Competition between Mutually Exclusive Object States in Event Comprehension

Sarah H. Solomon[1], Nicholas C. Hindy[2], Gerry T. M. Altmann[3], and Sharon L. Thompson-Schill[1]

Abstract

■ Successful language comprehension requires one to correctly match symbols in an utterance to referents in the world, but the rampant ambiguity present in that mapping poses a challenge. Sometimes the ambiguity lies in which of two (or more) types of things in the world are under discussion (i.e., lexical ambiguity); however, even a word with a single sense can have an ambiguous referent. This ambiguity occurs when an object can exist in multiple states. Here, we consider two cases in which the presence of multiple object states may render a single-sense word ambiguous. In the first case, one must disambiguate between two states of a single object token in a short discourse. In the second case, the discourse establishes two different tokens of the object category. Both cases involve multiple object states: These states are mutually exclusive in the first case, whereas in the second case, these states can logically exist at the same time. We use fMRI to contrast same-token and different-token discourses, using responses in left posterior ventrolateral prefrontal cortex (pVLPFC) as an indicator of conflict. Because the left pVLPFC is sensitive to competition between multiple, incompatible representations, we predicted that state ambiguity should engender conflict only when those states are mutually exclusive. Indeed, we find evidence of conflict in same-token, but not different-token, discourses. Our data support a theory of left pVLPFC function in which general conflict resolution mechanisms are engaged to select between multiple incompatible representations that arise in many kinds of ambiguity present in language. ■

INTRODUCTION

Language is rife with ambiguity. Consider the following sentences: (1) *I swung a bat on the baseball field, and then I saw a bat fly overhead.* (2) *I cracked the wineglass, and then I drank from the wineglass.* Both of these sentences include ambiguity: (1) lexical ambiguity, in which the type of concept is ambiguous (i.e., a baseball bat vs. a nocturnal bat), and (2) state ambiguity, in which the state of a particular exemplar is ambiguous (i.e., the wineglass before or after it was cracked). The presence of these kinds of ambiguity, even when the context is fully disambiguating, presents a problem during language comprehension, because successful comprehension requires an individual to correctly match the symbols in an utterance to the intended referents and/or their intended states. In the case of lexical ambiguity in (1), each instance of "bat" can refer to either a baseball bat or a nocturnal bat, but it cannot simultaneously refer to both—that is, the referent of "bat" cannot simultaneously be two different kinds of bat. In other words, the type of object is mutually exclusive. Similarly, in the case of state ambiguity in (2), each instance of "wineglass" can be intact or cracked, but not both. In this way, states of a particular exemplar are also mutually exclusive. In both of

these cases, correctly representing and retrieving the intended referent and its state, given that there may be more than one referent or it may be in one of several distinct states, presents a challenge to the cognitive system.

The resolution of lexical ambiguity recruits prefrontal mechanisms (Bedny, McGill, & Thompson-Schill, 2008) associated with the need to select between alternative semantic interpretations (Thompson-Schill, D'Esposito, & Kan, 1999; Thompson-Schill et al., 1998; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997); this is referred to as semantic conflict. We have recently reported that resolution of state ambiguity recruits these same prefrontal mechanisms (Hindy, Solomon, Altmann, & Thompson-Schill, 2015; Hindy, Altmann, Kalenik, & Thompson-Schill, 2012). Here we ask whether state ambiguity engenders conflict because of the mutual exclusivity of object states. We test this by including events like in (2) above, in which one object appears in two different states, as well as events that include multiple exemplars of the same object type (e.g., *I cracked the wineglass, and then I drank from another wineglass*). Although one object can only be in one state at any given point in time (one wineglass can, at any one time, only be intact or cracked, but not both), if more than one exemplar is present, multiple states are no longer mutually exclusive: It is readily possible to have one wineglass that is cracked, and another that remains intact. If comprehension of state

[1]University of Pennsylvania, [2]Princeton University, [3]University of Connecticut

ambiguity engenders conflict because of the mutual exclusivity of states, then comprehension of sentences involving multiple states of an object should recruit prefrontal mechanisms when only one exemplar is present, but not when two are present. Alternatively, if conflict arises because of the inherent similarity of two representations of the same object type in two different states, then prefrontal mechanisms should be engaged when two states are present, both when there is one exemplar (e.g., a glass that changes from intact to cracked) and when there are two exemplars (e.g., a glass that is intact and another glass that is cracked).

Selecting a representation from among competing, incompatible alternatives recruits regions of left pFC (Thompson-Schill et al., 1997, 1998, 1999). Stroop color–word interference is a popular example of conflict that recruits these prefrontal mechanisms (Banich et al., 2000). In the Stroop task (MacLeod, 1991; Stroop, 1935), the participant must respond on the basis of one representation (the color in which a word is displayed) while ignoring another concurrently available representation associated with the same stimulus (the color to which the word refers). In incongruent trials, the color in which the word is displayed (e.g., blue) does not match the color term presented (e.g., "red"). Selecting the correct response involves selecting between competing representations, and left pFC is implicated in this resolution process. Likewise, selecting the correct word meaning during lexical ambiguity engenders this neural conflict (Bedny et al., 2008); in this case, two concept representations are competing, and an individual must select the appropriate one to successfully comprehend an utterance. In cases of state ambiguity, a single, otherwise unambiguous word can refer to a single object in either of two (or more) different states—research on language-mediated eye movements suggests that these distinct representations of an object can compete with each other during language comprehension (Altmann & Kamide, 2009). Recent fMRI evidence indicates that similar neural regions are involved in resolution of these two very different kinds of competition—that is, between different concepts associated with a single lexical item and between different states of a single token of a concept (Hindy et al., 2012, 2015). In other words, resolution of state ambiguity engenders conflict similar to that seen in the Stroop color–word interference task, which makes sense, because in both cases an individual must select one representation at the expense of the other. Thus, though the Stroop task and state-change comprehension tasks differ in some respects (e.g., stimuli, response options), we consider both tasks to involve a domain-general conflict resolution process in which the appropriate representation is chosen among competing alternatives. It is this domain-general conflict resolution process, recruited during language comprehension, that is the focus of this article.

Previous research on event comprehension has revealed a cortical network that is sensitive to distinct states of an object as it changes during an event. When an object is substantially changed by an agent's action, increased activity is observed in left posterior ventrolateral prefrontal cortex (pVLPFC; Hindy et al., 2012, 2015). This response is not due to the action word itself (Hindy et al., 2012), but rather to the extent to which the object changes in state (e.g., smashing a wine glass changes its state more than cracking it). In other words, the more an object changes in state, the greater the response in left pVLPFC. The supramarginal gyri (SMG) also emerged as sensitive to object-change in Hindy et al. (2012); this region has been implicated in tasks involving action simulation (Grezes & Decety, 2001), and it is possible that the SMG are sensitive to the action-specific components of events that cause changes in state. Additionally, regions of visual cortex are implicated in comprehension of object change: Hindy et al. (2015) analyzed the neural patterns in early visual cortex that corresponded to imagining the pre- and post-change state of an object and found (i) that similarity between these patterns depended on the degree of the described changes and (ii) that the more dissimilar these patterns, the greater the response in left pVLPFC. Thus, it appears that left pVLPFC, SMG, and regions of visual cortex are part of a cortical network involved in the comprehension of object change.

The correspondence between visual cortex and left pVLPFC during representation of distinct object states (Hindy et al., 2015) suggests that the left pVLPFC may select among representations in visual cortex in this task, at least when changes of state are accompanied by changes in visual form. There are known reciprocal connections between left pVLPFC and posterior domain-specific cortical regions, including visual cortex (Miller & Cohen, 2001), and left pVLPFC is thought to select between both perceptual and conceptual representations (Kan & Thompson-Schill, 2004). Furthermore, regions of visual cortex are recruited during semantic tasks for which context requires the representation of detailed or specific information (Hsu, Kraemer, Oliver, Schlichting, & Thompson-Schill, 2011). Given this previous work, here we assume that, in certain contexts, (i) visual cortex may be recruited to represent objects during sentence comprehension; (ii) these objects are represented in terms of their visual features; (iii) the relevant features are modulated by events involving object state change; and (iv) left pVLPFC is recruited to select between these features or representations in successful comprehension of these events.

