Co-localization of Stroop and Syntactic Ambiguity Resolution in Broca's Area:

Implications for the Neural Basis of Sentence Processing

David January, John C. Trueswell and Sharon L. Thompson-Schill

University of Pennsylvania

3401 Walnut Street

Suite 400A

Philadelphia, PA 19104

Corresponding Author:

David January

3401 Walnut St.

Suite 400A

Philadelphia, PA 19104

Abstract

For over a century, a link between left prefrontal cortex and language processing has been accepted, yet the precise characterization of this link remains elusive. Recent advances in both the study of sentence processing and the neuroscientific study of frontal lobe function suggest an intriguing possibility: The demands to resolve competition between incompatible characterizations of a linguistic stimulus may recruit top-down cognitive control processes mediated by prefrontal cortex. We use functional magnetic resonance imaging to test the hypothesis that individuals use shared prefrontal neural circuitry during two very different tasks – color identification under Stroop conflict and sentence comprehension under conditions of syntactic ambiguity – both of which putatively rely on cognitive control processes. We report the first demonstration of within-subject overlap in neural responses to syntactic and non-syntactic conflict. These findings serve to clarify the role of Broca's area in, and the neural and psychological organization of, the language processing system.

Co-localization of Stroop and Syntactic Ambiguity Resolution in Broca's Area:

Implications for the Neural Basis of Sentence Processing

What mechanisms are involved in understanding sentences? What mechanisms enable us to change or override our characteristic responses to a given situation? A recent proposal (Novick, Trueswell, & Thompson-Schill, 2005) suggests that some of these mechanisms might be shared. The work presented here addresses this question using functional magnetic resonance imaging.

When perceiving or interacting with the world, multiple interpretations of a stimulus are often available. Cognitive control refers to the ability to mediate among these incompatible, competing representations in a goal- or context-relevant manner. A growing body of research associates such control mechanisms with the prefrontal cortex (Badre & Wagner, 2007; Feredoes, Tononi, & Postle, 2006; Miller & Cohen, 2001). At the most general level, these and related studies indicate that prefrontal regions selectively respond to situations of conflict, in which task-specific characterizations of an input are at odds with other characterizations. As proposed by Kan and Thompson-Schill (2004b), a mechanism analogous to that described by the biased competition model for visual selective attention (Desimone & Duncan, 1995) might serve to modulate more abstract (or conceptual) representations.

As an example of the biased competition mechanism operating in higher-order cognition, Thompson-Schill, D'Esposito, and Kan (1999) reported that activity in the posterior left inferior frontal gyrus (PLIFG, specifically Brodmann Area 44) was associated with demands to regulate competition among multiple semantic representations: when subjects generated a verb associated with a noun, PLIFG was less active in response to those nouns for which they had previously generated a verb, but more active to nouns for which they had previously generated a color. This pattern contrasted with the temporal lobe, where activation was reduced for the second

3

presentation of the noun regardless of the prior task. This pattern suggests that even though the retrieval of the semantic information associated with the noun was facilitated by a second presentation (indicated by decreased activation in the temporal lobe), the need for control to override active but task-irrelevant semantic information was increased (indicated by the increased PLIFG activation). Similar findings of PLIFG activity in response to the need to manipulate the activity of task and stimulus representations appear in other domains, such as proactive interference resolution in simple memory tasks (Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Jonides & Nee, 2006; Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003) and the Stroop task, where participants must bias attention toward the color representation of a printed word instead of its meaning (Milham et al., 2001) .

The region within PLIFG that is commonly associated with increasing cognitive control demands is the same brain region that has historically been referred to as Broca's area (BA 44 and 45). In keeping with this long-standing linguistic association, many studies have characterized activation in Broca's area as the result of syntactic and/or morphological processing (e.g., Grodzinsky, 2000; Musso et al., 2003; Sahin, Pinker, & Halgren, 2006). For instance, processing more syntactically complex sentences, such as center-embeddings or object-relative clauses, has been shown to elicit greater Broca's area activity (Caplan, Alpert, & Waters, 1998, and references therein). Such findings have been taken as evidence for Broca's area as the seat of syntactic working memory or syntactic representations themselves.

Such syntactic accounts of Broca's area have difficulty incorporating findings from the cognitive control literature, especially given that many of the cognitive control tasks do not require syntactic processing. Recently, Novick et al. (2005) offered a unifying account of these findings, by noting that sentence processing may frequently require cognitive control. Sentence

processing research has shown that readers and listeners take into account probabilistic evidence from a variety of domains when they structure and interpret an incoming sentence, including frequency of alternative structures in which a word appears (Snedeker & Trueswell, 2004; Trueswell, 1996), fit between a noun's attributes and the role it must play in an event (Garnsey, Pearlmutter, Myers, & Lotocky, 1997), and even the referential context in which the sentence is heard (Altmann & Steedman, 1988; Crain & Steedman, 1985; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell, Sekerina, Hill, & Logrip, 1999). These different evidential sources modulate the activation of representations of the sentence structure and interpretation, often with multiple representations being active simultaneously, with representations receiving support early on sometimes turning out to be incorrect. Cognitive control may be necessary in these situations to bias attention toward one representation and away from another (for further discussion, see Novick et al., 2005). Consistent with this view are past findings from the fMRI and patient literature showing an important role for PLIFG in ambiguity resolution in discrimination of phonetic categories (Blumstein, Myers, & Rissman, 2005), word sense ambiguity (Bedny, Hulbert, & Thompson-Schill, 2007; Bedny, McGill, & Thompson-Schill, 2008; Mason & Just, 2007; Rodd, Davis, & Johnsrude, 2005; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007) and syntactic ambiguity (e.g., Mason, Just, Keller, & Carpenter, 2003) .

Taken together, these literatures suggest that syntactic and non-syntactic conflict resolution rely on general cognitive control mechanisms subserved by PLIFG. It is possible however that there exists specialization within PLIFG for different types of conflict and/or different types of linguistic operations. Recent anatomical investigations of the structure of Broca's area indicate that there are at least three and possibly four distinct cytoarchitectural

patterns (i.e., "areas") within this region of cortex (Amunts et al., 1999). Indeed, there is some controversy regarding whether the localization of cognitive control abilities is in PLIFG or instead in the caudally adjacent left inferior frontal junction (Derrfuss, Brass, & von Cramon, 2004; Derrfuss, Brass, Neumann, & von Cramon, 2005). Additionally, there are several proposals for a specialization by content within PLIFG, with, for example, some regions responsible for semantic processing and some regions responsible for phonological processing (e.g., Devlin & Watkins, 2007; Fiez, 1997; Gough, Nobre, & Devlin, 2005; though see Barde & Thompson-Schill, 2002).   In light of these facts, we decided to examine the extent to which substantially different sorts of conflict co-localize within BA 44/45 *within each individual*, rather than relying on previous descriptions of task localization.

