Neural substrates of attentive listening assessed with a novel auditory Stroop task

Thomas A. Christensen*, Julie L. Lockwood, Kyle R. Almryde and Elena Plante

Laboratory for Brain Imaging of Language, Attention and Memory, Department of Speech, Language and Hearing Sciences, The University of Arizona, Tucson, AZ, USA

A common explanation for the interference effect in the classic visual Stroop test is that reading a word (the more automatic semantic response) must be suppressed in favor of naming the text color (the slower sensory response). Neuroimaging studies also consistently report anterior cingulate/medial frontal, lateral prefrontal, and anterior insular structures as key components of a network for Stroop-conflict processing. It remains unclear, however, whether automatic processing of semantic information can explain the interference effect in other variants of the Stroop test. It also is not known if these frontal regions serve a specific role in visual Stroop conflict, or instead play a more universal role as components of a more generalized, supramodal executive-control network for conflict processing. To address these questions, we developed a novel auditory Stroop test in which the relative dominance of semantic and sensory feature processing is reversed. Listeners were asked to focus either on voice gender (a more automatic sensory discrimination task) or on the gender meaning of the word (a less automatic semantic task) while ignoring the conflicting stimulus feature. An auditory Stroop effect was observed when voice features replaced semantic content as the "to-be-ignored" component of the incongruent stimulus. Also, in sharp contrast to previous Stroop studies, neural responses to incongruent stimuli studied with functional magnetic resonance imaging revealed greater recruitment of conflict loci when selective attention was focused on gender meaning (semantic task) over voice gender (sensory task). Furthermore, in contrast to earlier Stroop studies that implicated dorsomedial cortex in visual conflict processing, interference-related activation in both of our auditory tasks was localized ventrally in medial frontal areas, suggesting a dorsal-to-ventral separation of function in medial frontal cortex that is sensitive to stimulus context.

Keywords: anterior cingulate, anterior insula, attention, auditory Stroop, conflict processing, fMRI, medial frontal gyrus

INTRODUCTION

In the classic visual Stroop test, interference arises when behavioral responses are contingent upon selecting the task-relevant dimension (ink color) over the task-irrelevant information (word meaning) embedded in an incongruent sensory stimulus. Our ability to process word meaning is faster and more automatic than our skill at naming colors, and Stroop proposed that our natural tendency to read words must therefore be suppressed in favor of color naming to successfully complete the task (Stroop, 1935). Since its introduction, many variations of this popular cognitive test for attentional control have been devised (see Lezak et al., 2004), but with very few exceptions, research has focused on reading words or viewing objects (MacLeod, 1991). A growing number of neuroimaging studies have also explored the neural underpinnings of the Stroop effect, but again, most have tested only visual stimuli (Banich et al., 2000; MacDonald et al., 2000; Botvinick et al., 2001; Milham et al., 2001; Matthews et al., 2004; van Veen and Carter, 2005; Carter and Van Veen, 2007; Roberts and Hall, 2008). Results consistently point to the involvement of a frontal lobe network comprising the anterior cingulate cortex, lateral prefrontal cortex, and anterior insula, but they also indicate some regional variation across different Stroop tasks and studies. In addition, these same brain regions are often identified in studies that use other paradigms to measure response

inhibition, such as go/no-go, flanker, and stimulus-response incompatibility tasks (MacDonald et al., 2000; Buchsbaum et al., 2005; Thompson-Schill et al., 2005; Wager et al., 2005; Simmonds et al., 2008). An open question, therefore, is whether these frontal lobe regions consistently associated with Stroop and other conflict tasks are components of a domain-general network, or if these areas can be further dissociated through manipulations that target specific cognitive processes.

Concerns have been raised that the activation patterns reported in the imaging literature to date may have less to do with conflict processing *per se* than with other cognitive or affective operations related to the analysis of other aspects of the Stroop test itself. These concerns include such factors as the specific sensory domain being tested, the particular stimulus feature being presented, or the specific requirements of the behavioral response (Carter et al., 1995; MacDonald et al., 2000; Barch et al., 2001; Wager et al., 2005; Roberts and Hall, 2008; Simmonds et al., 2008). For example, the human anterior cingulate cortex serves diverse functions and has been subdivided into dorsal (dACC) and ventral (vACC) regions based upon a proposed dissociation between their cognitive (dorsal) and emotional (ventral) processing functions, respectively (Mayberg, 1997; Bush et al., 2002). More recently, using a visual counting Stroop task, the dACC has been shown to play a role

in response inhibition, whereas activation of vACC was found to vary with heart rate, a measure of its role in integrating emotional information (Matthews et al., 2004).

To address the issue of whether the sensory domain may influence neuroimaging results, one recent study used an auditory version of the Stroop test known as the "high/low" paradigm (Haupt et al., 2009). In this test, listeners are presented with the words "high" and "low" in either a high or low-pitched voice (Hamers and Lambert, 1972; Shor, 1975). In the SEMANTIC portion of the test, attention is focused on word meaning while ignoring voice pitch, whereas in the VOICE task, the focus is on voice pitch irrespective of word meaning. In accordance with the traditional visual Stroop literature, results using this auditory paradigm led to the conclusion that a stronger interference effect observed in the VOICE task resulted from listeners having to suppress the more automatic processing of semantic information in favor of the voice-pitch component (Haupt et al., 2009). This task recruited strong activation in left lateral prefrontal cortex, a region also believed to be important in implementing attentional control in visual Stroop tasks (MacDonald et al., 2000; Milham et al., 2001; Roberts and Hall, 2008). The "high/low" auditory VOICE task also led to stronger activation of a caudal portion of dACC situated above the corpus callosum and posterior to the anterior commissure (approximate peak Talairach coordinates: −3, −5, 25). This finding is significant because previous visual Stroop studies have found that the rostral portion of dACC is consistently recruited in conflict tasks involving suppression of the reading response (MacDonald et al., 2000; Milham et al., 2001; Roberts and Hall, 2008). One possible conclusion to be drawn from this result is that the caudal dACC may be selectively engaged in auditory task-related interference processing (Haupt et al., 2009). However, ascribing a modality-specific role to any portion of cingulate cortex is problematic because one cannot be certain whether activation in this region is due to auditory conflict processing *per se*, or whether activation is related to some other task dimension. For example, another recent study using the same "high/low" paradigm did not report interference-related activity in the caudal dACC, but found similar patterns of activity in the rostral dACC related to Stroop interference when they compared activation in both the auditory and visual domains (Roberts and Hall, 2008). These results highlight the need for further research regarding localization of function in dACC, vACC, and associated regions of the medial frontal wall (Carter and Van Veen, 2007).

To help clarify the unresolved issues surrounding task- vs. modality-dependent neural activation in Stroop tasks, we developed a novel auditory Stroop-conflict paradigm for use in conjunction with functional magnetic resonance imaging (fMRI). The test consists of two tasks that use gender as a common construct that we varied along two dimensions. We chose gender-based tasks because gender is a highly salient social and emotional construct, and gender cues can be varied in terms of the perceptual and cognitive demands placed on the listener (Most et al., 2007). The two gender tasks we report here share the same sensory modality (auditory), as well as the behavioral requirement to inhibit an inappropriate response. Briefly, participants listened through headphones to a series of common English nouns that were either congruent (e.g., the word "princess" spoken by a woman, or "bull" spoken by a man), incongruent ("father" spoken by a woman, or "lioness" spoken by a man), or

