



Published in final edited form as:

Lang Cogn Process. 2010 July ; 25(6): 866–901. doi:10.1080/01690961003672447.

Task effects on BOLD signal correlates of implicit syntactic processing

David Caplan

Neuropsychology Lab, M.G.H

Abstract

BOLD signal was measured in sixteen participants who made timed font change detection judgments in visually presented sentences that varied in syntactic structure and the order of animate and inanimate nouns. Behavioral data indicated that sentences were processed to the level of syntactic structure. BOLD signal increased in visual association areas bilaterally and left supramarginal gyrus in the contrast of sentences with object- and subject-extracted relative clauses without font changes in which the animacy order of the nouns biased against the syntactically determined meaning of the sentence. This result differs from the findings in a non-word detection task (Caplan et al, 2008a), in which the same contrast led to increased BOLD signal in the left inferior frontal gyrus. The difference in areas of activation indicates that the sentences were processed differently in the two tasks. These differences were further explored in an eye tracking study using the materials in the two tasks. Issues pertaining to how parsing and interpretive operations are affected by a task that is being performed, and how this might affect BOLD signal correlates of syntactic contrasts, are discussed.

Functional neurovascular imaging has been applied to the study of the neural basis for parsing and sentence interpretation for over 20 years (see Bornkessel and Schlesewsky, 2006; Grodzinsky and Friederici, 2006; Caplan, 2006, 2007, for recent reviews). Much of this research has been interpreted as providing evidence about the neural basis of parsing and interpretive processes in isolation from other operations that occur during the comprehension process. It is, however, clear from studies using behavioral measures that many cognitive operations, such as scene analysis (Tanenhaus et al., 1995, 2000; Spivey and Grant, 2005; Altman and Kamide, 2007), motor planning (Farmer et al. 2007), and assessment of plausibility (Trueswell et al., 1994; Pearlmutter and MacDonald, 1995; Garnsey et al., 1989), occur at the same time as parsing and interpretation and are influenced by the sentences being processed. This poses a problem for the interpretation of BOLD signal correlates of sentence contrasts: how can neural effects due to parsing and interpretation be distinguished from those due to processes that run concurrently (Page, 2006)?

Results from our lab illustrate the problem (Caplan et al., 2008a). We reported BOLD signal effects of the contrast of the same object-vs-subject extracted sentences, shown in (1) and (2), in two tasks -- a plausibility judgment task, where foils consisted of implausible sentences, and a non-word detection task, where foils consisted of sentences containing an orthographically and phonologically legal non-word. In the non-word detection task, the contrast of object-vs-subject extracted sentences evoked increased BOLD signal only in left

inferior frontal gyrus, BA 44. The same contrast evoked widespread, multifocal BOLD activity in the plausibility judgment task.

1. Object Extracted

The deputy that the newspaper identified chased the mugger

2. Subject Extracted

The newspaper identified the deputy that chased the mugger

These results raise the question above: which parts of the activation in these two studies is due to parsing and interpretation and which is due to other cognitive activities? There are two main possibilities to consider.

One is that the results in the plausibility judgment task reflect parsing and interpretation and the activity in the nonword detection task does not reflect the totality of the BOLD signal effect of these processes. This would occur if participants only engaged in parsing and interpretation in a few of the trials in the non-word detection task. However, multiple analyses of the results indicated that participants processed the majority of sentences syntactically and semantically in that task: in sentences without non-words (sentences (1) and (2)), the interaction of structure with order of animacy of the first two nouns was significant in the analysis by items, which strongly suggests parsing and interpretation occurred in most sentences, and histograms of RTs by deciles showed unimodal distributions for all sentence types, not the bimodal distributions that would be expected if some sentences were inspected superficially for non-words and other processed deeply for meaning. In sentences with non-words, reading times were longer the further into the sentence the non-word occurred, suggesting that visual search was controlled by language-related processes (Reichle et al, 2003), and reading times for object-extracted sentences with non-words in embedded verb position (the most syntactically demanding position in the most syntactically demanding sentence type) were significantly longer than those for subject extracted sentences with non-words in either embedded verb or noun positions (the controls for the demanding positions in the more complex sentences) while reading times for object- and subject-extracted sentences with non-words in other positions did not differ, strongly suggesting that subjects assigned some aspects of structure and/or meaning even to the sentences with non-words. We concluded that the limited number of areas activated was not a reflection of participants' not implicitly processing most sentences syntactically and semantically in this task.

The second interpretation of the discrepancy between the results of the two studies is that the BOLD signal in the non-word detection task reflects parsing and interpretation and some of the BOLD signal in the plausibility judgment task reflects other cognitive activities. This raises the question of what these other cognitive activities are. One possibility is that they consist of assessing the sentences' plausibility and using results of this assessment to choose a response. It might be thought that, because the subject- and object-extracted sentences are synonymous, these operations do not differ in these two sentence types and therefore could not be responsible for BOLD signal effects of the contrast between them. In fact, we considered this to be the case in multiple studies using the plausibility judgment task with these sentences (Stromswold et al, 1996; Caplan et al., 1998, 1999, 2000; Waters et al., 2003; Chen et al., 2006). However, this view ignores the fact that these operation take place incrementally; that is, that assessment of plausibility and response selection occur as soon as they can, not after the sentence is totally processed (Trueswell et al., 1994; Pearlmutter and MacDonald, 1995; Garnsey et al., 1989). Because the number of thematic roles that can be assigned at the embedded verb of an object-extracted relative clause (*identified* in (1)) is greater than the number that can be assigned at any point in a subject-extracted sentence, the

number of thematic roles that are assessed for plausibility and the complexity of the response selection process are also greater at the embedded verb of an object-extracted relative clause than at any point in a subject-extracted sentence. We favored the possibility that these cognitive operations led to some of the BOLD signal effects seen in the plausibility judgment task and not in the nonword detection task.

If this account of these findings is plausible, it highlights the need to eliminate these ancillary processes from tasks that focus on parsing and interpretation *per se*. One approach to the problem of eliminating the effects of ancillary cognitive operations on BOLD signal in studies of parsing and interpretation has been to use tasks in which syntactic processing is “implicit.” In tasks that utilize this approach, participants perform a task that does not require sentence processing, such as cross-model lexical decision (Swinney, 1979) or word monitoring (Marslen Wilson and Tyler, 1980). Non-word detection has these features (to some extent – see below). Effects of syntactic variables on these measures have been taken as evidence for parsing and interpretive operations that are uninfluenced by the use of the products of comprehension to perform a task. Several neurovascular effects of sentence contrasts in these types of tasks have been reported.

Hashimoto and Sakai (2002) compared a “sentence short-term memory” task, in which participants read Japanese sentences presented in rapid serial visual presentation and made judgments about whether the order of two words in a subsequent probe was the same as in the target sentence, to a “baseline STM” task, in which strings of unrelated words were presented, followed by a two word probe in which the same order judgment was made. The comparison of the sentence versus the baseline STM tasks yielded BOLD signal effects in left dorsolateral prefrontal cortex, localized to Brodman areas 6, 8, and 9. The authors argue that the sentence STM task involves implicit syntactic processing. However, as the authors noted, differences in case markings in the two conditions may have been responsible for the BOLD signal effect; in addition, the effects in the “sentence STM” task may be due to maintaining the sentence in memory, and, last, we cannot know whether participants assigned structure or meaning to the sentences in the “sentence STM” task.

Kang et al. (1999) had subjects silently read syntactically and semantically acceptable two word phrases (drove cars) in one of two lists: intermixed with syntactically deviant phrases (forgot made) or with semantically deviant phrases (wrote beers). Relative to the syntactically and semantically acceptable stimuli, activation was found for both the syntactically and semantically deviant stimuli in left and right inferior frontal gyrus (BA 44/45), inferior parietal lobe (supramarginal and angular gyri: BA 40/19) and anterior cingulate (BA 32). Isolated activation for the syntactic violations relative to the syntactically and semantically acceptable stimuli occurred in posterior left MFG (BA 46) and for semantic violations in anterior right MFG (BA 10, BA46). In left BA 44, the activation was greater for the syntactic than for the semantic violations. The authors argued that the BOLD signal correlates of syntactic violations reflect implicit syntactic processing and the correlates of semantic violations reflect implicit semantic processing. However, this interpretation ignores the facts that 1) syntactically anomalous phrases are also semantically anomalous and differences between the “syntactic” violations and either semantically acceptable phrases and or “semantic” violations may reflect the nature of the semantic violations in syntactically anomalous phrases; 2) the differences between processing the “syntactic” and “semantic” violations cannot reveal areas of the brain in which “syntactic” and “semantic” processing occurs, only areas in which each of these types of processes leads to more activity than the other (see Caplan, 2008, for discussion); and 3) BOLD signal in the contrast of anomalous and ill-formed meaningful phrases relative to the syntactically and semantically acceptable stimuli may have been due to the detection of the anomaly, not the assignment of syntactic structure or semantic meaning (see Caplan, 2007, for discussion).

Noppeney and Price (2004) had participants silently read blocks of five nine-word sentences presented with rapid serial visual presentation. Blocks contained sentences either with the same syntactic structure or with a mixture of four syntactic structures. BOLD signal was reduced in left temporal pole for blocks with identical compared to non-identical sentences. The authors interpret this as a result of implicit syntactic priming, and conclude that this area supports parsing and interpretation. However, the BOLD signal reduction may have occurred because priming occurred for superficial features of the stimuli, even such low-level features as the patterns of punctuation marks, not syntactic structures. Aspects of their data are hard to reconcile with the view that syntactic processing took place in this area. If this were the case, one would expect increased BOLD signal effects for syntactically more complex sentences, but this was not the case: ambiguous sentences that were ultimately assigned unpreferred structures (e.g., The child left by his parents played table football) appear to have generated less BOLD signal in this area than their unambiguous counterparts (e.g., The child, left by his parents, played table football).¹

Love et al (2006) had participants perform one of three tasks in association with auditorily presented subject- and object-extracted relative clauses: passive listening (with an end-of-sentence button push), word probe verification (determining whether a word following the sentence appeared in the sentence), and thematic role verification (determining whether the first word in sentence was the theme of a probe word). The passive listening task and the word probe verification task are tasks in which syntactic processing is implicit. The authors focused on Broca's area, and found that the thematic role verification task produced more BOLD signal in this area than the other tasks. However, there were no effects of sentence type and no task by sentence type interaction in either behavioral measures (accuracy) or in BOLD signal in this region or in a whole brain analysis. The authors consider the length of the sentences, the auditory presentation, and the possible role of different portions of Broca's area in supporting operations found in either of the sentence types, as possible reasons for these null effects, but, regardless of what led to them, the absence of a sentence type effect in the BOLD signal make this study uninformative regarding the neural basis for implicit syntactic processing.

