Emotional and Temporal Aspects of Situation Model Processing during Text Comprehension: An Event-Related fMRI Study

Evelyn C. Ferstl¹,², Mike Rinck³, and D. Yves von Cramon¹,²

Abstract

Language comprehension in everyday life requires the continuous integration of prior discourse context and general world knowledge with the current utterance or sentence. In the neurolinguistic literature, these so-called situation model building processes have been ascribed to the prefrontal cortex or to the right hemisphere. In this study, we use whole-head event-related fMRI to directly map the neural correlates of narrative comprehension in context. While being scanned using a spin–echo sequence, 20 participants listened to 32 short stories, half of which contained globally inconsistent information. The inconsistencies concerned either temporal or chronological information or the emotional status of the protagonist. Hearing an inconsistent word elicited activation in the right anterior temporal lobe. The comparison of different information aspects revealed activation in the left precuneus and a bilateral frontoparietal network for chronological information. Emotional information elicited activation in the ventromedial prefrontal cortex and the extended amygdaloid complex. In addition, the integration of inconsistent emotional information engaged the dorsal frontomedial cortex (Brodmann's area 8/9), whereas the integration of inconsistent temporal information required the lateral prefrontal cortex bilaterally. These results indicate that listening to stories can elicit activation reflecting content-specific processes. Furthermore, updating of the situation model is not a unitary process but it also depends on the particular requirements of the text. The right hemisphere contributes to language processing in context, but equally important are the left medial and bilateral prefrontal cortices.

INTRODUCTION

Everyday language serves a communicative purpose. Understanding the meanings of words and the structure of sentences is an important prerequisite for a correct interpretation. However, pragmatically appropriate comprehension requires the reader or listener to go beyond the word and sentence onto a contextual level.

Psycholinguistic theories assume that the comprehension process involves the formation of representations on different levels. Beyond the verbatim form of the text information, a so-called text base is formed that encodes the semantic meaning of the text. In the text base, the text information is supplemented by additional propositions from general world knowledge needed for establishing local coherence, that is, for sensibly linking successive utterances. These so-called bridging inferences are considered mandatory and automatic. In addition, comprehenders engage in a variety of other inference processes, for instance, elaborative, predictive, or causal inferences (Graesser, Singer, & Trabasso, 1994; Singer, 1994).

On a level more removed from the verbatim input, comprehenders form a global representation or a mental model of text or discourse information. This so-called situation model integrates the current language input with both general world knowledge and the prior discourse context (Kintsch, 1998; Zwaan & Radvansky, 1998; Graesser, Millis, & Zwaan, 1997; van Dijk & Kintsch, 1983). In contrast to bridging inferences, elaborative and predictive inferences are considered part of the situation model. The situation model is retained longer than the verbatim form of the text, and it is used for application of and learning from the text information (Ferstl, 2001; Ferstl & Kintsch, 1999). The situation model of narrative texts has been shown to contain information about several important dimensions of the story, including the where, when, who, what, and why of the events (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995). Its representation is not necessarily verbal or propositional, but flexibly tailored to the specific text contents (e.g., Zwaan & Yaxley, 2003; Glenberg & Kaschak, 2002; Stanfield & Zwaan, 2001).

Inference processes have been studied extensively in patients with brain damage. There is overwhelming

¹ Max-Planck-Institute of Human Cognitive and Brain Sciences, Leipzig, ² Day Clinic of Cognitive & Brain Sciences, University of Leipzig, ³ Dresden University of Technology

© 2005 Massachusetts Institute of Technology
evidence for the right hemisphere to play a crucial role (Beeman, 1998; Brownell, Gardner, Prather, & Martino, 1995; but see Lehmann & Tompkins, 2000; McDonald, 1993, for critical reviews). In addition, there is evidence for inference deficits in patients with left prefrontal lesions (Ferstl, Guthke, & von Cramon, 2002; Zalla, Phipps, & Grafman, 2002). Less is known about the neuropsychology of situation model processing. For instance, Hough (1990) showed that patients with right-sided lesions had difficulty with story comprehension when they needed to infer the theme of the situation. When a theme sentence in the beginning provided an anchor for setting up an appropriate situation model, their performance improved. Zalla et al. (2002) showed that patients with prefrontal lesions had problems with both on-line inference processes and extracting the sequential order of story events, a process that reflects situation model building.

Neuroimaging studies provide ample evidence for a left-dominant frontotemporal network involved in language processing at the word and sentence level (see Gernsbacher & Kaschak, 2003; Bookheimer, 2002, for recent reviews). In contrast, relatively few descriptions of the functional neuroanatomy of language processing at the text and discourse level have been attempted. The activation patterns differ depending on the stimuli and the task, but there seems to be no doubt about the candidate components of the comprehension network. In addition to the perisylvian cortex of the dominant hemisphere, there is a contribution of the anterior temporal lobes (aTLs; Crinion, Lamber-Ralph, Warburton, Howard, & Wise, 2003; Ferstl & von Cramon, 2001; Maguire, Frith, & Morris, 1999; Mazoyer et al., 1993), the frontomedial and parietomedial/posterior cingulate cortices (Ferstl & von Cramon, 2001, 2002; Maguire et al., 1999; Fletcher et al., 1995), and of course the right hemispheric frontal and/or temporal areas (Caplan & Dapretto, 2001; Robertson et al., 2000; St. George, Kutas, Martinez, & Sereno, 1999).

When inspecting these imaging studies, it becomes apparent that an attribution of the various activations to specific subprocesses postulated in psycholinguistics is difficult at best. In particular, in most imaging studies, local coherence building (i.e., inference processes) and situation model building are inseparable. This is the case when connected text is compared with unrelated sentences (e.g., Fletcher et al., 1995; Mazoyer et al., 1993) or with a nonlanguage baseline (Crinion et al., 2003). More controlled studies manipulating the situation model by providing a theme or a title for a passage (Maguire et al., 1999; St. George et al., 1999) also introduce a concurrent difference in local coherence: Without the title, the easily comprehensible passages become a list of loosely connected sentences. Similarly, by varying the definiteness of articles, Robertson et al. (2000) manipulated the lexical cohesion of texts, thereby affecting both local coherence and the resulting situation model. Even in studies specifically targeting local coherence processes using short two-sentence texts (Mason & Just, 2004; Caplan & Dapretto, 2001; Ferstl & von Cramon, 2001, 2002), the situation model representation clearly varies with pragmatic coherence. Thus, even in the case of minimal texts, local coherence and situation model processes are confounded.

