Text Comprehension After Brain Injury: Left Prefrontal Lesions Affect Inference Processes

Evelyn C. Ferstl
Max-Planck-Institute of Cognitive Neuroscience

Thomas Guthke
University of Leipzig

D. Yves von Cramon
Max-Planck-Institute of Cognitive Neuroscience and University of Leipzig

Comprehending language in context requires inferencing, particularly for the establishment of local coherence. In the neurolinguistic literature, an inference deficit after right hemisphere brain damage has been postulated, but clinical observation and imaging data suggest that left-frontal lesions might also result in inference deficits. In the present experiment, 25 nonaphasic patients performed a coherence judgment task requiring them to indicate a pragmatic connection between 2 successively presented sentences. Patients with left-temporal or right-frontal lesions performed the task well. In contrast, patients with left- and bifrontal lesions exhibited the most severe deficit. Both error rates and response times were elevated for coherent trials as compared with incoherent trials. These results confirm that the left-frontal lobe contributes to inference processes.

Efficient communication is dependent on the listeners' or readers' ability to fill in information left implicit by the speaker or writer. This inferencing process takes place continually during language comprehension and most often without the awareness of the comprehender.

Neurolinguistic research has almost exclusively concentrated on inference deficits in patients with right hemisphere brain lesions (for overviews, see Beeman, 1993; Brownell & Martino, 1998). Because clinical observations of some of these patients have shown that they have inappropriate discourse behavior, they are seen to have a particular deficit in sufficiently taking into account the context. The types of processes hypothesized to be impaired include as varied a set as the activation of semantic associations (Beeman, 1993), the revision of initial interpretations, social pragmatics, nonliteral language, and indirect speech (e.g., Brownell, Gardner, Prather, & Martino, 1995).

Recently, Lehman and Tompkins (2000) argued that the empirical evidence is not as clear-cut as it seems. After reviewing the relevant neurolinguistic literature, they concluded that inferencing deficits in right-brain damaged (RBD) patients have not been documented conclusively. Furthermore, McDonald (1993) pointed out that language

deficits after right hemisphere brain damage closely resemble those described after prefrontal lesions, irrespective of the lateralization. In particular, patients with prefrontal brain damage, either focal or after traumatic brain injury (TBI), often fail to take into account the communicative context, fail to take into consideration the listeners' needs, or have difficulties with structuring coherent discourse (cf. Kaczmarek, 1984; Novoa & Ardila, 1987; Prigatano, Roueche, & Fordyce, 1986). One possible explanation for these frontal nonaphasic language deficits is that discourse production and text comprehension require the use of executive functions, such as structuring, monitoring, and problem solving.

Two main reasons for these inconsistencies in the neurolinguistic literature have been proposed. The first issue concerns the patient selection. McDonald (1993) notes that many samples of RBD patients tested in inferencing studies include patients with prefrontal brain damage or subcortical lesions that might also cause frontal dysfunction, so that right hemisphere and prefrontal dysfunction cannot be cleanly separated. Lehman and Tompkins (2000), on the other hand, point out a sampling bias. Some researchers, they argue, specifically select RBD patients exhibiting symptoms of a nonaphasic communication deficit and thus increase the likelihood of documenting problems with text-level processes. Finally, in many lesion studies, researchers compare RBD patients to a control group without brain injury, so that it becomes difficult to separate the general effects of illness and brain damage from specific effects due to lesion lateralization (cf. Brownell & Martino, 1998; see also Snow, Douglas, & Ponsford, 1997). This problem becomes particularly apparent in studies focusing on patients with infarction of the middle cerebral artery, typically causing rather large lesions.

The second issue concerns the selection of materials and comprehension tasks. Despite the extensive psycholinguis-

tic literature on inference processes, there are few systematic investigations of different inference types in a brain-injured population, and few studies include a systematic and careful analysis of the task requirements (see Beeman, 1993, for an exception). Moreover, it is often overlooked that many studies taken as demonstration of a right hemisphere inferencing deficit require the processing of affective and emotional, social and pragmatic, or prosodic information (for overviews, see Brownell et al., 1995; Brownell & Martino, 1998; Lehman & Tompkins, 2000). These processes, most of which have been postulated to be realized in right hemisphere brain regions, need to be carefully separated from inference processes during text comprehension as they have been studied in psycholinguistic research (e.g., Kintsch, 1998; Singer, 1994). Therefore, there still is a need for carefully conducted empirical studies on inference processes after brain injury.

An alternative approach to studying the neuroanatomical bases of inference processes is the use of recent neuroimaging techniques for directly measuring the relative involvement of various brain regions during text-level processes in healthy participants. However, only a few imaging studies investigating language processing in context are available, and they have produced conflicting results as well. When comparing connected discourse to single, unrelated sentences, Fletcher et al. (1995) and Mazoyer et al. (1993) found activation in anterior temporal and superior and medial left-frontal brain areas. Similarly, when presenting stories with and without illustrating pictures to render locally incoherent texts comprehensible, Maguire, Frith, and Morris (1999) reported activation in the median wall of the left-frontal lobe but no right-sided activation foci. Using a similar paradigm in which a title was used for rendering the stories coherent, St. George, Kutas, Martinez, and Sereno (1999) reported right-temporal areas to be particularly sensitive to coherence differences. For their functional magnetic resonance imaging (fMRI) study, Robertson et al. (2000) induced coherence processes by varying cohesion, that is to say, by using lexical information to signal connections between sentences. On the basis of the presupposition that definite articles refer to previously known or mentioned entities, comprehenders are more likely to connect a list of sentences into a coherent story representation when definite articles are used than when indefinite articles are used. This processing difference was reflected in an increase of right-sided prefrontal activation for the coherent condition compared with the incoherent condition.

Ferstl and von Cramon (2001) used whole-head event-related fMRI to study local coherence and its interaction with lexical cohesion. The task was to judge whether two successively presented sentences had "something to do with each other," that is to say, whether the sentence pair was pragmatically coherent or not. Compared to a baseline, the language trials elicited clearly left dominant activation, a result that is consistent with numerous previous studies on language processing. The activation foci included anterior and posterior regions along the left superior temporal sulcus, a left-sided, inferior prefrontal region including the pars triangularis and left-sided superior and median prefrontal

regions (BA 8/9). The right-sided homologues of the temporal and inferior prefrontal activation foci were also active, but to a considerably lesser extent. When comparing coherent sentence pairs to incoherent ones, the student volunteers showed activation in the frontomedian wall. Once more, this activation was left lateralized. There was no evidence for a modulation of right hemisphere activation by coherence.

The second factor investigated in this study was the cohesion of the sentence pairs. *Cohesion* refers to the lexical connection between subsequent sentences. It is signaled by cohesive ties, as, for instance, definite articles, pronouns, anaphora, or conjunctions (see Halliday & Hasan, 1976). There is plenty of evidence that cohesion facilitates comprehension (cf. Gernsbacher, 1990) and that these effects have neuropsychological manifestations (e.g., Robertson et al., 2000, for an fMRI experiment; and Münte, Schiltz, & Kutas, 1998, for a study using evoked potentials). In discourse production, it has been shown that neurological patients, including TBI patients, Alzheimer patients, and patients with aphasia, have difficulties with the sufficient and unambiguous use of cohesive ties (e.g., Coelho, Liles, & Duffy, 1995; Glosser & Deser, 1990; Lock & Armstrong, 1997).

In the fMRI study by Ferstl and von Cramon (2001), the coherence and cohesion factors were crossed (see Table 1 for examples of the resulting four conditions). Behavioral data confirmed that for the coherent sentence pairs, cohesion facilitated processing. For the incoherent pairs, the cohesive ties produced a mismatch between lexical and pragmatic information. The cohesive ties indicated a connection, but there was no pragmatic solution to the inference problem. In this conflict condition, processing times and error rates increased. In the imaging data, left-sided prefrontal activation at the junction of the precentral and the inferior frontal sulci was seen. This area has been implicated as having a functional role for a variety of tasks targeting executive functions, such as the Stroop task, go/no-go tasks, dual-task, or task-switching paradigms (Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Zysset, Müller, Lohmann, & von Cramon, 2000). To summarize, this study provided evidence for the contribution of left-lateral prefrontal regions to text-level language processing, and its results were consistent with findings of frontomedian and superior frontal regions being involved during language comprehension in context (Fletcher et al., 1995; Maguire et al., 1999). In contrast, it did not support the claim that the right hemisphere has a special role for inference processes as required in Ferstl and von Cramon's (2001) coherence judgment task (cf. Robertson et al., 2000).

The goal of the present study was to find converging evidence for these results in a patient study. Using the same materials and the same task, we examined nonaphasic patients with brain lesions resulting from various etiologies. By measuring reading and judgment times as well as accuracy data, we attempted to identify patients who had problems with the coherence judgment. The patient group in the focus of our study had prefrontal brain lesions. Within this group, we distinguished left-sided, bilateral, and right-lateralized brain injuries. A further group of participants had

Table 1
Example Materials for the Four Conditions of the Experiment

Coherent	Incoherent
Incohesive	
Mary's exam was about to start. The palms were sweaty.	Laura got a lot of mail today. The palms were sweaty.
Laura got a lot of mail today. Some friends had remembered the birthday.	Mary's exam was about to start. Some friends had remembered the birthday.
Sometimes a truck drives by the house. The dishes start to rattle.	The lights have been on since last night. The dishes start to rattle.
The lights have been on since last night. The car doesn't start.	Sometimes a truck drives by the house. The car doesn't start.
Cohesive	
Mary's exam was about to start. *Therefore, her* palms were sweaty.	Laura got a lot of mail today. *Therefore, her* palms were sweaty.
Laura got a lot of mail today. *Her* friends had remembered *her* birthday.	Mary's exam was about to start. *Her* friends had remembered *her* birthday.
Sometimes a truck drives by the house. *That's when* the dishes start to rattle.	The lights have been on since last night. *That's when* the dishes start to rattle.
The lights have been on since last night. *That's why* the car doesn't start.	Sometimes a truck drives by the house. *That's why* the car doesn't start.