Here we explore this possibility by looking for fMRI signatures of cognitive control during the comprehension of ambiguous sentences in a visual referential context. Specifically, we presented sentences containing a prepositional phrase (PP) that could either denote an instrument with which to carry out an action or serve as a modifier describing the object to act on. We varied the scene to modify contextual support for each analysis, thereby creating a parametric manipulation of the amount of conflict among the interpretations, as in Figure 1. The weakest level of the series pairs a syntactically unambiguous sentence (using a relative clause instead of a PP) in condition S1 with a context containing no good instruments for the verb. The next level, condition S2, introduces the ambiguous PP with the same context. The strongest level, condition S3, changes the context to include a good instrument of the verb, lending support to an instrument analysis of the PP. If cognitive control mechanisms are deployed during syntactic ambiguity resolution, this situation would require cognitive control to bias activation away from the instrument interpretation when bottom-up information increases such activation.
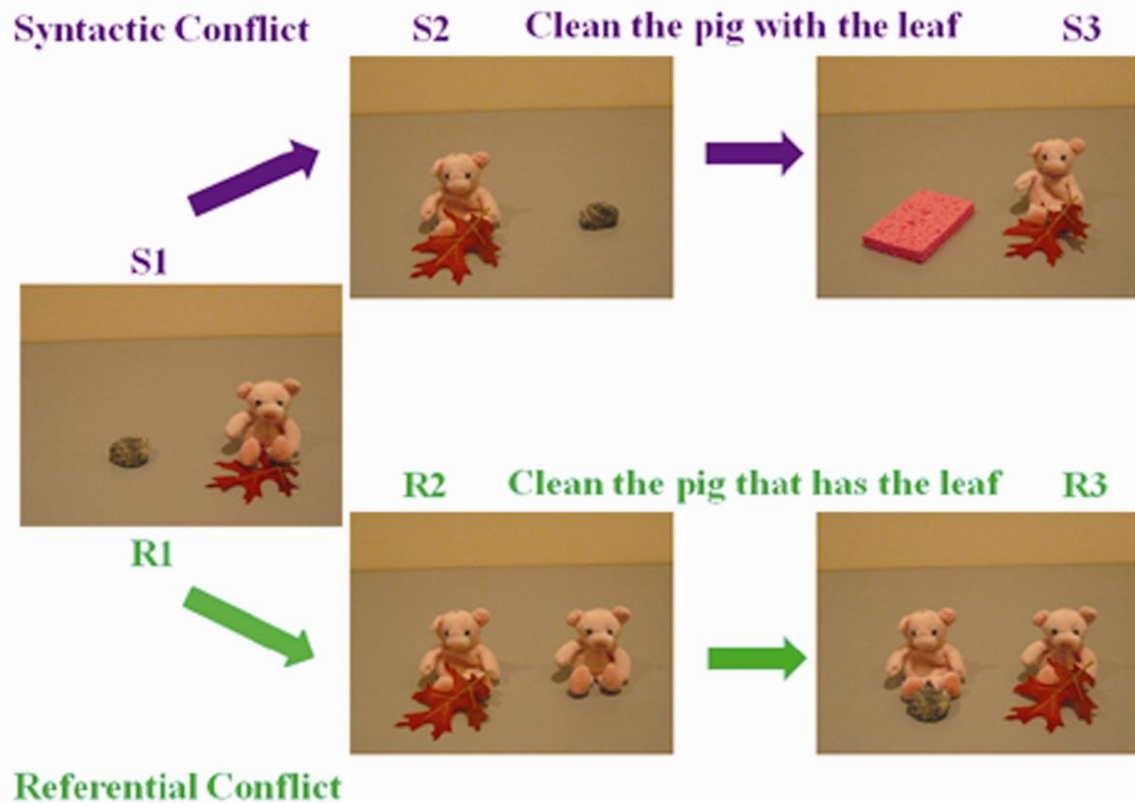
Figure 1. Sample visual display accompanying each sentence for each condition. See text for description of each trial.

It is an open question how domain general the proposed PLIFG conflict resolution might be. The bulk of the neuroimaging work to date advocating for such a view of PLIFG function has been concerned with control of semantic memory (e.g., Badre & Wagner, 2007; Kan & Thompson-Schill, 2004a; Thompson-Schill et al., 1999) or working memory (e.g., Feredoes et al., 2006; Jonides & Nee, 2006). To explore the breadth of the proposed conflict resolution mechanism, we also generated a parametric series of conflict in the referential domain. The weakest level, condition R1, is identical to S1. Condition R2 introduces a temporary ambiguity for the referent of the noun used in the sentence by introducing another potential referent (but crucially keeps the sentence syntactically unambiguous). Condition R3 extends the referential ambiguity yet further in time by making the final word in the relative clause the only

7

distinguishing characteristic (where the existence of the relative clause itself in R2 can serve to disambiguate). The proposed cognitive control mechanism would operate in these conditions to push the system to choose a referent when the bottom-up information is insufficient to do so (Botvinick, Braver, Barch, Carter, & Cohen, 2001).

We additionally administered the Stroop task, long recognized as a prototypical cognitive control task, to look for co-localized activity within each subject. Specifically, we administered the modified Stroop task from Milham et al. (2001). In this task, subjects use a button box to indicate the color that a word is displayed in, with only three buttons available. The response set is therefore limited to only three colors. The words subjects see fall into four conditions. In the first, the response-ineligible incongruent condition, the printed word corresponds to a color not in the response set, which creates conflict at the level of the representation of the stimulus. In the second, the response-eligible incongruent condition, the printed word corresponds to a color that is in the response set, which creates conflict at both the representational level and also the response level. These incongruent conditions are compared to a neutral condition in which the distracting word is a length- and frequency-matched non-color term. This design allows for the isolation of conflict at the representational level, the level at which we predict conflict during language processing (see also Badre & Wagner, 2006).[1]

Method

Participants

Seventeen members of the University of Pennsylvania community (11 female; aged from 18.5 to 34.5 years) participated in the study. All were right handed, had normal or corrected-to-normal

---

[1] The Milham et al. (2001) Stroop task has received some criticism, suggesting that the response-ineligible incongruent manipulation does not isolate representational conflict (van Veen & Carter, 2005). We delay discussion of this concern to the General Discussion.

vision, had no history of head trauma, were not on psychoactive medications, and spoke only English through the age of 5. Participants were paid $20/hr for participation.

