Syntactic and Thematic Effects on BOLD Signal associated with Comprehension and Determination of Plausibility of Sentences with Relative Clauses

David Caplan [1]
Evan Chen [1,2]
Gloria Waters [3]

[1] Neuropsychology Laboratory,
Department of Neurology,
Massachusetts General Hospital

[2] MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging,
Massachusetts General Hospital

[3] Department of Communication Disorders,
Boston University

Address Correspondence to:
David Caplan, M.D., Ph.D.,
Neuropsychology Laboratory,
175 Cambridge Street, Suite 340,
Boston, MA, 02114
Tel: 617-726-3274
FAX: 617-724-7836
Email: dcaplan@partners.org

Abstract

Behavioral and BOLD signal measures were obtained while fifteen participants performed two tasks when presented sentences with more complex "subject-object (SO)" relative clauses and less complex "object-subject (OS)" relative clauses. In Experiment 1 (plausibility judgment), participants made speeded judgments regarding the plausibility of written sentences. In experiment 2 (non-word detection), participants made speeded judgments regarding whether a written sentence contained a non-word. The naturalness of thematic role assignment and the syntactic structure of a sentence affected behavioral and neurovascular results and the effects of these factors were increased when subjects made decisions about sentence plausibility. The results bear on the neural location of operations involved in assigning the syntactic structure of sentences and the interaction of these processes with thematic role assignment in the processes of comprehension and determination of plausibility of sentence meaning.

Syntactic and Thematic Effects on BOLD Signal associated with Comprehension and Determination of Plausibility of Sentences with Relative Clauses

Introduction

Sentences containing "wh words," such as the relative clauses illustrated in (1) – (3), have been studied with functional neuroimaging to investigate the neural basis of processing syntactic structure.

1. The reporter who the senator attacked admitted the error. (SO)[1]

2. The reporter who attacked the senator admitted the error. (SS)

3. The senator attacked the reporter who admitted the error. (OS)

Object-extracted structures such as (1) are more demanding than subject-extracted structures (2) and (3) for several reasons. In (1), the reader/listener must retain two noun phrases in memory before encountering a verb that allows thematic roles to be assigned. In (2) and (3), s/he encounters a single noun phrase, then a verb, and then a second noun phrase. Assuming that thematic roles are assigned at the verb, there is a greater memory load in (1) than in (2) and (3); assuming that assigning two thematic roles simultaneously is more costly than assigning one, there is also a greater computational load in (1) than in (2) or (3) (Gibson, 1998). In addition, there is a memory load associated with the center-embedded structures (1) and (2) that is not present in the right branching structure (3), since the sentence-initial noun must be retained in memory over the embedded relative clause for later assignment as the subject and agent of the main verb in (1) and (2), which is not required in (3).

[1] Regarding nomenclature, in sentence (1), the "head" of relative clause (*the reporter*) is the subject of the main clause and the object of the verb of the relative clause, and (1) can be called a "subject–object (SO)" structure. (2) is called a "subject- subject (SS)" structure, because the head of relative clause (*the reporter*) is the subject of the main clause and the subject of the verb of the relative clause. (3) is called an "object- subject (OS)" structure, because the head of relative clause (*the reporter*) is the object of the main clause and the subject of the verb of the relative clause. Similar nomenclature can be extended to other sentence types.

Neurovascular responses to sentences such as (1) – (3) have been taken as evidence regarding the localization of the operations and memory systems that support these memory and computational demands during sentence comprehension (Ben Shachar et al, 2003, 2004; Caplan et al , 1998, 1999, 2000, 2002; Cooke et al, 2001; Fiebach et al. 2001; Indefrey et al, 2001; Just et al, 1996; Stromswold et al, 1996; Waters et al, 2003). Studies presenting these structures have interpreted such responses as reflections of the memory load associated with maintaining items in memory while assigning syntactic structure (Just et al, 1996), the computational demands associated with assigning thematic roles to two noun phrases when a verb is encountered (Feibach et al., 2001), the combination of these demands with those of center-embedding (Caplan et al , 1998, 1999, 2000, 2002, 2003; Stromswold et al, 1996; Waters et al, 2003), establishing the connection between the position of a noun phrase in a sentence and the position that it occupies in an underlying syntactic representation (Ben Shachar et al, 2003, 2004), and the combination of a load on short-term memory and a memory system specialized for syntactic processing (Cooke et al., 2001).

Two features of these interpretations of these data may be questioned.

First, these models attribute the neurovascular effects found in these studies to the processes of constructing syntactic structures, inserting lexical items into positions in these structures that allow thematic roles to be assigned, and the assignment of thematic roles. We might say that these models are “syntactic” models, in the sense that, although they attribute some part of the neurovascular response to the assignment of a semantic representation (thematic roles), the particular semantic values that are assigned are not thought to account for the neurovascular response. Put in somewhat different terms, these models attribute the neurovascular responses seen in these sentences to the nature of the syntactic representations that are built and the timing of the assignment of thematic roles to noun phrases in those structures, with later and/or multiple simultaneous assignments of thematic roles leading to increased neurovascular response. They do not take into account the possibility that the nature of the thematic roles that are assigned might affect the neurovascular responses to these sentences. Recent results have raised questions about whether this is the case, or whether the thematic roles that are assigned significantly affect the neurovascular responses to these sentences.

Traxler et al. (2002) studied sentences such as (1) and (2) and found that the animacy of the nouns in these sentences greatly affected eye fixations. SO sentences in which the sentence-initial noun was animate and the subject of the relative clause was inanimate (4) were much harder to process than sentences in which the sentence-initial noun was inanimate and the subject of the relative clause was animate (5). Chen et al (in press) showed a similar effect of noun animacy on neurovascular effects. There was a large increase in BOLD signal in sentences such as (4) compared to sentences such as (5).

(4) The deputy that the newspaper identified chased the mugger (SO - AI) [2]

(5) The wood that the man chopped heated the cabin (SO - IA)

In addition, Chen et al (in press) reported that SO sentences with animate sentence-initial nouns and inanimate subjects of the object-extracted relative clause (4) produced greater BOLD signal than OS sentences with subject extracted relative clauses ((6) and (7)), while the difference in BOLD signal between SO sentences with inanimate sentence-initial nouns and animate subjects of the object-extracted relative clause (5) and OS sentences ((6) and (7)) was much smaller and was found exclusively outside the left perisylvian association cortex where syntactic operations

[2] We use the same terminology as previously, adding the letter A to indicate an animate noun and I to indicate an inanimate noun, with the order AI indicating the first noun in the sentence is animate and the second inanimate and IA indicating the opposite order.

associated with sentence comprehension are widely thought to take place (Grodzinsky, 2000; Ben Shachar et al, 2003, 2004).

(6) The newspaper identified the deputy that chased the mugger (OS - IA)

(7) The man chopped the wood that heated the cabin (OS AI)

In Chen et al (in press), the effect of the order of animacy of nouns on BOLD signal in the comparison of sentences with identical syntactic structure ((4) and (5)) was as great as the effect of syntactic structure ((4) compared to (6) and (7)). The importance of the order of animacy of nouns in determining BOLD signal in sentence comprehension was further documented by Caplan and Chen (in press), who reported widespread increases in BOLD signal for plausibility judgments to sentences such as (6) compared to (7).

Noun animacy is thought to affect the ease of assignment of thematic role by generating expectations regarding the thematic and grammatical roles that a noun phrase will be assigned (Trueswell et al., 1994; Pearlmutter and MacDonald, 1995). Animate nouns are more likely to initiate actions than inanimate nouns, and therefore more likely to be agents of verbs and to occupy the position of grammatical subject than inanimate nouns. The increased difficulty and neurovascular responses seen in the SO sentences that differ in animacy ((4) compared to (5)) and in the OS sentences that differ in animacy ((6) compared to (7)) can be attributed to the fact that the inanimate noun "the newspaper" is less-preferred as the subject and agent of the verb "identified" than the animate noun "man" is of the verb "chopped."[3] Semantic factors such as noun animacy that make for naturalness of thematic role assignment (which we shall call "thematic role fit") thus influence BOLD signal responses to sentences, and appear to interact with syntactic features such as object- compared to subject-extraction in relative clauses in provoking

[3] "Newspaper" may not be an agent in (4) and (6); we use the term simply to indicate that "newspaper" is the perpetrator of the act of "identifying" in these sentences.

neurovascular responses. Many of the effects attributed to structure alone in previous studies may thus be due to these interactions.

An effect of syntax on neurovascular responses that is not affected by these semantic factors appears to have been documented in studies that have reported increased neurovascular responses to object-extracted relative clauses in sentences in which the nouns are all animate (e.g., Just et al., 1996). However, here the second issue that affects the interpretation of previous studies arises.

Previous studies attribute the neurovascular responses found in association with object- and subject-extracted sentences to processes that arise during comprehension of the sentences. That is, these responses have been interpreted as arising during the assignment of structure and meaning when a sentence is first encountered, and not during the performance of the task that participants engage in. We note that "comprehension" is itself a complex process that may be performed in several stages, as discussed below, but we are here considering all such possible stages as part of one process that yields the propositional meaning of a sentence, to be distinguished from processes that use that meaning to accomplish a task.

One reason that it seems reasonable to attribute sentence type effects to comprehension rather than other aspects of a task is that the tasks that have been used in neurovascular studies, such as making plausibility judgments or verifying whether a probe matches the meaning of a target sentence, typically require participants to use the meaning of the sentences that have been presented and do not require reference to the form of a sentence that has been understood. The simplest model of how these tasks are performed would seem to be a two-step model in which sentences are initially assigned their structure and meaning (i.e., are comprehended) and then that meaning is used to undertake the task. If this is the way these tasks are performed, the effect of syntactic structure can only arise during the comprehension process since only propositional meaning, not syntactic form, is used during the performance of the task itself. Some support for this view comes from experimental results that show that the syntactic form of a sentence is hard to retrieve soon after a sentence is understood (Bransford and Franks, 1971), suggesting that syntactic differences between stimulus sentences would not influence neurovascular activity during the performance of a task. However, the two-step model may be wrong. Sentence form may be available for short periods of time after a sentence is understood, and participants may refer back to the syntactic structure of presented sentences, or partially reconstruct those structures, while they undertake the tasks in these studies. If so, neurovascular effects of sentence types may arise in association with both the comprehension of the sentence and the performance of the tasks that participants undertake.

Results of studies using the verification task provide evidence that this does in fact occur. Caplan et al. (2005a, b) and Caplan et al. (under review) reported that BOLD signal increases in object- compared to subject-extracted relative clauses arose solely in association with processing and responding to the probe and not in association with the presentation and processing of the target in a verification task. Stanczak et al. (2005) compared complex sentences with embedded object-extracted relative clauses with passivized verbs to simpler sentences with embedded sentential complements and reported increases in BOLD signal in association with both the presentation of the target sentences and the probes. These studies thus show that syntactic structure can affect neurovascular responses to a task that is performed after a sentence has been understood.