As noted above, our study using nonword detection mentioned above (Caplan et al, 2008a) may also be considered one in which parsing and interpretation were implicit, since responses do not require sentence-level processing. However, although nonword detection does not *require* assigning the structure and meaning of a sentence to select a response, the task *can* be accomplished on the basis of the semantic meaning of the sentence: participants can decide that a stimulus contains a non-word when the propositional meaning of the sentence is incomplete. It is therefore not a perfectly designed task if the goal is to study the BOLD signal effects of implicit parsing and interpretation. What is required for this purpose to ensure that a task *cannot* be performed on the basis of a syntactic structure and its interpretation.

Font change detection is a task that has these properties. Unlike non-word detection, in which the presence of a non-word can be inferred from the fact that the meaning of a sentence is incomplete, if only well-formed, semantically plausible sentences are presented

¹Noppeney and Price (2004) attempted to verify that participants encoded and remembered the sentences by testing them in an unannounced recognition task in which 29 sentences from the stimulus set used in the scanner and 23 new sentences (10 with syntactic structures similar to those presented in the fMRI environment) were presented. Participants were above chance at discriminating old and new sentences, even when only new sentences with structures similar to those in the fMRI scanner were analyzed, but this does not show that there were syntactic priming effects during the presentation of the sentences in the scanner. The occurrence of syntactic priming might be inferred from the finding that recognition differed for sentences presented at the end of blocks of identical and different sentences, since sentences whose structure was primed might be encoded and retained differently than sentences whose structure was not primed, but the small number of sentences used in the recognition task makes it impossible to determine whether this was the case.

in a font change detection task, the font in which a word occurs cannot be inferred from the structure or meaning of the sentence. Font change detection is thus a useful task with which to replicate and extend the results of the non-word detection task. We here report the use of a font change detection task to examine BOLD signal effects of syntactic processing.

Methods

We conducted both a behavioral study using eye tracking and an fMRI study. The eye tracking study was done after the fMRI study (because we did not have an eye tracker at the time the fMRI study was run). However, for ease of exposition, we first present the materials (which were used in both the behavioral and fMRI studies), then the eye tracking study, and finally the fMRI study.

Materials

We used the syntactically and semantically acceptable sentences that were used in the Caplan et al (2008a) study in a font change detection task. The experimental items consisted of 144 pairs of SO and OS sentences with the structures shown in (3) – (6).

3. Object Extracted - incongruent noun animacy order (SO AI)²:

The deputy that the newspaper identified chased the mugger

4. Subject Extracted - incongruent noun animacy order (OS IA):

The newspaper identified the deputy that chased the mugger

5. Object Extracted - congruent noun animacy order (SO IA):

The wood that the man chopped heated the cabin

6. Subject Extracted - congruent noun animacy order (OS AI):

The man chopped the wood that heated the cabin

Sentences were based on scenarios, each of which appeared once as an SO and once as an OS sentence. The animacy of subject and object noun phrases and the plausibility of the sentences was systematically varied within each sentence type.

These sentence types were selected because they have been used in previous studies; in particular, in the non-word detection study referred to above (Caplan et al, 2008a). They incorporate the following factors.

The subject-object extraction factor (SO vs OS) varies a syntactic factor. Matched SO and OS sentences have the same thematic roles, but different syntactic structures. SO sentences are more demanding syntactically than OS sentences, for many reasons. In Gibson's (1998) terms, the category that is required because of the presence of the relative pronoun (the relative clause verb) is further displaced from the relative pronoun in SO than in OS sentences, leading to increases in what he terms "storage costs." There is an NP intervening between the head of the relative clause and the position to which it is related, leading to increased "integration" costs, in Gibson's terms; in other models, the presence of this NP leads to increased interference in memory (Lewis et al, 2006). Two thematic roles can be assigned at the object position of the verb of the relative clause in SO sentences and only one at any point in OS sentences, leading to a higher local computational load; the sentence-initial noun phrase must be retained in memory over the relative clause to serve as

²Labels are as in previous work. SO: subject-object relative clause – an object-extracted relative clause whose head NP is the matrix sentence subject; OS: object-subject relative clause – a subject-extracted relative clause whose head NP is the matrix sentence object; AI: animate first NP, inanimate second NP; IA: inanimate first NP, animate second NP.

the subject of the main clause in SO sentences, also increasing memory demands. SO sentences may also be harder to process at the discourse level.

The animacy order factor is semantic, also affects processing load, and interacts with the syntactic factor. Traxler et al (2002,2005) have shown that SO sentences in which the first noun is animate and the second noun inanimate (sentences such as (3)) lead to increased regressions from and first pass fixations on the relative clause. Traxler et al. argued that an animate sentence-initial NP leads to the expectation that that NP will be the subject of the next verb encountered, and is therefore incongruent with the syntactic structure of the sentence, which requires this NP to be the object of that verb. Traxler et al (2002,2005) argued that this led to revisions of a structure that had been assigned when readers first fixated in the relative clause (a “syntax-first, garden-path” parsing model). These results can also be explained by constraint satisfaction models (e.g., MacDonald et al, 1994) that maintain that the expectations about thematic roles that are established by animacy of nouns in particular syntactic interact immediately with parsing.

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.

Sentences were presented in eleven fonts, each used in 2 - 5 sentences: Andale Mono 33; Arial 33; Bodoni Svty 36; Comic Sans MS 36; Courier 36; Geneva 36; Georgia 33; Georgia 36; Helvetica 33; New York 30; Times 36. In half the sentences of each type, one word was changed to a similar font, as follows (fonts shown in 12 point type; type size in stimuli are indicated): Andale Mono 33 --> Courier 36; Arial 33 --> Andale Mono 33; 36 --> Times 36; Comic Sans MS 36 --> Verdana 36; Courier 36 --> Andale Bodoni Svty Mono 33; Bodoni Svty > Geneva 36 --> Trebuchet MS 36; Georgia 33 --> Geneva 30; Georgia 36 --> 36; Helvetica 33 --> Trebuchet MS 33; New York 30 --> Times 33; Times 36 -- Svty 36). The stimuli were designed to make the task sufficiently difficult so that the changed Bodoni font did not “pop out.” This was accomplished through the use of different fonts, which did not allow participants to establish a single template against which a font change could be established, and using visually similar fonts for changes (the changed font was similar to the font of rest of the sentence in serif style and spacing). Each of the pairings was used 1 - 6 times to create sentences in which a word appeared in a different font (the two that were only seen once each were Andale Mono/Courier and Geneva/Trebuchet; for these pairings, it was difficult to change an entire word without it being very obvious where the change was). The location of the change was counterbalanced within the different sentences and the nature of the font and the font change randomly assigned across all sentences.

The task for the participants during the experimental trial was to read the sentence and judge whether all the words were in the same font as quickly and accurately as possible.

Behavioral Study

Font change detection in visually presented sentences – the task in the scanner -- provides information regarding (a) the effects of syntactic structure and noun animacy order on accuracy and RT to entire sentences without font changes, and (b) the effects of the position of words with font changes on accuracy and RTs. To add information about processing of specific phrases in sentences without font changes, we undertook a study of eye fixations in the font change task. Because we contrasted the BOLD signal results in the font change detection study with those in the non-word detection study, we also carried out a separate study of eye fixations in a non-word detection task, using the materials from Caplan (2008a), which, as noted, contain the same “normal” sentences as in the font change detection study.

Participants

Eleven subjects (9 females, 2 males; mean age 23 years) participated in the nonword detection study. One subject was dropped from participation due to difficulties maintaining accurate calibration. All subjects were native speakers of English with normal or corrected-to-normal vision, and were naïve to the purposes of the study. Subjects were paid for their participation.

Materials and Tasks

Materials and tasks were identical to those used in the font change detection study described above and those used in the nonword detection study in Caplan et al (2008a).

Procedure

The experiment was programmed and run using Experiment Builder software, and subject gaze data were collected using a head-mounted Eyelink II system, both products of SR Research Ltd. (Ontario, Canada). Average gaze-position error was less than 0.5 degrees after calibration, and the pupil locations were sampled in the right eye at a rate of 250 Hz. Participants were placed approximately 1 meter from a computer screen, and their head position was tracked in relation to the screen using an infrared camera positioned on the headset. One degree of visual angle subtended approximately 4.9 characters on average.

In order to present sentences containing more than one font type within the experimental program, all stimuli in the font change detection study were converted into .png image files of uniform pixel height and width. In both studies, sentences were aligned to the center left of each image, as in the fMRI studies.

The experiments began with a calibration procedure and a set of practice items. Before each trial, subjects were presented with a fixation point at the center of the screen, which both corresponded to the fixation cross presented during the fMRI study and also served to maintain calibration accuracy. The experimenter waited for subject gaze to stabilize near the fixation point, then pressed a key to present the sentence. Each sentence appeared in the vertical center of the screen, with the first word of each placed near the left screen edge.

In the font change detection study, for each sentence, subjects were told they would see a sentence and that one word in the sentence might appear in a different font than the others, and that the task was to determine if all the words appeared in either one or more than one font. They were asked to be as accurate as possible, and also to respond as quickly as possible by pressing one of two buttons. In the nonword detection study, for each sentence, subjects were told some sentences would contain a made-up word, and were asked to determine whether or not all the words in the sentence were real words of English, and to indicate their judgment as quickly and accurately as possible with a button press. The order of studies was counterbalanced across participants.

In both studies, each sentence appeared for a total of 5000ms, and only the first button response was recorded for each trial. Items were divided into lists in the same pseudo-randomized order used in the fMRI studies. Recalibrations were conducted and short breaks were provided between each list. In addition, the experimenter qualitatively observed participant gaze data online during the experiment using a separate display, and if the calibration appeared to have become inaccurate, a new one was performed between trials.

Gaze results were categorized using Dataviewer software (SR Research, Ltd.). Interest areas were created for phrases in each sentence (NP1, NP2, NP3, V1, V2). We analyzed four measures in those interest areas: the order of first fixations, first fixation time (the duration

of the first fixation in an interest area), second fixation time (the duration of the second fixation in an interest area), and total fixation time (the sum of the duration of all fixations in an interest area); a fifth measure we report is the difference between total and first fixation time.