Note. Items have been translated from the original German. Cohesive ties are italicized.

damage within the left-temporal lobe, an area in which the fMRI study uncovered large and stable regions of activation. As a control group, patients whose lesions spared frontal and left-temporal areas were included.

Although it is known that both right hemisphere patients (Leonard, Waters, & Caplan, 1997a, 1997b) and left hemisphere patients (Chapman & Ulatowska, 1989) use contextual information for the disambiguation of pronouns, we did not have specific predictions about the influence of cohesive ties on coherence processes in the brain-injured population. Nevertheless, the error distributions across the four conditions of the experiment were expected to shed light on the relative contributions of lexical factors to inference processes. For the influence of pragmatic coherence on the judgment task, we had the following predictions: The hypothesis that inference processes are mediated by the right hemisphere predicts that patients with right-sided lesions would show a deficit in the coherence judgment task. In contrast, according to the imaging results (Ferstl & von Cramon, 2001) and following clinical descriptions of non-aphasic language disorders, we expected patients with prefrontal lesions, and in particular those with left-sided lesions, to be more likely to show such a deficit.

Method

Participants

Thirty-one patients admitted to the Outpatient Clinic for Cognitive Neurology at the University of Leipzig, Leipzig, Germany, participated in the experiment. Patients with more than mild aphasic symptoms were excluded. None of the patients had a severe vision deficit or acquired dyslexia, so they were able to read and comprehend sentences on a computer screen. Twenty-one patients were right handed, 1 was ambidextrous, 2 were nonfamilial left-handed, and for 1, there was no information about handedness available. No patient exhibited evidence of right-lateralized language dominance.

For an evaluation of the effects of the lesion site on coherence processes, D. Yves von Cramon performed a radiological diagnosis based on anatomical MRI scans (axial T1- and T2-weighted, and 3D data set) taken at least 3 months after brain damage. Detailed information about the medical diagnosis is provided in Table 2. The diagnosis included the scoring of the presence and laterality of both frontal lesions and temporal lesions. Following this lesion evaluation, 25 patients were assigned to one of five subgroups. Crucial for this assignment was the presence of a frontal or left-temporal lesion. Among the 15 patients with frontal brain damage, we identified 7 with unilateral right-sided lesions (RF), 4 with unilateral left-sided lesions (LF), and 4 with bilateral frontal lesions (BF). In addition to the frontal patients, two further patient groups were formed. Five patients had left-temporal lesions but no involvement of frontal areas (LT). The fifth group of 5 patients had brain damage sparing frontal as well as left-temporal regions; this last group was considered the patient control group (C). Additional cortical lesions were not taken into account for group assignment, but they are included in Table 2. Two of the 8 patients in the LF and BF groups had additional cortical lesions. In each of the five groups, about half of the patients had sustained TBI, whereas the other half were patients with lesions resulting from vascular or other etiologies. For 2 TBI patients in Group C, the MRI scan showed no visible brain lesions. In the acute stage, however, both of them had suffered multiple microbleeds that provided indirect evidence for diffuse axonal injury.

For 6 patients, an unambiguous assignment to one of the groups was not possible. Two of these patients had pure subcortical lesions whose effect on frontal functions was unclear; for 1 patient, no MRI scan was available; and for 3 patients, the lesions involved

Table 2
Demographic Characteristics, Etiology of Condition, Radiological Diagnosis, and Group Classification for 25 Participants

Patient	Age	Gender	TSL	Etiology	Left hemisphere lesion	Right hemisphere lesion	Group
033	50	M	35	TBI	Frontopolar; frontolateral (F1, F2, F3)	Frontopolar, temporopolar	BF
044	38	M	4	ICH	Frontodorsal (excl. WM)	Frontodorsal (excl. WM); internal capsule (genu)	BF
051	26	M	39	TBI	Frontopolar, anterior fronto-orbital	Frontopolar, anterior fronto-orbital	BF
934	62	M	4	INF	Frontolateral (posterior F2, F3; fronto-opercular cortex, anterior insula)	Anterior orbital gyrus, calcarine lip, occipital gyri	BF
170	35	M	71	TBI	Frontodorsal (anterior F1, excl. WM)	—	LF
201	50	M	45	INF	Frontolateral (posterior F3); anterior insula	—	LF
363	50	M	6	INF	Frontolateral (posterior F3, fronto-opercular cortex); insula; striatum (putamen, caudate nucleus)	—	LF
804	41	M	11	TBI	Frontodorsal (anterior F1, excl. WM); anterior insula	—	LF
072	50	M	21	TBI	—	Frontolateral (anterior F3); temporolateral (anterior T1, T2, T3)	RF
073	37	M	28	TBI	—	Frontopolar, frontolateral (anterior F3); temporopolar, temporolateral (anterior T2, T3)	RF
514	35	M	296/149	ICH/MF	—	Frontodorsal (excl. WM)	RF
561	39	F	23	SAH	—	Frontoorbital (G. rectus, medial orbital gyrus)	RF
743	49	F	11	TUM	—	Frontodorsal (middle F1, F2)	RF
853	18	M	8	TBI	—	Frontolateral (posterior F3; temporopolar, temporolateral (anterior T1, T2)	RF
933	51	M	8	SAH	—	Frontodorsal (excl. WM); corpus callosum (body, isthmus)	RF
252	45	M	14	TUM	Anterior temporal lobectomy	—	LT
393	58	M	6	INF	Temporolateral (posterior T2)	—	LT
504	29	M	8	SAH	Temporopolar, temporolateral (anterior T1, T2, T3)	—	LT
713	54	F	24	TBI	(Temporopolar), temporolateral (T1, T2, T3, T4)	—	LT
774	34	M	16	ICH/MF	Temporolateral (posterior T1, posterior STS)	—	LT
053	24	F	8	TBI	—	—	C
113	33	M	229	TBI	—	Temporolateral (posterior T3)	C
263	37	F	5	TBI	—	—	C
753	49	M	10	INF	—	Temporolateral (posterior T2)	C
834	50	M	3	INF	Internal capsule (genu)	—	C

Note. Dashes indicate no visible lesions. TSL = time since lesion (in months); M = male; TBI = traumatic brain injury; F1–F3 = frontal gyri (superior to inferior); BF = bifrontal; ICH = intracerebral hemorrhage; excl. WM = exclusively white matter; INF = infarction; LF = left frontal; T1–T4 = temporal gyri; RF = right frontal; MF = arterio-venous malformation; F = female; SAH = subarachnoid hemorrhage; TUM = tumor; LT = left temporal; STS = superior temporal sulcus; C = control.

both left-temporal as well as frontal brain areas. The data for these 6 patients were excluded from further analysis, so that 25 patients remained in the sample.

Design and Materials

The language trials were based on 120 sentence pairs in which the second sentence (the target) was pragmatically related to the first (the context). The sentence pairs were from a wide variety of topics and used different syntactic structures, but care was taken that only half of the sentences mentioned people, and the other half were about inanimate subjects. The relation between sentences could not be derived solely on the basis of associative links between content words; rather, comprehension required the use of general world knowledge. Each target occurred in two versions: the cohesive version contained one or two lexical items (e.g., pronouns or conjunctions) that explicitly signaled the connection between the sentences. In the incohesive version, these so-called "cohesive ties" were omitted or replaced, so that the relationship between the two sentences had to be inferred based on pragmatic information alone. The incoherent conditions were created by switching the context sentences of two coherent trials. As for the coherent conditions, the target sentences in the incoherent conditions appeared both in cohesive and incohesive versions. Thus, the experiment used a 2×2 within-subjects design, with the variables being cohesion (yes or no) and coherence (yes or no).

Examples for the resulting four conditions of the experiment are shown in Table 1. As can be seen, the target sentences in the coherent and incoherent conditions are identical, so that syntactic, lexical, and semantic features are kept constant. Moreover, in German, word order is more flexible than in English, so that the cohesive ties were not necessarily at the beginning of cohesive target sentences, as in the translated versions. Most important, the examples clearly illustrate that lexical cohesion has opposing effects, depending on the pragmatic coherence. In the incoherent trials, cohesion leads to a pragmatic garden-path effect, whereas in the coherent trials, cohesion facilitates inferencing.

