From Perception to Sentence Comprehension: The Convergence of Auditory and Visual Information of Language in the Left Inferior Frontal Cortex

Fumitaka Homae,* Ryuichiro Hashimoto,* Kyoichi Nakajima,†
Yasushi Miyashita,† and Kuniyoshi L. Sakai*,‡,1

*Department of Cognitive and Behavioral Science, Graduate School of Arts and Sciences, The University of Tokyo, Komaba,
Tokyo, Japan; †Department of Physiology, The University of Tokyo School of Medicine, Tokyo, Japan; and
‡CREST, Japan Science and Technology Corporation, Kawaguchi-shi, Japan

We used functional magnetic resonance imaging (fMRI) to characterize cortical activation associated with sentence processing, thereby elucidating where in the brain auditory and visual inputs of words converge during sentence comprehension. Within one scanning session, subjects performed three types of tasks with different linguistic components from perception to sentence comprehension: nonword (N_{AV}; auditory and visual), phrase (P; either auditory or visual), and sentence (S; either auditory or visual) tasks. In a comparison of the P and N_{AV} tasks, the angular and supramarginal gyri showed bilateral activation, whereas the inferior and middle frontal gyri showed left-lateralized activation. A comparison of the S and P tasks, together with a conjunction analysis, revealed a ventral region of the left inferior frontal gyrus (F3t/F3O), which was sentence-processing selective and modality-independent. These results unequivocally demonstrated that the left F3t/F3O is involved in the selection and integration of semantic information that are separable from lexico-semantic processing. © 2002 Elsevier Science (USA)

Key Words: language; sentence processing; lexico-semantics; modality; frontal cortex; fMRI; conjunction analysis.

INTRODUCTION

Previous models of language processing, which were based on neuropsychological data, have proposed that auditory and visual inputs partially converge during word processing (Geschwind, 1965; Patterson and Shewell, 1987). Recently, Price (2000) reviewed functional neuroimaging studies on single word processing, and proposed a model in which auditory and visual inputs of single words converge on the left extrasylvian temporo-parietal regions. This model did not incorporate the left frontal cortex as an essential component, because the role of this area remains still unclear. Nevertheless, previous studies have reported activation in the left inferior frontal gyrus (IFG) during word recognition under either auditory or visual conditions (e.g., Mazoyer *et al.,* 1993; Bookheimer *et al.,* 1995). It is also highly controversial whether the left IFG critically subserves sentence processing (Sakai *et al.,* 2001). Sentence comprehension involves the processes of recognizing individual spoken or written words and appreciating their meanings and syntactic features, as well as assigning semantic values of propositional and discourse-level meanings (Caplan and Waters, 1999). The representations constructed during sentence comprehension are constrained by grammatical phrase-formation information, which delimits the range of possible combinations and partially determines the interpretations for the combinations (Gibson and Pearlmutter, 1998). Therefore, sentence processing critically involves selection and integration of various semantic attributes into more specific concepts, which would be probably amodal processes. The goal of the present study is to search for neural correlates of sentence comprehension that subsist in the convergence of auditory and visual inputs of words.

To accomplish this goal, it would be necessary to address two fundamental issues. First, any tasks to target language processing with sensory inputs, such as lexical decision tasks, interact with sensory factors. Second, sentence comprehension tasks interact with lexical processing, because sentence processing inherently involves lexical factors. As to the first issue, baseline conditions are required to be equated between auditory and visual inputs, to examine commonalities and differences of their effects on cortical responses. In

a positron emission tomography (PET) study, Petersen *et al.* (1989) contrasted a verb generation task and a repetition task under either auditory or visual condition, and found overlapping, but anatomically dissimilar, activation in the inferior frontal cortex [Brodmann's area (BA) 47]. This difference may be due to the difference in sensory factors involved not only in the generation tasks but in the repetition tasks, which were not entirely equated. Similarly, Howard *et al.* (1992) contrasted word repetition under the auditory condition with hearing spoken words presented backwards, while word reading was contrasted with seeing stimuli written in false fonts. In an fMRI study, Chee *et al.* (1999) contrasted an auditory semantic task (abstract/concrete words) with syllable counting, while a visual semantic task was contrasted with case identification. In these previous studies, the differential activation patterns found in word tasks under the two modality conditions confounded differences in baseline tasks, because relative activation patterns in one experimental task critically depends on the choice of baseline tasks (Newman *et al.,* 2001; Gusnard and Raichle, 2001). Therefore, it would be ideal to use an identical task with basic language processing as the baseline task under both auditory and visual conditions. To meet this requirement in the present fMRI study, a cross-modal nonword (N_{AV}) task for matching auditory and visual nonwords was newly introduced as an identical baseline task.

As to the second issue, it should be noted that sentence processing involves various cognitive factors, which are absent in simpler baseline tasks. In the study of Michael *et al.* (2001), they compared listening and reading comprehension of sentences, while sentence tasks were contrasted with an identical fixation condition. However, any differential activation patterns found in their study were not free from lexical factors as well as sensory factors. In contrast, Bottini *et al.* (1994) directly compared a sentence comprehension task (plausible/implausible sentences) with a lexical decision task under the visual condition, and reported widespread activation in both hemispheres. A simple direct comparison between the two tasks might reveal regions related to sentence processing, but it is also likely that lexical processing may be simply enhanced in the sentence task because of their interactions. Generally speaking, there is an inherent problem in the cognitive subtraction method, such that many studies overlooked an "interaction" between newly introduced and already existing cognitive components (Price and Friston, 1997; Jennings *et al.,* 1997; Sidtis *et al.,* 1999). The present study employed similar sentence and lexical tasks that were established by Bottini *et al.* (1994), together with a conjunction analysis to eliminate possible activation in cortical language areas, which may reflect interactions between sentence and lexical processing.

While overcoming these problems in the study of sentence processing, we devised a new paradigm to reliably identify sentence-processing selective regions under the auditory and visual conditions. The baseline N_{AV} task involves both speech-sound processing and letter-form processing, in which phonologically equivalent nonwords were presented auditorily and visually at the same time. By using nonwords that were phonologically different between the modalities as probe stimuli, we ensured that the subjects paid full attention to both the auditory and visual stimuli. In contrast, a phrase (P) task primarily involves lexico-semantic processing. In the phrase task under the auditory condition (P_A), we auditorily presented the phrases from dialogues in a completely randomized order. Similarly, in the phrase task under the visual condition (P_V), the same sets of phrases used in the P_A task were visually presented. We used nonwords as probe stimuli, so that the P_A and P_V tasks explicitly required lexical decisions. Furthermore, a sentence (S) task requires comprehension of sentence meaning. In the sentence task under the auditory condition (S_A) or visual condition (S_V), phrases used in the P tasks were presented in the order of the sentences in dialogues. We used semantically anomalous (implausible) phrases as probe stimuli, so that the S_A and S_V tasks explicitly required decisions on consistency about sentence meaning. The S_A and the P_A tasks were linguistically equivalent to the S_V and the P_V tasks, respectively. We first established the equivalence of cortical activation between the P_A and P_V tasks, by comparing each task with the identical N_{AV} task. We then tried to unequivocally identify sentence-processing selective regions, by directly comparing the S task with the P task in each modality, together with a conjunction analysis.

METHODS

Subjects

Nine male native Japanese speakers (ages: 23–27) participated in the experiment. All subjects showed right-handedness (laterality quotients: 65–100) by the Edinburgh inventory (Oldfield, 1971). None had a history of neurological or psychiatric disease. During the experiment, individual subjects with headphones and prism glasses were in a supine position in the magnet, and their heads were immobilized with padding inside the radio-frequency coil. Informed consent was obtained from each subject after the nature and possible consequences of the studies had been explained. Approval for these experiments was obtained from the institutional review boards of the University of Tokyo, Graduate School of Arts and Sciences.

Stimuli

Three sets of dialogues between two persons were prepared in Japanese, and the same sets were used either as auditory or visual stimuli. Each sentence in the dialogues was divided into three phrases, each of which consisted of one to four words. Under both the auditory and visual conditions, the subjects were asked to open their eyes and to fixate on a red central cross on a screen.

Under the auditory condition, all speech sounds were digitized (16 bit; the normal audio cut-off, 11025 Hz) using speech synthesis software (Oshaberi-mate, Fujitsu, Tokyo, Japan), which converts Japanese written texts into voice waveforms, faithfully replicating the prosody of Japanese speech. In order to distinguish two narrators of the dialogues, we alternated each sentence in the dialogues between a male and a female synthesized voice. The auditory stimuli were presented binaurally to the subjects through plastic tubes (inner diameter: 9 mm; length: 6.5 m) with silicone-cushioned headphones specifically designed to isolate the subjects from the scanner noise (Resonance Technology Inc., Northridge, CA). The maximum intensity of stimuli was an 80.6 dB sound pressure level within the headphones, and the duration of each stimulus ranged from 400 to 700 ms. A similar sound delivery system and auditory stimuli were used in our previous imaging study on auditory areas (Hashimoto *et al.,* 2000).

Under the visual condition, we presented each stimulus as a row of either green (for a male narrator in the dialogues) or orange (for a female narrator in the dialogues) letters against a dark background, using four to eight letters from the Japanese *hiragana* syllabary, which consists of a type of syllabic symbols known as *kana*. We did not use the logographic characters known as *kanji*, which contain both phonetic and morphemic information, but rather used *hiragana* containing only phonetic information, thereby restricting direct access only to the phonological form. Each visual stimulus (maximum visual angle: 1.6° × 12°) was back-projected onto a translucent screen near the subject's feet with a liquid crystal projector. The presentation time of each phrase under the visual condition was 700 ms, which corresponded to the maximum duration of auditory stimuli. The subjects read the stimuli covertly through prism glasses.

Tasks

Using a block design protocol, we tested three types of language tasks: sentence (S) tasks, phrase (P) tasks, and nonword (N) tasks. In the S tasks under the auditory (S_A) or visual (S_V) condition, phrases were presented in the order of the original sentences. One phrase in a sentence was randomly replaced with a probe stimulus at a low frequency (one to six). These probe stimuli belonged to the same grammatical category as the phrases they replaced, but were contextually unrelated to the dialogue: for example (in English), "May I take a picture here?"—"Yes, if you can break the flash," wherein the original phrase "do not use" was replaced with "can break." The subjects were asked to follow sentences of the dialogue, and to respond to a probe stimulus by pressing one of two buttons. The subjects pressed the other button, if there was no probe stimulus. We equated the number of syllables in original phrases (mean ± SD: 5.6 ± 1.1) with those in probe stimuli (5.4 ± 1.1).

In the P tasks under the auditory (P_A) or visual (P_V) condition, we presented the same phrases used in the S tasks (except probe stimuli), but in a completely randomized order. Because these phrases cannot be integrated into a sentence, automatic sentence processing was prohibited. As a probe stimulus, we randomly replaced a phrase with a pronounceable nonword at a rate of one to six. The nonwords were made from an original phrase by changing the sequence of syllables, and they conformed to the rules of Japanese phonotactics, but had no lexical meaning. The subjects judged whether a probe stimulus was present or not, and responded by pressing one of two buttons. In the N_{AV} task, phonologically equivalent nonwords were presented auditorily and visually at the same time, using the same set of the nonwords in the P tasks. As a probe stimulus, we randomly presented phonologically different nonwords at a rate of one to six.

