RAPID COMMUNICATION

Auditory Language Comprehension: An Event-Related fMRI Study on the Processing of Syntactic and Lexical Information

Angela D. Friederici, Martin Meyer, and D. Yves von Cramon

Max Planck Institute of Cognitive Neuroscience, Leipzig, Germany

The functional specificity of different brain areas recruited in auditory language processing was investigated by means of event-related functional magnetic resonance imaging (fMRI) while subjects listened to speech input varying in the presence or absence of semantic and syntactic information. There were two sentence conditions containing syntactic structure, i.e., normal speech (consisting of function and content words), syntactic speech (consisting of function words and pseudowords), and two word-list conditions, i.e., real words and pseudowords. The processing of auditory language, in general, correlates with significant activation in the primary auditory cortices and in adjacent compartments of the superior temporal gyrus bilaterally. Processing of normal speech appeared to have a special status, as no frontal activation was observed in this case but was seen in the three other conditions. This difference may point toward a certain automaticity of the linguistic processes used during normal speech comprehension. When considering the three other conditions, we found that these were correlated with activation in both left and right frontal cortices. An increase of activation in the planum polare bilaterally and in the deep portion of the left frontal operculum was found exclusively when syntactic processes were in focus. Thus, the present data may be taken to suggest an involvement of the left frontal and bilateral temporal cortex when processing syntactic information during comprehension.

Key Words: functional magnetic resonance imaging (fMRI); syntax; language; speech.

INTRODUCTION

It has been known for more than a century that the cerebral organization of language processing is represented predominantly in the left hemisphere.
The network underlying auditory language comprehension and its functional specification, however, is not yet fully described. The processing system responsible for language comprehension must guarantee the analysis of the phonological properties of the input and access to the meaning of words as well as the analysis of the structural relations between words. Electrophysiological approaches have provided convincing evidence that the temporal structures of phonological, semantic, and syntactic processing are different, since these distinct aspects of language comprehension are reflected by unique electrical brain responses (Friederici, 1998). Recent brain imaging studies also suggest a neurotopological difference for phonological and semantic aspects of word processing: The perception of auditorily presented single words activates the primary auditory cortices bilaterally as well as cortical regions in the upper left temporal lobe (Binder, Rao, & Hammeke, 1994; Démonet, Chollet, Ramsay, Cardebat, Nespoulous, Wise, Rascol, & Frackowiak, 1992; Fiez, Raichle, Balota, Tallal, & Petersen, 1996; Price, Wise, Ramsay, Friston, Howard, Patterson, & Frackowiak, 1992). Processing of phonological information was reported to occur in the left frontolateral cortex close to the border between the upper pars opercularis (Broca’s area) and the inferior precentral gyrus (Price, Moore, Humphreys, & Wise, 1997; Zatorre, Evans, Meyer, & Gjedde, 1992), whereas the analysis of word meaning was found to be associated with an increase of metabolic processes in a more widely distributed cortical network, including the pars triangularis of the left and right inferior frontal cortex and the left superior temporal gyrus and involving the banks of the superior temporal sulcus at posterior sites (Démonet et al., 1992; Shaywitz, Pugh, Constable, Shaywitz, Bronen, Fulbright, Shankweiler, Katz, Fletcher, Skudlarski, & Gore, 1995; Warburton, Wise, Price, Weiller, Hadar, Ramsay, & Frackowiak, 1996).