Care was taken that the incoherent sentence pairs did not yield unintended pragmatic relationships. To confirm this and to evaluate the hypothesized interaction of coherence with cohesion, Ferstl and von Cramon (2001) pretested the sentence pairs in an experiment with a group of 24 healthy students. The procedure of this experiment was identical to that of the present study. Three dependent variables converged on similar conclusions: An analysis of the error rates, as well as the reading times for the target sentences and the judgment times, confirmed the effects of the experimental variables. In particular, there was a main effect of coherence in the error rates; coherent trials were more difficult than incoherent trials. The reading times and the judgment times for the incoherent, misleadingly cohesive condition were significantly longer than for the incohesive version. When correcting for sentence length, it could be seen that the facilitative effect of cohesion in the coherent trials was also obtained. In summary, the pretest confirmed the hypothesized interaction between coherence and cohesion. For ease of comparison, the respective results from this pretest are displayed together with the results of the current study in the Results section.

For the present study, we shortened the experiment to 80 sentence pairs. Predominantly, we omitted those trials that contained infrequent content words, that were particularly long, or for which the control group's responses were inconsistent. For the resulting selection, the average number of words in the context sentences was 6.7 ($SD = 2.1$). The cohesive target sentences were slightly longer ($M = 6.6$, $SD = 1.4$) than the incohesive sentences ($M = 6.1$, $SD = 1.4$). In the pretest, the overall error rate for this selection was 3.7%. Four different lists of 80 trials were used for counterbalancing. Each list contained each target sentence exactly once. Twenty trials appeared in each of the four conditions in the Coherence $\times$ Cohesion design. Across the four lists, each target sentence occurred once in each of the four conditions. The trials in each block were ordered pseudorandomly, with the constraint that not more than three trials of the same condition appeared successively. Finally, the lists were further subdivided into four blocks of 20 trials, so that the presentation order could be varied, and breaks within the experiment were possible.

Procedure

The experiment was carried out in individual sessions, lasting about 20 to 40 min. The participants were seated in a quiet room in front of a computer screen, and a response box was placed into their dominant hand.[1] After the experimenter had given a short introduction, she explained the three response keys on the button box, all of which were to be pressed using fingers of the dominant hand. The function of the middle key was to proceed with the experiment (i.e., it was used to record the reading times). The left and right buttons were used for the coherence judgments, and they were randomly assigned to yes and no responses. In each of the five patient groups, the key assignment was approximately equally distributed. Further instructions, examples, and practice trials were then presented on the computer screen. To practice the key assignment, the patient was first asked to press the correct key after seeing one of the cues, yes or no. This task was repeated six times. The instructions for the coherence judgment task then told the patients that sentences would be shown one at a time and that the task was to read them carefully for comprehension. Examples were used to illustrate the coherence judgment task. The patient was to respond "yes" whenever two subsequent sentences had "something to do with each other" and respond "no" when there was no connection whatsoever.

After the experimenter had ensured that the instructions were fully understood, the patient proceeded with the experiment. Four blocks of 20 trials were presented. Before the beginning of each block, two practice trials were shown, and after each block, the patients could take a short break if they wished to do so. One of the four lists of trials was randomly assigned and combined with one of four different block orders. The presentation of each trial was self-paced. Following a fixation cross, the context sentence appeared. By pressing the middle button, the patients indicated that they had understood the context sentence and that they were ready to proceed. The target sentence then appeared. Once more, pressing the middle button indicated that the sentence was read, and a question mark appeared to request the coherence judgment. Participants gave their response by pressing the left or the right button on the button box. The accuracy of this response and the reaction times for the three button presses were recorded.

Results

Data Analysis

For each of the 25 participants and each of the four experimental conditions, we calculated the error rates. The

---

[1] One patient (804) suffered from hemiparesis, preventing him from using his dominant hand, but inspection of his data show that with 2,808 ms on average, his total response times were in the lower range.

reading and judgment times were both individually corrected for outliers by truncating the times at a cutoff value of two standard deviations above or below the participant's mean. Fewer than 5% of the observations were affected by this procedure. After this correction, overall means of the reading and judgment times were calculated for each patient. We used only those trials that were followed by correct responses. For analyzing the influence of the within-subjects factors on the response times, it was necessary to reduce the large interindividual variability. Using the individual means and standard deviations, we standardized each patient's data reading times and judgment times separately. For each condition, the resulting z scores were averaged. Once more, only trials followed by correct responses contributed. Thus, an overall mean of less than 0 indicates that the processing of correct trials was faster than that of incorrect trials. A positive mean for one of the four conditions indicates that trials in this condition were processed more slowly than average; in other words, these trials were more difficult than trials of the other conditions.

For the statistical analysis of group differences, we used four orthogonal contrasts for coding the lesion variables. Contrast codes have the advantage of testing more focused hypotheses than the variable group, so that post hoc comparisons are unnecessary (Judd & McClelland, 1989). Furthermore, contrast codes include all observations in each of the focused comparisons, yielding an increase in power compared with pairwise comparisons of subgroups. The first contrast, *frontal*, compared the three frontal with the two nonfrontal groups (LF, RF, and BF vs. C and LT); the *temporal* contrast compared the control group with the left-temporal participants (C vs. LT); the *left* contrast compared frontal patients with left-sided or bilateral frontal lesions with those with unilateral right-sided frontal lesions (LF and BF vs. RF); and the final contrast, *bilateral*, compared those with unilateral left-frontal lesions with patients with bilateral frontal lesions (LF vs. BF).

Error Rates

The mean error rates, as a function of group and condition, are shown in Figure 1. Individual data for both coher-
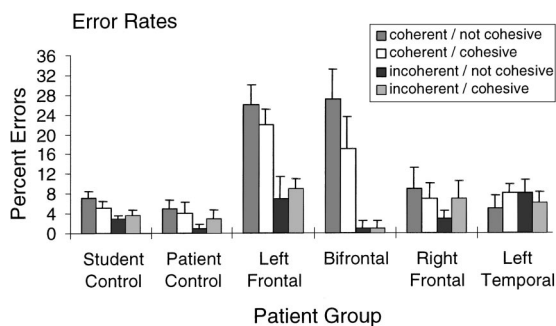

Figure 1. Error rates as a function of coherence and cohesion for the five patient groups and, for ease of comparison, the data from a student control group (Ferstl & von Cramon, 2001). The error bars indicate one standard error above the mean.

ent and incoherent trials are provided in Table 3. Overall, the patients responded in 91.7% of the trials correctly ($SD = 5.8\%$). None of the effects involving cohesion were significant, but across the entire sample of patients, there was a highly significant main effect of coherence, $F(1, 20) = 27.7$, $p < .0001$. It was more difficult to establish a pragmatic connection than to detect a coherence gap. In addition, there were clear and highly significant group differences. Patients with lesions involving the left-frontal lobe made more errors (12% for BF vs. 16% for LF) than the other three patient groups who showed almost perfect performance (7% for both RF and LT and 3% for C). Correspondingly, there were significant main effects of frontal, $F(1, 20) = 14.9$, $p < .001$, and left, $F(1, 20) = 12.3$, $p < .01$. The group differences in performance were due mainly to errors in the coherent trials, resulting in highly significant interactions of coherence with these two group contrasts; for Coherence × Frontal, $F(1, 20) = 14.6$, $p < .01$; and for Coherence × Left, $F(1, 20) = 14.3$, $p < .01$.

The main results of this analysis are that patients with left-frontal lesions had problems performing this task and that they showed a clear and highly reliable coherence effect.

Response Times

The mean reading and judgment times for each participant are also shown in Table 3. The reading times for the target sentences were 4,045 ms ($SD = 1,440$) on average. As expected, these times were considerably longer than the times for the student control group, who needed about 2,500 ms for comprehending the target sentences. The reading times did not differ systematically across lesion groups, $F(4, 20) < 1$, but they varied greatly across patients (range $= 1,874–8,009$ ms). Similarly, the judgment times ranged from 231 ms to 1,436 ms, with an overall mean of 641 ms ($SD = 345$). Once more, these times were somewhat longer than those of the students, who needed about 460 ms on average. A marginally significant difference for the contrast frontal, $F(1, 20) = 4.2$, $p = .05$, showed that the judgment times were slightly longer for patients with frontal lesions.

For an evaluation of the effects of the experimental factors, the data were transformed into z scores and analyzed in a mixed analysis of variance (ANOVA). In the student control group, the reading times were more sensitive to the experimental factors than were the very fast judgment times, indicating that the coherence judgment was made online, while the target sentence was read. However, in our patient group, both measures were more variable, within as well as between participants. Some patients followed the control group's pattern, whereas others' long judgment times indicated that they shifted part of their decision process into the interval after the presentation of the target sentence. To capture both of these patterns and to evaluate whether these differences might depend on lesion location, we included both measures in one overall analysis. Thus, the ANOVA contained the additional within-subjects factor of variable (reading time score vs. judgment time score)

Table 3
Individual Data and Group Means for Errors and Response Times

Patient ID	Error rate (%): coherent	Error rate: incoherent	Reading time (ms)	Judgment time (ms)
		Bifrontal group		
033	.25	.05	4,986	865
044	.08	.00	1,874	465
051	.23	.00	2,905	344
934	.35	.00	8,009	1,181
M	.23	.01	4,443	714
		Left-frontal group		
170	.30	.10	4,119	1,436
201	.25	.03	3,245	486
363	.25	.15	4,858	905
804	.18	.05	2,533	275
M	.24	.08	3,689	775
		Right-frontal group		
072	.15	.13	4,708	1,088
073	.13	.00	5,283	399
514	.15	.00	2,996	309
561	.03	.03	4,238	877
743	.00	.00	4,136	855
853	.10	.08	7,635	1,302
933	.03	.13	3,534	585
M	.08	.05	4,647	773
		Left-temporal group		
252	.05	.10	5,198	380
393	.08	.05	2,738	231
504	.00	.13	2,597	327
713	.08	.08	4,350	556
774	.13	.00	3,692	625
M	.07	.07	3,715	424
		Control group		
053	.03	.00	3,684	374
113	.08	.03	3,876	386
263	.03	.00	2,976	706
753	.00	.00	3,537	291
834	.10	.08	3,420	777
M	.05	.02	3,499	507
Overall *M*	.12	.05	4,045	641

besides the factors of coherence, cohesion, and the four contrasts coding the group factor.

