactic trees. The complexity of this operation was manipulated to investigate the cognitive and neural processes underlying the extraction of syntactic structure from a sentence.

More precisely, on each trial, we presented a sentence (9–13 words long) followed by a probe phrase that always contained three words (determiner + noun + verb). The participants were asked to decide whether or not the probe matched the preceding sentence. Complexity was manipulated by changing the structure of the initial sentence, which could either consist of two adjunct clauses or contain a center-embedded clause (Fig. 1). Extensive psycholinguistic research has shown that center-embedded constructions are more complex to process than nonembedded constructions (Gibson & Pearlmutter, 1998; Miller & Isard, 1964). It is worth noting that most studies that have investigated this issue have used sentences with object-relative construction as the prototypical case of embedded structure (Caplan, Alpert, & Waters, 1998; Cooke et al., 2002; Dapretto & Bookheimer, 1999; Love, Haist, Nicol, & Swinney, 2006; Michael, Keller, Carpenter, & Just, 2001; Stromswold, Caplan, Alpert, & Rauch, 1996). To examine the impact of a syntactic structure that was uncontaminated by noncanonical word order, we decided to use a subject-relative clause in the center-embedded condition and contrast it to an adjunct construction.

Conceptually, the verification task could be performed in at least two ways. First, the participants might transform the input sentence into a language-independent semantic code, and then compare the meaning of the probe with the meanings expressed by the main sentence. In this case, at the time of the probe, the matching operation would no longer be affected by the structure of the initial sentence. Second, on reading the probe, they might return to the syntactic representation of the main sentence and extract the relationship between the noun and the verb contained in the probe. In this case, the participants' performance as well as the underlying brain activation would reflect a tree-extracting operation, and we would expect it to be influenced by whether, in the sentence, the embedded tree structure conflicts with the linear sequence of incoming words. In the simpler, *adjunct*, condition, the noun that comes just before the verb is also the subject of the verb, rendering its matching with the probe quite immediate (see example in Fig. 1, upper part). In the more complex, *center-embedded*, condition, the presence of an embedded relative clause brings another noun closer to the verb. We expected that the presence of this noun would create interference in deciding whether the probe matches the sentence or not. In the above example "*The kids who exhausted their parents slept*", it would be hard to decide whether it was the kids or the parents who slept.

We also manipulated an additional variable: the size of the two nouns phrases (NP) included in each sentence. The NP size could be varied by adding optional words (an adjective or an adjective preceded by an adverb), leading to a three-level factor: a noun without adjective (NP = 1, *enfant-parents*), a noun with one adjective (NP = 2, *enfants agités-parents fatigués*), or a noun with an adjective and its modifier (NP = 3, *enfants très agités-parents très fatigués*). These variables also had the effect of increasing the distance between the verb and its subject noun, again rendering the match response partially distinct from the linear sequential structure of words. We assumed that the impact of the linear distance between different elements within the same constituent would be strongly reduced in the situations that require the participants to extract syntactic structure of the sentence in comparison with the situations where shallow analysis of the superficial sequence of words is sufficient. Therefore, we expected the distance between the noun and the associated verb to specifically affect the adjunct structure: the "match" decision would be harder when the distance between the probed noun and the verb increases as well as when the distance between the competing noun and the verb decreases. On the contrary, the decision on "mismatch" trials would be facilitated when the distance between the probed noun and the verb increases.

2. Methods

2.1. Participants

Twenty young adults (12 women, age: mean = 23 years, SD = 3.02, range = 20–33 years) participated in the experiment. All were right-handed native speakers of French. The experiment was approved by the regional ethics committee, and written informed consent was obtained from all participants. They received 85€ for their participation.

2.2. Stimuli

The stimuli were pairs of sentences and probes. Fig. 1 shows the complete design and illustrates how the stimuli were constructed. We started by generating probes, that are, triplets of two nouns and one verb (Fig. 1, lower part). Both nouns (e.g., *kids*, *parents*) were semantically compatible with the verb (e.g., *slept*). For each triplet (e.g., *kids/parents/slept*), we constructed 18 sentences following the scheme described in Fig. 1 (lower part). Half of the sentences possessed an adjunct structure and the other half an embedded structure. The length of the sentences was manipulated by adding one adjective or an adverb followed by an adjective to the noun phrases such that the total length of the sentences varied from 9 to 13 words.

At the highest level of detail, there were 36 types of items, corresponding to the combination of syntactic structure (embedded *vs* adjunct), relation between main sentence and probe (match *vs* mismatch), size of noun phrase 1 (hereafter,

NP1: 1, 2 or 3) and size of noun phrase 2 (hereafter, NP2: 1, 2 or 3).

This material was used to create experimental lists of 252 trials for each participant. Each list comprised 45 critical trials from each of the four combinations of structure and relation. Among these trials, there were 5 trials per combination of NP1 and NP2 lengths. In addition to the critical trials, 36 blank trials (used as implicit baseline in the functional magnetic resonance imaging – fMRI – analyses) and 36 unrelated trials were added. The unrelated trials corresponded to the trials where the nouns and verbs in the probe did not match those in the main sentence. Therefore, the participants could easily recognize that the main sentence and the probe were semantically and lexically unrelated. We expected this condition to provide a low-level baseline where syntactic tree operations would be bypassed.

2.3. *Procedure*

Each trial started with a 500 msec fixation cross followed by the rapid serial visual presentation (300 msec per word) of a main sentence. Five-hundred milliseconds after the last word, the probe was presented and remained on the screen for 700 msec. Stimulus-onset asynchrony was maintained constant at 7.8 sec. The participants were required to read the main sentence and the probe and to judge whether the proposition expressed in the probe matched a proposition expressed in the main sentence. They responded by pushing one of two response buttons placed in their right and left hand. For each participant, the 252 trails were divided into 4 sessions. The stimuli from the different experimental conditions were distributed equally across sessions. No main sentence–probe combination was repeated within the same participant. The assignment of hands to match *vs* mismatch responses was systematically changed after two sessions with counterbalanced order across participants. Before the experiment, a short familiarization block of 6 trials was presented to the participants outside the scanner.