Syntactic processing has been investigated during reading (Stromswold, Caplan, Alpert, & Rauch, 1996; Just, Carpenter, Keller, Eddy, & Thulborn, 1996) as well as during auditory comprehension (Mazoyer, Tzourio, Frak, Syrota, Murayama, Levrier, Salamon, Dehaene, Cohen, & Mehler, 1993; Müller, Rothermel, Behen Muzik, Mangner, & Chugani, 1997; Schlosser, Aoyagi, Fulbright, Gore, & McCarthy, 1998). Studies varying syntactic complexity in sentence reading using positron emission tomography (PET) revealed a selective activation of the left Broca’s area (Stromswold et al., 1996; Caplan, Alpert, & Waters, 1998), and a study using functional magnetic resonance imaging (fMRI) observed a modulation of the activation level by sentence complexity in Broca’s and Wernicke’s areas bilaterally though with a preponderance in the left hemisphere (Just et al., 1996). The comprehension of auditorily presented sentences, however, was found to correlate with either no or only weak activation in Broca’s area but strong responses in the temporal cortex bilaterally in two studies (Mazoyer et al., 1993; Müller et al., 1997). The present study aimed to identify specific cerebral regions mediating the processing of syntactic and semantic information during comprehen-
sion of running speech by systematically varying the absence and presence of structural information and content words in the stimuli. This was realized by means of an event-related design which allows the presentation of the different conditions in a randomized order avoiding the confounds of a blocked design.

**MATERIALS AND METHODS**

Eighteen right-handed subjects with normal hearing (7 males; age range 20–26) were presented with four auditory stimulus types: (1) normal speech, i.e., sentences which are syntactically and semantically correct; (2) syntactic speech, i.e., syntactically correct utterances in which content words are replaced by phonologically legal pseudowords; (3) real-word lists containing no syntactic structure but semantic information; and (4) pseudoword lists containing phonologically legal pseudowords with neither syntactic nor semantic information. Word lists were constructed in the following manner: the number of words matched the number of words in the sentences; the first word was an article in order to keep the first word of each stimulus in an identical category; and the following words in the real-word lists were uninflated adjectives, adverbs, and nouns. Verbs were not used, as they are always inflected in German, and since German, in contrast to English, is a language with free word order, the inclusion of verbs would have counteracted the intended absence of syntactic processes. In the pseudoword-list condition the words following the first word (article) were phonologically legal pseudowords. Note that conditions (1) and (2) may differ from (3) and (4) in their prosodic parameters in addition to syntactic parameters. Given the possible confound between syntactic and prosodic information, a main effect of syntax, i.e., sentences [conditions (1) and (2)] versus word lists [conditions (3) and (4)], would have to be interpreted with caution. An interaction, however, could shed a particular light on syntactic processes under different conditions (e.g., normal speech versus syntactic speech).

All stimuli were spoken by a trained female speaker and were equal in loudness. Mean number of syllables for the sentence conditions was 16.5 and for the word-list conditions was 14.5. The mean maximum of amplitude in its absolute measure Pascal (Pa) was 0.2 Pa (SD = 0.04) for the normal speech condition and 0.23 Pa (SD = 0.04) for the syntactic speech condition. Expressed in the more common measure, decibel (dB SPL), the normal prose condition has a mean of 20 dB (SPL), whereas for the syntactic prose condition a mean of approximately 21 dB (SPL) is quantified. The mean length of the stimuli in the normal speech condition was 3.4 s (SD = 0.36) and in the syntactic speech condition 3.6 s (SD = 0.35). Each subject was presented with a total of 144 stimuli. In an event-related design single trials were presented in a pseudorandomized order to balance sequence effects. Each trial started with a warning tone 1500 ms long prior to the speech input. Single-trial presentation did not exceed 6 s. Successive trials were separated by 8 to 10 s of rest to allow the fMRI signal to return to baseline adequately. Informed consent was obtained from all participants.

To maintain the subjects’ attention while listening to the auditory input, they were required to perform a judgment task after each trial. In order to keep the task constant for the different input conditions, subjects were required to indicate whether the stimulus heard contained a syntactic structure or content words (Table 1).