Stimuli

Sentence comprehension task. Sentence comprehension stimuli were photographs of toy objects on a neutral background paired with recorded instructions directing the participant to act on the objects in the photographs. Target stimuli consisted of 150 sentence—picture pairings divided evenly among 5 conditions designed to parametrically manipulate the degree of conflict: in the weak conflict condition (S1), a syntactically unambiguous sentence such as "clean the pig that has the leaf" was paired with a picture of a pig holding a leaf, sitting next to a rock (see Figure 1). For the middle level of conflict (S2), the target sentence was replaced with an instruction containing a PP attachment ambiguity ("clean the pig with the leaf"), where the PP denotes either an instrument of the verb (meaning "use the leaf to clean the pig") or a modifier of the noun (meaning "clean the pig that has the leaf"). In both S1 and S2, the other objects on screen were poor instruments of the verb (Snedeker & Trueswell, 2004), thus providing poor support for the instrument interpretation of the PP. For level 3 of syntactic conflict (S3), the second object (the rock) was replaced with a good instrument of the verb (a sponge), providing contextual support for the instrument interpretation of the PP. Since all the target verbs were biased to appear with a subsequent PP denoting an instrument (Snedeker & Trueswell, 2004), this series increased conflict by increasing the support for, and therefore activation of, an instrument interpretation of the sentence from incompatible (S1) to strongly contextually supported (S3). Normative data collected from an independent group of subjects instructed to act out these instructions (N = 30) confirm that these trials elicited both interpretations of the ambiguity, with greater conflict as the proportion of responses consistent with the instrument

9

interpretation of the PP increased with level (S1: 11%, S2: 38%, S3: 48%). All other actions were modifier actions (in which the participant used his/her hand to act on the object rather than using the instrument). These norming data also demonstrate that our stimuli were not disambiguated by prosody. The speaker who recorded our stimuli was trained to avoid major prosodic breaks and thus avoid disambiguation (e.g., Snedeker & Trueswell, 2003).

For the referential conflict series, the weak conflict condition (R1) was identical to S1. For the middle level of referential conflict (R2), the second object (the rock) was replaced with another animal of the same category that was not holding anything (another pig). This introduced a temporary referential ambiguity for the phrase "the pig" which was resolved when the sentence continued "that has", indicating that the target was an animal holding something. For the strongest level of referential conflict (R3), the second animal also held something, delaying the disambiguation of the referent of "the pig" until the last word of the sentence is encountered ("the leaf" versus "the rock").

Each subject saw all target trials in all conditions.

Because the target trials always directed the subject to act on an animal holding an object, we included 124 filler trials that were designed to direct the subject to act on animals not holding anything and also on inanimate objects to maintain unpredictability (and thus ambiguity) in the target trials. Finally, on 31 catch trials, the target of the action described by the sentence was not present in the picture (see Procedure).

Stroop task. Items for the Stroop task were based on those of Milham et al. (2001). Responses were made via a button box and were restricted to yellow, green, and blue. Stimuli were composed of four types: response-eligible conflict, response-ineligible conflict, and two groups of neutral trials. In response-eligible conflict trials, the distracting word denoted a color

that was a potential response (i.e., YELLOW, GREEN, or BLUE). In response-ineligible trials, the distracting word denoted a color that was not a potential response (i.e., ORANGE, BROWN, or RED). This manipulation allows for the separation of conflict at a response level, where the meaning of the printed word might lead a participant to push a wrong button, and conflict at the representational level, where the meaning of the printed word is not in the response set but still competes with the font color. Neutral trials were composed of non-color terms length- and frequency-matched for the terms used in the conflict trials. For example, the words PLENTY, HORSE, and DEAL were paired with the response-eligible incongruent trials (and are henceforth called response-eligible neutral) and the words FARMER, STAGE, and TAX were paired with the response-ineligible trials (and henceforth called response-ineligible neutral). For all conditions, the font color was part of the target response set.

Procedure

Subjects completed the sentence comprehension task first, divided into four imaging runs each lasting approximately ten minutes[2]. Order of presentation of conditions (including null events) was pseudo-randomized for each subject, optimized for statistical power by OptSeq, with the constraint that no two conditions involving the same item appeared within 6 items of each other. Subjects were told to listen to the sentences and imagine carrying out the actions on the pictured objects as vividly as possible. They were told that they would occasionally encounter an instruction they could not carry out because an object would be missing from the picture and to push a button on the response pad when this occurred (catch trials). Each trial lasted six seconds (two TRs) to allow adequate time to imagine carrying out the task. Stimulus presentation was controlled by E-Prime software (Psychology Software Tools, Inc). Subjects were shown two

---

[2] 3 Subjects completed only 3 runs of the sentence comprehension task.

sample trials, including one catch trial, neither of which were used in the main experiment, before entering the magnet. After the sentence comprehension task, subjects completed the Stroop task, keeping as close as possible to the procedure in Milham et al. (2001). Trials in the Stroop task were arranged in a mixed blocked and event-related design such that subjects saw one block of 48 trials composed of only response-eligible conflict trials and neutral trials and then, following a 12 second break, one block of 48 trials composed only of response-ineligible conflict trials and neutral trials. Within each block, the order of presentation of conditions was individually determined for each subject using OptSeq. Following another 12 second break, these blocks were repeated, with new optimized sequences. As in Milham et al., the sequence of blocks was the same for all subjects.

In the scanner, sentences were played over Confon electrodynamic headphones and images projected onto a screen at the top of scanner bore, viewed on a mirror mounted on the headcoil.

fMRI image acquisition

Following the acquisition of axial T1-weighted localizer images (TR = 1620 msecs, TE = 3 msecs, TI = 950 msecs), gradient echo, echoplanar fMRI was performed in 46 contiguous 3-mm axial slices (TR = 3000 msecs, TE 30 msecs, 64 x 64 pixels, Field of view 19.2 cm, voxel size 3 x 3 x 3 mm) using a 3T Siemens Trio system and 8 channel array headcoil. Twelve seconds preceded data acquisition in each run to approach steady-state magnetization.

Image processing

Off-line data processing was performed using VoxBo software (www.voxbo.org). After image reconstruction, normalization, and motion correction, the data were sinc interpolated in time to

correct for the fMRI acquisition sequence and spatially smoothed with a kernel with FWHM of 3 voxels.