2.4. *Imaging*

The acquisition was performed on a 3 Tesla Siemens Tim Trio system equipped with a twelve-channel coil. For each participant, an anatomical image was first acquired using a three-dimensional gradient-echo sequence and voxel size of $1 \times 1 \times 1.1$ mm. Then, a total of 1496 functional scans were acquired during four sessions of 374 scans each, using an echo-planar sequence sensitized to the blood oxygen level–dependent (BOLD) effect (time repetition = 2.4 sec, time echo = 30 msec, matrix = 64×64, voxel size = $3 \times 3 \times 3$ mm, 40 slices in ascending order).

2.5. *Analyses*

2.5.1. *Behavioral data*

The log-transformed reaction time for trials with correct responses were analyzed with the R software (R Core Team, 2014) using a mixed-effects model with subjects and items as random factors and the fixed factors, structure and relation, also entered as random slopes. The model was fit with the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) and the *p* values (Type III) were computed with the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2014). When a similar model on error data (with a binomial link function) failed to converge, we performed a classic within-subject analysis of variance (ANOVA) on the arcsin-transformed error rates averaged per subject and per condition.

2.5.2. *Imaging data*

Data processing was performed with SPM8 (Wellcome Department of Cognitive Neurology, software available at http://www.fil.ion.ucl.ac.uk/spm). The anatomical scan was spatially normalized to the avg152 T_1-weighted brain template defined by the Montreal Neurological Institute (MNI) using the default parameters (nonlinear transformation). Functional volumes were realigned to correct for motion, spatially normalized to the MNI space using the parameters obtained from the normalization of the anatomy and smoothed with an isotropic Gaussian kernel with a full-width at half maximum (FWHM) equal to 5 mm. For the main analysis, a general linear model was generated for each participant, which included, for each session, eight event-related regressors modeling the two types of main sentences (adjunct *vs* embedded structure) and their associated probes that corresponded to the six possible combinations of the structure of the main sentence and its relation with the probe (i.e., adjunct/unrelated, adjunct/match, adjunct/mismatch, embedded/unrelated, embedded/match, embedded/mismatch). The sentences were modeled by box-car functions that were on for the full duration of the sentence, while the probe was modeled by an event of fixed duration (700 msec). Data were high-pass filtered with a cutoff of 128 sec. Six motion regressors corresponding to translation and rotation in each *xyz* were included in the design matrix: within-session translations ranged from .08 mm to 3.04 mm (mean = .67 mm) and rotations ranged from .1 to 3.8° (mean = .62°). For the group analysis, individual contrast maps representing coefficients to the eight regressors were smoothed with a Gaussian filter (FWHM of 8 mm) and entered in an ANOVA model (sphericity not assumed) with the same eight regressors as in the first-level analysis and one regressor per subject.

In addition to whole-brain analyses, we performed analyses within a priori regions of interest (ROIs) relevant to syntactic processing. The location of the ROIs is illustrated in Fig. 2a. These ROIs included the temporal pole (TP; MNI coordinates of center of mass −48 15 −27), the anterior superior temporal sulcus (aSTS; −54 −12 −12), the posterior superior temporal sulcus (pSTS; −51 −39 3), temporo-parietal junction (TPJ; −39 −57 18), the inferior frontal gyrus orbitalis (IFG orb; −45 33 −6), inferior frontal gyrus triangularis (IFG tri; −51 21 21) and inferior frontal gyrus opercularis (IFG oper; −42 10 22). All regions, except IFG oper, were obtained from a contrast comparing sentences to lists of words in a previous study (see Pallier, Devauchelle, & Dehaene, 2011). Although IFG oper was not found in Pallier et al.'s study (2011), it is considered by many studies to be crucial in syntax processing (Friederici, 2011; Santi & Grodzinsky, 2010). We, therefore, added this region to those listed in Pallier et al. (2011) by using the Talairach coordinates (−42, 11, 20) reported in a study by Amunts et al. (2004), converting them to MNI, defining a

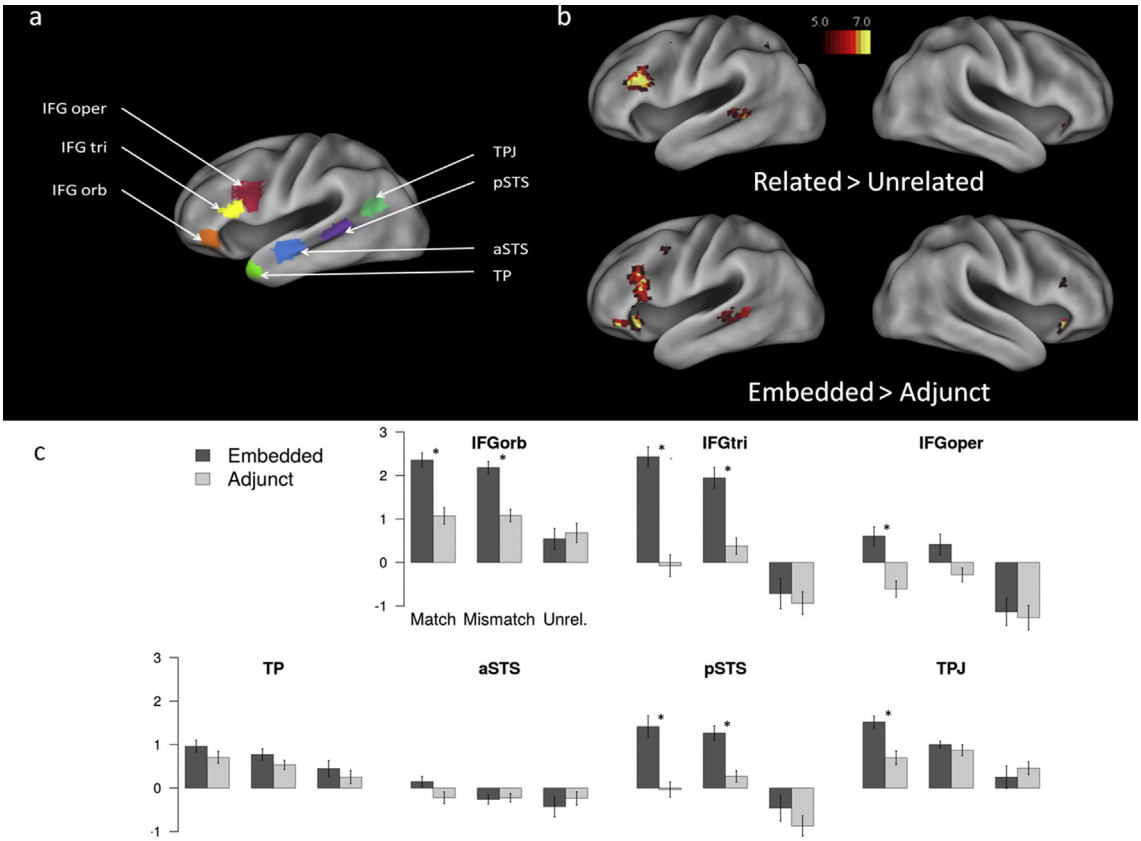

Fig. 2 – (a) The brain regions included in the ROIs analyses, (b) SPM maps of the effects of relation and structure on related trials observed during probe presentation, (c) responses to probes in seven ROIs as a function of the experimental conditions defined by (structure * relation). The vertical bars show the within-subject standard errors (Cousineau, 2005). The significant effects of structure were indicated by *.

sphere of 15-mm radius and intersecting it with the gray matter mask provided along with SPM8.