In these tasks, we presented three consecutive stimuli every five seconds, after which the subjects judged the presence of a probe stimulus by pressing one of two buttons as quickly as possible. Both accuracy and reaction time (measured from the beginning of the first stimulus of three consecutive stimuli) were measured during the fMRI experiments, and the stimulus presentation and behavioral data collection were controlled by LabVIEW software and interface (National Instruments, Austin, TX). In a single session (420 s), following the first N_{AV} task block, two sequences of S_A-N_{AV}-P_A-N_{AV} and S_V-N_{AV}-P_V-N_{AV} (N_{AV}: 20 s, S and P: 30 s) were alternated twice. Initiation of each task block was cued by presenting the name of the task on the screen. The choice of S_A or S_V as the first S task block was counterbalanced within a subject.

We performed a follow-up experiment (with six out of the nine subjects) to examine which regions were active in the N_{AV} task, by employing three additional tasks: the N tasks under the auditory (N_A) or visual (N_V) condition, and a cross-modal control (C_{AV}) task for matching auditory and visual nonlinguistic stimuli. In the N_A and N_V tasks, we presented the same nonwords used in the N_{AV} task for each modality. A probe stimulus in the N_A task was a nonword, in which one syllable was replaced with white noise (the low-pass cut-off, 5512 Hz). In the N_V task, a probe stimulus was a nonword, in which one letter was replaced

with random dots in the same color. In each task, the subjects judged whether a probe stimulus was presented or not, and responded by pressing one of two buttons. In the C_{AV} task, which served as a nonlinguistic control of the N tasks, we simultaneously presented an auditory stimulus and a visual stimulus: auditory stimulus, white noise alone (W) or white noise with a 300-Hz pure tone (W′); visual stimulus, green random dots alone (R) or random dots with a green horizontal line (R′). The "matching" condition of this task was either absence or presence of tone/line, i.e., W/R or W′/R′. The other combinations were used as probe stimuli, presented at a rate of one to six. In a single session (320 s), following the first C_{AV} task block, one of two sequences of N_A-C_{AV}-N_V-C_{AV}-N_{AV}-C_{AV} and N_V-C_{AV}-N_A-C_{AV}-N_{AV}-C_{AV} (C_{AV}: 20 s, N: 30 s) was presented twice. The choice of N_A or N_V as the first N task block was counterbalanced within a subject.

fMRI Data Acquisition and Analyses

The present study was performed using a 1.5-Tesla MRI system (STRATIS II, Premium; Hitachi Medical Corporation, Tokyo, Japan). With a gradient echo echo-planar imaging (EPI) sequence (TR = 5 s, TE = 50 ms, flip angle = 90°, field of view = 192 × 192 mm², resolution = 3 × 3 mm²), we scanned over 18 horizontal slices of 4-mm thickness, covering the range of $z = -12$ to 60 (or 18 slices of 5-mm thickness, covering the wider range of $z = -25$ to 65 in the follow-up experiment). The scanning sounds were confined within the interstimulus interval by using a clustered volume acquisition sequence (acquisition time = 2250 ms). In a single scanning session, we obtained 85 volumes following the three dummy images, which allowed for the rise of the BOLD signal. We obtained corresponding high-resolution structural images in each experiment. For normalizing individual brains into a standard brain, a three-dimensional structural image of each subject's whole brain was obtained using a gradient echo sequence (TR = 30 ms, TE = 8 ms, flip angle = 60°, field of view = 192 × 192 mm², resolution = 1.5 × 1.5 × 1.5 mm³).

We performed group analyses with statistical parametric mapping software (Friston *et al.,* 1995) (SPM99; Wellcome Department of Cognitive Neurology, London, UK) in MATLAB (Math Works, Natick, MA). We realigned the functional volume data in multiple sessions and removed sessions which included data with a translation of more than 2 mm in one of the three directions. Each individual brain was spatially normalized to the standard brain space as defined by the Montreal Neurological Institute (MNI) with resampling every 3 mm using sinc interpolation. The present study was based on fMRI time-series data of 19 to 24 sessions per subject. Multiple sessions were collapsed into two sessions for each subject, which were classified according to the choice of S_A or S_V as the first S task block (N_A or N_V as the first N task block). These averaged data were then smoothed with an isotropic Gaussian kernel of 8 mm full width at half maximum. Low-frequency noise and global changes in activity were further removed. Task-specific effects were estimated with a general linear model (fixed effect model) with a delayed (5 s) boxcar waveform. The significant activation was determined by using the t statistics on a voxel-by-voxel basis.

In group analyses, statistical parametric maps in each comparison were thresholded at a corrected $P < 0.05$ ($t = 4.6$, $df = 1278$). We used the same statistical threshold ($t = 4.6$) for the exclusive masks to the statistical parametric maps (Table 5). In all the comparisons, we removed clusters that were smaller than nineteen voxels ($k = 19$ for an 8-mm cube). We used the atlases of Talairach and Tournoux (1988) and of Duvernoy (1991) for the anatomical identification of activated regions.

RESULTS

Task Performance

In all the tasks during fMRI experiments, the subject's performance was evaluated on-line (see Methods). Table 1 shows the accuracy and reaction time (RT) for each task, and the mean accuracy was greater than 90%. An F test showed that there was the main effect of tasks in accuracy [$F(4,40) = 8.6$, $P < 0.0001$]. Accuracy for the N_{AV} task was significantly lower than that for the P_A, P_V, S_A, and S_V tasks (Fisher's protected least significant difference, $P < 0.0005$). RT was not significantly affected by tasks [$F(4,40) = 0.7$, $P > 0.5$]. According to an analysis of variance (ANOVA) with two variables [processing (S, P) × modalities (audition, vision)], the main effects and the interaction for the four tasks were not significant in both accuracy and RT

TABLE 1

Behavioral Data in Language Tasks under the Auditory and/or Visual Condition

Task	Modality	Accuracy mean ± SD (%)	Reaction Time mean ± SD (ms)
S_A	Audition	96.3 ± 2.6	2495 ± 217
S_V	Vision	95.2 ± 2.5	2400 ± 269
P_A	Audition	95.9 ± 2.3	2522 ± 183
P_V	Vision	96.2 ± 2.0	2443 ± 230
N_{AV}	Audition and vision	90.5 ± 3.0	2392 ± 133

Note. Nine subjects participated in all the tasks. S_A, the sentence (S) task under the auditory condition; S_V, the S task under the visual condition; P_A, the phrase (P) task under the auditory condition; P_V, the P task under the visual condition; and N_{AV}, the nonword (N) task, in which auditory and visual stimuli are simultaneously presented.

TABLE 2

Modality-Selective Activation in the S, P, and N_{AV} Tasks

Region	BA		x	y	z	t	x	y	z	t
			Audition: $(S_A + P_A) - (S_V + P_V)$				Audition: $N_{AV} - P_V$			
Heschl's g, planum temporale,	41/42/22/21	L	−57	−18	6	31.7	−57	−18	6	39.4
superior and middle temporal g		R	60	−15	3	30.3	60	−15	3	33.7
Angular g	39	R	51	−72	36	6.3				
Middle occipital g	19	R	51	−78	15	7.2				
		L	−45	−81	21	6.2				
			Vision: $(S_V + P_V) - (S_A + P_A)$				Vision: $N_{AV} - P_A$			
Precentral s	6/8	L	−54	12	42	7.6	−57	9	39	10.4
Precentral g	4/6	L	−36	−9	45	6.5	−39	−9	45	7.4
		R	42	−12	54	6.5	42	−9	54	6.5
Intraparietal s	7	R	27	−66	48	6.8	30	−57	51	10.2
		L	−30	−60	60	6.7	−24	−66	54	9.4
Cuneus, middle and inferior	17/18/19	R	24	−99	6	22.9	24	−99	6	28.7
occipital g		L	−15	−102	9	22.3	−15	−102	12	25.9

Note. Stereotactic coordinates (x, y, z) in Montreal Neurological Institute (MNI) space are shown for each voxel with a local maximum of t values in the contrasts indicated ($P < 0.05$, corrected). The same statistical threshold was used throughout all analyses (Tables 2–5). BA, Brodmann's area; L, left hemisphere; R, right hemisphere; g, gyrus; and s, sulcus.

($P > 0.1$). While the N_{AV} task used as a baseline was the most demanding task, the other four tasks were equally balanced in terms of behavioral control.

Modality-Selective Activation in the S, P, and N_{AV} Tasks

In either the S or P tasks, the auditory and visual conditions were linguistically equivalent, and the difference between the two conditions was the modality of stimulus presentation alone. In order to examine the cortical regions related to modality-selective processing, we contrasted $(S_A + P_A)$ and $(S_V + P_V)$. Figure 1A shows the auditorily selective activations [$(S_A + P_A) - (S_V + P_V)$, red-colored regions], as well as the visually selective activations [$(S_V + P_V) - (S_A + P_A)$, green-colored regions]. The most prominent activation under the auditory condition was found in the early auditory areas, which extended from the bilateral Heschl's gyri to the superior and middle temporal gyri (Table 2). In contrast, the most prominent activation under the visual condition was found in the early visual areas, which extended from the bilateral cuneus to the middle and inferior occipital gyri. These early sensory areas as well as the adjacent association areas showed highly symmetrical activation under each modality condition. The additional activated regions were found in the frontal, parietal, and occipital lobes, but these regions showed much weaker activation than the sensory areas as indicated by their smaller t values and restricted localization. These results confirm that the S and P tasks involved modality-dependent perceptual processing of language inputs and subsequent language processing.

Next, we examined activated regions that reflected each of auditory and visual components included in the N_{AV} task. As we simultaneously presented auditory and visual stimuli in the N_{AV} task, the contrast $N_{AV} - P_V$ extracts auditory components in the N_{AV} task, while the contrast $N_{AV} - P_A$ extracts visual components in the N_{AV} task. In the contrast $N_{AV} - P_V$, activation was found in the bilateral early auditory areas (Table 2; red-colored regions in Fig. 1B); in the contrast $N_{AV} - P_A$, activation was found in the bilateral early visual areas (Table 2; green-colored regions in Fig. 1B). It is striking to note that these activation patterns are almost identical to those in the direct comparison between the auditory and visual conditions in the S and P tasks (Fig. 1A). Moreover, the auditory and visual components in the N_{AV} task did not show overlapping activation. Taken together, the activations in these contrasts reflected modality-selective processing in the N_{AV} task, which involved either speech-sound processing or letter-form processing of nonwords.

