Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study

Colin Humphries, Kimberley Willard, Bradley Buchsbaum and Gregory Hickok[CA]

Department of Cognitive Sciences, University of California, Irvine, CA 92697, USA

[CA]Corresponding Author

Received 7 March 2001; accepted 29 March 2001

Recent neuropsychological and functional imaging evidence has suggested a role for anterior temporal cortex in sentence-level comprehension. We explored this hypothesis using event-related fMRI. Subjects were scanned while they listened to either a sequence of environmental sounds describing an event or a corresponding sentence matched as closely as possible in meaning. Both types of stimuli required subjects to integrate auditory information over time to derive a similar meaning, but differ in the processing mechanisms leading to the integration of that information, with speech input requiring syntactic mechanisms and environmental sounds utilizing non-linguistic mechanisms. Consistent with recent claims, sentences produced greater activation than environmental sounds in anterior superior temporal lobe bilaterally. A similar speech > sound activation pattern was noted also in posterior superior temporal regions in the left. Environmental sounds produced greater activation than sentences in right inferior frontal gyrus. The results provide support for the view that anterior temporal cortex plays an important role in sentence-level comprehension. *NeuroReport* 12:1749–1752 © 2001 Lippincott Williams & Wilkins.

Key words: Anterior temporal lobe; Auditory sentence processing; fMRI; Language; Speech perception

INTRODUCTION

Over the past few decades, Broca's area has frequently been implicated as a site critical for syntactic processing not only in production, but also in comprehension [1–4]. Several lines of evidence, however, have cast doubt on at least the strongest version of this view. First, despite poor comprehension of syntactically complex sentences in Broca's aphasia [2,5,6], such patients are often able to make grammaticality judgments quite accurately, suggesting relatively preserved syntactic processing ability [7]. Second, the agrammatical pattern of sentence comprehension often observed in Broca's aphasia (which is largely responsible for the proposed link between syntax and Broca's area) is not exclusive to patients of that clinical group [8,9]. Third, functional activation studies involving the perception of sentence stimuli do not uniformly report activation in left frontal areas [10]. Finally, even if a connection between syntactic comprehension and Broca's area can be firmly established, it is generally acknowledged that it will involve only a relatively small sub-component of syntax [4] or perhaps the working memory routines which support certain forms of syntactic comprehension. If Broca's area is playing only a minor role in sentence-level comprehension, where are the neural systems which support the bulk of sentence-level comprehension?

One region which has emerged from recent neuropsychological and neuroimaging evidence as a candidate for supporting aspects of sentence-level comprehension is the anterior temporal lobe. For example, in a study of aphasic subjects it was shown that patients with the most severe deficits on a morphosyntactic comprehension task had lesions involving the left anterior superior temporal lobe, whereas those patients who performed well on the task had lesions outside of this region. In that study, having a lesion in Broca's area was not a good predictor of morphosyntactic comprehension difficulty [11]. Corroborating evidence comes from a study of patients with fronto-temporal dementia which found that relative cerebral perfusion in anterior temporal regions (as well as inferior frontal regions) correlated with sentence comprehension difficulty [12].

Several functional imaging studies have found evidence implicating the anterior temporal lobe in sentence processing by comparing activations resulting from sentence perception with various types of control conditions. Two studies for example have contrasted listening to sentences with listening to sentences in an unfamiliar language, a comparison which should subtract out all those activations related to sublexical processing [13,14]. When such a contrast is carried out, areas of residual activation (i.e. native language > unfamiliar language) include anterior superior temporal regions (STS/STG) bilaterally, as well as lateral posterior temporal lobe (STS, MTG predominantly on the left), and Broca's area. Additional studies have contrasted listening to sentences and listening to word lists, the relevant contrast being that sentences, but not word lists, contain syntactic information. Both Mazoyer and colleagues [13] and Friederici *et al.* [10] found that sentences activated anterior superior temporal lobe regions to a greater extent than did word lists. This difference between sentences and word lists held up even when the sentence stimuli were composed of strings in which the content words were replaced with pseudowords (jabberwocky sentences). A similar study comparing sentences and word lists in the visual modality also showed greater activation in anterior superior temporal cortex for sentences than word lists [15], providing converging evidence and suggesting further that the differences observed in the auditory modality cannot be explained in terms of auditorily perceived prosodic differences between sentences and word lists. Thus, the imaging literature corroborates the neuropsychological evidence noted above suggesting a connection between sentence-level processing and anterior temporal lobe regions.