The stimuli in Hindy et al. (2012) were two-sentence events in which an agent (e.g., a chef) acted upon an object (e.g., an onion). In the first sentence of each event, the object was either minimally or substantially changed by an action (e.g., *"The chef will weigh/chop the onion."*). In the second sentence, the object was mentioned again (e.g., *"And then, she will smell the onion."*). Thus, the first sentence modulated the degree of object change, and the second sentence required that the participant represent the object in its updated state. Hindy et al.

found that substantial change trials (the "chop" condition) resulted in greater left pVLPFC activity than minimal change trials (the "weigh" condition) and that the left pVLPFC response varied continuously with the degree of change. In that study, the event sequences involved only one object, so the multiple states involved in each event always belonged to the same object token. Thus, it could not be determined whether the conflict was due specifically to selecting between mutually exclusive states or due to maintaining multiple object representations that are highly similar to one another. To resolve this, we include in this study events involving both single and multiple object tokens (i.e., one onion or two). We contrast single and multiple object tokens by contrasting *"The chef will weigh/chop the onion. And then she will smell the/another onion."*

We adopted the experimental paradigm and sentence stimuli from Hindy et al. (2012), but modified the design to specifically test whether or not representational conflict in object change comprehension is due to competition between mutually exclusive token states. While fMRI data were collected, participants read two-sentence event sequences in which an object was changed either minimally or substantially by an agent's action. The events differed in the number and type of object referents. For example, a participant might read: *"The chef will chop the onion. And then, she will smell…the onion"* or *"… another onion"* or *"…a piece of garlic."* Our predictions are as follows: If it is the case that competition arises specifically between mutually exclusive states, then we should observe increased activity in left pVLPFC in substantial change events in which the same token (S-Token) is presented in the second sentence (*"the onion"*), but not when a different token of the same type (D-Token) is presented (*"another onion"*). This is because the former case requires the participant to represent a particular object token in one of two mutually exclusive states, whereas the latter does not. In the S-Token condition, by the time you need to represent "the onion" in the second sentence, it is either whole or it is chopped—it cannot be both. But in the D-Token condition, the state of the second onion is not constrained by the events in the preceding sentence; the word "another" indicates that the second onion need not inherit the episodic characteristics of the first-mentioned onion. Thus, the main analysis of interest is whether left pVLPFC responses are similar or different in the S-Token and D-Token conditions. If the left pVLPFC response increases for substantial change events involving only one object token (S-Token), but not for events with two tokens of the same type (D-Token), then this would be evidence that, in comprehension of events that introduce state ambiguity, competition arises when the states belong to a single object token and are therefore mutually exclusive, but not when they belong to different object tokens.

If we observe a left pVLPFC response to substantial change in the D-Token condition as well as the S-Token condition, this would suggest that conflict arises not because the states are mutually exclusive, but because the underlying representations are inherently similar (e.g., an intact wineglass looks similar to another, cracked wineglass). To confirm that this conflict is due to representational similarity, we would want to measure left pVLPFC response to events that include two different types of objects—two representations that correspond to two different object types (e.g., wineglass vs. soda can) will generally be more dissimilar than two representations that correspond to one object type (e.g., two different wineglasses). Thus, if state ambiguity engenders conflict because of maintaining multiple representations, then we should observe a left pVLPFC response to S-Token and D-Token events, but not (or less so) to events in which two different object types are presented (D-Type condition). To test these hypotheses, we adopted a factorial design, with two levels of object change (minimal, substantial), and three different object-referent conditions (S-Token, D-Token, D-Type). We focused on an anatomical ROI in the left pVLPFC and further restricted our ROI to the voxels that were most sensitive to conflict trials in a Stroop color–word interference task (MacLeod, 1991; Stroop, 1935) to ensure that we were targeting conflict-sensitive voxels. We also examined neural responses to object change in visual cortex and bilateral SMG, regions implicated in representing object states and simulating actions, respectively (Hindy et al., 2015; Grezes & Decety, 2001).

METHODS

Participants

Twenty-four right-handed, native English speakers (16 women) were recruited from the University of Pennsylvania community. Ages ranged from 18 to 41 years ($M = 22.6$). All had normal or corrected-to-normal vision and reported no history of neurological disease or damage. All fMRI participants were compensated $20 per hour. Three additional individuals participated in the experiment but were excluded from the analysis and replaced. One participant's data were not analyzed because of excessive movement in the scanner. Two participants' neural data were not analyzed because they failed to follow task instructions. An additional 85 individuals participated in the study by providing ratings used for stimulus-norming.

Design

We tested the interaction of object change (minimal, substantial) and referent: same token (S-Token), different token of the same type (D-Token), and different token of a different type (D-Type). This resulted in six trial types (3 object-referent conditions × 2 levels of object change).

Event Stimuli

The event stimuli were adapted from Hindy et al. (2012) and modified to accommodate the additional experimental conditions. There were a total of 120 event frames, which could be seen in any of the six different conditions (each participant saw 20 trials within each condition). An example of an event frame and its different versions can be seen in Table 1. The six versions of each event frame were identical except for the action in the first sentence (manipulation of object change) and the object described in the second sentence (manipulation of referent). Each participant saw only one version of each event. Events were two sentences long and described an agent acting upon an object. The object-altering action always occurred in the first sentence; the degree of change that this action elicited was varied, such that the object was either minimally changed or substantially changed (e.g., *"The chef will weigh/chop an onion."*). Across the different examples of substantial change, the actual degree to which the object underwent change (that is, would change in the corresponding real-world event) varied; see below. The second sentence included a reference either to the same token that appeared in the first sentence (S-Token; e.g., *"And then, she will smell the onion"*), a different token of the same type (D-Token; "*…smell another onion*"), or a different token from a different type (D-Type; "*…smell a piece of garlic*"). The different object in each D-Type event was chosen to be as semantically related to the first object as possible while still maintaining plausibility. No substantial change occurred in the second sentence. The data analyses described below ensure that any differences in number of characters as a function of reference type are not confounded. The first-sentence verb for each item was matched across trial types on lexical ambiguity, measured as the number of distinct meanings ($t(238) = 0.75$, $p = .45$; Burke, 2009) and on frequency of use ($t(238) = 1.00$, $p = .32$; Brysbaert & New, 2009). In addition to the events described above, each participant read 18 nonsensical events, which were used as catch trials to be detected in a task that was designed to be orthogonal to comparisons of interest. These events were nonsensical either because the second sentence was implausible given the first sentence (e.g., *"The man will shave off his beard. And then, he will comb his beard."*) or because the word "another" was used with an implausible object in the second sentence (e.g., *"The lion will devour a buffalo. And then, it will chase after another astronaut."*).

Object Change and Action Imageability Ratings

The norming data described in this section are the same norming data reported by Hindy et al. (2012). Online surveys were completed by 85 Penn undergraduates and were used to assess the degree of object change in the first sentence of each event. Each of the survey participants rated only one version of each event. For object change ratings, participants were presented with the sentence and then asked to rate "the degree to which the depicted object will be at all different after the action occurs than it had been before the action occurred" on a 7-point scale ranging from "just the same" (1) to "completely changed" (7). The mean change rating for the "substantial change" items was 4.64 ($SD = 0.84$) and was reliably greater ($t(119) = 27.63$, $p < .001$) than the mean change rating for the "minimal change" items (1.97; $SD = 0.57$). For action imageability ratings, participants rated how much each sentence "brings to mind a clear mental image of a particular action" on a 7-point scale ranging from *not imageable at all* (1) to *extremely imageable* (7). The mean imageability rating for minimal change items was 4.89 ($SD = 0.64$), and the mean for substantial change items was 5.46 ($SD = 0.41$). Because there was a reliable difference of action imageability between substantial and minimal change items ($t(119) = 8.92$, $p < .001$), we included the imageability of each action as a covariate when comparing substantial and minimal change items.

Event Comprehension Task

The event comprehension task consisted of 120 experimental events and 18 catch events distributed across five scanner runs (~6.5 min per run). Each run included

Table 1. An Example Event Frame with Its Six Different Versions

	Condition	The chef will [action] an onion. And then, she will smell [object].	
A	Same-token (min)	The chef will *weigh* an onion.	And then, she will smell *the onion*.
B	Different-token (min)	The chef will *weigh* an onion.	And then, she will smell *another onion*.
C	Different-type (min)	The chef will *weigh* an onion.	And then, she will smell *a piece of garlic*.
D	Same-token (sub)	The chef will *chop* an onion.	And then, she will smell *the onion*.
E	Different-token (sub)	The chef will *chop* an onion.	And then, she will smell *another onion*.
F	Different-type (min)	The chef will *chop* an onion.	And then, she will smell *a piece of garlic*.