Results

Data were analysed in SAS using repeated measures ANOVAs. Accuracy and RTs for sentences without font changes or nonwords were analyzed in 2 (structure) X 2 (noun animacy order) ANOVAs, for the font change and nonword detection studies separately (Figure 1). There were no significant effects of structure or noun animacy order in the accuracy data. For font change, the interaction of structure and animacy order was significant in the RT analysis ($F(1, 9) = 18.4, p < .001$). RTs were longer for sentences in which the animacy order of noun was incongruent with the syntactically licensed interpretation of the sentence (SO-AI and OS-IA sentences) than for sentences in which animacy order and syntactic structure were congruent. In the nonword detection study, the interaction of structure and animacy order was significant in the RT analysis ($F(1, 9) = 22.3, p < .001$). RTs were longer for SO-AI sentences than for all other sentences.

The distribution of RTs to sentences of each type was examined for evidence of the presence of more than one process in determining response times. To eliminate individual differences in overall RTs, all RTs for each participant were converted to z -scores, and the histograms of the entire set of z -scores by deciles were constructed for each sentence type in each condition, separately for sentences to which participants responded correctly and incorrectly (8 analyses). These histograms were assessed for bimodality using the bimodality co-efficient (b), in which evidence for bimodality is manifest by $b > .55$ (Farmer et al, 2007). All bimodality co-efficients (b) were below this value.

To explore the question of whether participants controlled eye movements using mechanisms associated with sentence processing, the order of first fixations on regions in the sentences without font changes or nonwords was examined. Fixations frequently occurred first on the mid and late segments of the sentences (Figure 2, left panels). This pattern is attributable to the effects of the pre-sentence fixation cue, which appeared above these positions in the sentences. First fixations occurred more frequently on phrases further to right in the font change than in the nonword study because many of the fonts occupied less horizontal space. To eliminate effects of the pre-trial fixation point, we removed the first first fixation in each trial from the analysis. After the initial first fixation, first fixations moved from left to right in both tasks (Figure 2, right panels), with minor differences between tasks.

Fixation duration data for sentences without font changes or non-words were analyzed in 2 (structure: SO, OS) X 2 (animacy: AI, IA) X 5 (phrase (N1, N2, N3, V1, V2)) ANOVAs separately for the font change and nonword detection tasks. Because of the effect of the pre-fixation cue on first fixations, we considered second fixations (on which there were no effects of the variables manipulated), total fixation time, and the difference between total and first fixation time. Total fixation time gives a measure of the overall demand made by each phrase; the difference between total and first fixation duration reflects processes that occur on each segment after a word is first encountered (where the effects of sentence-level variables may have been affected by the unusual order of first encounter). These data are shown in Figure 3.

For the font change task, for total fixation times, the interaction of structure and phrase was at the level of a trend ($F(4, 36) = 2.4, p = 0.07$) and three way interaction of structure, animacy order and phrase was not significant. To explore possible effects of sentence

variables that exerted significant effects on BOLD signal in the non-word detection task, the terms of the three way interaction were examined using Tukey's tests. This showed that total fixation times were longer on the relative clause subject noun (N2) in SO-AI sentences than on the embedded verb (V1) of OS-AI sentences ($p = .04$). For the difference between total fixation and first fixation time, the interaction of structure, animacy order and segment was at the level of a trend ($F(4, 36) = 2.1, p = .09$). The difference between total and first fixation times was marginally greater at the relative clause subject (N2) of SO-AI sentences than the relative clause verb (V1) in OS-AI sentences ($p = .1$).

For nonword detection, for total fixation time, the three-way interaction of structure, animacy order and phrase was significant ($F(4, 36) = 7.3, p < 0.01$). Total reading times were longer for V1 in SO than for V1 and N2 in OS sentences regardless of animacy order, and for N2 in SO-AI sentences than for V1 and N2 in OS-AI sentences and N2 in OS-IA sentences. For the difference between total and first fixation time, the interaction was also significant ($F(4, 36) = 4.5, p < .01$). The difference between total and first fixation times was greater at the relative clause subject (N2) and verb (V1) of SO-AI sentences than the relative clause subject or verb of OS sentences

Discussion of Eye Tracking Experiment

The RT data show effects of sentence-level variables (syntactic structure and animacy order) that are consistent with the participants processing the sentences to the level of structure and meaning. The distribution of RTs provide evidence that they processed the majority of stimuli in similar ways. The eye tracking data add information about the processing of these sentences.

First, the eye tracking study showed that the pre-sentence fixation point led to many initial eye fixations occurring on the portion of the sentence adjacent to the fixation cue. In almost all eye-tracking studies of sentence processing, participants have seen sentences without a fixation cue, and, in languages in which words are displayed left-to-right, eye fixations have moved in that order from sentence onset, providing initial information about words in sentences that is similar to that available with auditory input. One consequence of the effect of the pre-stimulus fixation is that the interpretation of first fixation durations cannot be easily compared to interpretation of these measures in studies in which fixations move left-to-right from sentence onset, because participants in our studies had an early exposure to later portions of a sentence. A point we make in passing is that the use of a pre-sentence fixation cue is very common in fMRI studies of sentence processing; one implication of this study is that future studies should consider eliminating pre-sentence fixation cues, or having them appear in the spatial location of the onset of the stimulus sentences.

The patterns of first fixations from which the first first fixation was and was not excluded demonstrate the effects of control of eye fixations by attentional and linguistic factors, respectively. Once allowance was made for the position of the eyes at the start of each trial, eye fixations moved from left to right in sentences without font changes or nonwords. As noted, a left-to-right path of first fixations is the normal pattern of eye movements in reading English (Reichle et al, 2003), and its presence points to the top-down control system in these tasks being the one that controls eye movement in sentence reading. The fact that first fixations were heavily influenced by the position of the pre-stimulus fixation cue and subsequent first fixations on interest areas were not is informative: it shows that eye fixations were initially affected by control mechanisms associated with processing of non-linguistic stimuli (the effect of bottom-up attraction of attention to a cued location) and that these control mechanisms were subsequently overridden by ones related to sentence processing.

Second, there were effects of structural and semantic factors on fixation times in sentences without font changes or nonwords. In the font change detection study, total fixation times and the difference between total and first pass fixations, which are relatively uninfluenced by the effect of the pre-stimulus fixation point on first fixations, were longest on the subject of the relative clause in SO-AI sentences. Because these words were in the same font as the other words in the sentence, the longer eye fixations on these words cannot be due to determining that no font change was present. Similarly, because these words were the same as those that produced shorter fixations in less demanding sentences, we conclude that the extra time spent in fixating them was not determined by lexical access. In the nonword detection study, total fixation times and the difference between total and first fixation times were increased at the verb and subject noun of the relative clause of SO-AI sentences. Again, because these words were the same as those that produced shorter fixations in less demanding sentences, we conclude that the extra time spent in fixating them was not determined by lexical access. We can thus conclude that the longer fixations on words in these positions was due to implicit sentence-level processing of these words, a conclusion that is reasonable given the fact that these words are in the positions of greatest parsing and interpretive demand in the most demanding sentences; that is, that the longer fixations for these words in these positions reflects participants' not moving their eyes until this sentence-level processing reached some point of completion. We consider what these implicit and obligatory sentence-level parsing and interpretive operations might have been in the discussion of the results of the fMRI study.

To summarize, the eye tracking results in sentences without font changes or non-words (i.e., in "normal" sentences) point to effects of aspects of the experimental set-up, and suggest that the subject NP of the relative clause in SO-AI sentences was the point of greatest consistent load in the font change study, and the embedded verb and subject NP of SO-AI sentences were the points of greatest load in the nonword detection study. As noted, we discuss the implications of these loci of load below.

MR Study

Participants—Sixteen participants (9 males and 7 females; mean age 21 years) 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 naive as to the purposes of the study. Participants were paid for their involvement.

MR Imaging Parameters—Participants were scanned in a single session in a 3.0T whole body Siemens Trio scanner (Siemens Medical Systems, Iselin, NJ). Two sets of high-resolution anatomical images were acquired using a T1-weighted MP-RAGE sequence (TR = 2530 ms, TE = 3 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 × 1.0 mm (256 × 256 matrix, 256 mm Field of View (FOV)). 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 × 3.13 mm (64 × 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.

Item Presentation—In all experiments, each stimulus was visually displayed in its entirety on a single line in the center of the screen. The sentences were projected to the back of the scanner using a Sharp LCD projector and 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 a proprietary software package (Stimpres) was used to both present the stimuli and record the accuracy and reaction times.

A given experimental trial consisted of a brief 300 ms fixation cross, a 100 msec blank screen, the sentence presented for 5 sec, and a final 600 ms blank screen, for a total trial length of 6 sec. Fixation trials (0-12 second centered +) were randomly interspersed between each 6 second sentence trials. A pseudo-randomized item presentation order was determined for the event-related design by a computer program (Optseq) developed to randomize trial types and vary the duration of inter-stimulus fixation trials for optimum efficiency in the deconvolution and estimation of the hemodynamic response (Burock et al., 1998; Dale, 1999; Dale and Buckner, 1997). The 144 stimulus items interspersed with the fixation trials were divided into 4 runs. No pair of matched SO and OS sentences were presented in the same run. Participants were given a short break after each 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 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)—Group statistical activation maps were constructed for contrasts of interest using a *t* statistic. 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 the 20 second period beginning with the onset of each sentence, which contains the time points at which vascular responses were expected to peak (Bandettini, 1993; Turner, 1997).

To correct for multiple spatial 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 mm² and threshold for rejection of the null hypothesis at $p < .05$. These functional activations were 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 (i.e., the voxel with the greatest significance level).

Conjunction Analysis—The uncorrected significance maps from the contrast that yielded significant effects (SO-AI/OS-IA) in the non-word and the font detection tasks were thresholded at $p < .01$. These binarized maps were combined in an “AND” conjunction (intersection).

Results

Behavioral Results—Mean RT and accuracy for sentences without font changes are displayed in Figure 4. Analyses of variance (2 (sentence type: SO, OS) X 2 (animacy order: AI, IA)) were performed over participant and item means for accuracy and RTs that had been trimmed for outliers ± 3 sd from the condition mean for each participant. Significant terms of significant interactions were identified using Tukey's test.

There were no significant effects in the accuracy data. For the RT data, for sentences without font changes, there was an effect of structure ($F_1(1, 15) = 21.2, p < .001$; $F_2(1, 68) = 5.6, p < .05$); responses were faster for OS than for SO sentences.