The goal of the present study was to investigate the functional neuroanatomy of situation model building and updating in the context of locally coherent texts. We used the inconsistency paradigm (Rinck, Hähnel, & Becker, 2001) for an event-related functional magnetic resonance imaging (fMRI) study at 3 T. Short stories were presented that contained information contradicting the global text content. Direct comparison with consistent versions of the same stories enabled us to specifically study the effects of processing inconsistent information. At the same time, both word- and sentence-level features as well as local coherence were kept constant.

Behavioral research has shown that globally inconsistent information induces two separable processes. First, readers of narratives exhibit increased reading times of the inconsistent word(s), indicating on-line detection of the situation model violation (Rinck et al., 2001). Second, the detection is quickly followed by a temporally extended repair process. Comprehenders attempt to salvage a plausible situation model by integrating the inconsistent information into the ongoing representation. Readers employ a variety of integration processes (Rinck et al., 2001; Otero & Kintsch, 1992); for instance, they discard the previous context information (“Oh, I might have misunderstood”) or infer a potential explanation reconciling even explicit, mutually exclusive pieces of information (“Maybe first it goes up and then it goes down again”).

In addition to varying consistency, we used different story types for inducing qualitatively different integration processes. Sample stories for the two types are presented in Table 1. The target information, that is, the word or words at which the consistency or inconsistency can be first evaluated, are printed in bold face.

From the various aspects represented in narrative situation models (cf. Zwaan, Langston, et al., 1995), two particularly salient ones were selected: Violations of temporal or chronological aspects of the plot (Rinck, Gámez, Díaz, & de Vega, 2003; Rinck et al., 2001) were compared with violations concerning the affective status of the protagonists (cf. Gernsbacher, Goldsmith, & Robertson, 1992). These aspects were realized in different texts because incorporating both aspects into the same one renders the resulting stories very artificial (Rinck & Weber, 2003).

A comparison of the two different story types was expected to shed light on the information-specific, possibly nonpropositional, representation of the situa-
Today, Markus and Claudia would finally meet again. Markus’s train arrived at the station 20 minutes after / before Claudia’s train. Markus was very excited when his train stopped at the station on time. He tried to think of what he should say when he met her. Many people were crowding on the platform. Claudia was already waiting for him when he got off the train with his huge bag. They were both very happy.

The semester was finally over and Sarah wanted to celebrate. A lot of her friends had shown up for her end-of-school party. It was one of these parties with everything being just perfect. Sarah’s best friend gave her a hug and told her how much fun she was having. Sarah couldn’t remember that she had ever been so happy / sad before. She put her favorite record into the CD player and started dancing by herself.

### Table 1. Sample Stories for the Two Story Types

<table>
<thead>
<tr>
<th></th>
<th><strong>Chronological</strong></th>
<th><strong>Emotional</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>The words in bold face constitute the target information. The words in italics provide the inconsistent versions. Each participant listened to only one version of each story.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The integration of the inconsistencies was also expected to differ for the two story types. The inconsistencies in the chronological stories were all or none. For instance, either Markus’s train arrived before Claudia’s or the other way around. Moreover, the relevant context information is mentioned explicitly at the beginning of the story. In contrast, in the emotional stories, the relevant context information is implied by several pieces of information and the emotional inconsistencies lie on a fuzzy continuum. Ambivalent feelings are perfectly possible. In the example describing the cheerful party, for instance, the student may be sad because her boyfriend is absent or because she will not see her college friends during the summer. Thus, integration in the chronological stories is expected to involve memory reaccess of the previous discourse information and a comparison procedure, probably reflected in prefrontal activation; integration in the emotional stories will more likely consist of an explanatory inference, which is expected to engage right frontotemporal regions (Mason & Just, 2004; Beeman, 1998) or the dorsomedial prefrontal cortex (dmPFC; Zysset, Huber, Samson, Ferstl, & von Cramon, 2003; Zysset, Huber, Samson, Ferstl, & von Cramon, 2002; Ferstl & von Cramon, 2001, 2002). In contrast, the processing of emotional information is known to engage medial prefrontal, orbitofrontal, and limbic structures, including the amygdalae (Dolan, 2005). Whether listening to stories suffices for eliciting activation in either of these networks was an open question.

To reduce susceptibility artifacts in the orbitofrontal and ventromedial prefrontal cortices—regions likely to be involved during emotion processing (Dolan, 2002; Luan Phan, Wager, Taylor, & Liberzon, 2002)—a spinecho echo-planar imaging (EPI) sequence was employed (Zysset, Huber, Samson, et al., 2003; Norris, Zysset, Mildner, & Wiggins, 2002). Although this sequence suffers from a considerable reduction in signal-to-noise ratio, the possibility of uncovering qualitative differences between the two story types was important.

To separate the two processes, detection and integration, changes in the BOLD contrast were analyzed based on two distinct periods. First, activations elicited by local processing of the consistent or inconsistent target information were considered indicative of on-line comprehension of the current input and possibly detection of inconsistencies. Second, the activations observed during the entire period from presentation of the target information up to the end of the story were considered to reflect global integration of the target information into the final situation model.

### RESULTS

#### Behavioral Data

The reaction times from one participant were not recorded because of a technical error. One participant did not provide any correct responses in one of the four conditions. Thus, the response time analyses are based on the data from the remaining 18 participants. The behavioral data were analyzed in a 2 × 2 ANOVA with the within-subjects factors Consistency and Story Type. The means of the error rates and the response times are displayed in Table 2. For the error rates, there was a highly significant main effect of consistency, \( F(1,19) = 32.6, p < .0001 \); it was more difficult to detect inconsistencies (\( M = 64\% \) correct) than to accept consistent stories (\( M = 85\% \) correct). This effect was slightly smaller for the chronological stories, but the interaction between Consistency and Story Type did not reach significance (\( p > .11 \)).
The response times reflect off-line processing because they were measured from the end of the story presentation. Nevertheless, there was an interaction between Story Type and Consistency, $F(1,17) = 6.0; p < .05$. When the target information concerned a chronological aspect, the inconsistency was indicated more quickly. When the target information concerned an emotional aspect, the responses were faster after consistent stories. No other effects reached significance. There was no systematic relationship between speed and accuracy ($r = -.1, ns, n = 18$).