There are, however, other interpretations of the imaging data. Sentences and word lists differ with respect to the semantic content of the utterance, with sentences yielding a coherent integrated meaning and word lists yielding a disjointed sequence of meanings. The fact that the anterior temporal lobe activation for sentences holds up even with jabberwocky sentences argues that semantic factors do not drive the activation in that region, but additional evidence in this respect would be helpful given that it is not clear whether subjects attempt a semantic analysis of jabberwocky sentences. In addition, it is possible that anterior temporal activation simply reflects the requirement in sentences but not word lists, to integrate distinct auditory events (i.e., words) over time. This is a requirement of both the natural and jabberwocky sentences, but not of the word list stimuli.

We sought to address these issues by contrasting sentence comprehension with a control which was matched as closely as possible in semantic content and which required temporal integration. In the experiment reported below, subjects listened to envirornmental sound sequences which described a coherent event (tires squealing followed by a crash). Activations resulting from listening to environmental sound events were contrasted with those resulting from listening to simple content-match sentences (the car skidded and crashed). If anterior temporal lobe activations reported in other studies reflect in some way the semantic content of a sentence or the requirement to integrate a sequence of individual events into a coherent whole, then we should observe no differences between the two conditions in this study. If differences are observed, then this result, when combined with previous observations, would lend further support to the view that anterior temporal lobe activations reflect sentence-level linguistic processing.

MATERIALS AND METHODS

Subjects: Seven subjects (four female, three male; ages 18–28 years) participated in this experiment. Subjects gave informed consent under a protocol approved by the Institutional Review Board of the University of California, Irvine.

Materials: Two conditions were presented. On different trials, subjects heard either a sentence describing an event (e.g. there was a gunshot and then someone ran away) or a series of environmental sounds denoting the same event (e.g. the sound of a gun followed by the sound of footsteps quickly fading into the distance). Sentences were digitally recorded by a male speaker. The environmental sound sequences were made by concatenating together various sounds taken from an environmental sound library. The duration of each sentence and environmental sequence was edited to the same length (3 s) and matched for subjective volume. The stimuli were then presented during fMRI scanning using a Macintosh Powerbook. Subjects listened to the sound stimuli through headphones that were attached to an air conductance sound delivery system.

Procedure: The experiment was organized in an event-related design [16,17]. Three separate runs were presented each containing 16 trials. Each trial lasted 18 seconds with either a sentence or an environmental sound event played during the first 3 s. The conditions were randomly ordered. The first trial in each run began 15 s after scanning was initiated. The corresponding images during this initial time period were later discarded.

Sixteen axial slices were collected using a 1.5 T Siemens Vision scanner. Images were recorded using an EPI sequence (FOV = 256 mm, matrix = 64×64, size = 4×4 mm, TE = 40, flip angle = 90°, thickness = 6 mm) with a TR of 3 s. Anatomical images were acquired in the sagittal plane using an MPRAGE sequence.

To correct for head motion artifact the image volumes of each subject were aligned to the first volume in the series using a 3D rigid body, six parameter model in the AIR 3.0 program [18]. Using the subjects anatomical data, the functional volumes were then warped into a space defined by a standard Talaraich-aligned atlas using a 3D 5th -order polynomial model in the AIR3.0 programs [18]. A Gaussian spatial filter (6 mm FWHM) was applied to each volume. After alignment the timecourse of each voxel was bandpass filtered between 0.01 Hz and 0.1 Hz.

A group analysis was performed on the data for all seven subjects using a regression analysis on each individual voxel. The first explanatory variable included trials during which sentences were presented and the second variable included trials during which environmental sounds were presented. Each explanatory variable was convolved with a standard hemodynamic response function taken from the SPM99 toolbox to account for the hemodynamic response lag. T-statistics were calculated for three conditions: the sentence condition *vs* rest, the environmental sound condition *vs* rest, and the contrast between the sentence and environmental sound conditions. The resulting statistical images were thresholded at a probability level of $p < 0.0001$ (uncorrected). To reduce further the probability of Type I error, voxels which were not part of a cluster of $\geqslant 4$ adjacent voxels were excluded from the probability maps [19].

RESULTS

The sentence condition when contrasted with rest showed significant activation in anterior portions of the temporal lobe (STG, MTG), middle portions of the temporal lobe (STG, MTG, Heschl's gyrus (HG)), and posterior portions of the temporal lobe (STG, STS) in both the left and right hemispheres (Fig. 1). Activations were also observed in the left frontal lobe (Broca's area).

