Evelyn C. Ferstl: The Functional Neuroanatomy of Text Comprehension.
Leipzig: Max Planck Institute for Human Cognitive and
Brain Sciences, 2006 (MPI Series in Human Cognitive and Brain
Sciences; 74)

The Functional Neuroanatomy of Text Comprehension

The Functional Neuroanatomy
of Text Comprehension

Der Fakultät für Biowissenschaften, Pharmazie und Psychologie
der Universität Leipzig
eingereichte

HABILITATIONSSCHRIFT

zur Erlangung des akademischen Grades

DOCTOR PHILOSOPHIAE HABILITATUS
(Dr. phil. habil.)

vorgelegt von

Dr. Evelyn Christina Ferstl

geboren am 29. April 1962 in München

Gutachter

Prof. Dr. D. Yves von Cramon

Prof. Dr. Franz Schmalhofer

Prof. Dr. Jörg Jescheniak

Bibliographische Beschreibung

Dr. Evelyn C. Ferstl

The functional neuroanatomy of text comprehension
[Die funktionelle Neuroanatomie des Textverstehens]

Universität Leipzig, Habilitationsschrift

99 Seiten, 15 Abbildungen, 259 bibliographische Referenzen

Referat:

Das Verstehen von Sprache im pragmatischen und situativen Kontext erfordert kognitive Leistungen, die weit über die Analyse von Einzelwörtern und –sätzen hinausgehen. In der psycholinguistischen Forschung werden vor allem Inferenz- und Situationsmodellbildung als für das Textverstehen wesentlich betrachtet. Diese beiden Prozesse dienen der lokalen und globalen Kohärenzbildung und ermöglichen eine integrierte Textrepräsentation, die sowohl Vorwissen und Kontext als auch die neu verarbeitete sprachliche Information enthält.

In dieser Arbeit wurde die funktionelle Neuroantomie dieser Prozesse untersucht. Mit einer Kombination aus behavioralen, experimentellen Studien, Läsionsstudien und funktioneller Magnetresonanztomographie (fMRT) sollten die Hirnregionen beschrieben werden, die beim Textverstehen beteiligt sind. Basierend auf neuropsychologischen Theorien und klinischer Beobachtung war vor allem die Rolle der rechten Hemisphäre und des Frontallappens von Interesse.

Eine Reihe von behavioralen Experimenten zeigten für eine einfache Kohärenzentscheidungsaufgabe, dass sie unabhängig vom Alter der Rezipienten ist und auch von Patienten mit rechtshemisphärischen und links-temporalen Läsionen gut bewältigt werden kann. Im Gegensatz zu diesen Patientengruppen hatten hirngeschädigte Personen, deren Läsion in den linken Frontallappen reichte, deutliche Einbussen. In einer bildgebenden Studie konnte die Beteiligung lateraler und vor allem medianer linksseitig präfrontaler Gebiete bestätigt werden. Dass die medianen Aktivierungen nicht nur durch gleichzeitig ausgelöste Theory-of-Mind Prozesse ausgelöst wurden, wurde in einer weiteren fMRT-Studie bestätigt.

Situationsmodellbildung wurde mit dem Inkonsistenz-Paradigma untersucht, bei dem lokal kohärente Geschichten mit global Inkonsistenzen versehen werden. Die Fehler, die entweder emotionale oder zeitliche Information betrafen, lösten Verständnisschwierigkeiten aus. Die fMRT-Daten legten eine Funktion des rechten Temporallappens für die Verarbeitung der sprachlichen Inkonsistenzen nahe. Ausserdem bewirkten sie eine qualitative Trennung der Informationsarten. Emotionale Information löste Aktivierung im ventro-medianen präfrontalen Kortex aus, während die Verarbeitung zeitlicher Information laterale fronto-parietale Regionen beteiligte. Die Integration der emotionalen Information war in einer ähnlichen dorsalen fronto-medianen Region sichtbar, wie die, die auch zu lokaler Kohärenzbildung beiträgt.

Diese Studien bestätigen die Funktion des sogenannten *erweiterten Sprachnetzwerkes* für das Textverstehen. Fronto-temporale Gebiete jenseits des perisylvischen Sprachkortex sind konsistent aktiviert. Wichtig sind neben den erwähnten fronto-medianen Arealen vor allem lateral präfrontale Areale und der anteriore Temporallappen. Für die klinische Diagnostik und Therapie von erworbenen Kommunikationsstörungen nach Hirnschädigung sind diese Ergebnisse von grosser Bedeutung. Sie unterstreichen die Notwendigkeit, auch bei nichtaphasischen Patienten, - vor allem nach traumatischer Hirnschädigung oder Frontalhirnläsionen -, die Kommunikationsfähigkeiten auf Text- und Diskursebene zu berücksichtigen.

Vorwort

Acknowledgments

Nach meiner Dissertation hatte ich schon viele Berufe. Je nach Anlass schlüpfte ich in die Rolle der Mathematikerin, der Linguistin oder der Psychologin. Interdisziplinarität bedeutete für mich damals, dass ich überzeugt war, mich in vielen verschiedenen Bereichen so einigermassen gut auszukennen. Meine Lehrzeit als Neurowissenschaftlerin und Neurolinguistin hat mir gezeigt, dass Interdisziplinarität erst da anfängt, wo ein Kopf allein nicht mehr ausreicht. Das Max-Planck-Institut für Kognitions- und Neurowissenschaften und die Tagesklinik für kognitive Neurologie in Leipzig boten ein aussergewöhnliches Umfeld für einen echten Austausch, ein echtes Zusammenführen von Fähigkeiten und Kompetenzen aus ganz unterschiedlichen Feldern.

Danken möchte ich also den zahlreichen Menschen, die auf so vielfältige Weise an meiner Arbeit Anteil hatten. Mein ganz besonderer Dank gilt Herrn Prof. Dr. von Cramon, der mir sowohl die klinische Welt, als auch die der Bildgebung als Forschungsgebiet eröffnet – und kontinuierlich offen gehalten - hat. Ohne seine Unterstützung und seine neuroanatomische Expertise wäre meine Forschung nicht möglich gewesen.

Für die intensive inhaltliche Zusammenarbeit danke ich allen Kolleginnen und Kollegen am MPI, - dort vor allem den Mitstreitern in der EXEC-Gruppe -, und in der Tagesklinik, - dort insbesondere den SprachtherapeutInnen. Besonders wichtig waren für mich die Text-Kollegen. Allen voran danke ich Thomas Guthke für viele inhaltliche Diskussionen, pragmatische Lösungsvorschläge und klinische Einsichten. Florian Siebörger hat neuen Schwung in unser Textgrüppchen gebracht und mich oft dazu verleitet, nochmal anders nachzudenken. Mike Rinck und Franz Schmalhofer danke ich für ihre Kooperation.

Nicht zu vergessen sind die StudentInnen, die zahlreichen PraktikantInnen und DiplomandInnen, die an verschiedenen Projekten mitgewirkt haben. Hier nenne ich stellvertretend Katrin Walther, eine meiner ersten Praktikantinnen, und Carsten Bogler, der mir als studentische Hilfskraft immer tatkräftig zur Hand gegangen ist.

Massgeblich unterstützt wurde meine Arbeit von MitarbeiterInnen in der Graphik, der Lipsia-Gruppe und der EDV, der Bibliothek, dem Reaktionszeitlabor, und von den MTAs und Physikern im fMRT-Labor. Stellvertretend erwähne ich hier besonders Stephan Liebig und vor allem Gabi Lohmann.

Aber auch das beste Team ist nicht erfolgreich ohne die Anfeuerungen seiner Fans. Daher danke ich – *last but not least* – meinem grössten Fan, Max: für die Begeisterung, sogar wenn mal ein Spiel verloren ging, für die Geduld bei der Nachbesprechung aller Partien, und für das Vertrauen, dass die Championsleague nicht weit ist...

Leipzig, September 2005

Inhaltsverzeichnis

Table of Contents

PART I:

Summary in German
Zusammenfassung auf Deutsch

Theoretischer Hintergrund

Aphasische und nicht-aphasische Sprachstörungen

Während die Erforschung von aphasischen Sprachstörungen schon auf eine hundertjährige Geschichte zurückblicken kann, sind *nicht-aphasische Kommunikationsstörungen* (NAKS) erst seit Anfang der 80er Jahre in das Blickfeld neuropsychologischer Forschung gerückt. Nicht-aphasische Sprachstörungen werden dabei als Defizite der verbalen Kommunikation verstanden, die *nicht* bedingt sind durch linguistische Defizite auf phonologischer, syntaktischer und semantischer Ebene, sondern erst auf der Ebene zusammenhängender Sprache, also auf Text- oder Diskursebene sichtbar werden (vgl. Glindemann & von Cramon, 1995).

