

# Neuroanatomic Organization of Sound Memory in Humans

Michael A. Kraut<sup>1</sup>, Jeffery A. Pitcock<sup>2</sup>, Vince Calhoun<sup>3</sup>, Juan Li<sup>2</sup>,  
Thomas Freeman<sup>2</sup>, and John Hart, Jr.<sup>4</sup>

## Abstract

■ The neural interface between sensory perception and memory is a central issue in neuroscience, particularly initial memory organization following perceptual analyses. We used functional magnetic resonance imaging to identify anatomic regions extracting initial auditory semantic memory information related to environmental sounds. Two distinct anatomic foci were detected in the right superior temporal gyrus when subjects identified sounds representing either animals or threatening items. Threatening animal stimuli elicited signal changes in both foci, suggesting a distributed neural repre-

sentation. Our results demonstrate both category- and feature-specific responses to nonverbal sounds in early stages of extracting semantic memory information from these sounds. This organization allows for these category-feature detection nodes to extract early, semantic memory information for efficient processing of transient sound stimuli. Neural regions selective for threatening sounds are similar to those of non-human primates, demonstrating semantic memory organization for basic biological/survival primitives are present across species. ■

## INTRODUCTION

Assigning meaning to the sounds we hear engages stages of auditory processing up to and including the determination of whether the listener has heard the sound before and what it represents. There have been few studies addressing the early stages of access to semantic memory stores occurring after initial perceptual processing of the acoustic properties of auditory stimuli. Investigations of the neural organization of semantic memory for *visually* presented objects (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Martin, Wiggs, Ungerleider, & Haxby, 1996) have demonstrated that access to basic levels of information about different categories of objects is associated with functional magnetic resonance imaging (fMRI)–detectable signal changes in distinct loci in the ventral temporal lobes, although there is good evidence that the information is organized on the basis of object categories or their component features (Haxby et al., 2001). Another system, the lexical–semantic system, is also organized both functionally and anatomically by categories (e.g., animals, fruit and vegetables, tools) (Caramazza & Shelton, 1998; Hart, Berndt, & Caramazza, 1985; Warrington & Shallice, 1984; Nielsen, 1946) and/or features (e.g., visual perceptual) (Hart & Gordon, 1992).

The semantic memory network associated with object sound recognition has been explored to some degree. Studies have shown that multimodality semantic categorization or identification tasks pairing environmental sounds with visual object representations are associated with fMRI signal changes detected in the frontal lobes bilaterally (Adams & Janata, 2002). However, such studies did not allow for direct isolation of regions involved with auditory semantic object memory. In a study of environmental sound recognition, Lewis et al. (2004) assessed patterns of brain activation during evaluation of sounds from multiple object categories, detecting an extensive network of brain regions predominantly in the left hemisphere (including bilateral posterior middle temporal gyri), which were imputed to have roles in processing multimodal semantic information about objects.

Overall, the auditory system has not been interrogated to determine the organization of object sound semantic memory as distinct from (1) auditory linguistic representation for object names and (2) primary auditory perceptual processing of sounds. Anatomic organization of these linguistic representations is subdivided by phonemes/speechlike sounds in the middle of the left superior temporal gyrus, with higher order representations for word perception in the posterior left superior temporal gyrus (Specht & Reul, 2003; Wise et al., 2001; Zatorre & Belin, 2001; Binder et al., 2000; Scott, Blank, Rosen, & Wise, 2000; Hugdahl et al., 1999; Mummery et al., 1999; Belin et al., 1998). These language-related signal changes in the left superior temporal gyrus have

<sup>1</sup>The Johns Hopkins University School of Medicine, <sup>2</sup>University of Arkansas for Medical Sciences, <sup>3</sup>Yale University, <sup>4</sup>University of Texas at Dallas

been attributed to this region's specialization for the analysis of speech sounds (Binder, Frost, Hammeke, Rao, & Cox, 1996; Binder et al., 1995). A comparable system has been detected in nonhuman primates in the superior temporal gyrus for species-specific vocalizations (Rauschecker & Tian, 2000; Rauschecker, 1998).

From the perspective of auditory perceptual processing, there is a well-defined anatomic organization imputed to the processing of the characteristics of nonverbal object sounds at a basic acoustic/perceptual level, localized within the right superior temporal gyrus (Tzourio et al., 1997). Analyses of neural processing beyond the basic perceptual acoustic level have shown differential activation in (1) the left posterior superior temporal region for complex auditory-temporal features in speech and nonspeech stimuli, (2) middle left superior temporal region for the acoustic profile of human voices, (3) anterior left superior temporal sulcus to intelligible speech, and (4) right superior temporal gyrus region to pure tones and environmental sounds in general (Specht & Reul, 2003). Thus, environmental sounds have been investigated as a single, undifferentiated stimulus group to explore relatively early perceptual stages of auditory processing or links to other semantic representations or networks.

There have been no previous studies investigating if there are subdivisions (e.g., categories, features) to the organization of environmental object sounds per se in nonprimary auditory cortical regions, where initial neural access to auditory semantic memory might be mediated. Because rapid evaluation of environmental sounds has been a fundamentally important tool for survival, we postulated that there are brain regions that respond to environmental sound stimuli, and respond differentially to them on the basis of their semantic valence. We further postulated that at least some of these regions are close to the primary auditory cortex to facilitate efficient initial assignments of semantic significance. This study aimed to find some of those regions. In addition to detecting brain regions that demonstrate activation patterns reflecting early access to semantic memory through the auditory system, we assessed if these regions are organized by categories and/or features. We used animals and nonliving objects as categories, and threatening versus nonthreatening as distinguishing features. Category/featural distinctions were chosen because such stimulus properties have been previously isolated in other semantic memory systems (e.g., animals in the lexical-semantic system) and/or perhaps are evolutionarily important to detect and classify. To isolate the signal changes specifically related to initial accessing of auditory semantic memory, task design consisted of the subjects being asked to determine whether they recognized a sound to represent a real item. This task construction sought to avoid preferential targeting of particular categories, objects, or features. To reduce primary auditory cortical activation

as a factor, each real stimulus was temporally scrambled, so that the frequency spectrum of the scrambled sound resembled that of the original sound. We contrasted fMRI signal changes elicited by specific categories of objects (animals vs. nonliving objects), and by specific features (threatening vs. nonthreatening).

## METHODS

### Subjects

There were 18 right-handed participants, 13 men and 5 women between the ages of 20 and 51 years (mean = 32.2 years,  $SD = 10.2$ ) who had no neurological or psychiatric impairments.