Whether such effects arose in association with sentence comprehension or task performance in other studies that have studied object- and subject-extracted structures in which all nouns are animate is often not clear. In studies using the verification task, Just et al (1996) and

Fiebach et al (2002) did not separate BOLD signal responses to targets from those to probes, and it is not clear what stage of processing was responsible for BOLD signal sentence type effects. Ben Shachar et al. (2004) did report BOLD signal effects in association with target presentation for the comparison of the combination object- and subject-extracted wh-questions compared to embedded yes/no questions; however, separate contrasts of object- and subject-extracted wh-questions individually against embedded yes/no questions and against each other were not reported.[4]

Results in other tasks that have been used to examine the effects of the contrast between object- and subject-extracted structures also suggest that sentence type effects may arise in connection with the performance of a task. Cooke et al (2001) found increased BOLD signal in the contrast of SO and SS sentences when subjects viewed sentences presented one word at a time and made judgments regarding the gender of the agent of the embedded verb. Participants' RTs for determining the gender of the agent of the relative clause in the criticial "long object-extracted" sentences were over 2.5 seconds, making it likely that the sentence type effects partially reflected the maintenance of sentence form in a short-term memory system. Indefrey et al (2001) had participants correct grammatical errors in sentences with subject and object extracted relative

[4] Whether there are syntactic effects in verification tasks, and whether such effects arise in association with the target or the probe, appears to be affected by many factors and may be subject to strategic influences. One factor may be how often probes are presented. Just et al (1996), who reported bilateral activation of inferior frontal and inferior parietal regions in the comparison of object- and subject-extracted relative clauses, presented probes after each target sentence, while Fiebach et al (2002), who did not find such effects, presented probes only on rare occasions. Participants may have strategically encoded sentence meaning into memory less frequently in the Fiebach et al (2002) study than in the Just et al (1996) study; if so, the difference in BOLD signal results in the two studies suggests that syntactic effects occur during encoding of sentence meaning, not comprehension. How frequently semantic values in the portion of the sentence whose structure is varied may also affect neurovascular responses. Another factor is the complexity of the target. Caplan et al. (2005a, b) and Caplan et al. (under review) presented SO and OS sentences. Participants may have encoded these stimuli superficially on initial exposure (e.g., they may have tried to remember the first few words in serial order) and worked backwards from the probe to (re)construct sentence form and extract meaning when verification was required. This would have lead to syntactic effects at the time of probe but not target presentation. In the Stanczak et al. (2005) study, the sentences were more complex and it may have been beyond participants' abilities to retain enough of the target for such a strategy to be useable.

clauses in which all nouns and verbs were replaced with nonsense words (so-called Jabberwocky).[5] These authors did not find BOLD signal differences between these two structures. The absence of an effect may have been related to the unusual nature and difficulty of the task and the stimuli, which may have lead to ceiling effects.

All told, the existing literature does not clearly establish the existence and location of an increase in neurovascular activity associated with the construction of object- compared to subject-extracted structures and the assignment of thematic roles in those structures, that is not influenced by semantic and pragmatic factors, during sentence comprehension. The present study re-explores the determinants of neurovascular responses to object- and subject-extracted relative clauses.

We tested a group of subjects in two tasks: plausibility judgment and detection of nonwords. In each task, we presented object- and subject-extracted relative clauses in which we varied the animacy order of nouns to examine the effect of thematic role fit, as discussed above. The two tasks were used because they differ with respect to the level of sentence processing they require.

Plausibility judgment requires that a sentence be comprehended and its propositional content be matched against propositions in semantic memory. Given the materials used in these studies, comprehending the sentences requires assigning syntactic structure. However, any neurovascular effects of sentence type in this task may not arise solely as a result of the comprehension process. Just as syntactic form may be re-activated when the propositional

---

[5] The use of Jabberwacky has been advocated as a way to identify neurovascular and neurophysiological effects of syntactic processing that are not influenced by semantic variables (Indefrey et al., 2001), but may be inappropriate for this purpose. Readers impute thematic roles in Jabberwacky, which is why works such as Jabberwacky achieve their effects. The use of Jabberwacky leaves the effect of thematic fit and other semantic factors totally in the hands of the study participants, and thus makes it impossible to know what part of a neurological response is due to syntactic factors, semantic factors, or their interaction. The use of stimuli in which all lexical items are real but propositional content is anomalous (e.g., Colourless green ideas sleep furiously; Roder et al, 2002), sometimes called "syntactic prose," also leads to uncontrolled effects of thematic fit and does not constitute an appropriate means of isolating syntactic effects on measures of neural function.

meaning of a target sentence is matched to the meaning of a probe in the verification task, syntactic form may be activated when propositional meaning is matched to representations in semantic memory in plausibility judgment (see General Discussion).

Non-word detection does not require comprehension of propositional meaning or construction of syntactic form, but there is strong behavioral evidence that both are activated during the performance of this task (Stormwold et al, 1996). Effects of syntax and/or propositional meaning that occur in non-word detection mostly likely arise as a result of processes that are obligatorily activated when readers or listeners recognize words that happen to form sentences (Fodor, 1982). Therefore neurovascular effects of these variables offer a fairly direct window on the neural structures that support the initial, largely unconscious, obligatory construction of these types of representations in comprehension.

Converging data from the two tasks would increase confidence regarding the location of the areas of the brain in which syntactic and related thematic operations take place during comprehension. Differences in neurovascular responses across the two tasks would point to the location of brain regions that reflect the use of syntactic information to accomplish task demands.

Methods

Participants:

Fifteen participants (11 female, 4 male; mean age 21.8 years, range 19-26; all college undergraduates) took part in the research. The study was conducted with the approval of the Human Research Committee at the Massachusetts General Hospital and informed consent was obtained for all participants. All participants were right-handed, native speakers of English and naïve as to the purposes of the study. Participants were paid for their involvement.

Experiment 1: Plausibility Judgment

Materials

The experimental items consisted of 144 pairs of SO and OS sentences (see Table 1). Sentences were based on scenarios, each of which appeared once as an SO and once as an OS sentence. Each matching pair of sentences had identical lexical items. All noun phrases were singular, common, and definite to ensure that subjects would not be influenced by the referential assumptions made by the noun phrases in a sentence in different ways in the two conditions. Half the items of each structure in each condition were plausible and the other half were implausible. The violations in plausibility were the result of mismatches in the animacy required by the matrix verb and the animacy of the subject or object noun of the matrix clause. The animacy of subject and object noun phrases and the plausibility of the sentences was systematically varied within each sentence type. Sentences became implausible at various points in the relative clauses and the main clauses, to ensure that subjects had to read each sentence in its entirety before they could decide if it was plausible. Overall, the point at which SO sentences became implausible was earlier than the point at which OS sentences became implausible, reducing the possibility that subjects could decide on strategic grounds that an OS sentence was plausible at an earlier point than was possible in a SO sentence.

Table 1 here

Procedures:

Stimulus Presentation

Each stimulus sentence item was visually displayed in its entirety on a single line in the center of the screen. A given experimental trial consisted of a brief 300 ms fixation cross (a centered "+"), a 100 msec blank screen, the sentence presented for 5sec, and a final 600 ms blank screen, for a total trial length of 6 sec. The task for the participants during the experimental trial was to read the sentence and judge the plausibility of the presented item as quickly and accurately as possible. A plausible sentence was described as a sentence that had a meaning the participant could imagine happening in the real world.

Fixation trials of 0, 2, 4, 6, or 8 seconds were randomly interspersed between each 6 sec sentence trial. Stimuli were presented in a pseudo-randomized order that was determined by a computer program developed to achieve optimum efficiency in the deconvolution and estimation of the hemodynamic response (Burock et al., 1998; Dale, 1999; Dale and Buckner, 1997).

The 288 stimulus items interspersed with the fixation trials were divided into 6 runs. No pair of matched SO and OS sentences were presented in the same run. The sentences were projected to the back of the scanner using a Sharp LCD projector and were viewed by the participants as a reflection in a mirror attached to the head coil. Responses were recorded via a custom-designed, magnet compatible button box. A Dell Inspiron 4000 computer running proprietary experiment set-up software (Stimpres) was used to both present the stimuli and record response accuracy and reaction time.

MR Imaging Parameters.

Participants were scanned in three separate sessions on separate days. In a structural session, two sets of high-resolution anatomical images were acquired in a 3T whole-body Siemens

Sonata scanner (Siemens Medical Systems, Iselin, NJ) using a T1-weighted MP-RAGE sequence (TR = 2530 ms, TE = 3.0 ms, TI = 1100 msec, and flip angle = 7°). Volumes consisted of 128 sagittal slices with an effective thickness of 1.33mm. The in-plane resolution was 1.0 mm x 1.0 mm (256 x 256 matrix, 256 mm Field of View (FOV)).

Participants were scanned in three functional sessions, one for each experiment,[6] with the order of experiments randomized across participants. The functional sessions utilized a 3.0T head-only Siemens Allegra scanner (Siemens Medical Systems, Iselin, NJ). The functional volume acquisitions utilized a T2*-weighted gradient-echo pulse sequence (TR = 2000 msec, TE = 30 ms, and flip angle = 90°). The volume was comprised of 30 transverse slices aligned along the same AC-PC plane as the registration volume. The interleaved slices were effectively 3mm thick with a distance of 0.9mm between slices. The in-plane resolution was 3.13 x 3.13 mm (64 x 64 matrix, 200 mm FOV). Each run consisted of 200 such volume acquisitions for a total of 6000 images. By definition, the 30 slices of a single volume took the entire TR (2s) to be fully acquired and a new volume was initiated every TR. An initial 8 second (4 TR equivalent) buffer of RF pulse activations, during which no stimulus items were presented and no functional volumes were acquired, was employed to ensure maximal signal during the length of the functional run.

Cortical Surface Reconstruction.

The high-resolution anatomical MP-RAGE scans were used to construct a model of each participant's cortical surface. An average of the two structural scans was used to maximize the signal to noise ratio. The cortical reconstruction procedure involved: (1) segmentation of the cortical white matter; (2) tessellation of the estimated border between gray and white matter, providing a geometrical representation for the cortical surface of each participant; and (3) inflation

[6] The same participants were tested in the verification task reported in Caplan et al (under review).

of the folded surface tessellation to unfold cortical sulci, allowing visualization of cortical activation in both the gyri and sulci simultaneously (Dale et al., 1999; Fischl et al., 1999a, 2001).

For purposes of inter-subject averaging, the reconstructed surface for each participant was morphed onto an average spherical representation. This procedure optimally aligns sulcal and gyral features across participants, while minimizing metric distortion, and establishes a spherical-based co-ordinate system onto which the selective averages and variances of each participant's functional data can be resampled (Fischl et al., 1999a, 1999b).

Functional Pre-processing.

Pre-processing and statistical analysis of the functional MRI data was performed using the FreeSurfer Functional Analysis Stream (FS-FAST) developed at the Martinos Center, Charlestown, MA (Burock & Dale, 2000). For each participant, the acquired native functional volumes were first corrected for potential motion of the participant using the AFNI algorithm (Cox, 1996). Next, the functional volumes were spatially smoothed using a 3-D Gaussian filter with a full-width half-max (FWHM) of 6mm. Global intensity variations across runs and participants were removed by rescaling all voxels and time points of each run such that the mean in-brain intensity was fixed at an arbitrary value of 1000.

The functional images for each participant were analyzed with a General Linear Model (GLM) using a finite impulse response model (FIR) of the event-related hemodynamic response (Burock and Dale, 2000). The FIR gives an estimate of the hemodynamic response average at each TR within a pre-stimulus window. The FIR does not make any assumption about the shape of the hemodynamic response. Mean offset and linear trend regressors were included to remove low-frequency drift. The autocorrelation function of the residual error, averaged across all brain voxels, was used to compute a global whitening filter in order to account for the intrinsic serial autocorrelation in fMRI noise. The GLM parameter estimates and residual error variances of each participant's functional data were resampled onto his or her inflated cortical surface and into the spherical coordinate system using the surface transforms described above. Each participant's data were then smoothed on the surface tessellation using an iterative nearest-neighbor averaging procedure equivalent to applying a two-dimensional Gaussian smoothing kernel with a FWHM of approximately 8.5 mm. Because this smoothing procedure was restricted to the cortical surface, averaging data across sulci or outside gray matter was avoided.

Voxel-wise Analysis (or Statistical Activation Maps).

Contrasts of interest were tested at each voxel on the spherical surface across the group using a random effects model of the cross-participant variance of the FIR parameter estimates. Contrasts were constructed over a window of post-stimulus delays in the FIR model corresponding to the delays at which vascular responses were expected to be peaking. BOLD signal changes follow electrophysiological events associated with elementary sensory stimuli and simple motor functions by as little as 2 seconds, with an established response by 4-6 seconds (Bandettini, 1993; Turner, 1997). The hemodynamic response was collapsed across the post stimulus delay intervals from 4 to 10 sec, covering the peak of the BOLD signal.