(A) S-Token, minimal change. (B) D-Token, minimal change. (C) D-Type, minimal change. (D) S-Token, substantial change. (E) D-Token, substantial change. (F) D-Type, substantial change.

equal numbers of each of the six trial types and either three or four catch events. The order of experimental events and catch trials was randomized within each run, with one exception: in runs with four catch events, the last catch event occurred on the last trial. This was to ensure that participants would have to attend to all events in a run because they could not predict whether or not another catch trial would be presented. The participants were instructed to press the outermost buttons of a handheld response pad when they read one of the catch events, which were designed specifically such that participants would have to read both sentences of each event carefully to respond correctly on each trial. No response was required for plausible events. Each sentence was presented for 3 sec, with the second sentence immediately following the first sentence of each event. Trials were separated by 3–15 sec of jittered fixation, optimized for statistical power using the OptSeq2 algorithm (surfer.nmr.mgh.harvard.edu/optseq/). Stimuli were presented using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA).

Stroop Color–Word Interference Task

Following the event comprehension task, participants completed a 10-min Stroop color–word interference task that was used to localize the regions of left pVLPFC most sensitive to conflict (Hindy et al., 2012, 2015; Milham et al., 2001). Participants were instructed to press the button on the response pad (blue, yellow, green) that corresponded to the typeface color of the word displayed on the screen. Stimuli included four trial types: response-eligible conflict, response-ineligible conflict, and two groups of neutral trials. In response-eligible conflict trials, the word presented on the screen was a color term that matched a possible response (*"blue," "yellow," "green"*) but mismatched its typeface color (blue, yellow, or green). In response-ineligible conflict trials, the color term was not a possible response (*"orange," "brown," "red"*) and mismatched the typeface color. In neutral trials, noncolor terms were used (e.g., *"farmer," "stage," "tax"*); these neutral trials were intermixed with the aforementioned conflict trials across a total of four blocks. Both response-eligible and response-ineligible conflict trial types have been demonstrated to induce conflict at nonresponse levels, and response-eligible conflict trial types also induce conflict at the level of motor response (Milham et al., 2001). To optimize power for identifying subject-specific conflict-sensitive voxels, we collapsed across these two types of conflict trials.

Imaging Procedure

Structural and functional data were collected on a 3-T Siemens Trio system (Erlangen, Germany) and 32-channel array head coil. Structural data included axial T1-weighted localizer images with 160 slices and 1 mm isotropic voxels

(repetition time = 1620 msec, echo time = 3.87 msec, inversion time = 950 msec, field of view = 187 × 250 mm, flip angle = 15°). Functional data included echo-planar fMRI performed in 42 axial slices and 3-mm isotropic voxels (repetition time = 3000 msec, echo time = 30 msec, field of view = 192 × 192 mm, flip angle = 90°).

Data Analysis

Image preprocessing and statistical analyses were performed using FMRIB Software Library (FSL). Functional data processing was carried out using FEAT (fMRI Expert Analysis Tool) Version 6.00. Preprocessing included motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002), interleaved slice timing correction, spatial smoothing using a Gaussian kernel of FWHM 5 mm, grand-mean intensity normalization of the entire 4-D data set by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 sec). Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). Each two-sentence trial was modeled as a 6-sec boxcar function convolved with a double-gamma hemodynamic response function, with motion outliers and imageability ratings as covariates of no interest. Functional data were then registered with each participant's high-resolution anatomical data set, normalized to a standard template in Talairach space, and scaled to percent signal change.

Analyses focused on a conflict ROI, which was defined as the voxels in left pVLPFC that were most sensitive to Stroop-conflict as follows: Left pVLPFC was anatomically constrained based on probabilistic anatomical atlases (Eickhoff et al., 2005) transformed into Talairach space and was defined as the combination of pars triangularis (Brodmann's area 45), pars opercularis (Brodmann's area 44), and the anterior half of the inferior frontal sulcus. On average, these anatomical ROIs comprised 883 voxels. Within these anatomical boundaries, subject-specific conflict ROIs were created by restricting the ROI to the 50 voxels that had the highest t statistics in a subject-specific contrast of conflict > neutral trials in the Stroop color–word interference data. Although this ROI was defined by a Stroop comparison, this region has been shown, even on an individual participant level, to be sensitive to multiple forms of conflict that pertain to semantic retrieval. Analyses were also performed in both 123-voxel and 57-voxel spherical ROIs in the left and right SMG. The left SMG ROI was centered on the peak voxel (x: 40.5, y: 43.5, z: 35.5; Talairach coordinates) of a cluster in the left SMG that arose in a group-level substantial change > minimal change contrast in Hindy et al. (2012). The right SMG ROI was centered on the peak voxel (x: −46.5, y: 43.5, z: 41.5; Talairach coordinates) of a cluster in the right SMG that arose in a group-level contrast of Experiment 1 > Experiment 2 in Hindy et al.

(2012). In Experiment 1, the objects in the substantial and minimal change events were fixed, but the action varied (as in the current study), whereas in Experiment 2, the action was fixed and the objects varied. All statistical tests for each ROI were assessed at the two-tailed $p < .05$ level of significance.

For the psychophysiological interaction (PPI) analysis of correlations across brain regions, we used the 50-voxel Stroop-conflict ROI for each participant as the seed region to identify which neural regions have responses that are correlated (i.e., are "functionally connected") with the left pVLPFC during each of the different object-referent conditions of the event comprehension task. The physiological regressor for each participant was the time course of activation in that participant's Stroop-ROI, and three psychological regressors marked the event time points for the S-Token, D-Token, and D-Type conditions. A separate PPI regressor for each object-referent condition modeled the interaction between events of that condition and the time course of the Stroop-conflict ROI. For all whole-brain analyses, including the PPI analyses, statistical maps were corrected for multiple comparisons across the whole brain at $p < .05$, based on a voxelwise alpha of $p < .01$ and a cluster-forming threshold of 24 voxels determined using 3dClustSim (Cox, 1996; afni.nimh.nih.gov/pub/dist/doc/program_help/3dClustSim.html).

RESULTS

Behavioral Analysis

Stroop Color–Word Interference Task

Participants performed with an average accuracy of 96.4% correct on the Stroop task. RT was significantly greater for incongruent trials ($M = 724$ msec; $SD = 143$ msec) than for congruent trials ($M = 657$ msec; $SD = 127$ msec) in a paired t test ($t(23) = 7.97$, $p < .0001$).

Event Comprehension Task

Participants correctly identified an average of 85.3% of catch trials and made false alarms on an average of 2.3% of the experimental trials. There was no significant difference between the false alarm rates for substantial change trials ($M = 2.6\%$) and minimal change trials ($M = 2.1\%$; $p = .30$).

Stroop-conflict ROI in the Left pVLPFC

To create subject-specific Stroop-conflict ROIs, we identified the 50 voxels within the anatomical boundaries of left pVLPFC with the highest t statistics in a contrast of conflict > neutral trials in the Stroop interference task. The location of the top 50 conflict-responsive voxels varied across participants, with greatest cross-subject overlap in the most posterior area of left pVLPFC (Figure 1).
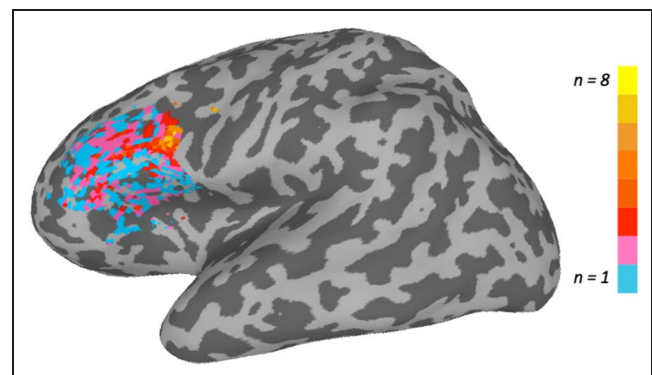
Within these subject-specific conflict ROIs, neural sensitivity to Stroop-conflict (mean t statistics in the incongruent > congruent contrast) was predicted by behavioral Stroop-conflict (incongruent RT – congruent RT); *Spearman's rho*$(23) = 0.49$, $p = .02$, which further indicates that neural activation in this ROI could be interpreted as a conflict response. Within these subject-specific conflict ROIs in left pVLPFC, we examined how the BOLD signal was affected by state change and object referent.

### Left pVLPFC Response to Same Token versus Different Tokens

To determine whether conflict-sensitive voxels in left pVLPFC are specifically sensitive to token-state competition, we compared events in which the second sentence included a reference to the same token (S-Token condition) with those that included a reference to a different token of the same type (D-Token condition). For each, we treated object change (based on the action in the first sentence) as a categorical as well as a continuous variable.