### Stimuli

The stimuli used in the study consisted of 32 environmental sounds chosen from a set of 120 sounds normed in a previous study for familiarity, pleasantness, confidence in naming, naming accuracy, duration, and reaction time (Marcell, Borella, Greene, Kerr, & Rogers, 2000). The 120 sounds as well as the normative data were downloaded from the Internet site, [www.cofc.edu/~marcellm/confrontation%20sound%20naming/confront.htm](http://www.cofc.edu/~marcellm/confrontation%20sound%20naming/confront.htm).

Each of these sounds was then modified to create a corresponding meaningless sound. Simple reversal of the time courses of the sound stimuli left them recognizable in many cases and thus plausibly identifiable as exemplars of real items. Similarly, binning of the time courses into short (tens of milliseconds) segments and scrambling of those segments still resulted in stimuli that might be identified as real items. We ultimately chose to modify the original sounds by creating a spectrum-deforming template, using Cool Edit Pro 2.1 (Syntrillium Software Corp., Phoenix, AZ), and applying the same template to each real sound, changing the overall time frequency envelope until the sounds were not identifiable. The duration of each modified sound was the same as its parent sound, as was the overall area under each time-power curve. Two-tailed  $t$  tests revealed no significant differences between the root-mean-squared power of any of the real sounds and their scrambled derivative nonreal sound. Likewise, analysis of the two main comparison groups revealed no significant differences between the root-mean-squared power of the animal sounds compared to nonliving object sounds, or threatening sounds to nonthreatening sounds (two-tailed  $t$  tests). The peak amplitudes of all the real and nonreal sounds were normalized to a common peak amplitude and adjusted so that there was no clipping of the waveform on final sound output. This normalization of peak amplitudes was performed to ensure that detected signal changes for a stimulus or group of stimuli are not related to the loudness of the stimuli, which is

particularly relevant for the semantic distinctions investigated (e.g., animal, threatening).

There were 32 real sounds in the final stimulus set. These were derived by taking the 120 sounds from Marcel et al. (2000) and segregating them into two categories of animals and living objects that were identifiable when heard in isolation. Each of these two categories was then further divided into groups of threatening and nonthreatening stimuli. This threatening designation was derived in the following fashion. These items were equally divided into two groups based on the rating on the pleasantness rating scale of Marcel et al. (items ranged between 1.52 and 6.36 on this scale) forming one “unpleasant” group and another “pleasant” group. From the “unpleasant” stimuli, a set of items was chosen that represented unpleasant and also threatening stimuli (tiger roar, gunshots, etc.). From the “pleasant” group, stimuli were chosen that were also nonthreatening stimuli (cow moo, bicycle bell, etc.). The threatening and nonthreatening groups were then segregated into items that filled the following cells: (1) threatening animals, (2) threatening objects, (3) nonthreatening animals, and (4) nonthreatening objects. The groups contained 8 sounds each with the resultant threatening groups collapsed across category having a mean rating of 4.42,  $SD = .67$  and the nonthreatening group having a mean rating of 2.65,  $SD = .84$ .

## Procedures

The 64 sounds were pseudorandomized and presented using E-Prime (Psychology Software Tools, Inc., Pittsburgh, PA) with an interstimulus interval of 8.2 sec via MRI-compatible, pneumatically driven headphones. The frequency transmission characteristics of the amplifier/pneumatic tube combination was such that the response was maximal at about 2 kHz (102 dBc), remaining at about 90 dBc at 4 kHz, and, at lower frequencies, dropping to 86 dBc at 250 Hz.

Subjects were instructed to use a button box placed in their left hand to indicate whether they perceived a sound to be real (sound of an item they have previously heard) or nonreal (not a sound of an item with which they were familiar). An index finger button push indicated a perceived real object sound, and a middle finger button push indicated a perceived nonreal object sound. Subjects were monitored for test performance.

fMRI data were acquired on a 1.5-T MRI instrument, using a standard head coil. The data were acquired in the coronal plane, using a TR of 3 sec, a TE of 35 msec, and a flip angle of 90°. Section thickness was 4 mm, with a 0.5-mm gap between scan sections. Field of view was 24 cm, with a 64 × 64 acquisition matrix.

Image analysis was performed by using SPM99. Imaging data for each individual was adjusted for section acquisition time, corrected for motion, normalized into a standardized Talairach template, and spatially smoothed

(10 × 10 × 10 Gaussian kernel) using SPM99. Signal changes were modeled as delta functions temporally coincident with the onset of each stimulus and convolved with a canonical hemodynamic response function. Initially, event-related analyses were conducted using a random-effects model excluding all false-positive and false-negative responses from the model, comparing responses reflecting detection of real items to responses to nonreal items. We established four regressors for our model to represent animal, nonliving, threatening, and nonthreatening/pleasant stimuli, and used these regressors to detect differences in the responses to animals versus nonliving items and also between threatening and nonthreatening stimuli, using the same random-effects schema. We subsequently subjected each focus of signal change that emerged with either the animal/nonliving or the threatening/nonthreatening contrast to a small volume correction (SVC) analysis, run under SPM99, to determine the degree to which a region of signal change that emerged with one contrast (e.g., animal/nonliving) also responded to the distinction drawn in the other contrast (threatening/nonthreatening).

## RESULTS

### Behavioral Results

Performance results showed that the average correct response rate for all stimuli was 82% (false-negative rate of 31% and false-positive rate of 6%). Behaviorally, the findings from these analyses showed that the proportions of correct responses to real items, regardless of category, were all approximately 70%, and indicate that the subjects were able to hear the stimuli throughout the experimental sessions (Table 1).

### Imaging Results

None of the signal-change contrasts performed, including those comparing any of the classes of “real” stimuli

**Table 1.** Grouped Subject Performance for the Different Classes of Stimuli Delivered during these Experiments

<i>Sound Stimuli</i>	<i>Correct Response (%)</i>	<i>Stimulus Duration, Average ± SD (msec)</i>
All	82	1772 ± 670
Real	69	1772 ± 670
Nonreal	94	1772 ± 669
Animals	67	1765 ± 827
Objects	71	1778 ± 489
Threatening items	72	1444 ± 430
Nonthreatening items	67	2101 ± 682

to their “nonreal” modified forms, showed statistically significant signal change in the primary auditory cortex, suggesting that the degree of primary auditory cortical activity elicited by the real and nonreal stimuli was approximately equivalent.

### Primary Effects

1. Responses to threatening sounds, as contrasted with nonthreatening sounds, thresholded at  $p < .001$  uncorrected, are focused in (1) the right superior temporal gyrus (Brodmann’s area [BA] 41), posterior to the typical boundaries of the right primary auditory cortex as delineated by Rademacher et al. (2001); (2) the right inferior frontal gyrus (BA 47) extending medially toward the lateral orbitofrontal region; and (3) the right superior parietal lobule (BA 7) (Figure 1 and Table 2).