Group statistical activation maps were constructed for contrasts of interest using a *t* statistic. To correct for multiple comparisons, we identified significant clusters of activated voxels on the basis of a Monte Carlo simulation (Doherty et al, 2004). A volume of Gaussian distributed numbers was generated for each subject, and was processed in the same manner as the real data, including volumetric smoothing, resampling onto the sphere, smoothing on the spherical surface, random effects analysis, and activation map generation. A clustering program was run on these maps to extract clusters of voxels whose members each exceeded a specified voxel-level *p* value threshold and whose area was equal to or greater than a specified size. This process was repeated 3500 times, allowing us to compute the likelihood of one or more clusters of a given size and voxel-level threshold occurring under the null hypothesis. The real data was then subjected to the same clustering procedure as applied to the simulated data using a cluster size threshold of 200 mm^2 and threshold for rejection of the null hypothesis at $p < .05$. These functional activations are displayed on a map of the average folding patterns of the cortical surface, derived using the surface-based morphing procedure (Fischl et al., 1999a, 1999b). The accompanying Talairach coordinates correspond to the vertices within each cluster with the minimum local p-value.

Results

Behavioral Results

The behavioral data are displayed in Figure 1. Accuracy and RT for correct responses were analyzed in 2 (Syntactic Structure: SO, OS) X 2 (Response: Plausible; Implausible) ANOVAs by subjects (F_1) and items (F_2). RTs, trimmed for outliers $\pm$ 3 sd from the condition mean for each subject, were analyzed for sentences with correct responses.

Figure 1 here

There was a significant effect of structure in accuracy ($F_1(1, 14) = 14.8$, $p < .001$; $F_2(1, 71) = 8.2$, $p < .01$) and RTs ($F_1(1, 14) = 216.2$, $p < .001$; $F_2(1, 71) = 38.1$, $p < .001$). Subjects were more accurate and faster in responding to OS than SO sentences. There was a significant effect of response in accuracy in the subject analysis ($F_1(1, 14) = 5.6$, $p < .05$; $F_2(1, 71) = 1.5$, *ns*) and RTs in both analyses ($F_1(1, 14) = 5.6$, $p < .05$; $F_2(1, 70) = 10.1$, $p < .01$). Subjects were more accurate and faster in responding to plausible than to implausible sentences. The interaction was only significant in the subject analysis of the accuracy data (F_1 (1, 14) = 12.0, $p < .01$; F_2 (1, 71) = 3.4, p = .07); for RTs (F_1 (1, 14) = 2.6, *ns*; F_2 (1, 71) = 0.5, *ns*).

We further analyzed the accuracy and RT data in plausible sentences alone for possible effects of order of animacy of nouns in 2 (Syntactic Structure: SO/OS) X 2 (Animacy of first two nouns: animate first noun and inanimate second noun "AI"/ inanimate first noun and animate second noun "IA") X 2 (Animacy of sentence-final noun: A/I) ANOVAs.

In the accuracy data there were significant effects of structure (F_1 (1, 14) = 18.2, $p < .001$; F_2 (1, 136) = 38.4, $p < .001$), of animacy order of the first two nouns (F_1 (1, 14) = 11.3, $p < .01$; F_2 (1, 136) = 19.1, $p < .001$), and an interaction of these factors (F_1 (1, 14) = 36.3, $p < .001$; F_2 (1, 136) = 55.8, $p < .001$). For OS sentences, responses were equally accurate for AI and IA orders, and for SO sentences responses were more accurate for IA than for AI orders. There were more errors on SO sentences with inanimate relative clause subjects (SO AI) than on OS sentences with inanimate main clause subjects (OS IA), but accuracy was equal on the other two sentence types (SO IA = OS AI). The effect of animacy of the sentence-final noun was at the level of a trend in the subject analysis and significant in the item analysis ((F_1 (1, 14) = 2.92, $p = .1$; F_2 (1, 136) = 3.9, $p < .05$). Accuracy was lower on sentences with animate sentence-final nouns.

In the RT data there were significant effects of structure (F_1 (1, 14) = 63.8, $p < .001$; F_2 (1, 136) = 41.9, $p < .001$), of animacy order of the first two nouns (F_1 (1, 14) = 20.9, $p < .001$; F_2 (1, 136) = 10.1, $p < .001$) and of animacy of the sentence-final noun (F_1 (1, 14) = 8.8, $p < .05$; F_2 (1, 136) = 12.0, $p < .001$). The order of animacy of the first two nouns interacted with sentence type (F_1 (1, 14) = 49.6, $p < .001$; F_2 (1, 136) = 97.7, $p < .001$). For SO sentences, responses were faster with the IA than with the AI order, and for OS sentences, the RTs were in the opposite direction. RTs were longer for

SO-AI than for OS-IA sentences and equal in SO-IA and OS-AI sentences. RTs were longer for sentences with animate sentence-final nouns.

fMRI Results.

Figure 2 displays the statistical activation maps of the contrasts for plausible sentences that attained statistical significance. Table 2 is a listing of the Talairach coordinate locations of the local minimum p-value within each cluster.

Figure 2 and Table 2 here

We will discuss comparisons of interest in relation to structural and thematic fit factors, as follows (A, B, C, D refer to panels in Figures 2 and 5):

I. Effects of syntactic structure (controlling for thematic fit)

A) Effect of syntactic structure across sentences with preferred thematic roles: SO-IA vs OS-AI

B) Effect of syntactic structure across sentences with unpreferred thematic roles: SO-AI vs OS-IA

II. Effects of thematic fit (controlling for syntactic structure)

C) Effect of thematic fit across sentences with a simple syntactic structure: OS-IA vs OS-AI

B) Effect of thematic fit across sentences with a complex syntactic structure: SO-AI vs SO-IA

The first comparison of syntactic structure -- in thematic-role-preferred sentences ([SO IA] - [OS AI]; Figure 2a) -- showed increases in BOLD signal in the middle frontal (BA 6), middle and inferior temporal (BA 37, 20) occipital (BA 18, 19), and medial parietal (BA 7, 31) regions of the left hemisphere and superior and medial parietal (BA 7), insular, and inferior temporal (BA 8, 19, 36, 37) regions of the right hemisphere.

The second comparison across syntactic structure -- in thematic-role-unpreferred sentences ([SO AI] - [OS IA]; Figure 2b) -- showed extensive increases in BOLD signal in the left hemisphere in the inferior, middle, and superior gyri (BA 44, 45, 47, 6), the inferior and superior parietal regions (BA 40,7), middle and inferior temporal areas (BA 21,19, 37) and the precentral gyrus. In the right hemisphere, the contrast showed increases in BOLD activity in the three frontal gyri (BA 45 and 6).

The first comparison of sentences that differ in thematic-role preference -- in syntactically simple sentences ([OS IA] - [OS AI]; Figure 2c) -- resulted in many areas of activation throughout both hemispheres. In the left hemisphere, these included the inferior, medial and superior frontal gyri (BA 47, 6), the superior, middle and inferior temporal gyri (BA 22, 21, 39, 19, 20) and the precuneus (BA 7). In the right hemisphere, they included the inferior and superior frontal gyri (BA 45, 47, 6), the superior, middle and inferior temporal gyri (BA 22, 21, 36, 37, 19), the superior parietal lobe (BA 7), the cingulate (BA 23, 31), and the precuneus (BA 7, 31).

The second comparison of sentences that differ in thematic-role preference -- in syntactically complex sentences ([SO AI] - [SO IA]; Figure 2d)—also yielded BOLD signal increases in a large number of areas: the left inferior, middle and superior frontal gyri (BA 44, 45, 6, 9, 10), the inferior parietal gyrus (BA 40), the middle and inferior temporal gyri (BA 21, 37), the occipital gyrus (BA 18) and the right inferior, middle and superior frontal gyri (BA 44, 6, 9) and the insula.

Discussion of experiment 1

The behavioral data showed effects of syntactic structure and animacy order, and an interaction of these factors. These results replicate those previously documented in accuracy and RTs in plausibility judgment (Chen et al, in press) and parallel those in eye fixation measures in a verification task (Traxler et al., 2002). Our data also show longer RTs for OS sentences in which the first noun is inanimate and the second noun is animate compared to OS sentences with the opposite animacy order, consistent with the results in Caplan and Chen (in press). The importance of the animacy factor is reinforced by the previously unreported finding that accuracy was lower and RTs longer for sentences with animate sentence-final nouns, which are less likely to be recipients of actions, and hence objects of verbs, than inanimate nouns.

The neurovascular data are broadly consistent with these behavioral findings. BOLD signal was increased for SO compared to OS sentences, markedly so for the comparison of the thematically unpreferred [SO AI] and [OS IA] sentences. There was a strong effect of animacy order in the syntactically simple OS sentences ([OS IA] > [OS AI]). The BOLD signal responses are relevant to the neural basis for syntactic processing and the interaction of syntactic and semantic processing. The interpretation of the neurovascular data in relation to the processes that underlie the task depends upon what model of those processes is adopted. We will begin by considering models of the comprehension process, and then turn to models of the use of syntactic and thematic information in the judgment operation.

As we indicated above, though we have been considering sentence comprehension as a single process for purposes of contrasting it with performing a task on the basis of what is understood, sentence comprehension is a complicated cognitive function. We cannot discuss all models of sentence comprehension here, but will discuss the neurovascular results in relationship to two main types of theories of the comprehension process.

The first type are two-stage models that maintain that syntactic structure and thematic role fit are computed separately and interact only in a second stage of processing (e.g., Frazier, 1989; Frazier and Clifton, 1996). Traxler et al. (2002) developed a model along these lines to account for their eye fixation results. They found that initial eye fixation duration in the relative clause did not differ for [SO AI] and [SO IA] sentences and was longer on both than in SS sentences. Subsequent regressions to the first noun, later fixations on the relative clause and fixations on the main verb were longer for the [SO AI] sentences than for any other sentence type, and did not differ for the other sentence types. Traxler et al. (2002) concluded that the first operations subjects undertook when they encountered the relative clause of an SO sentence were uninfluenced by the animacy of the nouns. These operations are therefore syntactic and could consist of inserting words into grammatical positions in which thematic roles are assigned. These operations demanding for reasons outlined in the Introduction to this paper. Upon experiencing difficulty in the initial structuring the object-extracted clause in SO sentences, participants return to the sentence-initial noun. At this point and going forward, syntax and semantics interact; we may consider subsequent operations to occur at a second stage of processing. In the [SO AI] sentence, the animacy of the sentence-initial noun biases against the object-relativized structure, leading to increased processing load at the second stage of processing. In the [SO IA] sentence, the animacy of this noun supports the object extracted relative structure, making it easy for subjects to structure and interpret the sentence. The fact that all fixation measures after first-pass eye fixations on the relative clause were the same in [SO IA], [OS AI] and [OS IA] sentences indicates that second stage processing is equally (un)demanding in these sentence types.

Relating the neurovascular effects to this model, the difference in BOLD signal in thematic-role-preferred sentences that differ in syntactic structure ([SO IA] - [OS AI]) provides evidence regarding the neural loci of the syntactic operations that underlie the initial difficulty associated with constructing object- compared to subject-extracted relative clauses. An analysis of the areas that overlap in the two SO/OS comparisons -- ([SO IA] - [OS AI]) and ([SO AI] - [OS IA]) -- is also relevant to the localization of this process. Differences between [OS IA] and [OS AI] sentences, in which processes associated with assigning syntactic structure are simple in both cases, presents the most direct view of the areas of the brain associated with establishing thematic role fit. Differences between [SO AI] and [SO IA] sentences provide evidence regarding the loci of the interaction of these factors in second-stage processing.

The first of these contrasts -- the comparison of thematic-role-preferred sentences that differ in syntactic structure, ([SO IA] - [OS AI]) -- yielded multiple areas of activity, consistent with the view that sentences with the more complex object-extracted syntactic structure require more neural activity to structure and understand. However, the location of the neurovascular activity is surprising. None of the areas activated in this contrast lie in the left perisylvian cortical regions widely thought to be associated with syntactic operations.