The effects of the size of NP1 and NP2 were considered separately in additional analyses. For both NP1 and NP2, a general linear model was created for each participant. Given that each session contained a small number of trials from each combination of main sentence structure, its relation with the probe (match *vs* mismatch) and the size of NP (NP = 1, 2 or 3), we decided to pool the data from the four sessions together. The model included 16 event-related regressors modeling the two types of main sentences (adjunct structure *vs* embedded structure), the two unrelated probes (those preceded by adjunct- and those preceded by embedded-structure main sentence) and the 12 related probes that corresponded to the combinations of main sentence structure, its relation with the probe (match *vs* mismatch) and the size of NP (NP = 1, 2 or 3). The same 16 regressors were included in the ANOVA model of the second-level group analysis. In addition to the whole-brain analysis, the effects of the size of NP were also explored in ROI analyses.

3. Results

For both behavioral and brain imaging data, the analyses were performed in two stages. First, we looked at the effect of structure (embedded *vs* adjunct) and relation (match, mismatch, unrelated), without taking into account the manipulation of sizes of NP1 and NP2. In a second stage, we explored the effects of NP size.

3.1. *Behavioral data*

3.1.1. *Effects of structure and relation*

Data from two participants were excluded from behavioral and fMRI analyses due to high error rates (both at ~32%). The remaining 18 participants had an average error rate of 8.7% (SD = 5.8%). The mean reaction times (RTs) on correct trials and the mean error rates are displayed on the left panel of Fig. 3 as a function of structure (embedded *vs* adjunct) and relation (match, mismatch and unrelated). The analysis of reaction times, using a mixed-effects model with structure and relation defined both as fixed factors and as random slopes nested inside subject and item random factors, revealed a significant interaction between structure and relation [F=(1, 3668) = 23.4, $p < .001$] as well as significant main effects for each factors [structure: $F(1, 23) = 34.4$, $p < .001$; relation: $F(2, 19) = 60.8$, $p < .001$]. Participants were significantly slower in the embedded condition than in the adjunct condition in the match and mismatch conditions [95% confidence interval – CI – for the difference embedded–adjunct restricted to match = (169, 277) msec; difference

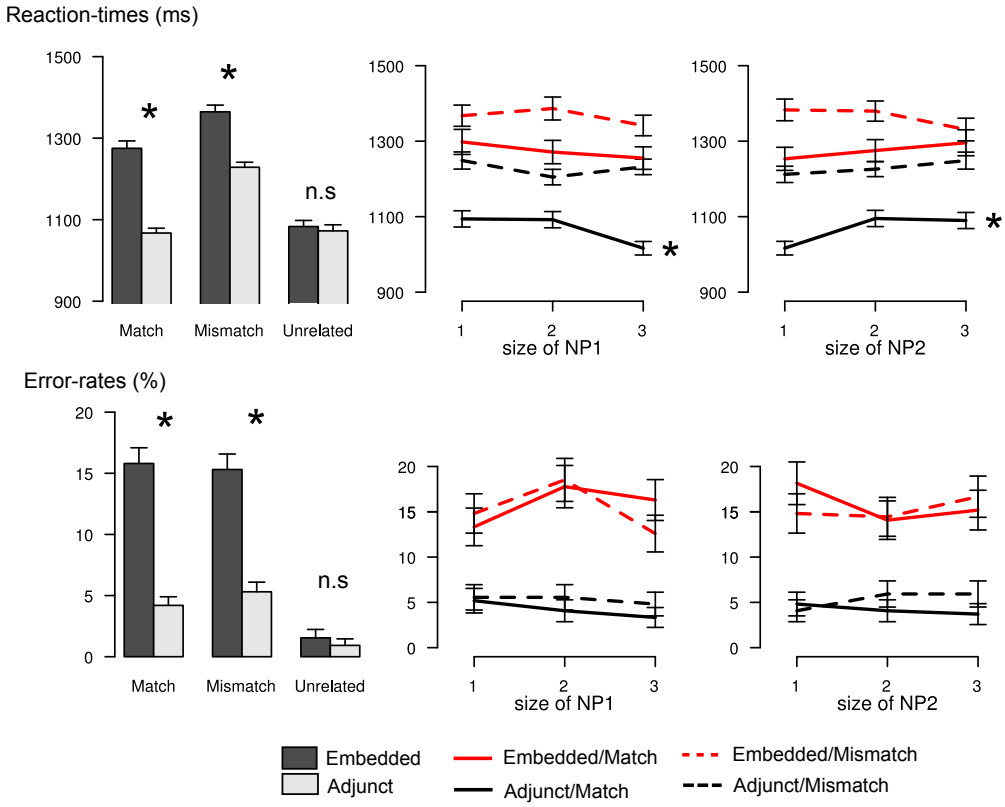

Fig. 3 – Reaction times and error rates obtained in the probe-matching task illustrating the effects of structure and relation (left panel) and the effects of the size of NP1 and NP2 for the following conditions: adjunct/match, adjunct/mismatch, embedded/match and embedded/mismatch (right panel). The vertical bars show the within-subject standard errors (Cousineau, 2005). The significant effects of structure and NP size were indicated by *.

embedded–adjunct restricted to mismatch: 95% CI = (73, 189) msec] but not in the unrelated condition [95% CI = (−41, 51) msec]. The pattern of results for error rates was similar: structure and relation interacted significantly [$F(2, 34) = 18.1$, $p < .001$] and yielded main effects [structure: $F(1, 17) = 40.9$, $p < .001$; relation: $F(2, 34) = 29.8$, $p < .001$]. Embedded structure produced more errors than adjunct structure in the match [95% CI = (8.4, 14.8) %] and mismatch conditions [95% CI = (5.3, 14.7) %] but not in the unrelated condition [95% CI = (0, 1.2) %]. Fewer errors were found in the unrelated compared with match [$F(1, 17) = 42.4$, $p < .001$] and mismatch conditions [$F(1, 17) = 38.8$, $p < .001$].