2. The converse contrast, nonthreatening sounds compared to threatening sounds, revealed a single focus of signal change, in BA 19, in the left precuneus (Table 2).

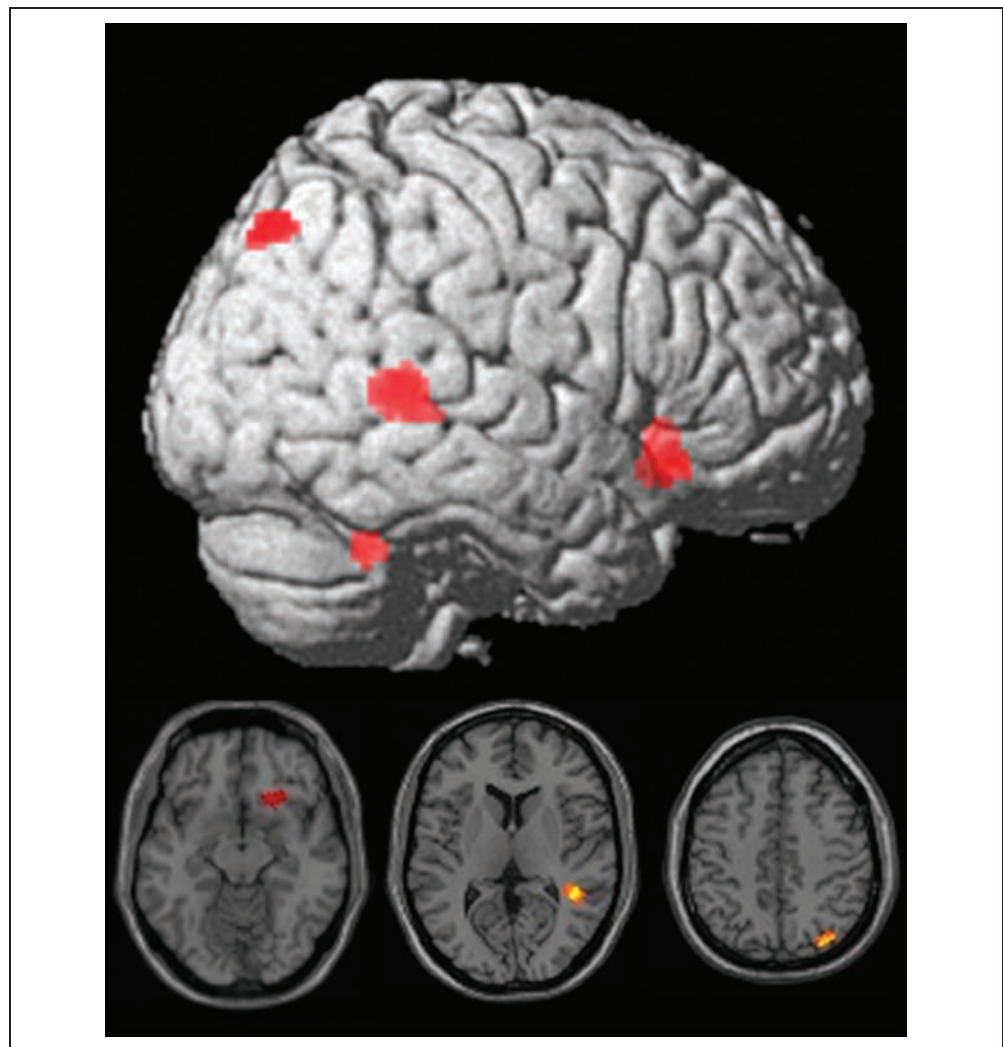
3. Contrasting responses to stimuli representing animals with those of nonliving objects revealed foci of

signal change in (1) the right superior temporal gyrus (BA 22) but rostral to the focus evident in the threatening/nonthreatening contrast (Figure 1) and also outside of the noted boundaries of the primary auditory cortex (Rademacher et al., 2001), (2) the left middle temporal gyrus (BA 21 and 44), (3) the left superior frontal gyrus (BA 8), and (4) the medial aspect of the left frontal lobe (BA 32) (random effects analysis,  $p < .001$ , uncorrected; Figure 2 and Table 2). Contrasting nonliving objects with animals resulted in no significant signal changes.

### Secondary Effects: Small Volume Correction Analyses of Primary Effects

1. The SVC analysis of the regions of signal change that emerged with the threatening/nonthreatening contrast showed that the animal subset of the threatening sounds demonstrated signal changes in the right superior temporal gyrus cluster at a level of  $p < .002$  corrected (voxel level), and that there were no areas of signal change that could be identified for the nonliving subset of threatening sounds at the same thresh-

**Figure 1.** fMRI signal changes that emerge with the threatening/nonthreatening contrast, projected onto axial images, and a surface-rendered view of the lateral aspect of the right cerebral hemisphere. The focus in the right superior temporal gyrus is posterior to the expected extent of the primary auditory cortex (Rademacher et al., 2001).



**Table 2.** Locations of fMRI Signal Change Evident upon Contrasting Responses to Threatening vs. Nonthreatening Stimuli and to Animals vs. Nonliving Items

Cluster Size	<i>t</i>	Talairach Coordinates			Anatomical Location
		<i>x</i>	<i>y</i>	<i>z</i>	
<i>Threatening–nonthreatening random effects analysis (p &lt; .001 uncorrected)</i>					
151	5.69	44	−40	8	Right STG (BA 41)
109	4.64	18	21	−8	Right frontal (BA 47)
	4.14	26	25	−11	Right IFG (BA 47)
	3.72	28	17	−9	Right extranuclear (BA 13)
54	4.62	32	−72	46	Right superior parietal (BA 7)
17	4.51	36	−52	−26	Right culmen
25	4.27	−32	9	35	Left precentral gyrus (BA 9)
	3.89	−36	2	39	Left MFG (BA 6)
17	4.22	−32	23	23	Left subgyral
5	4.18	38	51	7	Right MFG (BA 10)
4	4.17	−57	−31	5	Left MTG (BA 22)
2	4.1	−34	44	−11	Left MFG (BA 11)
2	3.98	−30	44	−11	Left MFG (BA 11)
<i>Nonthreatening–threatening random effects analysis (p &lt; .001 uncorrected)</i>					
19	5.36	−22	−83	41	Left parietal (BA 19)
<i>Animal–nonliving random effects analysis (p &lt; .001 uncorrected)</i>					
50	5.05	57	−15	4	Right STG (BA 22)
5	4.95	−65	−37	−2	Left MTG (BA 21)
12	4.49	−50	−39	0	Left MTG (BA 22)
9	4.29	−12	12	47	Left MTG (BA 32)
10	4.11	0	16	49	Left SFG (BA 8)
1	3.94	−63	−14	−3	Left MTG (BA 21)
1	3.7	−57	5	−9	Left STG (BA 38)
1	3.67	−12	−5	9	Left thalamus (ventral anterior)

BA = Brodmann area; IFG = inferior frontal gyrus; MFG = middle frontal gyrus; MTG = middle temporal gyrus; SFG = superior frontal gyrus; STG = superior temporal gyrus.