Image analyses

At each voxel, general linear models modified for serially correlated error terms (Worsley & Friston, 1995) and containing estimates of intrinsic temporal autocorrelation (Aguirre, Zarahn, & D'Esposito, 1997) and a covariate capturing global signal variation were applied to data from both the sentence comprehension task and the Stroop task. The model for the sentence comprehension task included a covariate for each target condition, randomly assigning 15 of the 30 1-level conflict items to either the syntactic conflict series or the referential conflict series, as well as a covariate for catch trials and filler trials. Error trials (incorrectly pressing a button in response to an instruction that could be carried out, failing to press a button in response to a catch trial) were assigned to a common separate covariate[3]. For the Stroop task, we constructed a separate model that included covariates for each condition (response-eligible incongruent, neutral for response-eligible, response-ineligible incongruent, neutral for response-ineligible) and a covariate for errors.

Individualized ROI Definition

For all subjects, a language-related region of interest for the sentence comprehension task was defined by isolating voxels that passed a threshold of $F = 3.0$ for an F-test testing for all six covariates simultaneously differing from zero ("The sentence comprehension ROI"). We then applied an anatomical constraint requiring voxels to fall in either pars opercularis or pars triangularis (BA 44/45) within PLIFG. These anatomical regions were defined by manually selecting all voxels in opercularis and triangularis on high-resolution anatomical scans for each

---

[3] For one subject, the button box was not connected to the computer during the sentence comprehension task. It was connected before starting the Stroop task.

subject blind to the activation pattern within the subject. This resulted in one subject having no mask, and he was subsequently dropped from all sentence-comprehension analyses. We did not include in this ROI voxels in pars orbitalis (BA 47) as this region has been implicated in other types of cognitive control (Badre & Wagner, 2007).

Additionally, we defined two Stroop-related ROIs. The first ("the Stroop ROI") was defined as those voxels in PLIFG which passed a threshold of F= 3.0 for all target Stroop conditions against baseline. Again, one subject did not have any voxels meeting this criterion and was dropped from analysis in it. Finally, to localize representational conflict, we isolated those regions of PLIFG responsive in the Stroop task to response-ineligible conflict (response-ineligible incongruent minus response-ineligible neutral), thresholding at t = 1.8 ("the representational conflict ROI").

Summary statistics for the ROIs are presented in Table 1.

ROI	N	Mean Voxels (SD)	Median Voxels	Min. Voxels	Max. Voxels
Sentence comprehension	16	142 (102)	145.5	10	328
Stroop	16	189 (226)	81	6	633
Representational conflict	17	73 (71)	72	1	253

Table 1. The size of each of the individualized ROIs used in the analysis. The sentence comprehension ROI was defined as those voxels in PLIFG that exhibited an F-score greater than or equal to 3 for a contrast of all target conditions (R1, R2, R3, S1, S2, S3) compared to fixation baseline. The Stroop ROI was defined as those voxels in PLIFG that exhibited an F-score greater than or equal to 3 for a contrast of all target Stroop conditions compared to fixation baseline. The representational conflict ROI was defined as those voxels in PLIFG that exhibited a t-score of 1.8 or greater for the contrast of response-ineligible incongruent trials minus neutral trials in the Stroop task. One subject did not have any voxels in PLIFG with an F greater than or equal to 3 for the language task and was dropped from all analyses in this region.

Whole-brain analysis

To test for the specificity of our co-localization effects, we conducted a post-hoc random-effects analysis of the incongruent minus neutral contrast and tested for effects of the syntactic and referential manipulations using all regions that passed a threshold of t = 3.25 ($p < .01$, uncorrected) with a minimum cluster size of 15 voxels (see Table 3 for a list of region locations,

and Figure 4 for the location of two particularly relevant ROIs). These regions derived from the group-level analysis were used as ROIs parallel to the individualized ROIs used in the analyses described above.

Results and Discussion

Behavioral Results

In the sentence comprehension task, all subjects correctly responded to at least 74% of catch trials. All target trials on which a subject incorrectly pressed a button, less than 3% of trials in any target condition, were removed from subsequent analyses. The only behavioral measure for the subjects in the scanner is the rate at which they incorrectly press a button, which they are only supposed to do if an instruction is impossible to follow, in response to a target instruction (all of which were possible). A repeated-measures two-way ANOVA on the log-odds transform of this error rate with conflict type and level as factors revealed no significant effect of condition ($F(1,16) = .09$, $p > .7$) or level ($F(2,32) = 1.67$, $p > .2$) on false alarm rate.

In the Stroop task, all subjects responded correctly to 90% or more of trials. RTs (excluding errors) that were more than 3SD above each subject's grand mean were trimmed to the cutoff value, affecting 0.2% of the data. Mean RTs (response-eligible/incongruent 769; response-eligible/neutral 670; response-ineligible/incongruent 719; response-ineligible/neutral 665) were entered into a repeated-measures 2 x 2 ANOVA with Response Type (response-eligible, response-ineligible) and Trial Type (incongruent, neutral) as factors, revealing a main effect of Trial Type ($F(1,16) = 53.90$, $p < .01$), with no effect of Response Type ($F(1,16) = 1.85$, $p > .1$). Additionally, the ANOVA revealed a significant interaction between Response Type and Trial Type ($F(1,16) = 7.10$, $p < .05$), with eligible/incongruent trials slower than ineligible/incongruent ($t(16) = 2.67$, $p < .05$).