Aphasien sind hauptsächlich Folge von Läsionen der sprach-dominanten, meist linken Hemisphäre. Von NAKSs betroffen sind ätiologisch überwiegend Schädel-Hirn-Traumatiker (SHT), und lokalisatorisch bevorzugt Patienten mit Frontalhirnläsionen oder rechts-hemisphärischen Läsionen. Die freie Sprachproduktion im Kontext zeigt bei NAKS vielfältige Symptome, die von reduzierter zu überschüssiger Sprache reichen, und von inkohärenter, über unstrukturierte zu inhaltlich redundanter Redeweise (Chapman, Levin, & Culhane, 1995; Hartley & Jensen, 1992; Glindemann & von Cramon, 1995; Ferstl & Guthke, 1998); die Kommunikation schlägt fehl, weil nicht auf das Vorwissen des Gesprächspartners geachtet wird oder eine für die soziale Situation unangemessene Sprachebene gewählt wird. Vor allem im beruflichen Umfeld kann dies zu erheblichen Problemen führen. Für eine erfolgreiche Rehabilitation ist somit die Diagnostik und Therapie dieser Störungen unerlässlich.

Psycholinguistik des Textverstehens

Obwohl anzunehmen ist, dass sich NAKSs auch auf rezeptive Sprachleistungen in ähnlicher Weise auswirken, sind Textverstehensprozesse der betroffenen

Patienten noch relativ wenig untersucht. *Text* wird dabei ganz allgemein definiert als *Sprache im Kontext*, die über einzelne Wörter oder Sätze hinausgeht, d. h. als zusammenhängende Sprachäusserung mit einer kommunikativen Intention. Da zu deren Verstehen und Interpretation eine Vielzahl von Teilleistungen zusammenspielen, herrscht in der Literatur Konsens darüber, dass eine genaue begriffliche Einordnung der untersuchten Teilprozesse nötig ist. Wir beziehen uns hierzu auf die Theorie von Kintsch und van Dijk (1978; van Dijk & Kintsch, 1983; Kintsch, 1998; Graesser, Millis & Zwaan, 1997). Sie postulieren u.a. zwei kognitive Prozesse als wesentlich, um aus Einzelsätzen ein sinnvolles Ganzes zu erstellen: die *Kohärenz*- und die *Situationsmodell*-Bildung.

Kohärenz wird durch sogenannte *Inferenzen* erreicht. Mit Inferenz bezeichnet man den kognitiven Prozess, der unter Einbeziehen des Kontexts, des Weltwissens und des aktuell Gehörten die Verbindung zwischen den Einzelsätzen herstellt (Graesser, Singer & Trabasso, 1994; Singer, 1994; Beyer, Guthke & Pekrul, 1996). Ein Beispiel ist der kurze Text: *Charlotte ist bei Max zum Geburtstag eingeladen. Sie schlachtet ihr Sparschwein*. Der Leser wird eine kausale und eine intentionale Beziehung inferieren: Charlotte schlachtet das Sparschwein, weil sie Geld braucht, um ein Geschenk zu kaufen, das sie Max mitbringen will – eine Schlussfolgerung, die nur getroffen werden kann, wenn das Weltwissen über übliche Geburtstagsfeiern in die geschilderte Situation mit eingebracht wird.

Während Inferenzen oft der lokalen Kohärenz dienen, also dem Sinnzusammenhang zwischen aufeinanderfolgenden Äusserungen, handelt es sich bei der *Situationsmodellbildung* (van Dijk & Kintsch, 1983; Ferstl & Kintsch, 1999; Kintsch, 1998; Ferstl, 2001; Zwaan, 2004, Rinck, 2000) um einen Prozess, der eine globale Repräsentation des gesamten Textes erstellt. Diese Repräsentation bildet die Basis für weiterführende Nutzung der gelernten Textinformation (Kintsch, 1998; Ferstl, 2001). Auch hier ist der Abruf von Weltwissen nötig, um die präsentierte Information in das Vorwissen des Lesers oder der Hörerin einzubinden. In dem obigen, sehr kurzen Beispiel entspricht das Situationsmodell dem Schema über Kindergeburtstage. Für Erzähltexte im allgemeinen ist bekannt, dass das Situationsmodell u. a. zeitliche, räumliche, kausale und emotionale Aspekte einer Geschichte enkodiert (Rinck, Hähnel & Becker, 2001; Zwaan, Langston & Graesser, 1995; Zwaan, Magliano & Graesser, 1995).

Textverstehen nach Hirnschädigung

Während es Hinweise darauf gibt, dass aphasische Patienten ihr Sprachverständnis in geeignetem Kontext mithilfe von Weltwissen verbessern können (Pierce, 1988; Germani & Pierce, 1992; Chapman & Ulatowska, 1989), werden Probleme im Bereich des Textverstehens häufig der rechten Hemisphäre (RH) zugeschrieben (s. Brownell, Gardner, Prather, & Martino, 1995; Brownell & Martino, 1998; Beeman, 1998). Mit der Inferenzbildung haben RH-Patienten vor allem dann Schwierigkeiten, wenn soziale, situative oder kommunikative Implikationen berücksichtigt werden müssen. In jüngerer Zeit wurde jedoch vermehrt darauf hingewiesen, dass die empirische Evidenz keineswegs eindeutig ist (cf. Lehmann & Tompkins, 2000), dass die Läsionsorte manchmal ungenügend spezifiziert sind, um Konfundierungen auszuschließen (McDonald, 1993) oder dass qualitativ unterschiedliche Teilprozesse (z.B. Verstehen von Ironie vs. Erkennen einer Schlußfolgerung) mit dem gleichen Begriff *Inferenz* bezeichnet werden. Daraus ergibt sich ein weiterhin erhöhter Forschungsbedarf. Eine offene Frage ist, welche RH-Läsionen Inferenzdefizite verursachen. Eine alternative Erklärung postuliert, dass frontale Läsionen, unabhängig von ihrer Lateralisierung, Inferenzdefizite bedingen, die dann als Symptome einer allgemeineren Exekutivfunktionsstörung gesehen werden (Alexander, Benson & Stuss, 1989, Fuster, 1997, 1999; Novoa & Ardila, 1987).

Funktionelle Bildgebung von Textverstehensprozessen

Die Beschreibungen der aphasischen und nicht-aphasischen Defizite sowie eine Reihe von Läsionsstudien legen nahe, dass die für das Textverstehen nötigen Leistungen nicht im Sprachkortex der dominanten Hemisphäre realisiert sind, sondern rechts-hirnige und/oder präfrontale Regionen eine Rolle spielen. Um diese neuroanatomischen Hypothesen zu untersuchen und weiter zu spezifizieren, ist besonders die funktionelle Bildgebung geeignet. Jedoch sind bislang erst wenige Studien veröffentlicht, die spezifische Fragestellungen aus diesem Bereich angehen (Maguire, Frith & Morris, 1999; Caplan & Dapretto, 2001; St. George et al., 1999; Robertson et al., 2000; for reviews see Mar, 2004, Bookheimer, 2002; Gernsbacher & Kaschak, 2003). Aus Experimenten mit hirngesunden Probanden, in denen Sprache im Kontext präsentiert wurde, ist jedoch ersichtlich, dass wir ein erweitertes Sprachnetzwerk betrachten müssen (Ferstl, im Druck, a; Ferstl & von Cramon, im Druck). Neben dem perisylvischen Sprachkortex in der dominanten Hemisphäre spielen vor allem anterior temporale Areale beidseits, sowie der linksseitige fronto-dorsale Kortex und fronto-mediane Kortex (BA 8/9/10) eine besondere Rolle (Mazoyer et al., 1993; Xu et al., 2005; Bavelier et al., 1997; Crinion et al., 2003). Außerdem wurde vorgeschlagen, dass die Region des posterioren cingulären Kortex und unteren Präcuneus (PCC/prec) an der Situationsmodellbildung beteiligt sei.

Inwieweit auch die rechte Hemisphäre deutlichen Anteil hat, ist auch aus den Ergebnissen der Bildgebung noch nicht eindeutig zu ersehen. Das erweiterte Sprachnetzwerk ist eindeutig links-dominant, vor allem für anterior temporale und laterale präfrontale Areale wird jedoch oft eine Beteiligung der rechtsseitig homologen Areale berichtet. Darüberhinaus gibt es eine Reihe von Studien, die spezifische RH-Aktvierung für Textverständnisaufgaben nachweisen (St. George et al., 1999; Robertson, et al., 2000; Nichelli et al., 1995; Bottini et al., 1994); andere Studien mit teils sehr ähnlichen Fragestellungen betonen jedoch die Funktionen des linksseitigen fronto-temporale Kortex (Fletcher et al., 1995; Maguire et al., 1999, Rapp et al., 2004).