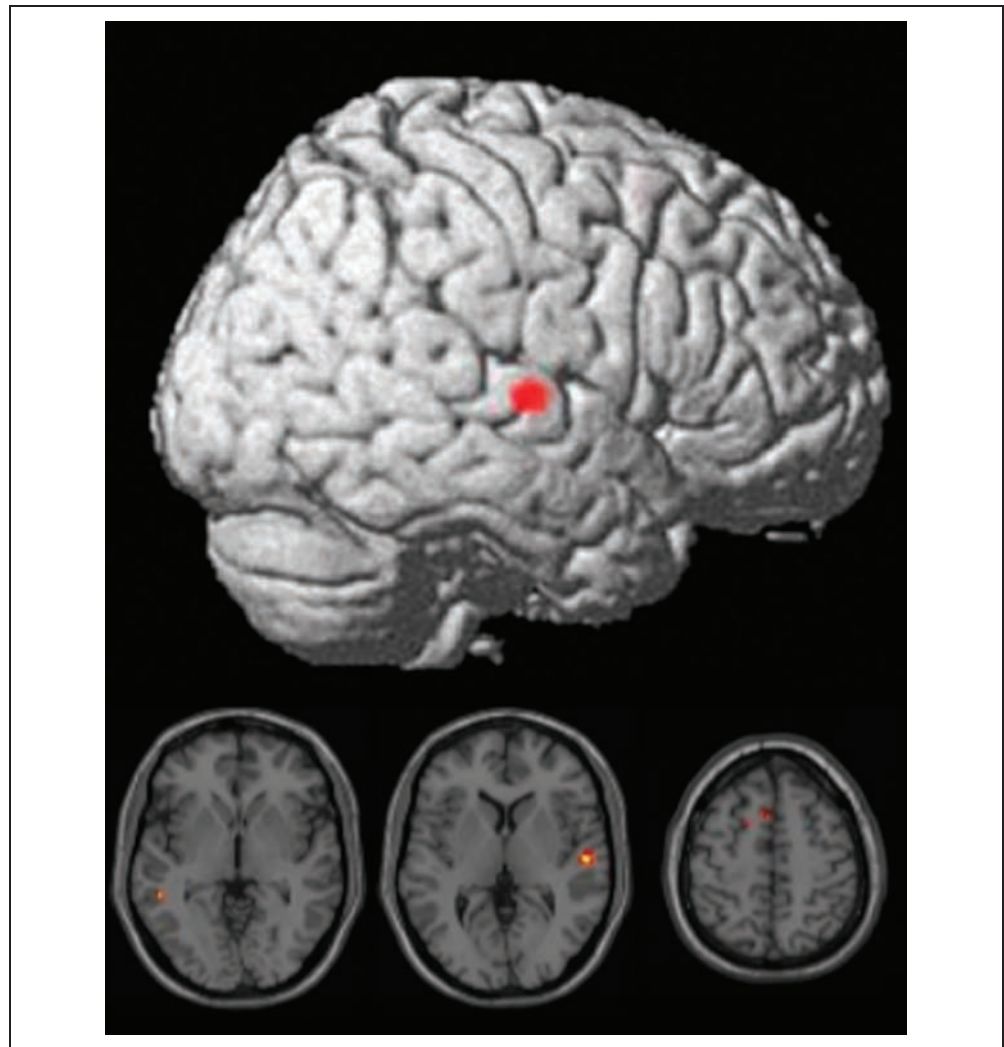
old. Thus, the focus of signal change revealed with the threatening versus nonthreatening (featural) contrast is driven in substantial measure by the distinction between animal and object (category) aspects of the stimuli. The right inferior frontal gyrus region likewise demonstrates significantly greater signal change as subjects evaluate stimuli that represent threatening animals compared to the responses to threatening but nonliving objects ( $p < .001$  corrected, voxel level). Conversely, SVC analysis indicates that the right parietal focus evident in the threatening/nonthreatening contrast exhibits statistic-

ally significant signal change only with identification of stimuli that represent nonliving objects ( $p < .002$  corrected, voxel level).

2. SVC-based evaluation of the BA 19 focus that emerged with the nonthreatening/threatening contrast showed no preference for the animal versus the nonliving stimuli.

3. SVC analysis of the more rostral right temporal and of the lateral left temporal foci of signal change found with the animal/nonliving contrast showed signal changes for the threatening animals at  $p < .002$

**Figure 2.** Signal changes specific to stimuli that represent animals, again superimposed upon axial images and a right hemisphere surface-rendered view. The main focus again is along the right superior temporal gyrus, but rostral to the main right STG focus in Figure 1 and again outside the expected extent of the primary auditory cortex.



corrected (voxel level). There were no areas of statistically significant signal change in either of these regions under SVC analysis for the nonthreatening animals at this threshold, indicating that in these regions, the signal change revealed with the animal/object (category) contrast reflects primarily the threatening (featural) aspect of the animal stimuli. None of the other regions of signal change that emerged with the animal/nonliving contrast demonstrated statistically significant signal change for either the threatening or the nonthreatening stimuli analyzed in isolation.

## DISCUSSION

The present findings demonstrate patterns of fMRI signal change that reflect predominantly whether the perceived stimuli represent animals or threatening items, with nonthreatening items and nonliving objects eliciting signal changes that indicate less extensive activation. The most striking findings are the two foci of signal change in the right temporal lobe in the immedi-

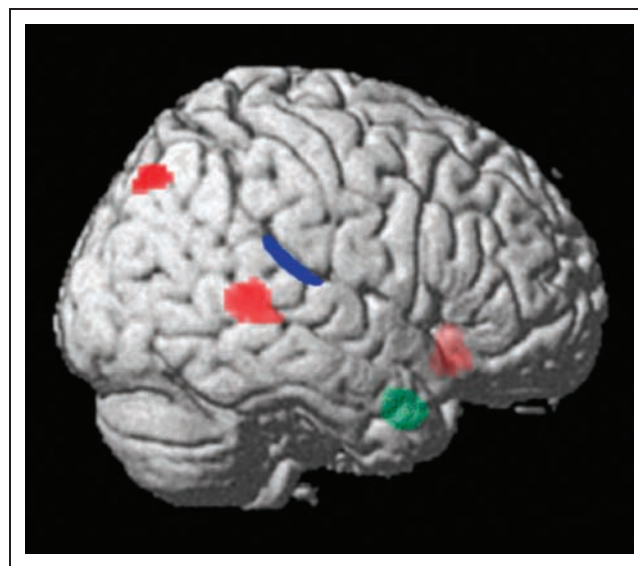
ate vicinity of but not within the primary auditory cortex, which become evident with evaluation of stimuli that represent animals (category-level distinction) or threatening (feature level distinction) items. The collocation of signal changes (responses to both categorical and featural attributes) in the right superior temporal region suggests that early access to semantic information about items in the auditory semantic system is likely accomplished through the coactivation of the multiple attributes (both categorical and featural) of an item, similar to the visual system (Haxby et al., 2001).