**fMRI Data**

The fMRI data were analyzed in two time windows (see Methods for details). The first analysis used the offset of the target word as a distinct event (event related), and the second analysis used the period from the target word until the end of the story as a short epoch (epoch related).

**Event-Related Analysis**

The first analysis investigated the brain activation caused by encountering the target word. The results for the $2 \times 2$ analysis using the factors Story Type and Consistency are presented in Table 3 and Figure 1.

The main effect of Story Type compared the chronological and emotional target words independent of their Consistency. Emotional information elicited more activation than chronological information in the posterior ventromedial prefrontal cortex (vmPFC) bilaterally and the left-sided extended amygdaloid complex. The time course diagram, included in Figure 1, shows the percent

### Table 2. Behavioral Data

<table>
<thead>
<tr>
<th></th>
<th>Temporal Stories</th>
<th>Emotional Stories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Consistent</td>
<td>Inconsistent</td>
</tr>
<tr>
<td><strong>Errors (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>16.0</td>
<td>25.9</td>
</tr>
<tr>
<td><strong>Reaction times (msec)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>990</td>
<td>799</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>430</td>
<td>267</td>
</tr>
</tbody>
</table>

The response times reflect off-line processing because they were measured from the end of the story presentation. Nevertheless, there was an interaction between Story Type and Consistency, $F(1,17) = 6.0; p < .05$. When the target information concerned a chronological aspect, the inconsistency was indicated more quickly. When the target information concerned an emotional aspect, the responses were faster after consistent stories. No other effects reached significance. There was no systematic relationship between speed and accuracy ($r = -.1, ns, n = 18$).

### Table 3. Regions of Activation for the Event-Related Analysis Time Locked at the Target Information

<table>
<thead>
<tr>
<th>Event-Related Inconsistency Detection</th>
<th>Area</th>
<th>Size</th>
<th>Max</th>
<th>Side</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional &gt; chronological</td>
<td>Ventromedial prefrontal cortex/supraorbital sulcus</td>
<td>945</td>
<td>4.19</td>
<td>R</td>
<td>2</td>
<td>20</td>
<td>−12</td>
<td>15.9**</td>
</tr>
<tr>
<td></td>
<td>Extended amygdala</td>
<td>371</td>
<td>3.94</td>
<td>L</td>
<td>−22</td>
<td>2</td>
<td>−12</td>
<td>11.8*</td>
</tr>
<tr>
<td></td>
<td>Pons</td>
<td>823</td>
<td>4.65</td>
<td>L</td>
<td>−8</td>
<td>−25</td>
<td>−27</td>
<td>20.9**</td>
</tr>
<tr>
<td>Chronological &gt; emotional</td>
<td>Precuneus</td>
<td>2,927</td>
<td>4.81</td>
<td>L</td>
<td>−8</td>
<td>−61</td>
<td>44</td>
<td>21.9**</td>
</tr>
<tr>
<td></td>
<td>Precuneus</td>
<td>354</td>
<td>4.01</td>
<td>R</td>
<td>4</td>
<td>−49</td>
<td>41</td>
<td>27.5***</td>
</tr>
<tr>
<td></td>
<td>Intraparietal sulcus</td>
<td>474</td>
<td>3.72</td>
<td>L</td>
<td>−40</td>
<td>−55</td>
<td>50</td>
<td>21.4**</td>
</tr>
<tr>
<td></td>
<td>Intraparietal sulcus/ temporoparietal junction area</td>
<td>1,804</td>
<td>4.11/3.93</td>
<td>R/R</td>
<td>35/43</td>
<td>−55/−53</td>
<td>44/26</td>
<td>31.7***</td>
</tr>
<tr>
<td></td>
<td>Parieto-occipital</td>
<td>790</td>
<td>4.10</td>
<td>L</td>
<td>−34</td>
<td>−82</td>
<td>35</td>
<td>17.5**</td>
</tr>
<tr>
<td></td>
<td>Frontopolar</td>
<td>1,677</td>
<td>4.22</td>
<td>L</td>
<td>−23</td>
<td>47</td>
<td>12</td>
<td>32.8***</td>
</tr>
<tr>
<td></td>
<td>Frontopolar</td>
<td>1,179</td>
<td>4.09</td>
<td>R</td>
<td>23</td>
<td>41</td>
<td>15</td>
<td>26.5***</td>
</tr>
<tr>
<td></td>
<td>Anterior superior insula</td>
<td>636</td>
<td>3.89</td>
<td>L</td>
<td>−29</td>
<td>14</td>
<td>15</td>
<td>17.6**</td>
</tr>
<tr>
<td>Inconsistent &gt; consistent</td>
<td>Anterior temporal/anterior inferior insula/posterior orbital gyrus</td>
<td>523</td>
<td>3.79</td>
<td>R</td>
<td>40</td>
<td>−1</td>
<td>−12</td>
<td>18.2**</td>
</tr>
</tbody>
</table>

*p < .01; **p < .001; ***p < .0001

$Z > 3.09; p < .001$, uncorrected; extent threshold, $324 \text{mm}^3$; cluster level, $p < .05$, uncorrected. Shown are the Talairach coordinates and the $Z$ value for the region’s peak as well as its extent. The interaction contrast did not yield significant regions of activation. The rightmost column shows the $F$ values [$df(1,19)$] for the corresponding effect from the second-level ROI analysis (Bosch, 2000).
signal change as a function of the four conditions, time locked at the target word. As can be seen, the activation is due to a signal increase primarily for the inconsistent emotional condition, whereas for both chronological conditions the signal decreases. The reverse comparison showed a frontoparietal network to be engaged during the processing of chronological target information. In particular, chronological target information elicited activation in bilateral regions in the anterior portion of the middle frontal gyrus, superior parietal, and, most prominently, the dorsal part of the precuneus. For the latter region, Figure 1 also includes a time course diagram. The diagram shows a signal increase for both chronological conditions starting with the presentation of the target information. All activations are superimposed on a white-matter rendition of the participants’ average anatomy. For descriptive details see Table 3. Areas in red/yellow were more active during the processing of emotional information; the areas in green were more active during the processing of chronological information. Two time course diagrams are added, displaying the percent signal change time locked at the presentation of the target information. The activation in the amygdaloid complex peaked at 6 sec and was mostly due to the inconsistent emotional stories. The activation in the dorsal precuneus increased more gradually and persisted throughout the remainder of both chronological stories, independent of their consistency.