Empirische Studien

Die in dieser Arbeit vorgestellten empirischen Studien befassen sich mit den zwei vorgestellten Teilprozessen des Textverstehens. Zuerst wird ein Paradigma zur Kohärenzbildung vorgestellt, das für die Nutzung in Patienten- und Bildgebungsstudien entwickelt wurde. Dieses Paradigma erlaubt, die Interaktion zwischen lexikalischer und pragmatischer Information während Inferenzprozessen zu untersuchen. Kapitel 2 und 3 legen die behavioralen Grundlagen. Hier wird gezeigt, dass die Aufgabe sowohl für studentische als auch für ältere Versuchspersonen ohne Mühe bewältigt werden kann, und dass die gewünschten Effekte der lexikalischen Information auf die Inferenzbildung messbar sind. Kapitel 4 hat eine Patientenstudie zum Thema, in der der Einfluß des linken präfrontalen Kortex auf die Kohärenzbildung deutlich wird. Kapitel 5 berichtet über eine Studie, in der mittels funktioneller Magnetresonanztomographie (fMRT) zusätzlich Teilregionen differenziert werden. Der laterale präfrontale Kortex wird für die Integration inkonsistenter lexikalischer Information benötigt, während der dorso-mediane präfrontale Kortex die pragmatische Kohärenzbildung unterstützt. In Kapitel 6 wird ein weiteres fMRT-

Experiment besprochen, in dem eine alternative Erklärung für die in Kapitel 5 berichtete Aktivierung des dorso-medianen präfrontalen Kortex (dmPFC) ausgeschlossen wird. Schließlich stellt Kapitel 7 ein Experiment vor, in dem die Situationsmodellbildung in global kohärenten Texten untersucht wird. Neben der Konsistenz der Information werden zwei qualitativ unterschiedliche Informationsaspekte untersucht.

Kapitel 2:
Kohärenzbildung: Lexikalische und pragmatische Information

Um satzübergreifende Verarbeitung zu untersuchen, aber gleichzeitig größtmögliche Kontrolle über die beteiligten Teilprozesse zu behalten, wurde eine einfache Aufgabe entwickelt. Bei dem sogenannten Kohärenzparadigma werden „minimale" Texte verwendet, die nur aus zwei Sätzen bestehen (dem *Kontextsatz* und dem *Targetsatz*). Die Aufgabe ist, mittels einer JA/NEIN-Entscheidung anzugeben, ob die beiden Sätze in inhaltlichem Zusammenhang stehen oder nicht. Im ersten Experiment (**Kapitel 2;** Ferstl & von Cramon, 2001a; *Cognitive Brain Research*) wurden 120 kohärente Satzpaare verwendet, die durch eine erklärende Inferenz zu verbinden waren. In den meisten Fällen gelang dies nicht schon durch einfache assoziative Verknüpfung, sondern es war nötig, zusätzliches Weltwissen einzubeziehen. Die nicht-kohärenten Satzpaare wurden dadurch erzeugt, dass die Kontextsätze von je zwei Satzpaaren ausgetauscht wurden. Beispiele für zwei kohärente Satzpaare und die sich daraus ergebenden nicht-kohärenten Durchgänge werden in Tabelle 1 in der ersten Zeile gezeigt.

Als zweiter Faktor wurde die lexikalische *Kohäsion* verwendet (Halliday & Hasan, 1976). In kohäsiven Sätzen gibt es ein spezifisches Wort (z. B. Pronomen, Konjunktionen, Wiederholungen), das sich explizit auf das Vorhergehende bezieht. Diese sogenannten *Kohäsionsmarkierungen* haben unmittelbare Auswirkungen auf die Interpretation und erleichtern die inhaltliche Inferenz (vgl. Robertson, Gernsbacher, Guidotti, et al., 2000; Münte, Schiltz, & Kutas, 1998; Halliday & Hasan, 1976; Schwarz, 2000). Im Gegensatz dazu wird die Verarbeitung von nicht-kohärenten Inhalten durch falsch eingesetzte Kohäsionsmarkierungen erschwert. Durch Kreuzen der beiden Faktoren Kohärenz und Kohäsion kann somit die Interaktion zwischen lexikalischer und pragmatischer Verarbeitung, also zwischen der Wort- und der Textebene untersucht werden.

Beispiele für die kohäsiven Versionen der Satzpaare sind in der zweiten Zeile von Tabelle 1 enthalten.

Tabelle 0.1. Satzbeispiele der vier Bedingungen.

	kohärent	**nicht kohärent**
nicht kohäsiv	Ein Lastwagen donnert um die Ecke. Die Gläser klirren im Schrank. Das Licht war die ganze Nacht an. Das Auto springt nicht an.	Das Licht war die ganze Nacht an. Die Gläser klirren im Schrank. Ein Lastwagen donnert um die Ecke. Das Auto springt nicht an.
kohäsiv	Ein Lastwagen donnert um die Ecke. *Dann* klirren die Gläser im Schrank. Das Licht war die ganze Nacht an. Das Auto springt *deshalb* nicht an.	Das Licht war die ganze Nacht an. *Dann* klirren die Gläser im Schrank. Ein Lastwagen donnert um die Ecke. Das Auto springt *deshalb* nicht an

Ein behaviorales Experiment stellte sicher, dass in den nicht-kohärenten Satzpaaren keine ungewollten inhaltlichen Verknüpfungen enthalten waren (**Kapitel 2**; Ferstl & von Cramon, 2001a; *Cognitive Brain Research*). Die Fehlerrate einer studentische Kontrollgruppe lag bei nur 7%. Wichtig ist außerdem, dass vor allem in den Reaktionszeiten, aber - trotz des Bodeneffektes - auch in den Fehlerraten die erwartete Interaktion zwischen Kohärenz und Kohäsion sichtbar wurde. Kohäsion erleichtert die Inferenzbildung, erschwert jedoch das Erkennen von Kohärenzbrüchen.

Kapitel 3:
Altersunterschiede bei Inferenzprozessen

Das zweite Experiment, beschrieben in **Kapitel 3** (Ferstl, im Druck; *Aging, Neuropsychology & Cognition*) hatte das Ziel, die Aufgabe an einer nicht-studentischen Versuchsgruppe zu testen. Da das Paradigma auch für Patientenstudien geeignet sein sollte, war wichtig, dass auch Personen anderer Bildungs- und Berufsgruppen die Kohärenzentscheidung leicht treffen können. Außerdem sollte geprüft werden, ob in dem Altersbereich, der unserer Patientenpopulation entspricht, Alterseffekte auftreten.

Die gerontologische Forschung zum Textverstehen zeigt, dass ältere und alte Versuchspersonen bei sprachlichen Aufgaben wenig Schwierigkeiten haben, solange sie ihr Weltwissen nutzen können und die Anforderungen an Arbeitsgedächtnis, Geschwindigkeit und Aufmerksamkeit gering sind (Kliegl, Mayer, Junker & Fanselow 1999). Insbesondere zeigen Experimente zur Inferenzbildung, dass Textverständnisprozesse auf der Situationsmodellebene auch bei Probanden der Altersgruppe 65-90 weitgehend intakt sind. Es wurden also für die Kohärenzaufgabe keine altersbedingten Einschränkungen erwartet.

39 Versuchspersonen aus der Altersgruppe von 26–64 Jahren nahmen an einer auditorischen Version des Experimentes teil. Die Ergebnisse waren sehr vergleichbar mit denen der studentischen Probanden. Insbesondere lagen die Fehlerraten für fast alle Personen unter 10%, und die Interaktion zwischen Kohäsion und Kohärenz war sogar leicht stärker ausgeprägt. Auch bei den Antwortzeiten waren keine bedeutenden Alterseffekte zu beobachten. Es gab nur einen Hinweis darauf, dass einige wenige ältere Personen für die kohärenten Durchgänge länger brauchten, als für die inkohärenten.

Diese beiden Experimente bestätigten, dass die Aufgabe für neurowissenschaftliche Studien gut geeignet ist. Insbesondere hat das Paradigma die folgenden Vorteile: a) durch die Kürze der Texte sind Gedächtniseffekte oder Aufmerksamkeitsprobleme minimiert, b) durch die JA/NEIN Entscheidung ist keine Sprachproduktion nötig, c) die Aufgabe ist leicht, schnell und ohne die Anwesenheit eines Versuchsleiters durchzuführen, d) die behavioralen Effekte sind stabil und zeigen die erwartete Interaktion zwischen lexikalischer und pragmatischer Information, e) durch Erheben der Reaktionszeiten gewinnt man Information, die über die Fehlerraten hinausgeht.