Across the entire patient group, there was a main effect of cohesion, $F(1, 20) = 8.5$, $p < .01$, indicating that cohesive trials took longer to process than incohesive trials. The Coherence × Cohesion interaction, $F(1, 20) = 4.3$, $p = .05$, indicated that across the entire group, the patients' processing times showed a similar pattern as those of the student control group (cf. Ferstl & von Cramon, 2001). Both of the stimulus-dependent effects were more apparent during the reading phase than during the judgment phase of the trials: for the Variable × Cohesion interaction, $F(1, 20) = 13.7$, $p < .01$; for the three-way interaction Variable × Coherence × Cohesion, $F(1, 20) = 5.2$, $p < .05$. These findings are depicted in Figure 2, and they confirm that as a group, the patients were sensitive to the lexical factor cohesion and

its impact on coherence building. In particular, it was most difficult to read the target sentences when there was a mismatch between cohesion and coherence (i.e., when a cohesive tie falsely indicated a connection between two pragmatically incoherent sentences).

In addition to these general effects, there were clear and reliable group differences concerning the impact of the coherence factor. As can be seen in Figure 3, the left frontal and the bifrontal groups took longer to process the coherent trials, whereas the other three patient groups showed the opposite pattern. This result was reflected in the significant Coherence × Frontal interaction, $F(1, 20) = 5.6$, $p < .05$, and, most important, the Coherence × Left interaction, $F(1, 20) = 8.5$, $p < .01$. The lack of an interaction of these effects with variable suggests that both reading and judgment time scores yielded analogous patterns. And, indeed,
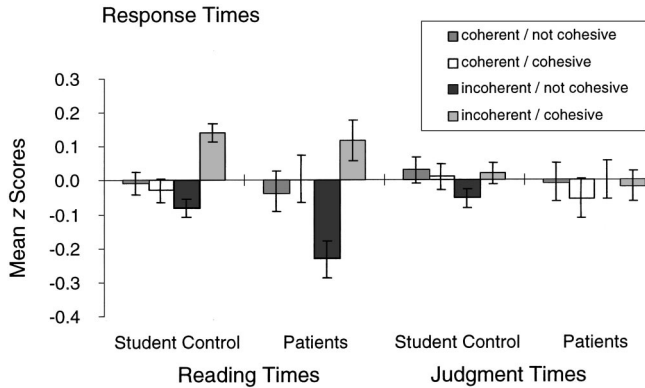
Figure 2. Response times as a function of coherence and cohesion for the entire patient group. Shown are the *z* scores for the reading times and the judgment times. The error bars are one standard error above the mean. The student control group's data from the pretest (Ferstl & von Cramon, 2001) are displayed for comparison.

the coherence effect for the reading time scores, that is, the difference between coherent and incoherent trials, was 0.33 (*SD* = 0.53) for the patients with left-frontal lesions (Groups LF and BF), compared with −0.10 (*SD* = 0.41) for the patients without left-frontal lesions (Groups C, RF, and LT), *t*(23) = 2.2, *p* < .05. The corresponding result for the judgment time scores was 0.30 (*SD* = 0.42) for the patients with left-frontal lesions, compared with −0.18 (*SD* = 0.37) for patients without left-frontal lesions, *t*(23) = 2.9, *p* < .01. Thus, both measures converged on the finding that patients with left-frontal lesions needed longer to process coherent trials than to process incoherent trials.

Relationship Between Speed and Accuracy

The results of the accuracy data and the response times converged on the conclusion that the group factor indicating the presence of a left frontal lesion was crucial for predicting difficulties with coherence building. Independent of the
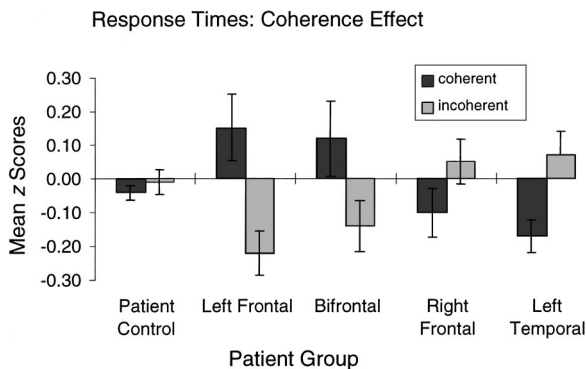
particular patient classification, we evaluated the relationship between the variables. By calculating rank order correlations, we reduced the influence of outliers.

For the group of 25 patients, there was a positive relationship between the error rates and the judgment times (*r* = .41, *p* < .05). Participants taking longer for the coherence judgment task made more errors. The reading times were highly correlated with the judgment times as well (*r* = .63, *p* < .001). The relationship between the error rates and the target reading times was also positive, but not significantly so (*r* = .28, *ns*). This pattern of results clearly rules out a speed–accuracy tradeoff. All three variables capture processing difficulty in the coherence judgment task.

A second question was whether the coherence effect (i.e., the difference between coherent and incoherent trials) was reflected similarly in all three dependent variables. There were highly significant effects of the Coherence × Left interaction for both error rates and response times. Correspondingly, the rank order correlation between the coherence effect of the error rates and that of the response times was also reliable (*r* = .67, *p* < .001). To illustrate this correlation and to show its relationship to the lesion groups, Figure 4 presents a scatterplot of the individual data.

Neuropsychological Control Variables

To ensure that the patient groups in this experiment were comparable with respect to overall level of cognitive functioning and for evaluating relationships between experimental performance and other cognitive processes, we now consider the patients' neuropsychological profiles. Besides demographic characteristics, the available data from the standard diagnosis carried out in the Outpatient Clinic are

Figure 3. Response times as a function of coherence and patient group. Shown are the overall mean *z* scores (±1 *SE*), including both the reading and the judgment phases of the trials.
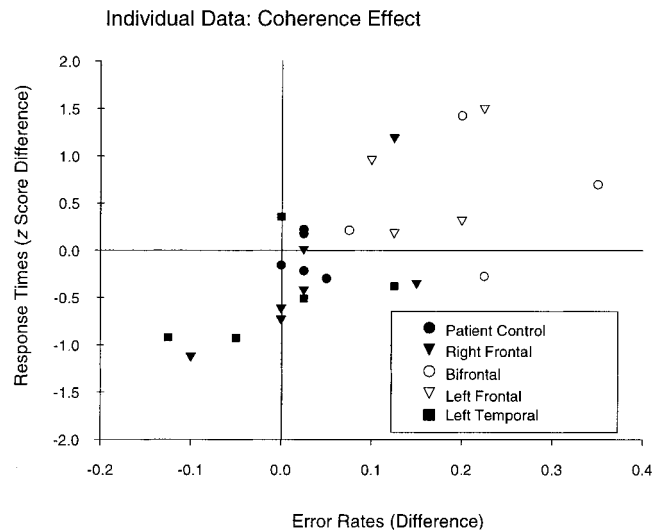

Figure 4. Scatterplot illustrating the relationship between the error rates and the response times. Shown is the coherence effect in the response time scores as a function of the coherence effect as reflected in the error rates. For both variables, the *coherence effect* is defined as the difference between coherent and incoherent trials.

presented and analyzed. Because the tests were part of the clinical assessment, not all tests had been administered to each patient. From the available test results, only those directly relevant to the inferencing task were selected, such as tests assessing verbal memory and verbal learning and tests related to executive functions. To screen for severe attentional deficits, we include some test results from this area as well.

The patients' information processing speed, an aspect of attention, was assessed using the Alertness subtest from the Testbatterie zur Aufmerksamkeitsprüfung (Test of Attentional Processes, or TAP test; Zimmermann & Fimm, 1992). This subtest measures simple reaction times to briefly presented stimuli. We also report omissions from the subtest evaluating divided attention. In this test, the patients are required to monitor auditory and visual stimuli simultaneously for targets.

Verbal fluency was assessed using two tests. For measuring literal fluency, we chose the format of the LPS-6, which is a subtest of the German intelligence test Leistungsprüfsystem (Performance Assessment System, or LPS; Horn, 1983). As usual, the task is to list as many words as possible starting with a given letter. Semantic fluency was assessed using the "supermarket" test, in which the participants are first asked to list as many items as possible that can be bought in the supermarket. After this free production, four subcategory labels (e.g., "vegetables") are presented for cued production.

An aphasia diagnostic was carried out using the Aachen Aphasia Test (AAT; Huber, Poeck, & Willmes, 1984). However, only those patients whose screening showed impairments on the word and sentence level were tested. The screening was based on an evaluation by speech therapists and the age-corrected error scores of the Token Test. This test is included in the AAT as a subtest and has been shown to be diagnostic for aphasic language deficits (De Renzi & Faglioni, 1978).