RTs to sentences with font changes were analyzed for the effect of the position of the font change as a function of sentence structure and the order of animate and inanimate nouns. The data are displayed in Figure 5. ANOVAs (2 (sentence type: SO, OS) X 2 (animacy order: AI, IA) X 5 (region: NP1, SONP2/OSV1; SOV1/OSNP2; V2; NP3) showed a significant interaction of sentence type, animacy order, and position of font change (F_1

(4,48) = 5.2, $p < .001$; $F_2(4, 52) = 2.2, p = .08$). For sentences in which the order of animacy of noun phrases biased against syntactically determined meaning (“incongruent sentences,” SO-AI and OS-IA -- Figure 5, top panel), there were significantly longer RTs for sentences in which the font change occurred at the embedded verb of SO sentences compared to font changes at either the first verb or the main clause object NP of OS sentences and a trend for longer RTs for SO sentences in which the font change occurred at the embedded subject noun phrase than for OS sentences in which the change occurred at the first verb; RTs were also longer for the final noun in OS-IA than in SO-AI sentences. For sentences in which the order of animacy of noun phrases biased towards syntactically determined meaning (“congruent sentences,” SO-IA and OS-AI - Figure 5, bottom panel), there were no phrases at which RTs differed in the sentence types.

Histograms of RTs by deciles, shown in Figure 6, were constructed for each sentence type in each condition for each subject separately for sentences to which participants responded correctly and incorrectly, and assessed for bimodality using the bimodality co-efficients (b), as described above. All bimodality co-efficients (b) were below .55.

fMRI Results—BOLD signal increased for SO compare to OS sentences in extrastriate cortex bilaterally. The interaction of the SO/OS and animacy order factors yielded no areas of activation. To compare results with those of the non-word detection study, we examined the BOLD signal effect of the SO and OS sentence contrasts in which the order of animacy of noun phrases was incongruent with the syntactically derived meaning of the sentence (SO-AI vs OS-IA); Figure 7 and Table 1. There were significant increases in BOLD signal in visual association areas bilaterally and in left supramarginal gyrus, and significant decreases in BOLD signal in two right frontal areas. Time course data show that the decreases in BOLD signal for SO-AI compared to OS-IA sentences in right frontal lobe were due to decreases in BOLD signal for SO-AI sentences below the pre-stimulus baseline. The comparison of SO and OS sentences in which the order of animacy of noun phrases was congruent with the syntactically derived meaning of the sentence (SO-IA vs OS-AI) showed a very small area of activation in extrastriate cortex bilaterally.

There were no areas of activation in the conjunction analysis of the SO-AI vs OS-IA contrasts in this study and in the nonword detection study reported in Caplan et al (2008a).

Discussion of fMRI study—There is evidence from the effects of the syntactic and semantic variables we manipulated that participants processed stimuli to the level of syntactic structure and meaning. First, for sentences without font changes, RTs were longer for object- compared to subject-extracted relative clauses, indicating that participants engaged in more linguistic processing of object- than of subject-extracted sentences (which led to longer times to indicate whether a font change occurred). Second, for sentences with font changes, RTs were longer (a) when the font change occurred at the embedded verb of SO-AI compared to either the first verb or the main clause object NP of OS-IA sentences and (b) when the font change occurred at the embedded subject noun phrase of SO sentences than when it occurred at the first verb of OS sentences (the latter at the level of a trend). This indicates that participants engaged in more linguistic processing of words in these than in other locations in these sentences, leading to longer times to indicate whether a font change occurred.

The results provide evidence that participants processed most stimuli syntactically and semantically: the sentence type effect was significant by items and that the distribution of RTs for all sentence types was unimodal. If participants assigned syntactic structure to only some sentences and only analyzed fonts in others, one might expect that the effect of syntax

would not be significant in the item analyses and that the distributions of RTs would be bimodal.

With these behavioral results in mind, we can attempt to relate the neurovascular effects to psychological operations. There are two related issues to consider: the location of areas in which BOLD signal increased in the more complex sentences, and the psychological operations that caused these increases.

Beginning broadly, we note that it is unlikely that all the areas of BOLD signal increase in the SO-AI/OS-IA contrast are ones in which parsing and interpretation took place. It is much more likely that the increased BOLD signal in visual association areas reflects increased visual processing that occurred in SO sentences because they are fixated for longer periods of time than that it reflects the assignment of syntactic and semantic structure. The lower activity for BOLD signal for SO-AI compared to OS-IA sentences in right frontal lobe is not explicable as due to parsing and interpretive processes that apply to OS and not SO sentences, because the latter are more complex and harder to process than the former. Since the BOLD signal for SO-AI sentences fell below the pre-stimulus baseline, it is possible that this effect is due to inhibition of this area by the area in which parsing and interpretation takes place, which would be greater for the more demanding sentence type. The increase in BOLD signal in the SO-AI/ OS-IA contrast in left supramarginal gyrus is most likely to reflect parsing and interpretation of the sentences.

This area is very close to that activated in some previous studies of the object-/subject-extraction contrast (Caplan et al, 2002), but differs from that in other studies (Just et al, 1996, 2004; Reichle et al., 2000; Keller et al., 2001; Ben Schachar et al, 2003, 2004; Fiebach et al., 2004; Prat et al., 2007), which also show a large set of activated areas. The differences in the exact areas activated in these different studies are unexplained. They may indicate a degree of variability within the post-Rolandic peri-Sylvian association cortex for supporting syntactic operations (see Caplan et al, 2007, for discussion). Alternatively, they may reflect different localizations of particular parsing and interpretive operations: the exact contrast used here (SO-AI *versus* OS-IA sentences) differs from those used in other studies, and the SO-AI sentences are likely associated with operations not found in other experimental sentences (see discussion below).

Most important from the point of view of the rationale for this study, the areas activated in this study did not overlap with the region in the left inferior frontal lobe area activated by the same sentence contrast in the non-word detection study reported by Caplan et al (2008a). If the same psychological operations were applied by participants in both studies, there should be overlap between regions activated in the two tasks (assuming that the sensitivity of the detection of BOLD signal effects of these operations did not differ in the two studies, which were performed on the same equipment and analyzed the same way). We can reject the possibility that one area of activation reflects the psychological operations directly involved in sentence comprehension that differ in the SO-AI/ OS-IA contrast and the other reflects strategic or ancillary cognitive operations that differed between these sentence types, because strategies are triggered by the effort to structure and understand sentences and we would have no correlate of the parsing and interpretive operations that differ in SO-AI and OS-IA sentences in the task in which we attribute BOLD signal effects to strategic or ancillary cognitive operations. This leaves one of two possible interpretations of the different activations for the same contrasts in the two tasks.

One is that the psychological operations that apply in SO-AI and not OS-IA sentences are the same in the two tasks, and are localized in different areas in different individuals (see Caplan et al, 2007, for discussion of variable localization). This interpretation is possible,

but it would require that one study happened to test primarily participants in whom these operations were localized posteriorly and that the other tested primarily participants in whom these operations were localized anteriorly. We thus consider it unlikely.

A second possibility is that different parsing and interpretive operations were applied implicitly in the two tasks. The feature of the behavioral data that provides the best clues as to what those different operations might be is the evidence that the subject of the object relative clause in SO-AI sentences was consistently the point of greatest demand in the font change detection study and the verb of this clause the point of greatest demand in the non-word detection study. These results are summarized in Table 2, both for RTs for sentences with font changes or nonwords in the scanner and for sentences without such changes in the eye fixation studies.³

Two types of operations can lead to increased processing load at the subject of the object-extracted relative clause and not at the verb of such clauses. One are operations that predict upcoming words, which generate “entropy” or “surprisal” effects -- the negative log probability of the occurrence of a word in a context. These models predict long reading times and fixations on the relative clause subject of object extracted clauses (Levy, 2008).⁴ The second type of operation are ones that assign generalized thematic roles (prototypical agent, theme, instrument, experiencer), which can be assigned at the subject of the relative SO clause (Bornkessel and Schlesewsky, 2006). The evidence that the subject of the relative SO clause was the point of greatest load in the font change detection is consistent with the view that participants’ sentence processing led to the generation of surprisal effects and/or to the assignment of generalized thematic roles in the font change detection study.

In contrast, an operation that can only occur at the verb of the relative clause and that is more demanding in object- than in subject-extracted sentences (and most demanding in SO-AI sentences) is the assignment specific thematic roles (agent, theme, etc., specified for a particular type of entity and action). As discussed above, either a constraint satisfaction process that integrates conflicting syntactic and semantic cues (MacDonald et al., 1996) or a revision process (Traxler et al, 2002, 2005) makes SO-AI a more resource-demanding sentence than the other sentence types. These processes utilize information in the relative clause verb, and therefore their operation could underlie the increased RTs for nonwords in that position and the BOLD signal for SO-AI sentences seen in the non-word detection study.

To summarize, the dissociation between the BOLD signal response to the same sentences in the font change and non-word detection studies indicates that different parsing and interpretive operations applied in the two tasks. Behavioral data suggest that the effects of

³In sentences with font changes, there were differences in the effects of position of font changes on RTs in the eye tracking study and the MR environment; for the 2 (environment: eye tracking; MR) X 2 (structure: SO, OS) X 5 (region) X 2 (animacy: AI, IA) X 5 (location of font change/nonword: N1, N2 N3, V1, V2) ANOVA, the interaction of environment, structure, animacy, and location was significant ($F(4, 96) = 7.2, p < 0.001$). For sentences with font changes, mean RTs were faster in the eye tracker than in the scanner (2162 msec in the eye tracker vs 3776 msec in the scanner), possibly due to increased attention to visual processing in the head-mounted eye tracking environment, and the differences in response times as a function of sentence type, animacy order and location of font change found in the scanner were not found in the eye tracking environment, likely a floor effect related to the fast RTs. In the nonword detection in study, RTs for sentences with nonwords were the same in the two environments (2107 msec in the eye tracker vs 2112 msec in the scanner) and were longer for SO sentences in which the nonword occurred in the position of the relative clause verb in both environments; the interactions of environment with structure and location and with structure, animacy, location were not significant.