The areas that overlap in the comparisons of ([SO AI] - [OS IA]; Figure 2a) and ([SO IA] - [OS AI]; Figure 2b) are also relevant to the localization of the area responsible for initial processing of object relative clause. On Traxler et al.’s (2002) model, both subtractions would be expected to lead to increased BOLD activity associated with the difficulty of initial processing of the object-extracted structure, and the second is also associated with revision and integration processes. Figure 3 shows the areas activated in each of these two comparisons alone and those activated in the intersection of the comparisons. As can be seen, the area of intersection is much smaller than that seen in the ([SO IA] - [OS AI]; Figure 1b) comparison alone. Necessarily, the areas activated in intersection of the two subtractions do not lie in the left perisylvian cortex, since those found in the first subtraction do not fall in this area.

Figure 3 here

Though not predicted by many models of the regional functional neuroanatomy of syntactic processing, the finding of increased BOLD signal outside the perisylvian cortex in a comparison of object- and subject-extracted relative clauses has been reported in some previous studies. Using the same materials as those used here, Caplan et al. (1998, 1999, 2000) reported extra-perisylvian as well as left perisylvian activity in the comparison of all SO and OS sentences, and Chen et al (in press) found only non-left-perisylvian activation for the contrast of [SO IA] sentences against all OS sentences. Cooke et al (2001) also found only extra-perisylvian activation (bilaterally in the inferior temporal gyri) in a comparison of SO and SS sentences without "padding" in their gender monitoring task. These results present challenges to the view that operations involved in structuring object-extracted relative clauses are only localized in Broca's area and other left perisylvian regions (e.g., Ben Shachar et al, 2003, 2004).

The different areas in which BOLD signal increases are found in association with object-extracted relatives may be responsible for different computations and memory loads. As briefly discussed in the introduction to this paper, several memory systems and computational operations are thought to be activated to a greater extent when an object-extracted structure is assigned and interpreted (see also Traxler et al., 2002, for other operations, related to the use of heuristics, that make for difficulty in object-extracted relative clauses). Which of these proposed operations is carried out in which area cannot be determined from these or any other neuroimaging data currently available, and remains a subject for further detailed research.

If the logic underlying the interpretation of the overlap analysis is correct, the BOLD signal associated with the ([SO IA] - [OS AI]) contrast should be a proper subset of that associated with the ([SO AI] - [OS IA]) contrast. However, that is not the case: there are areas of increased BOLD signal in the ([SO IA] - [OS AI]) contrast that are not found in the ([SO AI] - [OS IA]) contrast. The most likely candidate for the operations that generate BOLD signal in the ([SO IA] - [OS AI]) and not in the ([SO AI] - [OS IA]) subtraction is the process of assigning thematic roles in the main clause. The sentence-initial inanimate noun is conducive to the assignment of structure and thematic roles in the object-extracted relative clause in the [SO IA] sentence, but it is not conducive to assigning thematic roles around the main verb. Much of the BOLD signal increase in the ([SO IA] - [OS AI]) comparison does not occur in the ([SO AI] - [OS IA]) subtraction, suggesting that much of the BOLD signal increase in the ([SO IA] - [OS AI]) comparison is due to the goodness of fit of thematic roles around the main verb.

Turning to the assignment of thematic roles on the basis of thematic fit, the comparison of sentences with simple syntactic structure that differ in thematic role fit -- [OS IA] - [OS AI]; Figure 2c-- yielded BOLD signal increases in a large number of structures in both hemispheres. This result, which replicates that reported by Caplan and Chen (in press), again provides evidence that thematic role fit greatly affects the neural activity associated with sentence processing.

The extent of the BOLD signal effect contrasts with the sparse behavioral evidence for an effect of goodness of thematic role fit in simple syntactic structures. Traxler et al. (2002), for instance, did not find differences between eye fixations in [SS IA] and [SS AI] sentences. It is possible that the difference between the results is due to the location of the relative clause. A sentence-initial inanimate noun might be hard to assign as the "agent" of the verb of an OS sentence, which it immediately precedes, and less difficult to assign as the "agent" of the verb of an SS sentence, from which it is separated by a relative clause. This possibility is consistent with the observations that the area of BOLD signal increase in the ([OS IA] - [OS AI]) subtraction (Figure 2c) is considerably more extensive than the area of increased BOLD signal that is found in the ([SO IA] - [OS AI]; Figure 2a) and not in the ([SO AI] - [OS IA]; Figure 2b) subtraction, which we suggested above reflects the thematic role fit of the sentence-initial noun and the verb of the main clause in the SO structure, and that the areas activated in the ([SO IA] - [OS AI]; Figure 2a) and not in the ([SO AI] - [OS IA]; Figure 2b) subtraction are all found in the ([OS IA] - [OS AI]; Figure 2 c) subtraction.

Although these data strongly suggest that the nature of subject noun phrases affects the ease of sentence comprehension, exactly what features of that noun phrase do so remains to be explored. Animacy may be just one feature that makes a noun a likely agent. Thus, sentences with inanimate subject nouns like "car," which are reasonably likely to move, might be easier to process than sentences with inanimate subject nouns like "chair," which are less likely to do so. Alternatively, the factor that affects processing may be less general than animacy: sentences may be easier to process when their subjects are unlikely agents of their particular verbs, not unlikely subjects in general (see Traxler et al, 2002, experiment 2). Given the robustness of the neurovascular effects of the animacy order variable in these simple sentences, such measures may help distinguish between these finer-grained theories.

Moving on to second stage integration processes, we suggested above that the comparison - of [SO AI] to [SO IA] sentences (Figure 2b) is relevant to the neural loci of these interactions. Following Traxler et al.'s (2002) model of how SO sentences are processed, outlined above, participants experience initial difficulty structuring the object-extracted clause in both these sentences, and return to the sentence-initial noun. In the [SO AI] sentence, the animacy of this noun biases against the object-relativized structure, and in the [SO IA] sentence it supports this structure. Thus more work is required at the second stage of processing to integrate the semantic and syntactic representations in the [SO AI] than in the [SO IA] sentences. This comparison yielded extensive BOLD signal activation, as noted in the Results section above.[7]

Turning to the second major framework for modeling sentence comprehension, this view of sentence comprehension maintains that syntactic structure is never created without an influence of semantics (e.g., MacDonald et al., 1994). In such models, although information about the plausibility or likelihood of thematic roles is accessed independently, there are no first-stage operations that build syntactic structures independently. Rather, every pairing of syntactic structure and thematic fit leads to an interaction of these features. The BOLD signal responses reflect the complexity of these interactions.

The results of this study are partially consistent with these models. BOLD signal increased to the greatest extent in [SO AI] sentences, in which syntactic structure was complex and thematic

---

[7] The comparison of [SO AI] and [OS AI] sentences also might seem relevant to this contrast. However, this contrast is complex. Structuring [SO AI] sentences is demanding at both the putative initial and second stages of processing; in contrast, [OS AI] sentences are easily interpretable at the initial stage and require little integration effort. The difference between [SO AI] and [OS AI] sentences thus does not isolate second stage processing, but involves increased effort at both stages. To eliminate the effect of the initial syntactic analysis of [SO AI] sentences from the contrast of [SO AI] and [OS AI] sentences, one would need to subtract the difference between [SO IA] and [OS AI] from that between [SO AI] and [OS AI] sentences; that is, one must compute the subtraction of subtractions ([SO AI] - [OS AI]) – ([SO IA] - [OS AI]). When this is done, the two measures of second stage processing that we are considering are algebraically equivalent: ([SO AI] - [OS AI]) – ([SO IA] - [OS AI]) = [SO AI] - [OS AI] – [SO IA] + [OS AI] = [SO AI] – [SO IA]. It is thus uninformative that the results of this complex subtraction were identical to those discussed above.

role fit was unpreferred, consistent with these theories. The results also provide data that could constrain such models. For instance, BOLD signal increased much more in the comparison of sentences that differed in thematic role fit and had simple syntactic structure ([OS IA] - [OS AI]; Figure 2c) than in sentences that differed in syntactic structure and had preferred thematic role fit ([SO IA] - [OS AI]; Figure 2a). This suggests that it is more difficult to integrate thematically unpreferred information with a simple syntactic structure than it is to integrate thematically preferred information with a complex syntactic structure. This could be modeled by interactive models.

To this point, we have taken the BOLD signal effects as reflections of the comprehension process. As noted above, it is also possible that they arise during the process of judging whether the proposition extracted from a sentence is plausible. The differences in BOLD signal associated with thematic role fit can easily be related to the plausibility judgment process, since propositions with better thematic fits are likely to be more easily recognized as being plausible.[8] The syntactic effects and the effects of the combination of syntactic and semantic factors may also arise during the plausibility judgment process. BOLD signal effects may be associated with operations of the sort discussed above in reference to the comprehension process, only now repeated (at least partially) when the plausibility judgment is made, differences in extent to which different sentence types are rehearsed, or to other mechanisms that arise in connection with the judgment phase of the task. More discussion of aspects of the plausibility judgment operation that might involve syntactic structures is found in the General Discussion.

In summary, Experiment 1 found robust effects of thematic role fit and of the combination of syntactic complexity and unpreferred thematic role fits upon BOLD signal responses to

[8] Evidence that this is the case would come from the analysis of implausible sentences, which should be harder to recognize as being implausible if they have poor thematic role fits. We cannot test this prediction in our data, because the point of implausibility was deliberately varied to reduce strategic effects in making positive judgments.

sentence comprehension in a plausibility judgment task. Effects of syntactic complexity alone were seen entirely outside the left perisylvian cortex. These effects can be related to one quite detailed two-stage model of sentence comprehension and are not inconsistent with interactive activation, constraint-satisfaction, models of this process, though they cannot be related to such models in as much detail as to two-stage models. The results could also reflect a sentence type effect on BOLD signal associated with the judgment aspects of the plausibility judgment task. Some of these possible loci of these effects will be partially clarified by the results of Experiment 2.

Experiment 2: Non-word Detection

Plausibility judgment requires a participant to consciously decide whether a sentence is plausible, and, given the materials used in experiment 1, participants must have assigned syntactic structure to extract sentence meaning in that experiment. Non-word detection does not require either of these operations. Effects of syntactic structure and/or noun animacy order in the non-word detection task would show that participants did not restrict their cognitive processes to those minimally required in that task but rather constructed syntactic representations and assigned meaning to some extent. Effects of these variables in non-word detection are thus evidence for some degree of implicit, obligatory processing of syntax and sentence-level meanings in this task, and BOLD correlates of these variables point to those brain areas involved in these processes in an unconscious, obligatory fashion.

Materials

The experimental items consisted of the 144 pairs of SO and OS sentences from the plausibility judgment task with changes to the implausible sentences to render them plausible. Half the pairs were altered to include a phonologically and orthographically legal non-word, as illustrated in Table 3. Non-words were located in positions occupied by both nouns and verbs in both the relative and the main clauses.

Table 3 here

Methods

The stimulus presentation, fMRI methods and data analysis were identical to those in Experiment 1 except that subjects were told to detect the presence of a non-word as accurately and quickly as possible.

Results

Behavioral Results

The behavioral data are displayed in Figure 4. Accuracy and RT for correct responses were analyzed in 2 (Syntactic Structure: SO, OS) X 2 (Response: All real words; non-word) ANOVAs by subjects (F_1) and items (F_2).

Figure 4 here

There was a trend towards an effect of structure in accuracy ($F_1(1, 14) = 3.3$, $p = .09$; $F_2(1, 71) = 2.6$, $p < .1$). The effect in RTs was significant by subjects ($F_1(1, 14) = 9.6$, $p < .01$) and at the level of a trend by items ($F_2(1, 71) = 2.7$, $p = .1$). Subjects were more accurate and faster in responding to OS than SO sentences. There was an effect of response in accuracy ($F_1(1, 14) = 11.1$, $p < .01$; $F_2(1, 71) = 15.7$, $p < .001$) and in RTs ($F_1(1, 14) = 7.0$, $p < .01$; $F_2(1, 71) = 17.3$, $p <$

.001). Subjects were more accurate at indicating that a sentence did not contain a non-word than at detecting a non-word, but took longer to respond to sentences without non-words. The interaction was only significant in RTs in the subject analysis RTs (F_1 (1, 14) = 6.2, $p < .05$; F_2 (1, 71) = 1.6, *ns*); for accuracy, (F_1 (1, 14) = 0.1, *ns*; F_2 (1, 71) = 0.0, *ns*).