The environmental sound condition when contrasted with rest showed activation in middle (STG, HG), and posterior portions of the temporal lobe (STG, STS) also

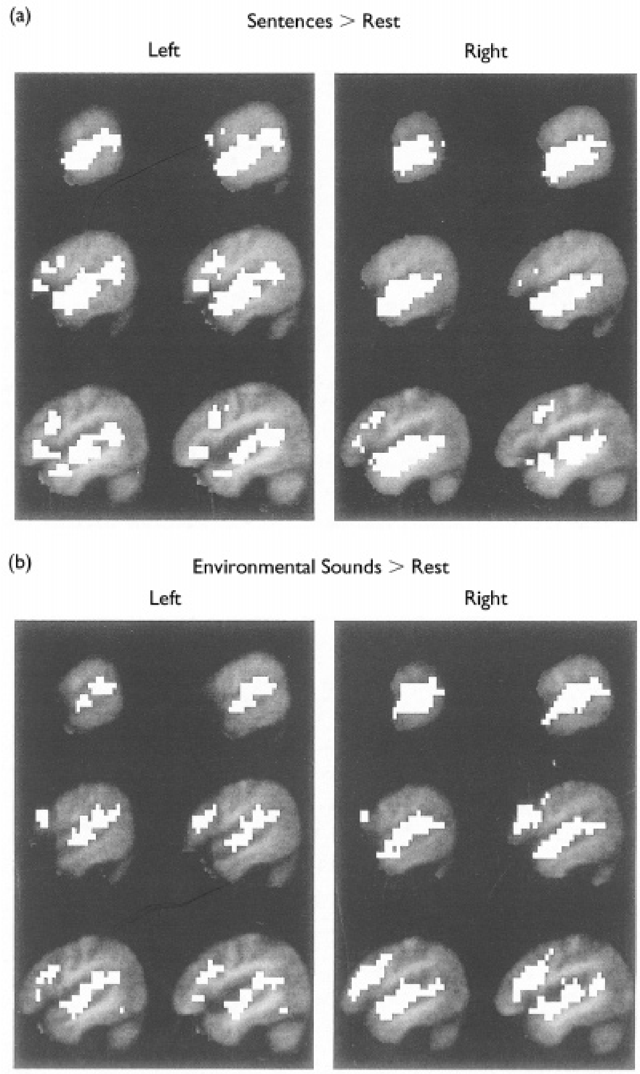

Fig. 1. (**a**) Sagittal slices of the group averaged data showing voxels with greater activation for sentences than rest. The main sites of activation are seen bilaterally along the temporal lobe (posterior and anterior) as well as areas of the frontal lobe. (**b**) Voxels showing greater activation for environmental sounds than rest. Activations are seen bilaterally along the temporal lobe and in frontal areas.