Substantial versus Minimal Change

To determine whether left pVLPFC responded similarly to object change in the S-Token and D-Token events, we calculated mean percent signal change (compared to a fixation baseline) in the minimal and substantial change trials within the S-Token and D-Token conditions, for each participant. The mean percent signal change for each of the four trial types (substantial/minimal change, S-Token/D-Token referent), averaged across participants, is shown in Figure 2A. A two-way within-subject ANOVA revealed a significant main effect of State change ($F(1, 23) = 8.30$, $p = .01$) and a significant interaction between State change and Referent ($F(1, 23) = 12.17$, $p = .002$). There was no main effect of Referent

Figure 1. A visualization of the extent to which the Stroop-conflict ROIs overlapped across participants. Each subject-specific ROI included the 50 left pVLPFC voxels with the highest within-subject t statistics for the Stroop contrast. The left pVLPFC voxel with the greatest overlap across participants included 8 of the 24 total participants.
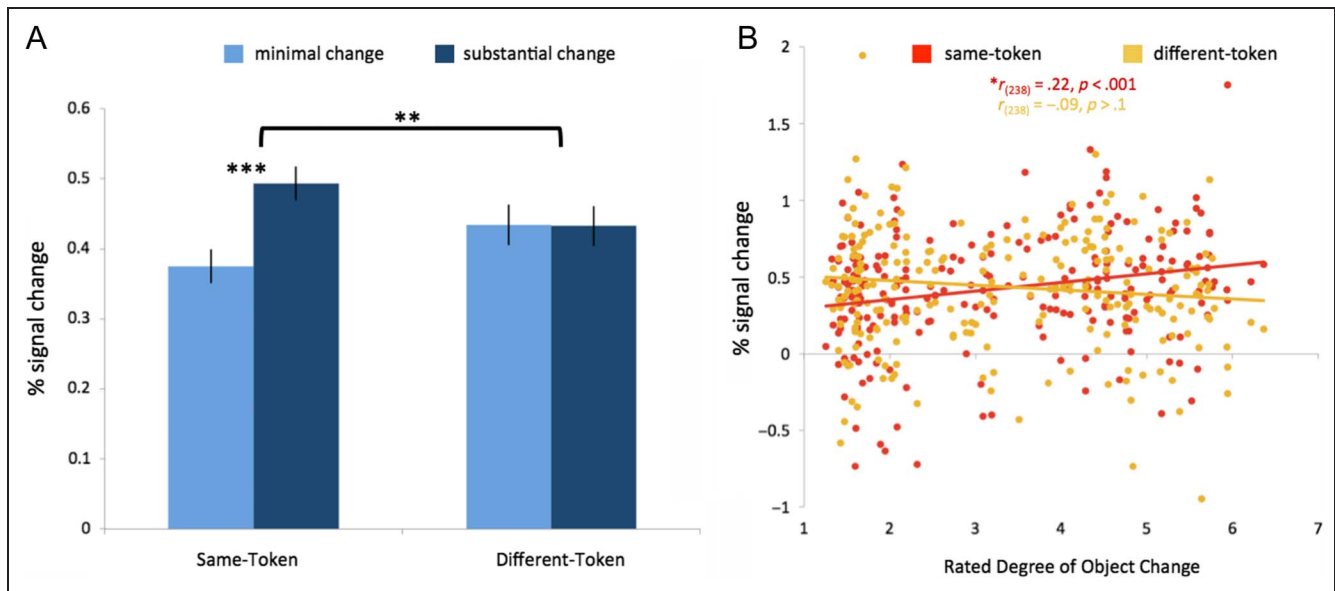
Figure 2. Comparison of S-Token and D-Token conditions. (A) There is a significant effect of object change in the S-Token condition ($p < .001$), but not in the D-Token condition ($p = .97$). Error bars reflect the difference of the means. (B) Item analysis using object change as a continuous variable. Rated degree of object change predicts percent signal change for each item in the S-Token condition ($p < .001$) but not in the D-Token condition ($p = .16$).

($p = .97$). Paired t tests revealed that, in the S-Token condition, the left pVLPFC response was greater for substantial change than for minimal change trials ($t(23) = 4.92$, $p < .0001$). The difference between substantial- and minimal change trials was not significant in the D-Token condition ($p = .97$). The pattern of results was identical when action imageability was not covaried out, with a reliable difference between minimal and substantial state change for the S-Token condition ($t(23) = 4.30$, $p < .001$), and no such difference for the D-Token condition ($p = .87$). Although we used a 50-voxel conflict ROI for this and all following analyses, this result was robust at a wide range of ROI sizes. In fact, the same pattern of results held for all ROI sizes within the anatomically defined left pVLPFC ROI.

Item Analysis

Within each of the substantial and minimal change conditions, there was considerable variability in rated degree of change across items. We calculated effects of state change for each item (following from Bedny, Aguirre, & Thompson-Schill, 2007) to determine whether the left pVLPFC response to each item could be predicted by the magnitude of the object change for that item. Because there were 120 event frames, each with substantial and minimal change versions for each referent, there were 240 items within each object-referent condition. Percent signal change for each of these items was collapsed across participants. We then looked within the S-Token and D-Token conditions to determine whether the neural response for each item could be predicted by degree of object change.

We measured the Pearson correlation between the rated degree of object change for each of the 240 items and the mean percent signal change in the S-Token and D-Token conditions (Figure 2B). Rated degree of object change predicted activity in the Stroop-conflict ROI in the S-Token condition ($r(238) = 0.22$, $p < .001$), but not in the D-Token condition ($p = .16$). A Steiger's Z test on the difference between these dependent correlations (Hoerger, 2013; Steiger, 1980) revealed these correlations to be significantly different from each other ($Z_H = 3.44$, $p < .001$). We also ran partial correlations to remove variance from the imageability ratings, and the same pattern was found. Degree of change predicted the left pVLPFC response to S-Token items ($r(238) = 0.22$, $p = .001$), but not to D-Token items ($p = .14$).

Left pVLPFC Response to Different Types

If the S-Token and D-Token conditions had not differed from each other, the D-Type condition would have been useful to determine whether left pVLPFC discriminates between different types, if not different tokens. Because we found a reliable difference between the S-Token and D-Token conditions, the data from the D-Type condition are not central to tests of the primary hypothesis, but we report the data here for the sake of completeness.

Substantial versus Minimal Change

We calculated mean percent signal change (compared to a fixation baseline) in the minimal and substantial change trial types within the D-Type condition for each participant. The effect of Object change was marginally significant

when imageability was not included as a covariate ($p = .06$) but was significant when imageability was included in the model ($t(23) = 2.50$, $p = .02$). The left pVLPFC response to substantial change did not differ between the D-Type condition ($M = 0.52$) and the S-Token condition ($M = 0.49$; $t(23) = 0.89$, $p = .38$). The results of the Object-change × Object-referent ANOVA, reported above for the S-Token and D-Token conditions, remain the same when the D-Type condition is added to the model.

Item Analysis

In the item analysis, we asked whether the left pVLPFC response to each item in the D-Type condition could be predicted by the object change ratings. Because of a programming error, one of the event frames was not seen in the substantial change D-Type condition, resulting in 239 (instead of 240) D-Type items. Rated degree of object change did not reliably predict activity in the Stroop-conflict ROI for the D-Type condition, although the effect was marginal ($r(237) = 0.12$, $p = .08$). When the variance from imageability ratings was removed, this relationship became significant ($r(237) = 0.13$, $p = .04$). We return to this effect below.

D-Type versus D-Token

We previously contrasted left pVLPFC responses in the S-Token and D-Token conditions to discriminate between the mutual exclusivity and similarity-based interference hypotheses (the idea that the conflict we observed in previous studies was due to the similarity of the distinct states regardless of whether it was one token in different states or two tokens each in a different state). We predicted that if the effect of state change in the S-Token condition was due to similarity-based interference, then we should observe this same effect in the D-Token condition; this result was not found. However, another prediction put forward by proponents of the similarity-based interference view might be that when two tokens of the same object type are present in a discourse, similarity-based interference should occur both in cases of minimal and substantial change, thus a difference in left pVLPFC response would not be expected between these conditions. That is, it is possible that instantiating two distinct, yet otherwise identical tokens induces interference. (Note that this would also predict no difference between substantial and minimal change in the S-Token condition, which was not the case—however, it could be that the presence of different representations introduces greater interference, which masks any more subtle differences that arise due to state change.) To test this, we compared the minimal change D-Token condition with the minimal change D-Type condition: according to this interpretation of the similarity-based interference hypothesis, there should be increased left pVLPFC response for minimal change D-Token events relative to minimal

change D-Type events, because two representations of the same object-type are more similar than two representations of objects from different categories. No such difference was found ($t(23) = 0.19$, $p = .85$).