neutral (words that carry no gender meaning, e.g., "orange"). In the VOICE task, listeners were asked to attend to the voice while ignoring the gendered meaning attached to the word, while in the SEMANTIC task, they were asked to attend to the word gender irrespective of the voice. This auditory Stroop paradigm allowed us to generate specific predictions that address several ongoing concerns. First, are modality-specific effects the best explanation for the discrepancies observed in the behavioral and neuroimaging data currently associated with top-down conflict processing? If not, we would expect to find little difference between our results and those reported in previous visual Stroop studies. If true, however, our results should more closely match those of the few existing auditory Stroop studies published to date. Our Stroop test was also designed to address a second concern, namely, that disparate results may also be due to task-specific effects. A common explanation for the interference effect observed in visual Stroop experiments is that the more automatic semantic response (reading the word) must be suppressed in favor of the less automatic and slower sensory response (naming the text color; Stroop, 1935; MacLeod, 1991; Lezak et al., 2004). Although this hypothesis is supported by visual Stroop experiments, it remains unclear whether automatic processing of semantic information is sufficient to explain the effect in other conflict situations. We tested this by developing a Stroop paradigm in which the relative dominance of semantic and sensory feature processing is reversed. The unique feature of this paradigm is that VOICE is the more automatically processed component of the incongruent stimulus, and this information must be suppressed in order to successfully complete the SEMANTIC discrimination task. The paradigm therefore offers a novel test of whether the patterns of frontal lobe activation as previously reported are associated with conflict resolution in general, or instead, are related to the specific demands of cognitive tasks that require suppressing a prepotent response to semantic content.
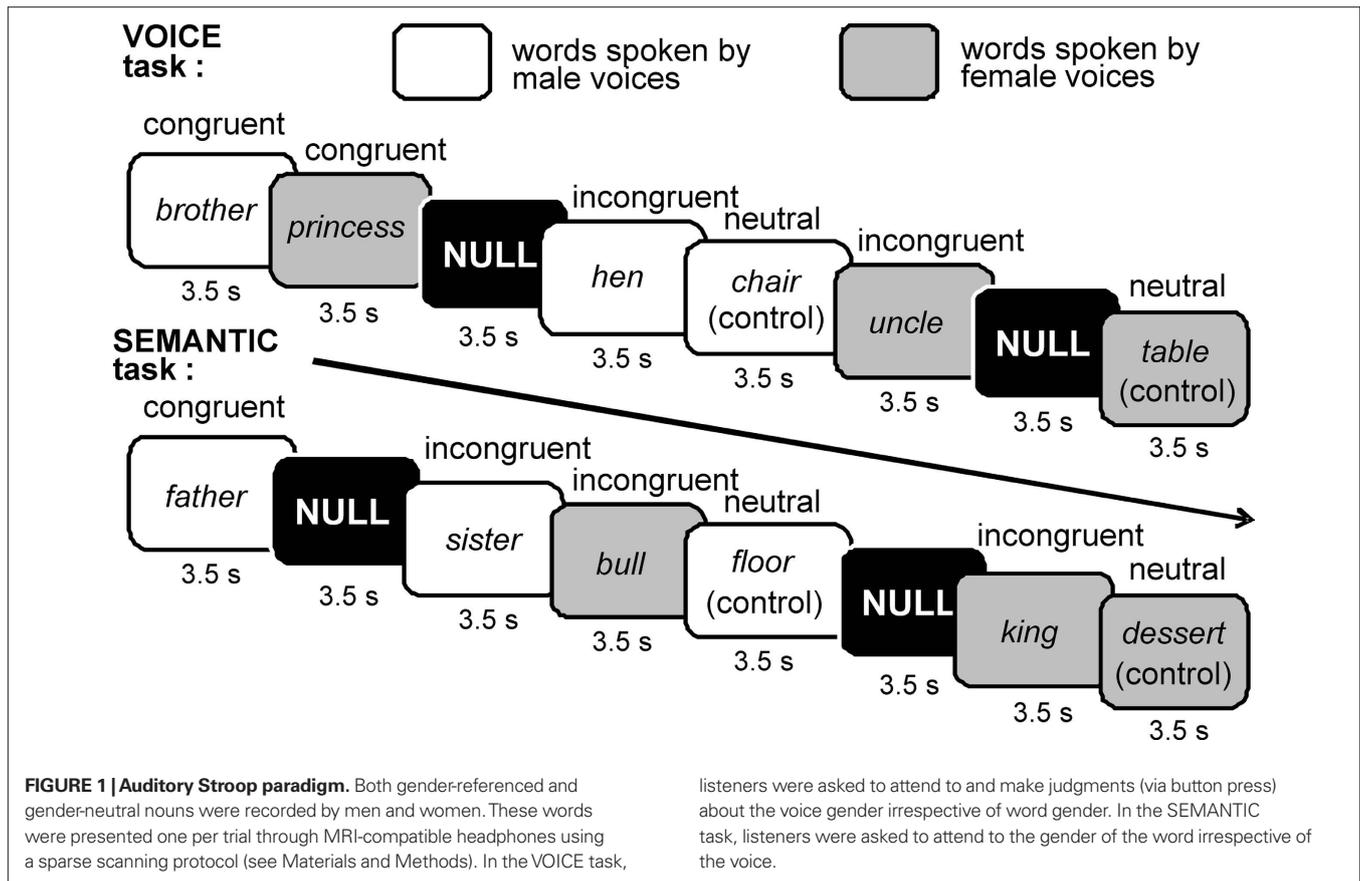
MATERIALS AND METHODS
PARTICIPANTS
The study included 26 healthy volunteers (16 women), mean age was 25.9 years (range 19–53), and all were right-handed, native English speakers recruited from the Tucson community. All participants gave informed consent and were paid for their time. All procedures used in this study were approved by the University of Arizona Human Subjects Protection IRB.

STIMULI
We used an auditory variant of the Stroop test in which words that referenced male or female gender were spoken by either men or women in order to assess attentional control in the presence of either semantic or sensory interference (**Figure 1**). Unlike previous tests, multiple exemplars of both words and voices were used to control for effects of stimulus habituation. Participants listened to a sequential list of common English nouns that were separated into three categories (60 words each) distinguishable by their gender-referenced meaning: (1) typically masculine words (e.g., "brother"), (2) typically feminine words (e.g., "princess"), and (3) semantically neutral words (e.g., "chair") that carry no gender meaning. For each category, 30 nouns were spoken by a male voice and 30 by a female voice, for a total of 180 word trials per task (**Figure 1**). Words were presented in pseudorandom order along with 30 null trials.

FIGURE 1 | Auditory Stroop paradigm. Both gender-referenced and gender-neutral nouns were recorded by men and women. These words were presented one per trial through MRI-compatible headphones using a sparse scanning protocol (see Materials and Methods). In the VOICE task, listeners were asked to attend to and make judgments (via button press) about the voice gender irrespective of word gender. In the SEMANTIC task, listeners were asked to attend to the gender of the word irrespective of the voice.

Stimuli were digitally recorded and edited using Sound Forge software (Sony Creative Software Inc., New York, NY, USA). To control for potential effects of habituation to a single male or female voice, five men and five women recorded the words. The word stimuli were tested in a pilot study involving eight listeners to ensure that the gender of each speaker was clear and consistent. To control for potential effects of word length, all categories included one to three-syllable words, and the average stimulus durations for the three categories were as follows (means ± SEM): male = 667 ± 21 ms; female = 680 ± 19 ms; neutral = 677 ± 18 ms. Recordings were edited to assure maximal signal-to-noise ratios without peak clipping, and were adjusted to equalize loudness percept across individual speech stimuli. The edited recordings were then delivered as individual sound files in the scanner environment using E-Prime software (Psychology Software Tools Inc., Pittsburg, PA, USA), and participants listened to the words through MRI-compatible headphones (Resonance Technology Inc., Northridge, CA, USA). Behavioral responses were collected using two response pads (Lumina System, Cedrus Corp., San Pedro, CA, USA), placed in the participant's right and left hands.

PROCEDURE

While undergoing an fMRI scan (see below), listeners performed two tasks in which they had to differentially use gender-based cues to classify the word presented in each trial. In the VOICE task, they had to attend to the voice gender while ignoring the gender-associated meaning, while in the SEMANTIC task, they had to attend to gender meaning irrespective of the voice. In the VOICE task, participants were asked to press the button in one hand if the word was spoken by a man, and the button in the other hand if it was spoken by a woman. In the SEMANTIC task, they were asked to press the button in one hand if the word meaning was masculine (or feminine), and the button in the other hand for all other words. The 60 gender-congruent trials served as a measure of facilitation, the 60 gender-incongruent trials were used to measure interference, and the 60 trials with neutral words served as controls. While some previous studies have used the contrast between incongruent and congruent responses as a measure of conflict, it is important to note that this contrast includes cognitive components relating to both interference and facilitation (Roberts and Hall, 2008). In order to disambiguate these two effects, we defined interference as the contrast between incongruent and neutral responses, and facilitation as the contrast between congruent and neutral responses, for both our behavioral and neuroimaging data.

Task order in the scanner was counterbalanced across subjects, and button-press responses (accuracy rates and reaction times) were recorded by E-Prime. Immediately following each functional scan, we asked each participant to rate the difficulty of the task using a subjective five-point rating scale, with one being very easy and five being very difficult.