Results of the individualized ROI analysis are summarized in Figure 2 and Table 2. We found a nearly linear increasing response to predicted syntactic conflict in the sentence
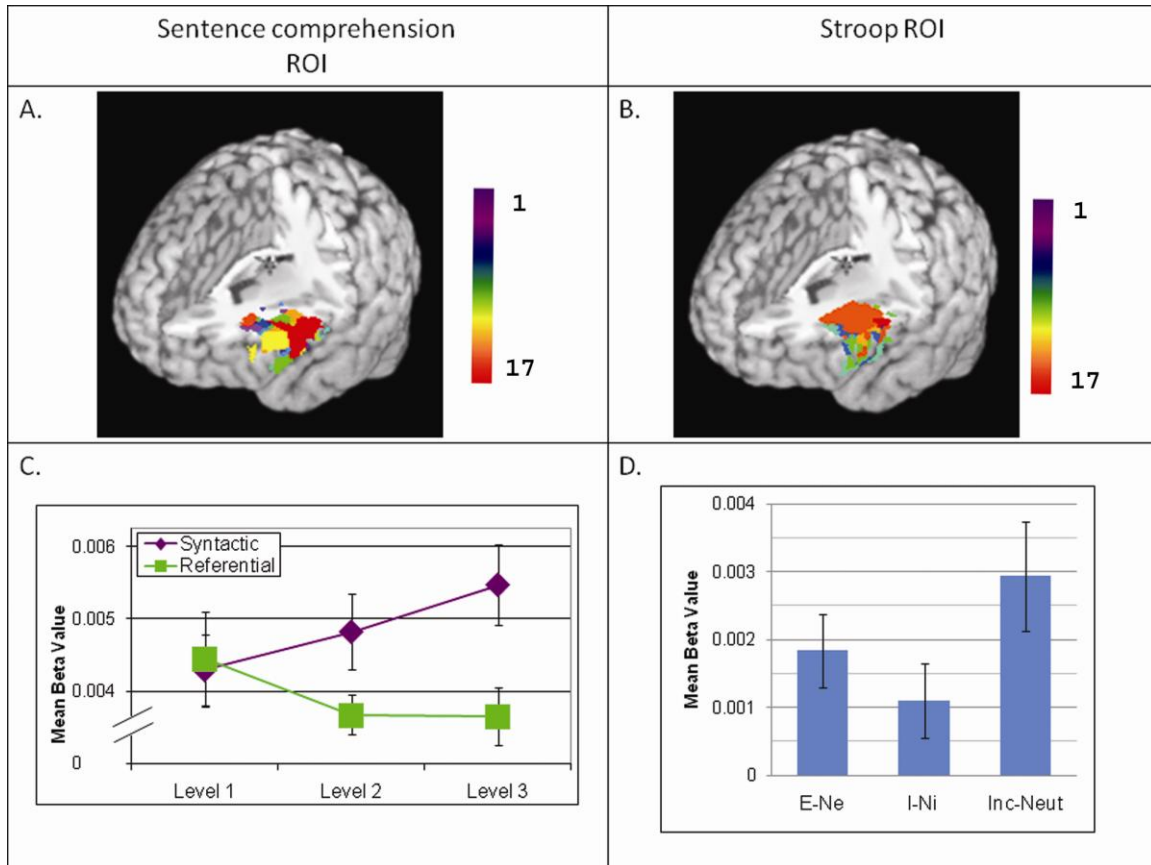

Figure 2. Results for the ROI analysis, primary effects. A. Map showing overlap across subjects for sentence comprehension ROI on a standardized brain. Color code indicates how many subjects had that voxel included in ROI. B. Map showing overlap across subjects for representational conflict ROI on a standardized brain. Color code indicates how many subjects had that voxel included in ROI. C. Mean beta estimates for each sentence comprehension target condition compared to fixation baseline in sentence comprehension ROI. Error bars indicate standard error of the mean. D. Mean beta estimates for each conflict type in the Stroop task. E-Ne: Response-eligible incongruent minus neutral. I-Ni: Response-ineligible incongruent minus neutral. Inc-Neut: Incongruent (averaged over response eligibility) minus neutral. Error bars indicate standard error of the mean.

comprehension ROI, averaging over all suprathreshold voxels, with significantly higher response in S3 than in S1. An ANOVA on the beta values for each of the syntactic conflict target conditions revealed a significant linear trend ($F(1,15) = 16.58$, $p < .01$), indicating a statistically reliable increase in PLIFG activation as the level of conflict increased. We additionally found an

effect of representational conflict in the Stroop ROI with activation in response to response-ineligible incongruent Stroop trials significantly higher than response-ineligible incongruent trials. These results are consistent with the findings of Milham et al. (2001) and previous literature finding effects in PLIFG of syntactic complexity (e.g., Caplan et al., 1998).

However, it is possible that there is specialization within PLIFG for these different conflict types (Syntactic Conflict vs. Stroop Conflict). We therefore tested for the syntactic conflict effect within the representational conflict ROI from the Stroop task and found a similar increase in activation with increasing syntactic conflict (Fig. 3b). An ANOVA on the beta values in this ROI for each of the syntactic conflict conditions revealed a significant linear trend ($F(1,16) = 6.25$, $p < .05$). Additionally, we found a marginally significant effect for response-ineligible incongruent minus response-ineligible neutral in the sentence comprehension ROI. We additionally found significant effects for the response-eligible incongruent minus neutral and for

ROI	N	Contrast	Mean Beta	t-value	p-value	
Sent. comp.	16	R2-R1	-0.001	-1.28	> .20	
Sent. comp.	16	R3-R2	-0.00002	-0.06	> .90	
Sent. comp.	16	R3-R1	-0.001	-1.07	> .30	
Sent. comp.	16	S2-S1	0.001	1.14	> .25	
Sent. comp.	16	S3-S2	0.001	1.20	> .20	
Sent. comp.	16	S3-S1	0.001	4.07	< .01	*
Sent. comp.	16	E-Ne	0.002	3.60	< .01	*
Sent. comp.	16	I-Ni	0.001	1.84	< .09	~
Sent. comp.	16	Inc-Neut	0.002	3.60	< .01	*
Stroop	16	E-Ne	0.002	4.99	< .05	*
Stroop	16	I-Ni	0.001	2.00	< .05	*
Stroop	16	Inc-Neut	0.003	5.72	< .05	*
Rep. conf.	17	S3-S1	0.001	2.50	< .05	*
Rep. conf.	17	R3-R1	0.0001	0.39	> .70	

Table 2. Summary table of results for target contrasts. *Sentence comprehension task*: **R1: referential conflict, level 1. R2: referential conflict, level 2. R3: referential conflict, level 3. S1: syntactic conflict, level 1. S2: syntactic conflict, level 2. S3: syntactic conflict, level 3.** *Stroop task*: **E: response-eligible incongruent. Ne: neutral paired with response-eligible incongruent. I: response-ineligible incongruent. Ni: neutral paired with response-ineligible incongruent. Inc: incongruent trials pooled across response eligibility. Neut: neutral trials pooled across response eligibility. *: significant at $p < .05$. ~: marginally significant. Sent. comp.: sentence comprehension ROI. Rep. conf.: representational conflict ROI.**

all incongruent minus neutral in the sentence comprehension ROI, which is not surprising given that in all of these conditions there is representational conflict, with the response-eligible conflict trials adding response conflict as well.