Whole-brain Analysis

To further compare the effects of object change across object-referent conditions, we performed a whole-brain, group level contrast of substantial > minimal change within the S-Token, D-Token, and D-Type conditions. In the S-Token condition, reliable clusters were found in left pVLPFC (pars triangularis), right superior frontal gyrus, right inferior frontal gyrus (pars triangularis), and left SMG. No voxels survived threshold in the D-Token condition. In the D-Type condition, reliable clusters were found in the left thalamus, left SMG, left superior frontal gyrus, left cingulate gyrus, right lentiform nucleus, left insula, left pVLPFC (pars opercularis), and right inferior parietal lobule. Although both the S-Token and D-Type conditions appear to recruit left pVLPFC during object change comprehension, the S-Token condition recruits a larger cluster of left pVLPFC voxels (396 voxels) than the D-Type condition (26 voxels; Figure 3; Table 2). If we restrict the analysis to the anatomical left pVLPFC ROI, we can statistically compare the number of significant voxels (not clusterized) between conditions. The S-Token condition resulted in 221/895 voxels sensitive to state change, whereas this number was only 36/895 voxels in the D-Type condition; this difference was significant ($\chi = 155.5$, $p < .0001$).

SMG

Regions of posterior parietal cortex have been implicated in previous work on object change (Hindy et al., 2012, 2015). The left SMG emerged in a contrast of substantial > minimal change in Hindy et al. (2012). Additionally, Hindy et al. (2012) compared two experiments: In Experiment 1, substantial and minimal change events involved the same objects but different actions (as in the current experiment); in Experiment 2, the action was fixed but the object varied. The only significant cluster that arose in a contrast of these two experiments was focused on the right SMG. We thus centered a 123-voxel sphere (9-mm radius) on the peak coordinate in the left (Talairach coordinates: 40.5, y 43.5, z 35.5) and right SMG (Talairach coordinates: −46.5, y 43.5, z 41.5) that arose in these contrasts in Hindy et al. (2012) and calculated the mean percent signal change for each of the six trial types (Figure 4). Similar to the Stroop-conflict ROI, a two-way within-subject ANOVA revealed a significant main effect of State change in the right SMG ($F(1, 23) = 7.03$, $p = .01$); this effect was marginal in the left SMG ($F(1, 23) = 3.28$, $p = .08$). There was no main effect of Referent in the right ($p > .35$) or left ($p > .65$) SMG. However, unlike the Stroop-conflict ROI, the interaction between State change and

Figure 3. Whole-brain group results of object change. Clusters are thresholded at $p < .05$, corrected for multiple comparisons. (A) Reliable clusters in the S-Token condition included left pVLPFC (pars triangularis), left superior frontal gyrus, right pVLPFC (pars triangularis), and left SMG. (B) No clusters survived threshold in the D-Token condition. (C) Reliable clusters shown in the D-Type condition include left SMG, left superior frontal gyrus, left pVLPFC (pars opercularis), and right inferior parietal lobule. The full list of reliable clusters (including ones not shown) can be found in Table 2.
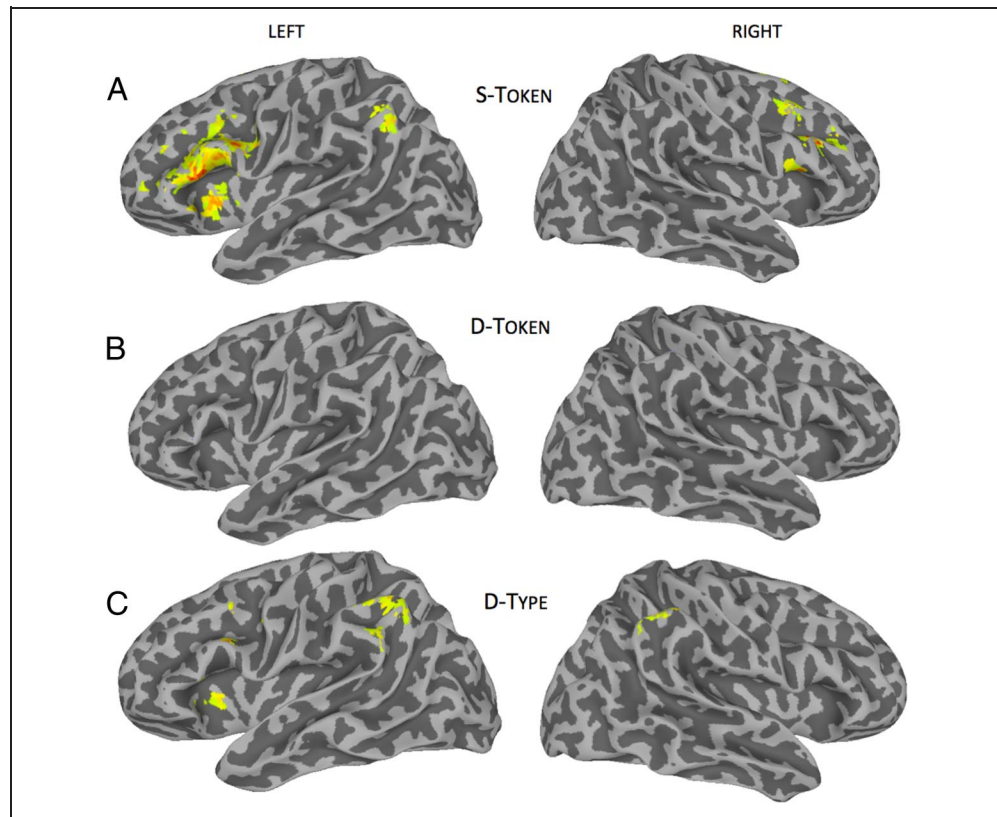

Referent was not significant in either the right SMG ($p > .71$) or the left SMG ($p > .56$). Individual pairwise analyses revealed a reliable effect of State change for the S-Token ($t(23) = 2.12$, $p = .045$) and D-Type ($t(23) = 2.50$, $p =$ .02) conditions in the right SMG. No other pairwise comparisons were reliable. We repeated these analyses in 57-voxel spherical ROIs placed over the same peak coordinates to make sure this finding was robust. Results were identical.

Table 2. Clusters of Voxels with Reliably Greater Signal Amplitude for Substantial Change Trials than Minimal Change Trials

	Voxels, n	Peak z	x	y	z	Brain Region
Same token	396	3.92	52.5	−16.5	8.5	L. pVLPFC (pars triangularis)
	236	3.83	4.5	−16.5	50.5	L. superior frontal gyrus
	160	3.87	−43.5	−19.5	29.5	R. pVLPFC (pars triangularis)
	30	3.28	43.5	49.5	35.5	L. SMG
Different token	–	–	–	–	–	–
Different type	185	3.85	7.5	4.5	14.5	L. thalamus
	85	3.43	55.5	43.5	35.5	L. SMG
	81	3.54	7.5	−22.5	47.5	L. superior frontal gyrus
	64	3.41	1.5	22.5	29.5	L. cingulate gyrus
	40	3.28	−13.5	−1.5	−3.5	R. lentiform nucleus
	32	2.99	31.5	−19.5	2.5	L. insula
	26	3.42	52.5	−7.5	29.5	L. pVLPFC (pars opercularis)
	24	3.04	−49.5	−37.5	47.5	R. inferior parietal lobule