Repeated measures ANOVA was used to evaluate accuracy and reaction times using TASK (focusing attention toward voice or meaning) and CONGRUENCY (congruent, conflict, and neutral trials) as repeated measures across all listeners. However, a

number of neuroimaging studies have found significant gender differences in both speech production and comprehension tasks (Buckner et al., 1995; Sokhi et al., 2005; Tomasi et al., 2008). We therefore also examined the effects of subject and stimulus gender as additional independent variables in our analysis.

MRI IMAGE ACQUISITION

MRI data were acquired with a 3.0T GE Signa VH/i scanner (General Electric Medical Systems, Milwaukee, WI, USA) equipped with an eight-channel RF head coil. Each session began with a T1-weighted structural volume in the axial plane that covered the entire brain (fast-spin echo protocol: TR = 300 ms, TE = minimum, flip angle = 30°, in-plane resolution = 3.44 mm × 3.44 mm × 5 mm). Next, functional images were acquired using a spiral-in/out protocol (TR = 3500 ms, TE = 30 ms, flip angle = 90°, in-plane resolution = 3.75 mm × 3.75 mm × 5 mm) that reduces susceptibility artifacts and spatial distortion (Glover and Law, 2001). A "sparse" scanning sequence collected each volume in the first 1800 ms of the 3500 ms TR, thus eliminating gradient noise during the remainder of the TR (Nebel et al., 2005; Schmidt et al., 2008). The timing of the auditory stimuli was carefully adjusted in E-Prime so that words were presented only during these silent periods. Following the two functional scans, the session concluded with another high-resolution structural volume in the sagittal plane (SPGR protocol: TR = 500 ms; TE = minimum; flip angle = 30°; voxel volume = 1.5 mm × 1 mm × 1 mm). The two structural volumes (one acquired before, and one after, the functional scans), were used to facilitate the localization and co-registration of functional data across participants.

fMRI DATA ANALYSIS

Image analysis was performed with AFNI (http://afni.nimh.nih.gov/afni/; version 2009_12_31_1431) in accordance with the event-related study design. Stimulus timing files were created relative to response times, and only trials resulting in correct responses were included in our GLM model. Data files for individual participants were analyzed separately, followed by a group analysis. Whole-brain inclusion masks were constructed on the basis of peak intensity values using *3dAutomask*, retaining only the largest connected component of the supra-threshold voxels. Standard procedures for data preprocessing were as follows: the first four volumes of each functional scan were discarded to allow for magnetization equilibration, low-frequency drift was corrected, all slices in a volume were aligned to the same temporal origin, all volumes were realigned to the base volume in the series, and the resulting data were spatially smoothed using a 6-mm Gaussian kernel. Outlier data points (e.g., brief movement artifacts) were also excluded from analysis. Data were then normalized to a scale of 0–100%, and functional images were co-registered to the structural data, followed by transformation into common Talairach space.

To extract parameter estimates, the preprocessed fMRI data from each participant was convolved with multiple regressors of interest using a gamma-spline hemodynamic response function using *3dDeconvolve*. The model used a set of seven regressors of interest: two for congruent trials (female word with female voice and male word with male voice), two for incongruent trials (female word with male voice and male word with female voice), two for

neutral trials (neutral word with female voice and neutral word with male voice), and one regressor for nulls. The model also used an additional set of nuisance regressors for six movement parameters: yaw, pitch, roll, and linear displacement along each coordinate axis. Only correct responses were modeled to assure that the patterns of extracted BOLD activity were representative of a positive behavioral outcome.

Brain masks for each participant were corrected to account for possible susceptibility artifacts that can occur near air sinuses, resulting in distortion of the BOLD signal in regions of interest in this study, particularly areas in the ventromedial frontal lobe (Ojemann et al., 1997). Following mask correction, Monte Carlo simulation was used to correct the remaining data for false positive activation. Using a single-voxel threshold of $p = 0.005$, a cluster connection radius of 5.1 mm, and Gaussian smoothing at 6 mm, it was calculated that activation clusters exceeding 17 contiguous voxels (in original space) could be considered significant at a corrected cluster-wise threshold of $p < 0.01$. Clusters that survived this analysis are reported. Cluster coordinates are reported in the space of the Talairach and Tournoux (T–T) atlas included in AFNI (Talairach and Tournoux, 1988; Cox, 1996).
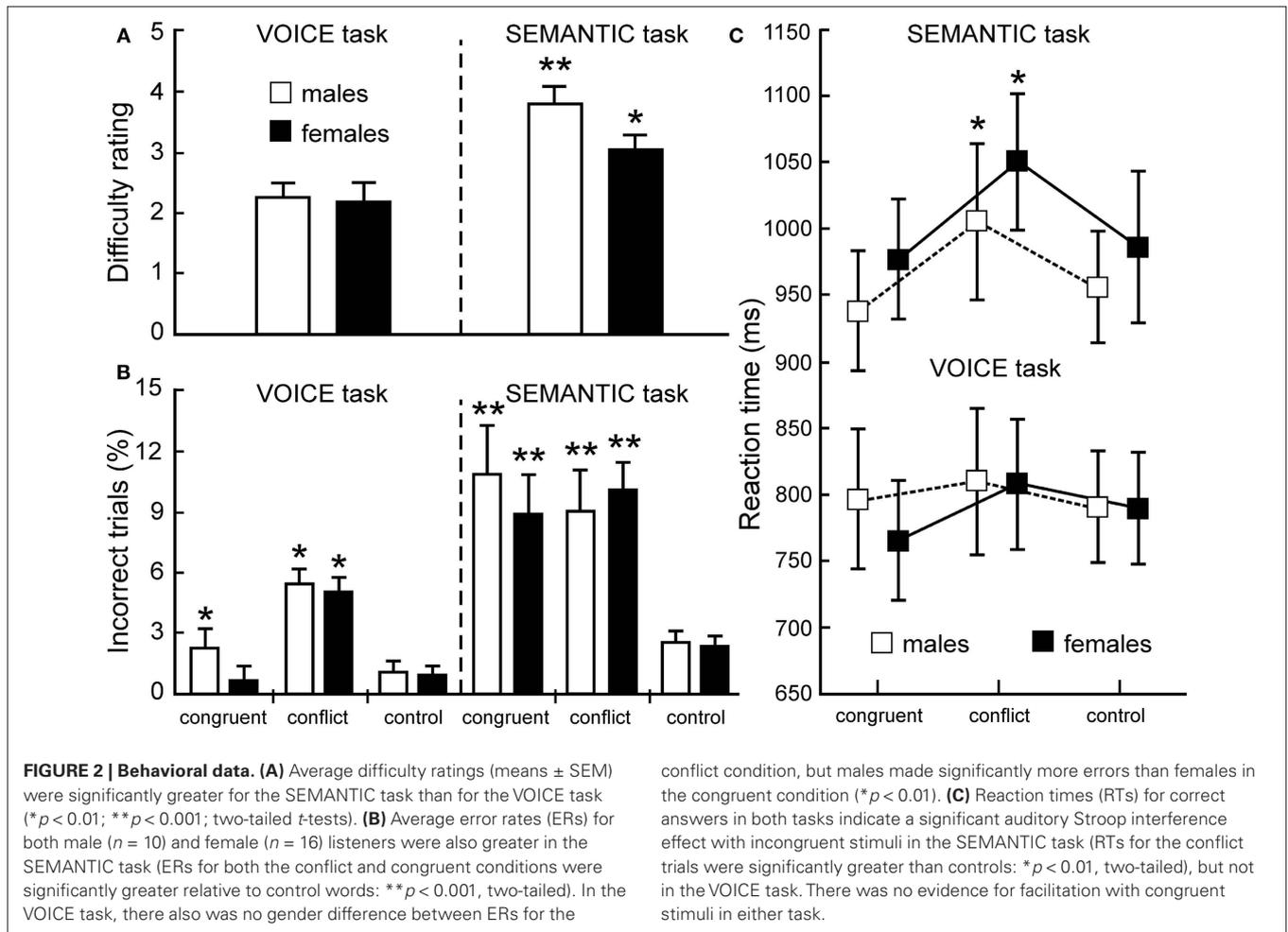
RESULTS

BEHAVIORAL PERFORMANCE DATA

Difficulty ratings and error rates

As shown in **Figure 2A**, both male and female listeners (open and closed bars, respectively) rated the SEMANTIC task as more challenging than the sensory-based VOICE task, but no significant gender effects were observed. In accordance with the difficulty ratings, accuracy measures revealed a robust interference effect, but error rates were significantly higher for the SEMANTIC task (**Figure 2B**). Across all 26 listeners, error rates for incongruent stimuli averaged 9.7% incorrect in the SEMANTIC task compared to 5.5% in the VOICE task ($p < 0.02$, two-tailed t-test). A two-way repeated measures ANOVA of the accuracy data, using a 2 (TASK) by 3 (CONGRUENCY) design, revealed main effects of TASK ($F_{1,25} = 57.26$; $p < 0.0001$) and CONGRUENCY ($F_{2,50} = 17.22$; $p < 0.0001$), as well as a significant TASK × CONGRUENCY interaction ($F_{2,50} = 7.99$; $p < 0.001$).