This co-localization *within each subject* strongly argues for a shared mechanism across these tasks. Figure 3c shows the number of subjects who had a given voxel pass a threshold of t = 1.8 for both the S3-S1 contrast and the response-ineligible incongruent minus response-ineligible neutral contrast; every highlighted voxel indicates that at least one subject showed co-localization in that voxel. As can be seen in the figure, a fairly sizable region of the superior portions of BA 44 had, for any particular voxel, 14 or more of the 17 subjects (orange to red)
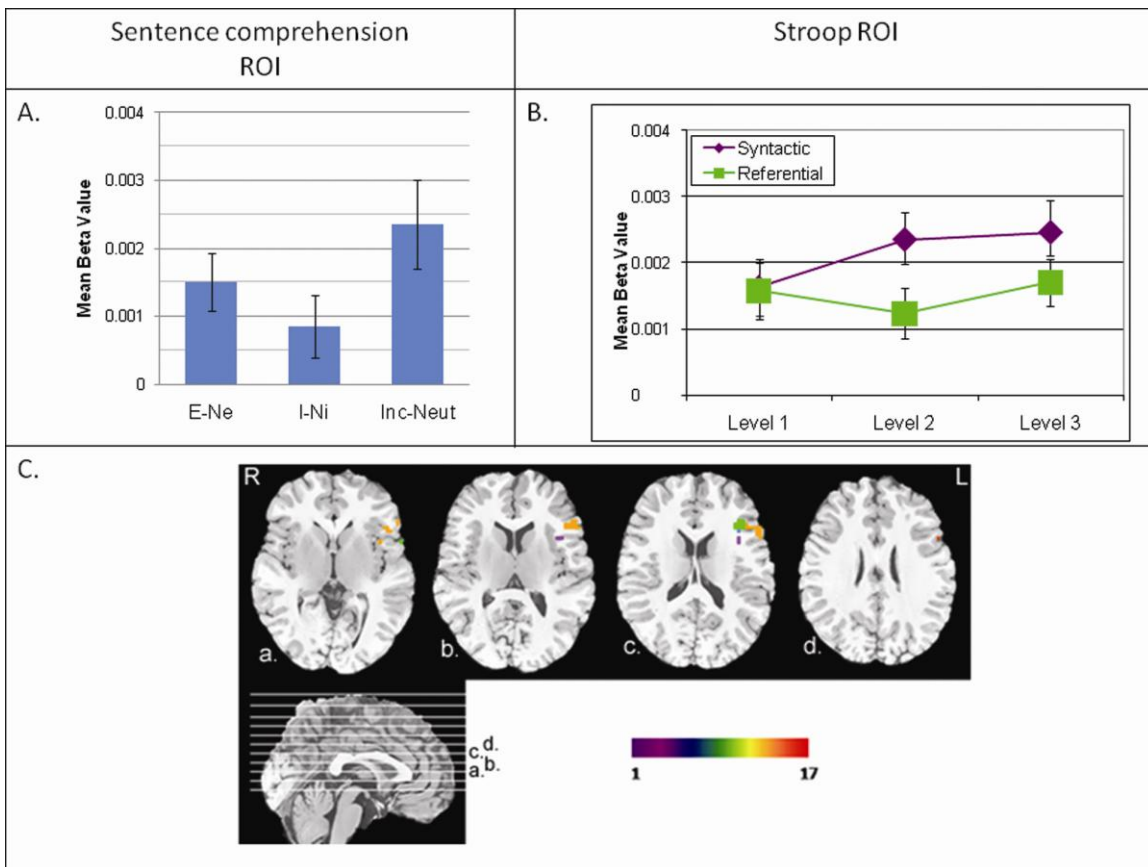

Figure 3. Figure 3. Co-localization analyses. A. The effects of Stroop conflict in the sentence comprehension ROI. B. The effects of conflict in the sentence comprehension task in the representational conflict ROI. C. Overlap of sentence comprehension ROI and representational conflict ROI. Letters next to sagittal slice indicate location of corresponding axial slices displayed above. Color code indicates the number of subjects who had a given voxel pass a threshold of t = 1.8 for both S3-S1 and I-Ni.

showing co-localization of syntactic and Stroop conflict.

Interestingly, we did not find a monotonic effect for increasing referential conflict in any ROI, suggesting a different mechanism is employed in these conditions. The conclusion that referential conflict resolution does not rely on the same mechanism as syntactic and Stroop conflict resolution must be tentative at this time, however. Among the non-target trials in the sentence comprehension task, we included nine items that were globally referentially ambiguous (e.g., *pat the dog* paired with a context with two dogs, one holding a tube) to distract subjects from the target manipulation. Since this condition represents an extreme form of referential ambiguity, we analyzed these trials to test for effects of referential ambiguity in PLIFG at its extreme. We found significantly more activation to these globally ambiguous fillers than any level of referential conflict in both the sentence comprehension ROI (all mean differences: 0.002, all $t(14)$s > 2.4, all ps < .05) and representational conflict ROI (all mean differences: 0.002, all

Location	Number	BA	Size (voxels)	X	Y	Z	S3-S1	R3-R1
L. cerebellum, inferior semi-lunar lobule			175	-18	-69	-39	> .6	> .3
R. cerebellum, inferior semi-lunar lobule			74	39	-75	-39	> .4	> .8
L. precentral gyrus		6	358	-30	2	33	> .6	> .6
L. inferior/ middle frontal gyrus	1	44/46	18	-42	33	15	< .05	> .1
L. inferior frontal gyrus		47	42	-63	13	24	> .3	> .8
R. medial frontal gyrus	2	8	15	9	26	46	< .005	> .5
R. inferior frontal gyrus		47	248	48	11	27	> .2	> 5
R. insula			22	33	18	2	> .8	> .2
			38	69	15	2	> .3	> .2
			51	54	25	-40	> .3	> .1
L. precuneus		7	38	-30	-45	41	> .1	> .7
R. precuneus		7	85	6	-65	50	.054	> .5
R. superior parietal lobule		7	34	33	-54	47	> .9	> .3

Table 3. Talairach coordinates of peak coordinates of ROIs identified in post-hoc analysis and p-values of syntactic and referential contrasts in them. ROIs are all regions identified via the incongruent minus neutral contrast from the Stroop task surpassing a threshold of t = 3.252 (p = .0025, uncorrected) with a minimum cluster size of 15. Number indicates the label for the region on Figure 4.