Figure 1. fMRI results for the main effect of story type in the event-related analysis, time locked at the target information. All activations are superimposed on a white-matter rendition of the participants’ average anatomy. For descriptive details see Table 3. Areas in red/yellow were more active during the processing of emotional information; the areas in green were more active during the processing of chronological information. Two time course diagrams are added, displaying the percent signal change time locked at the presentation of the target information. The activation in the amygdaloid complex peaked at 6 sec and was mostly due to the inconsistent emotional stories. The activation in the dorsal precuneus increased more gradually and persisted throughout the remainder of both chronological stories, independent of their consistency.

Figure 2. fMRI results for the main effect of consistency in the event-related analysis, time locked at the target information (cf. Table 3). (I) The area in green was more active when an inconsistency was heard. The temporal lobe is cut at x = 43. (II) Parametric second-level analysis of the consistency effect as a function of the participants’ performance. A scatter plot displays the difference in percent signal change (inconsistent – consistent) as a function of the proportion of correct responses. Participants with lower performance engage the dmPFC more when processing inconsistent trials, whereas the reverse is true for participants with higher performance.
information and lasting about 13 sec. During the same period, the curve levels off for both emotional conditions.

Consistent stories did not activate any of the regions more than inconsistent stories. In the reverse comparison, there was an activation in the right temporal lobe, connected with a small activation in the neighboring frontal operculum and anterior insula.

The interaction contrast did not reveal any regions in which story type varied with the consistency of the story.

Further Consistency Analyses

Because of the variability in the performance data, we could not be sure whether the observed effects of consistency properly reflected processing difficulties. A restriction of the analyses on the correct trials was not possible because of the small number of stories presented. For further evaluation, two parametric analyses were carried out. The first analysis focused on the variability due to possible differences between the stories, whereas the second took into account the variability due to individual differences between participants.

For the item-based parametric analysis the consistency was recoded in a graded way. Instead of the predefined categories, the frequency of yes responses across all 20 participants was taken as the degree of consistency. This measure varies from 0 (definitely inconsistent) to 1 (all participants agree on the consistency) and yields continuous frequency values. Using this measure in a parametric analysis, we confirmed the previous main effect in the right aTL. The same region proved to be significantly related to the consistency of the story, and its extent was even larger ($Z_{\max} = 4.95$, peak coordinates $[40, 5, -12]$, extent $1,290 \text{ mm}^3$).

For evaluating the between-subject variability, a second-level analysis was conducted. The participant’s overall performance, in percent correct, was entered as a covariate. The effect of interest was the interaction between consistency and the overall performance measure. We hypothesized that the consistency effect would be more pronounced in participants who detected most of the inconsistencies accurately. However, the activation in the right aTL did not vary with performance level. Instead, as shown in Figure 1, activation in the dmPFC (Brodmann’s area [BA] 8) varied significantly with both the consistency of the story and the participant’s performance (peak $[-8, 36, 42]$, $Z = 4.5$; extent $= 3,656 \text{ mm}^3$).

Taking the mean percent signal change in the window of 3–9 sec after the target event as the dependent measure, a highly significant negative correlation between the consistency effect—defined as the difference between inconsistent and consistent stories—and the subjects’ performance (in percent correct) results ($r = -0.80$). Even when an extreme value is removed, the correlation is still $r = -0.69$, $p < .001$. The correlation is displayed in a scatter plot in Figure 2, showing the difference in percent signal change between inconsistent and consistent trials, averaged in a window of 3–9 sec after target presentation, as a function of performance.

Epoch-Related Analysis

A further analysis was used for evaluating the effects of integrating the target information into the unfolding story context. For each story, this analysis identified activation present during the interval from the target information until the end of the story. In particular, we took 1 sec before the target offset as the beginning of the epoch, and the presentation of the response beep as the end (i.e., 1 sec after the end of the story). During this period, the participants heard the crucial target information; they integrated it into their accumulating text representation, further checked the plausibility of the remainder of the story, and they prepared their decision. The difference from the previous event-related analysis is mainly the length of the crucial target. In the event-related analysis a discrete point in time was defined, whereas in the epoch-related analysis a short period was included.

The average length of the epochs was 8.4 sec. The target information in both consistent and inconsistent versions was the same, so that across the four lists the epoch lengths were equal. The epoch length was also matched between chronological and emotional stories ($T = 8.53 \text{ sec}$, $E = 8.37 \text{ sec}$). As in the event-related analysis, a factorial design with the factors Consistency and Story Type was conducted. The results are displayed in Figure 3 and Table 4.

There were no significant areas of activation for the emotional stories compared with the chronological stories. The reverse comparison yielded bilateral activation of the precuneus in regions overlapping with those found in the event-related analysis.

Inconsistent stories elicited prefrontal activation bilaterally, which was slightly more extended on the left. This activation was located in the anterior and lateral orbital gyrus bilaterally and, on the left side, encompassed the orbital part of the inferior frontal gyrus (IFG; BA 47/11). These activations were more ventral and posterior than the frontopolar regions activated by the chronological target information in the event-related analysis. There were no significant areas of activation for consistent as compared with inconsistent stories.