Reaction times

Correct responses also revealed significantly longer reaction times for SEMANTIC trials than for VOICE trials (**Figure 2C**). Mean reaction times for control words across all listeners were significantly longer for the SEMANTIC task than for the VOICE task ($970 ± 56$ ms vs. $789 ± 50$ ms; $p < 0.001$, two-tailed t-test). A significant auditory Stroop interference effect (defined as the difference between incongruent and neutral responses) was also observed, but once again, only in the SEMANTIC condition. Across all listeners, reaction times for the incongruent words averaged only 19 ms longer than controls in the VOICE task, but this difference increased to 58 ms in the SEMANTIC task (**Figure 2C**). The interference effect was also slightly greater for female than for male listeners, but this difference was not significant. A two-way ANOVA of the reaction time data once again revealed main effects of TASK ($F_{1,25} = 81.01$, $p < 0.0001$) and CONGRUENCY ($F_{2,50} = 42.75$; $p < 0.0001$).

FIGURE 2 | Behavioral data. (A) Average difficulty ratings (means ± SEM) were significantly greater for the SEMANTIC task than for the VOICE task (*$p < 0.01$; **$p < 0.001$; two-tailed t-tests). (B) Average error rates (ERs) for both male ($n = 10$) and female ($n = 16$) listeners were also greater in the SEMANTIC task (ERs for both the conflict and congruent conditions were significantly greater relative to control words: **$p < 0.001$, two-tailed). In the VOICE task, there also was no gender difference between ERs for the conflict condition, but males made significantly more errors than females in the congruent condition (*$p < 0.01$). (C) Reaction times (RTs) for correct answers in both tasks indicate a significant auditory Stroop interference effect with incongruent stimuli in the SEMANTIC task (RTs for the conflict trials were significantly greater than controls: *$p < 0.01$, two-tailed), but not in the VOICE task. There was no evidence for facilitation with congruent stimuli in either task.

In contrast to a strong interference effect from our auditory Stroop protocol, we did not observe a significant facilitation effect (defined as the difference between congruent and neutral responses; **Figure 2C**) in either the VOICE or SEMANTIC task. Reaction times for correct responses to the congruent words averaged only 11 ms less than those for neutral words in the VOICE task, and 13 ms less in the SEMANTIC task. There also was no significant performance difference between male and female listeners in either task, which justified combining all subjects for the group fMRI analysis.

fMRI DATA

Main effects of the VOICE and SEMANTIC tasks

We first investigated regions related to the main effects (all stimulus trials combined) for each of the two auditory tasks separately, using a group ANOVA ($N = 26$; $p < 0.01$, corrected for multiple comparisons). Overall, both tasks resulted in a similar distribution of activated sites across the brain (**Figure 3**; **Table 1**). Although unilateral clusters were found at a corrected threshold of $p < 0.01$, homologous structures in the contralateral hemisphere were also activated at a more relaxed threshold ($p < 0.05$, corrected).

For both tasks, the most striking result was the enhanced activation in several regions of the medial frontal cortex situated below the genu of the corpus callosum. For the VOICE task, this activity

was located in the ventral anterior cingulate cortex (vACC; BA32). vACC activity was right-lateralized, but present in both hemispheres (**Figure 3A**; **Table 1**). This task also led to bilateral activity in other "conflict" loci, including the anterior insular cortex (AIC; BA13). While peak activation was localized to the anterior insula in both hemispheres, it also spread medially through the claustrum and into ipsilateral putamen (**Figure 3A**; **Table 1**). The VOICE task also produced significant increases in dorsal middle frontal gyrus (BA6), superior and inferior parietal lobules (BA7/39/40), and primary auditory cortex (BA41/42), all lateralized to the left hemisphere (**Figure 3A**). Another large cluster was situated in the dorsolateral prefrontal cortex. This cluster crossed through several Brodmann areas (BA9, 46, and 10) and showed a peak of activity along the anterior portion of the inferior frontal sulcus (first cluster in **Table 1**).

For the SEMANTIC task, the largest cluster was localized to the ventral portion of the medial frontal gyrus in BA10/32 (**Figure 3B**; **Table 1**). Activation was also observed in several ventral regions of the lateral prefrontal cortex. The largest cluster was found in the right homolog of Broca's area (pars triangularis; BA45) as well as in regions of AIC lying deep to this area. Activation did not spread into subcortical sites. In the left hemisphere, inferior frontal activity was greatest in
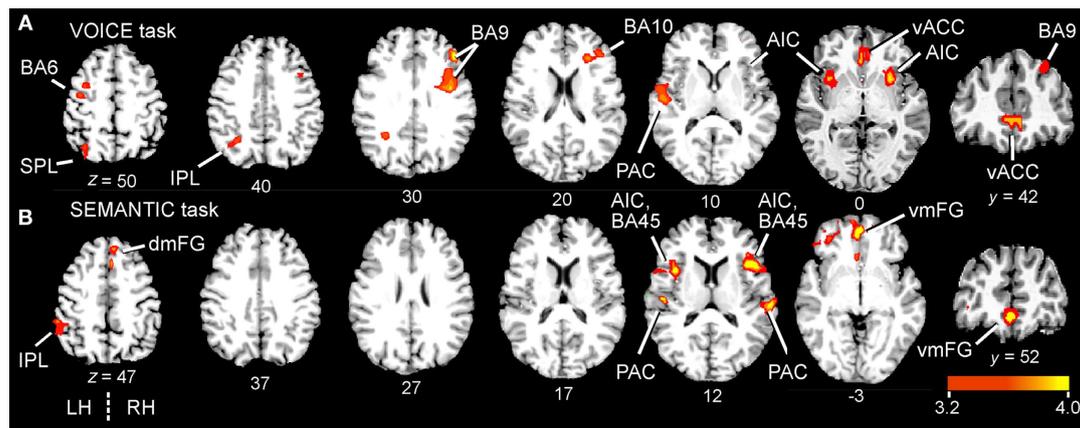
FIGURE 3 | Group-averaged BOLD activation maps comparing the main effects of attentive listening (all correct word trials combined) in the two attentive listening tasks. Statistical maps of positive BOLD activity ($N = 26$) overlaid onto a canonical brain (six axial and one anterior coronal slice; color scale = t-values). All clusters contain at least 17 contiguous voxels at $p < 0.01$ (corrected). Talairach coordinates are in millimeters and right hemisphere (RH) is to the right. Axial slices are from superior to inferior, with the z-coordinate shown beneath each image. The VOICE (sensory) task **(A)** activated both dorsal and ventral frontal regions, while the SEMANTIC task **(B)** engaged mainly ventral areas. Complete lists of activated clusters are given in **Table 1**. AIC, anterior insular cortex; BA, Brodmann area; dmFG, dorsomedial frontal gyrus; IPL, inferior parietal lobule; PAC, primary auditory cortex; SPL, superior parietal lobule; vACC, ventral anterior cingulate cortex; vmFG, ventromedial frontal gyrus.
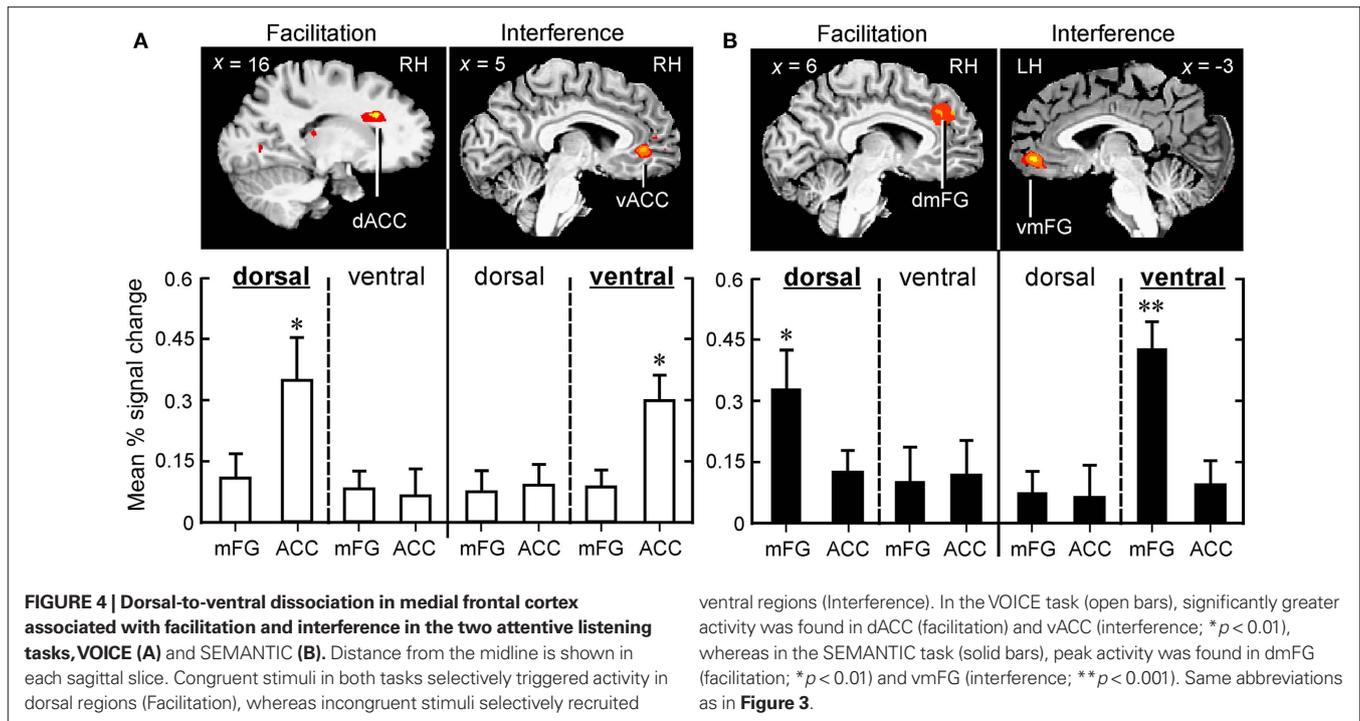
Table 1 | Brain regions, peak Talairach coordinates and volumes of activated clusters showing main effects of attentive listening in the two auditory Stroop tasks.