Overall, the existence of the effect of structure clearly suggests that, on match and mismatch trials, the participants relied on a syntactic representation of the main sentences. The absence of this effect on unrelated trials confirmed that this condition acted as a baseline where syntactic analysis was bypassed.

3.1.2. *Effects of NP size*

We investigated the effects of increasing the size of NP1 and NP2 to assess whether an increase in the distance between the verb and its subject noun, thus rendering the match response partially distinct from the linear sequential structure of words, made the matching operation more difficult. The right panel of Fig. 3 presents the effects of the size of NP1 and NP2 separately for the four conditions: adjunct/match, adjunct/mismatch, embedded/match and embedded/mismatch. The effects of NP size were assessed in each condition by running mixed-effect linear models with logRT as the dependent variable and the size of either NP1 or NP2 as fixed factors (1–3). Item and subject random factors were included with random slopes for NP1 size and NP2 size. As expected, these analyses revealed significant effects of size of NP1 [$t(785) = -2.6$, $p < .01$] and NP2 [$t(785) = 2.8$, $p < .01$] in the adjunct/match condition, with opposite signs as predicted in the Introduction. Similar analyses on error rates (using generalized mixed-effect models with binomial link function) did not reveal a significant ($p < .05$) effect on either NP1 or NP2, in any condition.

3.2. *Imaging data: whole-brain analyses*

All participants showed a similar global network of brain areas recruited by the task, which included occipital and basal temporal areas (regions involved in reading), parietal, mid-temporal (anterior and posterior) and frontal regions (both inferior and median gyri).

Unless stated otherwise, the threshold of $p < .05$ with family-wise error (FWE) correction for multiple comparisons

($t > 4.97$) was applied in the whole-brain analyses presented below.

3.2.1. *Activation during sentence presentation*

3.2.1.1. Effect of structure. We first examined the effect of sentence complexity during main sentence presentation. The coefficients associated to the regressors modeling the period during which the main sentence was presented did not differ significantly in the embedded and in the adjunct conditions (neither with FWE threshold of $p < .05$, $t > 4.97$ nor with an uncorrected voxel-based threshold at $p < .001$ combined with a cluster size FWE corrected at $p < .05$).

This absence of a syntactic complexity effect may be surprising since a clear behavioral difference emerged between embedded and adjunct conditions. However, further analyses of brain activity, described below, showed that the effect of complexity occurred while the probe was presented, when the participant had to explicitly extract information from the sentence.

3.2.2. *Activation during probe presentation*

3.2.2.1. Effect of structure. In contrast to what happened during sentence presentation, the effect of syntactic complexity was observed when participants processed the probes. The embedded > adjunct contrast revealed stronger activations in the left IFG tri, IFG orb, insula and mid-temporal lobe. More localized activation was found in the right IFG tri and insula (Table 1).

3.2.2.2. Effect of relation. *Related vs unrelated probes*: In the related trials (match or mismatch), participants had to analyze the main sentence, whereas in the unrelated trials, a superficial analysis based on the presence of lexical items sufficed. We, therefore, expected the basic contrast between related *vs* unrelated trials to show the brain regions involved in the extraction of information from the structural representation of the sentence. This contrast indeed revealed stronger activations in the left IFG tri, IFG orb, pSTS, supplementary motor area (SMA), precentral gyrus, inferior parietal lobe, the bilateral insula and caudate nucleus (see the upper part of Fig. 2b and Table 1). The opposite contrast did not yield any significant result.

Mismatch vs match probes: Although the difference between mismatch and match trials was significant at behavioral level, no corresponding difference was observed in brain activation in either direction.

3.2.2.3. Interaction between structure and relation. An examination of the interaction of the two factors (*embeded vs adjunct*) * (*related vs unrelated*) suggests a dissociation between the effects of structure on related and unrelated probes. Effects failed to pass the significance threshold of FWE-corrected $p < .05$ voxel-based, but an effect was detected in the left IFG orb (−45 35 −5) when a threshold of $p < .001$ voxel-based (uncorrected) was combined with a cluster-level correction for multiple comparison (FWE-corrected $p < .05$).

We further analyzed the effect of structure separately on related and unrelated trials. Focusing on the related trials only, and collapsing match and mismatch conditions, we contrasted the responses to probes in trials with embedded *vs* adjunct sentences. We found that embedded trials produced stronger brain activation than adjunct trials in a network of regions quite similar to the one involved in the related *vs* unrelated contrast, notably the left inferior frontal regions, the left pSTS and the left SMA (see the lower part of Fig. 2b and Table 1). The opposite contrast (adjunct *vs* embedded) did not detect any significant activity.

The analysis of the structure effect performed on unrelated trials was consistent with the behavioral data, i.e., no effect of structure was observed. The absence of a structure effect on this contrast provides a highly relevant piece of information

Table 1 – Regions showing significant effects in embedded probe > adjunct probe, related probe > unrelated probe and embedded-related probe > adjunct-related probe contrasts. Anatomical labels are obtained with the Anatomical Automatic Labeling toolbox (http://www.cyceron.fr/web/aal_anatomical_automatic_labeling.html).