⁴The BOLD signal effect only occurred in SO-AI sentences, and surprisal is expected to occur at the relative clause subject of all SO sentences. Categorizing nouns as animate or inanimate will lead to different difficulty predictions if the frequency of occurrence of animate and inanimate NPs differs in object relative clauses and the corpus as a whole (see Levy, 2008, for discussion). SO-AI sentences are likely to be less common than SO-IA sentences, leading to increased surprisal effects at N2 in SO-AI than in SO-IA sentences.

sentence type in the font change study are due to surprisal or assignment of generalized thematic roles and that these effects are due to the assignment of specific thematic roles in the nonword detection study. In the General Discussion, we take up the implications of the conclusion that different parsing and interpretive operations appear to have occurred in two different tasks in which parsing and interpretation occur implicitly.

General Discussion

The most important result of this study is that there are BOLD signal effects of syntactic and semantic variables in the font change detection study reported here, and that these BOLD signal effects do not overlap with those associated with the same sentence contrast in the nonword detection study reported in Caplan et al (2008a). The different BOLD signal effects of the same sentence contrast in these two studies indicate that “implicit” parsing and interpretation differed in these two tasks. This raises a number of questions about tasks in which parsing and interpretation are implicit.

The rationale for, and advantage of, tasks in which parsing and interpretation are implicit is that they eliminate possible interactions of sentence processing and response selection from affecting behavioral and neural parameters. This advantage comes at the cost of not having a direct measure of the process of interest – sentence comprehension. There is no way around this limitation if one uses a task that eliminates effects of task-related operations on BOLD signal. Even occasional probing of meaning in an “implicit” task (e.g., by asking whether a sentence was plausible, or asking a question about its meaning) would force participants to process all sentences to meet this task demand, defeating the purpose of using a task in which the products of the comprehension process are not used to accomplish the experimental task. In addition, rare sampling of comprehension could be misleading: if a participant made errors on a few probes, s/he might yet be processing many other sentences to the level of propositional content and, conversely, if s/he did well on a few probes, these might be the exceptions to his/her general level of processing. It would be possible to conduct unannounced post-MR recall or recognition tests, but they also provide limited information because participants’ not remembering sentences that appeared early in the study could be because they did not process them or because they forgot them.

One might imagine that one cost of using tasks in which sentence comprehension is implicit is that it is hard to know exactly what operations participants use to generate the effects of sentence-level variables. This is the case in the studies presented here. Based on RTs in the scanner and eye fixation patterns outside the scanner, we suggested that the left inferior parietal activation seen in the present study might be due to surprisal effects in sentence comprehension and/or to the assignment of generalized thematic roles, and that the left inferior frontal activation seen in the non-word detection study using the same materials might be due to assignment of specific thematic roles, but our suggestions regarding the parsing and interpretive operations that participants applied in these studies are clearly tentative. However, we do not believe the principal reason for the difficulty in interpreting the behavioral and neural measurements here, or in other studies that use implicit measures, is solely, or even mainly, that they are made in tasks in which comprehension is implicit.

We have attempted to infer the nature of participants’ syntactic and semantic processing from the on-line effects of syntactic and semantic variables on performance on a task, but this is required in tasks that measure comprehension as well. What is gained by using tasks in which comprehension is measured is re-assurance that comprehension has taken place to the point that the task can be accomplished, but this does not tell us what operations participants employed to gain that understanding. For instance, the effects of incongruity between syntactic and animacy information that Traxler et al (2002,2005) attributed to revision of an assigned syntactic structure can also be explained as effects of low level

variables on first fixations and subsequent interactions of a large set of variables affecting later fixations. Measuring comprehension does not help say which of these operations underlies the eye fixations reported in Traxler's studies. Nor does it help us verify many hypotheses about what operations occur at particular points in sentences because some types of representations that are postulated to be constructed cannot be observed directly. For instance, using a comprehension task does not allow a researcher to verify the hypothesis that generalized thematic roles are assigned at a point in a sentence (see, e.g., Bornkessel and Schleewsky, 2006), because it is not possible to test for comprehension of generalized, as opposed to specific, thematic roles.

The difficulties associated with determining the operations that drive BOLD signal in tasks in which parsing and interpretation are either explicit or implicit leads to consideration of the benefits of the absence of any task at all ("passive" viewing or listening). However, this is not an attractive option in the effort to eliminate task-related operations that co-occur with parsing and interpretation from affecting BOLD signal correlates of sentence contrasts. These tasks solve none of the problems discussed here, and raise many of their own. In the absence of a task, sentence type effects on neural variables are very difficult to relate to psycholinguistic operations because of the great latitude available to participants regarding the use of particular processing operations. Without behavioral data, there is nothing to suggest hypotheses regarding the operations that determine these BOLD responses except the location of the responses themselves except the areas that are activated, which is at present insufficient for this purpose. Progress in determining the neural basis for syntactic processing will require correlation of patterns of neurovascular responses with behavioral data that give clues as to what specific operations might have been activated in a study, which are unavailable if no task is used.

Returning to the use of tasks in which parsing is implicit, we have come to the conclusion that, although such tasks eliminate psychological operations associated with the use of the products of the comprehension process to accomplish a task from affecting either behavioral or neural correlates of sentence contrasts, they do not eliminate the effects of tasks on parsing and interpretation itself. The reason that tasks in which parsing and interpretation are implicit affect parsing and interpretation themselves that different implicit tasks are associated with different deployments of attention. Two aspects of attention are relevant: the processes whereby attention is directed towards parts of a stimulus (conceiving of attention as a spotlight; Treisman and Gelade, 1980) and the processes whereby attention is allocated to processing a stimulus dimension (conceiving of attention as a resource; Kahneman, 1973). Both these aspects of attention can affect parsing and interpretation, the first by leading to different sequences of perception of the language input and thereby allowing operations to occur in specific orders, and the second by allowing different amounts of processing to be applied to that input.

In the case of the tasks under consideration here, attention was initially visually directed to different phrases in the font change and nonword detection tasks because of the effect of the pre-sentence fixation cue, affecting the initial stage of implicit sentence processing. With respect to the "resource" demands of the tasks, the fact that different fonts were used in different sentences in the font change task and only one font was used in the nonword detection task adds to the processing required by low level visual perceptual operations, and the process of determining whether all words are in the same font or not differs from, and makes different demands than, the process of determining whether all letter strings are existing words. The RT data in the scanner suggest that the font change task was more difficult than the nonword detection task. If so, the resources available for implicit sentence processing may have been sufficient to support surprisal affects and to assign generalized thematic roles, which depend upon syntactic position and noun animacy but not specific

verb information, but not sufficient to assign specific thematic roles, which depend upon syntactic position and specific verb and noun semantic information.

The effect of deployment of attention on parsing and interpretation has been explored through studies of comprehension in the “unattended channel” (e.g., Lackner and Garrett, 1972). This literature shows that sentence-level processing differs as a function of the task in the attended channel. For instance, Potter et al (2007) reported no effect of scrambling a written sentence presented in rapid serial visual presentation upon reporting two words presented in red, and concluded that readers did not process the sentence in these detection conditions. The difference between the absence of a sentence structure effect in that task and the presence of sentence type effects in the font change detection task used here probably is due to the fact that color changes are much more easily detected than the font changes used here, and the time taken to determine whether words have a different font allows for parsing and interpretation. It is likely that detecting obvious font changes would be associated with very fast RTs and not with syntactic or semantic effects. What features of a task allow the “attention filter” to be set after lexical access but before parsing and interpretation remain to be explored.

The effect of task on attention is not restricted to tasks in which parsing and interpretation are implicit. Tasks in which the products of the comprehension process are relevant to response selection also demand attention. In essence, all comprehension tasks create dual task conditions, with one task consisting of the assignment of sentence meaning and the second consisting of perceptual-motor functions such as picture (or, in the real world, scene) analysis, planning actions, encoding and maintenance of propositional information in memory, rating plausibility, etc. Differences in aspects of tasks such as the quality of drawings, the length of time between presentation of a sentence and a probe, the subtlety of plausibility judgments, etc., affect the extent to which these concurrent functions demand attention and resources, and thereby affect on-line sentence comprehension.

Neuroimaging can combine with behavioral measures to advance the interpretation of behavioral measures associated with parsing and interpretation as well as neural correlates of these operations. The effects of interactions of task-related processing and parsing and interpretation are inextricably interwoven with those of parsing and interpretation *per se* in any one behavioral measure but may be manifest by neural activity in different brain areas. In the study reported here, the fact that there were different areas of BOLD signal response to the same sentence contrast was the first indication that different parsing and interpretive operations occurred in two tasks. The joint use of behavioral and neural data can move inquiry forward in a number of ways. Conjunction analyses that identify areas of the brain in which the same sentence contrast leads to neurovascular effects in a set of tasks provide candidates for areas that support task-independent parsing and interpretive operations. The corresponding behavioral data are sentence type effects that do not interact with task. The combination of such neural and behavioral data would provide evidence for the existence and neural location of parsing and interpretive operations that apply independent of task. In combination with behavioral information, effects of syntactic and semantic factors on neural parameters in areas that are found in such conjunction analyses afford information about the task-independent parsing and interpretation operations that take place in those areas (Caplan, 2009).

Returning to theme with which this paper began – the search for the neural basis of parsing and interpretive processes in isolation from other operations that occur during the comprehension process -- it is unclear at present what the set of task-independent parsing and interpretive operations is. Tanenhaus et al (2000) suggest that it may be very limited: “It is becoming increasingly clear that even the earliest moments of linguistic processing are

modulated by context ... call[ing] into question the long-standing belief that perceptual systems create context-independent perceptual representations.” (p. 577) It may be that only the most local combinations of words occur without interacting with tasks. In that case, there will no common areas that are activated in several tasks by sentence contrasts that incorporate more abstract aspects of parsing and interpretation, and there will be interactions of most sentence factors with task in behavioral measures. Regardless of how the facts turn out, consideration of the effects of task on syntactic and interpretive is needed to understand both the behavioral and neurological correlates of sentence type contrasts.

Acknowledgments

This research was supported by a grant from the NIDCD (DC02146). I am indebted to Gayle DeDe, David Gow, Gina Kuperberg, Louise Stanczak, and Gloria Waters for discussion of the issues raised here. Jung Min Lee was responsible for fMRI data collection, she and Rebecca Woodbury for fMRI data analysis, and Jennifer Michaud for creation of the font change materials and analysis of behavioral data. Will Evans set up, ran and analyzed the data in the eye tracking study. I am indebted to two anonymous reviewers of this paper for raising important issues about the interpretation of the results and helping me convey them and my interpretation of them coherently.