The interaction contrast (comparing the conditions chronological/consistent and emotional/inconsistent with chronological/inconsistent and emotional/consistent) showed that the dmPFC of the left hemisphere was differentially sensitive to consistency for the two story types. For an evaluation of this interaction, the contrasts coding the consistency of the stories were calculated separately for the emotional stories and for the chronological stories. For the chronological stories, integrating an inconsistency required the contribution of the prefrontal areas also seen in the overall consistency con-
In contrast. The time course diagram for the right-sided region is shown in Figure 3. It can be seen that both inconsistent conditions elicit an early increase in signal change, but that this effect is more pronounced for the chronological stories in a later time window. Nevertheless, neither the interaction contrast nor the simple main effect for emotional stories yielded a significant result in the bilateral prefrontal regions.

In contrast, inconsistent emotional information elicited activation in the dmPFC. This activation included three peaks, reaching from the most posterior pre-SMA (BA 6/32) and the BA 8 region also found in the event-related analysis to the dorsomedial part of BA 9. The time course diagram plotting the percent signal change time locked at the target information shows that for the inconsistent emotional condition, the activation in BA 9
remains on a high level up to 14 sec after hearing the target word. In contrast, the other three conditions level off at about 8–10 sec after target presentation. In addition, the upper portion of the right IFG, close to the junction of the inferior frontal sulcus and the pre-central sulcus, was engaged during the integration of emotional inconsistencies.

In neither of the comparisons did we find regions more strongly activated for the consistent stories than for the inconsistent stories.

The second-level analysis, taking into account the individuals’ performance level, did not yield any significant regions of activation.

DISCUSSION

In this study we investigated language processing in context, and in particular the on-line updating of the situation model (Rinck et al., 2001; van Dijk & Kintsch, 1983). Using a factorial design, we independently varied the consistency of global target information and the aspect of this information. The first analysis evaluated the immediate effect of encountering and detecting a consistent or inconsistent target word.

The local detection of the inconsistencies was related to the activation of a right anterior temporal region. This right-hemispheric activation seems to be in line with neurolinguistic theories of discourse comprehension attributing global text level processes to the right hemisphere (Robertson et al., 2000; St. George et al., 1999; Brownell et al., 1995; Beeman, 1993, 1998). However, previous imaging results supporting the right-hemisphere hypothesis, both in comprehension and production, have found right-hemisphere involvement in prefrontal (Robertson et al., 2000), and posterior temporal areas (Nathaniel-James & Frith, 2002; Kircher, Brammer, Andreu, Williams, & McGuire, 2001; St. George et al., 1999), rather than in the anteriormost temporal region active in the present study.

The contribution of the aTL to language processing is undisputed. Many studies on sentence and text-level comprehension have shown bilateral involvement of the aTL (Crinion et al., 2003; Ferstl & von Cramon, 2001, 2002; Humphries, Willard, Buchsbaum, & Hickok, 2001; Bavelier et al., 1997; Mazoyer et al., 1993). Despite a number of hypotheses (semantics: Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; auditory meaning integration: Humphries et al., 2001), its specific function has not yet been agreed upon. For the right aTL, few comparably specific proposals are available, but it has been suggested that the right-hemispheric homologues support left-hemispheric language areas when processing difficulties arise (Meyer, Friederici, & von Cramon, 2000). We believe the left aTL to be responsible for propositionalization (cf. Kintsch, 1998), that is, for forming a meaning-based representation independent of the particular wording (Long & Baynes, 2002; Humphries et al., 2001). When encountering an inconsistency,
Ventromedial Prefrontal Cortex

In the main effect of story type, the emotional target information directly engaged two regions in the vmPFC, along the supraorbital sulcus bilaterally and in the left extended amygdaloid complex. These regions are known to be related to emotion processing (Kringelbach & Rolls, 2004; Luan Phan et al., 2002; Davidson & Irwin, 1999). More specifically, the vmPFC has been implicated to have a role in decision making (Bechara, Tranel, & Damasio, 2000; Damasio, 1995) or during interpersonal interactions (Davidson, Putnam, & Larson, 2000; Anderson, Bechara, Damasio, Tranel, & Damasio, 1999). Consequently, lesions in the vmPFC may cause a lack of empathy (Damasio, 1995). In the present experiment, the emotional information presented in the stories always concerned the protagonist’s feelings. Thus, the evaluation of this information involved empathy with the story characters, a process consistent with the vmPFC activation (cf. Decety & Chaminade, 2003).

Moreover, the results presented here are novel in two important respects. First, the present study is one of the first to apply spin–echo EPI at 3 T to avoid susceptibility artifacts previously preventing the use of high-field-strength fMRI for investigating these ventral cortical areas (Zysset, Huber, Samson, et al., 2003; Norris et al., 2002). Most importantly, the study is one of the very few that were successful in eliciting emotion-related activation using a simple language-comprehension paradigm (Maratos, Dolan, Morris, Henson, & Rugg, 2001; Tabert et al., 2001; Elliott, Rubinsztein, Sahakian, & Dolan, 2000; Isenberg et al., 1999; Whalen et al., 1998; Maddock & Buonocore, 1997). Most studies targeting emotion processing use emotion elicitation procedures and/or visual materials such as pictures or faces. To obtain measurable emotional reactions, it is usual to employ highly arousing materials, such as pictures of war scenes or highly offensive or threatening verbal material. In our study, comprehending simple, everyday stories probably did not elicit the described emotions in the comprehenders, but it was sufficient for eliciting activation in regions known to be related to emotion processing, including the extended amygdaloid complex. Consistent with this finding, Glenberg, Havas, Becker, and Rinck (2005) have shown behaviorally that the reader’s current emotional state affects the processing of verbal statements that imply emotional reactions. Thus, merely listening to fictional stories engages the recipient’s affective system. This conclusion is in obvious agreement with everyday experience: our emotional involvement is one of the main reasons why we enjoy reading narrative texts.