---

1 Maximum of amplitude was measured by means of the PRAAT software (Institute of Phonetic Sciences, University of Amsterdam; http://www.fon.hum.uva.nl/praat/). Transformation of Pascal into dB (SPL) was done by applying the common formula $20 \times \log_{10}(x/r)$, whereby $x$ means the sound pressure amplitude and $r$ means reference sound, normally the absolute hearing threshold (Johnson, 1997).
TABLE I
Different Stimulus Types Used in the Study

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Semantics+</th>
<th>Semantics−</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Normal prose</td>
<td>(2) Syntactic prose</td>
</tr>
<tr>
<td></td>
<td><em>Die hungrige Katze jagt die flinke Maus.</em></td>
<td><em>Das mumpfge Fööföl föngert das apoldige Trekon.</em></td>
</tr>
<tr>
<td>The hungry cat chased the fast mouse.</td>
<td>The mumpfy folofel fonged the apole trecon.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Real-word lists</td>
<td>(4) Pseudoword lists</td>
</tr>
<tr>
<td></td>
<td><em>Der Koch stumm Kater Geschwindigkeit doch Ehre.</em></td>
<td><em>Der Norp Burch Orlont Kinker Deftey Glauch Legerei.</em></td>
</tr>
<tr>
<td>The cook silent cat velocity yet honor.</td>
<td>The norp burch orlont kinker deftey glauch legerly.</td>
<td></td>
</tr>
</tbody>
</table>

Note. The + indicates the presence of a particular information type; −, the absence of a particular information type. The English translations given are approximations in the case of conditions (2) and (4). Note, that half of the sentences were realized in active mode and half of the sentences were realized in passive mode to increase the variability of experimental stimuli. Due to the free word order in the German language, position of grammatical subject, grammatical object, and verbal phrase was completely balanced across all sentences applied in the experiment.

Imaging

The alignment of structural and functional slices was based on the acquisition of scout spin echo sagittal scans, which defined the anterior and posterior commissures on a midline sagittal section. Eight contiguous horizontal slices with an in-plane resolution of 1.9 × 3.8 mm (FOV = 19.2 cm) were collected at 3.0T using a Bruker 30/100 Medspec (Bruker Medizintechnik GmbH, Ettlingen, Germany) roughly along the bicommissural plane (AC–PC plane) (thickness = 4 mm, gap = 2 mm). A single-shot EPI sequence (matrix = 64 × 64, TE = 40 ms, flip angle of 40°) was used with a TR of 2 s for completing one image. One trial (single stimulus and ISI) lasted 16 s, during which eight gradient echo planar T2* images were obtained.

Data Analysis

The initial four sentences were excluded from analysis to avoid influence of vascular arousal caused by the onset of the scanner noise. For data analysis the BRIAN software package was used (Kruggel & Lohmann, 1996). Preprocessing included movement correction in 2D and a baseline correction by which functional data were filtered in the temporal domain using a low-pass filter. Furthermore, a spatial Gaussian filter (FWHM = 1.34 pixels) was applied to the single data sets (Kruggel, von Cramon, & Descombes, 1999).

Prior to multisubject averaging, functional activation within each individual data set was calculated for each condition separately. Changes in local blood supply were detected by computing a pixelwise Pearson correlation of the MR signal and a 4-s delayed boxcar reference waveform. The correlation coefficient for each single voxel was transformed to Z values (via Fisher’s Z transformation). Individual activation maps were calculated without thresholding (z > 0, p = 1).

High-resolution whole-brain images were acquired in a separate session to assist localization of significant activation foci using a T1-weighted 3D MDEFT for all 18 participants. A trans-
formation matrix was computed that coregisters the 2D anatomical slices with the 3D data set. In the next step this transformation matrix was applied to the 2D activation map. Prior to averaging, the individual registered Z maps were linearly normalized into standard space. Normalized Z maps were averaged by computing pixelwise averages of individual input images. Each voxel in the output image contained the average Z value of the corresponding voxels in the input image. The Z map average was multiplied by a Z-map correction factor of the square root of the current number of subjects (n = 18) (Bosch, 2000).