We further analyzed the accuracy and RT data in sentences with real words only for possible effects of order of animacy of nouns in 2 (Syntactic Structure: SO/OS) X 2 (Animacy of first two nouns: animate first noun and inanimate second noun "AI"/ inanimate first noun and animate second noun "IA") X 2 (Animacy of sentence-final noun: A/I) ANOVAs.

There were no effects in the accuracy data. In the RT data, there was an effect of structure (F_1 (1, 14) = 20.0, $p < .001$; F_2 (1, 108) = 4.3, $p < .05$) and an interaction of structure with order of animacy of the first two nouns (F_1 (1, 14) = 35.0, p < .001; F_2 (1, 108) = 6.2, $p < .01$). For SO sentences, responses were significantly faster with the IA than with the AI order, and for OS sentences, the RT effect was significant in the opposite direction. RTs were significantly longer both for SO IA than for OS AI sentences and for SO AI than for OS IA sentences.

Differences in performances between the two experiments were examined in 2 (Task) X 2 (Sentence Structure: SO/OS) X 2 (Animacy of first two nouns: animate first noun and inanimate second noun "AI"/ inanimate first noun and animate second noun "IA") X 2 (Animacy of sentence-final noun: A/I) ANOVAs for accuracy and RTs. The three-way interaction of task X structure X animacy order of the first two nouns was significant both for accuracy (F_1 (1, 14) = 4.7, p < .05; F_2 (1, 37) = 5.2, $p < .05$) and RTs (F_1 (1, 14) = 8.0, p < .05; F_2 (1, 37) = 12.7, $p < .01$), confirming the different effects of the sentence variables in the two tasks. In addition, this analysis showed that accuracy was higher and RTs were faster in non-word detection than in plausibility judgment.

fMRI Results

Figure 5 displays the statistical maps of the contrasts for plausible sentences that attained statistical sgnificance. Table 4 is a listing of the Talairach coordinate locations of the local minimum p-value within each cluster.

Figure 5 and Table 4 here

Comparisons across syntactic structures controlled for thematic role fit showed significant BOLD signal in the left inferior frontal lobe in the comparison of [SO AI] and [OS IA] sentences, and no areas of increased BOLD signal in the comparison of [SO IA] and [OS AI] sentences.

The effect of noun animacy order on BOLD signal when syntactic structure was held constant was seen in a very small area of increased BOLD signal -- less than 200 mm^2 and therefore not listed in Table 4 -- in the left posterior inferior temporal lobe at the junction with the occipital lobe in the subtraction of [OS IA] minus [OS AI] sentences. There was no difference between [SO AI] and [SO IA] sentences.

Discussion of Experiment 2

Beginning our discussion with the behavioral results, the fact that accuracy was higher and responses were longer to sentences without non-words would ordinarily suggest a speed-accuracy trade-off, but in this case may simply indicate that participants read the entirety of sentences without non-words and stopped reading as soon as they detected a non-word. If the speed/RT pattern does not reflect a speed-accuracy trade-off, the accuracy data indicate either that the non-words were hard to detect, which would be a desirable consequence of how the stimuli were constructed, and/or that participants had a response bias. The interpretation of syntactic and thematic fit effects as due to obligatory implicit processing would not change if either a speed-accuracy trade-off or a response bias were present.

The fact that there were effects of syntactic structure and animacy order and an interaction of animacy order and structure indicates that participants processed the sentences beyond the lexical level despite the fact that dependencies between words are irrelevant to the task. This suggests that these levels of representation are subject to obligatory processing; that is, that a reader cannot help but construct at least a partial syntactic and sentential representation when s/he attends to a sequence of words that is in fact a sentence (Fodor, 1982). From the perspective of theories of selective attention, either readers did not selectively attend to lexical identity and set an attentional filter "shallowly," that is, before syntactic and semantic representations were activated (Broadbent, 1958), or syntactic and semantic representations were activated in an unattended channel.

The effects of syntax and semantics might arise from a subset of the stimuli; that is, participants might have been capable of selectively attending to the words and non-words without constructing syntactic and semantic representations for many stimuli and the behavioral and BOLD signal effects might be due to the smaller number of trials on which they failed to selectively attend. This is rendered less likely by the fact that histograms revealed unimodal, fairly symmetrical, somewhat platykurtic distributions of the RTs to sentences containing real words. Excluding the first and last deciles, which together contained less than 5% of responses, for OS sentences, kurtosis = -1.3, skew = - 0.05; for SO sentences, kurtosis = -.5, skew = - 0.45. These features of the distributions suggest that the processing of most stimuli was similar, and subject to noise. In addition, it is likely that subjects inspected all the words of the sentences with all real words, since the high accuracy rates could not have been achieved had they ignored the sentence-final words, which were loci of non-word insertions in some stimuli.

Accuracy was higher and RTs faster in non-word detection than in plausibility judgment. One might think that it is easier to detect a non-word than to make a plausibility judgment, but the relative difficulty of the tasks is likely to depend upon the complexity of the searches and matches that are needed in semantic memory in plausibility judgment and in the lexicon in the case of non-word detection. The relative difficulty of the tasks may well differ depending upon these features. For instance, making plausibility judgments about propositions such as "dogs like bones" might be easier than making judgments about whether sentences with items such as “squill,” “skib,” “squit,” and “skeg” do or do not contain non-words.[9]

The BOLD signal effects seen in the sentence contrasts in the non-word detection task reflect processes associated with assigning sentence structure and meaning. Of these, only the second stage integration of syntactic and thematic information (on the two stage model) produced reliable BOLD signal effects in Experiment 2. These effects were found in the left inferior frontal region, strongly implicating this region in these integrative processes. The posterior part of the left inferior temporal gyrus was marginally activated in the contrast of thematic-role-unpreferred and thematic-role-preferred syntactically simple sentences, suggesting that the use of thematic fit information in sentence comprehension involves this region.

General Discussion

The present study provides data that are relevant to the neural basis of aspects of syntactic processing in sentence comprehension. The two most important results presented here that bear on

[9] “Squill” and “skeg” are both words according to Webster's dictionary.

this problem are: 1) that neurovascular responses differed considerably in different tasks; and 2) that thematic role fit greatly affected these responses to sentences.

The difference in neurovascular responses in the two experiments can be explained in two ways. Neurovascular effects of sentence differences could arise at different stages of processing in the tasks and result in more BOLD signal when they arise at one stage than when they arise at another. Alternatively, they could arise at the same stage of processing in the two tasks but that stage of processing could be less engaged in one task than another. Both of these possibilities require that the stages of processing in these tasks be identified. For this purpose, we will initially continue to consider that the plausibility judgment task requires sentence comprehension and matching of propositional content against semantic memory in order to make the plausibility judgment. We will assume that sentence-level effects in non-word judgment arise at the stage of comprehension; that is, that the sentences were in some sense "understood" in the non-word detection task but not matched to semantic memory, encoded into semantic memory, or used in any subsequent cognitive operation.

Within this framework, we first consider the second of the possibilities listed above: that processing of the sentences in the non-word detection task at the comprehension level may have been "shallow" in some sense. This may have happened even if such processing was obligatory and occurred for (almost) all the sentences, as we suggested was the case.

One possibility is that there is a separate mode of unconscious, obligatory processing that leads to activation of sentence-level representations that occurred in the non-word detection task, which differs from the type of processing of these representations that occurs when a task requires overt judgments about meanings (or other uses of sentence meaning). However, the idea that the BOLD signal sentence type effects differed in non-word detection and plausibility judgment because the comprehension process itself is in some way less "deep," or differs in some aspects of its processing from the comprehension process in the plausibility judgment task leads to a number of problems. It fractionates the unconscious, obligatory sentence comprehension process into at least two "modes" of processing, one that yields one type of output, that may be "shallow" but nonetheless includes thematic and syntactic structure, and a second type of "deeper" output with the same types of representations. This seemingly simple claim thus adds at least one major theoretical construct to models of sentence processing. It raises the difficult question of what the shallow output is, if it is not the same as the one that forms the basis for comprehension. Moreover, it contradicts the simplest and most widely adopted view of sentence comprehension, namely that it is unconscious and obligatory in all tasks (i.e., that it is necessarily triggered when a language user attends to a sequence of words that is a sentence; Fodor, 1982). On the usual view, the process of comprehension only becomes deliberate and conscious when sentences are beyond the capacity of the normal parser/interpreter. In short, the idea that the computation of syntactic structure and sentence meaning differs in plausibility judgment and non-word detection leads to significant complications of models of how sentence comprehension takes place.

A modification of this idea, however, seems to us to be more promising. The modified claim is based on the fact that attentional processes affect the neural response to computing representations. Neurons increase their firing rates to effective stimuli in the focus of attention compared to stimuli that are outside the focus of attention (Moran and Desimone, 1985; see Hillyard et al (1973) for a similar ERP effect in humans). Attention increases the gain without affecting the tuning curves for some such cells (Treue and Martinez Trujillo, 1999), indicating that, for these cells, the representations that determine whether a cell will fire do not change as a function of attention. If sentence-level thematic and syntactic representations fall within the focus of attention in plausibility judgment and not in non-word detection, an effect of attention of this sort could lead to greater neural responses to the same representations in the plausibility judgment than in the non-word detection task. This view maintains that the effects of thematic fit and syntactic structure in both tasks are the result of the same processes, which occur in attended or unattended channels.

The second account of the differences in sentence type effects in the two tasks – that they arise at different stages of processing in the tasks -- does not raise the problem of introducing a new entity into the theory of sentence processing, but it faces other, perhaps equally serious, challenges. If sentence-level variables produce equivalent BOLD signal effects at the stage of comprehension in both non-word detection and plausibility judgment, the greater BOLD signal sentence-type effects in plausibility judgment must arise at the stage of checking propositional content against semantic memory for purposes of making plausibility judgments, a possibility that we considered in the discussion of Experiment 1. The idea that they arise at this level may seem counter-intuitive, and we shall explore this possibility in two ways: first, by reviewing some of the data in this study that suggest it is reasonable, and, second, by briefly outlining a plausible model of the comprehension and judgment processes that calls for such effects.

Two features of the neurovascular results provide reason to believe that sentence-level effects arise at this judgment stage of processing. First, there were extensive neurovascular responses to the noun animacy order variable. These were seen in the contrast of ([OS IA] - [OS AI]) sentences, and also in the areas of BOLD signal found in the areas activated solely in the ([SO IA] - [OS AI]) contrast and not the ([SO AI] - [OS IA]) sentences. As mentioned in the discussion of Experiment 1, these effects are easily attributed to the judgment process, since they affect the likelihood that a proposition is plausible. Second, the BOLD signal effects of these variables contrast with the eye tracking results of Traxler et al. (2002), who, the reader will recall, found that, except for a brief initial fixation on the relative clause in SO sentences, the three sentence types [SO IA], [SS AI] and [SS IA] did not differ. The difference between the extent of the animacy-order-related BOLD signal effects in the plausibility judgment task and the paucity of effects of these variables in the eye tracking data is explained if the eye fixation data reflect processes that occur during the comprehension of these sentences and the BOLD signal effects arise during the process of making judgments about the plausibility of comprehended propositional representations.

These observations do not provide a basis for suggesting that syntactic effects, or the combination of syntactic and thematic fit effects, occur in judgment process and, as noted above, the claim that they do seems counter-intuitive. However, we suggest this is only the case if one thinks of the comprehension and judgment processes as occurring sequentially. If we think of them as occurring in an interactive cascade system, the intuitions are reversed, and it becomes almost impossible to imagine that syntactic effects are *not* to be found in the "judgment" part of the plausibility judgment task.