Outside of the right temporal lobe, signal changes associated with animal sound stimuli are evident in the left middle temporal gyrus (BA 21 and 44), left superior frontal gyrus (BA 8), and the medial aspect of the left frontal lobe (BA 32) including the anterior cingulate region. Lewis et al. (2004) detected left middle temporal gyrus activation in a task similar to ours, but that included environmental sounds from multiple object categories, and suggested that the signal changes detected in these regions were attributable to multimodal semantic processing. This contention, with which we

agree, fits well with a schema under which access to lexical-semantic aspects of visually presented stimuli that represent animals is subserved by regions within the ventral left temporo-occipital region (Martin et al., 1996), and multimodal (auditory/visual) processing of the semantic stimulus attributes for stimuli from multiple categories is localized more dorsolaterally on the temporal convexity (Beauchamp, Lee, Argall, & Martin, 2004; Hart & Gordon, 1990). The anterior cingulate region, which comprises parts of BA 32 and BA 24, is interconnected with the auditory association cortex (Barbas, 1988), as well as with other limbic structures, for example, the amygdala (see Devinsky, Morrell, & Vogt, 1995, for a review), and has been imputed to play a role in modulating affect (Devinsky et al., 1995) and attention (Morecraft, Geula, & Mesulam, 1993; Posner & Petersen, 1990). Although it is somewhat surprising given the limbic connections that the signal changes in this region were not differentially evident with evaluation of stimuli that represent either threatening animals or objects, it may be that auditory stimuli are not frightening (unless they are very loud, in which case they may be more startling than frightening) when they are experienced in an experimental environment.

The right superior parietal lobule (BA 7) typically has been associated in functional imaging studies with motor attention (Rushworth, Johansen-Berg, Gobel, & Devlin, 2003) and with planning motor movements (Heilman & Gonzalez Rothi, 1993); however, recent studies have suggested that this region is engaged in long-term episodic memory retrieval (Shannon & Buckner, 2004; Rugg & Wilding, 2000). These findings showed that the regional fMRI signal changes were significantly greater during correct retrieval versus correct rejection in the bilateral posterior parietal cortex (right > left). Our findings showed that this same region was activated selectively on the right for threatening sounds and on the left for nonthreatening sounds. This suggests that although these regions are bilaterally activated for correct recall trials versus distractors in episodic memory, there is lateralization of activity related to specific properties of the stimuli retrieved. In our experimental paradigm, the main differentiating feature, across both the animal and nonliving objects, is whether the sound represents an item that is threatening or, perhaps more generally, unpleasant. Dominance of right-sided brain structures in the processing of stimuli perceived as threatening has been observed in rodents (Wall, Blanchard, Yang, & Blanchard, 2004; Sullivan & Gratton, 2002), perhaps on the basis of dopaminergic circuitry (Wall et al., 2004), as well as in nonhuman primates (Fernandez-Carriba, Loeches, Morcillo, & Hopkins, 2002; Hauser, 1993) and humans (Compton, Heller, Banich, Palmieri, & Miller, 2000; Cooley, Stringer, & Hodnett, 1997; Otto, McNally, Pollack, Chen, & Rosenbaum, 1994). The lateralization of signal changes for threatening sounds to the right hemisphere sup-

ports the contention of Davidson (1998), who has suggested that whereas the left hemisphere predominantly mediates positive emotional responses, the right hemisphere mediates negative emotional responses. Right-hemisphere lateralization for negative affective responses have also been demonstrated for fear recognition (Adolphs, Damasio, Tranel, & Damasio, 1996) and the relative activation of the right hemisphere (as measured by positron emission tomography and fMRI) in adults who received repeatedly painful stimuli (Hsieh, Stone-Elender, & Ingvar, 1999), experience unpleasant tastes (Zald, Lee, Fluegel, & Pardo, 1998), or view photographs of negative emotional scenes (Canli, Desmond, Zhao, Glover, & Gabrieli, 1998). These findings together with ours provide concordant evidence of a neural circuit dedicated to the evaluation of auditory stimuli that represent threatening or unpleasant stimuli, with this circuit being lateralized predominantly to the right hemisphere (see Figure 3), which includes the right superior temporal gyrus, superior parietal lobule, inferior frontal lobe, and amygdala. The amygdala did not demonstrate significant activation for threatening sounds in this study; however, the inferior frontal lobe did show significant signal change and it has been associated with the amygdala in processing of fearful visual stimuli in several studies (Benuzzi et al., 2004; Buchel, Dolan, Armony, & Friston, 1999). We posit that although the amygdala is associated with our described circuit for



**Figure 3.** Schema of brain regions engaged with threatening sounds, projected onto lateral view of the right hemisphere. The regions in red are the actual signal changes from this study that were isolated by comparing threatening to nonthreatening sounds. The area in blue is drawn in as an approximation of the rostrocaudal extent of primary auditory cortex (Rademacher et al., 2001), and the green structure is drawn in as an approximation of the amygdala, which has been shown in other studies to be engaged in threatening emotional processing.

threatening sounds, it did not show significant signal changes in our paradigm either because auditory stimuli delivered in the relatively detached context of an fMRI experiment may not be perceived as compellingly frightening (e.g., a lion's roar heard while performing an experiment during fMRI vs. in the jungle), as we hypothesized above, or that auditory threatening stimuli are not as robust as visual threatening stimuli in engaging the limbic fear system. The right inferior frontal lobe is ideally situated to mediate the association of sensory stimuli with these limbic/emotional representations, and the orbitofrontal region, into the lateral aspect of which the ventral frontal signal changes appear to have spread, has been shown to have strong reciprocal connections with the amygdala in monkeys (Ghashghaei & Barbas, 2002). Ventral frontal structures are thus plausibly in good position to modulate, and perhaps dampen, amygdala responses to stimuli that represent potentially threatening items. Overall, this selective circuit for threatening auditory stimuli, including a threatening-specific detector neighboring primary auditory cortex, is in keeping with recent findings in nonhuman primates of a similar circuit for threatening animal "screams," further suggesting that this represents a basic aspect of semantic memory organization that extends to other primate species (Gil-da-Costa et al., 2004).