**Statistical Analysis**

The statistical analysis, called Blob analysis, was conducted as proposed in Bosch (2000). The regions of interest (ROIs) were selected under the constraint that a significant activation (Z value = 3.09) was identified in one of the four conditions. The ROIs were defined as spheres (as in Meyer, Friederici, & von Cramon, 2000). In all cases the spherical ROI contained the local maxima of the summed activation of all conditions. In a second step normalized Z values were averaged within each ROI separately for each task and each participant, allowing the calculation of ANOVAS (Figs. 1 and 2).

**RESULTS**

**Behavioral Data**

The syntactic judgment performance revealed correctness of 97.74% (SE = 1.22) for the normal speech condition, 94.62% (SE = 1.47) for the syntactic speech condition, 95.66% (SE = 2.97) for the real-word lists, and 96.35% (SE = 1.32) for the pseudoword lists. The semantic judgment performance revealed correctness of 96.61% (SE = 1.26) for the normal speech condition, 89.58% (SE = 4.94) for the syntactic speech condition, 96.96% (SE = 1.46) for the real-word lists, and 96.06% (SE = 1.38) for the pseudoword lists.

**Imaging Data**

The fMRI results of multisubject averaging indicate that the pattern of activation varied as a function of the specific type of auditory language input. The cerebral network subserving auditory language processing was activated by all auditory stimulus types, since outstanding hemodynamic responses were observed bilaterally in Heschl’s gyri as well as in adjacent temporal areas, in particular the mid- and posterior STG. Systematic comparisons between the different conditions allow a functional specification of different brain areas associated with sentential processes and those processes involved in single-word comprehension. To obtain comparable Z values from all individuals, we defined two temporal and two frontal regions of interest (ROIs) in distinct regions of the brain: (a) the posterior portion of the STG (Heschl’s gyrus and planum temporale), (b) the anterior portion of the STG (planum polare), (c) the inferior frontal cortex (deep frontal operculum), and (d) the junction of the inferior frontal sulcus with the inferior precentral sulcus. Within all ROIs ANOVAS with factors hemisphere (two levels), syntax (sen-
FIG. 1 Activation as a function of stimulus type is illustrated in a sagittal view of the left hemisphere. The multisubject average \((N = 18)\) revealed significant activation for each condition separately superimposed onto a T1-weighted individual brain. (A) Normal speech; (B) syntactic speech; (C) real-word lists; (D) pseudoword lists. The figure displays voxels exceeding a fixed \(Z\) threshold \((Z \geq 12)\) against a white-matter segmentation. Thus the brain’s white matter is separated from gray matter so that the cortical layers (the outermost 3 to 5 mm) are removed. The particular activation of the deep left frontal operculum in the syntactic speech condition is indicated by a white circle.

2 Temporal regions. For the posterior portion of the STG the ANOVAS revealed a main effect of hemisphere, \(F(1, 17) = 8.72, p < .01\), indicating the superiority of the left hemisphere in auditory language comprehension.

\[^2\text{A global ANOVA with factors hemisphere \times ROI \times syntax \times semantics revealed a significant main effect of hemisphere, } F(1, 17) = 39.14, p < .0001; \text{ of ROI, } F(3, 51) = 24.16, p < .0001; \text{ of syntax, } F(1, 17) = 48.52, p < .0001; \text{ and of semantics, } F(1, 17) = 46.05, p < .0001. \text{ In addition several interactions, hemisphere \times syntax, } F(1, 17) = 40.53, p < .0001; \text{ ROI \times syntax, } F(3, 51) = 8.55, p < .0001; \text{ hemisphere \times semantics, } F(1, 17) = 23.10, p < .0005; \text{ semantics \times syntax, } F(1, 17) = 56.55, p < .0001; \text{ hemisphere \times ROI \times syntax, } F(3, 51) = 10.58, p < .0001; \text{ and hemisphere \times syntax \times semantics, } F(1, 17) = 5.20, p < .05, \text{ point to differential involvement of syntactic and semantic processes in distinct brain areas. Thus, ANOVAs with factors hemisphere \times syntax \times semantics could be performed separately within distinct ROIs. Note that all main effects or interactions with two or} \]
A main effect syntax was also found, $F(1, 17) = 41.85, p < .0001$, indicating that processing sentences elicited more activation as compared to word lists. A hemisphere $\times$ syntax interaction, $F(1, 17) = 10.82, p < .005$, reflects the finding that the left STG is more strongly involved in the two sentence conditions than in the two word list conditions. No significant main effect could be found for the factor semantics. There was, however, a significant syntax $\times$ semantics interaction, $F(1, 17) = 89.44, p < .0001$. This interaction is due to the syntactic speech condition, which elicited maximal activation in the mid- and posterior position of the STG bilaterally.