	Cluster size	T (df = 119)	x	y	z
Analysis at $p < .05$ family-wise error corrected					
Embedded probe > Adjunct probe					
Frontal Inf Tri L	57	5.81	−48	20	28
Frontal Inf Tri L		5.66	−54	20	19
Frontal Inf Tri L		5.20	−48	29	22
Frontal Inf Tri L	1	5.03	−57	20	1
Frontal Inf Tri R	3	5.02	54	26	28
Frontal Inf Orb L	2	5.42	−48	26	−8
Frontal Inf Orb L	3	5.27	−48	38	−8
Insula L	8	5.48	−27	23	−5
Insula R	4	5.39	33	26	−5
Temporal Mid L	34	5.80	−57	−37	−2
Temporal Mid L		5.67	−51	−46	4
Related probe > Unrelated probe					
Frontal Inf Tri L	173	7.75	−48	20	25
Frontal Inf Tri L		6.24	−54	26	19
Frontal Inf Tri L	7	5.50	−54	20	4
Frontal Inf Tri L	1	5.16	−42	26	−2
Frontal Inf Orb L	2	5.11	−48	44	−8
Supp_Motor_Area L	56	6.48	−6	11	55
Supp_Motor_Area L		6.25	−3	5	67
Frontal Sup Medial L		5.89	−3	26	49
Precentral L	11	6.12	−45	2	52
Insula L	2	5.14	−30	26	1
Insula R	1	5.28	30	23	−2
Caudate L	37	6.02	−9	8	7
Caudate R	17	6.00	9	8	10
Temporal Mid L	37	7.02	−57	−40	1
Parietal Inf L	5	5.39	−30	−52	40
Embedded-related probe > Adjunct-related probe					
Insula L	339	7.38	−27	23	−2
Frontal Inf Orb L		7.37	−48	38	−8
Frontal Inf Tri L		7.09	−54	23	19
Frontal Inf Oper L		6.23	−51	17	10
Frontal Mid L		5.49	−51	14	40
Insula R	27	7.09	33	23	−5
Frontal Inf Oper R	45	6.35	54	23	34
Frontal Mid L	1	5.12	−39	53	1
Frontal Sup Medial L	68	6.33	0	32	49
Supp Motor Area L		5.87	−3	14	55
Supp Motor Area L		5.73	−3	5	67
Precentral L	8	5.29	−39	2	43
Temporal Mid L	74	6.31	−57	−43	1
Parietal Inf L	4	5.33	−36	−55	46
Precuneus L	2	5.11	3	−64	46

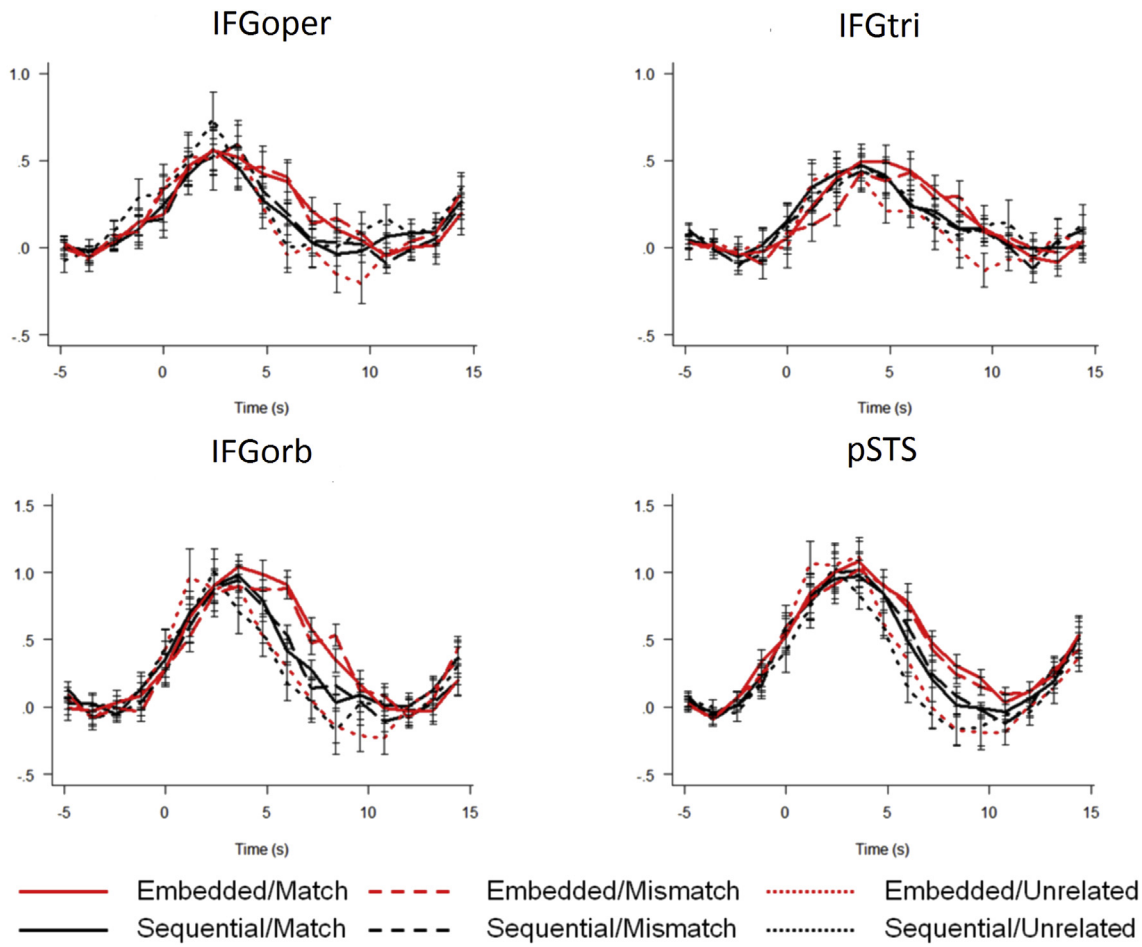

Fig. 4 – Time course of activation at four locations observed during probe presentation.

since it confirms that the significant structure effect obtained on related trials was due to the matching operation performed on the probes rather than to a mere carryover effect of activation evoked by the preceding sentence.[1]

Fig. 4 plots the time course of activation in the IFG tri, IFG orb, pSTS and IFG oper regions belonging to the core syntax network (Friederici & Kotz, 2003; Friederici, Makuuchi, & Bahlmann, 2009; Pallier et al., 2011; Saur et al., 2010; Tyler & Marslen-Wilson, 2008). The first three ROIs were reported in Pallier et al. (2011) as sensitive to constituent structure and the last area is centered on BA44 as reported in Amunts et al. (2004), a ROI for syntactic processing. The plots illustrate all the effects identified above: an initial activation unrelated to sentence complexity, followed by a divergence of the curves for the unrelated condition, which dropped quickly, and the development of an effect of syntactic structure (higher activation on center-embedded trials) only for related probes.