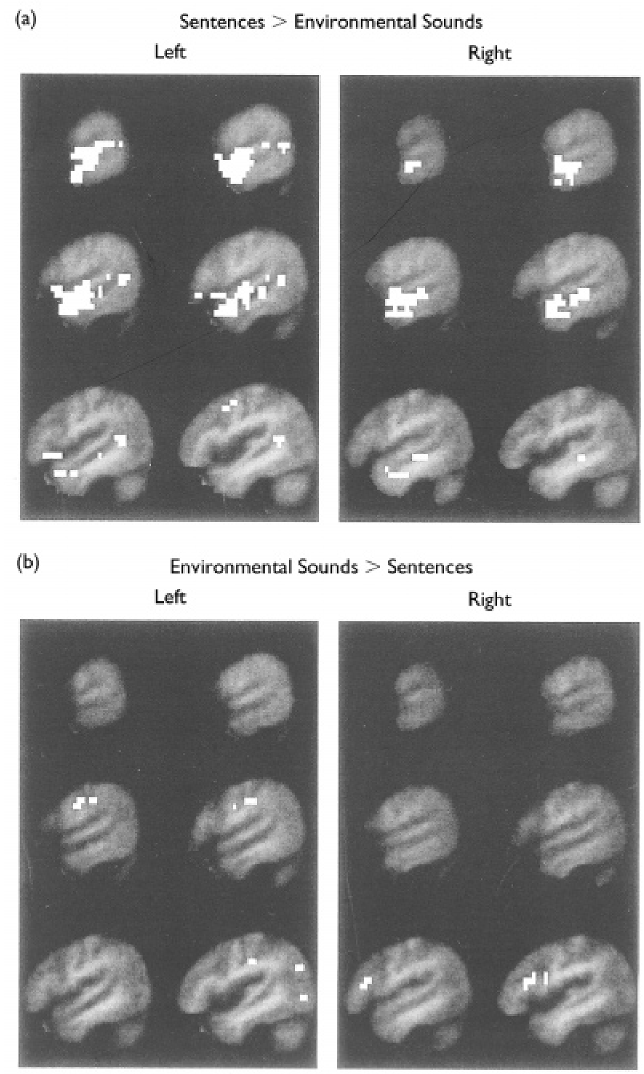

Fig. 2. (**a**) Sagittal slices of the group-averaged data showing voxels with greater activation for sentences than environmental sounds. The main sites of activation are seen in regions of bilateral anterior temporal lobe, as well as in left posterior superior temporal sulcus. (**b**) Sagittal slices of the group-averaged data showing voxels with greater activation for environmental sounds than sentences. The main sites of activation are seen in the right inferior frontal gyrus.

bilaterally (Fig. 1). Activations were also present in left and right inferior frontal gyrus in this contrast.

The contrast between sentences and environmental sounds showed significant activation for sentences over sounds in lateral regions of left and right anterior temporal lobe (STG, MTG; Fig. 2). Activations were also seen more posteriorly in the left lateral temporal lobe including a site in middle portion of the STS lateral to Heschl's gyrus and another in posterior STS . Activations for sounds over sentences appeared in the right inferior frontal gyrus (Fig. 2). Additional patchy right posterior temporal activations were also observed (sounds > sentences), but these did not survive the four cluster thresholding procedure. Importantly, no difference between the two conditions (in either direction) was observed in early auditory cortices (dorsal STG) suggesting that the two classes of stimuli did not differ radically in terms of low-level acoustic properties.

DISCUSSION

The results showed that both the sentence condition and the environmental sound condition activated areas in superior and lateral temporal cortex bilaterally. Both conditions also activated areas in the frontal lobe including left inferior frontal gyrus. Direct contrasts between the two conditions revealed that sentences produced greater activation in superior anterior temporal cortex bilaterally, as well as regions in left posterior superior temporal cortex. environmental sound events, on the other hand produced greater activation in right inferior frontal gyrus. Both classes of stimuli activated early auditory areas to a similar degree suggesting that the differences found downstream in auditory perception are not likely to be due to low-level acoustic differences between sentence and sound items.

A previous study has suggested the existence of voice-specific areas in the superior temporal gyrus [20], that is, regions which respond more strongly to sounds produced by a human vocal tract than to sounds emitted from other sources. Most of our environmental sound stimuli were non-voice sounds, raising the possibility that some of the activation differences between the two conditions can be attributed to voice-specific regions in STG. Indeed, we did find activation, greater for sentences than environmental sounds, in the middle portion of the STG which is a site that was strongly associated with voice-specific activations [20]. However, the anterior STG site identified in the present study has been shown, in previous studies, to activate more for auditorily presented sentences than for non-sentence stimuli even when the non-sentence stimuli did not differ from the sentence stimuli in terms of the presence of vocal tract sounds (foreign language speech, auditory word-lists). Further, anterior STG activation has been found when visually presented sentences were contrasted with visually presented word lists [15], showing that auditory stimulation is not necessary to drive these anterior STG sites. Voice-specific responses, therefore, cannot account for the present observation of greater response in anterior temporal regions for sentences.

The present finding, together with previous neuropsychological and functional imaging data converge on the view that anterior superior temporal cortex is a site which plays an important role in sentence-level comprehension. The activation of this region appears fairly selective for sentence-level stimuli: It does not respond robustly to unstructured meaningful speech stimuli (word lists), or to meaningful sequences of environmental sounds, but it does respond both to meaningful sentences and meaningless pseudoword sentences (jabberwocky). It is an open question whether other hierarchically organized stimuli, such as music will activate anterior temporal cortex. The left inferior frontal gyrus may also play a role in sentence comprehension, but available evidence, including the present finding that Broca's area activation in the sentence minus rest contrast was virtually non-existent in the sentence-environmental sound contrast, suggests a relatively minor contribution