There are several factors that influence our results. One is that, in general, there is a limited number of sound stimuli that are strongly or uniquely characteristic of their associated item. Thus, the number of recognizable sounds that could be tested and the number of categories and features among which our subjects could distinguish were limited. Also affecting the number of stimuli available for testing in the auditory semantic system is the variability of individual subjects' previous exposure to specific stimuli or classes of stimuli, and thus possible expertise in evaluating them. Within the visual system, expertise/exposure factors clearly influences both behavioral results as well as patterns of fMRI signal change, both in the field of expertise and more generally (Kronbichler et al., 2004; Tarr & Gauthier, 2000). In addition, potentially confounding our assignment of stimulus groupings and labels to the identified brain regions is our imposition of a classification or processing hierarchy (category/object/feature) upon the auditory system. Although our schema is intuitive and consistent with previously proposed schemata for organization in semantic memory systems, it may not be consistent with the classification system inherent to neural organization in auditory semantic memory, which could conceivably reflect properties more closely associated with acoustic distinctions (other than peak loudness, for which we controlled) such as rate-of-loudness or rate-of-frequency change, or properties/groupings that we did not consider.

Even in the face of these caveats, animals have been demonstrated to represent a functionally and anatomi-

cally distinct category in both the lexical-semantic and possibly visual systems, with several theories posited for its compelling distinct representations. The sensory-functional hypothesis (Gainotti, 2000) suggests that animals are encoded as a category because its exemplars are readily distinguished from items in other categories, and from other animals within the category, by the visual-perceptual properties each animal possesses. This visual-perceptual distinction has been supported by the anatomic localization of neural access to information about the animal category to the posterior inferior temporo-occipital region (Martin et al., 1996). These perceptual systems distinctions may be paralleled in the auditory system, as animals can be distinguished from items in other categories and from each other by their vocalizations (e.g., barks, growls, roars). The categorical organization of animals in the auditory semantic system is also consistent with the theory that items of evolutionary significance, such as animals, are likely to be encoded categorically in semantic systems (Caramazza & Shelton, 1998). Similarly, threatening items have an equally compelling evolutionary significance of representing dangerous stimuli, which would initiate a sensorimotor (flight/fight) response, and thus merit efficient processing as a salient subcategory of living things. Another possible organizational framework is that suggested by Jaaskelainen et al. (2004) in a study focused on delineating the neural generators of the auditory mismatch negativity (MMN) response. They found evidence for two auditory cortical generators for the MMN in the superior temporal gyrus region, with the posterior one more active when an auditory stimulus was novel. They speculated that the adaptive properties of the posterior source make it well suited for rapid analysis of novel stimuli, such as would be useful in detection of stimuli that represent dangerous or threatening items. Given the specificity of our stimuli, we would suggest that the posterior generator reported by Jaaskelainen et al. may well function primarily for detecting or assessing threatening stimuli.

More generally, our data provide support for an organizational schema within the auditory semantic system similar to that which may be present in the visual system, such that the representation of an item sound in the auditory semantic system is instantiated by the coactivation of regions that encode for the characteristics (e.g., category, features) of that item. In addition to interactions within the auditory system, output from the regions we have delineated could potentially project to or merge with outputs from brain regions associated with the same attribute in other semantic systems (e.g., verbal, visual) or a common amodal semantic system.

A previous study of nonverbal sounds (Specht & Reul, 2003) did not investigate subdivisions for different categories of items, and localized responses to all environmental sounds to the right superior temporal gyrus in general. Because that study did not distinguish between,



or assess for, subtypes of environmental sounds, it was not possible to further localize stimuli by categorical or other groupings. However, because many of the environmental sounds used in that study were animal sounds, their similar localization may have been largely the consequence of the animal-sound stimuli. Our localization in this study for animals is more posterior in the right superior temporal gyrus than is the general environmental sound localization Specht and Reul (2003) proposed, likely due to our categorical stimulus grouping consisting of only animal sounds. Our work builds upon that of Lewis et al. (2004) and evaluates the system at more specific levels of semantic memory organization and processing in that the Lewis et al. study focused on localizing patterns of brain activity associated with the classification of environmental item sounds in general. The present study's experimental design and analyses allowed for probing patterns of neural activity related to the specific category and featural representations encoded in auditory semantic memory for distinct classes of items.

Previous investigations have not focused primarily on functional subdivisions and organization of nonverbal sound memory in the auditory association cortices. The patterns of signal change revealed in our data, demonstrating evidence for category and feature-specific activation, represent components of a framework for the organization of the functional–anatomic interface between perception and memory in the auditory sound system. The auditory schema may possess similarities to that in the visual system, but there are differences that likely reflect the spatial and temporal properties of auditory as distinct from visual stimuli. Processing of information in the visual system proceeds from the primary visual cortex, through multiple anatomic regions that extract and process higher order form, color, orientation, and motion information, to a visual object memory system (the “what” system) located bilaterally in the ventral temporo-occipital cortices (Ungerleider & Haxby, 1994; Iwai & Mishkin, 1968). These latter temporo-occipital regions extend along the rostrocaudal axis of the basal convexities of the brain.

Visual stimuli, unless they are only rapidly and transiently presented, can be scanned repeatedly as necessary for analysis of their features. In addition, the processing of visual stimuli necessitates that considerable perceptual processing be performed to analyze the diverse number of objects in the visual world, the variety of “tokens” that represent the same object, and the almost innumerable possible spatial orientations of presentation. These aspects of visual stimuli plausibly necessitate the large numbers of cortical regions dedicated to primary perceptual and higher order complex perceptual processing. As a consequence, the number of regions of the brain that have been imputed to mediate access to visual memory, the ventral temporal regions, are relatively far removed from the primary visual cortex.

In contrast, the serial and transient nature by which auditory stimuli are momentarily present mandates rapid, online identification of the objects/items producing those sounds. This efficient, rapid processing can be expeditiously accomplished through close proximity between regions that perform the early stages of auditory perceptual processing and those regions that perform higher order operations, such as evaluating perceptual constructs in auditory memory in the context of the object identification process. Based on the nature of auditory stimuli, we propose that auditory information processing likely proceeds in a cascaded sequence (McClelland, 1979), where partial products of the perceptual processing can be forwarded to the memory system prior to completion of perceptual processing in order for the memory system to start and narrow the choices/search in that system. Given the left-hemisphere auditory system's relatively predominant role in language and speech assessment, the right hemisphere is more likely to be predominantly dedicated to nonverbal sound.