Activation in the P and S Tasks in Contrast to the N_{AV} Task

The P tasks required not only speech-sound and letter-form processing but lexico-semantic processing, which was not involved in the N_{AV} task. The contrasts $P_A - N_{AV}$ and $P_V - N_{AV}$ exhibited very similar activation patterns in each hemisphere (Table 3; Figs. 2A and 2B). In both contrasts, the angular gyrus (AG) and the supramarginal gyrus (SMG) showed bilateral activation. In addition, bilateral activation of the superior temporal sulcus (STS) and the middle temporal gyrus (MTG) was observed in $P_A - N_{AV}$, but not in $P_V - N_{AV}$.
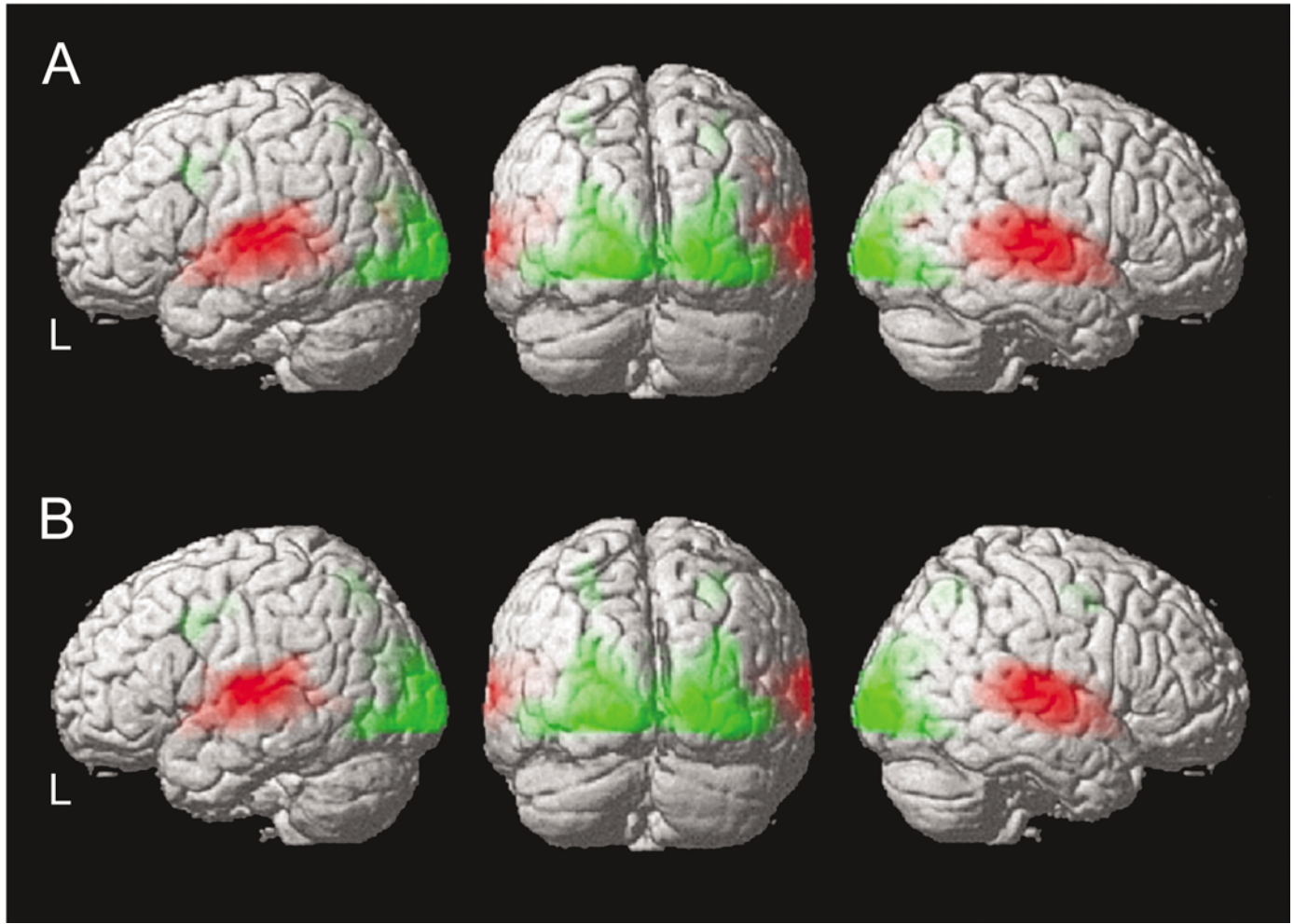
FIG. 1. Modality-selective activation in the S, P, and N$_{AV}$ tasks. Based on group analyses with SPM99 software, all activated regions are rendered on the surface of a standard brain. Stereotactic coordinates for local maxima of activation are listed in Table 2. (A) Activated regions in the direct comparison between the auditory and visual conditions in the sentence (S) tasks and the phrase (P) tasks. Activated regions in $(S_A + P_A) - (S_V + P_V)$ and in $(S_V + P_V) - (S_A + P_A)$ are shown in red and green, respectively. Left to right: the left hemisphere, the posterior view, and the right hemisphere. Note that activated regions were clearly segregated into early auditory areas and early visual areas, as well as adjacent association areas. (B) Activated regions in the nonword (N$_{AV}$) task in contrast to the P tasks. Activated regions in $N_{AV} - P_V$ and in $N_{AV} - P_A$ are shown in red and green, respectively. These activation patterns are almost the same as those in the direct comparison between two modality conditions (A).

On the other hand, the inferior frontal sulcus (BA 10/46 or IFS) and the anterodorsal part of the pars triangularis showed activation in both contrasts. This activation was more prominent in the left hemisphere than in the right hemisphere. In order to confirm that these activation patterns reflected lexico-semantic components in the P tasks, we performed conjunction analyses. While activation in STS and MTG slightly decreased in the conjunction analysis with $P_A - N_{AV}$ and the exclusive mask of $N_{AV} - P_V$, other regions showed the same activation patterns as in $P_A - N_{AV}$. In addition, a conjunction analysis with $P_V - N_{AV}$ and the exclusive mask of $N_{AV} - P_A$, resulted in almost the same activation patterns as in $P_V - N_{AV}$. Because the effect of exclusive masks was minor, the comparison

between the P and N$_{AV}$ tasks revealed cortical regions related to lexico-semantic processing, not the simple enhancement of perceptual processing. Consequently, we established the equivalence of cortical activation between the P_A and P_V tasks, by comparing each task with the identical N$_{AV}$ task.

We further examined the contrasts $S_A - N_{AV}$ and $S_V - N_{AV}$. These two comparisons showed similar activation patterns each other (Table 4; Figs. 2C and 2D). The overall activated regions were wider in $S_A - N_{AV}$ and $S_V - N_{AV}$ than in $P_A - N_{AV}$ and $P_V - N_{AV}$. In contrast to the absence of significant activation in the left STS and the bilateral MTG in $P_V - N_{AV}$ (Fig. 2B), $S_V - N_{AV}$ showed activation in these regions. Moreover, in $S_A - N_{AV}$ and $S_V - N_{AV}$, we found more prominent
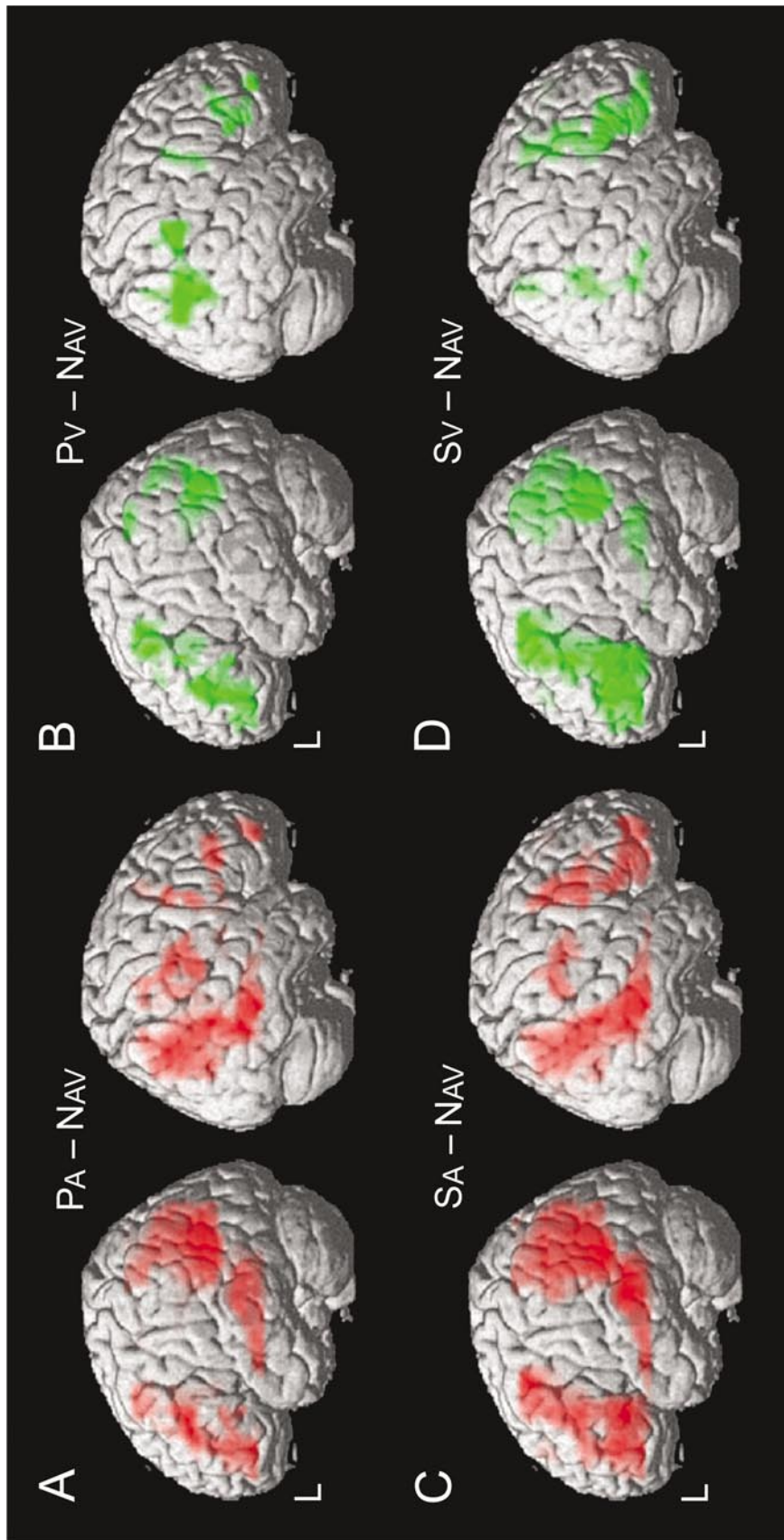
FIG. 2. Activated regions in the P and S tasks in contrast to the N_{AV} task. (A and B) Activated regions in $P_A - N_{AV}$ (A) and in $P_V - N_{AV}$ (B) are rendered on a standard brain. Talairach coordinates for local maxima of activation are listed in Table 3. These contrasts showed similar activation patterns in the cortical language areas. (C and D) Activated regions in $S_A - N_{AV}$ (C) and in $S_V - N_{AV}$ (D) are rendered on a standard brain. Talairach coordinates for local maxima of activation are listed in Table 4. Note that activation of the bilateral inferior frontal gyrus (IFG) in $S_A - N_{AV}$ and $S_V - N_{AV}$ extended to more ventral regions than in $P_A - N_{AV}$ and $P_V - N_{AV}$.