Verbal and visual memory was assessed with the Wechsler Memory Scale—Revised (WMS–R; Wechsler, 1987). In addition to the general memory quotient, we report two subtests especially relevant for recall of verbal material. As an assessment of verbal short-term memory deficits, results from the Digit Span subtest are reported for both forward and backward recall. Long-term memory for complex verbal material is captured by the Logical Memory subtest of the WMS–R. In this test, participants recall two short stories immediately after they were read to them.

For an evaluation of verbal learning abilities, a list-learning test similar to the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Obler, 1987) was used. Patients are required to memorize a list of 16 auditorily presented nouns from four different categories. The list is presented a total of five times, so that in addition to the initial recall ability, the course of learning can be evaluated. From the variety of measures this test provides, we selected the overall score, that is, the T value based on the total number of words recalled during the five learning trials, the

number of words recalled in the first trial, the number of false alarms during the recognition test, and the semantic clustering score.

For an evaluation of executive dysfunction, we report two tests covering distinct aspects. For a subgroup of patients, results from the Modified Card Sorting Test (MCST; Nelson, 1976) were available. Reported is the proportion of perseverative errors. As a measure of basic inhibitory processes, we used the reaction times in a modified version of the Stroop Color–Word Test (Wolfram, Neumann, & Wieczorek, 1986). A more comprehensive evaluation of executive functions was carried out using the Behavioral Assessment of the Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie, & Evans, 1996). From this test, we report the standard score.

The available test scores were used for two analyses: First, we evaluated whether there were systematic differences between the five patient groups with respect to the patients' neuropsychological profiles. Second, we evaluated with correlational analyses the respective contributions of cognitive deficits to the behavior shown in the coherence judgment task.

Relationship Between Lesion Location and Neurospychological Diagnosis

The median values, the arithmetic means, and the ranges of the scores are shown in Table 4 for each of the five groups. One-way ANOVAs with the between-subjects factor of group using the four group contrasts defined above yielded a significant difference for two test scores only. The Token Test results were slightly different between the patient groups: contrast bilateral, $F(1, 13) = 4.8$, $p < .05$. When inspecting the data, it can be seen that this significant result was caused by the fact that patients with bilateral frontal lesions made almost no errors, whereas patients with unilateral left-frontal lesions had more difficulties with this task. However, a more thorough aphasia diagnostic using the AAT had yielded an aphasia classification for only 1 patient. In the Language Comprehension subtest, this patient reached the 88th percentile, so we are confident that the aphasic deficit did not impair performance in the coherence judgment task. Furthermore, the two groups most different with respect to the Token Test, the bilateral and left-sided frontal patients, performed similarly on the coherence judgment task. Thus, the Token Test results, which are indicative of language deficits on the word and sentence level, are not likely to account for the coherence effect particularly evident for these patients.

The second difference was that patients with bilateral frontal lesions made more omission errors in the attention test than did patients in the other groups, leading to significant effects of frontal, $F(1, 17) = 5.2$, $p < .05$, and bilateral, $F(1, 17) = 11.0$, $p < .01$. Once more, this latter effect concerns a difference between those two groups who performed similarly in the coherence judgment task, so it is unlikely that attentional deficits were confounded with difficulties in inferencing.

To confirm this conclusion, while reducing the eventual impact of outliers (which is particularly likely for the TAP results), we conducted additional, nonparametric comparisons of the neuropsychological tests for the patients with left-frontal involvement (Groups LF and BF) with those of the patients without left-frontal lesions (RF, LT, and C). According to Wilcoxon two-sample tests, neither the Token test ($z = 0.90$) nor the omission errors ($z = 0.37$) differed. The only test showing a difference approaching significance was literal fluency. Here, patients with lesions involving left-frontal regions tended to produce fewer words than patients without left-frontal lesions ($z = -1.76$, $p = .08$).

These results show that the groups were well matched with respect to neuropsychological deficits. None of the test results alone could account for the differences found in the coherence judgment data.

Relationship Between Experimental Performance and Neuropsychological Diagnosis

For an evaluation of the relationship between neuropsychological deficits and the performance in the coherence judgment task, we calculated correlations between all neuropsychological test results and four indexes of experimental behavior. Because of the differences in scales and distributions for the various test and experimental variables, nonparametric Spearman rank order correlations were used throughout. From the measures of processing time, the overall means of the judgment and target reading times were taken.[2] For these two measures, the respective means for coherent and incoherent trials were very highly correlated ($r = .95$ for the judgment times; $r = .82$ for the reading times). In contrast, the error rates for coherent and incoherent trials were not at all correlated with each other ($r = .06$), a result consistent with the large interindividual differences in the coherence effect. Thus, for the correlational analysis, the error rates for both coherent and incoherent trials contribute independent information and are therefore considered separately. For each correlation with the neuropsychological test scores, only those participants were included for whom the respective test score was available. The resulting correlation coefficients are shown in Table 5.

The only effect of the demographic characteristics was that education contributed to processing speed—the higher the educational level, the faster the processing times.

The judgment times were related to verbal working memory—the lower the digit span, both forward and backward, the longer the judgment times for the target sentences. The fact that the error rates did not follow this pattern was not unexpected. Simple short-term memory measures, such as the Digit Span subtest, are generally unrelated to reading comprehension (e.g., Daneman & Carpenter, 1980).

We did not expect attentional factors to contribute strongly to inferencing ability. Consistent with this hypothesis was the finding of no reliable correlations between the measures in the TAP test and the errors in the coherence judgment task. In contrast, simple reaction times from the alertness subtest corresponded with reading times and judgment times. Of course, this result is due to the fact that increases in reaction times to simple stimuli also affect the reaction times in the context of a more complex task.

The correlations with measures of verbal long-term memory, executive functions, and language will be discussed in the following section.

Discussion

In this study, we used a straightforward coherence judgment task for evaluating inference processes in brain-injured patients. When asked to indicate whether a pair of visually presented sentences was pragmatically related or not, the entire group of 25 patients had more difficulties with deriving correct inferences than with detecting breaks in coherence. For *yes* answers the error rates were considerably higher than for *no* answers. Moreover, the analysis of the reading times showed that cohesion affected the patients' comprehension in the expected manner: Cohesive ties (i.e., lexical connections) facilitated the comprehension of coherent target sentences, whereas they hindered the comprehension of incoherent target sentences. These results confirm that the task was appropriate for the patients studied, and that despite the group's heterogeneity, even the reading times were sufficiently sensitive to capture the effects of a linguistic factor.

The central question of interest was how lesion location would influence the patients' inferencing ability. Because of previous fMRI results (Ferstl & von Cramon, 2001), we focused on patients with frontal and left-temporal lesions. The main finding was that patients with left-frontal and bifrontal lesions had difficulties with the inferencing task. They made more errors, and these occurred mostly on coherent trials. Thus, it was more difficult for these patients to draw an inference than to detect a coherence gap. This coherence effect was also seen in the analyses of the response times. Before returning to the theoretical and clinical implications of this finding, we first discuss the results for the other patient groups and then evaluate the validity of several alternative explanations.

The good performance of the patients with left-temporal lesions was unexpected, because the imaging data had shown considerable areas of activation in both anterior and posterior portions of the left-temporal lobe. Posterior left-temporal regions are involved in lexical and semantic processing (e.g., Price, 1997), functions that clearly are a prerequisite for text comprehension and inference generation (Kintsch, 1988). However, because of the exclusion of patients with aphasia, we cannot generalize our results to patients with large posterior temporal lesions or severe lexico–semantic deficits. Anterior temporal regions, on the other hand, have been related to language processing in context (Bavelier et al., 1997; Mazoyer et al., 1993), and in particular to conceptual semantics. The fact that patients with lesions in this region performed the inferencing task

---

[2] Here, we used the overall mean in milliseconds rather than the standard scores. The z scores have a mean close to 0 for all participants and are therefore not informative as a measure of performance.

Table 4
Demographic Characteristics and Neuropsychological Test Results for the Five Patient Groups