References

- Altman GTM, Kamide Y. The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language* 2007;57(4):502–518.
- Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS. Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine* 1993;30:161–173. [PubMed: 8366797]
- Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, Grodzinsky Y. The neural reality of syntactic transformatons: Evidence from fMRI. *Psychology Science* 2003;14:433–440.
- Ben Shachar M, Palti D, Grodzinsky Y. The neural correlates of syntactic movement: converging evidence from two fMRI experiments. *Neuroimage* 2004;21:1320–1336. [PubMed: 15050558]
- Bornkessel, Schlesewsky. The Extended Argument Dependency Model: A Neurocognitive Approach to Sentence Comprehension Across Languages. *Psychological Review* 2006;113:787–821. [PubMed: 17014303]
- Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *Neuroreport* 1998;9:3735–3739. [PubMed: 9858388]
- Burock MA, Dale AM. Estimation and detection of event-related fMRI signals with temporally correlated noise: A statistically efficient and unbiased approach. *Human Brain Map* 2000;11:249–260.
- Caplan D. fMRI studies of syntactic processing. *Current Medical Imaging Reviews* 2006;2:443–451.
- Caplan D. Functional Neuroimaging Studies of Syntactic Processing in Sentence Comprehension: A Critical Selective Review. *Language and Linguistic Compass* 2007;1:32–47.
- Caplan D. Experimental Design and Interpretation of Functional Neuroimaging Studies of Cognitive Processes. *Human Brain Mapping* 2009;30:59–77. [PubMed: 17979122]
- Caplan D, Alpert N, Waters G. Effects of syntactic structure and propositional number of patterns of regional cerebral blood flow. *J Cogn Neurosci* 1998;10:541–552. [PubMed: 9712683]
- Caplan D, Alpert N, Waters G. PET studies of sentence processing with auditory sentence presentation. *Neuroimage* 1999;9:343–351. [PubMed: 10075904]
- Caplan D, Alpert N, Waters G, Olivieri A. Activation of Broca’s area by syntactic processing under conditions of concurrent articulation. *Human Brain Map* 2000;9:65–71.
- Caplan D, Vijayan S, Kuperberg G, West C, Waters G, Greve D, Dale AM. Vascular responses to syntactic processing: event-related fMRI study of relative clauses. *Hum Brain Mapp* 2002;15(1): 26–3. [PubMed: 11747098]

- Caplan D, Waters G, Kennedy D, Alpert N, Makris N, DeDe G, Michaud J, Reddy A. A study of syntactic processing in aphasia II: Neurological aspects. *Brain and Language* 2007;101:151–177. [PubMed: 16997366]
- Caplan D, Chen E, Waters G. Task-dependent and task-independent neurovascular responses to syntactic processing. *Cortex* 2008a;44:257–275. [PubMed: 18387556]
- Caplan D, Stanczak L, Waters G. Syntactic and thematic constraint effects on BOLD signal correlates of comprehension of relative clauses. *Journal of Cognitive Neuroscience* 2008b;20:643–656. [PubMed: 18052788]
- Chen E, West C, Waters G, Caplan D. Determinants of BOLD Signal Correlates of Processing Object-Extracted Relative Clauses. *Cortex* 2006;42:591–604. [PubMed: 16881269]
- Cooke A, Devita C, Gee J, Aslop D, Detre J, Chen W, Grossman M. Neural basis for sentence comprehension: grammatical and short-term memory components. *Human Brain Map* 2001;15:80–94.
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* 1996;29:162–173. [PubMed: 8812068]
- Dale AM. Optimal experimental design for event-related fMRI. *Human Brain Map* 1999;8:109–114.
- Dale AM, Buckner RL. Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Map* 1997;5:329–340.
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis – I: Segmentation and surface reconstruction. *Neuroimage* 1999;9:179–194. [PubMed: 9931268]
- Doherty CP. Question/statement judgments: An fMRI study of Intonation Processing. *Human Brain Mapping* 2004;23:85–98. [PubMed: 15340931]
- Farmer TA, Anderson SE, Spivey MJ. Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language* 2007;57(4):570–595. [PubMed: 18037980]
- Ferreira F. The misinterpretation of noncanonical sentences. *Cogn Psych* 2003;47:164–203.
- Fiebach CJ, Schlesewsky M, Lohmann G, Von Cramon DY, Friederici AD. Revisiting the Role of Broca's Area in Sentence Processing: Syntactic Integration versus Syntactic Working Memory. *Human Brain Mapping* 2005;24:79–91. [PubMed: 15455462]
- Fiebach CJ, Schlesewsky M, Friederici AD. Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *J Psycholing Research* 2001;30:321–338.
- Fiebach CJ, Vos SH, Friederici AD. Neural Correlates of Syntactic Ambiguity in Sentence Comprehension for Low and High Span Readers. *Journal of Cognitive Neuroscience* 2004a; 16:1562–1575. [PubMed: 15601519]
- Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis – II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 1999a;9:195–207. [PubMed: 9931269]
- Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Map* 1999b;8:272–284.
- Fischl B, Liu A, Dale AM. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans Med Imaging* 2001;20(1): 70–80. [PubMed: 11293693]
- Garnsey SM, Tanenhaus MK, Chapman RM. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 1989;18:51–60. [PubMed: 2926696]
- Grodzinsky Y. The neurology of syntax: language use without Broca's area. *Behavioral and Brain Sciences* 2000;23:1–71. [PubMed: 11303337]
- Grodzinsky Y, Friederici A. Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology* 2006;16:240–246. [PubMed: 16563739]
- Hashimoto R, Sakai K. Specialization in the left prefrontal cortex for sentence comprehension. *Neuron* 2002;35:589–597. [PubMed: 12165479]
- Indefrey P, Hagoort P, Herzog H, Seitz RJ, Brown CM. Syntactic processing in the left prefrontal cortex in independent of lexical meaning. *NeuroImage* 2001;14:546–555. [PubMed: 11506529]
- Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR. Brain activation modulated by sentence comprehension. *Science* 1996;274:114–116. [PubMed: 8810246]

- Just MA, Newman SD, Keller TA, McEleney A, Carpenter PA. Imagery in sentence comprehension: An fMRI study. *NeuroImage* 2004;21:112–124. [PubMed: 14741648]
- Kahneman, D. *Attention and Effort*. Prentice Hall; Englewood Cliffs, NJ: 1973.
- Kang AM, Constable RT, Gore JC, Avrutin S. An event-related fMRI study of implicit phrase-level syntactic and semantic processing. *Neuroimage* 1999;10(5):555–61. [PubMed: 10547332]
- Keller TA, Carpenter PA, Just MA. The neural bases of sentence comprehension: An fMRI examination of syntactic and lexical processing. *Cerebral Cortex* 2001;11:223–237. [PubMed: 11230094]
- Konieczny, L.; Doring, P. Anticipation of clause final heads. Evidence from eye-tracking and SRNs; The Joint International Conference on Cognitive Sciences; Sydney, Australia. 2003, July;
- Lackner JR, Garrett MF. Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition* 1972;1:359–372.
- Levy R. Expectation-based syntactic comprehension. *Cognition* 2008;106:1126–1177. [PubMed: 17662975]
- Lewis RL, Vasishth S, Van Dyke JA. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences* 2006;10:447–454. [PubMed: 16949330]
- Love T, Haist F, Nicol J, Swinney D. A functional neuroimaging investigation of the roles of structural complexity and task-demand during auditory sentence processing. *Cortex* 2006;42:577–590. [PubMed: 16881268]
- Marslen-Wilson W, Tyler LK. The temporal structure of spoken language understanding. *Cognition* 1980;8:1–71. [PubMed: 7363578]
- Noppeney U, Price CJ. An fMRI Study of Syntactic Adaptation. *Journal of Cognitive Neuroscience* 2004;16:702–713. [PubMed: 15165357]
- Pearlmutter NJ, MacDonald MC. Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language* 1995;34:521–542.
- Potter, MC.; Nieuwenstein, M.; Strohminger, N. Whole versus partial report: The attentional blink in RSVP sentences; paper presented at Psychonomic Society 48th meeting; Long Beach, CA. 2007;
- Prat C, Keller TA, Just MA. Individual differences in sentence comprehension: A functional Magnetic Resonance Imaging investigation of syntactic and lexical processing demands. *Journal of Cognitive Neuroscience* 2007;19:1950–1963. [PubMed: 17892384]
- Reichle ED, Carpenter PA, Just MA. The neural basis of strategy and skill in sentence-picture verification. *Cognitive Psychology* 2000;40:261–295. [PubMed: 10888341]
- Reichle ED, Rayner K, Pollatsek A. The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences* 2003;26:445–476. [PubMed: 15067951]
- Reichle ED, Rayner K, Pollatsek A. E-Z Reader: A cognitive-control, serial-attention model of eye-movement control during reading. *Cognitive Systems Research* 2006;7:4–22.
- Reichle ED, Warren T, McConnell K. Using E-Z Reader to model effects of higher level language processing on eye-movements in reading. *Psychonomic Bulletin & Review* 2009;16(1):1–20. [PubMed: 19145006]
- Santi, Andrea; Grodzinsky, Yosef. Working memory and syntax interact in Broca's area. *NeuroImage* 2007;37:8–17. [PubMed: 17560794]
- Spivey MJ, Grant ER. Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology* 2005;45:447–481. [PubMed: 12480476]
- Stromswold K, Caplan D, Alpert N, Rauch S. Localization of syntactic comprehension by positron emission tomography. *Brain Language* 1996;52:452–473.
- Swinney D. Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 1979;18:645–659.
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC, et al. Integration of visual and linguistic information in spoken language comprehension. *Science* 1995;268:1632–4. [PubMed: 7777863]

- Tanenhaus M, Maghuson JS, Dahan D, Chambers C. Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *J. Psycholing Res* 2000;29:557–580.
- Traxler M, Morris R, Seely R. Processing Subject and Object Relative Clauses: Evidence from Eye Movements. *Journal Memory Language* 2002;47:69–90.
- Traxler M, Williams R, Blozis S, Morris R. Working memory, animacy, and verb class in the processing of relative clauses. *Journal Memory Language* 2005;53:204–224.
- Treisman A, Gelade G. A feature-integration theory of attention. *Cognitive Psychology* 1980;12:97–136. [PubMed: 7351125]
- Trueswell JC, Tanenhaus M, Garnsey S. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal Memory Language* 1994;33:285–318.
- Turner R. Signal sources in BOLD contrast fMRI. *Advance Experiment Medical Biology* 1997;413:19–25.
- Waters G, Caplan D, Alpert N, Stanczak L. Individual differences in rCBF correlates of syntactic processing in sentence comprehension: Effects of working memory and speed of processing. *Neuroimage* 2003;19:101–112. [PubMed: 12781730]