3.2.2.4. Effects of NP size. We performed whole-brain analyses searching for effects of size of NP1 and NP2 for each combination of structure by relation. No significant effect was detected at a FWE threshold of $p < .05$ ($t > 4.97$).

3.3. *Imaging data: analyses by ROIs*

To complement the whole-brain analyses presented above, we examined the brain responses to the sentences and to the probes in seven ROIs known to be involved in syntactic processing (cf. the Method section). Each contrast was examined with a threshold of $p < .05$ corrected for multiple comparisons

[1] The presence of a structure effect only in the related but not in the unrelated condition is an important argument suggesting that the extraction of the syntactic structure might take place only when this operation is necessary for the task. Therefore, it was important to ascertain that the null result obtained in the unrelated condition was not due to a lack of statistical power. In the analyses presented above, the number of events for the related trails was twice as many as the one for the unrelated trials and this might explain the null result in the latter condition. To address this issue, we split the related condition into match and mismatch trials, obtaining the same number of events for the three types of trials (match, mismatch and unrelated). We then computed the structural effect contrasts separately for the three types of probes. These analyses showed reliable activations in the left IFG and the pSTS in the match and mismatch conditions and no activation in the unrelated condition (at $p < .001$ voxel-wise uncorrected for the unrelated). Moreover, the interaction contrast between structure (embedded *vs* sequential) and relation (related *vs* unrelated) also detected activation in the left IFG.

across the seven ROIs using the false detection rate (FDR) method. The aim of the ROI analysis was to ascertain the main result pattern obtained in the whole-brain analysis: (1) the absence of the syntactic structure effect during the processing of the main sentences and the unrelated probes and (2) the presence of the effect during the processing of the related probes could be replicated even when the analyses were focused on small sets of brain regions that are relevant to syntactic processing.

3.3.1. *Activation during sentence presentation*

3.3.1.1. Effect of structure. Consistent with the result of the whole-brain analyses, no effect of structure on the main sentences was detected.

3.3.2. *Activation during probe presentation*

3.3.2.1. Effect of structure. The contrasts between the probes that examined the effects of main sentence structure on match, mismatch and unrelated probes are reported in Fig. 2c displaying the coefficients of the probe regressors according to condition. A significant effect of structure was found in the left IFG orb, IFG tri, pSTS for both match and mismatch probes and in IFG oper and TPJ for match probes. No region showed a significant effect of structure for unrelated probes.

3.3.2.2. Effect of NP size. We performed analyses looking at the linear effects of size of NP1 and NP2 in each ROI for each combination of structure and relation. No effect survived multiple correction threshold (FDR, $p < .05$). However, it is worth noting that in the adjunct/match condition, the pattern of brain activations were parallel to the response time pattern, especially in the IFG tri, IFG oper, IFG orb and pSTS (see Figs. 3 and 5 for the effects of NP size on response time and brain activation, respectively). Activations and response times both decreased with NP1's size and increased with NP2's size.

In summary, the ROI analyses confirmed the main finding obtained in the whole-brain analyses: the syntactic complexity effect was not observed during online sentence processing but rather during the processing of the probe, and only when the verification could not rely upon a superficial lexical analysis.

4. Discussion

Behavioral and brain imaging data indicated that participants were sensitive to the syntactic structure of the sentence. Behaviorally, indeed, performance was better with adjunct than with center-embedded structure. Nevertheless, this effect of syntactic complexity only appeared in related trials that required extracting relational information from the syntactic structure. In unrelated trials, where a shallow verification of lexical items sufficed to respond, no effect of structure was found. Thus, the results suggest that structural complexity affects primarily the hypothesized "tree-extraction" operation when subjects are forced to extract information from a sentence, and that this effortful operation is bypassed when the probe is obviously unrelated to the original sentence.

Brain imaging analyses further allowed us to dissociate the effects of structure during online parsing of the main sentence and during the matching operation performed on the probes.[2] They confirmed that the effect of the structure was present only at the time of the probes and only when the probes were related to the main sentences. The absence of any effect of structure on unrelated trials provides evidence allowing us to rule out the possibility that the effect of structure found on related probes was a late carryover effect from the parsing of the main sentence.

Our manipulation of syntactic complexity did not produce a focal effect but affected a set of areas that have been repeatedly identified as forming a "core syntax" network (Friederici & Kotz, 2003; Friederici et al., 2009; Pallier et al., 2011; Saur et al., 2010; Tyler & Marslen-Wilson, 2008). On both match and mismatch trials, the related probes preceded by center-embedded structures yielded stronger activations within the left IFG pars opercularis, triangularis and orbitalis, the pSTS as well as regions within the insula bilaterally and the left SMA (Table 1). This network was also identified in the related > unrelated contrast that compared trials where the syntactic analysis was necessary with trials where a shallow lexical strategy sufficed. The latter finding indicates that the large neural network identified here is triggered not only by the higher complexity of center-embedded sentences, but also whenever there is a need for an internal manipulation of syntactic representations. It is also likely that this more complex and time-consuming language processing situation increases activity in the brain areas outside the classic syntactic network, for instance, those that are involved in lexico-semantic (e.g., IFG orb, middle temporal lobe) or even speech processing (e.g., SMA, precentral cortex, insula).

The fact that the effect of structure was restricted to the processing of the probes can be accounted for in two different

[2] Because the probes systematically followed the sentences at a short time interval and because of the hemodynamic response's sluggishness, the regressors of the probes and sentences are collinear, which makes the estimation of effects less precise and this negatively impacts the detectability of effects. An index of the impact of collinearity is the variance inflation factor (VIF). The VIF associated to the sentence regressors in our individual design matrices had a value of 1.8, meaning that the standard error of estimates of the sentences are inflated by 34% (sqrt(1.8) = 1.34). This is a relatively small amount (in the literature, collinearity has been considered to have an important impact when the VIF is more than 5 or 10).We could have modeled only the duration of the last verb of the sentence, rather than the full sentence. The rationale was that the effect of syntactic complexity may be the strongest there (although the complexity effects could emerge earlier in the sentence, from the beginning of the embedded relative clause). A difficulty with this proposal is that it would have resulted in an even stronger correlation between the sentence and the probe regressors, and in a lack of fit of the model. We acknowledge that because of the proximity of the probe to the sentence, one cannot exclude the possibility that a complexity effect on the sentence could carry over to the regressors associated the probes. It is, therefore, noteworthy that when the probe is unrelated to the sentence, no complexity effect was detected (neither on the sentence nor on the probe), while these effects are robust when the probe is related to the sentence. This is the central observation that makes us believe that the full syntactic tree is computed only when needed.