Kapitel 4:
Kohärenzbildung: Läsionen des linken präfrontalen Kortex

Die in **Kapitel 4** beschriebene behaviorale Patientenstudie testete die Hypothese, dass Frontalhirnschädigungen die Inferenzbildung beeinflussen (Ferstl, Guthke & von Cramon, 2002; *Neuropsychology*). Insgesamt nahmen 25 Patienten teil, wovon etwa die Hälfte Patienten mit traumatischen Hirnschädigungen waren. Die 25 Teilnehmer wurden gemäß ihrer Läsion einer von fünf Untergruppen zugeordnet. Die folgenden Gruppen wurden betrachtet: Patienten mit Läsionen im Frontallappen (rechts-, links- oder bifrontal), Patienten mit links-temporalen Läsionen und eine Kontrollgruppe von Patienten mit Läsionen, die diese Gebiete aussparten, die also keine sprach- oder text-relevanten Schädigungen zeigten. Zusätzlich zu den erhobenen experimentellen Daten standen aus der klinischen Diagnostik umfangreiche neuropsychologische Testungen zur Verfügung.

Die Ergebnisse zeigten eindeutig, dass vor allem Patienten, deren Läsion in links-frontale Gebiete reichte, mit der Kohärenzaufgabe Schwierigkeiten hatten. Interessanterweise bestätigte sich, dass spezifisch die Entscheidungen für die kohärenten Satzpaare erschwert waren. Unrelatierte Satzpaare zu erkennen, war dagegen auch für die links-frontalen Patienten ohne Probleme möglich. Dieser *Kohärenzeffekt* legt nahe, dass links-frontale Läsionen ausreichen, um ein Inferenzdefizit zu erzeugen. Dagegen zeigten weder Patienten mit links-temporalen, noch Patienten mit rechts-frontalen Läsionen deutliche Einbussen. Ihre Performanz war mit der der Kontrollgruppe von Patienten ohne relevante Läsionen vergleichbar.

Kapitel 5:
Die Neuroanatomie der Kohärenzbildung: eine fMRT-Studie

Die oben beschriebenen Materialien wurden in zwei fMRT-Experimenten mit hirngesunden Probanden verwendet um Hirnareale zu identifizieren, die bei der Sprachverarbeitung im Kontext beteiligt sind. In der in **Kapitel 5** zusammengefassten Studie (Ferstl & von Cramon, 2001a; *Cognitive Brain Research*) verwendeten wir die visuelle Version in einer ereignis-relatierten, pseudo-randomisierten Reihenfolge.

Im Vergleich der Aktivierungsmuster aller Satzpräsentationen mit denen einer perzeptuellen Kontrollaufgabe fand sich ein Netzwerk von perisylvischen Arealen, ganz ähnlich dem eingangs beschriebenen erweiterten Sprachnetzwerk. Im Vergleich von kohärenten zu nicht-kohärenten Satzpaaren ergaben sich zwei Gebiete in der medianen Wand der linken Hemisphäre. Das hintere Areal, am Übergang vom retrosplenialen, posterior-zingulären Kortex zum Präcuneus (PCC/prec) wird mit Langzeitgedächtnisprozessen und Situationsmodellerstellung in Zusammenhang gebracht (z.B. Krause et al., 1999). Das vordere Areal im präfrontalen Abschnitt des fronto-medianen Kortex (BA 9/10, dmPFC) scheint mit der Initiierung und Aufrechterhaltung von nicht-automatischen kognitiven Prozessen in Beziehung zu stehen (vgl. auch Zysset, Huber, Ferstl & von Cramon, 2002). Schließlich zeigte sich eine Interaktion der beiden Faktoren. Wenn kohäsive Mittel die Kohärenz fälschlicherweise suggerierten, fand sich Aktivierung im linken lateralen präfrontalen Kortex, genauer im inferioren frontalen Sulcus. Da dieses Areal auch bei Exektivfunktionsaufgaben, wie z. B. Doppelaufgaben, Wechselaufgaben oder Inhibitionsparadigmen beteiligt ist, interpretierten wir diese Aktivierung als einen Index der Integration von lexikalischer mit pragmatischer Information.

Kapitel 6
Kohärenz oder Theory-of-Mind?

Das wichtigste Ergebnis dieser ersten Bildgebungsstudie war die Beteiligung der beiden medianen Hirnareale, dmPFC und PCC/prec. Dieses Ergebnis wurde in der Zwischenzeit mehrmals repliziert. Auditive statt visueller Präsentation, andere Instruktionen, oder abgewandelte Materialien ändern an diesem allgemeinen Muster wenig. Studien über induktives Schließen (Goel, Gold, Kapur & Houle et al., 1997), oder evaluatives Urteilen (Zysset, Huber, Ferstl & von Cramon, 2002; Zysset et al., 2003) bestätigen die Rolle dieser beiden medianen Areale für verbal basierte Inferenzen. Jedoch wird vor allem das fronto-mediane Areal mit einer Klasse von qualitativ anderen Prozessen in Verbindung gebracht, nämlich Emotion (Greene et al., 2001), selbst-relevante Prozesse (Northoff & Bermpohl, 2004; Decety & Chaminade, 2003) oder, sehr prominent, Theory-of-Mind (Frith & Frith, 1999, 2003; Saxe, Carey & Kanwisher, 2004).

Die in **Kapitel 6** berichtete Studie (Ferstl & von Cramon, 2002; *NeuroImage*) sollte die alternative Erklärung ausschließen, dass die spezifische Aktivierung in der fronto-medianen Wand das funktionelle Korrelat sogenannter Theory-of-Mind (ToM) Prozesse sein könnte (vgl. Fletcher et al., 1995; Vogeley et al., 2001). Die Szene über Charlotte's Geburtstagseinladung löst natürlich Gedanken über ihre Pläne, Motive und Gefühle aus. Um diese gleichzeitigen ToM-Prozesse in der Kohärenzaufgabe zu minimieren, wurden die Satzpaare gamäß ihrem Inhalt aufgeteilt. Die Sätze, in denen von Menschen die Rede war, wurden zusammen mit einer expliziten ToM-Instruktion in der zweiten Hälfte des Experimentes präsentiert. Diejenigen, in denen keine Menschen vorkamen, also eine Brückeninferenz ohne ToM-Inhalt erforderlich war, wurden im ersten Teil des Experimentes mit einer betont sachlichen Logik-Instruktion präsentiert. Von Interesse war vor allem, ob diese Änderung der Instruktionen zu einer Modulation der dmPFC-Aktivierung führen würde. Die Ergebnisse replizierten den Kohärenzeffekt für den ersten Teil des Experimentes, währen im zweiten Teil sowohl kohärente, als auch nicht-kohärente Satzpaare eine dmPFC-Aktivierung auslösten.

Dieses Ergebnis ist in Einklang mit einer Funktion des dmPFC für einen übergeordneten, domänen-unabhängigen Prozess, der sowohl ToM als auch die Kohärenzbildung beinhaltet. In Anlehnung an das philosophische Konzept des Selbst-Modells (Metzinger, 2000) und mit Blick auf die klinische Symptomatik von Patienten mit fronto-medianen Läsionen (Marin, 1991) postulierten wir, dass dieser Prozess der ständigen Interaktion zwischen interner Welt und externer Stimulation dient. Er wird also vor allem in Aufgaben sichtbar, die ideosynkratische Kriterien unter Nutzung von Weltwissen erfordern. Unabhängig von der genauen Interpretation ist jedoch die Schlußfolgerung, dass eine dmPFC-Aktivierung als Indikator von nicht-automatischen, erklärungsbasierten Inferenzen betrachtet werden kann.

Kapitel 7:
Situationsmodellbildung: Integration emotionaler und zeitlicher Information

Von der Vielzahl komplexer Teilleistungen, die zur Textverarbeitung notwendig sind, werden im Kohärenzparadigma nur lokale Inferenzprozesse erfasst. Um das Verstehen längerer Texte und die damit verbundene Erstellung einer globalen Repräsentation erfassen zu können, betrachten wir nun die Situationsmodellbildung (Kintsch, 1998; van Dijk & Kintsch, 1983; Rinck, 2000; Zwaan & Radvansky, 1998; Ferstl & Kintsch, 1999; Ferstl, 2001). Das Situationsmodell ist ein mentales Modell, das die im Text enthaltenen Informationen mit dem Vorwissen integriert. Somit ist diese Repräsentation geprägt durch die individuelle Erfahrung, die die Rezipientin in den Verstehensprozess einbringt. Zahlreiche behaviorale Studien haben jedoch nachgewiesen, dass fast alle Leser von Erzähltexten Informationen z.B. über die handelnden Personen, sowie eine Reihe von Dimensionen der Handlung on-line im Situationsmodell enkodieren. Unter anderem wurde die psychologische Relevanz von räumlichen, zeitlichen, kausalen, intentionalen und emotionalen Aspekten gezeigt (Albrecht & O'Brien, 1993; Gernsbacher, Goldsmith & Robertson, 1992). Es gibt noch keine neuropsychologischen Untersuchungen über diese Situationsmodell-Aspekte.