	Patient group					
Demographics and test scores	Control (n = 5)	Left-temporal (n = 5)	Left-frontal (n = 4)	Bifrontal (n = 4)	Right-frontal (n = 7)	Total (N = 25)
Age (years)						
Mdn	37.0	45.0	45.5	44.0	39.0	41.0
M	38.6	44.0	44.0	44.0	39.9	41.8
Range	24–50	29–58	35–50	26–62	18–50	18–62
Available n	5	5	4	4	7	25
Time since lesion (months)						
Mdn	8.0	14.0	28.0	19.5	21.0	11.0
M	51.0	13.6	33.3	20.5	56.4	37.3
Range	3–229	6–24	6–71	4–39	8–296	3–296
Available n	5	5	4	4	7	25
Education (category)						
Mdn	10.0	10.0	10.0	10.0	10.0	10.0
M	10.0	10.4	10.0	10.0	10.0	10.1
Range	8–12	10–12	8–12	8–12	8–12	8–12
Available n	5	5	4	4	7	25
TAP, Alertness (RTs in ms)						
Mdn	290.0	234.0	277.0	231.0	302.0	287.0
M	283.8	269.2	345.3	242.3	316.0	293.1
Range	210–321	207–437	237–590	200–307	244–420	200–590
Available n	5	5	4	4	7	25
TAP, Divided Attention (no. of misses)						
Mdn	3.0	2.0	1.0	13.0	0.0	2.0
M	2.8	2.6	1.3	9.0	1.6	3.0
Range	2–3	1–4	0–3	1–13	0–5	0–13
Available n	4	5	3	3	7	22
CVLT total score						
Mdn	31.0	45.0	35.0	30.0	29.0	35.0
M	36.2	45.4	32.7	32.3	34.1	36.6
Range	16–63	23–65	25–38	28–39	16–58	16–65
Available n	5	5	3	3	7	23
CVLT, first recall (A1) (no. of words)						
Mdn	6.0	5.0	5.0	5.0	6.0	6.0
M	6.0	6.2	5.7	5.0	5.9	5.8
Range	3–10	4–11	5–7	4–6	2–9	2–11
Available n	5	5	3	3	7	23
CVLT false alarms (no.)						
Mdn	0.0	0.0	0.0	3.0	2.0	0.0
M	0.8	2.6	1.7	4.0	3.0	2.4
Range	0–4	0–13	0–5	3–6	0–8	0–13
Available n	5	5	3	3	7	23
WMS–R Memory Quotient						
Mdn	83.0	86.0	87.0	101.0	74.0	84.0
M	92.0	86.6	85.3	96.0	83.3	87.8
Range	69–117	72–102	63–106	84–103	65–119	63–119
Available n	5	5	3	3	7	23
WMS–R Logical Memory I (no.)						
Mdn	23.0	17.0	20.0	29.0	18.0	22.0
M	26.2	20.0	21.0	28.0	22.4	23.2
Range	14–39	13–28	9–35	21–34	11–39	9–39
Available n	5	5	4	3	7	24
Digit Span forward (raw score)						
Mdn	8.0	7.0	6.5	7.5	6.0	7.0
M	8.2	7.4	7.3	7.8	6.6	7.4
Range	6–11	4–10	4–12	6–10	4–9	4–12
Available n	5	5	4	4	7	25
Digit Span backward (raw score)						
Mdn	6.0	7.0	5.5	6.0	7.0	6.0
M	7.6	7.4	6.5	6.5	6.3	6.8
Range	6–12	5–10	4–11	4–10	3–9	3–12
Available n	5	5	4	4	7	25

Table 4   (*continued*)

Demographics and test scores	Patient group					
	Control (*n* = 5)	Left-temporal (*n* = 5)	Left-frontal (*n* = 4)	Bifrontal (*n* = 4)	Right-frontal (*n* = 7)	Total (*N* = 25)
Semantic fluency, free (no. of words)						
Mdn	29.0	17.0	19.0	18.5	17.0	19.0
M	29.0	18.4	19.0	18.5	19.0	19.9
Range	27–31	16–23	19–19	14–23	10–33	10–33
Available *n*	2	5	2	2	6	17
Semantic fluency, cued (no. of words)						
Mdn	30.5	21.5	21.0	27.0	24.0	23.5
M	30.5	19.8	21.0	27.0	22.8	23.3
Range	30–31	9–27	19–23	18–36	17–26	9–36
Available *n*	2	4	2	2	6	16
Literal fluency (no. of words)						
Mdn	34.0	28.0	19.0	30.5	24.5	25.0
M	39.3	26.4	18.3	30.5	29.7	28.6
Range	33–51	15–35	14–22	15–46	21–49	14–51
Available *n*	3	5	3	2	6	19
Token Test (errors)						
Mdn	1.0	1.0	3.0	0.0	0.0	1.0
M	2.3	0.8	4.0	0.3	0.7	1.7
Range	0–6	0–2	1–9	0–1	0–2	0–9
Available *n*	3	5	4	3	3	18
BADS (standard score)						
Mdn	102.0	108.0	105.0	103.0	98.0	102.0
M	101.4	108.6	103.8	99.5	97.1	101.7
Range	93–113	102–117	97–108	68–124	59–118	59–124
Available *n*	5	5	4	4	7	25
MCST perseverations (prop. of errors)						
Mdn	0.0	0.3	0.0	0.2	0.1	0.0
M	0.0	0.3	0.1	0.2	0.2	0.2
Range	0–0	0–0.6	0–0.3	0.2–0.3	0–0.5	0–0.6
Available *n*	2	2	3	2	6	15
Stroop Test RTs (in milliseconds)						
Mdn	118.5	116.0	112.0	119.0	103.0	114.0
M	134.3	113.0	146.0	121.7	106.6	121.3
Range	90–210	95–128	104–222	117–129	91–129	90–222
Available *n*	4	3	3	3	7	20

Note. Available *n* = number of participants for whom the score was available; TAP = Test of Attentional Processes; RTs = response times; CVLT = California Verbal Learning Test; WMS–R = Wechsler Memory Scale—Revised; BADS = Behavioral Assessment of the Dysexecutive Syndrome; MCST = Modified Card Sorting Test; prop. = proportion.

well suggests that they were able to compensate for their focal lesion by using a distributed network of association cortices.

The performance of patients with right-frontal lesions was well within the normal range, and their response times, despite being elevated, showed the same qualitative pattern across the four conditions as the control group's data. Obviously, this group did not have any difficulties with the task, so our data do not support the hypothesis of an inferencing deficit related to RBD. However, it is important to note that because of the occurrence of nonaphasic language deficits after frontal lobe lesions, we focused on the possible involvement of right-frontal regions to inferencing. In contrast, Beeman (1993, 1998) argued for an RBD deficit after right-temporal lesions, in which inferencing fails because of reduced lexical activation. In our sample, there were 6 patients with lesions reaching into right-temporal areas, but these lesions were not considered for the assignment to the lesion group. To rule out that this way of grouping the participants in our study might have obscured possible

effects of hemisphere, and in particular, the effect of right-temporal brain damage, we calculated mean error rates for the 13 participants who had right-sided frontotemporal lesions, including the 4 bifrontal patients, as well as the 2 participants in the control group with lesions in right-temporal areas. The mean error rate of the resulting RBD group was 7.3%, which is clearly lower than the 9.5% error rate for the remaining 12 patients with exclusively left-sided lesions or without relevant lesions. Alternatively, we excluded both the patients without frontotemporal lesions and the patients with bifrontal lesions (i.e., we considered only those patients with unilateral frontotemporal lesions). The mean error rate for the 9 patients with left-sided lesions was 11.0%, about twice as high as the 5.3% error rate for the 9 patients with right-sided lesions. Taken together, both of these post hoc groupings supported the conclusion that in our sample, which consisted of patients with relatively small lesions, the presence of a left-frontal lesion was more predictive for a deficit in inferencing performance than was RBD.

Table 5
Rank Order Correlations Between Neuropsychological Control Variables and Experimental Variables From the Coherence Judgment Task

Test score	N	Error rate: coherent	Error rate: incoherent	Reading times	Judgment times
Age (years)	25	.16	.28	.17	.16
Time since lesion (months)	25	.29	−.07	.14	−.09
Education (category)	25	−.15	−.25	−.58**	−.53**
CVLT					
Total score	23	−.31	−.37[a]	−.14	−.16
A1	23	−.46*	−.38[a]	−.19	−.22
False alarms	23	.28	.31	.04	.01
Clustering score	22	−.43*	−.22	−.01	.05
WMS–R					
Memory Quotient	23	−.18	−.59**	−.44*	−.30
Logical Memory I	24	−.21	−.51**	−.47*	−.22
Digit Span					
Forward	25	−.26	−.17	−.35	−.57**
Backward	25	−.35	−.28	−.38[a]	−.51**
TAP					
Alertness	25	−.01	.09	.43*	.61**
Divided, misses	22	.20	.30	.13	.15
BADS, standard score	25	−.16	−.33	−.34	−.46*
MCST, perseveration score	15	.26	.56*	.30	.63*
Stroop Test	20	.29	.16	−.15	.16
Semantic fluency, free	17	−.39	−.45[a]	−.09	−.22
Semantic fluency, cued	16	−.37	−.50*	−.39	−.41
Literal fluency	19	−.62**	−.41	−.10	−.37
Token Test	18	.42[a]	.44[a]	.04	.34

Note. The test scores used are the same as those reported in Table 4. CVLT = California Verbal Learning Test; WMS–R = Wechsler Memory Scale—Revised; TAP = Test of Attentional Processes; BADS = Behavioral Assessment of the Dysexecutive Syndrome; MCST = Modified Card Sorting Test.
[a] Approached significance with values of $p < .08$.
* $p < .05$. ** $p < .01$.

Before discussing the left-frontal patients' apparent text-level language deficit, we now consider the possibility of alternative accounts for the data, in particular concurrent neuropsychological deficits.

The first possibility is that the group differences are not causally related to the lesion location, but that they are an artifact of the patient selection and lesion classification. The patient sample is somewhat heterogeneous, including different age groups, neuropsychological profiles, and disorder etiologies. However, we attempted to control these and other factors as carefully as possible. Eventual effects of age and neuropsychological deficits were minimized by ensuring that the lesion groups were comparable, and correlational analyses were used to describe the systematic influence of some of these factors (e.g., memory functions). To control for effects of etiology, we included patients with TBI in all five groups. TBI often causes diffuse axonal injury, rendering frontal pathology likely even in the absence of visible lesions. However, post hoc analyses show that for both the subgroup of TBI patients ($n = 11$), $F(1, 9) = 32.6$, $p < .001$, as well as the subgroup of patients with disorders having other etiologies ($n = 14$), $F(1, 12) = 9.9$, $p < .01$, the presence of a left-frontal lesion significantly predicts the size of the coherence effect on the error rates. Thus, the heterogeneity with respect to etiology only serves to increase the generalizability of our results.