TABLE 3

Activated Regions in the P Tasks in Contrast to the N_{AV} Task

Region	BA		Audition: $P_A - N_{AV}$				Vision: $P_V - N_{AV}$			
			x	y	z	t	x	y	z	t
Inferior frontal g (F3t)	45	L	−57	27	6	5.6	−54	27	6	5.5
		R	57	21	9	5.3	63	27	3	6.1
Inferior frontal s	10/46	L	−45	39	24	7.3	−51	27	27	5.7
		L	−39	60	−9	6.9	−42	57	0	6.4
		R	45	51	−9	8.0	48	51	−6	6.3
Middle frontal g	8/9	L	−27	18	51	7.6	−30	12	48	5.7
		R	30	18	48	5.2				
Precentral s	6/8	L					−39	6	45	5.7
		R	48	15	36	6.9	51	6	18	5.9
Insula		L	−30	15	12	6.0				
		R	45	−3	9	5.8				
Superior temporal s	21/22	L	−60	−12	−12	7.3				
		R	57	−18	−9	6.0				
Middle temporal g	21/37	L	−54	−36	−3	9.3				
		R	60	−39	−12	10.8				
Angular g and	39/40	L	−48	−69	21	10.4	−39	−60	24	7.9
supramarginal g		L	−45	−54	54	6.9	−42	−51	57	6.1
		R	48	−69	33	8.5	42	−72	30	7.2
		R	63	−21	21	7.6	54	−27	30	6.7
Cingulate g	23/24		6	−18	48	6.1				
			−3	−36	36	5.6				
Precuneus	19/7/31	L	−9	−81	33	7.0				
			−3	−72	24	7.3				

activation of the bilateral IFG, which extended to more ventral regions, compared with activation in $P_A - N_{AV}$ and $P_V - N_{AV}$. As a result, the left-lateralized activation of the frontal lobe in $S - N_{AV}$ became less distinct than in $P - N_{AV}$. These results suggest that the S tasks involved components for sentence comprehension other than lexico-semantics. Therefore, the use of the P tasks was essential in the present study to reveal the lexical-

TABLE 4

Activated Regions in the S Tasks in Contrast to the N_{AV} Task

Region	BA		Audition: $S_A - N_{AV}$				Vision: $S_V - N_{AV}$			
			x	y	z	t	x	y	z	t
Inferior frontal g (F3t/F3O)	45/47	L	−54	27	0	11.6	−54	30	0	14.1
		R	60	33	−3	11.4	60	33	−3	10.4
Inferior frontal s	10/46	L	−51	30	27	9.0	−51	27	30	10.0
		L	−39	60	−6	8.1	−48	51	−9	12.3
		R	45	54	−6	9.7	54	39	−12	9.2
Middle frontal g	8/9	L	−36	24	51	7.9	−45	21	45	10.9
		R	30	15	48	6.1	33	18	48	5.8
Superior frontal g	8/9	L	−9	48	39	5.9	−9	48	39	5.9
Precentral s	6/8	L	−45	9	51	11.7	−42	6	51	12.8
		R	51	15	39	8.5	51	15	39	7.1
Superior temporal s	21/22	L	−54	−21	−9	11.0	−51	−18	−12	5.5
		R	60	−6	−12	5.7				
Middle temporal g	21/37	L	−57	−36	−3	13.9	−54	−39	−6	9.4
		R	51	−36	−6	11.7	60	−57	−9	6.9
Angular g and	39/40	L	−60	−51	33	10.2	−42	−63	48	10.4
supramarginal g		R	51	−63	45	8.4	45	−63	30	5.2
Precuneus	7/31		−3	−60	39	9.2	−6	−63	39	5.6
	19		−9	−84	33	9.8				

TABLE 5

Activated Regions in the S Tasks in Contrast to the P Tasks

Region	BA		Audition: $S_A - P_A$				Vision: $S_V - P_V$			
			x	y	z	t	x	y	z	t
Inferior frontal g (F3t/F3O)	45/47	L	−54	33	−3	6.5*	−54	27	−6	9.0*
Precentral s	6/8	L	−45	9	48	6.4	−45	6	54	6.1
Middle temporal g	21/37	L					−66	−45	0	8.6
		L	−57	−60	12	5.3	−60	−54	15	7.3
		R					69	−39	−3	7.1

Note. An asterisk (*) indicates a region which remained to show activation in the conjunction analyses (Fig. 4).

processing related activation, thereby filling the gap between the S and N_{AV} tasks.

Activation in the S Tasks in Contrast to the P Tasks

In order to clarify cortical regions, which showed more prominent activation in the S tasks than in the P tasks, we directly compared these tasks for each modality condition. The contrasts $S_A - P_A$ and $S_V - P_V$ revealed similar activation patterns (Fig. 3), which were highly lateralized in the left hemisphere. As to the left IFG, the activated region extended from the inferior part of the pars triangularis (BA 45 or F3t) to the pars orbitalis (BA 47, F3O, or FOC), across the anterior horizontal ramus of the Sylvian fissure. This region did not extend to the pars opercularis (BA 44 or F3op). Additional activated regions were found in the left precentral sulcus (PrCS) and in the left MTG (Table 5). The left F3t/F3O activation in $S_A - P_A$ and $S_V - P_V$ was not apparent in $P_A - N_{AV}$ and $P_V - N_{AV}$ (Figs. 2A and 2B). Therefore, this region would be regarded as a candidate for a sentence-processing selective region.

### Sentence-Processing Selective Activation

In order to extract regions with sentence-processing selective activation from the activated regions in the contrasts $S_A - P_A$ and $S_V - P_V$, we employed a conjunction analysis. In this analysis, following regions in six contrasts were used as exclusive masks: regions with activation in modality-dependent perceptual processing [$(S_A + P_A) - (S_V + P_V)$ or $(S_V + P_V) - (S_A + P_A)$], regions with activation in speech-sound or letter-form processing ($N_{AV} - P_A$ or $N_{AV} - P_V$), and regions with activation in lexico-semantic processing ($P_A - N_{AV}$ or $P_V - N_{AV}$). In the conjunction analysis with $S_A - P_A$ and these exclusive masks, we found a single activated region in F3t/F3O (a red-colored region with maximum intensity projections in Fig. 4). Moreover, in the conjunction analysis with $S_V - P_V$ and the exclusive masks, we again found activation in the same region of the left F3t/F3O alone (a green-colored region in Fig. 4). Almost all voxels (23 out of 24 voxels) with activa-

tion in the left F3t/F3O under the auditory condition coincided with those under the visual condition. Moreover, the local maximum of the left F3t/F3O in the conjunction analysis of $S_A - P_A$ was within the extent of activation in the conjunction analysis of $S_V - P_V$, and the converse was also true (Table 5). There might be a possibility that the activation of the left F3t/F3O in one of the six contrasts, used as the exclusive masks, was just below the statistical threshold (a corrected $P < 0.05$). We reexamined each of the six contrasts at a lower threshold (an uncorrected $P < 0.001$), and confirmed the absence of activation at the local maxima found in the left F3t/F3O. These results demonstrated that the left F3t/F3O showed sentence-processing selective and modality-independent activation.

Next we performed individual analyses to confirm the consistency of activation in this region of the left F3t/F3O among the subjects. We used two coordinates of the local maxima (A: −54, 33, −3; V: −54, 27, −6; as shown in Table 5) for all subjects, after each individual brain was spatially normalized to the standard brain space. For either the S or P task blocks, we calculated the mean percent signal changes from the N_{AV} task blocks. Because the two coordinates of the local maxima were almost identical, we averaged the signal changes between these local maxima for each subject. As shown in Fig. 5, all subjects clearly showed larger signal changes in the S tasks than the P tasks under both modality conditions. Moreover, signal changes in the S_A and S_V tasks were comparable among all the subjects. We performed an ANOVA [processing (S, P) × modalities (audition, vision)] and found a significant main effect of processing [$F(1,32) = 28.8$, $P < 0.0001$]. Neither the main effect of modalities nor the interaction between processing and modalities was significant ($P > 0.5$). These results indicated that activation in the left F3t/F3O was consistent among all subjects.

Activation in the P and N_{AV} Tasks in Contrast to the S Tasks

In $P_A - S_A$, activation was restricted around the bilateral Heschl's gyri (Fig. 6A), while in $P_V - S_V$,
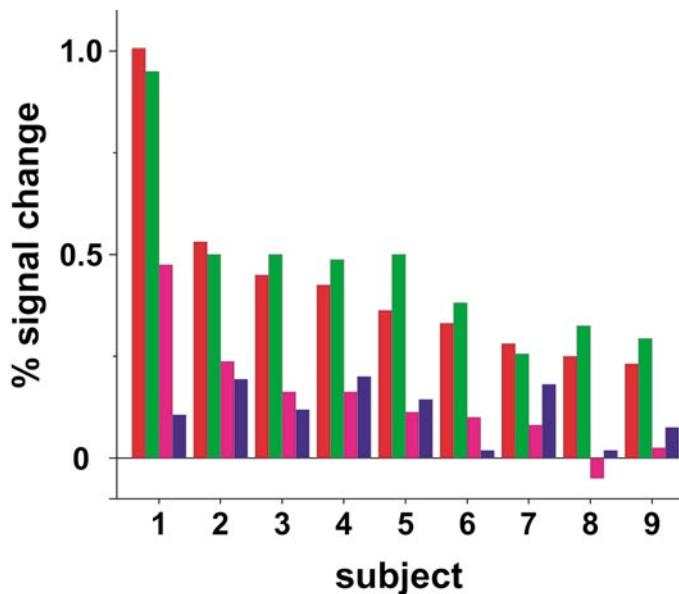
3