Each contrast has a threshold of $p < .05$, corrected for multiple comparisons. There were no reliable clusters for substantial > minimal change in the D-Token condition.
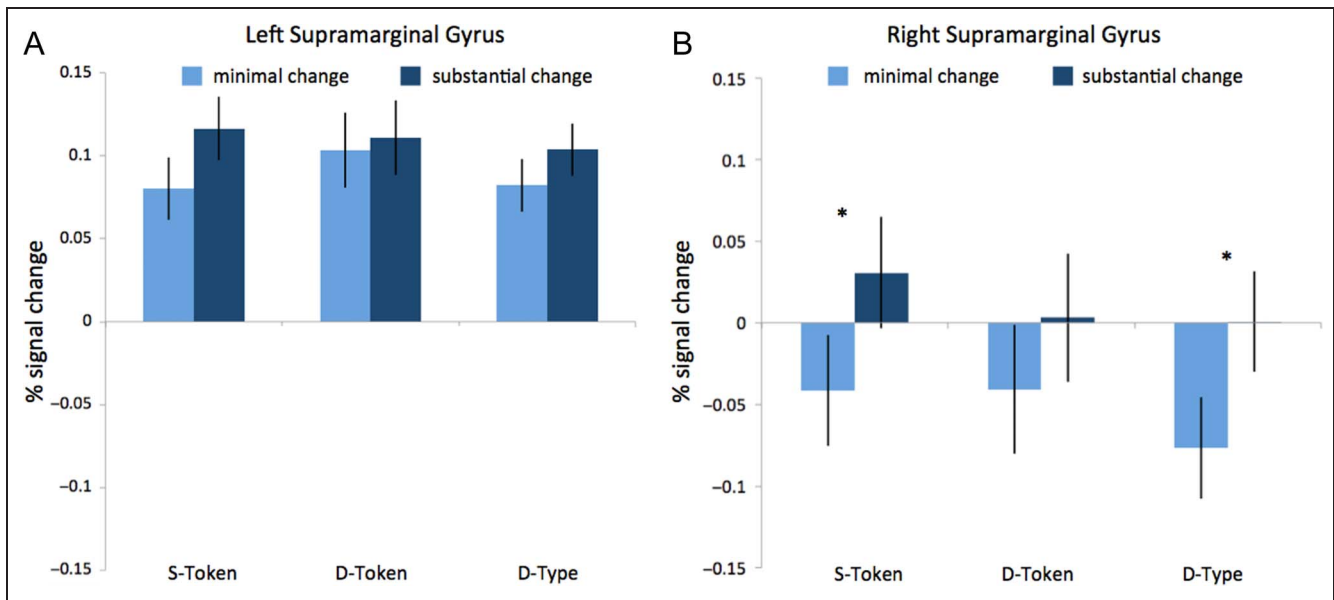
Figure 4. Effect of categorical object change in the S-Token, D-Token, and D-Type conditions in 123-voxel ROIs in left (A) and right (B) SMG. Error bars reflect the difference of the means. In right SMG, there is a main effect of Change ($p = .01$); this main effect is marginal in left SMG ($p = .08$). There is no interaction between Change and Object condition in right ($p = .71$) or left ($p = .56$) SMG.

Comparing S-Token and D-Type Conditions

One question that remains is why we are finding an effect of object change in the D-Type condition. Although this does not bear directly on our central hypothesis, it warrants exploration. One way to start is by examining the extent to which S-Token and D-Type events recruited the same neural regions and the extent to which they each recruited voxels sensitive to Stroop-conflict. To answer these questions, we binarized the group-level S-Token, D-Type, and Stroop z-stat maps at $p < .05$ (corrected for multiple comparisons) and overlaid the binarized maps to reveal the overlap between conditions. This visualization is shown in Figure 5. Comprehension of S-Token events recruited a relatively large cluster of Stroop-sensitive voxels (orange), whereas this overlap was minimal in the D-Type events (green). Across the voxels that survived the thresholding described above, overlap with Stroop-sensitive voxels occurred for 175 of 817 S-Token voxels (21%) and for 74 of 804 D-Type voxels (9%); a chi-square test for independence revealed the greater overlap between Stroop and S-Token voxels to be significant ($\chi = 33.45, p < .0001$). The three tasks (S-Token, D-Type, and Stroop) converged in left pVLPFC as well as left medial frontal gyrus (pink). Regions recruited for S-Token and D-Type event comprehension, but not for Stroop-conflict, include a distinct region of the left pVLPFC as well as the left SMG (purple).

These results, along with the ones previously reported, suggest that the particular regions of left pVLPFC recruited and the extent to which they are recruited differ between S-Token and D-Type conditions. This difference is also supported by the PPI analyses reported below.

A summary of these differences is shown in Table 3. However, the mechanisms underlying comprehension of D-Type events should be explored in future studies.

Connectivity between Left pVLPFC and Visual Cortex

Hindy et al. (2015) reported that ratings of visual similarity between pairs of object states predicted multivariate pattern similarity in early visual cortex when participants imagined those states. Furthermore, pattern similarity observed for each trial in early visual cortex predicted activation in left pVLPFC, beyond variance explained by the similarity ratings alone. This suggests that the pVLPFC may be functionally connected to visual cortex during comprehension of object-change events, perhaps due to the need to separate similar patterns belonging to alternate states of a single object. In Hindy et al. (2015), participants were explicitly asked to imagine the state of an object after it had changed in state, whereas in the current study only sentences were presented and no imagery was explicitly required. We ran PPI analyses using the subject-specific Stroop-conflict ROIs in left pVLPFC as the seed region. Results of the PPI analysis are shown in Figure 6. For both the S-Token and D-Type conditions, we found correlated activity between left pVLPFC and regions of right pFC. However, in the S-Token condition (across minimal and substantial change trials), we also found reliable correlated activity between left pVLPFC and bilateral early visual cortex. In the D-Type condition, we found reliable connectivity between left pVLPFC and the anterior part of the left superior temporal

Figure 5. A visualization of the relationship between the S-Token, D-Type, and Stroop-conflict activations. Condition overlap in left pFC is enlarged in the bottom circle for easier viewing. Group contrasts in each condition were thresholded at $p < .05$ (corrected), binarized, and projected onto a normalized brain in Talairach space. A larger percentage of significant S-Token voxels overlap with significant Stroop voxels than is the case for significant D-Type voxels ($p < .001$).
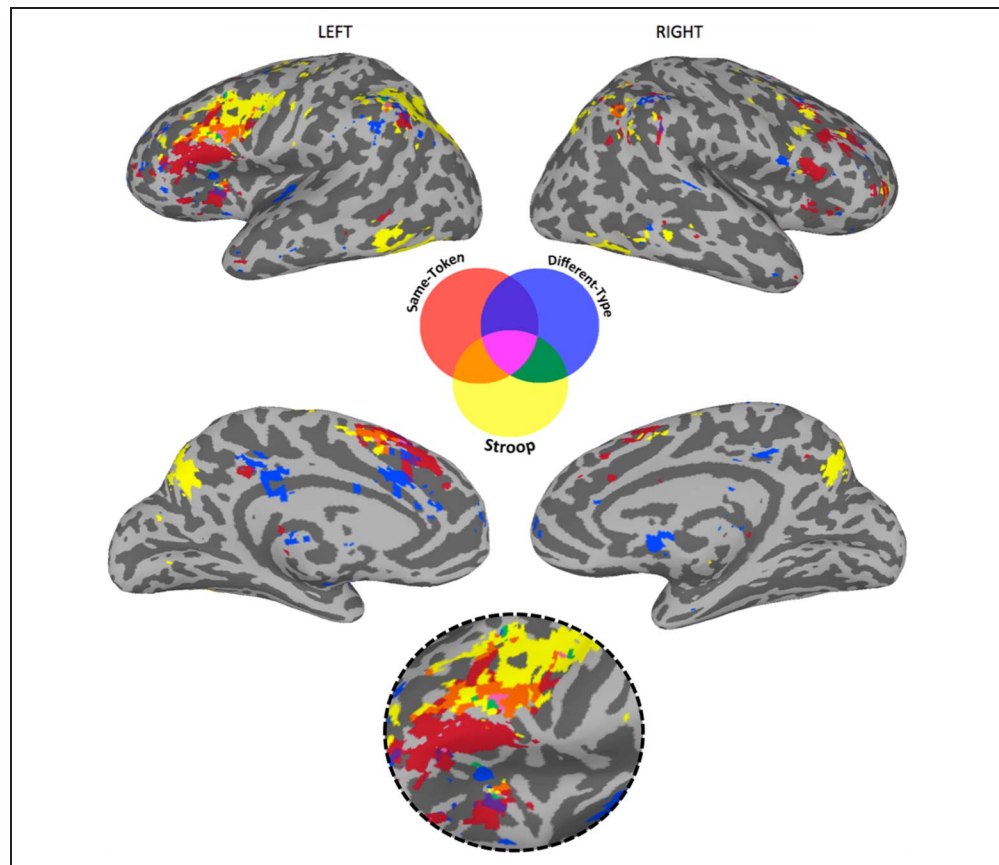
gyrus. Significant clusters for each condition are listed in Table 4.

DISCUSSION

Here we asked whether state ambiguity engenders representational conflict only when those states are mutually exclusive. To answer this question, we measured left pVLPFC response to events involving objects changing in state in which one (S-Token) or two (D-Token) object tokens were presented. By comparing the difference in left pVLPFC activation between substantial change and minimal change trials in the S-Token condition and D-Token conditions, we found that left pVLPFC was sensitive to object change in the S-Token condition, but not in the D-Token condition. Because a single object token can be in only one state at one point in time (a particular onion cannot be both whole and chopped at the same time), whereas two object tokens can be in different states (one onion can be whole, whereas another is chopped), our data provide evidence that, during comprehension of language with state ambiguity, competition arises specifically for mutually exclusive states of a single object token. Although the representations underlying the different object states are inherently similar (a whole onion is similar to a chopped onion, in general), simply maintaining multiple overlapping representations did not engender competition.