Region (Brodmann areas)	VOICE task					SEMANTIC task				
	x	y	z	Volume	t-value	x	y	z	Volume	t-value
DORSAL SITES										
Lateral frontal										
Dorsolateral prefrontal (9/46/10)	44	23	30	3797	3.99	–	–	–	–	–
Middle frontal gyrus (6)	−38	−1	53	1406	3.53	–	–	–	–	–
Medial frontal										
Medial frontal gyrus (8/32)	–	–	–	–	–	3	29	45	1477	3.79
Parietal										
Superior parietal lobule (7)	−34	−62	51	1547	3.44	–	–	–	–	–
Inferior parietal lobule (39/40)	**−37**	**−59**	**40**	**1758**	**3.51**	**−50**	**−50**	**45**	**1828**	**3.62**
VENTRAL SITES										
Inferior frontal										
Pars triangularis (45)	–	–	–	–	–	45	21	12[(b)]	2531	3.95
Left anterior insula (13)	**−34**	**17**	**7[(a)]**	**2039**	**3.80**	**−35**	**20**	**10[(c)]**	**1898**	**3.91**
Right anterior insula (13)	28	18	5	1969	3.91	–	–	–	–	–
Medial frontal										
Anterior cingulate cortex (32)	3	36	0	2391	3.85	–	–	–	–	–
Medial frontal gyrus (10/32)	–	–	–	–	–	−2	55	−6	2953	4.09
Temporal										
L. Primary auditory ctx (41/42)	**−55**	**−19**	**11**	**1969**	**3.69**	**−50**	**−14**	**10**	**1477**	**3.77**
R. Primary auditory ctx (41/42)	–	–	–	–	–	61	−16	10	1758	3.97

*Group-averaged data showing peak intensity values and coordinates (mm in Talairach space) for activated clusters (closest Brodmann areas in parentheses). Corrected activation threshold = $p < 0.01$, $N = 26$, and volume is in mm³. Areas showing effects common to both tasks are in **bold** type. [(a)]Activation extends through claustrum into putamen. [(b)]Activation extends into right anterior insula. [(c)]Activation extends into left BA45.*

AIC (**Figure 3B**; **Table 1**). Activated clusters were also observed more posteriorly, in primary auditory cortex (BA41/42; bilateral but right-lateralized) and in the inferior parietal lobule (BA39/40; left hemisphere only). A small cluster was also found in the dorsal portion of medial frontal gyrus (BA6) in the right hemisphere.

FIGURE 4 | Dorsal-to-ventral dissociation in medial frontal cortex associated with facilitation and interference in the two attentive listening tasks, VOICE (A) and **SEMANTIC (B).** Distance from the midline is shown in each sagittal slice. Congruent stimuli in both tasks selectively triggered activity in dorsal regions (Facilitation), whereas incongruent stimuli selectively recruited ventral regions (Interference). In the VOICE task (open bars), significantly greater activity was found in dACC (facilitation) and vACC (interference; *$p < 0.01$), whereas in the SEMANTIC task (solid bars), peak activity was found in dmFG (facilitation; *$p < 0.01$) and vmFG (interference; **$p < 0.001$). Same abbreviations as in **Figure 3**.

Contrast data: interference and facilitation in the VOICE task

Following the initial screening for main effects, contrast analysis was used to separate the activation patterns associated with interference and facilitation for each task. Interference processing (incongruent > neutral) in the VOICE task was associated with increased activation in a subset of the regions identified from the main effects analysis in **Figure 3**. These included specific loci in right vACC (**Figure 4A**) and bilateral AIC (**Figure 5A**). All activity in vACC was situated below the genu of the corpus callosum in the right hemisphere (**Figure 4A**, left). Activity in left AIC once again spread medially through the claustrum and into ipsilateral putamen, while activity in the right hemisphere was localized mainly to the putamen (**Table 2**). We found no significant activity in lateral prefrontal or parietal areas at a corrected threshold of $p < 0.01$.

In sharp contrast to the interference effect in the VOICE task, a facilitation effect (congruent > neutral words) was not associated with increased activity in vACC. Instead, activation associated with congruency was observed in a distinctly dorsal portion of the medial frontal wall, specifically right dACC (BA32; **Figures 4A and 5B**).

To examine these relationships in greater detail, individual subject analysis revealed that 92% (24 out of 26) of all listeners showed interference-related activity in vACC, whereas 80% showed increased activity in dACC associated with congruency. We then used linear regression to compare these two patterns of medial frontal activation with the behavioral measures for the Stroop effect examined in **Figure 2**. A significant correlation was found between peak BOLD signal in dACC and decreased reaction times to congruent stimuli, but not to the increased times observed with incongruent stimuli (**Table 3**). Conversely, significant correlations were observed between activity in vACC and both behavioral measures of Stroop interference, but not between vACC activity and the behavioral responses to congruent stimuli.