A second alternative explanation is that the results were due to a response bias shown mostly by the patients with left-frontal lesions. Both accuracy and judgment times are off-line measures. A tendency to respond "no" would lead to apparent problems with coherent sentence pairs while yielding better performance on incoherent sentence pairs. However, such a response bias would not account for the effects of cohesion, a factor unconfounded with response. Furthermore, the reading times for the second sentence, which is the dependent variable most prone to capture on-line effects and thus less likely to be influenced by an off-line response strategy, produced a similar pattern of results as the other two measures. In particular, the left- and bifrontal patients' reading times for coherent trials exceeded their times for incoherent trials. And most importantly, while a possible response bias explains the distribution of errors across the two conditions, it cannot account for the group differences in overall error rates. Replacing a bias by a random response strategy would increase the error rates on incoherent trials, but it would not reduce overall error rates. Thus, we can be confident that the left-frontal patients' inferencing deficit was not solely due to a response bias, but that it does reflect text-level comprehension difficulties.

The third possibility is that subtle aphasic deficits of the patients with left-sided lesions might have affected inferencing performance. The exclusion of aphasic patients had

been intended to reduce word- and sentence-level effects, but the diagnostic instruments can never be sufficiently fine-grained for definitely ruling out any residual deficits. And, indeed, there was a trend for a positive correlation of the experimental performance with the errors in the Token Test. However, at the same time, the Token Test was the one control variable for which we found a relevant group difference, and this difference concerned the two groups with increased error rates for coherent trials: Patients with bifrontal lesions had significantly better Token Test results than did patients with unilateral left-frontal lesions, although both groups showed a comparable inferencing deficit. A further argument against an account based on word- and sentence-level deficits is that the left-temporal patients were as likely as the left-frontal patients to have residual aphasic deficits, but they had considerably fewer problems with the coherence judgment task. And finally, there is empirical evidence showing that people with aphasia are able to make use of contextual information and general world knowledge for facilitating their sentence interpretation (e.g., Chapman & Ulatowska, 1989; Pierce, 1988). This context use requires successfully linking the previous discourse information to the current utterance—which is just the inferencing process evaluated in our coherence judgment task.

A fourth possibility is that memory deficits had an impact on inferencing. Of course, it is necessary to encode and retain the contents of the context sentence to establish its link to the target sentence. Correspondingly, there were relationships between verbal working and long-term memory scores with the processing times. More important, the overall memory quotient and the memory for complex verbal material was correlated with the error rates for incoherent trials only, but not those for coherent trials. Thus, successful inferencing requires more than just intact memory for the context sentence. In contrast, two verbal learning measures, recall during the first trial and the semantic clustering score, showed exactly the opposite pattern: They were correlated with the performance in coherent trials, but not in incoherent trials. An interpretation of this result, consistent with the conclusions from our imaging data (Ferstl & von Cramon, 2001), is that the initiation of non-automatic cognitive processes mediates both types of task. To successfully encode a word list, it is advantageous to immediately utilize a semantic clustering strategy, which leads to higher recall scores in the first trial and to higher semantic clustering scores throughout the test. In the coherence judgment task, the inference requires establishing a coherent situation model (i.e., a representation of "what the text is about"). Forming this representation consists of an active integration of the general world knowledge with the two presented sentences (cf. Kintsch, 1998). An initiation of this process becomes unnecessary as soon as a coherence gap is detected.

Finally, it has been suggested that text-level deficits are an epiphenomenon of executive dysfunction, which in turn is a symptom of frontal lobe damage (cf. Kaczmarek, 1987). Because the concept of executive functions comprises a variety of subcomponents and because the exact relation-

ship between executive dysfunction and frontal pathology is somewhat unclear, there is no single diagnostic test available (e.g., Boone, 1999; Burgess, Alderman, Emslie, & Wilson, 1998; Ettlin et al., 2000). Consequently, in our sample none of the test scores discriminated patients with and without frontal lesions. Nevertheless, several of the tests related to executive functions were significantly correlated with inferencing performance. The standard score from the BADS was moderately correlated with processing times, but not with the error rates. However, this test provided even a hint of a deficit (less than 95 points) for only 5 patients, so the range of test scores might not have been sufficient. Furthermore, it seems likely that text-level language processes are a component of executive functions that cannot be simply explained by more basic subprocesses (cf. Ettlin et al., 2000). The perseveration score from the MCST was related to the judgment times and to the errors in incoherent trials. Although this result is based on a small sample size and therefore needs to be considered with caution, it suggests that a lack of cognitive flexibility impairs the detection of coherence gaps. In incoherent trials, it is necessary to establish the representation of a new situation described by the target sentence, instead of updating the existing situation model representing the context sentence. Abandoning this prior situation model on the basis of externally presented information might be difficult for patients with a perseverative tendency.

The second set of executive function tests that proved to be related to inferencing performance were the verbal fluency measures. These are considered tests of frontal lobe function because they require the use of search strategies (Boone, 1999). The three scores considered here differ with respect to the amount of external guidance (cf. Crosson et al., 2001). Consistent with the previous interpretations, the errors in the incoherent trials were most strongly correlated with semantic fluency with category cues, a test in which external stimuli aid the search strategy. In contrast, the errors in coherent trials were most strongly correlated with the literal fluency score. Literal fluency is the test with the highest demand on the initiation and utilization of self-guided retrieval strategies.

In summary, these results show that the coherence judgment task tapped cognitive processes also reflected in neuropsychological test results. However, deficits in memory, sentence processing, or executive functions alone could not account for the group differences in inferencing performance. Moreover, the correlation analysis also confirmed that the errors in coherent and incoherent trials reflected different aspects of the coherence judgment performance. Thus, it is valid to conclude that patients whose lesions reached into the left-frontal lobe had difficulties with providing the correct inference, rather than showing a general performance decrement.

Taken together, the results of the present study provide further evidence for the claim that nonaphasic language deficits are related to prefrontal lesions, but that they are not necessarily concurrent with right-lateralized brain damage. In support of this claim, recent studies have shown that RBD patients do not always have problems with text-level

tasks (e.g., Leonard & Baum, 1998; Leonard et al., 1997a, 1997b). At the same time, there are studies documenting left-frontal patients' deficits for processing language in context (Channon & Crawford, 2000; Ferstl, Guthke, & von Cramon, 1999; Kaczmarek, 1984; Novoa & Ardila, 1987). Similarly, in addition to brain imaging studies supporting the right hemisphere hypothesis for text comprehension tasks (Robertson et al., 2000; St. George et al., 1999), there is increasing evidence for the importance of frontomedian and left-lateral prefrontal regions (Ferstl & von Cramon, 2001; Fletcher et al., 1995; Maguire et al., 1999).

For a theoretical interpretation of these apparently contradictory results, two main issues need to be considered (cf. Brownell & Martino, 1998; McDonald, 1993; Stemmer, 1999). First, the selection of appropriate patient and control groups is crucial. In our study, we based the grouping on whether a frontal or left-temporal lesion was present, rather than on behavioral factors; we carefully described lesion location within the hemispheres; and we included a brain-injured control group. Thus, it was possible to directly compare lesion groups rather than to compare a patient group to a healthy control group. This latter method tends to overestimate specific deficits because of the co-occurrence of unspecific performance impairments. Second—and this issue holds for both patient and imaging studies—the task analysis is important and often neglected. The concepts of *inferencing* or *text comprehension* subsume a number of subprocesses that vary with respect to their level of complexity, to their postulated automaticity, and to the type of information to be processed (cf. Gernsbacher, 1990; Kintsch, 1998). All of these factors have been drawn on to develop theoretical accounts of RH language deficits, but they cannot be readily adopted for making predictions about our coherence judgment task. The inferences require the nonautomatic use of general world knowledge, going beyond the activation and selection of associations (Beeman 1993, 1998), but they do not require taking into account nonliteral, pragmatic, affective, or social information (e.g., Alexander, Benson, & Stuss, 1989; Brownell et al., 1995).

In contrast to the variety of theories on right hemisphere inferencing and text comprehension deficits, less is known about the role of left hemisphere, and in particular, left prefrontal regions. However, it is undisputed that even in the absence of aphasia, these patients often exhibit symptoms of nonaphasic language production disorders (e.g., Kaczmarek, 1984). The results presented here add to the growing literature documenting problems of this population with receptive text-level tasks (Channon & Crawford, 2000; Ferstl et al., 1999). Accumulating evidence from imaging studies and clinical descriptions enables us to sketch a model of the specific role of the left-frontal lobe for comprehension processes. The overarching claim is that left-prefrontal regions are involved whenever the text comprehension task requires the use of nonautomatic, self-guided cognitive processes. This description encompasses executive functions as needed during processing of complex verbal material (e.g., structuring, sequencing, inhibition of inappropriate interpretations, goal-oriented language use, or integration of information), presumably realized in lateral prefrontal brain regions. In addition, the description encompasses a domain-independent set of functions, realized in frontomedian brain regions, that are related to the self-guided initiation and maintenance of the former strategic processes. To further strengthen this interpretation, additional patient studies are needed investigating the text-processing skills of patients with left-frontal brain damage.