4
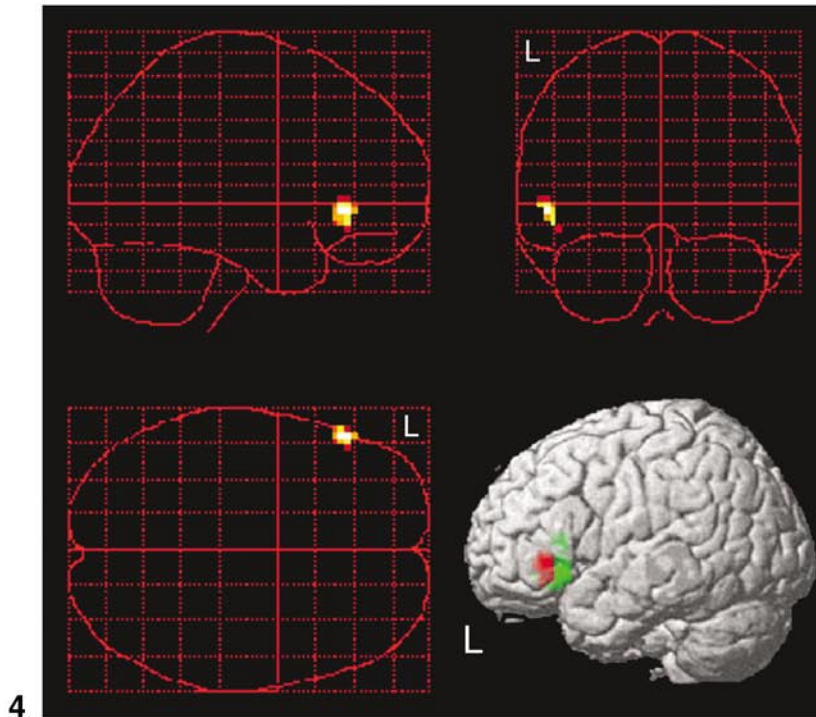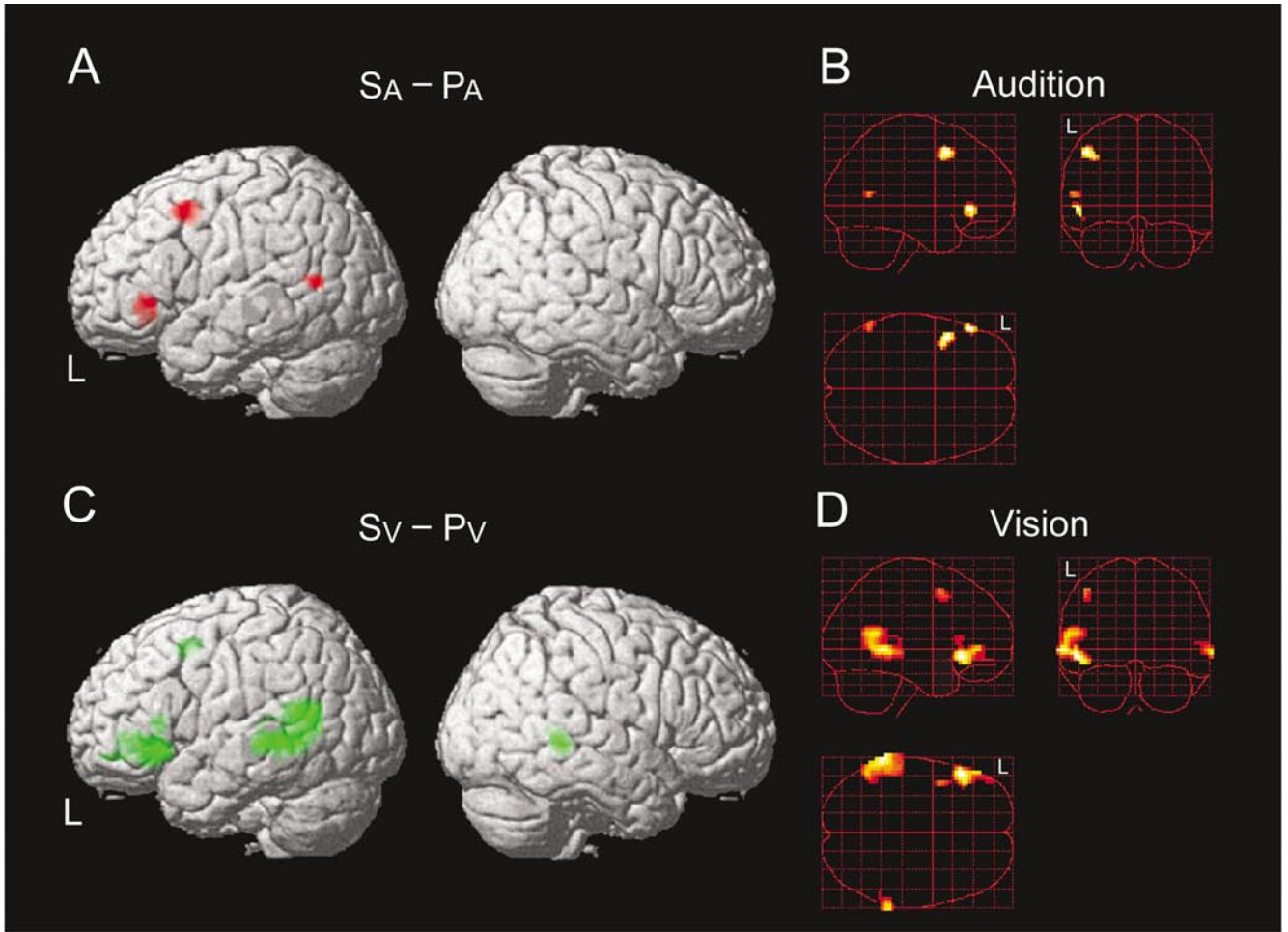
FIG. 5. Individual data of signal changes in the left F3t/F3O. As to the local maxima in the left F3t/F3O, mean percentage signal changes are shown for the S_A, S_V, P_A, and P_V tasks (red, green, magenta, and blue bars, respectively). Note that all subjects showed larger signal changes in the S tasks than the P tasks under both modality conditions.

activation was found in the right cuneus and the middle and inferior occipital gyri (Fig. 6B). No other regions showed significant activation in both contrasts. These results indicated that enhanced activation in the P tasks relative to the S tasks was confined in the early sensory areas, probably due to more attention to the stimuli in the P tasks.

We examined other contrasts $N_{AV} - S_V$ (Fig. 6C) and $N_{AV} - S_A$ (Fig. 6D), and obtained similar results as the contrasts $N_{AV} - P_V$ and $N_{AV} - P_A$, respectively (Fig. 1B). In $N_{AV} - S_V$, we found activation of the early auditory areas, which extended from the bilateral Heschl's gyri to the superior and middle temporal gyri. In $N_{AV} - S_A$, we found activation of the early visual areas, which extended from the bilateral cuneus to the middle and inferior occipital gyri. The coordinates of local maxima were close to those in the $N_{AV} - P_V$ and $N_{AV} - P_A$. As activation in the early sensory areas was enhanced in the P tasks relative to the S tasks (Figs. 6A

and 6B), it was weaker in the S tasks than in the P and N_{AV} tasks. This explains weak activation of the bilateral cuneus and the bilateral middle and inferior occipital gyri in the contrast $N_{AV} - S_V$, as well as that of the left Heschl's gyrus in $N_{AV} - S_A$. In addition, the anterior cingulate gyrus also showed activation in these contrasts ($N_{AV} - S_V$: -3, 57, 0; $N_{AV} - S_A$: 3, 57, -3). As to the left inferior frontal cortex, the posterodorsal part of the left F3op showed activation in $N_{AV} - S_V$ (-57, 3, 18), as well as in $N_{AV} - S_A$ (-63, 6, 15). In this region, the responses to the P tasks were intermediate between the N_{AV} and S tasks, although this region did not show significant activation in the contrasts $P_A - S_A$ and $P_V - S_V$. These results suggest that the left F3op is functionally distinct from the left F3t/F3O.

Activation in the N Tasks in Contrast to the C_{AV} Task

In order to ensure that the results in the left F3t/F3O were not affected by nonword processing in the N_{AV} task, we examined cortical activation in the N_{AV}, N_A, and N_V tasks in contrast to the C_{AV} task (see Methods). The mean accuracy for the N_{AV}, N_A, N_V, and C_{AV} tasks was 89.7, 97.6, 96.7, and 95.9%, respectively. An F test showed that there was the main effect of tasks in accuracy [$F(3,20) = 4.8$, $P < 0.05$]. This result indicates that the N_{AV} task remained the most demanding task in this follow-up experiment. In the contrast $N_{AV} - C_{AV}$, we found activation of the bilateral early auditory areas (Fig. 7A), which was similar to that in the direct comparison between the auditory and visual conditions in the S and P tasks (Fig. 1A). As to the early visual areas, however, this contrast showed more restricted activation in the ventral regions when compared with the direct comparison; the cuneus and the intraparietal sulcus did not show significant activation. In $N_A - C_{AV}$ and $N_V - C_{AV}$ (Fig. 7B), we found activations in the early auditory areas and the early visual areas, respectively, which were identical with those in $N_{AV} - C_{AV}$ (Fig. 7A). In accordance with the precuneus activation in $P_A - N_{AV}$ and $S_A - N_{AV}$ (Tables 2 and 3), the precuneus was activated only in the auditory condition ($N_A - C_{AV}$: -6, -63, 54).

As to the left frontal cortex, activations were found in PrCS ($N_{AV} - C_{AV}$: -54, 9, 39; $N_V - C_{AV}$: -48, 9, 48) and

FIG. 3. Activated regions in the S tasks in contrast to the P tasks. Activated regions in the comparison between the S and P tasks are rendered on a standard brain. Talairach coordinates for local maxima of activation are listed in Table 5. (A and B) Activated regions in $S_A - P_A$. (C and D) Activated regions in $S_V - P_V$. (B and D) Sagittal, coronal, and axial views of maximum-intensity projections of statistical parametric maps. In both contrasts, activated regions were found in the left pars triangularis/pars orbitalis (F3t/F3O), the left precentral sulcus (PrCS), and the middle temporal gyrus (MTG). Note that the activated regions were highly localized and lateralized in the left hemisphere.

FIG. 4. Sentence-processing selective activation in the left F3t/F3O. An activated region in the contrast $S_A - P_A$ with exclusive masks is shown in the maximum-intensity projections and rendered on a standard brain in red. An activated region in the contrast $S_V - P_V$ with the same exclusive masks is rendered on a standard brain in green. Almost all voxels of the activated region under the auditory condition, which were localized within the left F3t/F3O, coincided with those under the visual condition.

in IFS ($N_V - C_{AV}$: −54, 36, 12). The PrCS activation extended to the posterodorsal part of the left F3op, though there was no local maximum in F3op. In contrast, it is notable that the contrast $N_{AV} - C_{AV}$, as well as the contrasts $N_A - C_{AV}$ and $N_V - C_{AV}$, did not show significant activation in the left F3t/F3O (Fig. 7). This result indicates that the absence of activation of this region in $P - N_{AV}$ (Figs. 2A and 2B) was not due to the active baseline condition of the N_{AV} task. For the exclusive masks used in the conjunction analysis, we added the regions with activation in nonword processing ($N_{AV} - C_{AV}$, $N_A - C_{AV}$, or $N_V - C_{AV}$), and confirmed that the same portion of the left F3t/F3O remained active (Fig. 4). These results further established that the left F3t/F3O is not simply more active during sentence processing than lexical processing, but that it is selectively active for sentence processing.

DISCUSSION

In this study, we clarified the critical role of the left F3t/F3O in sentence comprehension. Sentence comprehension involves the selection of specific concepts from many possible semantic attributes, which are then integrated together into a sentence. The S tasks explicitly required such selection and integration of semantic information, while the P tasks required lexical decisions alone. By directly comparing the S and P tasks in both modality conditions, together with the conjunction analysis for eliminating lexico-semantic factors, it was unequivocally demonstrated that the left F3t/F3O plays a pivotal role in those amodal sentence-related processes beyond lexical processing.