In den hier relevanten eigenen Studien wurde das sogenannte Inkonsistenzparadigma angewandt (Rinck, Hähnel, & Becker, 2001; Rinck & Weber, 2003). In eine kurze Geschichte werden dabei Fehler eingestreut, die sich auf einen Aspekt der beschriebenen Situation beziehen. Ein Beispiel für Verletzungen der emotionalen und temporalen Dimensionen ist in Tabelle 2 gezeigt.

Tabelle 0.2. Beispiele für die emotionalen und chronologischen Geschichten. Die fett gedruckten Wörter beinhalten die Zielinformation, die kursiv gedruckten Wörter ergeben die inkonsistente Version.

emotional	**zeitlich**
Es war eine dieser Parties, wo einfach alles stimmte. Die Gäste lachten, waren ausgelassen und amüsierten sich großartig. ... Sarahs beste Freundin umarmte sie und lobte die tolle Party. Sarah konnte sich nicht erinnern, dass sie sich schon einmal so **glücklich / *traurig*** gefühlt hatte. ...	Alex war an diesem Tag ins Café gekommen, um sich mit Sabine zu treffen, **bevor / *nachdem*** er ein wichtiges Interview **zu führen/*geführt*** hatte. ... Heute wollten sie über eine Gerichtsreportage sprechen. Als Sabine endlich auftauchte, wollte Alex gerade **zu seinem Interview aufbrechen**. ...

In der in **Kapitel 7** beschriebenen fMRT-Studie (Ferstl, Rinck & von Cramon, 2005; *Journal of Cognitive Neuroscience*) wurden 20 Versuchspersonen gescannt, während sie die Inkonsistenzaufgabe bearbeiteten. Um ventro-mediane präfrontale Areale abbilden zu können, verwendeten wir eine Spin-Echo-Sequenz (Norris, Zysset, Mildner & Wiggins, 2002; Zysset et al., 2003), die Suszeptibilitäts-Artefakte vermindert.

Für die Entdeckung der Inkonsistenz fand sich eine rechts-anterior temporale Aktivierung, im Einklang mit der RH-Hypothese. Unabhängig von der Konsistenz lösten die temporalen Zielinformationen Aktivierung in einem fronto-parietalen Netzwerk aus, das mit Gedächtnis-, Exekutiv- und Aufmerksamkeitsprozessen zusammenhängt. Emotionale Information beteiligte dagegen den ventro-medianen präfrontalen Kortex und den Amygdala-Komplex, Bereiche, die mit Emotionsverarbeitung in Verbindung gebracht werden (Luan Phan, Wager, Taylor & Liberzon, 2002; Kringelbach & Rolls, 2004). Schließlich zeigte eine Analyse der Integration der Zielinformation, also eine Auswertung bei der der letzte Satz mit eingeschlossen war, dass die emotionalen Inkonsistenzen anders integriert wurden. Es fand sich eine Aktivierung des dmPFC, in Einklang mit der Beobachtung, dass sich auch inkonsistente emotionale Informationen durch erklärungs-basierte Inferenzen in ein plausibles Situationsmodell einbinden lässt.

Kapitel 8:
Schlussfolgerungen

Die in dieser Arbeit dargestellten Studien bestätigten, dass neurowissenschafltiche Methoden zu unserem Verständnis von Textverarbeitungsprozessen beitragen. Durch die Entwicklung eines für Patienten- und Bildgebungsstudien geeigneten Paradigmas konnte die Rolle des linksseitigen präfrontalen Kortex für Inferenzprozesse bestätigt werden. Laterale Regionen zeigten die Inkonsistenz von lexikalischer Kohäsionsinformation mit pragmatisch-inhaltlicher Information an. Eine besondere Beteiligung der rechten Hemisphäre für die Sprachverarbeitung im Kontext blieb auf den anterioren Temporallappen beschränkt. Sowohl in globalen Vergleichen, als auch bei der Entdeckung von Inkonsistenzen im Situationsmodell war dieses Areal aktiv. Da dieser Bereich auch bei Satzverarbeitung beteiligt ist, weisen wir ihm vorläufig eine Rolle für die sogenannte Propositionalisierung, also für das Zusammenfügen von Inhaltseinheiten zu. Weitere Forschung ist nötig, um diese Hypothese zu überprüfen.

Das wichtigste Ergebnis der hier vorgestellten Arbeiten war die wiederkehrende Beteiligung des linken dorso-medianen präfrontalen Kortex. Dieses Areal, das häufig zusammen mit dem postioren cingulären Kortex beobachtet wird, spielt eine Rolle bei der erfolgreichen Inferenzbildung, wenn eine Berücksichtigung des Weltwissens oder des Diskurskontextes erforderlich ist. Dass diese Aktivierungen Nebenprodukte eines gleich-laufenden Theory-of-Mind-Prozesses sein könnten wurde ausgeschlossen. Eine allgemeinere, domänen-unabhängige funktionelle Zuordnung stützt sich auf die Beschreibung von neuropsychologischen Defiziten von Patienten, die in dieser Hirnregion Einbussen erleiden. Sie zeigen verminderten Antrieb, und verminderte Sprachproduktion. Daher wurde die Rolle des fronto-medianen Kortex in Einklang mit neurophilosophischen Modellen als die kontinuierliche Integration von inneren, selbstbezogenen Prozessen mit der externen Stimulation beschrieben.

Diese Ergebnisse sind ein vielversprechender Anfang für die genauere Beschreibung der Neuroanatomie des Textverstehens. Weitere Forschung ist nötig, um die Resultate zu replizieren und die Interpretationen anhand unterschiedlicher Materialien und Aufgaben zu überprüfen. Das dadurch angzielte bessere Verständnis des Zusammspiels der verschiedenen Hirnregionen im erweiterten Sprachnetzwerk ist für die Diagnostik und Behandlung von erworbenen Textverständnisdefiziten unerlässlich.

PART II:
The neuroanatomy of text comprehension

Chapter 1
Introduction and theoretical background

Communication is at the core of our social lives. Without the ability to talk and converse, to understand instructions, to watch movies or read books and newspapers it is difficult to lead a full life in our modern world. In the last decades, the advent of new communication technologies has even increased the importance of spoken and written language. Talking on the telephone requires comprehension without the help of gestures or facial expressions, and using the internet necessitates the flexible and creative use of language cues for navigation and information extraction.

Brain damage can dramatically impede this seemingly effortless skill. After left-hemisphere lesions of the perisylvian language cortex, aphasic deficits may impair phonological, syntactic or lexico-semantic processes. However, when comprehending language in context, the linguistic analysis of the words and sentences interacts with a wide variety of other cognitive functions, including but not restricted to knowledge retrieval, attention, working memory, or executive functions, such as integration, inhibition or sequencing. It is therefore not surprising that even in the absence of aphasia, text comprehension or discourse production can be severely impaired. However, in clinical practice these so-called non-aphasic language deficits (Prigatano et a., 1986; Glindemann & von Cramon, 1995; Ferstl, Guthke & von Cramon, 1999) are still often overlooked. Specifically, diagnostic tools for assessing text comprehension abilites are few and far between (Ferstl, Walther, Guthke & von Cramon, 2005; Brookshire & Nicholas, 1993).

The most important reason for this lack of attention is that these deficits are not yet sufficiently understood. In particular, it is necessary to make exact predictions about the subprocesses required for a particular language task. In psycholinguistic research on text comprehension, a wealth of empirical data has been accumulated and computational models have been developed (see Schmalhofer & Perfetti, in press; Gernsbacher, 1994; Graesser, Gernsbacher &

Goldmann, 2003), so that a good understanding of the cognitive requirements of language processing in context can be assumed. Recently, these behavioral and computational approaches have been extended by neuropsychological methods, in the attempt to map the respective subprocesses onto brain regions likely to participate in their implementation.

In this volume I am going to present a series of neuroscientifc studies on language comprehension in context. In three sections of the following introductory chapter I will provide the relevant background for this research program. The theoretical part has the goal of introducing the most important concepts, embedded in a psycholinguistic theory of language comprehension in context. Second, findings from neuropsychological patient studies and relevant theories are introduced. The chapter concludes with a short overview of the functional neuroanatomy of language processing as it emerges from neuroimaging methods.

This introductory chapter draws on a review appearing in the volume by F. Schmalhofer and C. A. Perfetti (Eds.), *Higher Level Language Processes in the Brain* (Ferstl, in press, b).