$t(15)\text{s} > 3.1$, all $p\text{s} < .05$[4]. However, given the small number of these items, we are reluctant to draw strong conclusions from this post-hoc test. Nevertheless, consistent with the idea that referential conflict may recruit different mechanisms, a visual inspection of a map showing voxels responsive to both syntactic ambiguity (S3 minus S1) and this global referential ambiguity (compared to R1) revealed that these voxels were smaller in number and distributed more ventrally than the overlap effects in Fig. 3c.
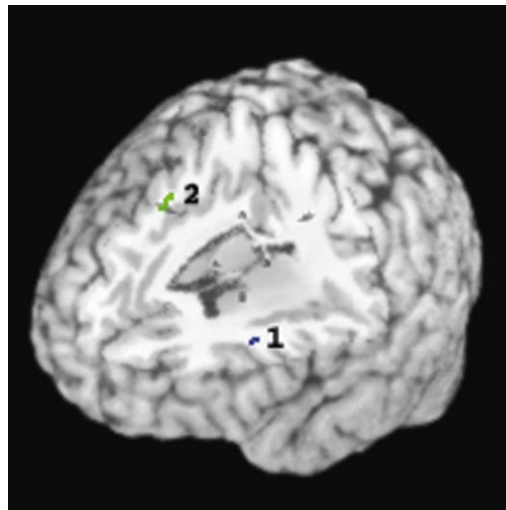

Figure 4. Locations of the left inferior frontal (1) and medial frontal (2) ROIs identified in the whole-brain analysis (see Table 3).

The results of the whole-brain analysis show that this co-localization of conflict resolution mechanisms is relatively restricted to the PLIFG, with the only significant co-localizations being on the dorsal extent of PLIFG, moving into the MFG, and in a medial frontal region very near the pre-SMA (see Table 3 and Figure 4 for results and ROI locations). Since the S3 condition had more behavioral responses using an instrument in the norming group that acted out these instructions, this medial frontal activation is consistent with findings that the pre-SMA plays a role in response planning and selection (e.g., Mostofsky & Simmonds, 2008).

---

[4] One subject responded to each globally ambiguous filler as a catch trial, leaving only 15 and 16 subjects for the language ROI and Stroop ROI, respectively.

General Discussion

These results strongly suggest that cognitive control mechanisms are recruited during syntactic ambiguity resolution. As the predicted need for cognitive control rose in the syntactic series, so too did activity in PLIFG, in the same region that responded to representational conflict in the Stroop task. This co-localization of activation places strong constraints on its interpretation. Given that there is little or no syntactic processing in the Stroop task (and no syntactic *conflict*), the finding of increased PLIFG activation in the syntactic manipulation is most naturally attributed to a shared conflict resolution mechanism and not to additional syntactic processing associated with ambiguity. Thus, at a minimum, the mechanism involved must at least apply beyond the syntactic domain and may be even more domain-general than we can test here. The null result in the target referential manipulation is potentially instructive of the limits of the domain generality of the PLIFG conflict resolution mechanism: it does not appear to extend to the conflict embodied in this manipulation. One possible explanation for the pattern of results is that BA 44/45 may be responsible for resolving conflict of linguistic representations (phonological, syntactic & semantic) but not conflicts in the mapping of linguistic input onto the visuo-spatial representations of the world. Recall that in the syntactic series there is no ambiguity regarding the mapping of the noun phrases in the utterance (e.g., "the bear" and "the leaf") onto the referent world (e.g., all scenes in this series contained exactly one bear and one leaf). Rather the ambiguity in this series was syntactic/semantic in nature, regarding the way in which the phrase "with the stick" is parsed and interpreted. Crucially, the fact that this syntactic series co-localizes with Stroop within each individual indicates that BA 44/45 is not exclusively responding to syntactic conflict.

21

As always though, a null result must be viewed with caution, as it does not necessarily offer evidence against the hypothesis being tested. For instance, the lack of a referential effect could also be due to the temporary nature of the referential ambiguity or the weakness of the manipulation. Such a tentative conclusion is bolstered by the fact that we did find increased PLIFG activation within the a priori sentence comprehension and representational conflict ROIs for filler sentences that had a referential ambiguity that was globally ambiguous (i.e., a referential ambiguity that was never resolved by the linguistic input or the context).

Our interpretation here is that the co-localization of Syntactic and Stroop effects within PLIFG occurs because both manipulations have in common increased representational conflict. Other types of conflict, such as response conflict, may also play a role, but not to the degree that representational conflict does (indeed, no action was required in the linguistic task). We should note that the Milham et al. (2001) Stroop task, which was used here, has received some criticism (e.g., van Veen & Carter, 2005) pertaining especially to the extent to which it taps representational conflict. On the other hand, Nelson et al. (2003) also find that PLIFG is responsive to representational conflict and not response conflict using a modified item recognition task. The outcome of this disagreement does not impact the main conclusion of our study, regarding the importance of conflict resolution mechanisms in sentence comprehension.

Our finding of co-localized processing from the Stroop task and the syntactic manipulation has important implications for the sentence comprehension literature. Specifically, we found that contextual factors known to influence real time syntactic ambiguity resolution in past eye tracking studies (see Snedeker & Trueswell, 2004) modulated activity in brain regions associated with cognitive control. These results are most consistent with interactive, constraint-based theories of sentence processing, according to which the parser integrates information from

multiple sources simultaneously to determine a coherent interpretation of the sentence (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994). Though the temporally imprecise nature of the BOLD signal prevents us from ruling out a serial process, where an initial representation of the structure is created based solely on syntactic category information, the connection to previous eye tracking studies corroborates an interactive, parallel model. Moreover, if we accept MacLeod's (1991) analysis of the Stroop task as relying on parallel processing of both the word and color information and the findings from the cognitive control literature on PLIFG function as mediating competition among simultaneously active representations, our results support the view that competitive processes operate during sentence comprehension. Findings of competition are inconsistent with non-competitive accounts of sentence parsing, such as the unrestricted race model proposed by van Gompel and colleagues (e.g., van Gompel, Pickering, Pearson, & Liversedge, 2005).

Our results also contribute to a theoretical unification of the diverse literature on the function of PLIFG (Novick et al., 2005). By attributing PLIFG involvement during language comprehension to cognitive control, we can account for inconsistent (Kaan & Swaab, 2002) or non-specific (Vigneau et al., 2006) findings of PLIFG involvement during sentence comprehension. Thus, our account would argue against the proposal that BA 44 or 45 is specialized for syntactic processing (Caplan et al., 1998) or for recovering the function of a phrase even though it is in a non-canonical position (e.g., logical object in subject position, as in passives; Grodzinsky, 2000). Rather, our proposal accounts for increased PLIFG functioning in syntactically complex sentences as a result of the need to alter the activation of multiple representations of the input in response to a number of potentially conflicting cues. Indeed, there are a number of accounts of syntactic complexity effects in the sentence parsing literature that

rely on mechanisms of ambiguity resolution that could be reinterpreted in terms of cognitive control (e.g., Gennari & MacDonald, 2008; Van Dyke & Lewis, 2003). For example, van Dyke & Lewis (2003) attribute failure to recover from an early misanalysis of a sentence to interference effects in memory, a domain in which the role of PLIFG as a locus of cognitive control was established (Jonides & Nee, 2006). Finally, these results strongly suggest that the prior findings of PLIFG activity in response to ambiguity (Mason et al., 2003; Mason & Just, 2007; Stowe, Paans, Wijers, & Zwarts, 2004) are related to the demands on cognitive control to direct activation from dominant to subordinate interpretations of an ambiguity.