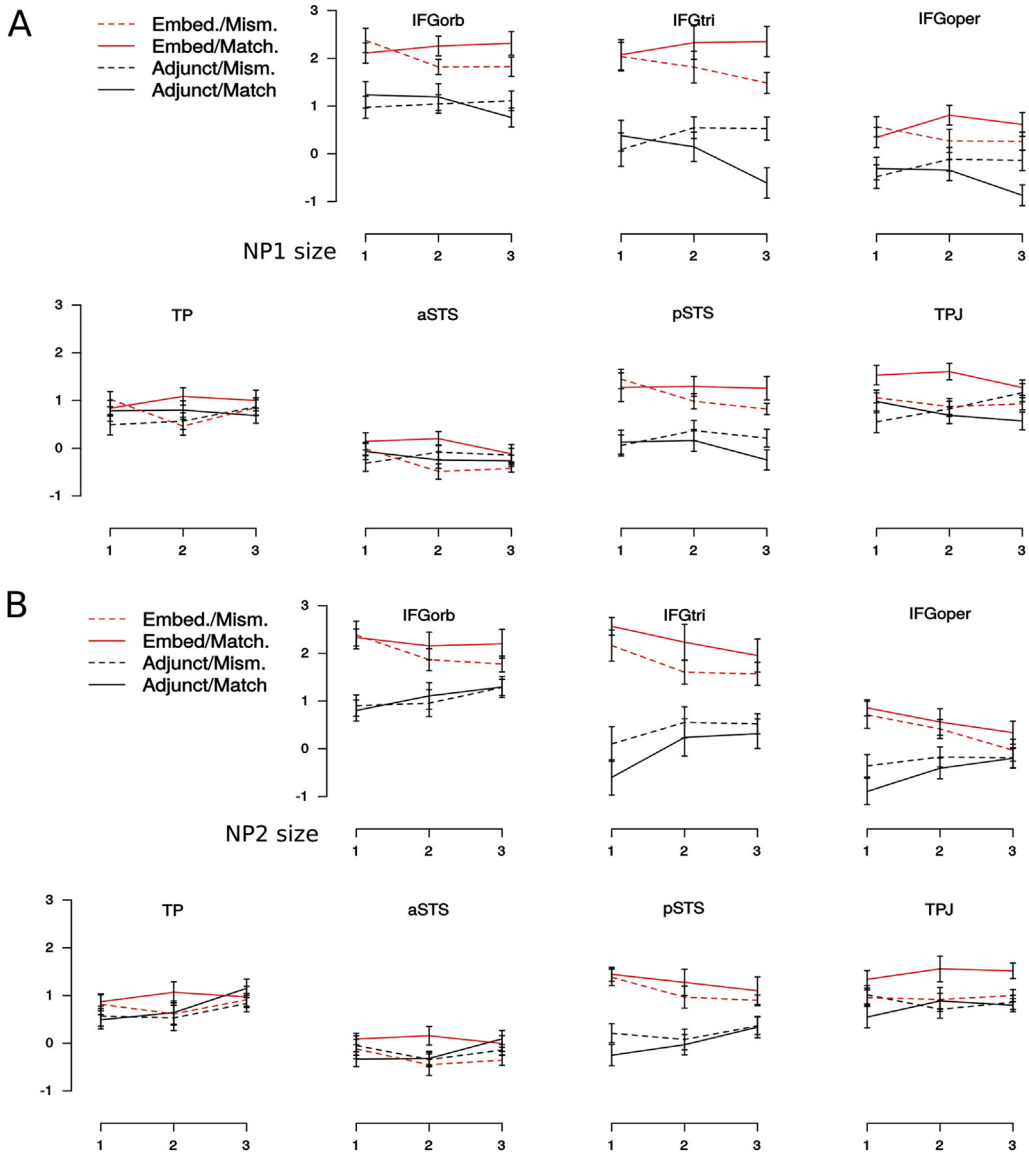

Fig. 5 – Response to probes in the seven ROIs as a function of the size of NP1 (A) and NP2 (B). The vertical bars show the within-subject standard errors (Cousineau, 2005).

ways. One possibility is that the full syntactic structure is computed online during the sentence presentation and reused during the verification task, but there is no differential cost for parsing embedded *vs* adjunct structures or such a cost is too small to be detected by our experiment. An alternative possibility is that the full syntactic structure is not computed online during sentence presentation but only later and only as necessary for the verification task. Although partial trees (phrasal fragments) might have been computed while processing the sentence, they might not be integrated into a complete syntactic tree that a complexity effect could be observed. This idea fits with the hypothesis of "good-enough" or "shallow" parsing according to which sentence processing can proceed without systematically computing a full analysis (Ferreira & Patson, 2007; Sanford & Sturt, 2002; Townsend & Bever, 2001). In our study, only when the subject actively performed the verification task would the full parse be generated and used to guide the tree-extraction process.

While our data do not resolve this alternative, they establish that a strong difference in activation to adjunct and embedded constructions occurs during the extraction of information necessary to perform the verification task. In the adjunct condition (e.g., *Even if the kids spoke loudly* ***their parents slept***), the transparent match between the syntactic structure and the linear sequence of words facilitates the identification of the match probe (*their parents slept*) and the rejection of the mismatch probe (*the kids slept*). The immediate mapping between the main sentence and the match probe at both semantic and syntactic levels could also result in a repetition suppression effect (Devauchelle, Oppenheim, Rizzi, Dehaene,

& Pallier, 2009). Although a difference between mismatch and match trials was indeed observed at the behavioral level, no corresponding neural correlate was found, perhaps due to a lack of statistical power in this subcomparison.

Conversely, in the center-embedded condition (e.g., ***The kids** who exhausted their parents **slept***) there is a conflict between the linear sequence of the words and the syntactic structure. The presence of an embedded relative clause brings another noun closer to the verb. The presence of this noun creates an interference when deciding that the "*the kids slept*" matches the main sentence and that the "*their parents slept*", which shares the same linear sequence of words as in the main sentence, does not. This behavioral effect is directly paralleled by a greater activation of the core syntax network, presumably to overcome the interference.