To elucidate how sentence comprehension is achieved by the convergence of auditory and visual information, we compared the effect of input modalities on the activity of cortical language areas, by newly introducing the N_{AV} task as an identical baseline task for both the auditory and visual conditions. The overall activation patterns were strikingly similar between the P_A and P_V tasks, as well as between the S_A and S_V tasks, when each task was compared with the identical N_{AV} task. In $S_A - N_{AV}$ and $S_V - N_{AV}$, we found more prominent activation in the bilateral IFG, which extended to more ventral regions of F3t/F3O, when compared with activation in $P_A - N_{AV}$ or $P_V - N_{AV}$. This finding clearly indicates that auditory and visual inputs of words converge on the F3t/F3O during sentence comprehension.

Lexical and Sentence Processing in the Left Inferior Frontal Cortex

The present study further suggests that F3op and other regions in the left frontal cortex are related to lexical processing. The left F3op showed activation in the contrasts $N_{AV} - S_V$ and $N_{AV} - S_A$ (Figs. 6C and 6D),

while its response to the P tasks, in which nonwords were infrequently presented as probe stimuli, was intermediate between responses to the N_{AV} and S tasks. It is possible that the F3op activation reflects the active processes of verbalization, further enhanced by unfamiliar nonwords. This possibility is consistent with previous PET studies of reading nonwords/words, which reported activation of F3op and adjacent regions (Herbster *et al.,* 1997; Mummery *et al.,* 1999; Fiez *et al.,* 1999). The functional separation between F3t/F3O and F3op extends the previous suggestion of anterior (BA 45/47) and posterior (BA 44/45) regions as reflecting semantic and phonological/lexical processing, respectively (Poldrack *et al.,* 1999). However, we found that the left IFS (BA 10/46) and the anterodorsal part of the left F3t showed activation in the contrasts $S - N_{AV}$ and $P - N_{AV}$ (Fig. 2). This activation may be related to automatic lexico-semantic processing, but not to the representation or effortful retrieval of semantic knowledge per se, in that there are hardly any lesion studies of semantic deficits associated with this region (Price, 1998). Indeed, activation of BA 44/45 has been often accompanied by BA 10/46 activation (Buckner *et al.,* 1995; Chee *et al.,* 1999). It may explain BA 10/46 activation in some nonword tasks (Tagamets *et al.,* 2000), as well as the activation in the N_V task (Fig. 7B), where visually presented nonwords may be automatically rearranged into possible real words. To our knowledge, the present study demonstrates for the first time that the left F3t/F3O is functionally separable from other regions involved in lexical processing, and that it subserves to link the meanings of different words together during sentence comprehension.

By introducing the non-linguistic C_{AV} task, we confirmed that the left F3t/F3O did not show significant activation in the N_{AV} task (Fig. 7A). Irrespective of the use of N_{AV} as contrasts, this region was clearly activated in $S_A - P_A$ and $S_V - P_V$ (Fig. 3), as much as in $S_A - N_{AV}$ and $S_V - N_{AV}$ (Figs. 2C and 2D). The activated regions in the N_{AV} task were essentially the addition of those in its component tasks, N_A and N_V. Any regions responsible for nonspecific divided attention in the N_{AV} task might be eliminated by the contrast with C_{AV} task. However, a recent fMRI study reported that IFG was not significantly activated by the load of divided attention, in which a word or nonword stimulus was presented in both modalities and subjects made lexical decisions (Shaywitz *et al.,* 2001).

It would be possible that a certain lexical task activates the left F3t/F3O if the task involves the selection and integration processes of words. Indeed, previous fMRI studies have proposed that the left IFG is related to the selection of semantic knowledge (Thompson-Schill *et al.,* 1997; Gabrieli *et al.,* 1998). Moreover, verb generation tasks, which resulted in activation of BA 47 and adjacent regions (Petersen *et al.,* 1989; Buckner *et al.,* 1995), may include integration processes, because

generating word associates is a part of sentence production. Even in lexical decision and semantic decision (e.g., abstract vs concrete words) tasks, subjects may associate and integrate some presented words together. Therefore, F3t/F3O activation in previous reports may reflect sentence-related processes other than intended lexical processing.

By directly comparing sentence processing with lexical processing, the present study identified the left F3t/F3O that is selectively involved in the selection and integration processes of semantic information. Bottini *et al.* (1994) contrasted a sentence comprehension task with a lexical decision task, which is similar to the contrast used in the present study. The widespread regions they observed would have been more localized if conjunction analyses had been employed, but they nevertheless included the left BA 45/47. Moreover, deficits in sentence comprehension and production in Broca's aphasia (Damasio and Geschwind, 1984) may be due to the impaired ability to integrate words into sentences. It was also suggested that the comprehension deficit of Broca's aphasics is at the level of postlexical integration processes (Hagoort, 1997). Our proposal is consistent with these studies, in that the selection and integration processes of sentence comprehension are separable from lexico-semantic processing.

Other Cognitive Factors Associated with Activations in the Left Inferior Frontal Cortex

It is also likely that the left F3t/F3O processes both syntactic and semantic information, a necessary part of sentence comprehension. The left IFG (BA 44/45) has been implicated in syntactic processing of sentences (Stromswold *et al.,* 1996; Just *et al.,* 1996; Dapretto and Bookheimer, 1999; Ni *et al.,* 2000; Embick *et al.,* 2000; Moro *et al.,* 2001), phrases (Kang *et al.,* 1999), and single words (Friederici *et al.,* 2000), though lexical and syntactic factors might interact (Keller *et al.,* 2001). Moreover, a region along the left IFS has been recently found to be selectively involved in syntactic processing of sentences (Indefrey *et al.,* 2001; Hashimoto and Sakai, 2001). Another PET study has reported a relatively anterior local maximum of BA 45 (-46, 36, 4) in sentence comprehension tasks (plausible/implausible sentences), which were performed with concurrent articulation to inhibit subvocal rehearsal (Caplan *et al.,* 2000). These results and our present results further suggest the neural organization of the left IFG, such that subregions are functionally combined to assign syntactic structures of sentences, as well as to select and integrate semantic information.

We can further exclude the possibility that activation of the left F3t/F3O is primarily related to the verbal short-term memory or general task difficulty (Just *et al.,* 1996). First, we recently observed that the left F3t showed prominent activation in syntactic decision tasks with sentence stimuli, even if they were directly compared with a verbal short-term memory task (Hashimoto and Sakai, 2001). Second, the behavioral results in the present study indicated that there were no significant differences in task difficulty between the S and P tasks. Third, fMRI studies have reported that task difficulty does not significantly affect activation of the left F3t/F3O as well as BA 46 in a semantic encoding task (Demb *et al.,* 1995), and in a verbal short-term memory task (Chein and Fiez, 2001).

Tasks and Analyses for Dissociating Lexical Processing and Sentence Processing

In the present study, automatic lexico-semantic processing in the S tasks was equated with that in the P tasks, by employing the same phrase sets except for the infrequently appearing probe stimuli. Moreover, a number of phrase stimuli in the P tasks were presented in a completely randomized order, which cannot be restored to the original sentences. In contrast, the S tasks cannot be performed correctly without linking the meanings of phrases in sentences. Thus, this paradigm successfully separates the processes necessary for sentence comprehension from those related to lexical processing.

In previous neuroimaging studies on language processing, subjects were asked to hear or view the presented stimuli passively, or to generate sentences silently during the scanning (Mazoyer *et al.,* 1993; Kim *et al.,* 1997; Kang *et al.,* 1999). However, passive hearing/viewing and internal speech do not thoroughly control the quality and depth of language processing. The N_{AV}, P, and S tasks used in the present study were not passive but active tasks, where their execution was constantly monitored with probe tests. In addition to the left frontal cortex, we found activation in the temporo-parietal regions in $P - N_{AV}$ and $S - N_{AV}$ under both modality conditions. As lesion studies and neuroimaging studies have suggested (Hart, Jr. and Gordon, 1990; Démonet *et al.,* 1992; Price, 2000), the function of the left posterior temporal and inferior parietal region would be semantic processing at the lexical level.