The Frontoparietal Network

The reverse comparison of the main effect of story type yielded a bilateral network of frontoparietal brain regions. Listening to the chronological target information elicited activation in the dorsal part of the anterior precuneus, in the intraparietal sulci and, in the polar aspect of the middle frontal gyri. Clearly, this network must be interpreted as reflecting working memory and attentional processes. In particular, activation of the dorsal portion of the precuneus is related to a switch from local to global stimulus aspects (Brass, Zysset, & von Cramon, 2001; Fink et al., 1997). In contrast, frontopolar regions at the anterior end of the intermediate frontal sulci, in concert with the inferior parietal lobules, have been shown to implement verbal working memory tasks under articulatory suppression (Gruber & von Cramon, 2003; Gruber, 2001). In our story-comprehension task, the verbal target information had to be held in working memory while the story presentation continued, that is, while the comprehension of the subsequent sentences constituted a secondary verbal task.

Moreover, the chronological target information required memory reaccess of the relevant context information mentioned at the beginning of the story. Behavioral evidence for this so-called reinstatement search (Kintsch, 1998; van Dijk & Kintsch, 1983; Kintsch & van Dijk, 1978) comes from eye-movement analyses. During reading of the chronological stories, regressions from the target information to the context sentence occur frequently (Rinck et al., 2003). The activation pattern elicited by the chronological target information indicates that hearing it triggers an attentional shift from local to global aspects of the discourse representation, associated with the retrieval of the relevant context information from long-term memory. In the emotional stories such a reinstatement search was not required. The relevant discourse information is repeatedly mentioned and is still available in working memory, when the target information is encountered.

In the right hemisphere, the IPS activation reached into the inferior parietal lobe or, more specifically, the temporoparietal junction area. This region has been identified as important for the processing of temporal information and for number processing (Rao, Mayer, & Harrington, 2001; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). More generally, Walsh (2003) proposed this region to be involved during the processing of magnitude information, independent of the specific domain. The present finding is consistent with this proposal. The chronological target information was realized in different ways. For instance, in the sample story in Table 1 chronological function words were used (before, after), whereas in some other stories the target information contained numerical information or required an estimation. The finding of the chronological
stories eliciting activation in the temporoparietal region is once more consistent with the suggestion that text comprehension includes processes that are representationally related to the information aspect of the specific content.

The second set of analyses provided information about integration processes after the target information was heard. In these analyses, based on short epochs encompassing the target information and the remainder of the stories, two further cortical regions were identified as important for situation model processing.

**The Ventrolateral Prefrontal Cortex**

The bilateral activations in the ventrolateral prefrontal cortex (vPFC; BA 47/11) during the integration of inconsistent information clearly differed from the frontopolar activations found in the event-related analysis. Left vPFC engagement, close to the most lateral left-sided peak found here, was also reported in two recent studies on language processing in context. Baumgartner, Weiller, and Büchel (2002) showed this region to be involved during the processing of semantically anomalous sentence completions compared with their expected counterparts. Caplan and Dapretto (2001) reported this area to be active during a comprehension task that required the detection and integration of unexpected topic changes in short two-sentence dialogues.

The sensitivity of the vPFC to violations of expectations is not dependent on the verbal domain (Petrides, Alivisatos, & Frey, 2002; Nobre, Coull, Frith, & Mesulam, 1999). The results of the present experiment are in line with the resulting view that the vPFC is required whenever explicit decisions are to be based on recent memory (Petrides et al., 2002). Verbal expectations were violated in the inconsistent condition and an explicit consistency judgment was required based on the discourse context represented in memory. The finding of this activation being somewhat stronger for the chronological stories is consistent with the higher likelihood of a reinstatement search in these stories.

**The Dorsomedial Prefrontal Cortex**

The dmPFC was the only region in which story type and consistency interacted during situation model integration. This region was active during the integration of emotional inconsistencies, but not when chronological inconsistencies were processed. Prior imaging research has shown the dmPFC to be engaged during inductive reasoning, evaluative processes, and inferencing (Zysset, Huber, Ferstl, et al., 2002; Ferstl & von Cramon, 2001; Goel, Gold, Kapur, & Houle, 1997). Thus, we interpret the dmPFC activation to reflect increases in inferencing demands when an emotional inconsistency was present. As predicted, the use of integrating inferences was more likely for the gradual emotional inconsistencies than for the all-or-none chronological inconsistencies. This conclusion is also reinforced by the fact that many participants judged more of the emotional than of the chronological stories as consistent. They succeeded in integrating the target information more often when listening to emotional stories.

The fMRI data support this conclusion as well. In the event-related analysis, the participants' performance level was related to the contribution of the dmPFC (BA 8) to the comprehension of the target information. Those comprehenders who frequently disagreed with the intended consistency judgment showed more frontomedial activation for inconsistent trials than for consistent trials. They were less sure about their answers and initiated an additional inferencing process. This individual difference is in line with an interpretation of the dmPFC activation not being directly caused by external stimulus properties, but rather being elicited by the internal evaluation process (cf. Ferstl & von Cramon, 2001, 2002).

Further evidence for this interpretation comes from analyzing the consistency effect during the end epochs for the emotional stories separately. Here, the dmPFC activation reached into a more posterior and ventral area in the pre-SMA (BA 6/32) and the neighboring BA 8. These regions are recruited when response conflicts arise (Ullsperger & von Cramon, 2001) or when the task induces uncertainty (Volz, Schubotz, & von Cramon, 2003), respectively. For the inconsistent emotional stories the participants carefully considered both responses and were not as confident about their final decision.

It is important to note, however, that in the present experiment the fuzziness of the inconsistency—probably related to the participants’ inferencing activities—and the story type were confounded. The dmPFC results can also be explained by the emotional inconsistencies inducing an affective component of mentalizing or Theory-of-Mind (ToM), a function attributed to the dmPFC (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Gallagher et al., 2000; Fletcher et al., 1995). Thus, in order to disentangle the respective contributions of ToM processes and inference demands, further research is needed (cf. Frith & Frith, 2003; Ferstl & von Cramon, 2002). However, the most important conclusion for theories of text comprehension is independent of this confound: The different story types did elicit qualitatively different and dissociable situation model integration processes.