We propose the *neural semantic detector* model to account for our findings in terms of the functional–anatomic organization of semantic memory for nonverbal sounds. The stimuli we employed preferentially activated two right superior temporal loci of fMRI signal change, which we posit to represent “nodes” or “detectors” that are engaged in early auditory semantic memory processing and which respond differentially to the animal category and the threatening feature of sound stimuli (Kronbichler et al., 2004). In this proposed framework, there are likely many of these detectors to allow for efficient processing of the wide variety of semantic features, objects, or categories likely to be encountered, and perhaps other detectors encoding for features that are not specifically semantic, but that may represent intrinsic attributes of certain classes of stimuli, such as motion. Thus, if several of these attributes are present in a sound, then several of these detectors will be activated simultaneously, and the summated and contemporaneous coactivation of the detectors will facilitate object recall. Coactivation of multiple detectors could itself represent the integrated auditory object memory, may elicit the memory by evoking or amplifying a synchronizing cortical gamma rhythm (Kraut, Calhoun, Pitcock, Cusick, & Hart, 2003; Pulvermuller, Lutzenberger, & Preissl, 1999), or might result in the activation of some as-yet unidentified unimodal synthesis region or store (Hart, Moo, Segal, Adkins, & Kraut, 2002). Neither the present study nor those in the literature have delineated such a region.

We predict that as future detectors are identified, many will reside in neighboring regions to the ones presently identified to minimize the temporal delay incurred by transmission of data between brain regions. As a consequence, these detectors that mediate access to the auditory semantic memory system will likely be

organized concentrically surrounding the posterior aspects of the right superior temporal gyrus (Figure 4).

These detectors appear to be coactivated with extra-auditory cortices that are associated with semantic processing of environmental sounds in general (e.g., right inferior frontal gyrus as noted in Lewis et al., 2004), or for specific groups of sound stimuli (threatening, non-threatening, animals). Although the detectors and the regions to which they are connected further along the processing chain can be engaged by the entire set of stimuli in the group, they may be more responsive to one subtype of stimuli in the group than to another. In our data, this is evident in the right superior temporal regions isolated in our primary analyses. The focus that demonstrates greater activity with threatening stimuli also shows greater signal change with animals than nonliving threatening stimuli, whereas the focus that exhibits greater signal change with animal stimuli in the primary analysis also shows greater activity with threatening than with nonthreatening animals. Activity in these early processing stages would presumably result in increased activity in other neural systems dedicated to multimodal semantic processing (e.g., inferior parietal-posterior temporal region that is located posterior to these foci; Beauchamp et al., 2004; Hart & Gordon, 1990) or retrieval of a multimodal semantic conceptual representation (see thalamic synchronization as a mechanism of semantic object recall in the neural hybrid model; Kraut et al., 2003).

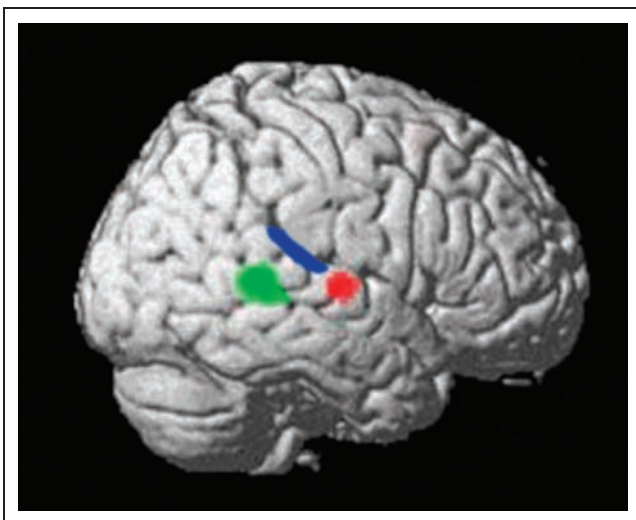
In conclusion, we have shown at least some of the brain regions associated with recognizing a nonverbal

sound and thus accessing semantic memory through the auditory system. The demonstration of both a categorical (animals) and featural (threatening) organization to the nonverbal sound semantic system is analogous to the lexical-semantic and visual semantic memory systems. The locations of at least two of these regions adjacent to the primary auditory cortex suggest that these regions in auditory association cortex likely represent early extraction of semantic content from acoustical/perceptual analysis from the auditory input stream. In the case of the auditory system, the signal changes for the category “animals” are more robust for animals that were threatening; conversely, signal changes associated with the “threatening” feature of sounds were of greater amplitude for threatening animals. This interaction between categorical and featural organization for the same items suggests that a nonverbal sound is encoded in the auditory semantic system via the coactivation of all of the representational nodes (e.g., category, feature) that pertain to the given sound. Clearly, however, higher level processing of auditory stimuli is not limited to dangerous animals, as is evident even within our data by the region in the right superior parietal lobe that appears to be driven by or responds preferentially to sounds that represent threatening, nonliving stimuli. Further studies of multiple other categories, and features, in the nonverbal sound system will likely reveal a more detailed organization of the auditory system for semantic memory. We propose a neural semantic detector functional-anatomic model of auditory semantic memory for sounds that encompasses these and other findings.

Reprint requests should be sent to John Hart, Jr., Center for BrainHealth, UTD-Callier, 1966 Inwood Road, Mail Station CD-A221, Dallas, TX 75235, or via e-mail: jhart@utdallas.edu.

## REFERENCES

- Adams, R. B., & Janata, P. (2002). A comparison of neural circuits underlying auditory and visual object categorization. *Neuroimage*, *16*, 361–377.
- Adolphs, R., Damasio, H., Tranel, D., & Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience*, *16*, 7678–7687.
- Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *Journal of Comparative Neurology*, *276*, 313–342.
- Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*, 809–823.
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., et al. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, *10*, 536–540.
- Benuzzi, F., Meletti, S., Zamboni, G., Calandra-Buonaura, G., Serafini, M., Lui, F., et al. (2004). Impaired fear processing



**Figure 4.** Right lateral projection, showing the signal changes from this study of the right temporal foci for threatening (designated by red) and animal (designated by green) sounds. Also drawn in are the approximate boundaries of primary auditory cortex (blue) from Rademacher et al., 2001. This representation shows the neural semantic memory detectors in proximity to the primary sensory cortex processing perceptual sound input.