A framework of comprehension

As the general framework, I adopt the psycholinguistic theory of text comprehension put forward by Kintsch and van Dijk (1978; van Dijk & Kintsch, 1983; Kintsch 1988, 1998). Text comprehension is conceptualized as the cognitive process of mapping the linguistic input onto a mental model of the text contents, called the situation model (Ferstl & Kintsch, 1999; Ferstl, 2001; Gernsbacher, 1990; Rinck, 2000; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). This global representation of "what the text is about" contains an integration of the text information with the reader's or listener's background knowledge. Thus, the situation model goes beyond the explicit information and includes elaborations and interpretations. Pragmatic interpretations and figurative language are special cases of situation model building in this framework. Although these aspects of language interpretation are not comprehensively included in the text comprehension framework, there are attempts, for instance, to account for metaphor comprehension (e.g. Kintsch, 2000).

To be able to develop the situation model, two lower level representations are required. First, the surface structure encodes the verbatim form of the text information. This representation of the exact wording preserves subtle differences in the syntactic structure or the specific choice of vocabulary. Building a surface structure is a requirement for the higher level memory representations, but the specific component processes (e.g., parsing or lexical access) are not actually part of the text comprehension framework. The second level, the so-called text base, is a semantic representation of the text. In a propositional format, this level represents content units independent of the exact wording, but still true to the specific information given. The text base is augmented by implicit information necessary for establishing local coherence, i.e, for sensibly connecting subsequent sentences to the discourse context.

One account for how the text base is derived is the construction-integration model (Kintsch, 1988). In this model, two basic mechanisms are combined to form a coherent propositional network. In the construction phase, associations to the text propositions are retrieved from the recipient's background knowledge and linked with the text propositions. Second, a constraint-satisfaction process integrates the loosely associated content units and deactivates those that are contextually inappropriate. Thus, the construction phase proposes a large number of possible interpretations and inferences, whereas the integration phase yields a coherent and plausible representation by an inhibitory process.

For each sentence, this process cycle is repeated with several most highly activated propositions being carried over from the previous cycle. Thus, both the influence of working memory and contextual integration are conceptualized as an associative, bottom-up activation process. However, without some minimal overlap between the immediately preceding and the current utterances, the network will remain incoherent. According to the proposal of the early text comprehension model (Kintsch & van Dijk, 1978), a reinstatement search for relevant information from the earlier discourse context, or an explanation-based inference (see Gueraud & O'Brien, 2005) using relevant information from long-term memory will be initiated.

Importantly, Kintsch (1998) argues for the domain-independence of this comprehension model. While it was developed specifically to describe the process sequence of language understanding in context, its constraint-satisfaction mechanism has also proven useful for modelling conceptual retrieval of categories and scripts, the comprehension of computer programming commands and problem solving processes (Doane, McNamara, Kintsch, Polson, & Clawson, 1992; Kintsch, 1998; Mannes & Kintsch, 1991).

Two very important research questions emerged from this framework and influenced much of the experimental research on text comprehension. First, there is considerable debate about the types and quality of inference processes. In particular, the discussion has centered around the question under which circumstances inferences are a mandatory part of the on-line comprehension process, as compared to inference processes elicited by a specific reading goal or comprehension task (McKoon & Ratcliff, 1992; van den Broek, 1990, 1994; Graesser, Singer & Trabasso, 1994; Singer, 1994). And second, the nature of the situation model has received much attention (Ferstl & Kintsch, 1999; Zwaan & Radvansky, 1998; van Dijk & Kintsch, 1983; Rinck, 2000; Graesser, Millis, & Zwaan, 1997). Specific issues include the psychological reality of information specific, non-verbal representations and the effects of adjunct information (e.g., illustrations or background knowledge on the ease of situation model building (cf. Ferstl & Kintsch, 1999; Gyselinck & Tardieu, 1999). We will return to these issues in more detail in chapters 2 and 7.

Neuropsychology of text comprehension

Ever since the ground breaking work of Paul Broca (1865) almost 150 years ago, it has been known that language functions are likely to be predominantly realized in the left hemisphere (LH), and more spefically, in the perisylvian cortex (shown in Figure 1.1). Lesions in these language-relevant areas are likely to result in aphasic symptoms, compromising phonological, syntactic or lexico-semantic language processes in comprehension and/or production (for more detail: Benson & Ardila, 1996; Brown & Hagoort, 1999; Kent, 2004; Stemmer & Whitaker, 1998; Caplan, 1987, 1992). The classic view attributed language production to the frontal language area, Broca's area, and comprehension to the posterior language area, Wernicke's area (Broca, 1865; Wernicke, 1874; Lichtheim, 1885; Geschwind, 1979). Although the functional attributions have since been differentiated from this dichotomy to linguistically based models, the Wernicke-Lichtheim-Geschwind conception of the functional neuroanatomy of language is still highly influential.

Despite word and sentence level deficits, aphasic patients are often surprisingly effective in using situational and contextual cues in communication (Pierce, 1988; Germani & Pierce, 1992; Chapman & Ulatowska, 1989; Guthke, Hauptmann & Ferstl, 2001). Their text comprehension abilities are sometimes better than predicted by the aphasia diagnostics, and their discourse production has been shown to reflect the use of structuring devices and to contain the standard elements of the specific text genre (Ulatowska & Chapman, 1994).

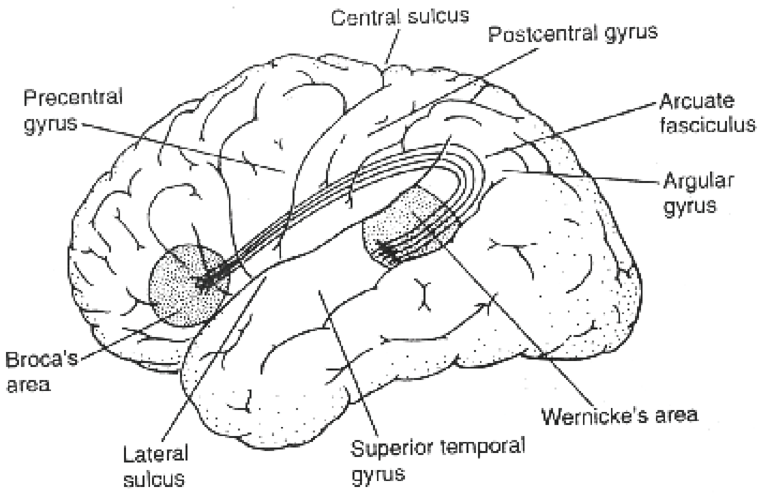


Figure 1.1. Schematic rendition of the left hemisphere of a single brain (cf. Geschwind, 1979), including the approximate locations of Broca's area in the inferior frontal cortex (BA 44 /45) and Wernicke's area in the posterior part of the superior temporal cortex

In contrast, so-called non-aphasic language or communication deficits (NALD; Prigatano et al., 1986; Glindemann & von Cramon, 1995)[1] specifically affect pragmatics and the text level more than the word and sentence levels. These deficits usually occur in patients without aphasia, i.e., in patients whose lesions spare the left perisylvian language areas. The two most well-studied groups

---

[1] The term non-aphasic language disorders was coined by Prigatano et al. (1986) for communication deficits in frontal lobe patients. Because the term is so general and underspecified I extend its use to encompass any non-aphasic language deficits after brain damage.

likely to suffer from NALD are with closed-head injury (CHI) patients and with right hemisphere (RH) brain damage.

Brownell and Martino (1998; Brownell, Gardner, Prather & Martino, 1995) describe the so-called RH syndrome, which typically occurs after infarctions of the right middle cerebral artery, as including but not being restricted to communicative symptoms. The patients are likely to be unkempt in their appearance, they are unaware of their deficits, their conversation is incoherent, tangential and socially inappropriate. With respect to language comprehension, RH patients often have considerable problems with non-literal, pragmatic interprations and fail to understand humor, irony, metaphors or indirect requests. Consequently, the dominant view emerging from experimental neuropsychology is that pragmatics is realized in the right hemisphere (e.g., Bookheimer, 2002). Apparently, this *RH hypothesis* is supported by a wealth of empirical data (see Brownell & Martino, 1998; Chiarello & Beeman, 1998; McDonald, 1993, 2000; Nicholas & Brookshire, 1995) and has been extended to text comprehension processes in general. RH-patients have been shown to exhibit difficulties with all aspects of higher level comprehension, such as inferencing (Brownell, Potter, Bihrle, & Gardner, 1986; Mc Donald & Wales, 1986), the derivation of a story's theme (Hough, 1990), or the revision of erroneous interpretations.