This study also revealed a focus of activation in the left posterior temporal lobe which responded more strongly to sentences than to environmental sound events. Left posterior temporal regions, roughly corresponding to Wernicke's area and surrounding fields, have been suggested to be important for associating sound-based representations of speech with lexical-conceptual representations [21,22]. These regions appear to be particularly important for sound-meaning interface at the word level. Transcortical sensory aphasics who have auditory comprehension difficulties, and posterior temporal-parietal-occipital lesions, have preserved abilities to repeat speech, and sometimes spontaneously correct syntactic errors in the sentences they repeat [23]. This suggests that posterior lesions impair word-level processes more so than sentence-level processes. A study which looked explicitly at word-level semantic deficits found that in the three cases they studied, left posterior temporal regions were damaged [24]. Finally, functional imaging studies of word-level semantic processing typically implicate posterior inferior temporal and posterior parietal cortex [25] sites which, when damaged, have been linked to transcortical sensory aphasia. These observations are consistent with the view that posterior regions near the temporal-parietal-occipital junction are important for sound-meaning interface at the level of words, whereas anterior temporal regions play a greater role in sentence-level processing. Understanding how and where these processing streams interact (which they must) will be an important issue to address in future studies.

CONCLUSION

Our results support the view that anterior temporal lobe plays a role in sentence-level processing. The fact that this region showed greater activation during sentences compared with environmental sounds suggests that this area is not simply involved in temporal integration of meaningful auditory stimuli. Broca's area, a site traditionally associated with sentence processing, did not show a robust difference between the two conditions. In addition, greater activations for sentences over environmental sounds were observed in left posterior temporal lobe, a site which we suggest functions in lexical-semantic processing at the word-level.

REFERENCES

1. Berndt RS and Caramazza A. Syntactic aspects of aphasia. In: Sarno MT, ed. *Acquired Aphasia.* New York: Academic Press; 1981, pp. 157–181.
2. Caramazza A and Zurif EB. *Brain Lang* **3**, 572–582 (1976).
3. Embick D, Marantz A, Miyashita Y *et al. Proc Natl Acad Sci USA* **97**, 6150–6154 (2000).
4. Grodzinsky Y. *Behav Brain Sci* **23**, 1–21 (2000).
5. Grodzinsky Y. *Theoretical Perspectives on Language Deficits.* MIT Press; 1990.
6. Hickok G and Avrutin S. *Brain Lang* **52**, 314–327 (1996).
7. Linebarger MC, Schwartz M and Saffran E. *Cognition* **13**, 361–393 (1983).
8. Goodglass H, Christiansen JA and Gallagher R. *Cortex* **29**, 289–300 (1993).
9. Hickok G. *Behav Brain Sci* **23**, 35–36 (2000).
10. Friederici AD, Meyer M and von Cramon DY. *Brain Lang* **74**, 289–300 (2000).
11. Dronkers NF, Wilkins DP, Van Valin Jr. RD *et al. Brain Lang* **47**, 461–462 (1994).
12. Grossman M, Payer F, Onishi K *et al. Neurology* **50**, 157–163 (1998).
13. Mazoyer BM, Tzourio N, Frak V *et al. J Cogn Neurosci* **5**, 467–479 (1993).
14. Schlosser MJ, Aoyagi N, Fulbright RK *et al. Hum Brain Mapp* **6**, 1–13 (1998).
15. Stowe LA, Paans AMJ, Wijers AA *et al. Psychophysiology* **36**, 786–801 (1999).
16. Buckner RL, Bandettini PA, O'Craven KM *et al. Proc Natl Acad Sci USA* **93**, 14878–4883 (1996).
17. Hickok G, Love T, Swinney D *et al. Brain Lang* **58**, 197–201 (1997).
18. Woods RP, Grafton ST, Watson JDG *et al. J Comput Assist Tomogr* **22**, 155–165 (1998).
19. Forman SD, Cohen JD, Fitzgerald M *et al. Magn.Reson. Med* **33**, 636–647 (1995).
20. Belin P, Zatorre RJ, Lafaille P *et al. Nature* **403**, 309–312 (2000).
21. Hickok G and Poeppel D. *Trends Cogn Sci* **4**, 131–138 (2000).
22. Mesulam MM. *Brain* **121**, 1013–1052 (1998).
23. Berthier M. *Transcortical Aphasias.* East Sussex: Psychology Press; 1999.
24. Hart Jr J and Gordon B. *Ann Neurol* **27**, 226–231 (1990).
25. Binder JB, Frost JA, Hammeke TA *et al. J Neurosci* **17**, 353–362 (1997).

Acknowledgments: This work was supported by NIH grant DC03681.