To see that this is the case, consider again the evidence from Traxler et al. (2002), Trueswell et al (1994), Pearlmutter and MacDonald (1995), and many other studies that shows that readers and listeners incorporate plausibility information into the comprehension process itself in an incremental fashion. (The point just made about the Traxler et al. results only shows that this process is not very demanding when this integration involves thematically preferred and syntactically simple representations, not that this does not occur.) This demonstrates that the comprehension process produces the very representations that are necessary for plausibility judgment to take place; namely, measures of the plausibility of various possible thematic roles that a presented sentence may express and of the meaning of the entire proposition that it expresses. Let us for expository purposes imagine that this system runs the thematic and propositional representations derived from the input through a neural net and computes an "error" score that is a combination of multiple factors (such as the extent to which they differ from some aggregate measure of those already in the net, the extent to which they differ from the closest match already in the net, etc.), with larger numbers corresponding to poorer matches and thus less plausible semantic values. The difference between making an overt (or covert) plausibility judgment and comprehension is then simply that, in the plausibility task, these products of the comprehension processes are fed into a system that is responsible for response selection on the basis of these values, and that this does not happen when processing stops with comprehension. If the intermediate products of comprehension are transferred to the response selection mechanism in an incremental fashion, they can be used incrementally for two parallel purposes in the plausibility judgment task – once to influence ongoing comprehension and once to weight the ultimate response -- and only once when the task is restricted to comprehension. Since these intermediate products of the comprehension process consist of integrated thematic and syntactic representations, both these information types will affect the plausibility judgment operation.

This analysis depends crucially on the assumption that plausibility judgment makes incremental use of thematic representations constructed in the course of comprehension. The considerations above argue that this is quite reasonable, as the relevant computations are already available to the comprehension system. In addition, there is very strong evidence that performance of other tasks and comprehension are cascaded in this fashion. Visual world tracking studies have shown that features of spatial layout interact incrementally with information derived from sentence comprehension during the performance an enactment task (Tanenhaus et al, 1995; Sedivy et al, 1999). The same is very likely to be true of all tasks.

We note that nothing in this model contradicts our previous statement that the unconscious, obligatory, "automatic" processing associated with comprehension does not differ in non-word detection and plausibility judgment. The claim here is that comprehension affects task performance, not *vice versa*. It remains possible that the ongoing construction of syntactic structure and thematic representations is uninfluenced by the concurrent use of the products of these computations to accomplish other tasks. As far as we can see, the same is true of the findings in the visual world tracking studies; these establish that the nature of the non-linguistic environment affects the use of information derived from a presented sentence to accomplish a task, but not that the way that information is derived differs as a function of the non-linguistic environment. It may, of course, be the case that, if computation of syntactic structure and thematic representations as part of comprehension and the use of these representations to accomplish a task go on concurrently, the latter also influences the former. This also does not entail that <u>some</u> aspects of sentence comprehension are not independent of task demands. How operations such as consolidation of representations in memory, retrieval of encoded information, logical and associative processes, decision making based on information in semantic memory, and others, interact with sentence comprehension remains a subject of study.

In short, the judgment phase of the plausibility task is likely to be ongoing throughout the presentation and comprehension of the sentence and to make use of the syntactic and propositional representations that are constructed by the comprehension apparatus. If so, the matching portion of the plausibility judgment taskwould be affected by many if not all of the syntactic and thematic operations that occur in comprehension, and the increased sentence-type-related BOLD signal effects seen in plausibility judgment could be due to the concurrent use of these representations in that aspect of the task. The many areas of activation found in many sentence contrasts in experiment 1 may thus reflect the effects of sentential variables on making plausibility judgments, not on comprehending the meaning of the presented sentences.

Turning to the second major finding in these studies, the fact that thematic fit greatly affected BOLD signal, our discussion can be brief. This finding shows that effects of syntactic variables are not easy to disentangle from those of semantic influences. This in turn raises questions about what drove neurovascular responses in virtually all studies that have been published regarding syntactic processing. We have touched on some of the issues that arise in existing studies: the collapsing across thematic role fit in previous studies using PET in our lab (Caplan et al , 1998, 1999, 2000, 2002; Stromswold et al, 1996; Waters et al, 2003); the use of proper names by Cooke et al (2001); the lack of control over thematic fit in studies using Jabberwacky and syntactic prose (Indefrey et al, 2001). All of these studies confounded syntactic and thematic processes, and the neurovascular effects in these studies may have been due to a particular combination of factors that was present in a study, not to a syntactic operation alone. Systematic control of these variables is needed to explore both how comprehension occurs and where it takes place in the brain.

Where does this lead us in regard to the neural localization for syntactic operations in sentence comprehension? The results of the present studies suggest that the neural locus of some integrated, possibly second stage, processing of syntactic and thematic representations is a part of the left inferior frontal lobe. Other neurovascular effects are seen in connection with many other brain regions, but these may reflect task-related processes that are closely related to, but nonetheless separate from, comprehension. Looking more generally at the literature, the presence of both task effects and effects of thematic variables in the present study raises questions regarding what determined neurovascular responses to sentence contrasts that have thus far been taken to reflect syntactic operations. The possibility that what have been interpreted as syntactic effects could have arisen at previously unexpected points of processing in tasks such as verification, plausibility judgment, and gender monitoring, and that they may reflect processing related to only one combination of thematic and syntactic representations, needs to be explored. The neural basis of much of syntactic comprehension remains to be studied.

Acknowledgements

This research was supported by a grant from the NIDCD (DC02146). This work was also supported in part by the National Center for Research Resources (P41RR14075) and the Mental Illness and Neuroscience Discovery (MIND) Institute. We thank Daphne Robakis and Doug Greve for assistance with the data collection and analysis.

References

Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993): Processing strategies for time-course data sets in functional MRI of the human brain. Magnetic Resonance in Medicine 30: 161-173.

Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, Grodzinsky Y (2003): The neural reality of syntactic transoformations: Evidence from fMRI. Psych Sci 14: 433-440.

Ben-Shachar M, Palti D, Grodzinsky Y (2004): The neural correlates of syntactic movement: converging evidence from two fMRI experiments. Neuroimage 21: 1320-1336.

Bransford JD, Franks JJ (1971): Abstraction of linguistic ideas. Cogn Psych 2: 331-350.

Broadbent, D. (1958) Perception and Communication, Pergamon: London

Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM (1998): Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. Neuroreport 9: 3735-3739.

Burock MA & Dale AM (2000): Estimation and detection of event-related fMRI signals with temporally correlated noise: A statistically efficient and unbiased approach. Human Brain Map 11: 249-260.

Caplan D, Chen E (in press): Using fMRI to discover cognitive operations. Cortex.

Caplan D, Alpert N, Waters G (1998): Effects of syntactic structure and propositional number of patterns of regional cerebral blood flow. J Cogn Neurosci 10: 541-552.

Caplan D, Alpert N, Waters G (1999): PET studies of sentence processing with auditory sentence presentation. Neuroimage 9: 343-351.

Caplan D, Alpert N, Waters G, Olivieri A (2000): Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Human Brain Map 9: 65-71.

Caplan D, Vijayan S, Kuperberg G, West C, Waters G, Greve D, Dale AM (2002): Vascular responses to syntactic processing: event-related fMRI study of relative clauses. Human Brain Map 15: 26-38.

Caplan D, Waters G, Alpert N (2003): Effects of age and speed of processing on rCBF correlates of syntactic processing in sentence comprehension. Human Brain Map 19: 112-131.

Caplan D, Chen E, Waters, G, Robakis D (2005a) Task effects on the localization of BOLD signal in syntactic processing, CUNY Sentence Processing Conference, Tucson, AZ

Caplan D, Chen E, Waters, G, Robakis D (2005b) The Effect of Task on the BOLD Signal Localization of Syntactic Processing, Cognitive Neuroscience Society, NY, NY

Caplan, D., Chen, E., Waters, G. (under review) Syntactic Effects on BOLD Signal Responses to Comprehension and Verification of Sentence Meaning

Caviness VS, Meyer J, Makris N, Kennedy DN (1996): MRI-based topographic parcellation of human neocortex: An anatomically specified method with estimate of reliability. J Cogn Neurosci 8: 566-587.

Chen E, West WC, Waters G, Caplan D (in press): Determinants of BOLD signal correlates of processing object-extracted relative clauses. Cortex.

Cooke A, Zurif EB, DeVita C, Alsop D, Koenig P, Detre J, Gee J, Pinãngo M, Balogh J, Grossman M (2001): Neural basis for sentence comprehension: grammatical and short-term memory components. Human Brain Map 15: 80-94.

Cox RW (1996): AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers & Biomedical Research 29: 162-173.

Dale AM (1999): Optimal experimental design for event-related fMRI. Human Brain Map 8: 109-114.

Dale AM & Buckner RL (1997): Selective averaging of rapidly presented individual trials using fMRI. Human Brain Map 5: 329-340.

Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis – I: Segmentation and surface reconstruction. Neuroimage 9:179-194.

Doherty CP, West WC, Dilley LC, Shattuck-Hufnagel S, Caplan D (2004). Question/statement judgments: An fMRI study of Intonation Processing. Human Brain Mapping, 23, 85-98.

Fiebach CJ, Schlesewsky M, Friederici AD (2001): Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. J Psycholing Research 30: 321-338.

Fiebach CJ, Schlesewsky M, Friederici AD (2002): Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. J Mem Lang 47: 250-272.

Fischl B, Sereno MI, Dale AM (1999a): Cortical surface-based analysis – II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9: 195-207.

Fischl B, Sereno MI, Tootell RBH, Dale AM (1999b): High-resolution intersubject averaging and a coordinate system for the cortical surface. Human Brain Map 8: 272-284.

Fischl, B., Liu, A., & Dale, A. M. (2001): Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans Med Imaging 20(1): 70-80.

Frazier, L. and C. Clifton (1996). Construal. Cambridge, MA, MIT Press.

Fodor (1982). The Modularity of Mind, MIT Press, Cambridge, MA

Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. Cognition, 68, 1-76.

Grodzinsky, Y (2000): The neurology of syntax: Language use without Broca'a area, Behavioral and Brain Sciences, 23, 47 - 117.

Hillyard, S.A., Hink, R.F., Shwent, V.L., Picton, T.W. (1973): Electrical signs of selective attention in the human brain, Science, 182, 177-180

Indefrey P, Hagoort P, Herzog H, Seitz RJ, Brown CM (2001): Syntactic processing in left prefrontal cortex is independent of lexical meaning. Neuroimage 14:546 –555

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99(1), 122-149.

Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR (1996): Brain activation modulated by sentence comprehension. Science 274: 114-116.

King, J.W., & Just, M.A. (1991). Individual difference in syntactic processing: The role of working memory. Journal of Memory and Language, 30, 580-602.

MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. Psychological Review, 101, 676-703.

Moran, J. and Desimone, R. (1985) Selective attention gates visual processing in extrastriate cortex, Science, 229, 782 - 784

Pearlmutter, N.J. & MacDonald, M.C. (1995). Individual differences and probabalistic constraints in syntactic ambiguity resolution. Journal of Memory and Language, 34, 521-542

Röder B, Stock O, Neville H, Siegfried B, Rösler F (2002): Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance imaging study. Neuroimage 15: 1003-1014.

Sedivy, J.C., Tanenhaus, M.K., Chambers, C.G., Carlson, G. N. (1999) Achieving incremental semantic interpretation through contextual representation , Cognition, 71, 109-147

Stanczak, L., Caplan, D., Waters, G., (2005) Similar and different patterns of activation in fast and slow sentence processors: An fMRI study, Cognitive Neuroscience Society, NY

Stromswold K, Caplan D, Alpert N, Rauch S (1996): Localization of syntactic comprehension by positron emission tomography. Brain Lang 52: 452-473.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M. & Sedivy, J.E. (l995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632-1634

Traxler M, Morris R, Seely R (2002): Processing Subject and Object Relative Clauses: Evidence from Eye Movements. Journal Mem Lang 47: 69-90.