Table 3. A Summary of the Differences between the S-Token and D-Type Conditions

	Same-Token	*Different-Type*
Condition analysis in left pVLPFC	Substantial change > Minimal change	Substantial change > Minimal change
Item analysis in left pVLPFC	Degree of object change predicts left pVLPFC response	Degree of object change predicts left pVLPFC response when imageability variance is removed
Whole-brain contrast of sub > min change	396 voxels in left pVLPFC	26 voxels in left pVLPFC
Whole-brain connectivity	Early visual cortex, right pFC	Left temporal lobe, right pFC

Figure 6. Whole-brain group results of the PPI analysis, using Stroop-conflict ROIs in left pVLPFC as seed regions. (A) In the S-Token condition, we find reliable connectivity between left pVLPFC and bilateral early visual cortex, as well as right pFC. (B) In the D-Token condition, left pVLPFC activation does not reliably covary with activation in any other regions. (C) In the D-Type condition, left pVLPFC is functionally connected to the right inferior frontal gyrus and the anterior part of the left superior temporal gyrus.
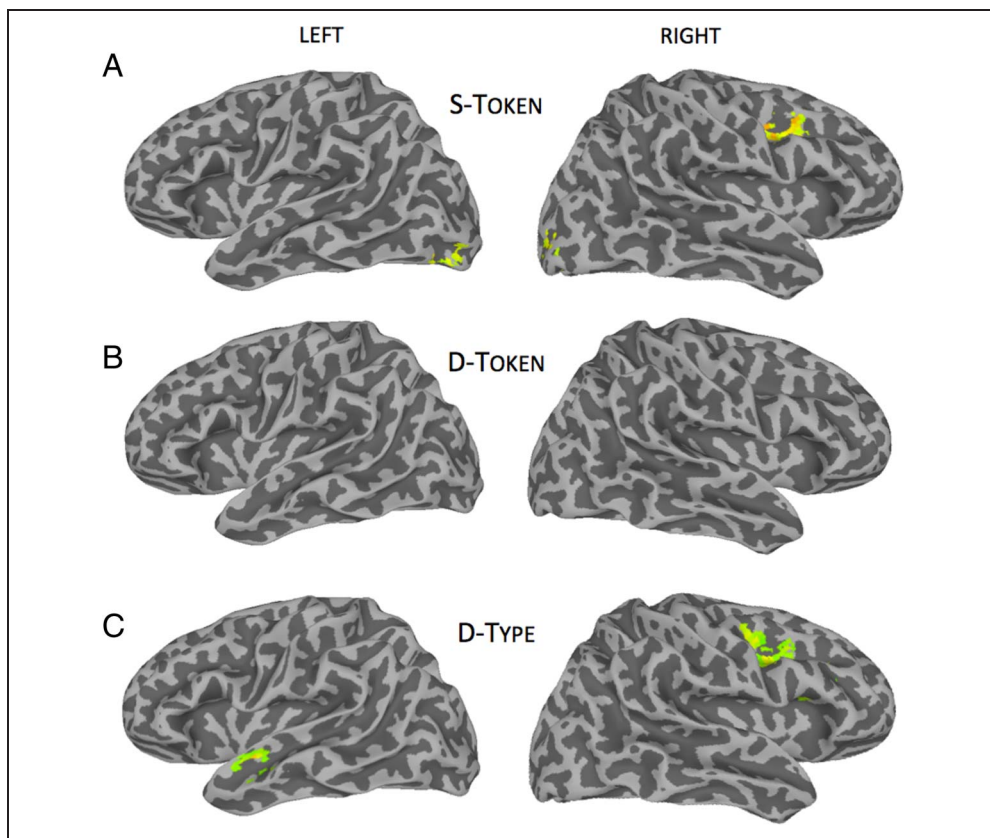
The data described here provide strong evidence that the object state-change pVLPFC response observed in this and previous studies (Hindy et al., 2012, 2015) reflects competition between incompatible, mutually exclusive states of a single object token and not the inherent similarities in the representations underlying those states. This gives us insight into how we process unfolding events with multiple objects in multiple states. Our claim here is that the conflict response we observe in left pVLPFC is indicative of retrieval processes for selecting among the alternative state representations—the first sentence in each two-sentence trial was identical across all conditions, and the fact that we find no evidence of conflict in the D-token condition ("The chef will chop/weigh an onion. And then, she will smell another onion.") suggests that the conflict response is not due to the introduction of two states in the first sentence (the onion before and after chopping) nor to their subsequent maintenance throughout the second sentence (which would result in similar conflict in both conditions). Rather, it is the retrieval of the appropriate state representation, at the end of the second sentence, that

Table 4. Clusters of Voxels that Were Reliably Correlated with Left pVLPFC in a PPI Analysis

	Voxels, n	Peak z	x	y	z	Brain Region
Same token	60	3.94	−31.5	−13.5	29.5	R. middle frontal gyrus
	42	3.53	−22.5	28.5	29.5	R. cingulate gyrus
	30	3.02	31.5	76.5	−9.5	L. middle occipital gyrus
	29	3.05	−31.5	85.5	−0.5	R. middle occipital gyrus
Different token	−	−	−	−	−	−
Different type	122	3.66	−37.5	−7.5	29.5	R. inferior frontal gyrus
	24	3.47	49.5	4.5	−9.5	L. superior temporal gyrus

Each cluster is thresholded at $p < .05$, corrected for multiple comparisons. Talairach coordinates and anatomical labels indicate the location of the peak voxel of each cluster.

engenders conflict. In other words, activation of left pVLPFC seems to reflect competition between multiple incompatible representations of an object.

Our results suggest that increased response in left pVLPFC to events involving an object state change is not due to similarity-based interference (in fact, we observe dissimilarity-based interference; the greater the change, the greater the conflict). As similarity-based interference has been observed behaviorally in other sentence comprehension studies (Fedorenko, Gibson, & Rohde, 2006; Gordon, Hendrick, Johnson, & Lee, 2006; Van Dyke & McElree, 2006; Gordon, Hendrick, & Johnson, 2001), why was it not observed here? Although similarity between different tokens raises the challenge of discriminating between those tokens to select the appropriate one (hence similarity-based interference), dissimilarity in the distinct states of the same token raises a different challenge—that of generalizing from one state to the other (and, in effect, binding each state exemplar onto the same token, such that certain characteristics are shared by both states)—hence dissimilarity-based interference.

Our finding that activity in the left pVLPFC was correlated with activity in early visual cortex exclusively in the S-Token condition is intriguing and follows from our previous findings regarding left pVLPFC involvement in the comprehension of object change (Hindy et al., 2015). If state-change comprehension can recruit visual cortex to represent the unfolding events and its objects and if it is the case that left pVLPFC "pulls apart" mutually exclusive representations (in this case, of objects in visual cortex), we would expect that the left pVLPFC is functionally connected to the visual cortex in the S-Token condition, but not the D-Token condition where no "pulling apart" is required. It is also possible that objects are only represented in visual cortex in the S-Token condition when multiple distinct, yet similar, representations correspond to a single object. Much like perceptual knowledge retrieval recruits visual regions during within-category discriminations (Hsu et al., 2011), perhaps visual regions are recruited for discriminations between within-object states (to the extent that these states are visually distinct). Either way, the present data suggest that in certain verbal contexts objects or object features are represented in modality-specific cortex (e.g., early visual cortex; primary or secondary motor cortex) during sentence comprehension. Additionally, the fact that we found a relationship between left pVLPFC and visual cortex is evidence for prefrontal domain-general control processes coordinating neural activity in domain-specific posterior regions during semantic tasks.

We focused our analyses on functionally defined ROIs based on an orthogonal Stroop task; these ROIs missed a significant portion of the language effect, which was centered on a more anterior portion of the left pVLPFC. Although here we are interested in the conflict process involved in state-change comprehension, many more processes may be involved in the comprehension of our events. The activity in anterior regions of left pVLPFC during event comprehension, but not Stroop-conflict, might reflect a memory retrieval process that is common across the event conditions (Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005), perhaps relating to the objects or actions present in the events.