Modality-Dependent Perceptual Processing during Language Tasks

The present results are consistent with previous studies as to activation in sensory areas and the adjacent association areas during language tasks. We found visually selective activation in the early visual areas and the adjacent association areas in both hemispheres. Activation in the bilateral early visual areas has been reported in a study using visual feature detection tasks that employed words and false fonts (Price *et al.,* 1996). In addition, the bilateral intraparietal sulcus and the transverse occipital sulcus
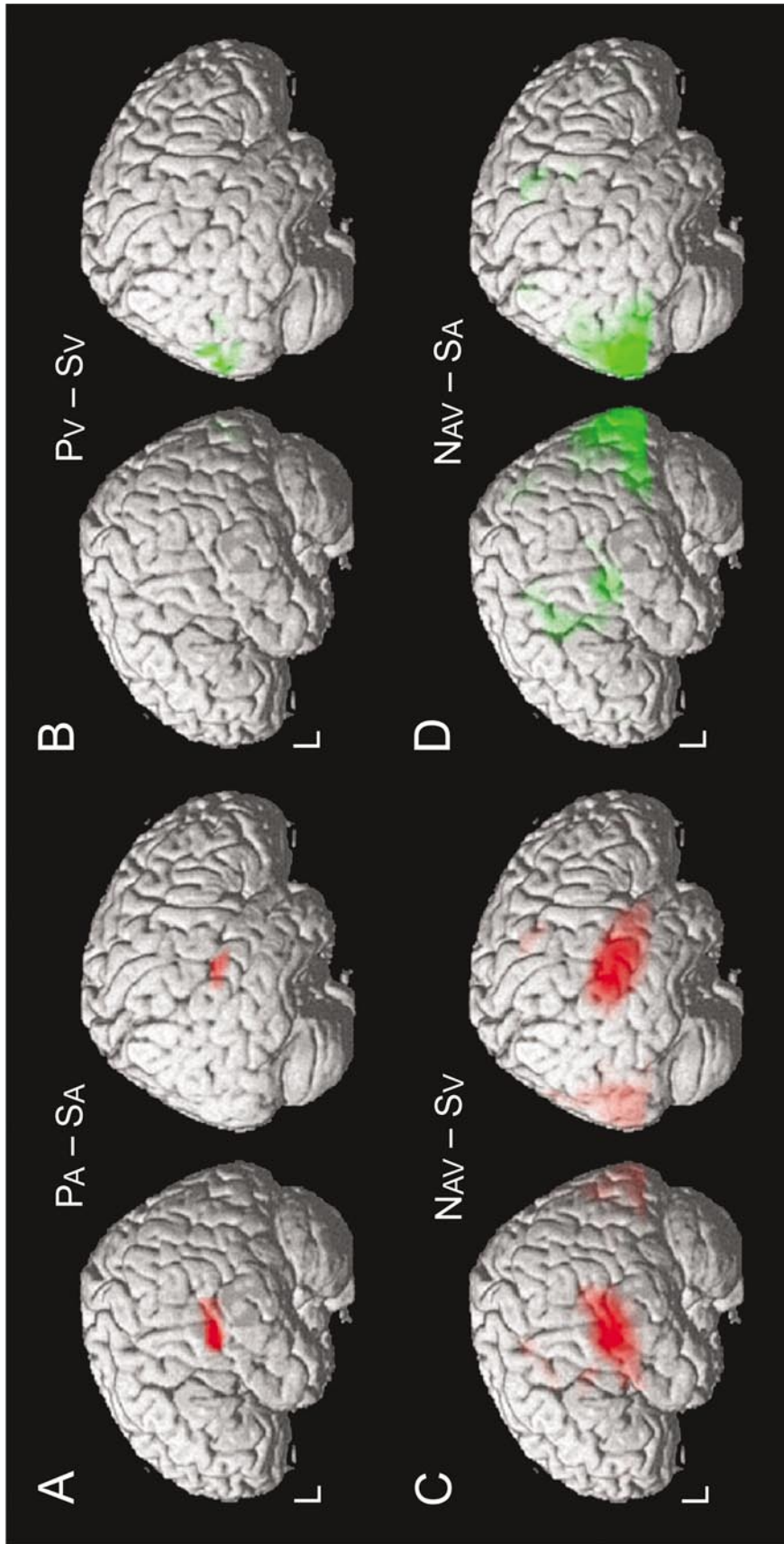
FIG. 6. Activated regions in the contrasts P − S and N_AV − S. (A and B) Activated regions in P_A − S_A (A) and in P_V − S_V (B) are rendered on a standard brain. Activation was found in the early sensory areas under each modality condition. Note that the cortical language areas did not show significant activation in both contrasts. (C and D) Activated regions in N_AV − S_V (C) and in N_AV − S_A (D) are rendered on a standard brain. Note that the early auditory and visual areas showed activation in each contrast (see Fig. 1).
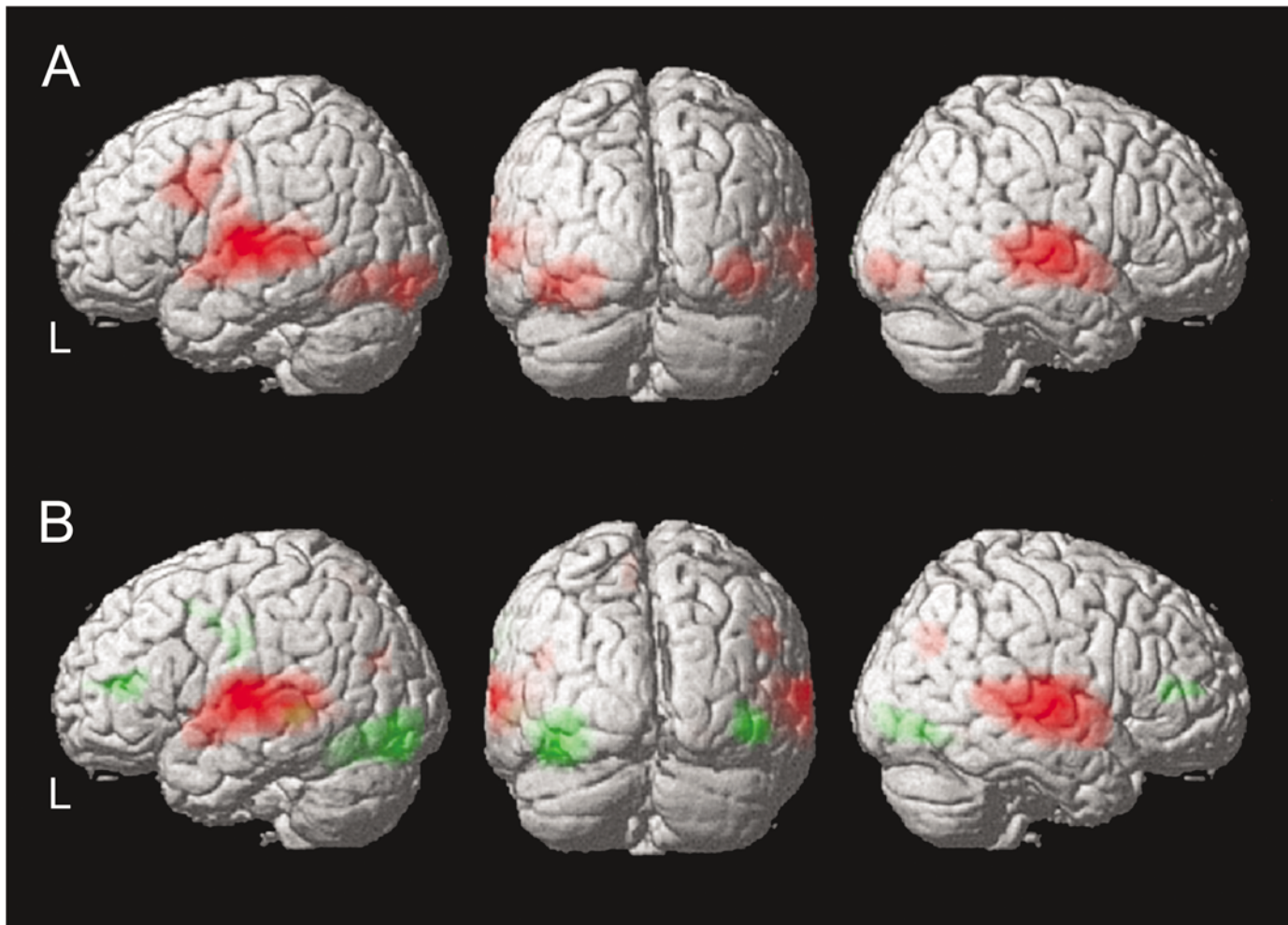
FIG. 7. Activated regions in the N tasks in contrast to a non-linguistic (C_{AV}) task. (A) Activated regions in $N_{AV} - C_{AV}$ are rendered on a standard brain. While the early auditory areas showed activation similar to that in the direct comparison between two modality conditions (see Fig. 1A), the early visual areas showed activation only in their ventral regions. (B) Activated regions in $N_A - C_{AV}$ and in $N_V - C_{AV}$ are shown in red and green, respectively. The activations in the early sensory areas were identical with those shown in (A). Note that the left F3t/F3O did not show significant activation in these contrasts.

showed significant activation in visual attention tasks with letters, as well as with faces and objects (Wojciulik and Kanwisher, 1999). We also found auditorily selective activation in the early auditory areas and the adjacent association areas in both hemispheres. Our recent fMRI study showed significantly higher activation in the bilateral early auditory areas under speech comprehension tasks than under a non-speech sound condition (Hashimoto *et al.,* 2000). This previous result indicates that these areas are related to the processing of the complex temporal and spectral features of speech sounds at the early stages of auditory perception.

Because the robust activations of the cuneus and the intraparietal sulcus in Fig. 1 disappeared in the contrasts $N_{AV} - C_{AV}$ and $N_V - C_{AV}$ (Fig. 7), the visual features of letters processed in these early visual areas may not be much different from those of random dots. In contrast, the auditory features of nonword speech

sounds processed in the early auditory areas differ significantly from those of white noise, as shown by the robust activations of these areas in $N_{AV} - C_{AV}$ and $N_A - C_{AV}$. This finding suggests more specialized role of the early auditory areas in speech recognition, rather than mere spectral decomposition of sounds. In a recent fMRI study, voice-selective regions were detected along the upper bank of the bilateral superior temporal sulcus (Belin *et al.,* 2000).

According to our behavioral data, the N_{AV} task was the most demanding task, which involved divided attention and spelling-to-sound correspondence. Thus it required more attention to the stimuli, which may have elicited greater activation of the sensory areas. Indeed, the early visual areas showed activation in $N_{AV} - S_V$ (Fig. 6C), though they showed activation both in the N_{AV} and S_V tasks (Fig. 1). Similarly, the early auditory areas showed activation in $N_{AV} - S_A$ (Fig. 6D), though

they showed activation both in the N_{AV} and S_A tasks. This result is consistent with previous fMRI studies of lexical decisions that required modality-selective attention (Hashimoto *et al.,* 2000; Shaywitz *et al.,* 2001).

On the other hand, the precuneus was activated in $P_A - N_{AV}$ and $S_A - N_{AV}$ (Tables 2 and 3), and weaker in $N_A - C_{AV}$, which is surprising in that the auditory tasks of P_A, S_A, and N_A had larger influences over this visual region than N_{AV} or C_{AV}, in spite of the fact that the N_{AV} task elicited greater activation of both early auditory and early visual areas than the S tasks (Figs. 6C and 6D). Indeed, enhanced activation of the precuneus in $S_A - N_{AV}$ was greater than that in $S_V - N_{AV}$ (Table 4). A previous PET study reported enhanced activation of the cuneus and precuneus under auditory attention conditions when compared with visual attention conditions, in which speech sounds and letters were simultaneously presented (O'Leary *et al.,* 1997). They suggested that this activation may be associated with inhibition of visual processing, and a recent fMRI study with similar selective attention tasks, as well as simple auditory or visual controls, has supported this possibility (Shaywitz *et al.,* 2001). However, it should be noted that only auditory stimuli were presented in our P_A, S_A, and N_A tasks, without any active visual processing to be inhibited. An alternative explanation for the present results is that speech stimuli in the P_A and S_A tasks automatically generated visual imagery of particular patterns of letter strings or contents of presented phrases, while nonwords in N_{AV} and N_A resulted in less concrete representations (Sakai and Miyashita, 1994).

Other Nonlinguistic Factors Possibly Involved in the Tasks

We observed that the cingulate gyrus was also activated in $P_A - N_{AV}$ (Table 3). This midcingulate region is known to be related to motor control (Petit *et al.,* 1998) and target detection (Braver *et al.,* 2001). According to our behavioral data, RT of the P_A task was longest, and that of the N_{AV} task was shortest among all the tasks (Table 1), and this difference was statistically significant ($t(8) = 5.1$, $P < 0.001$, paired t test, two-tailed). Thus, the cingulate gyrus activation may be due to response selection accompanied by probe detection in our tasks.

It has been pointed out that the masking effect of acoustic noise generated by an MR scanner may influence cortical responses, either by masking the auditory stimuli by the temporal overlaps with scanner noise (Edmister *et al.,* 1999) or by saturating cortical responses to auditory stimuli in the presence of scanner noise (Talavage *et al.,* 1999; Hall *et al.,* 1999). With respect to the first point, we minimized acoustic contamination due to scanner noise in the auditory stimuli (see Methods). As to the second point, signal enhance-

ment due to scanner noise has been observed mainly in the primary auditory cortex (Bandettini *et al.,* 1998; Talavage *et al.,* 1999), whereas MTG/STS, AG, SMG, and IFG were affected in less than half of subjects tested (Ulmer *et al.,* 1998). In the present study, all subjects showed a similar degree of signal changes in the left F3t/F3O between the S_A and S_V tasks (Fig. 5), indicating the absence of idiosyncratic effects of scanner noise under the auditory condition. It should also be noted that the physical characteristics of scanner noise were constant throughout all the tasks under the auditory and/or visual conditions.