Independent of the specific cause, the most important contribution of our study is to bring attention to comprehension deficits after left-frontal brain damage. The coherence judgment task as used in this study was appropriate for brain-injured patients and sufficiently sensitive for uncovering a specific inferencing deficit in patients with left- and bifrontal lesions. It is clear that neglecting coherence must have a detrimental effect on communication. Clinical observations from discourse production show that patients with nonaphasic language disorders often carry on with their own topics instead of responding appropriately to the conversation partner or that they tend to be insensitive to verbal or nonverbal cues to turn-taking. These deficits could be explained by the failure of sufficiently taking into account the context. Because communication skills are crucial for successful rehabilitation (cf. Brooks, McKinlay, Symington, Beattie, & Campsie, 1987), the diagnosis of language functions must not be restricted to aphasia testing. Particularly for frontal lobe patients, independent of lesion lateralization, a thorough assessment of text comprehension deficits is indispensable and should include the evaluation of inferencing abilities.

References

Alexander, M. P., Benson, D. F., & Stuss, D. T. (1989). Frontal lobes and language. *Brain and Language, 37,* 656–691.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V. P., Karni, A., et al. (1997). Sentence reading—A functional MRI study at 4 tesla. *Journal of Cognitive Neuroscience, 9,* 664–686.

Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain and Language, 44,* 80–120.

Beeman, M. (1998). Coarse semantic coding and discourse comprehension. In M. Beeman & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 255–284). Mahwah, NJ: Erlbaum.

Boone, K. B. (1999). Neuropsychological assessment of executive functions. In B. L. Miller & J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (pp. 247–260). New York: Guilford.

Brooks, N., McKinlay, W., Symington, C., Beattie, A., & Campsie, L. (1987). Return to work within the first seven years of severe head injury. *Brain Injury, 1,* 5–19.

Brownell, H. H., Gardner, H., Prather, P., & Martino, G. (1995). Language, communication and the right hemisphere. In H. S. Kirshner (Ed.), *Handbook of neurological speech and language disorders* (pp. 325–349). New York: Dekker.

Brownell, H. H., & Martino, G. (1998). Deficits in inference and social cognition: The effects of right hemisphere brain damage on discourse. In M. Beeman & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 309–328). Mahwah, NJ: Erlbaum.

Burgess, P. W., Alderman, N., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive functions. *Journal of the International Neuropsychological Society, 4,* 547–558.

Channon, S., & Crawford, S. (2000). The effects of anterior lesions on performance on a story comprehension test: Left anterior impairment on a theory of mind-type task. *Neuropsychologia, 38,* 1006–1017.

Chapman, S. B., & Ulatowska, H. K. (1989). Discourse in aphasia: Integration deficits in processing reference. *Brain and Language, 36,* 651–668.

Coelho, C. A., Liles, B. Z., & Duffy, R. J. (1995). Impairments of discourse abilities and executive functions in traumatically brain-injured adults. *Brain Injury, 9,* 471–477.

Crosson, B., Sadek, J. R., Maron, L., Gökcay, D., Mohr, C. M., Auerbach, E. J., et al. (2001). Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation. *Journal of Cognitive Neuroscience, 13,* 272–283.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Delis, D. C., Kramer, J. H., Kaplan, E., & Obler, B. A. (1987). *The California Verbal Learning Test: Adult Version.* San Antonio, TX: Psychological Corporation.

De Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex, 14,* 41–49.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C., & von Cramon, D. Y. (2000). *Cognitive Brain Research, 9,* 103–109.

Ettlin, T. M., Kischka, U., Beckson, M., Gaggiotti, M., Rauchfleisch, U., & Benson, D. F. (2000). The Frontal Lobe Score, Part I: Construction of a mental status of frontal systems. *Clinical Rehabilitation, 14,* 260–271.

Ferstl, E. C., Guthke, T., & von Cramon, D. Y. (1999). Change of perspective in discourse comprehension: Encoding and retrieval processes after brain injury. *Brain and Language, 70,* 385–420.

Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: An event-related fMRI study. *Cognitive Brain Research, 11,* 325–340.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57,* 109–128.

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Glosser, G., & Deser, T. (1990). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language, 40,* 67–88.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* London: Longman.

Horn, W. (1983). *Leistungsprüfsystem (LPS)* [Performance Assessment System]. Göttingen, Germany: Verlag für Psychologie.

Huber, W., Poeck, K., & Willmes, K. (1984). The Aachen Aphasia Test. In F. C. Rose (Ed.), *Advances in neurology: Progress in aphasiology* (Vol. 42, pp. 291–303). New York: Raven Press.

Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model comparison approach.* San Diego, CA: Harcourt Brace Jovanovich.

Kaczmarek, B. L. J. (1984). Neurolinguistic analysis of verbal utterances in patients with focal lesions of frontal lobes. *Brain and Language, 21,* 52–58.

Kaczmarek, B. L. J. (1987). Regulatory function of the frontal lobes. A neurolonguistic perspective. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 225–240). Hillsdale, NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95,* 163–182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, England: Cambridge University Press.

Lehman, M. T., & Tompkins, C. A. (2000). Inferencing in adults with right hemisphere brain damage: An analysis of conflicting results. *Aphasiology, 14,* 485–499.

Leonard, C. L., & Baum, S. R. (1998). On-line evidence for context use by right-brain damaged patients. *Journal of Cognitive Neuroscience, 10,* 499–508.

Leonard, C. L., Waters, G. S., & Caplan, D. (1997a). The use of contextual information by right brain-damaged individuals in the resolution of ambiguous pronouns. *Brain and Language, 57,* 309–342.

Leonard, C. L., Waters, G. S., & Caplan, D. (1997b). The use of contextual information related to general world knowledge by right brain-damaged individuals in pronoun resolution. *Brain and Language, 57,* 343–359.

Lock, S., & Armstrong, L. (1997). Cohesion analysis of the expository discourse of normal, fluent aphasic and demented adults: A role in differential diagnosis? *Clinical Linguistics & Phonetics, 11,* 299–317.

Maguire, E. A., Frith, C. D., & Morris, R. G. M. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain, 122,* 1839–1850.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience, 5,* 467–479.

McDonald, S. (1993). Viewing the brain sideways? Frontal versus right hemisphere explanations of non-aphasic language disorders. *Aphasiology, 7,* 535–549.

Münte, T. F., Schiltz, K., & Kutas, M. (1998, September 3). When temporal terms belie conceptual order. *Nature, 395,* 71–73.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12,* 313–324.

Novoa, O. P., & Ardila, A. (1987). Linguistic abilities in patients with pre-frontal damage. *Brain and Language, 30,* 206–225.

Pierce, R. S. (1988). Influence of prior and subsequent context on comprehension in aphasia. *Aphasiology, 2,* 577–582.

Price, C. J. (1997). Functional anatomy of reading. In R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, & J. C. Mazziotta (Eds.), *Human brain function* (pp. 301–328). San Diego, CA: Academic Press.

Prigatano, G. P., Roueche, J. R., & Fordyce, D. J. (1986). Nonaphasic language disturbances after brain injury. In G. P. Prigatano, D. J. Fordyce, H. K. Zeiner, J. R. Roueche, M. Pepping, & B. C. Wood (Eds.), *Neuropsychological rehabilitation after brain injury.* Baltimore: John Hopkins University Press.

Robertson, D. A., Gernsbacher, M. A., Guidotti, S. J., Robertson, R. R. W., Irwin, W., Mock, B. J., & Campana, M. E. (2000). Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science, 11,* 255–260.

Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479–515). San Diego, CA: Academic Press.

Snow, P., Douglas, J., & Ponsford, J. (1997). Conversational assessment following traumatic brain injury: A comparison across two control groups. *Brain Injury, 11,* 409–429.

Stemmer, B. (1999). Discourse studies in neurologically impaired populations: A quest for action. *Brain and Language, 70,* 402–418.

St. George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: Engagement of the right hemisphere during discourse processing. *Brain, 122,* 1317–1325.

Wechsler, D. (1987). *Wechsler Memory Scale—Revised.* New York: Psychological Corporation.

Wilson, B., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome.* Bury St. Edmunds, England: Thames Valley Test Company.

Wolfram, H., Neumann, J., & Wieczorek, V. (1986). *Psychologische Leistungstests in der Neurologie und Psychiatrie* [Psychological performance tests for neurology and psychiatry]. Leipzig, Germany: Thieme.

Zimmermann, P., & Fimm, B. (1992). *Testbatterie zur Aufmerksamkeitsprüfung (TAP)* [Test Battery for the Assessment of Attention]. Würselen, Germany: Psychologische Testsysteme.

Zysset, S., Müller, K., Lohmann, G., & von Cramon, D. Y. (2000). Color word matching Stroop Task: Separating interference and response conflict. *Neuroimage, 13,* 29–36.

Online Preview of *Neuropsychology* Articles

Are you an APA member or affiliate who subscribes to *Neuropsychology*? If so, you have online access to the most recently accepted articles before they appear in print. Articles accepted and scheduled for publication are available in the PsycARTICLES database (in draft form) at least 2 months prior to print publication. Access to this feature is available at *no charge* via http://www.apa.org/journals/neu.html to APA members and affiliates who subscribe to *Neuropsychology.*