Treue S. and Martinez Trujillo J.C (1999) Feature-based attention influences motion processing gain in macaque visual cortex, Nature, 399, 575-579

Trueswell JC, Tanenhaus MK (1994): Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. J Mem Lang 33: 285-318.

Turner R (1997): Signal sources in BOLD contrast fMRI. Adv Exp Med Biol 413: 19-25.

Waters G, Caplan D, Alpert N, Stanczak L (2003): Individual differences in rCBF correlates of syntactic processing in sentence comprehension: Effects of working memory and speed of processing. Neuroimage 19: 101-112.

Table 1: Examples of Sentence Types in Plausibility Judgment Task

Plausible sentences

SO-AI	The deputy that the newspaper identified chased the mugger
SO-IA	The wood that the man chopped heated the cabin
OS-IA	The newspaper identified deputy that the chased the mugger
OS-AI	The man chopped wood that the heated the cabin

Implausible sentences

SO-AI	The plumber that the hair extracted clogged the sink
SO-IA	The bill that the activist angered organized the march
OS-IA	The hair extracted the plumber that clogged the sink
OS-AI	The activist angered the bill that organized the march

Table 2: Talairach coordinates of peak activation, Brodman area location, and cluster size corresponding to the local minimum p-values for each cluster of activated vertices for comparison of sentences types in Plausibility Judgment. Cluster # for each contrast corresponds directly to cluster labels on the corresponding contrasts in Figure 2.

	Left Hemisphere						Right Hemisphere					
Contrast	**Cluster #**	**Region**	**BA**	**Talairach (x,y,z)**	**Size (mm2)**	**p-value**	**Cluster #**	**Region**	**BA**	**Talairach (x,y,z)**	**Size (mm2)**	**p-value**
A. SO-IA vs. OS-AI (Plausible Only)	1	Middle Frontal	6	(-42, 3, 43)	345	0.000013	8	Superior Parietal	7	(22, -50, 51)	1301	0.000009
	2	Middle Temporal	37	(-63, -55, -2)	576	0.000081	9	Insula		(25, 17, -6)	350	0.000033
	3	Occipital	19	(-55, -78, 7)	240	0.003155	10	Fusiform Gyrus	36/37	(40, -39, -9)	483	0.000002
	4	Occipital	18	(-29, -83, -5)	2681	0.000018	11	Lingual Gyrus	8	(11, -83, 2)	2823	0.000001
	5	Inferior Temporal	20	(-56, -44, -18)	276	0.000331	12	Pre-cuneus	7	(3, -49, 49)	488	0.000274
	6	Pre-cuneus	7	(-16, -56, 48)	397	0.000005	13	Pre-cuneus	7	(18, -67, 43)	381	0.000566
	7	Cuneus	31	(-17, -64, 16)	237	0.001545	14	Lingual Gyrus	19	(21, -54, 6)	567	0.000536
B. SO-AI vs. OS-IA (Plauisble Only)	2a	Inferior Frontal	44	(-49, 12, 20)	4605	0.000165	9	Middle Frontal	6	(29, 6, 40)	220	0.001208
	2b	Middle Frontal	6	(-38, -1, 50)	-	0.000008	10	Inferior Frontal	45	(35, 34, 14)	1066	0.000389
	2c	Pre- central	6	(-52, 1, 24)	-	0.000139	11	Inferior Frontal	45	(43, 33, 2)	436	0.000818
	2d	Inferior Frontal	44	(-56, 10, 10)	-	0.000004	12	Superior Frontal	6	(0, 19, 43)	363	0.000294
	2e	Superior Frontal	6	(-15, 12, 49)	-	0.000035						
	3	Inferior Parietal	40	(-43, -41, 34)	407	0.002037						
	4	Superior Parietal	7	(-35, -62, 39)	668	0.000130						
	5	Middle Temporal	21	(-56, -46, 4)	1313	0.000002						
	6	Fusiform Gyrus	37	(-49, -51, -19)	1931	0.000055						
	7	Fusiform Gyrus	19	(-32, -64, 4)	298	0.000067						
C. OS-IA vs. OS-AI (Plausible Only)	1	Inferior Frontal	47	(-47, 31, -1)	2108	0.000034	10	Occipital	19	(35, -68, 35)	477	0.000759
	2	Middle Frontal	6	(-43, 3, 42)	784	0.000010	11	Superior Parietal	7	(23, -53, 52)	3437	0.000008
	3	Middle Temporal	39	(-38, -68, 25)	1342	0.000113	12	Inferior Frontal	45	(43, 25, 16)	1675	0.000013
	4	Superior Temporal	22	(-58, -42, 9)	1317	0.000003	13	Inferior Frontal	47	(36, 26, -6)	812	0.000002
	5	Middle Temporal	21	(-54, -10, -9)	953	0.000001	14	Middle Temporal	21	(51, -15, -8)	888	0.000013
	6	Fusiform Gyrus	20	(-27, -32, -16)	2097	0.000020	15	Superior Temproal	22	(46, -29, 12)	367	0.000180
	7	Pre-cuneus	7	(-17, -48, 48)	610	0.000271	16	Superior Temporal	22	(47, -43, 15)	207	0.000066
	8	Superior Frontal	6	(-14, 14, 55)	401	0.000014	17	Fusiform Gyrus	36/37	(40, -40, -10)	3216	0.00000000001
	9	Lingual Gyrus	19	(-21, -44, -2)	525	0.000371	18	Pre-cuneus	7	(16, -66, 43)	589	0.000124
							19	Cingulate	23/31	(12, -51, 11)	629	0.000027
							20	Pre-cuneus	31	(16, -69, 18)	377	0.000497
D. SO-AI vs. SO-IA (Plauisble Only)	1a	Inferior Frontal	45	(-51, 33, 2)	7320	0.00000004	7a	Inferior Frontal	44/9	(28, 17, 24)	4576	0.000011
	1b	Inferior Frontal	44	(-47, 14, 16)	-	0.000015	7b	Middle Frontal	9	(26, 38, 30)	-	0.000012
	1c	Middle Frontal	9	(-53, 9, 40)	-	0.0000053	7c	Insula		(24, 24, -1)	-	0.0000005
	1d	Middle Frontal	6	(-26, -1, 51)	-	0.0000002	8	Superior Frontal	6	(0, 14, 52)	432	0.000050
	1e	Superior Frontal	6	(-18, 6, 62)	-	0.000001						
	2	Inferior Parietal	40	(-40, -41, 33)	2002	0.000006						
	3	Middle Temporal	21	(-54, -45, 3)	2910	0.0000003						
	4	Fusiform Gyrus	37	(-48, -51, -15)	640	0.000034						
	5	Occipital	18	(-17, -82, 0)	260	0.000910						
	6	Frontal	10	(-16, 63, 10)	221	-0.000369						

Table 3: Examples of Sentence Types in Non-word Detection Task

<u>Sentences with real words</u>

SO-AI	The deputy that the newspaper identified chased the mugger
SO-IA	The wood that the man chopped heated the cabin
OS-IA	The newspaper identified deputy that the chased the mugger
OS-AI	The man chopped wood that the heated the cabin

<u>Sentences with non-words</u>

SO-AI	The deputy that the haberfelt identified chased the mugger
SO-IA	The wood that the man dribed heated the cabin
OS-IA	The newspaper identified deputy that the chorried the mugger
OS-AI	The man chopped wood that the heated the gert

Table 4: Talairach coordinates of peak activation, Brodman area location, and cluster size corresponding to the local minimum p-values for each cluster of activated vertices for comparison of sentences types in Non-word Detection. Cluster # for each contrast corresponds directly to cluster labels on the corresponding contrasts in Figure 5.

	Left Hemisphere						Right Hemisphere					
Contrast	Cluster #	Region	BA	Talairach (x,y,z)	Size (mm2)	p-value	Cluster #	Region	BA	Talairach (x,y,z)	Size (mm2)	p-value
A. SO-IA vs. OS-AI (Real Words Only)				None						None		
B. SO-AI vs. OS-IA (Real Words Only)	1	Inferior Frontal	44	(-44, 6, 15)	453	0.000170				None		
C. OS-IA vs. OS-AI (Real Words Only)				None						None		
A. SO-AI vs. SO-IA (Real Words Only)				None						None		

Figure Captions

Figure 1: RT and accuracy in plausibility judgment

Figure 2: BOLD signal effects in plausibility judgment

Figure 3: BOLD signal effects in plausibility judgment in intersection disjunctions of ([SO IA] – [OS AI]) and ([SO AI] – [OS IA]) subtractions