FIGURE 5 | Insular and subcortical activation associated with interference (A) and **facilitation (B)** in the VOICE task. Cluster-wise activation threshold = $p < 0.01$, corrected. AIC, anterior insular cortex; dACC, dorsal anterior cingulate; Put, putamen; vACC, ventral anterior cingulate.
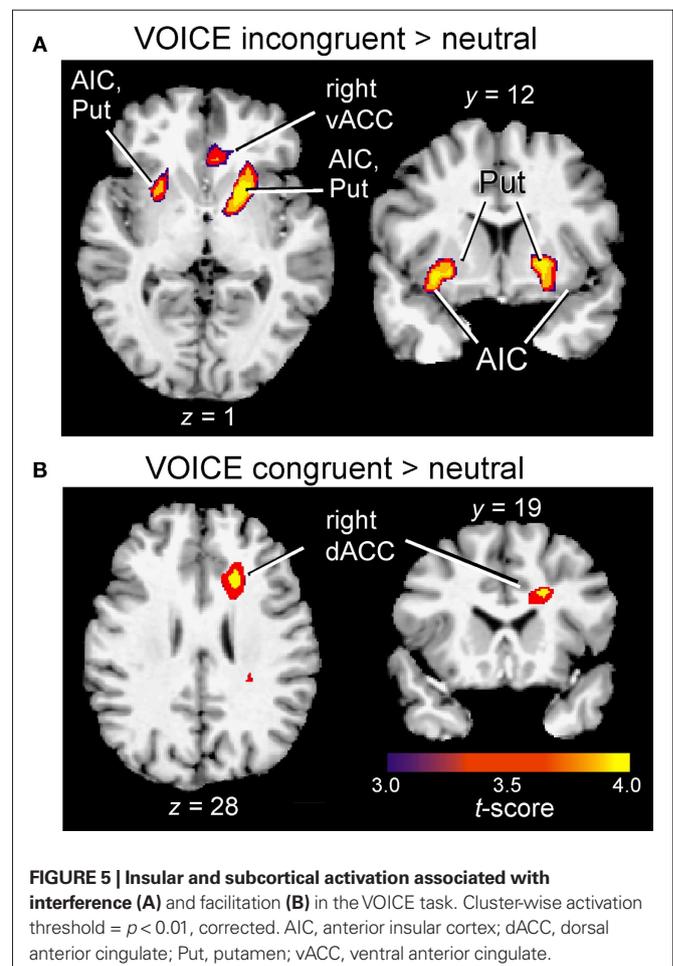
Table 2 | Brain regions, peak coordinates and volumes of activated clusters showing a significant effect of stimulus-level interference (incongruent > neutral words) and facilitation (congruent > neutral words) for the two attentive listening tasks.

Condition/region	VOICE task					SEMANTIC task				
	x	y	z	Volume	t-value	x	y	z	Volume	t-value
INTERFERENCE										
Ventral sites										
Inferior frontal										
L. pars triangularis (45)	–	–	–	–	–	–47	21	6	1266	3.75
R. pars triangularis (45)	–	–	–	–	–	39	31	6[b]	2461	3.98
Anterior insula (13)	**–36**	**12**	**–4**[a]	**1688**	**3.56**	**–41**	**11**	**–4**	**1477**	**3.76**
Medial frontal										
Anterior cingulate (24/32)	8	31	1	1898	3.73	–	–	–	–	–
Medial frontal gyrus (10/32)	–	–	–	–	–	–5	50	–3	2039	3.95
Subcortical										
Caudate head	–	–	–	–	–	11	13	5	1336	3.57
Putamen	21	13	0	1828	3.88	–	–	–	–	–
FACILITATION										
Dorsal sites										
Anterior cingulate (32)	16	22	28	1828	3.90	–	–	–	–	–
Medial frontal gyrus (8)	–	–	–	–	–	4	45	38	2180	3.59

*Group-averaged data showing peak intensity values and coordinates (mm in Talairach space) for activated clusters (Brodmann areas in parentheses). Activation threshold = $p < 0.01$ (corrected), N = 26, and volume is in mm^3. An interference effect common to both tasks was found in left anterior insula (**bold**). [a]Activation extends into left claustrum and putamen. [b]Activation extends into right anterior insula.*

Table 3 | Linear regression analysis of peak BOLD activity in medial frontal clusters vs. behavioral responsiveness in the two auditory Stroop tasks.

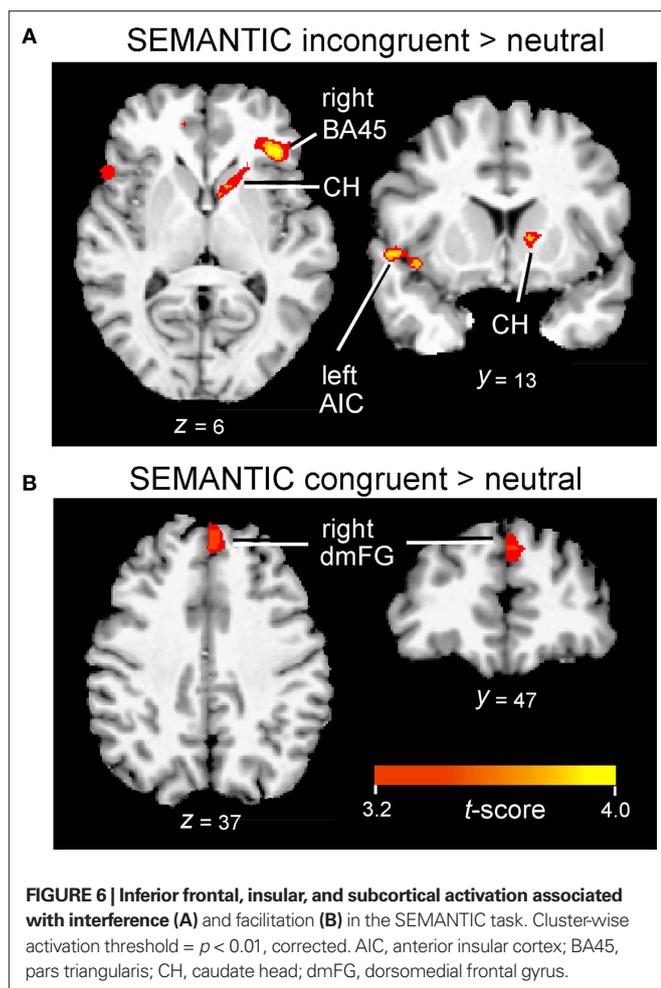| Medial frontal region	TASK/ condition	Behavioral measures			
		Accuracy		Reaction time	
		R^2	p-value	R^2	p-value
VOICE					
dACC	Congruent	0.275	0.071	**0.411**	**0.033**
	Incongruent	0.222	0.063	0.252	0.158
vACC	Congruent	0.044	0.217	0.179	0.140
	Incongruent	**0.288**	**0.031**	**0.313**	**0.047**
SEMANTIC					
dmFG	Congruent	**0.365**	**0.020**	**0.331**	**0.039**
	Incongruent	0.189	0.244	0.110	0.290
vmFG	Congruent	0.081	0.110	0.077	0.725
	Incongruent	**0.655**	**0.013**	**0.810**	**0.009**

*All listeners combined (N = 26). Abbreviations as in **Figure 3**. Significant correlations at $p < 0.05$ are in **bold**.*

Contrast data: Interference and facilitation in the SEMANTIC task

The interference contrast in the SEMANTIC task also revealed activation in several anterior regions, including clusters along the medial frontal wall, in anterior insula, and in subcortical regions, but the specific activation loci were distinctly different than those revealed in the VOICE task. Interference processing in the SEMANTIC task recruited a large cluster of activation more anteriorly in the ventral portion of medial frontal gyrus (BA10/32; **Figure 4B**, left). Also unlike the VOICE task, interference in the SEMANTIC task also recruited lateral prefrontal areas previously implicated in applying attentional control, with 80% of listeners showing increased activity in the inferior frontal gyrus, with a peak localized to right pars triangularis (BA45; **Figure 6A**; **Table 2**). A small cluster in the left anterior insula was also associated with interference in the SEMANTIC task (**Table 2**), but unlike the VOICE task data, no activity was observed in the claustrum or anterior putamen (**Figure 6A**). Instead, subcortical activation was located in the right caudate head (**Table 2**). As with the VOICE results, facilitation in the SEMANTIC task was associated with increased activity in a dorsal region of the right medial frontal wall, but this activity was localized to medial frontal gyrus (BA8; **Figure 6B**).

Individual subject analysis revealed that none of the 26 listeners showed significant interference-related activity along the dorsomedial frontal wall at a corrected threshold of $p < 0.01$, whereas 96% (25 out of 26) showed significant activation in the abovementioned ventral regions (**Table 2**). Regression analysis comparing medial frontal activation with the behavioral measures in the SEMANTIC task revealed significant correlations between peak BOLD signal in dmFG and behavioral responses to congruent stimuli, but not to the incongruent stimuli (**Table 3**). Conversely, significant correlations were observed between activity in vmFG and both behavioral measures of Stroop interference, but not between vmFG activity and the behavioral responses to congruent stimuli.

FIGURE 6 | Inferior frontal, insular, and subcortical activation associated with interference (A) and facilitation (B) in the SEMANTIC task. Cluster-wise activation threshold = *p* < 0.01, corrected. AIC, anterior insular cortex; BA45, pars triangularis; CH, caudate head; dmFG, dorsomedial frontal gyrus.