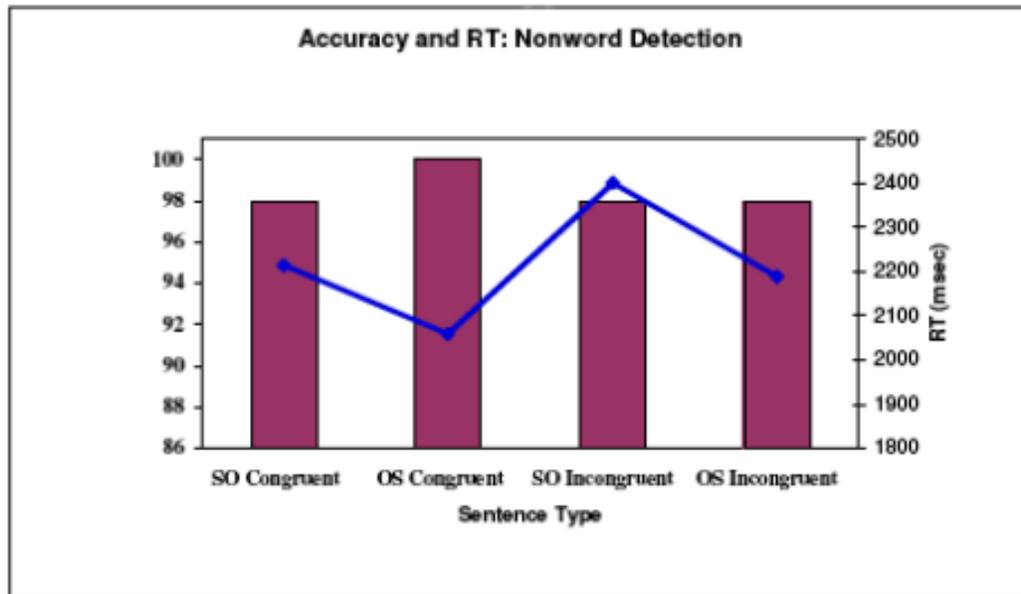
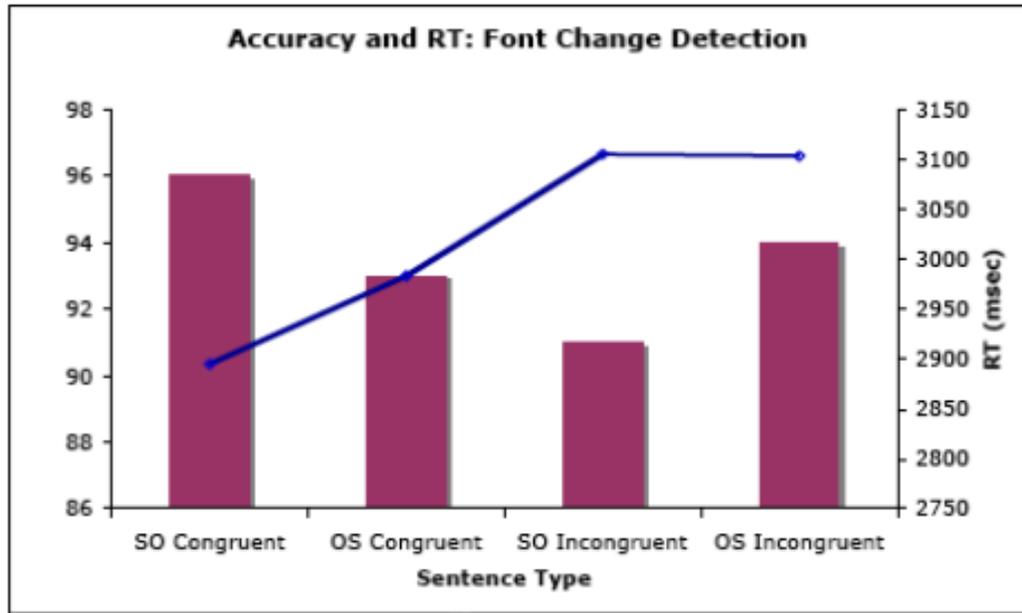


Figure 1. Accuracy (bars) and RTs (line) for different sentence conditions for sentences without font changes or nonwords in the eye tracking study. SO: subject object structure; OS: Object subject structure. Congruent: order of animacy of noun phrases biases towards syntactically determined meaning; Incongruent: order of animacy of noun phrases biases against syntactically determined meaning.

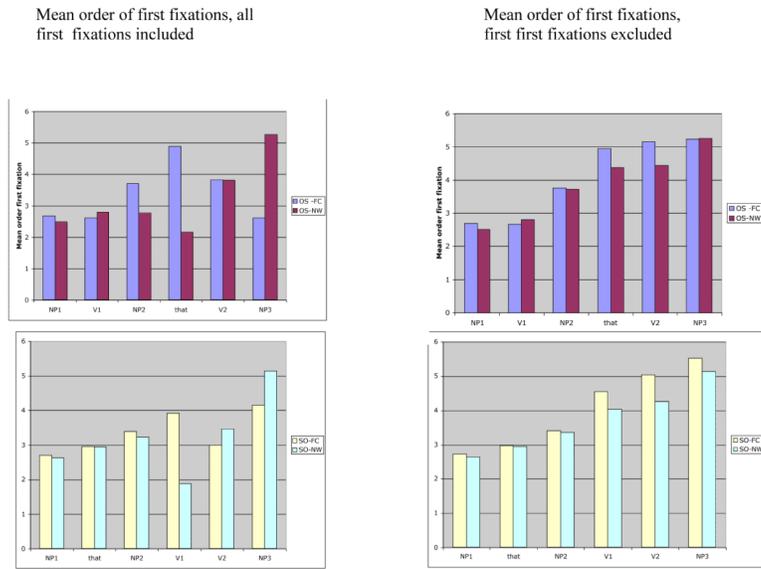


Figure 2.
Order of first fixations on phrases in sentences.

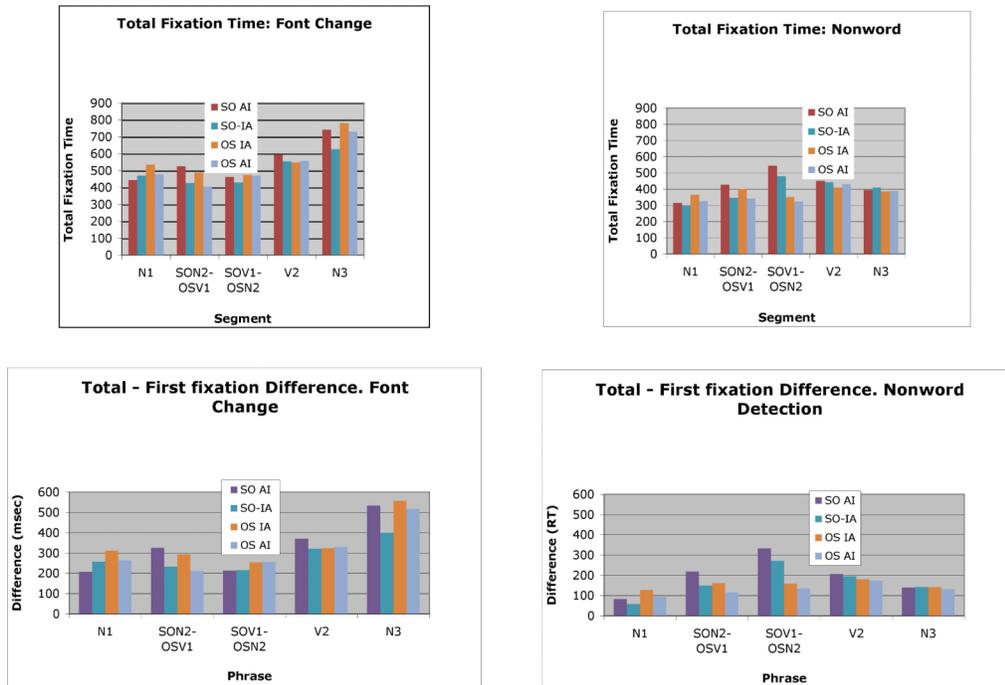


Figure 3. Total fixation time and difference between total and first fixation time on phrases in sentences.

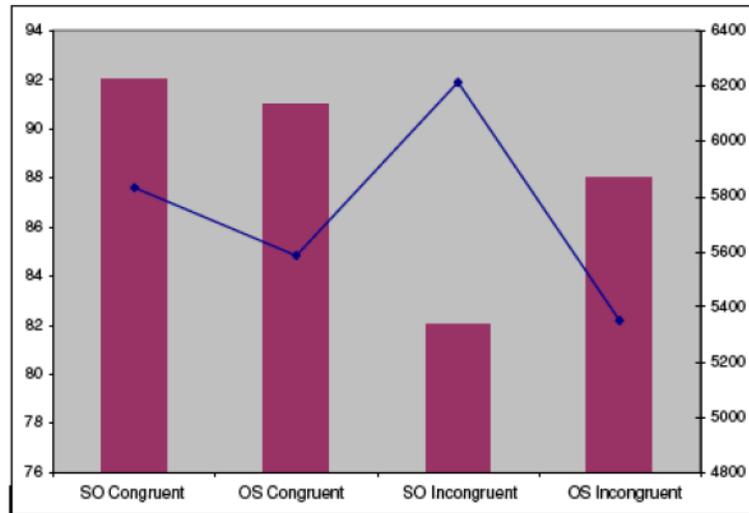


Figure 4. Accuracy (bars) and RTs (line) for different sentence conditions for sentences without font changes. SO: subject object structure; OS: Object subject structure. Congruent: order of animacy of noun phrases biases towards syntactically determined meaning; Incongruent: order of animacy of noun phrases biases against syntactically determined meaning.