Despite this overwhelming evidence, less attention has been spent on delineating the functional properties of the RH contribution to language (McDonald, 2000). The descriptions of the specific subprocesses realized in the right hemisphere vary widely. The proposals include a simple dichotomy between local and global, or analytic and holistic functions being supported by the LH vs. the RH (cf. Corbalis, 1997; Springer & Deutsch, 1997; Benowitz, Moya & Levine, 1990). According to this account, situation model building or "language synthesis" (McDonald, 2000) would be considered an RH function, while surface and textbase levels are LH functions. A second, rather general account is resource theory (Navon, 1984; Murray, 1999; Weissman & Banich, 2000). Monetta and Joanette (2003) suggested that the RH comes into play during complex and demanding language tasks (cf. Meyer, Friederici & von Cramon, 2000; Ferstl & von Cramon, 2001a). Interestingly, though, a mapping of the terms *complex* or *demanding* does not readily map onto the distinction between text and word level. Setting up a situation model for a newspaper article is not necessarily more resource demanding than, let's say, memorizing a poem. Using a very different explanation, higher level language deficits in RH patients have been attributed to underlying problems with social, pragmatic and affective aspects of communication (Brownell & Martino, 1998). Thus, it is necessary to carefully separate these "hot" aspects from the "cold" or cognitive ones (Goel & Dolan, 2004), such as inferencing or syntactic parsing (cf. Lehmann & Tompkins, 2000). The most specific proposal to account for RH language deficits was formulated by Beeman (1993; 1998). Based on the results of patient studies as well as priming studies using semi-field presentation (reviewed in Chiarello & Beeman, 1998), Beeman (1993) postulated a particular role for the RH during the activation of loosely structured semantic fields, a process accounting for inferencing, but also for the comprehension of metaphors or some non-literal meanings. This so-called *coarse coding hypothesis* has influenced the interpretations of many neurolinguistic findings.

The second group of patients affected by NALD are CHI patients (McDonald, Togher, & Code, 1999). Although aphasia affects only between about 2-4 % of all CHI patients, and about 20% of severely injured patients (Hartley & Levin, 1990), communication skills are predictive of rehabilitation outcome (Brooks et al., 1987). Consequently, the description of the communication deficits of CHI patients, whose brain damage often impairs frontal lobe (FL) functions (Adams & Victor, 1993), overlaps to a great deal with that of RH patients (Nicholas & Brookshire, 1995; McDonald, 1993). The American Speech and Hearing Association has summarized the symptoms of what they termed *Cognitive Communication Impairment* (ASHA, 1988). Once more, this disorder is described as a complex of symptoms in the domains of social behavior, discourse production and text comprehension. Patients may exhibit disorganized discourse, hyperverbosity, and socially inappropriate, tangential or imprecise language. Text comprehension difficulties are characterized as problems with comprehending extended language, with detecting main ideas, or with understanding abstract language, including indirect or implied meaning (ASHA, 1988; cited from Larkins, Worrall & Hickson, 2000).

A theoretical alternative to the RH hypothesis, consequently, is what I would like to call the *FL hypothesis*. And in fact, when listing cognitive processes required during text comprehension, such as sequencing, structuring, goal-directed reading, inferencing, integration, or monitoring, it can be seen that all of them fall in the domain of executive functions, which clearly engage the frontal lobes (Fuster, 1997, 1999). The view that non-aphasic communication deficits are closely related to the executive dysfunction (cf. Novoa & Ardila, 1987) is confirmed by the observation of left FL contributions to aspects of text comprehension (Channon & Crawford, 2000; Chapman, Levin, & Culhane, 1995; Ferstl, Guthke & von Cramon, 1999; 2002; Kaczmarek, 1984, 1987; Novoa & Ardila, 1987; Prigatano, Roueche, & Fordyce, 1986; Zalla, Phipps, & Grafman, 2002). Despite this relation, it is not yet clear whether communication deficits are secondary to more basic executive deficits (cf. ASHA, 1988), or whether they make an independent contribution (e.g., Ettlin, Beckson, Gaggiotti, Rauchfleisch & Benson, 2000).

The RH and FL attributions of discourse comprehension deficits are by no means mutually exclusive. They cover different aspects of the comprehension process, and the neuroanatomical specification is relatively imprecise. Thus, a co-existence of separable processes in different subregions is possible and likely. Neuropsychological research on text comprehension and more general, on non-aphasic communication deficits must therefore include both RH and FL patients. The goal is to compare and contrast the respective communication problems, so that an attribution of specific functions to brain regions becomes feasible.

The Extended Language Network

An alternative approach to lesion studies are functional neuroimaging techniques. These methods yield a description of the networks of brain regions engaged during listening to or reading of language in context. The most widely used technologies are functional magnetic resonance imaging (fMRI) and positron-emission tomography (PET). Based on different physical principles, both methods take advantage of the vascular properties of brain activation. The metabolism of activated neurons requires increased oxygen and glucose consumption. Thus, an increase in regional cerebral blood flow or volume (rCBF, rCBV) as measured by PET, or an increase in the BOLD (blood oxygen level dependent) contrast, as measured by fMRI, are taken as evidence for the activation of the region during task performance (for methodological introductions see: Schwartz, Kischka, & Rihs & 1997; Jezzard, Matthews & Smith, 2001; Cabeza & Kingstone, 2001).

The first step in studying the neuroanatomy of text comprehension is to identify all brain regions involved during reading of or listening to connected text. Although only few studies on this issue are available, the results consistently converge on a fronto-temporal network of brain regions including, but not restricted to the perisylvian cortex [2]. Most activation patterns include a subset of the regions in this network and are left-dominant, but partly bilateral. Figure 1.2 displays a schematic rendition of the left hemisphere with the relevant regions shaded and labelled. Specifically, this *extended language network* includes the prefrontal cortex, in particular the inferior and middle frontal gyri (IFG, alPFC, plPFC), the temporal lobes, in particular the anterior temporal lobe and temporal poles (aTL), the entire extent of the superior temporal sulcus (mTL, pTL), the temporo-parietal junction area and the inferior parietal lobe (IPL), in particular the supramarginal and angular gyri. In most studies, the pattern is bilateral but clearly left-dominant. In addtion, there is evidence for a contribution of medial structures, in particular the dorso-medial prefrontal cortex (dmPFC) and a parieto-medial region including the posterior cingulate cortex and lower precuneus (PCC/prec) (see Figure 1.3).

In their seminal PET study, Mazoyer et al. (1993) were the first to describe this extended language network. Comparing a number of different language conditions and employing an anatomically based region-of-interest analysis, the authors reported bilateral superior temporal gyrus activation (mTL/pTL), reflecting the auditory perception of the stimuli. In addition, story comprehension elicited bilateral, but left-dominant activation in the aTLs, including the temporal poles. The middle temporal gyrus (mTL and/or pTL) and the IFG were only active in the left hemisphere. Most importantly, there was a contribution of the left superior frontal gyrus (BA 8), possibly reaching into dmPFC. No activations were reported however, in other prefrontal regions, in medial parietal regions (PCC/prec) or in temporo-parietal regions, such as the supramarginal or angular gyri (IPL) (cf. Tzourio, Crivello, Mellet, Nkanga-Ngila, & Mazoyer, 1998; Papathanassiou, et al., 2000, for replications).

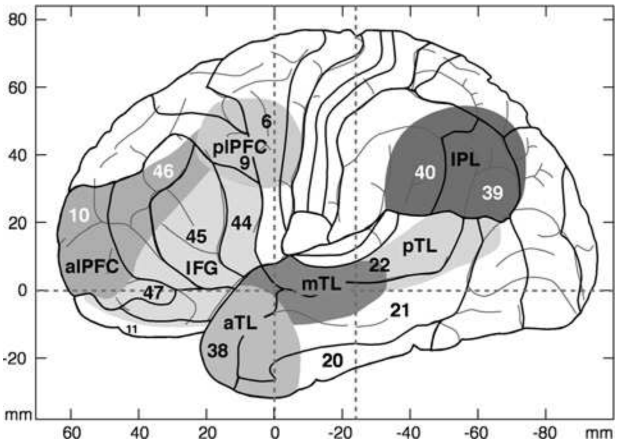


Figure 1.2. A left lateral view of the Talairach standard stereotaxic brain. Overlaid and shaded are regions of interest that indicate the most important brain regions for text comprehension in context (alPFC = anterior lateral prefrontal cortex, plPFC = posterior lateral prefrontal cortex, IFG = inferior frontal gyrus; aTL = anterior temporal lobe; mTL = middle temporal lobe; pTL = posterior temporal lobe; IPL = inferior parietal lobe). The numbers denote the cytoarchitectonic fields as described by Brodmann (1909).