**Summary and Conclusions**

The results of the present study are highly relevant for psycholinguistic theories of text comprehension. The first important issue concerns the representational format of the situation model. Elicited by the emotional target information, the orbitofrontal and ventromedial prefrontal activations, including the extended amygdala
complex, clearly show that the affective component of the stories directly induced processes beyond language comprehension. Thus, the situation model for these stories includes a nonverbal, nonpropositional representation of the affective dimension. Similarly, the parietal activation present during the processing of chronological target information indicates the buildup of information-specific situational representation.

Second, our results are consistent with the prediction of a reinstatement search for the chronological stories (Rinck et al., 2003; Kintsch & van Dijk, 1978). The memory component of this search was reflected in bilateral ventral prefrontal activation. In addition, the present data point to the importance of attentional components. Both the dorsal precuneus and the bilateral IPS activations suggest that reaccessing the discourse representation in memory requires a prior shift of attention from the local input to the global, contextual aspects. Attentional components have not yet received sufficient treatment in theories of text comprehension (cf. Gaddy, van den Broek, & Sung, 2001).

Finally, the frontomedial activation, sensitive to both performance level and information aspect, further confirms qualitative differences between the story types. Repair or integration of inconsistent words via inferential processing is attempted only if the consistency is a matter of degree, as in the emotional stories, but not when it is all or none. It is important to note that these explanations require a replication in which memory demands, that is, the distance of the target information to the relevant context information within the story, are manipulated independent of the information aspect. Furthermore, a replication using a larger number of trials, possibly using shorter stories, is desirable.

The results are also highly relevant for neuropsychological theories of text comprehension. There was clear evidence for right anterior temporal involvement when global inconsistencies were heard. In addition, we found activations in lateral and, most importantly, in medial prefrontal regions. Lesions in these frontal areas—even if they support domain-general cognitive processes—can cause so-called nonaphasic language deficits (Zalla et al., 2002; Ferstl et al., 1999, 2002; Novoa & Ardila, 1987; Kaczmarek, 1984). fMRI studies of text comprehension make a considerable contribution toward a better understanding of these deficits and of the interplay between language comprehension and higher level cognition.

**METHODS**

**Design**

The experiment varied the two factors Story Type (chronological vs. emotional) and Consistency (consistent vs. inconsistent) in a $2 \times 2$ within-subjects design.

**Materials**

Thirty-two stories were used in the experiment. Each of the stories consisted of seven sentences. In the emotional stories, a target word describing the emotional status of the protagonist appeared in the next to the last sentence. In the chronological stories, a target sentence mentioned information about a temporal sequence introduced in the second sentence. For both story types, an inconsistent version was created by switching the target sentence. The set of stories was fully balanced so that each target sentence appeared in one story in the consistent version and in another in the inconsistent version.

For the presentation, four different lists of 32 stories were created with 8 stories in each of the four cells of the design. Across lists, each story appeared twice in each consistency condition. The lists were then divided into two blocks, so that each of the stories in each consistency condition appeared once in the first block and once in the second block. Within the blocks, the order of presentation was pseudorandomized with the constraint of not more than three successive trials being from the same consistency or the same dimension condition.

Three additional stories resembling the experimental stories in length were added to each block. One was a practice trial at the beginning; one occurred in the first third of the block and contained a different situation model violation (e.g., a vegetarian eating meat) in an earlier position within the story. The third one, presented at the end of the block, ensured that the last experimental trial was presented under identical conditions as the previous ones.

**Tape Recording**

The stories were spoken by a female research assistant and directly recorded in digital format using the software package CoolEdit. To avoid eventual effects of prosodic cues, all stories were divided into three segments to be recorded separately. For recording, the sections from the different stories were randomly intermixed, so that the speaker was unaware of the experimental condition in which the segment was to appear during the experiment. The speaker read the segments with a slow and clear but natural articulation.

The duration of the stories was 43 sec, on average ($SD = 5.5$; range, 32–52 sec). Because only one or two words were different, the consistent and inconsistent versions were equally long. Similarly, there were no differences in duration between the two story types, $t(30) < 1.1$. Based on the recordings, the earliest point within the story was identified at which the inconsistency could be detected. For the emotional stories, this position was defined as the offset of the crucial word (e.g., hate vs. love); for the chronological story, it was defined as the first word in the target sentence that provided the complete chronological information (e.g., waiting).
On average, the target information was presented 36 sec ($SE = 5$) after the onset of the story, that is, about 7 sec before its end. Once more, there were no differences between the two story types with respect to the position of the target information, $t(30) < 0.2$.

**Participants**

Twelve women and eight men, all right-handers, received a small reimbursement for participating in the experiment. None of the 20 participants had any history of neurological disorder or other health problems preventing them from being exposed to the magnetic field. The median age was 25 years (range, 21–34). All participants had given informed consent and could withdraw from the experiment at any time.

**Procedure**

The instructions were presented in writing before the scanning session started. The participants were informed that short stories would be presented and that the task was to carefully screen for any inconsistencies concerning the content of the narratives. Three examples were provided to demonstrate the necessity to pay close attention. The participants were told that a short beep would indicate the beginning and end of each story. The response, a button press for either GOOD (consistent) or BAD (inconsistent), was to be given only after the end of the story. The right and left keys of a response box were randomly assigned to the two response types.

Before the start of the functional scan, one of the four presentation lists was randomly selected. The two blocks of 19 trials each were then presented with a short break between the blocks. Each trial started with the presentation of a warning tone of 200-msec duration to alert to the beginning of the story. One thousand milliseconds after the tone’s onset, the three story segments were presented with an intersegment interval of 500 msec. One second after the end of the story, a second warning tone requested the response for which a maximum response time of 5 sec was allocated. The reaction time and the accuracy of the response were recorded. Before the next trial started, there was an unfilled pause of variable length. The pause was 2 sec, on average, ranging from 0 to 4 sec after the maximal response time had elapsed. Thus, between the end of the presentation of one trial and the warning tone for the subsequent trial, there was a pause of 9 sec, on average (range, 7–11 sec, $SD = 1.23$). Because the participant’s button press occurred within this interval, the subjective intertrial interval was somewhat shorter. The variable length of the intertrial interval was intended for obtaining a better temporal resolution of the fMRI measurement and, at the same time, for making the procedure less predictable for the participant.