For the anterior portion of the STG (planum polare) there was also a main effect of syntax, $F(1, 17) = 77.61, p < .0001$, but no main effect of hemisphere or of a hemisphere $\times$ syntax interaction. The main effect of syntax reflects the finding that the planum polare bilaterally was more strongly recruited in the processing of sentences as compared to word lists.

**Frontal regions.** The activation in the deep frontal operculum was entered in an ANOVA with the factors hemisphere, syntax, and semantics, revealing a main effect of syntax, $F(1, 17) = 10.63, p < .01$, and a syntax $\times$ semantics
interaction, $F(1, 17) = 36.45$, $p < .001$. This interaction reflects the finding of a strong activation for the syntactic speech condition, but weak activations in the three other conditions. The outstanding contribution of the left deep frontal operculum to the syntactic speech might suggest that this region subserves particularly syntactic processes. However, the finding that there was no interaction between syntax and hemisphere indicates that the right deep frontal operculum also appears to play a role in processing syntactic speech.

Note that the deep frontal operculum only showed minimal activation for the normal speech condition as compared to syntactic speech. It seems that this area is involved above threshold only when the system is confronted with syntactic information in unusual sentences.

For the cortex at the junction of the inferior frontal sulcus with the inferior precentral sulcus an ANOVA with the factors hemisphere, syntax, and semantics yielded a main effect of syntax, $F(1, 17) = 11.39$, $p < .01$, a main effect of semantics, $F(1, 17) = 22.99$, $p < .001$, a syntax $\times$ semantics interaction, $F(1, 17) = 21.55$, $p < .001$, and a hemisphere $\times$ syntax interaction, $F(1, 17) = 5.48$, $p < .05$. The syntax $\times$ semantics interaction reflects the finding that the junction area bilaterally revealed a weak activation for the normal speech condition, but strong activations in the three other conditions. The hemisphere $\times$ syntax interaction is due to the finding that word lists (syntax absent) showed less activation in the right than the left hemisphere, whereas no such difference was observed for the two sentence conditions (syntax present). These effects seem to suggest that the left frontal junction area supports the processing of language stimuli deviant from normal speech, in general. The right frontal junction, on the other hand, seems to respond more strongly to lexical information in language stimuli deviant from normal speech, be it transferred by content words (as in the real-word list condition) or function words (as in the syntactic prose condition). Note that this frontal area again exposed lowest activation for the normal speech condition.

**DISCUSSION**

Differential activation patterns were found in temporal and frontal areas as a function of processing sentences and word lists. Whereas the different types of auditory language input displayed some localizational overlap, the activation sites and the clusters’ spatial extent clearly varied as a function of the specific type of information being actually processed. The cerebral network subserving auditory sentence processing includes temporal and frontal compartments in the left perisylvian cortex as well as in its right hemisphere homologs. *Auditory processing* occurs necessarily in all four conditions and stimulates Heschl’s gyri bilaterally, the left and the right planum temporale, and lateral segments of the midportion of the temporal gyrus, as observed in all four conditions that were investigated in the current study.
This result is compatible with earlier PET studies (Price et al., 1992; Wise, Chollet, Hadar, Friston, Hoffner, & Frackowiak, 1991) reporting similar activation patterns for auditory word processing.