Figure 4: RT and accuracy in non-word detection

Figure 5: BOLD signal effects in non-word detection

FIGURE 1

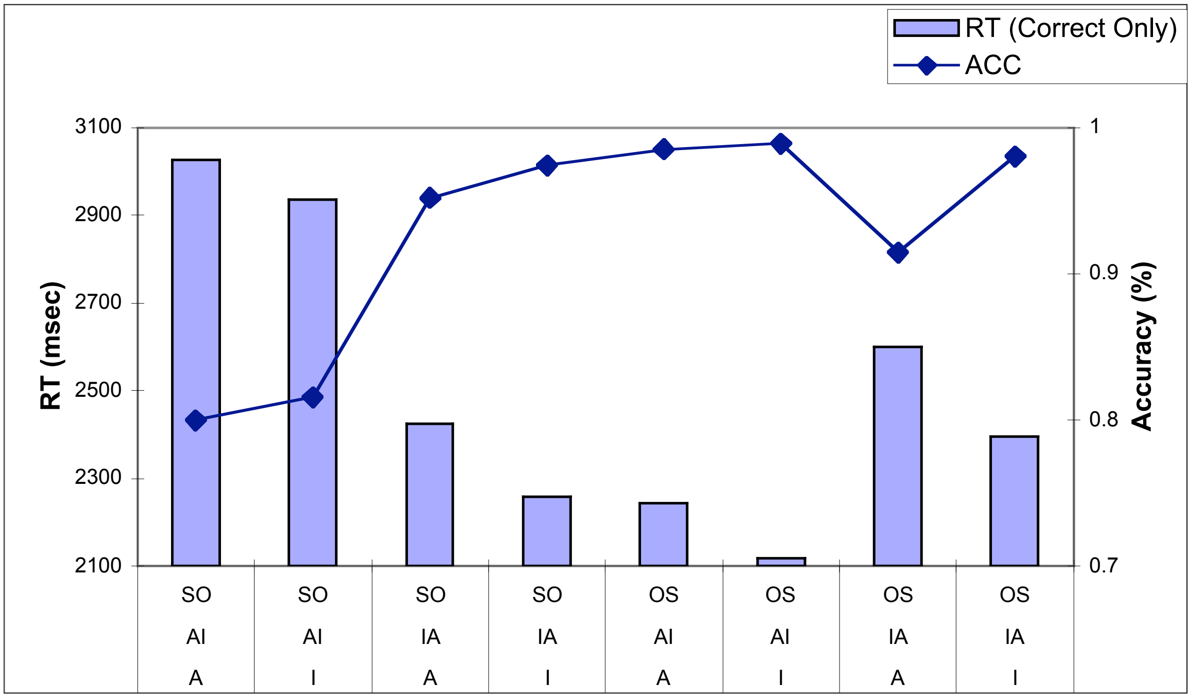
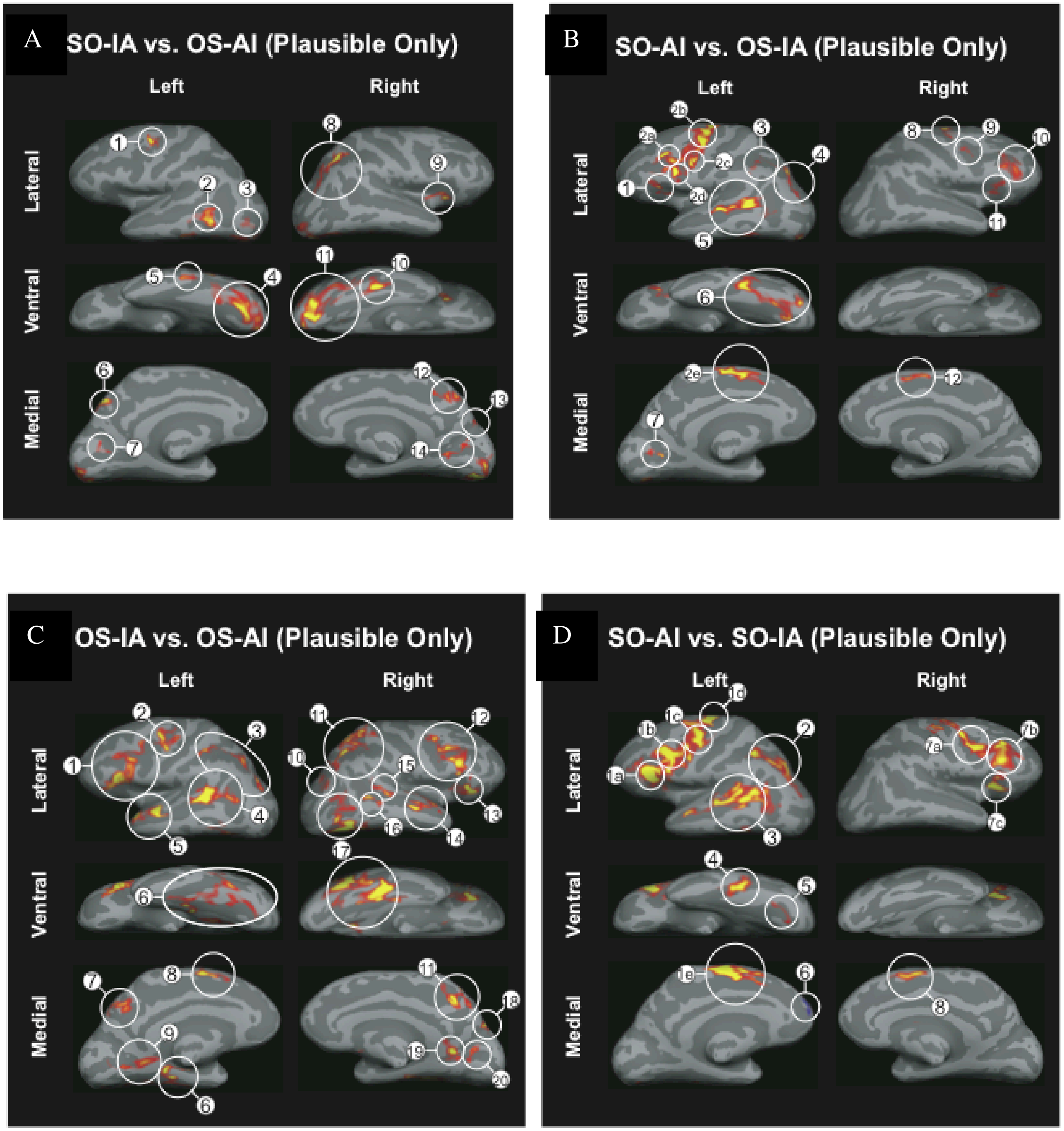
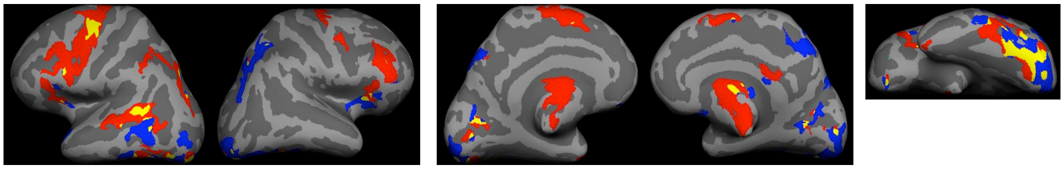

RT (Correct Only)
ACC
RT (msec)
Accuracy (%)
3100
2900
2700
2500
2300
2100
1
0.9
0.8
0.7
SO
AI
A
SO
AI
I
SO
IA
A
SO
IA
I
OS
AI
A
OS
AI
I
OS
IA
A
OS
IA
I

FIGURE 2



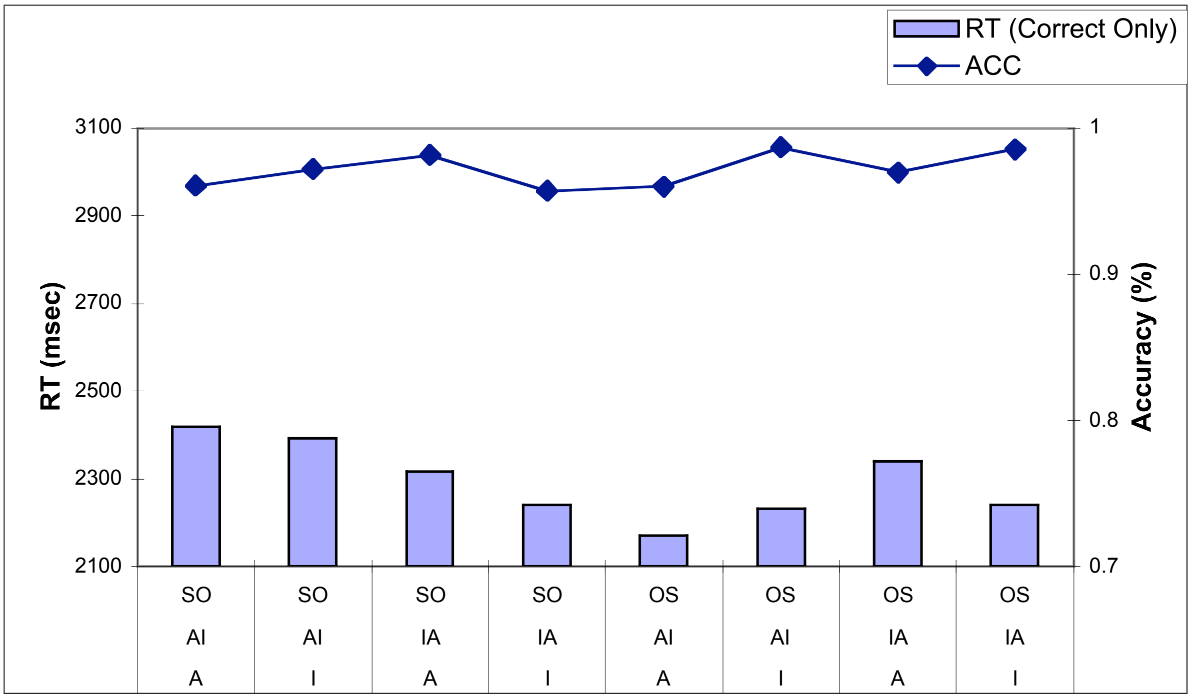
A
SO-IA vs. OS-AI (Plausible Only)
Left
Right
Lateral
Ventral
Medial
B
SO-AI vs. OS-IA (Plausible Only)
Left
Right
Lateral
Ventral
Medial
C
OS-IA vs. OS-AI (Plausible Only)
Left
Right
Lateral
Ventral
Medial
D
SO-AI vs. SO-IA (Plausible Only)
Left
Right
Lateral
Ventral
Medial

FIGURE 3



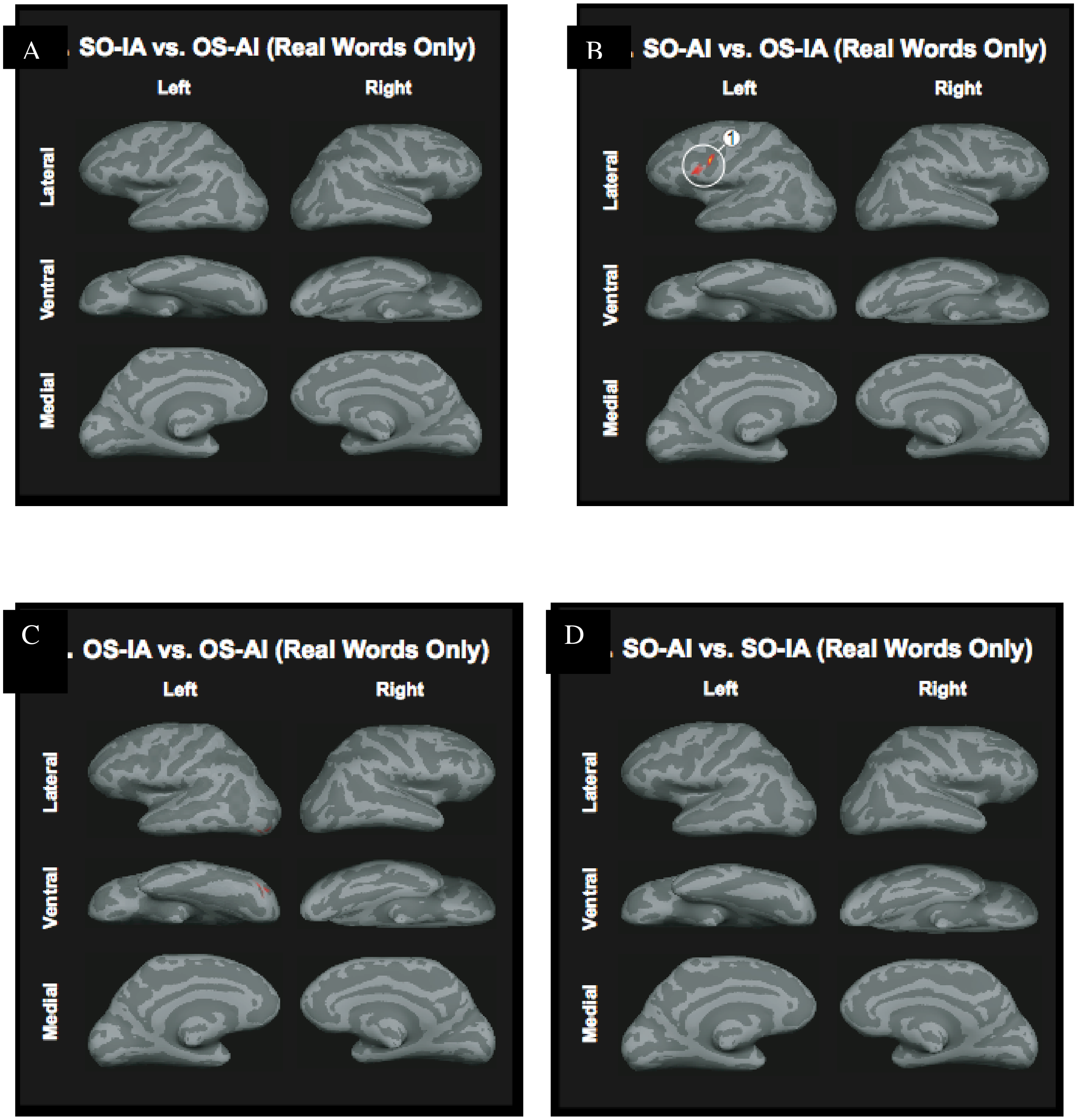
yellow=significant in both contrasts
red=significant contrast 1 only
blue=significant contrast 2 only

FIGURE 4



RT (Correct Only)
ACC
3100
2900
2700
2500
2300
2100
RT (msec)
1
0.9
0.8
0.7
Accuracy (%)
SO AI A
SO AI I
SO IA A
SO IA I
OS AI A
OS AI I
OS IA A
OS IA I

FIGURE 5


A
SO-IA vs. OS-AI (Real Words Only)
Left
Right
Lateral
Ventral
Medial
B
SO-AI vs. OS-IA (Real Words Only)
Left
Right
1
Lateral
Ventral
Medial
C
OS-IA vs. OS-AI (Real Words Only)
Left
Right
Lateral
Ventral
Medial
D
SO-AI vs. SO-IA (Real Words Only)
Left
Right
Lateral
Ventral
Medial