DISCUSSION

ORIGIN OF THE STROOP EFFECT

Since the classic Stroop effect was first described in the early twentieth century (Stroop, 1935), several theories have sought to explain the functional mechanisms underlying the robust interference effect in the color-naming (sensory) task, along with minimal interference in the word-reading (semantic) task (Treisman and Fearnley, 1969). The *selective attention to reading hypothesis* is based on the premise that interference occurs because attentional capacity is limited, and subjects are unable to fully suppress the automatic tendency to read words (Klein, 1964). An alternative view, the *response competition hypothesis* (as proposed by Stroop himself) states that the more familiar and practiced response (irrespective of stimulus modality) is more difficult to suppress than the less familiar and unpracticed response. The first of these two hypotheses is based mainly on sensory aspects of stimulus processing, whereas the second incorporates top-down response-inhibition mechanisms. These two hypotheses allowed us to make informed predictions about the outcome of the present auditory Stroop experiment.

Our data strongly favor the *response competition hypothesis* on several grounds. In accordance with Stroop's original hypothesis, our behavioral results show that the processing of voice information, not the semantic content of the words, was clearly the more prepotent

stimulus feature driving the interference effect in the present study. Acoustic variation in voice quality enables a listener to easily distinguish adult male and female voices (Klatt and Klatt, 1990), and our behavioral data confirmed that this was the easier of the two gender parameters to process. We found a more pronounced Stroop interference effect in the SEMANTIC task (in which listeners had to suppress incongruent voice information) than in the VOICE task (in which they had to suppress incongruent semantic information). These results demonstrate that automatic processing of semantic information *per se* cannot explain the Stroop interference effect in all cases. Furthermore, the effect of gender-based conflict does not depend on the gender of the listener, as these relationships held true for both male and female participants (**Figure 2C**). While our new findings do not resolve the ongoing debate regarding the origins of the Stroop effect (MacLeod, 1991), they point to the importance of looking beyond the classic visual Stroop paradigm introduced 75 years ago in order to properly address this longstanding issue.

NEURAL SUBSTRATES FOR AUDITORY STROOP INTERFERENCE EFFECTS

As shown in **Figure 3**, both auditory Stroop tasks revealed a distributed array of structures previously identified as attention network components, particularly lateral prefrontal cortex – an area believed to execute attentional control operations, anterior cingulate cortex – a region that may serve a more specific role in monitoring the occurrence of response conflict (MacDonald et al., 2000; Barch et al., 2001; Milham et al., 2001; Roberts and Hall, 2008), and anterior insular cortex – believed to play a key role in human awareness (Craig, 2009). A few neuroimaging studies using the auditory "high/low" conflict paradigm have suggested that auditory Stroop follows the same logic as visual Stroop, concluding that discriminating low from high-pitched spoken words is a sensory task that lacks automaticity, and thus produces Stroop-like interference in the presence of conflicting, prepotent semantic information (Roberts and Hall, 2008; Haupt et al., 2009). One of these studies also found common patterns of frontal BOLD activity when results from the "high/low" and visual Stroop tests were compared in the same subject group (Roberts and Hall, 2008). It was concluded that several areas, including left rostral anterior cingulate (BA32), bilateral inferior frontal gyrus (BA44), and anterior insula (BA13), serve as a supramodal network for conflict processing. It is important to note that, unlike the present auditory Stroop paradigm, these studies used a Stroop test in which the VOICE task was shown to be more difficult than the SEMANTIC task. In contrast, our experiment was designed so that voice gender, not word meaning, served as the "to-be-ignored" stimulus dimension in the more challenging SEMANTIC task (**Figure 2**). Thus, our neuroimaging results were obtained in a behavioral context that offers a stark contrast to previous Stroop studies, both auditory and visual. We also separated interference and facilitation effects in our tasks by employing event-related sparse scanning of incongruent, congruent, and neutral word stimuli using a selective-attention design across the two tasks.

Identification of conflict processing areas in ventral cingulate cortex and medial frontal gyrus

Using this new paradigm, we found strong behavioral evidence for Stroop interference when listeners focused on the more attentionally demanding stimulus dimension, gender-referenced meaning,

in the SEMANTIC task (**Figure 2C**). We also observed different patterns of interference-related BOLD activity in the medial frontal lobes under the two listening conditions. The most surprising finding, however, was that the interference-related activation in both tasks was distinctly more ventral than that typically associated with either the classical visual Stroop tests or previous auditory Stroop studies (Roberts and Hall, 2008; Haupt et al., 2009). Rather than the typical pattern of dorsal cingulate activation, we observed increased activity in vACC (BA24/32) for the less demanding VOICE task, and activation more anteriorly in ventral medial frontal gyrus (BA10/32) for the more challenging SEMANTIC task (**Figures 4A,B**, respectively). The ventral subdivision of ACC is typically associated with tasks that tap into human emotions and pragmatics (Simpson et al., 2001; Matthews et al., 2004). We therefore found it striking that the interference condition in both of our gender-based auditory Stroop tasks evoked increased activity in regions of the medial frontal wall that lie ventral to the genu of the corpus callosum. Gender provides a powerful social construct through which humans structure their everyday lives (Most et al., 2007), thus we believe that the selective engagement of these ventral regions is likely due to the strong gender-based context associated with our tasks.

In contrast, processing of congruent stimuli in both tasks led to increased activation only in dorsal portions of the right medial frontal cortex (**Figures 4A,B**). These results provide fresh support for the hypothesis that a dorsal subregion of the rostral cingulate zone is reliably engaged in a wide range of cognitive tasks because it performs a generalized monitoring function across modalities, evaluating the demand for attentional control by scanning for the occurrence of response conflict (Barch et al., 2001). Other regions of the cingulate/medial frontal cortex (ventral regions in our study, or caudal regions in "high/low" Stroop studies), however, appear to be more specialized in actively processing conflict when competition between incompatible stimulus features is present. Further empirical studies involving a wider range of sensory modalities, response domains, and stimulus features are needed to help resolve this important question regarding specialization of function across different regions of the medial frontal wall.

Inferior frontal cortex

In addition to medial frontal cortex, previous imaging studies of response inhibition have also consistently revealed activity in inferior frontal and insular cortices (Banich et al., 2000; Simmonds et al., 2008), but the question of whether the same brain regions are engaged in response inhibition across different tasks, especially in the auditory domain, remains unanswered. A recent meta-analysis of neuroimaging data compares the results from 21 experiments using traditional visual Stroop tasks with 19 studies of other visual conflict tasks, illustrating the key importance of inferior frontal and insular cortices in conflict processing (Roberts and Hall, 2008). In the present study, we also found support for significant involvement of both of these conflict regions using our new auditory Stroop tasks. In accordance with previous reviews of the conflict processing literature, incongruent but not congruent stimuli in our SEMANTIC task revealed strong activation of prefrontal cortex (focus in pars triangularis, BA45) with activity lateralized to the right hemisphere (**Figure 6A**). This activation lies just anterior to

BA44, a region previously identified as being active in a variety of visual conflict tasks (Roberts and Hall, 2008). The stronger right hemisphere involvement agrees with the conventional view that the right inferior frontal cortex is particularly important for tasks that require the processing of social and emotional elements of language (Mitchell et al., 2003; Beaucousin et al., 2007), whereas the homologous region in the left hemisphere (i.e., Broca's area) is more active during linguistic analysis (Bookheimer, 2002; Hickok and Poeppel, 2007).

Prefrontal and medial frontal regions are often activated concurrently, indicating a close functional link between these two attention areas (Carter et al., 1995; Thompson-Schill et al., 1997; Botvinick et al., 2001; Posner and Rothbart, 2007). It has been suggested that once the medial frontal cortex begins to actively monitor conflict, it then alerts a cognitive control system in prefrontal cortex that helps resolve conflict by first comparing the two conflicting inputs, and then biasing information processing toward the attended stimulus (Botvinick et al., 2001; Heekeren et al., 2004). This hypothesis may help explain concurrent activation of ventral medial frontal and inferior prefrontal areas in our SEMANTIC task, but not in the less demanding VOICE task (**Table 2**), possibly because the more automatic nature of the VOICE task would not require the same level of executive-control bias from prefrontal cortex to complete the task. In contrast to studies of visual Stroop interference, therefore, our data suggest that ventromedial frontal gyrus, deep to prefrontal language areas, may be selectively involved in exerting executive control over conflicting auditory information streams.