CONCLUSIONS

Here we employed a conjunction analysis with exclusive masks and the contrast either $S_A - P_A$ or $S_V - P_V$. Although this conjunction analysis might be considered conservative in terms of revealing sentence-processing selective activation, the determination of activation became solid by removal of a possible sentence-lexical interaction. Furthermore, we performed all possible comparisons among the N_{AV}, P, and S tasks, and confirmed that the left F3t/F3O showed consistent activation in sentence processing. Our innovative paradigm with strict linguistic controls, together with the conjunction analysis, successfully dissociated sentence processing from lexical processing, and further revealed a cortical region selectively involved in amodal sentence processing.

REFERENCES

Bandettini, P. A., Jesmanowicz, A., Van Kylen, J., Birn, R. M., and Hyde, J. S. 1998. Functional MRI of brain activation induced by scanner acoustic noise. *Magn. Reson. Med.* **39:** 410–416.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. 2000. Voice-selective areas in human auditory cortex. *Nature* **403:** 309–312.

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Gaillard, W., and Theodore, W. 1995. Regional cerebral blood flow during object naming and word reading. *Hum. Brain Mapp.* **3:** 93–106.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S. J., and Frith, C. D. 1994. The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain* **117:** 1241–1253.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., and Snyder, A. 2001. Anterior cingulate cortex and response conflict: Effects of frequency, inhibition and errors. *Cereb. Cortex* **11:** 825–836.

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.

Caplan, D., Alpert, N., Waters, G., and Olivieri, A. 2000. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Hum. Brain Mapp.* **9:** 65–71.

Caplan, D., and Waters, G. S. 1999. Verbal working memory and sentence comprehension. *Behav. Brain Sci.* **22:** 77–126.

Chee, M. W. L., O'Craven, K. M., Bergida, R., Rosen, B. R., and Savoy, R. L. 1999. Auditory and visual word processing studied with fMRI. *Hum. Brain Mapp.* **7:** 15–28.

Chein, J. M., and Fiez, J. A. 2001. Dissociation of verbal working memory system components using a delayed serial recall task. *Cereb. Cortex* **11:** 1003–1014.

Damasio, A. R., and Geschwind, N. 1984. The neural basis of language. *Annu. Rev. Neurosci.* **7:** 127–147.

Dapretto, M., and Bookheimer, S. Y. 1999. Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron* **24:** 427–432.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., and Gabrieli, J. D. E. 1995. Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *J. Neurosci.* **15:** 5870–5878.

Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., Rascol, A., and Frackowiak, R. 1992. The anatomy of phonological and semantic processing in normal subjects. *Brain* **115:** 1753–1768.

Duvernoy, H. 1991. *The Human Brain: Surface, Three-Dimensional Sectional Anatomy and MRI.* Springer-Verlag, Wien.

Edmister, W. B., Talavage, T. M., Ledden, P. J., and Weisskoff, R. M. 1999. Improved auditory cortex imaging using clustered volume acquisitions. *Hum. Brain Mapp.* **7:** 89–97.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., and Sakai, K. L. 2000. A syntactic specialization for Broca's area. *Proc. Natl. Acad. Sci. USA* **97:** 6150–6154.

Fiez, J. A., Balota, D. A., Raichle, M. E., and Petersen, S. E. 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* **24:** 205–218.

Friederici, A. D., Opitz, B., and von Cramon, D. Y. 2000. Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cereb. Cortex* **10:** 698–705.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. 1995. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2:** 189–210.

Gabrieli, J. D. E., Poldrack, R. A., and Desmond, J. E. 1998. The role of left prefrontal cortex in language and memory. *Proc. Natl. Acad. Sci. USA* **95:** 906–913.

Geschwind, N. 1965. Disconnexion syndromes in animals and man—Part I. *Brain* **88:** 237–294.

Gibson, E., and Pearlmutter, N. J. 1998. Constraints on sentence comprehension. *Trends Cogn. Sci.* **2:** 262–268.

Gusnard, D. A., and Raichle, M. E. 2001. Searching for a baseline: Functional imaging and the resting human brain. *Nat. Rev. Neurosci.* **2:** 685–694.

Hagoort, P. 1997. Semantic priming in broca's aphasics at a short SOA: No support for an automatic access deficit. *Brain Lang.* **56:** 287–300.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., Gurney, E. M., and Bowtell, R. W. 1999. "Sparse" temporal sampling in auditory fMRI. *Hum. Brain Mapp.* **7:** 213–223.

Hart, J., Jr., and Gordon, B. 1990. Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Ann. Neurol.* **27:** 226–231.

Hashimoto, R., Homae, F., Nakajima, K., Miyashita, Y., and Sakai, K. L. 2000. Functional differentiation in the human auditory and language areas revealed by a dichotic listening task. *Neuroimage* **12:** 147–158.

Hashimoto, R., and Sakai, K. L. 2001. Separate neural bases for syntactic processing and verbal short-term memory in the frontal cortex: An fMRI study. *Soc. Neurosci. Abstr.* **27:** 82.7.

Herbster, A. N., Mintun, M. A., Nebes, R. D., and Becker, J. T. 1997. Regional cerebral blood flow during word and nonword reading. *Hum. Brain Mapp.* **5:** 84–92.

Howard, D., Patterson, K., Wise, R., Brown, W. D., Friston, K., Weiller, C., and Frackowiak, R. 1992. The cortical localization of the lexicons. *Brain* **115:** 1769–1782.

Indefrey, P., Hagoort, P., Herzog, H., Seitz, R. J., and Brown, C. M. 2001. Syntactic processing in left prefrontal cortex is independent of lexical meaning. *Neuroimage* **14:** 546–555.

Jennings, J. M., McIntosh, A. R., Kapur, S., Tulving, E., and Houle, S. 1997. Cognitive subtractions may not add up: The interaction between semantic processing and response mode. *Neuroimage* **5:** 229–239.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., and Thulborn, K. R. 1996. Brain activation modulated by sentence comprehension. *Science* **274:** 114–116.

Kang, A. M., Constable, R. T., Gore, J. C., and Avrutin, S. 1999. An event-related fMRI study of implicit phrase-level syntactic and semantic processing. *Neuroimage* **10:** 555–561.

Keller, T. A., Carpenter, P. A., and Just, M. A. 2001. The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cereb. Cortex* **11:** 223–237.

Kim, K. H. S., Relkin, N. R., Lee, K. M., and Hirsch, J. 1997. Distinct cortical areas associated with native and second languages. *Nature* **388:** 171–174.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., and Mehler, J. 1993. The cortical representation of speech. *J. Cogn. Neurosci.* **5:** 467–479.

Michael, E. B., Keller, T. A., Carpenter, P. A., and Just, M. A. 2001. fMRI investigation of sentence comprehension by eye and by ear: Modality fingerprints on cognitive processes. *Hum. Brain Mapp.* **13:** 239–252.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., and Fazio, F. 2001. Syntax and the brain: Disentangling grammar by selective anomalies. *Neuroimage* **13:** 110–118.

Mummery, C. J., Shallice, T., and Price, C. J. 1999. Dual-process model in semantic priming: A functional imaging perspective. *Neuroimage* **9:** 516–525.

Newman, S. D., Twieg, D., and Carpenter, P. A. 2001. Baseline conditions and subtractive logic in neuroimaging. *Hum. Brain Mapp.* **14:** 228–235.

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., Shaywitz, B. A., Gore, J. C., and Shankweiler, D. 2000. An event-related neuroimaging study distinguishing form and content in sentence processing. *J. Cogn. Neurosci.* **12:** 120–133.

O'Leary, D. S., Andreasen, N. C., Hurtig, R. R., Torres, I. J., Flashman, L. A., Kesler, M. L., Arndt, S. V., Cizadlo, T. J., Ponto, L. L. B., Watkins, G. L., and Hichwa, R. D. 1997. Auditory and visual attention assessed with PET. *Hum. Brain Mapp.* **5:** 422–436.

Oldfield, R. C. 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9:** 97–113.

Patterson, K., and Shewell, C. 1987. Speak and spell: Dissociations and word-class effects. In *The Cognitive Neuropsychology of Language* (M. Coltheart, G. Sartori, and R. Job, Eds.), pp. 273–294. Lawrence Erlbaum, London, UK.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. 1989. Positron emission tomographic studies of the processing of single words. *J. Cogn. Neurosci.* **1:** 153–170.

Petit, L., Courtney, S. M., Ungerleider, L. G., and Haxby, J. V. 1998. Sustained activity in the medial wall during working memory delays. *J. Neurosci.* **18:** 9429–9437.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. E. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* **10:** 15–35.

Price, C. J. 1998. The functional anatomy of word comprehension and production. *Trends Cogn. Sci.* **2:** 281–288.

Price, C. J. 2000. The anatomy of language: contributions from functional neuroimaging. *J. Anat.* **197:** 335–359.

Price, C. J., and Friston, K. J. 1997. Cognitive conjunction: A new approach to brain activation experiments. *Neuroimage* **5:** 261–270.

Price, C. J., Wise, R. J. S., and Frackowiak, R. S. J. 1996. Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb. Cortex* **6:** 62–70.

Sakai, K., and Miyashita, Y. 1994. Visual imagery: An interaction between memory retrieval and focal attention. *Trends Neurosci.* **17:** 287–289.

Sakai, K. L., Hashimoto, R., and Homae, F. 2001. Sentence processing in the cerebral cortex. *Neurosci. Res.* **39:** 1–10.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., Marchione, K. E., Fletcher, J. M., Klorman, R., Lacadie, C., and Gore, J. C. 2001. The functional neural architecture of components of attention in language-processing tasks. *Neuroimage* **13:** 601–612.

Sidtis, J. J., Strother, S. C., Anderson, J. R., and Rottenberg, D. A. 1999. Are brain functions really additive? *Neuroimage* **9:** 490–496.

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

Tagamets, M. A., Novick, J. M., Chalmers, M. L., and Friedman, R. B. 2000. A parametric approach to orthographic processing in the brain: An fMRI study. *J. Cogn. Neurosci.* **12:** 281–297.

Talairach, J., and Tournoux, P. 1988. *Co-Planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging.* Thieme, Stuttgart.

Talavage, T. M., Edmister, W. B., Ledden, P. J., and Weisskoff, R. M. 1999. Quantitative assessment of auditory cortex responses induced by imager acoustic noise. *Hum. Brain Mapp.* **7:** 79–88.

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. USA* **94:** 14792–14797.

Ulmer, J. L., Biswal, B. B., Yetkin, F. Z., Mark, L. P., Mathews, V. P., Prost, R. W., Estkowski, L. D., McAuliffe, T. L., Haughton, V. M., and Daniels, D. L. 1998. Cortical activation response to acoustic echo planar scanner noise. *J. Comput. Assist. Tomogr.* **22:** 111–119.

Wojciulik, E., and Kanwisher, N. 1999. The generality of parietal involvement in visual attention. *Neuron* **23:** 747–764.