In the present study no main effect of semantic information operationalized as the presence/absence of content words in the stimuli was observed. However, lexical processes during unusual input conditions (real-word lists and syntactic speech), but not during normal speech, were correlated with an increase in activation in the right cortex lining the junction of the inferior frontal sulcus with the inferior precentral sulcus. The finding that the activation during normal speech was extremely weak in this region may be a reflection of a high automaticity of lexical processes during normal comprehension. The result that this region is activated by content words and by function words (but not to a same degree by pseudowords) receives partial support from a recent fMRI study on the processing of abstract and concrete content words and function words presented out of sentence context (Opitz, Friederici, & von Cramon, 1999; Friederici, Opitz, & von Cramon, 2000). In this study a right inferior frontal activation was found in addition to left-hemisphere activation as a function of concreteness, independent of word class.

It is important to note, however, that some earlier studies identified the left frontal cortex as well as the left inferior temporal cortex to be relevant for semantic processes (Démonet et al., 1992; Warburton et al., 1996; Wise et al., 1991). An engagement of the right inferior frontolateral cortex, however, was reported in conjunction with the processing of semantic meaning (Shaywitz et al., 1995). Based on these findings, it has been proposed that the left inferior frontal cortex is particularly involved when controlled or strategic semantic processes are processed, whereas the right homologue cortex comes into play during semantic memory processes (Nyberg, Cabeza, & Tulving, 1998).

Syntactic processing in the present study is correlated with an increased activation in the left posterior portion of the STG and in the anterior portion of the STG (planum polare) bilaterally, as observed in the two sentence conditions. Note, however, that these conditions also carried sentential prosodic information not present in the word-list conditions. Thus, this activation pattern may partly be due to prosodic information as well. The present data suggest that the deep frontal operculum in the left and possibly in the right hemisphere are part of the cerebral network mediating syntactic processes during auditory language comprehension, as it was significantly more strongly engaged in the syntactic speech condition than in all other conditions. The activation of the frontal operculum in the right hemisphere, not observed to a similar degree in studies of sentence reading, may be related to additional information, e.g., prosodic information, available in the auditorily presented stimulus material compared to visually presented material. A recent finding from an fMRI study which investigated the role of intonation parameters, e.g., accentuation and pitch contour in a sentence-comprehen-
sion paradigm, revealed an increase of the activation in the right frontal operculum when prosody is in focus (Meyer, Alter, Friederici, & von Cramon, 2000). The processing of normal speech, in contrast, did not elicit significant activation in the frontal operculum. This might be explained by the fact that comprehension of normal sentences may occur automatically so that only a minimal amount of local blood supply must be provided. The automaticity argument made for the processing of normal sentences with respect to the activation in the frontal operculum, on the one hand, and with respect to the right frontal activation in the cortex lining the junction of the inferior frontal sulcus and the inferior precentral sulcus, on the other, is based on the finding that the hemodynamic response decreases as familiarity of an item increases (Raichle, Fiez, Videen, MacLeod, Pardo, Fox, & Petersen, 1994).

When considering the present activation pattern, it appears that the frontal regions, but not the temporal regions, vary as a function of the input's deviance from normal speech. First, the left junction area (inferior frontal sulcus and inferior precentral sulcus) showed no significant activation in the normal speech condition but was significantly activated in the three other conditions. Second, the activation in the right junction area was particularly increased when subjects processed lexical information under unusual speech input, i.e., the presence of content words (real-word lists) and the presence of function words (syntactic speech). Third, the activation in the frontal operculum increased as a function of syntactic information under unusual input conditions, i.e., for syntactic speech. These data demonstrate that additional computational resources of frontal regions are required when processing speech stimuli which deviate from normal speech input.

CONCLUSION

The present findings are in general agreement with the view that the amount of activation in left inferior frontal cortices varies as a function of syntactic aspects (Stromswold et al., 1996; Just et al., 1996; Caplan et al., 1998). This view is further elaborated on the identification of the deep frontal operculum, situated nearby the anterior insula, in addition to parts of the superior temporal gyrus, as the relevant brain region in support of syntactic processes.

REFERENCES