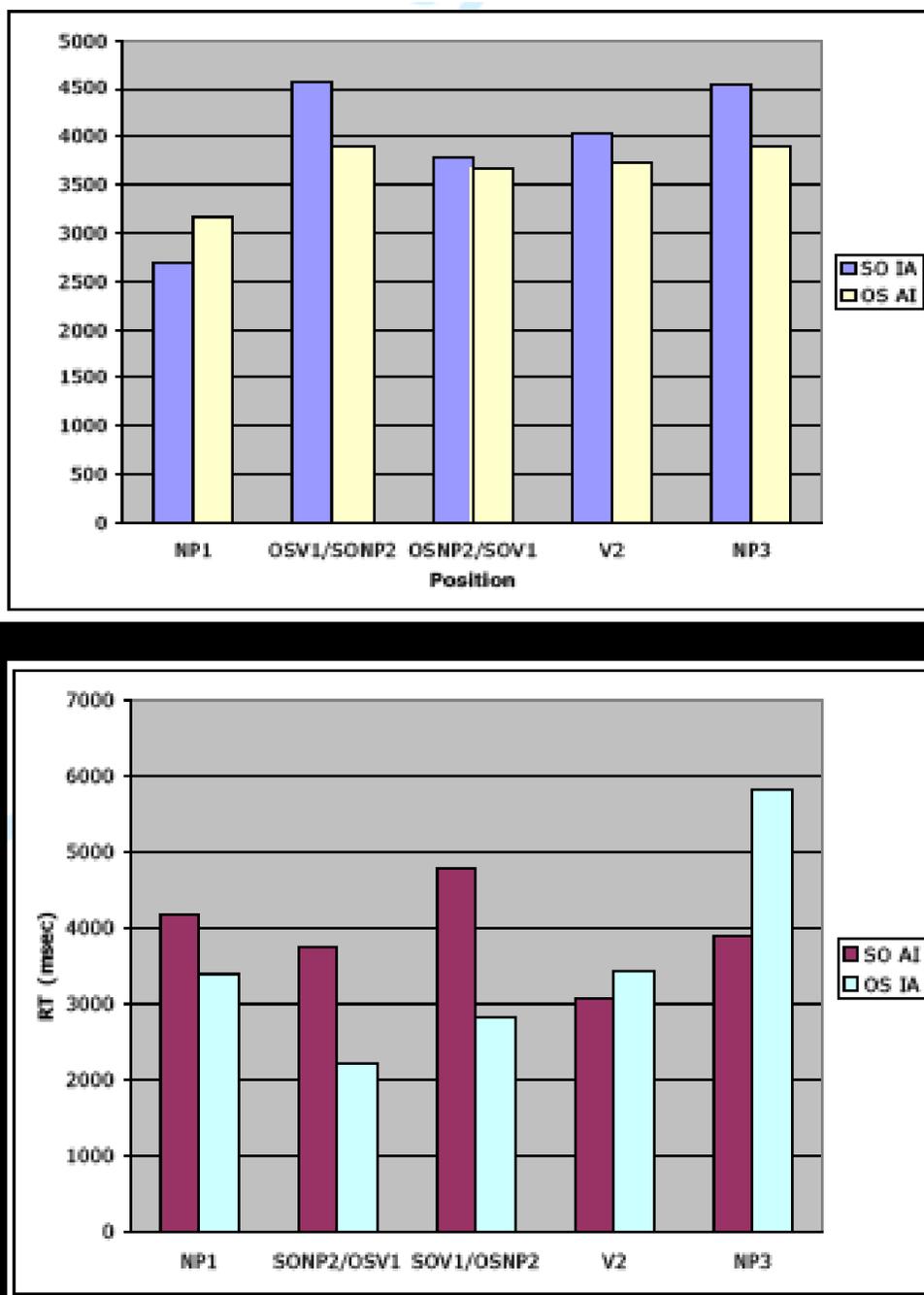


Figure 5. Mean RT for correct responses in for SO and OS sentences containing font changes. Top Panel: sentences in which the order of animate and inanimate nouns biased against the syntactically licensed meaning of the sentences (SO-IA vs OS-IA sentences); Bottom panel: sentences in which the order of animate and inanimate nouns biases towards the syntactically licensed meaning of the sentences (SO-IA vs OS-AI sentences);.

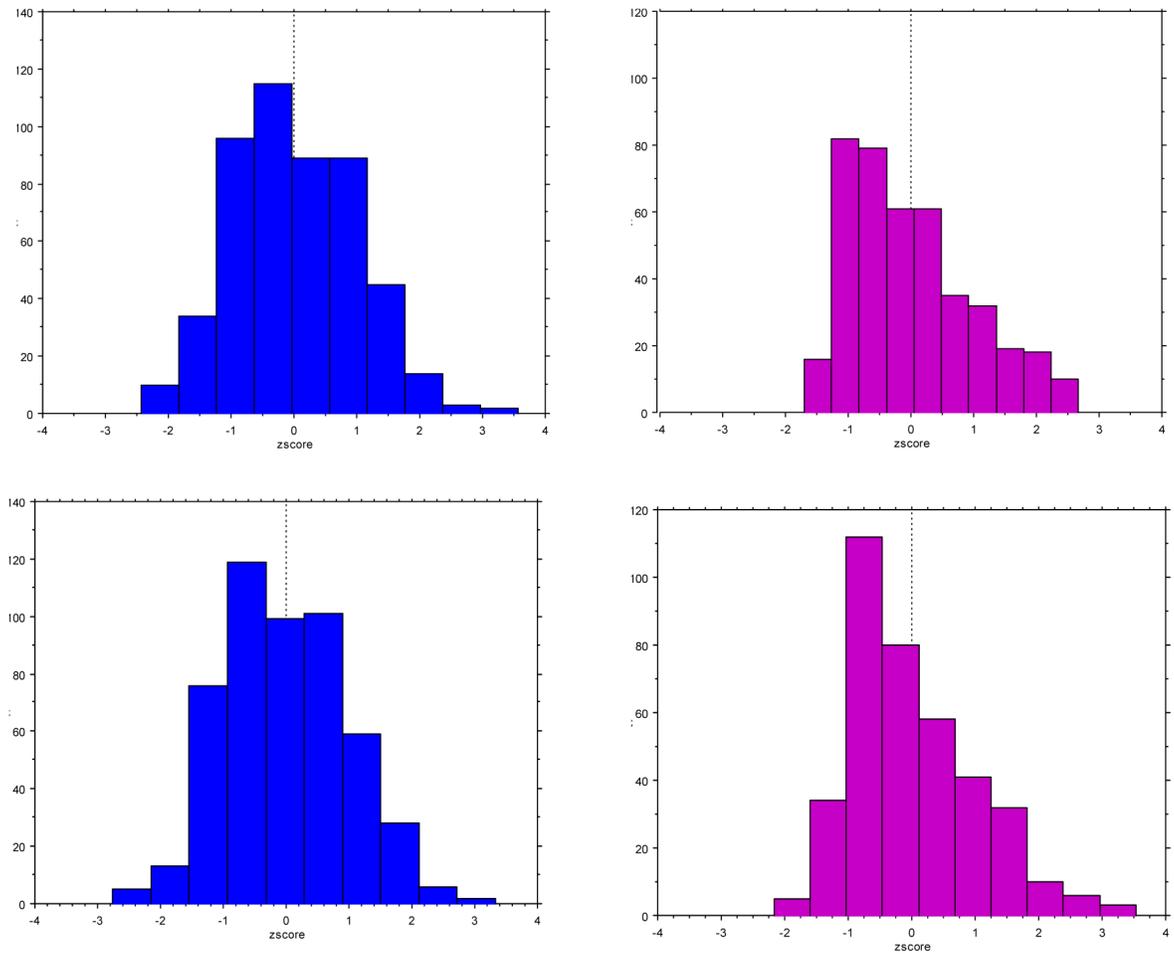


Figure 6. Histograms of RTs for correct responses. Top Panels: object extracted sentences. Bottom Panels: subject extracted sentences. Blue: sentences without font changes; Red: sentences with font changes.

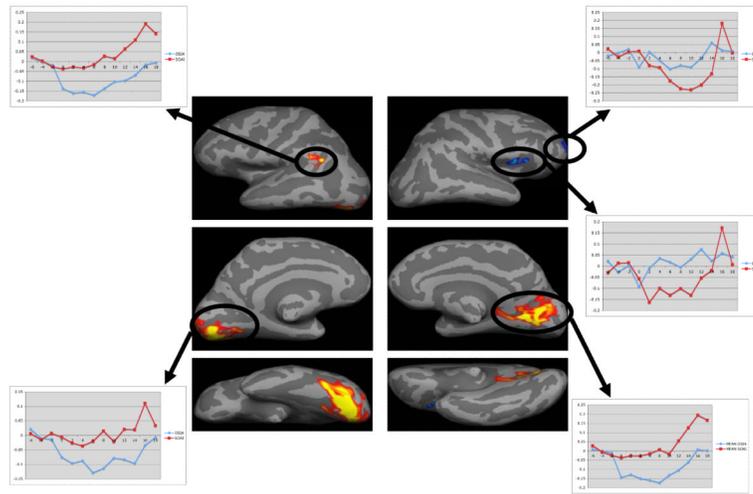


Figure 7. Areas of BOLD signal differences between SO – OS sentences without font changes in which noun animacy order biases against the meaning of the sentence (“incongruent” stimuli: SO-AI and OS-IA). Color overlays represent p-values of the contrast such that the color threshold (red) corresponds to $p = .01$ and ceilings (yellow) at $p = .001$. Each area has a minimum area of 200 mm^2 and a false-positive $p < .05$.

Table 1

Areas of significant activation and deactivation (*) in contrast of SO – OS sentences without font changes in which noun animacy order biases against the meaning of the sentence (“incongruent” stimuli)

	Gyral Location	BA	Max Z score	Size (mm ²)	TalX	Tal Y	TalZ
Left hemisphere	Middle Occipital gyrus	18	5.3	4112	-33	-80	1
	Supramarginal gyrus	39	3.6	357	-52	-55	31
Right hemisphere	Lingual Gyrus	18	5.3	2743	19	-57	5
	*Superior frontal gyrus	10	-3.1	319	20	62	2
	*Inferior frontal gyrus	13	-3.1	343	28	25	7

Table 2

Behavioral effects of sentence contrasts

Task	Measure	Effect at SO-AI Relative Clause Subject	Effect at SO-AI Relative Clause Verb
Font change detection (MR)	Position of font change	Trend	Present
Font change detection (Eye Tracking)	Position of font change	No ^a	No ^a
Font change detection (Eye Tracking)	Total fixation time	Present	No
Font change detection (Eye Tracking)	Total – first fixation time	Trend	No
Nonword detection (MR)	Position of font change	No	Present
Nonword detection (Eye Tracking)	Position of font change	No ^a	Present ^a
Nonword detection (Eye Tracking)	Total fixation time	Present	Present
Nonword detection (Eye Tracking)	Total – first fixation time	Present	Present

^aSee footnote 3