- in right mesial temporal sclerosis: A fMRI study. *Brain Research Bulletin*, *63*, 269–281.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119*, 1239–1247.
- Binder, J. R., Rao, S. M., Hammeke, T. A., Frost, J. A., Bandettini, P. A., Jesmanowicz, A., et al. (1995). Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. *Archives of Neurology*, *52*, 593–601.
- Buchel, C., Dolan, R. J., Armony, J. L., & Friston, K. J. (1999). Amygdala–hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 10869–10876.
- Canli, T., Desmond, J. E., Zhao, Z., Glover, G., & Gabrieli, J. D. (1998). Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport*, *9*, 3233–3239.
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain the animate–inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1–34.
- Compton, R. J., Heller, W., Banich, M. T., Palmieri, P. A., & Miller, G. A. (2000). Responding to threat: Hemispheric asymmetries and interhemispheric division of input. *Neuropsychology*, *14*, 254–264.
- Cooley, E. L., Stringer, A. Y., & Hodnett, C. E. (1997). Word stem priming in unilateral stroke patients: Word type and laterality effects. *Archives of Clinical Neuropsychology*, *12*, 71–79.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*, 499–505.
- Davidson, R. J. (1998). Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums. *Psychophysiology*, *35*, 607–614.
- Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, *118*, 279–306.
- Fernandez-Carriba, S., Loeches, A., Morcillo, A., & Hopkins, W. D. (2002). Asymmetry in facial expression of emotions by chimpanzees. *Neuropsychologia*, *40*, 1523–1533.
- Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: A review. *Cortex*, *36*, 539–559.
- Ghashghaei, H. T., & Barbas, H. (2002). Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience*, *115*, 1261–1279.
- Gil-da-Costa, R., Braun, A., Lopes, M., Hauser, M. D., Carson, R. E., Herscovitch, P., et al. (2004). Toward an evolutionary perspective on conceptual representation: Species-specific calls activate visual and affective processing systems in the macaque. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 17516–17521.
- Hart, J., Jr., Berndt, R. S., & Caramazza, A. (1985). Category-specific naming deficit following cerebral infarction. *Nature*, *316*, 439–440.
- Hart, J., Jr., & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology*, *27*, 226–231.
- Hart, J., Jr., & Gordon, B. (1992). Neural subsystems for object knowledge. *Nature*, *359*, 60–64.
- Hart, J., Jr., Moo, L. R., Segal, J. B., Adkins, E., & Kraut, M. A. (2002). Neural substrates of semantics. In A. E. Hillis (Ed.), *The handbook of adult language disorders* (pp. 207–228). New York: Psychology Press.
- Hauser, M. D. (1993). Right hemisphere dominance for the production of facial expression in monkeys. *Science*, *261*, 475–477.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.
- Heilman, K. M., & Gonzalez Rothi, L. J. (1993). Apraxia. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 131–150). New York: University Press.
- Hsieh, J. C., Stone-Elander, S., & Ingvar, M. (1999). Anticipatory coping of pain expressed in the human anterior cingulate cortex: A positron emission tomography study. *Neuroscience Letters*, *262*, 61–64.
- Hugdahl, K., Lundervold, A., Erslund, L., Smievoll, A. I., Sundberg, H., Barndon, R., et al. (1999). Left frontal activation during a semantic categorization task: An fMRI-study. *International Journal of Neuroscience*, *99*, 49–58.
- Iwai, E., & Mishkin, M. (1968). Two visual foci in the temporal lobe of monkeys. In N. Yoshii & N. A. Buchwald (Eds.), *Neurophysiological basis of learning and behavior* (pp. 209–210). Japan: Osaka University Press.
- Jaaskelainen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levanen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 6809–6814.
- Kraut, M. A., Calhoun, V., Pitcock, J. A., Cusick, C., & Hart, J., Jr. (2003). Neural hybrid model of semantic object memory: Implications from event-related timing using fMRI. *Journal of the International Neuropsychological Society*, *9*, 1031–1040.
- Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *Neuroimage*, *21*, 946–953.
- Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., & DeYoe, E. A. (2004). Human brain regions involved in recognizing environmental sounds. *Cerebral Cortex*, *14*, 1008–1021.
- Marcell, M. M., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology*, *22*, 830–864.
- Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.
- Morecraft, R. J., Geula, C., & Mesulam, M. M. (1993). Architecture of connectivity within a cingulo-fronto-parietal neurocognitive network for directed attention. *Archives of Neurology*, *50*, 279–284.
- Mummery, C. J., Patterson, K., Wise, R. J., Vandenberg, R., Price, C. J., & Hodges, J. R. (1999). Disrupted temporal lobe connections in semantic dementia. *Brain*, *122*, 61–73.
- Nielsen, J. M. (1946). *Agnosia, apraxia, aphasia: Their value in cerebral localization*. New York: Paul B. Hoeber.
- Otto, M. W., McNally, R. J., Pollack, M. H., Chen, E., & Rosenbaum, J. F. (1994). Hemispheric laterality and memory bias for threat in anxiety disorders. *Journal of Abnormal Psychology*, *103*, 828–831.

- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.
- Pulvermuller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex*, *9*, 497–506.
- Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H. J., et al. (2001). Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage*, *13*, 669–683.
- Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinion in Neurobiology*, *8*, 516–521.
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *97*, 11800–11806.
- Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, *4*, 108–115.
- Rushworth, M. F., Johansen-Berg, H., Gobel, S. M., & Devlin, J. T. (2003). The left parietal and premotor cortices: Motor attention and selection. *Neuroimage*, *20* (Suppl. 1), S89–S100.
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400–2406.
- Shannon, B. J., & Buckner, R. L. (2004). Functional–anatomic correlates of memory retrieval that suggest nontraditional processing roles for multiple distinct regions within posterior parietal cortex. *Journal of Neuroscience*, *24*, 10084–10092.
- Specht, K., & Reul, J. (2003). Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: An auditory rapid event-related fMRI-task. *Neuroimage*, *20*, 1944–1954.
- Sullivan, R. M., & Gratton, A. (2002). Behavioral effects of excitotoxic lesions of ventral medial prefrontal cortex in the rat are hemisphere-dependent. *Brain Research*, *927*, 69–79.
- Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*, 764–769.
- Tzourio, N., Massiou, F. E., Crivello, F., Joliot, M., Renault, B., & Mazoyer, B. (1997). Functional anatomy of human auditory attention studied with PET. *Neuroimage*, *5*, 63–77.
- Ungerleider, L. G., & Haxby, J. V. (1994). “What” and “where” in the human brain. *Current Opinion in Neurobiology*, *4*, 157–165.
- Wall, P. M., Blanchard, R. J., Yang, M., & Blanchard, D. C. (2004). Differential effects of infralimbic vs. ventromedial orbital PFC lidocaine infusions in CD-1 mice on defensive responding in the mouse defense test battery and rat exposure test. *Brain Research*, *1020*, 73–85.
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*, 829–854.
- Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within “Wernicke’s area.” *Brain*, *124*, 83–95.
- Zald, D. H., Lee, J. T., Fluegel, K. W., & Pardo, J. V. (1998). Aversive gustatory stimulation activates limbic circuits in humans. *Brain*, *121*, 1143–1154.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*, 946–953.