Anterior insular cortex

Interference-related activity in anterior insula was localized to the left hemisphere in both of our Stroop tasks (**Table 2**), consistent with the view that left anterior insula is an important component of a generic, supramodal network for conflict detection and monitoring (Wager et al., 2005). In primates, the anterior (agranular) insula, which integrates sensory information with inputs from the amygdala, has strong projections to the medial frontal cortex (Dupont et al., 2003), the region that was engaged when listeners responded to incongruent stimuli in our SEMANTIC task. Thus, if the rostral extent of the medial frontal cortex receives input from the insula, these two areas together could form the anatomical foundation for a frontal network that actively detects (insula) and monitors (medial frontal) conflict arising from lower-level sensory regions. Based on its other connections, this frontal network can then adjust the brain's limited attentional resources according to changing situational demands (Craig, 2009). For example, our previous data indicate that subcortical regions such as the putamen and caudate nucleus are also key to proper executive functioning in attentive listening (Christensen et al., 2008), and the differential activity we observed in these areas for the two tasks in the present study is also likely related to their involvement under different levels of demand for attentional control. While our data do not allow us to address this issue fully, they underscore the importance of considering the role of subcortical activity in cognitive functions that are traditionally associated with cerebral cortex (Kreitzer and Malenka, 2008; Lieberman, 2008; van Schouwenburg et al., 2010).

CONCLUSION

Using a novel auditory Stroop paradigm, we demonstrate a significant interference effect with gender-typical nouns spoken by gender-mismatched voices. Using several independent measures of conflict management, our results support the conclusion that suppression of semantic processing cannot explain all instances of the Stroop effect in the auditory domain. Rather, the behavioral results provide fresh evidence that Stroop interference is due to a failure to suppress whichever stimulus attribute is processed more automatically (in this case, voice gender). Secondly, our fMRI results differ in several major respects from previous neuroimaging studies. Specifically, we found no evidence for a selective role of dorsomedial frontal structures in auditory Stroop interference processing. Instead, our results provide support for a dorsal-to-ventral dissociation of function along the medial frontal wall, which links ventral regions to interference processing in emotionally salient, cognitively challenging tasks (in this case, discriminating nouns by their gender association), and dorsal regions to the more global task of conflict monitoring.

REFERENCES

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., Kramer, A. F., Liang, Z.-P., Wright, A., Shenker, J., and Magin, R. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *J. Cogn. Neurosci.* 12, 988–1000.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., and Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb. Cortex* 11, 837–848.

Beaucousin, V., Lacheret, A., Turbelin, M.-R., Morel, M., Mazoyer, B., and Tzourio-Mazoyer, N. (2007). fMRI study of emotional speech comprehension. *Cereb. Cortex* 17, 339–352.

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25, 151–188.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Buchsbaum, B. R., Greer, S., Chang, W.-L., and Berman, K. (2005). Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Hum. Brain Mapp.* 25, 35–45.

Buckner, R. L., Raichle, M. E., and Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *J. Neurophysiol.* 74, 2163–2173.

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., and Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proc. Natl. Acad. Sci. U.S.A.* 99, 523–528.

Carter, C. S., Mintun, M., and Cohen, J. D. (1995). Interference and facilitation effects during selective attention: an H215O PET study of Stroop task performance. *Neuroimage* 2, 264–272.

Carter, C. S., and Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cogn. Affect. Behav. Neurosci.* 7, 367–379.

Christensen, T. A., Antonucci, S. M., Lockwood, J. L., Kittleson, M., and Plante, E. (2008). Cortical and subcortical contributions to the attentive processing of speech. *Neuroreport* 19, 1101–1105.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.

Craig, A. D. (2009). How do you feel now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.

Dupont, S., Bouilleret, V., Hasboun, D., Semah, F., and Baulac, M. (2003). Functional anatomy of the insula: new insights from imaging. *Surg. Radiol. Anat.* 25, 113–119.

Glover, G. H., and Law, C. S. (2001). Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn. Reson. Med.* 46, 515–522.

Hamers, J. F., and Lambert, W. E. (1972). Bilingual interdependences in auditory perception. *J. Verbal Learn. Verbal Behav.* 11, 303–310.

Haupt, S., Axmacher, N., Cohen, M., Elger, C., and Fell, J. (2009). Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. *Hum. Brain Mapp.* 30, 3043–3056.

Heekeren, H. R., Marrett, S., Bandettini, P. A., and Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature* 431, 859–862.

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.

Klatt, D. H., and Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.* 87, 820–857.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *Am. J. Psychol.* 77, 576–588.

Kreitzer, A. C., and Malenka, R. C. (2008). Striatal plasticity and basal ganglia circuit function. *Neuron* 60, 543–554.

Lezak, M., Howieson, D., and Loring, D. (2004). *Neuropsychological Assessment.* New York: Oxford University Press.

Lieberman, P. (2008). Cortical-striatal-cortical neural circuits, reiteration, and the "narrow faculty of language." *Behav. Brain Sci.* 31, 527–528.

MacDonald, A. W. I., Cohen, J. D., Stenger, V. A., and Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1838.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203.

Matthews, S. C., Paulus, M. P., Simmons, A. N., Nelesen, R. A., and Dimsdale, J. E. (2004). Functional subdivisions within anterior cingulate cortex and their relationship to autonomic nervous system function. *Neuroimage* 22, 1151–1156.

Mayberg, H. S. (1997). Limbic-cortical dysregulation: a proposed model of depression. *J. Neuropsychiatry. Clin. Neurosci.* 9, 471–481.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., and Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cogn. Brain Res.* 12, 467–473.

Mitchell, R. L., Elliott, R., Barry, M., Cruttenden, A., and Woodruff, P. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia* 41, 1410–1421.

Most, S. B., Sorber, A. V., and Cunningham, J. G. (2007). Auditory Stroop reveals implicit gender associations in adults and children. *J. Exp. Soc. Psychol.* 43, 287–294.

Nebel, K., Stude, P., Wiese, H., Müller, B., de Greiff, A., Forsting, M., Diener, H.-C., and Keidel, M. (2005). Sparse imaging and continuous event-related fMRI in the visual domain: a systematic comparison. *Hum. Brain Mapp.* 24, 130–143.

Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., and Conturo, T. E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage* 6, 156–167.

Posner, M. I., and Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annu. Rev. Psychol.* 58, 1–23.

Roberts, K. L., and Hall, D. A. (2008). Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory Stroop tasks. *J. Cogn. Neurosci.* 20, 1063–1078.

Schmidt, C., Zaehle, T., Meyer, M., Geiser, E., Boesiger, P., and Jancke, L. (2008). Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. *Hum. Brain Mapp.* 29, 46–56.

Shor, R. E. (1975). An auditory analog of the Stroop test. *J. Gen. Psychol.* 93, 281–288.

Simmonds, D. J., Pekar, J. J., and Mostofsky, S. H. (2008). Meta-analysis of go/no-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* 46, 224–232.

Simpson, J. J., Snyder, A. Z., Gusnard, D. A., and Raichle, M. E. (2001). Emotion-induced changes in human medial prefrontal cortex: I. During cognitive task performance. *Proc. Natl. Acad. Sci. U.S.A.* 98, 683–687.

Sokhi, D. S., Hunter, M. D., Wilkinson, I. D., and Woodruff, P. W. R. (2005). Male and female voices activate distinct regions in the male brain. *Neuroimage* 27, 572–578.

Stroop, J. (1935). Studies of interference in serial verbal reaction. *J. Exp. Psychol.* 18, 643–662.

Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain.* New York: Thieme.

Thompson-Schill, S. L., Bedny, M., and Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Curr. Opin. Neurobiol.* 15, 219–224.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. (1997).

Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14792–14797.

Tomasi, D., Chang, L., Caparelli, E. C., and Ernst, T. (2008). Sex differences in sensory gating of the thalamus during auditory interference of visual attention tasks. *Neuroscience* 151, 1006–1015.

Treisman, A., and Fearnley, S. (1969). The Stroop test: selective attention to colours and words. *Nature* 222, 437–439.

van Schouwenburg, M. R., den Ouden, H. E. M., and Cools, R. (2010). The human basal ganglia modulate frontal-posterior connectivity during

attention shifting. *J. Neurosci.* 30, 9910–9918.

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

Wager, T. D., Sylvester, C.-Y. C., Lacey, S. C., Nee, D. E., Franklin, M., and Jonides, J. (2005). Common and unique components of response inhibition revealed by fMRI. *Neuroimage* 27, 323–340.