**Data Acquisition**

A Bruker (Ettlingen, Germany) Medspec 30/100 system was used for magnetic resonance imaging at 3 T. Before the functional scans, two anatomical scans were acquired for each participant using MDEFT sequences (modified driven equilibrium Fourier transform; Ugurbil et al., 1993). The first was a whole-brain image acquired with a T1-weighted 3-D segmented sequence: 128 sagittal adjacent slices, 1.5 mm thick, $256 \times 256$ pixel matrix per slice; TR = 1.3 sec, TE = 10 msec (Norris, 2000). Second, anatomical T1-weighted 2-D images were acquired, using the same number and orientation of slices as the functional scans.

Functional images were acquired using a single-shot, spin–echo EPI sequence (TR = 2,000 msec, TE = 75 msec; Norris et al., 2002). This sequence has the advantage of avoiding susceptibility artifacts in the ventromedial and orbitofrontal cortex, at the cost of a lower signal-to-noise ratio (Zysset, Huber, Samson, et al., 2003; Norris et al., 2002). Sixteen horizontal slices were measured (5-mm thickness, 2-mm spacing) parallel to the bicommissural plane (AC–PC). In-plane resolution was $3 \times 3$ mm and the repetition time was 2 sec (TR = 2). There were two functional runs of 508 time steps, each encompassing the presentation of 19 stories.

**Data Analysis**

Data analysis was conducted using the software package LIPSIA (Lohmann et al., 2001). Preprocessing began with motion correction consisting of a global affine linear transformation optimizing for each time step the linear correlation between the image at that time step and a constant reference image. A sinc-interpolation algorithm based on the Nyquist–Shannon theorem was used to correct for the temporal offset between the 16 slices of the acquired image (slice-time correction). A temporal high-pass filter with a cutoff frequency of 0.01 Hz was used for baseline correction, and a spatial Gaussian filter ($\sigma = 0.8$; corresponding to FWHM = 5.65 mm) was applied.

For the statistical analysis, each participant’s raw data were separately fitted into a standard stereotaxic space (Talairach & Tournoux, 1988) as follows: First, the 2-D data were rotated and shifted to map onto the 3-D whole-brain image and then linearly scaled to fit the standard size. The resulting coregistration matrices were finally applied to the raw data using trilinear interpolation (Lohmann, 1998). The data were rescaled so that all analyses were based on a voxel size of $3 \times 3 \times 3$ mm. The data were then analyzed using the general linear model (Friston, 1994), based on the within-subjects design Consistency $\times$ Story Type. Two different analyses were carried out: For the event-related analysis, the modeling was time locked at the offset of the target word. For the epoch-related analysis, the time between
the target word and the end of the story was defined as a short epoch. Note, however, that this analysis differs from the event-related analysis mainly with respect to the length of the window of analysis. The course of the experiment was still a pseudorandomized sequence of the different conditions.

For the event-related analysis, the design matrix was generated with a synthetic hemodynamic response function and its first derivative. For the epoch-related analysis, the matrix was generated by convolving a boxcar function with the hemodynamic response function. For both analyses, the model equation (observation data, design matrix, and error term) was convolved with a Gaussian kernel with a dispersion of 4 sec FWHM (Worsley & Friston, 1995). For each comparison of interest, a Z map was calculated for each participant based on t statistics. Because the signal-to-noise ratio is reduced by a factor of 2–3 in SE-EPI (Zysset, Huber, Samson, et al., 2003; Norris et al., 2002) the group statistics were based on the Gaussian test (Bosch, 2000). For this analysis, the individual Z maps are averaged and the resulting average is multiplied by the square root of N (number of subjects). This test is slightly more liberal than a random effects model. In the resulting average across all participants, those voxels with Z > 3.09 were considered significantly activated (p < .001). We defined an additional spatial extent threshold of 324 mm³, thereby ensuring an overall imagewise false-positive rate of p < .05 (cluster-level uncorrected; see Forman et al., 1995).

Three contrast images were calculated for both analyses, corresponding to the two main effects and the interaction of the factorial design. In particular, for the main effect of Story Type, we compared (emotional/consistent, emotional/inconsistent) with (chronological/consistent, chronological/inconsistent); for the main effect of Consistency, we compared (emotional/consistent, chronological/consistent) with (emotional/inconsistent, chronological/inconsistent), and for the interaction, we compared (emotional/inconsistent, chronological/consistent) with (emotional/consistent, chronological/inconsistent). In addition, for the epoch-related analysis, where there were significant interaction effects, pairwise comparisons were calculated, separately testing the influence of consistency for the emotional (emotional/inconsistent vs. emotional/consistent) and the chronological stories (chronological/consistent vs. chronological/inconsistent), respectively.

To control for interindividual variability, second-level statistics were conducted in all significant regions of activation (Bosch, 2000). Regions of interest (ROIs) were defined as neighborhoods of 189 mm³ around the peak coordinates (see Tables 3 and 4) of each activated area. Separately for each subject and each of the four conditions, the mean Z value within the ROI was calculated based on the contrast comparing the condition to rest (Bosch, 2000). These values were then analyzed in 2 × 2 ANOVAs, with the factors Consistency and Story Type for each ROI separately. The resulting F values and significance levels for the respective effect are presented in the rightmost columns of Tables 3 and 4.

The time course diagrams included for illustrating the most important effects were derived as follows: Percent signal change was defined for each voxel and each time step as the change from the average signal in this voxel across the entire scanning session. These values were averaged across all trials in each of the four conditions in a neighborhood of 26 voxels around the peak coordinate, that is, in a neighborhood of 729 mm³. Shown are the average percent signal change values from the presentation of the target information until 16 sec later. The end of the story presentation occurred approximately 8 sec after target presentation.

Acknowledgments

The data reported in this experiment have been deposited with the fMRI Data Center archive (www.fmridc.org). The accession number is 2-2004-11772.

Reprint requests should be sent to Evelyn C. Ferstl, Max-Planck-Institute of Human Cognitive and Brain Sciences, Stephanstrasse 1 A, 04103 Leipzig, Germany, or via e-mail: ferstl@cbs.mpg.de.

REFERENCES