Although not central to our hypothesis, we also found that left pVLPFC responded more to substantial changes than to minimal changes for events including tokens of two different concept types (D-Type condition), which reveals that the effects of object change on left pVLPFC are more nuanced than the story we provided above. However, we also found that (a) the whole-brain analysis reveals fewer voxels sensitive to the substantial change > minimal change contrast in the D-Type than the S-Token condition, (b) connectivity between left pVLPFC and visual cortex was only found in the S-Token condition, not the D-Type condition, and (c) connectivity between left pVLPFC and left superior temporal gyrus was only found in the D-Type condition, not the S-Token condition. This last observation may help explain the unexpected finding that substantial change items produced conflict (i.e., increased left pVLPFC response) in the D-Type condition. It is possible that, in substantial change events, objects are represented in more detail because of the activation of additional features or increased involvement of motor representations. If this is the case, then retrieving and representing information from a different, yet similar, object category may be more difficult. Because the left temporal lobe is associated with semantic knowledge (Binder & Desai, 2011; Lambon-Ralph et al., 2010; Patterson, Nestor, & Rogers, 2007), the increased left pVLPFC response in the D-Type condition could reflect increased difficulty in retrieving new category information from anterior temporal lobe. Here, too, we can interpret activation in left pVLPFC to reflect competition between multiple incompatible representations—whereas in the S-Token condition the competition is between incompatible object states, the competition in the D-Type condition is between incompatible object types. The connectivity between the left pVLPFC and visual cortex in the S-Token condition reflects the interaction between conflict resolution and visual properties of object states, whereas the connectivity between left pVLPFC and temporal cortex in the D-Type condition reflects the interaction between conflict resolution and category knowledge. However, it is clear that more research is required to understand how left pVLPFC and other semantic control mechanisms are recruited when multiple object types in multiple states are presented in the same event.

The observed pattern of results for the S-Token and D-Token events in left pVLPFC is evidence against the hypothesis that increased activation in pFC during comprehension of action sentences arises because of mental simulation of action (see Grezes & Decety, 2001). If left pVLPFC activation reflects simulating actions, as opposed

to resolving conflict between mutually exclusive token states, then we should have observed no differences between the S-Token and D-Token conditions, because the exact same actions were used in each.

The SMG, however, do not respond differentially to the distinct object referents (S-Token, D-Token, and D-Type) but show increased activation for events involving a substantial object change, especially the right SMG. In other words, the response in this region is determined by the action in the first sentence, irrespective of the referent in the second sentence. This result coheres with previous work suggesting that SMG is implicated in action simulation (Grezes & Decety, 2001), as well as gesture recognition and body schema (Chaminade, Metzoff, & Decety, 2005; Hermsdörfer et al., 2001). Our results also cohere with the finding in Hindy et al. (2012) that the right SMG was sensitive to the specific action (e.g., "stomp") and not the extent of state change that the action caused. By comparing our results in the SMG to those in left pVLPFC, we can see that these two regions are performing different roles in the comprehension process: Whereas the SMG may be involved in simulating actions described in the sentences, here we provide evidence that left pVLPFC is involved in detecting or resolving ambiguity between mutually exclusive states of individual objects.

It is almost surprising how effectual the word "another" was in the D-Token condition: It appears that adding this word served to block the new token from inheriting any (episodic) characteristics of the first token, resulting in no representational conflict whatsoever. Having this modifier directly in front of the new object label (e.g., "another onion") might explain the elimination of state conflict by virtue of specifying that this is a different object. Perhaps if we had added such a diagnostic modifier in the S-Token condition (e.g., "the chopped onion," where "chopped" resolves the state ambiguity of "the onion") we would also have observed decreased competition. In either case, these results highlight the power not just of the context of an utterance or phrase but of individual words within a phrase (i.e., "another") to influence semantic control mechanisms and object representations during language comprehension. Importantly, the distinct brain responses we observed here in pVLPFC and elsewhere map onto distinct behaviors—the linguistic behaviors that result in the interpretation of "another onion" as distinct from "the onion." Future studies should explore the ways in which the presence of different object types—that is, different object categories—interacts with this process.

Finally, our data imply an interaction between semantic and episodic memory, in which an object representation is determined both by semantic knowledge of an object category (e.g., what onions taste like and look like) as well as episodic knowledge of particular tokens and what they have been through (e.g., that your particular onion is chopped on the cutting board). This suggests that neural regions involved in the encoding of temporal information such as the medial-temporal lobe (MTL) or, more specifically, the hippocampus might be part of the network recruited during object state-change comprehension. Indeed, evidence suggests that structures in the MTL use causal relationships between actions and different states of an object to create structured representations of an object and its possible states (Hindy & Turk-Browne, 2015), although the relationship between these processes in MTL and control processes in left pVLPFC during event comprehension should be further explored. The challenge for future accounts of event comprehension and its neural underpinnings will be to better understand the interplay between episodic knowledge of the particular tokens and the states they were in and the semantic knowledge of the categories from which those tokens are drawn.

In summary, our findings support a theory of left pVLPFC function in which general conflict resolution mechanisms select between multiple, incompatible representations that arise in cases of lexical (Bedny et al., 2008), syntactic (January, Trueswell, & Thompson-Schill, 2009), and state (Hindy et al., 2012, 2015) ambiguity. The content of these representations will differ depending on the kind of ambiguity present, and the cortical regions housing these representations will likely be domain-specific and coactivate with left pVLPFC during the resolution process.

Conclusion

Here we have shown that state ambiguity results in representational conflict only when the multiple states correspond to mutually exclusive states of a particular object token. Left pVLPFC, the SMG, and regions of modality-specific cortex compose a network that underlies the comprehension of events in which objects undergo change, ensuring that different states of the same object token are kept distinct and that one token does not automatically inherit the properties of the other. This illuminates the way that the type/token/token-state distinction is maintained at the neural level, a cognitive capability that is crucial for accurately representing a constantly changing world.

REFERENCES

Altmann, G., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye

movements and mental representation. *Cognition, 111,* 55–71.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47,* 907–918.

Banich, M., Milham, M., Atchley, R., Cohen, N., Webb, A., Wszalek, T., et al. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience, 12,* 988–1000.

Bedny, M., Aguirre, G. K., & Thompson-Schill, S. L. (2007). Item analysis in functional magnetic resonance imaging. *Neuroimage, 35,* 1093–1102.

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic adaptation and competition during word comprehension. *Cerebral Cortex, 18,* 2574–2585.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15,* 527–536.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977–990.

Burke, R. J. (2009). *Roger's Reference: The complete homonym/homophone dictionary* (7th ed.). Queensland, Australia: Educational Multimedia Publishing.

Chaminade, T., Metzoff, A. N., & Decety, J. (2005). An fMRI study of imitation: Action representation and body schema. *Neuropsychologia, 43,* 115–127.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers in Biomedical Research, 29,* 162–173.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25,* 1325–1335.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language, 54,* 541–553.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1411.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1304.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping, 12,* 1–19.

Hermsdörfer, J., Goldengert, G., Wachsmuth, C., Conrad, B., Ceballos-Baumann, A. O., Bartenstein, P., et al. (2001). Cortical correlates of gesture processing clues to the cerebral mechanisms underlying apraxia during the imitation of meaningless gestures. *Neuroimage, 12,* 149–161.

Hindy, N. C., Altmann, G. T., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object changes on event processing: Do objects compete with themselves?. *The Journal of Neuroscience, 32,* 5795–5803.

Hindy, N. C., Solomon, S. H., Altmann, G. T., & Thompson-Schill, S. L. (2015). A cortical network for the encoding of object change. *Cerebral Cortex, 25,* 884–894.

Hindy, N. C., & Turk-Browne, N. B. (2015). Action-based learning of multistate objects in the medial temporal lobe. *Cerebral Cortex,* bhv030.

Hoerger, M. (2013). *Z_H: An updated version of Steiger's Z and web-based calculator for testing the statistical significance of the difference between dependent correlations.* Retrieved from www.psychmike.com/dependent_correlations.php.

Hsu, N. S., Kraemer, D. J., Oliver, R. T., Schlichting, M. L., & Thompson-Schill, S. L. (2011). Color, context, and cognitive style: Variations in color knowledge retrieval as a function of task and subject variables. *Journal of Cognitive Neuroscience, 23,* 2544–2557.

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience, 21,* 2434–2444.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17,* 825–841.

Kan, I. P., & Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 466–482.

Lambon-Ralph, M. A. L., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 2717–2722.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163.

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. *Brain Research. Cognitive Brain Research, 12,* 467–473.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8,* 976–987.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14792–14797.

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.

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 15855–15860.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55,* 157–166.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage, 14,* 1370–1386.