Acknowledgements

References

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI

statistics: II. Spatially smoothed data collected under null-hypothesis and experimental

conditions. *NeuroImage, 5*, 199-212.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence

processing. *Cognition, 30*(3), 191-238.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B. M., & Zilles, K. (1999).

Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of

Comparative Neurology, 412*, 319-341.

Badre, D., & Wagner, A. D. (2006). Computational and neurobiological mechanisms underlying cognitive flexibility. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 103*(18), 7186-7191.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Barde, L. H. F., & Thompson-Schill, S. L. (2002). Models of functional organization of the lateral prefrontal cortex in verbal working memory: Evidence in favor of the process model. *Journal of Cognitive Neuroscience, 14*(7), 1054-1063.

Bedny, M., Hulbert, J. C., & Thompson-Schill, S. L. (2007). Understanding words in context: The role of Broca's area in word comprehension. *Brain Research. Special Issue: Mysteries of Meaning, 1146*, 101-114.

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic adaptation and competition during word comprehension. *Cerebral Cortex,*

Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience, 17*(9), 1353-1366.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional

    number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience,*

    10(4), 541-552.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by

    the psychological parser. In D. Dowty, L. Karttunen & A. Zwicky (Eds.), *Natural language*

    processing: Psychological, computational, and theoretical perspectives (pp. 320-358).

    Cambridge, UK: Cambridge University Press.

Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the inferior

    frontal junction in cognitive control: Meta-analyses of switching and Stroop studies. *Human*

    Brain Mapping. Special Issue: Meta-Analysis in Functional Brain Mapping, 25(1), 22-34.

Derrfuss, J., Brass, M., & von Cramon, D. Y. (2004). Cognitive control in the posterior

    prefrontal cortex: Evidence from common activations in task coordination, interference

    control, and working memory. *NeuroImage, 23*, 604-612.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual*

    Review of Neuroscience, 18, 193-222.

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: Insights from TMS. *Brain, 130*,

    610-622.

Feredoes, E., Tononi, G., & Postle, B. R. (2006). Direct evidence for a prefrontal contribution to

the control of proactive interference in verbal working memory. *PNAS Proceedings of the*

National Academy of Sciences of the United States of America, 103(51), 19530-19534.

Fiez, J. A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex.

Human Brain Mapping, 5(2), 79-83.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of

verb bias and plausibility to the comprehension of temporarily ambiguous sentences.

Journal of Memory and Language, 37(1), 58-93.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses.

Journal of Memory and Language, 58(2), 161-187.

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left

inferior frontal cortex with transcranial magnetic stimulation. *Journal of Neuroscience,*

25(35), 8010-8016.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral*

and Brain Sciences, 23(1), 1-71.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working

memory. *Neuroscience, 139*, 181-193.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998).

> Inhibition in verbal working memory revealed

> by brain activation. *Proceedings of the National Academy of the Sciences, 95*, 8410-8413.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in*

> *Cognitive Sciences, 6*(8), 350-356.

Kan, I. P., & Thompson-Schill, S. L. (2004a). Effect of name agreement on prefrontal activity

> during overt and covert picture naming. *Cognitive, Affective & Behavioral Neuroscience,*
> *4*(1), 43-57.

Kan, I. P., & Thompson-Schill, S. L. (2004b). Selection from perceptual and conceptual

> representations. *Cognitive, Affective & Behavioral Neuroscience, 4*(4), 466-482.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of

> syntactic ambiguity resolution. *Psychological Review, 101*(4), 676-703.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review.

> *Psychological Bulletin, 109*(2), 163-203.

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain*

> *Research. Special Issue: Mysteries of Meaning, 1146*, 115-127.

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the brain:

   What brain imaging reveals about the processing of syntactically ambiguous sentences.

   Journal of Experimental Psychology: Learning, Memory, and Cognition, 29(6), 1319-1338.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The

   relative involvement of anterior cingulate and prefrontal cortex in attentional control

   depends on nature of conflict. *Cognitive Brain Research, 12*(3), 467-473.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual

   Review of Neuroscience, 24*, 167-202.

Mostofsky, S. H., & Simmonds, D. J. (2008). Response inhibition and response selection: Two

   sides of the same coin. *Journal of Cognitive Neuroscience, 20*(5), 751-761.

Musso, M., Moro, A., Glauche, V., Rijntes, M., Reichenbach, J., Buchel, C., et al. (2003).

   Broca's area and the language instinct. *Nature Neuroscience, 6*(7), 774-781.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. C., Jonides, J., & Smith, E. E. (2003).

   Dissociable neural mechanisms underlying

   response-based and familiarity-based conflict in working memory. *Proceedings of the

   National Academy of the Sciences, 100*, 11171-11175.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing:

   Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective &

   Behavioral Neuroscience, 5*(3), 263-281.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech

    comprehension: FMRI studies of semantic ambiguity. *Cerebral Cortex, 15*, 1261-1269.

Sahin, N. T., Pinker, S., & Halgren, E. (2006). Abstract grammatical processing of nouns and

    verbs in Broca's area: Evidence from fMRI. *Cortex, 42*(4), 540-562.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker

    awareness and referential context. *Journal of Memory and Language, 48*(1), 103-130.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The

    role of lexical-biases and referential scenes in child and adult sentence processing.

    Cognitive Psychology, 49(3), 238-299.

Stowe, L. A., Paans, A. M. J., Wijers, A. A., & Zwarts, F. (2004). Activations of "motor" and

    other non-language structures during sentence comprehension. *Brain and Language, 89*(2),

    290-299.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration

    of visual and linguistic information in spoken language comprehension. *Science, 268*(5217),

    1632-1634.

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and

    competition on activity in left prefrontal cortex during word generation. *Neuron, 23*, 513-

    522.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language, 35*(4), 566-585.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*(2), 89-134.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. *Perspectives on sentence processing.* (). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49*(3), 285-316.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 52*(2), 284-307.

van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage, 27*, 497-504.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., et al. (2006). Meta-analyzying left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage, 30*, 1414-1432.

Worsley, K. J., & Friston, K. (1995). Analysis of fMRI time-series revisited--again. *NeuroImage, 2*, 173-182.

Zempleni, M., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage, 34*, 1270-1279.