It is noteworthy that the computation of the full syntactic structure did not systemically take place, even though the task required it on 2/3rd of the trials. This absence of task-relevant anticipation is nevertheless coherent with the view that the full syntactic analysis took place only when it became necessary during the verification process. Our observation is consistent with findings previously reported by Caplan, Chen, and Waters (2008) and Newman, Lee, and Ratliff (2009). Using a similar verification task, Caplan et al. (2008) reported an effect of syntactic complexity (object-extracted *vs* subject-extracted structures) only in a late time period associated with the presentation of the probe and not in an early time period associated with the presentation of the main sentences. Newman et al. (2009) further dissociated the BOLD signals related to the sentence and the probe by adding a 6-sec delay between the two processing phases and observed a larger syntactic effect during the probe phase compared with the sentence reading phase.

Other studies have also reported that the syntactic complexity does not necessarily induce an increase in brain activity during online sentence processing. One striking example is the study by Indefrey, Hellwig, Herzog, Seitz, and Hagoort (2004) who observed a syntactic complexity effect in sentence production but not in sentence comprehension with the same materials. This observation makes a lack of power highly unlikely as an explanation of the null result for comprehension in their study. The authors concluded that the degree to which listeners recruit syntactic processing resources in language comprehension may be a function of the syntactic demands of the task. Similarly, Caplan et al. (2008) and Caplan (2010) compared sentences with center-embedded, object-extracted relative clauses and sentences with right-branching, subject-extracted relative clauses in both pseudoword and font change detection tasks. They did not observe an effect of syntactic complexity, at least when the sentence could be understood just based on the semantic properties of the words.

In addition to the manipulation of the syntactic structure, we also manipulated, in the main sentence, the distance between the verb and its subject noun. As mentioned in the Introduction, this manipulation rendered the match response partially distinct from the linear sequence of words. RT data showed that this manipulation of the linear distance mainly affected performance on the adjunct sentences. On the matched trials, participants were slowed down when the distance between the probed noun and verb increased, as well as when the distance between the competing noun and verb decreased. The same result pattern was found at the brain imaging level, although the effects were not significant.

The fact that the manipulation of the distance does not seem to affect the two types of syntactic structures in the same way is coherent with the idea of a late or optional computation of syntactic structure, as mentioned above. When searching for information, individuals may first base their judgments on a shallow analysis of the superficial sequence of words. Only if this analysis fails, would they switch to a syntactic tree-extraction mode, which is a more effortful and time-consuming process. Nevertheless, this is clearly a hypothesis, and studies using a higher temporal resolution method would be necessary to determine whether the sequential and syntactic modes operate in parallel or sequentially.

5. Conclusion

The findings reported here suggest that our cognitive system is equipped with a syntactic toolkit that comprises a tree-extraction operation, a process that takes a sentence structure as input and extracts useful information encoded in a subtree of the original structure. This operation imposes a measurable cost on response times, and its neural basis lies in a classical left-hemispheric core syntax network. The present findings also suggest a possibility that full parsing based on long-distance computations, which is a long and effortful operation, may not always take place spontaneously but only when a simpler operation based on surface information is insufficient to resolve the task. Further research is nevertheless needed to confirm this controversial observation.

Acknowledgments

This research was supported by a grant from the French Ministry of Research (ANR 2010 BLANC 1403 01).

REFERENCES

Amunts, K., Weiss, P. H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J. M., et al. (2004). .Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space—the roles of Brodmann areas 44 and 45. *NeuroImage, 22*(1), 42–56.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-6 http://CRAN.R-project.org/package=lme4.

Caplan, D. (2010). Task effects on bold signal correlates of implicit syntactic processing. *Language and Cognitive Processes, 25*(6), 866–901.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience, 10*(4), 541–552.

Caplan, D., Chen, E., & Waters, G. (2008). Task-dependent and task-independent neurovascular responses to syntactic processing. *Cortex, 44*(3), 257–275.
Cooke, A., Zurif, E. B., DeVita, C., Alsop, D., Koenig, P., Detre, J., et al. (2002). Neural basis for sentence comprehension: grammatical and short-term memory components. *Human Brain Mapping, 15*(2), 80–94.
Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to loftus and masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42–45.
Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron, 24*(2), 427–432.
Devauchelle, A. D., Oppenheim, C., Rizzi, L., Dehaene, S., & Pallier, C. (2009). Sentence syntax and content in the human temporal lobe: an fmri adaptation study in auditory and visual modalities. *Journal of Cognitive Neuroscience, 21*(5), 1000–1012.
Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass, 1*(1–2), 71–83.
Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews, 91*(4), 1357–1392. http://dx.doi.org/10.1152/physrev.00006.2011.
Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *NeuroImage, 20*(Suppl. 1), S8–S17.
Friederici, A. D., Makuuchi, M., & Bahlmann, J. (2009). The role of the posterior superior temporal cortex in sentence comprehension. *NeuroReport, 20*(6), 563–568.
Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences, 2*(7), 262–268.
Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J., & Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language, 89*(2), 312–319.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*. R package version 2.0-6 http://CRAN.R-project.org/package=lmerTest.
Love, T., Haist, F., Nicol, J., & Swinney, D. (2006). A functional neuroimaging investigation of the roles of structural complexity and task-demand during auditory sentence processing. *Cortex, 42*(4), 577–590.
Michael, E. B., Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). fMRI investigation of sentence comprehension by eye and by ear: modality fingerprints on cognitive processes. *Human Brain Mapping, 13*(4), 239–252.
Miller, G. A., & Isard, S. (1964). Free recall of self-embedded English sentences. *Information and Control, 7*(3), 292–303.
Newman, S. D., Lee, D., & Ratliff, K. L. (2009). Off-line sentence processing: what is involved in answering a comprehension probe? *Human Brain Mapping, 30*(8), 2499–2511.
Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences, 108*(6), 2522.
R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.
Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences, 6*(9), 382–386.
Santi, A., & Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage, 51*(4), 1285–1293. http://dx.doi.org/10.1016/j.neuroimage.2010.03.034.
Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., et al. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *NeuroImage, 49*(4), 3187–3197.
Sportiche, D., Koopman, H., & Stabler, E. (2013). *An introduction to syntactic analysis and theory*. Wiley-Blackwell.
Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language, 52*(3), 452–473.
Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules* (Vol. 1950). MIT Press.
Tyler, L. K., & Marslen-Wilson, W. (2008). Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 1037–1054.