When subtracting a perceptual baseline, the patterns change surprisingly little. Perani, Dehaene, Grassi et al. (1996), and Dehaene, Dupoux, Mehler et al. (1997) compared story comprehension to a foreign language unknown to the participants. There still was superior temporal activation, but it fell now more clearly into the superior temporal sulcus, rather than extending along the superior temporal gyrus. Moreover, the activation split up into a focus in the anterior (aTL) and one in the posterior superior temporal sulcus (pTL), while the mTL area was subtracted out by the auditory perceptual baseline. Crinion, Lambon-Ralph, Warburton et al. (2003), in a study comparing story comprehension to the same story played backwards, replicated these results. Activation was restricted to these temporal regions, with the left aTL activation reaching into the inferiormost IFG. Despite their rather conservative analysis

---

[2] The present discussion remains restricted to cortical regions, although some studies also show activations in of the basal ganglia, the hippocampus or the cerebellum. A functional interpretation of these activations is beyond the scope of this volume.

criteria, the authors attribute the lack of more dorsal prefrontal activation to the absence of task demands.

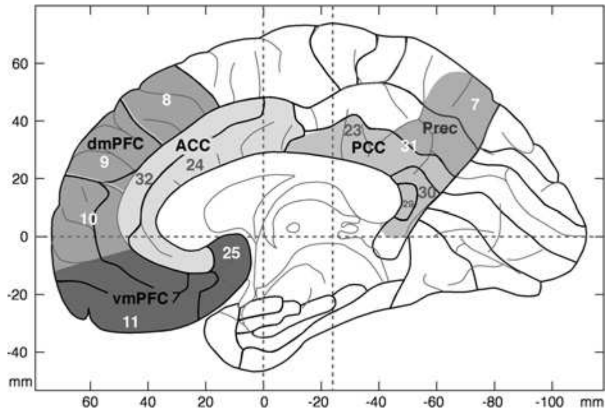


Figure 1.3. A right medial view of the Talairach standard stereotaxic brain. Once more, only those regions are shaded and labelled that are important for the present discussioin (Prec = precuneus; PCC = posterior cingulate cortex, ACC = anterior cingulate cortex; dmPFC = dorso-medial prefrontal cortex; vmPFC = ventro-medial prefrontal cortex). The numbers denote the cytoarchitectonic regions described by Brodmann (1909).

Most recently, Xu et al. (2005) conducted a carefully controlled study replicating the PET study by Mazoyer et al., (1993) with more advanced methods. Using whole brain analyses for fMRI at 3 Tesla, they identified the same regions as those described earlier. For word processing as compared to a perceptual baseline, bilateral posterior temporal and left inferior frontal regions were active. For the sentence condition, additional activation was seen in both anterior temporal lobes and the frontal activation was more extended. Finally, a narrative condition showed more extended and more bilateral activation, most notably in the temporo-parietal junctions, the dmPFC and the PCC/prec. Interestingly, and consistent with the RH hypothesis, right hemisphere regions were particularly active during the concluding episodes of the stories, when the incoming sentences had to be integrated into the previous story context.

Sentence comprehension can also activate parts of the extended language network if content based interpretations are relevant (e.g., Bavelier et al., 1997; Baumgärtner, Weiller & Büchel, 2002; Perani et al., 1996). Thus, not the text length is crucial, but whether processing includes the situation model level. Consequently, the extended network is not even specific to contextual language comprehension. Very similar results were reported for semantic decision tasks on the word level (Scott, Leff & Wise, 2003; Binder, Frost, Hammeke, et al., 1997), involving an internally guided, knowledge based categorization ("Can this word refer to a person?", "Is the lion a domestic animal?"). And even during rest, in the absence of any stimulation, a comparable pattern was found when compared to activation during a simple task capturing the participants' attention (Binder, Frost, Hammeke et al., 1999). Thus, it seems that at least part of the language comprehension network is continually at work, even when there is no external task demand.

Research Questions and Overview

Language processing in context requires brain areas outside the perisylvian language cortex. Both clinical descriptions of non-aphasic language deficits and imaging studies provide evidence for a contribution of right hemispheric and frontal brain regions, and more specifically, of an extended network of fronto-temporal brain regions. The open question is whether these regions map onto dissociable subprocesses of text comprehension. Of particular interest is the neuroanatomical realization of the two defining subprocesses: inferencing and situation model building. To investigate this research question, a combination of behavioral, neuropsychological and neuroimaging methods seems crucial. First, the operationalization of the postulated subprocesses requires a careful cognitive analysis of the materials and the task requirements. This step is best accomplished by behavioral experiments. Second, the identification groups of brain damaged patients who have difficulty with the task enables us to link the process to a particular lesion site. And third, fMRI provides the means for a more fine-grained description of the brain regions involved during task performance.

The series of studies summarized in the following chapters were designed for studying the functional neuroanatomy of inferencing and situation model building in more detail. With particular focus on the role of the right hemisphere and the frontal lobes, the goal was to describe the brain regions engaged during higher level language comprehension, taking into account the text properties and the task demands. To tackle this issue, I first developed a paradigm for assessing inference processes, appropriate for both patient and fMRI studies. The behavioral properties of this coherence judgment task are the topic of Chapters 2 and 3. In Chapter 4, I report on a group of brain injured patients tested with the coherence judgment task. The results implicated the left frontal lobe as relevant for bridging inferences based on general world knowledge. Chapter 5 presents a first fMRI study using the same materials and task. As expected, the results confirmed the importance of the left frontal lobe for inferencing. In particular, lateral prefrontal regions were active when the integration of inconsistent lexical and pragmatic informations was needed, and the dmPFC emerged as having a specific function for successful inferencing. The fMRI experiment summarized in Chapter 6 was carried out to exclude the alternative explanation that the dmPFC activation was solely due to concurrent Theory-of-Mind processes, as implied by previous functional attributions of the dmPFC. Chapter 7 provides the results of a further fMRI study in which the previous findings on local inference processes were compared to global inference processes needed for situation model building. Locally coherent, but inconsistent information was to be integrated in the on-going representation of stories. The crucial distinction, allowing for the separation of hot and cold aspects of cognition, was between emotional and chronological target information. Global inferences on emotional information once more activated the dmPFC, whereas the integration of chronological inconsistencies engaged a fronto-parietal network related to executive functions. The results are summarized and discussed in Chapter 8.

For all studies only the most relevant results are sketched. More detail is provided in the original articles.

Chapter 2
Coherence building: Lexical and pragmatic information

The results reported in this chapter were published in *Cognitive Brain Research* (Ferstl & von Cramon, 2001).

The most basic process distinguishing text comprehension from sentence comprehension is coherence building. Because speakers and writers leave much information implicit, inferences are needed for rendering texts coherent. Successive utterances are assumed to refer to the same topic. When the connection is not immediately obvious, an integration of the current language input with background knowledge and discourse context leads to a content-based link. In addition to these bridging inferences, a variety of other inference types are used for augmenting the explicitly presented text information (for reviews see Singer, 1994; van den Broek, 1994; Graesser, Singer & Trabasso, 1994; also Guéraud & O'Brien, 2005; Schmalhofer, McDaniel & Keefe, 2002). Consider, for instance, the classical example

Charlotte was invited to a birthday party.
She asked her mother for some money.

Although there is no direct associative overlap between the content words, the short text is immediately coherent. The comprehender adds the cultural knowledge or script knowledge about birthday parties (Schank & Abelson, 1977; Ferstl & Kintsch, 1999), and specifically, that birthday guests are expected to bring presents. In addition, a number of elaborative inferences augment the explicit information. In the example, the reader might infer that Charlotte is a child, and that she intends to use the money for buying a present. Dependent on the context, and on the reader's own experience with birthdays or with children, there might be inferences on the affective reactions of the protagonists: Charlotte is happy about being invited, or her mother does not want to give her any money.

Which of these inferences are mandatory and drawn automatically during comprehension has been an issue of considerable debate. The minimalist or memory-based hypothesis postulates that only those inferences are drawn online that are based on easily available knowledge and those that are required for establishing local coherence (McKoon & Ratcliff, 1992). In contrast, the constructionist or explanation-based view states that causal and goal related inferences are important building blocks for deriving a text representation, in particular for narrative texts (Singer, 1994; Graesser, Singer & Trabasso, 1994; van den Broek, 1994). Recently, these views have merged in the attempt to describe the circumstances and the text features leading to an interplay of either bottom-up or top-down processes (Guéraud & O'Brien, 2005).

For the series of experiments reported here, I focussed on bridging inferences that have the function of rendering texts coherent. Before going into the details of the task, let's once more consider the aforementioned example. In addition to the knowledge units or propositions from the birthday party script, the second sentence contains a linguistic marker signalling a connection between the sentences. In this case, the personal pronoun *she* indicates that Charlotte is the agent in both sentences, thus directly providing overlapping information. This so-called *cohesion* is a strong cue utilized by the comprehender (Halliday & Hasan, 1976; Robertson, Gernsbacher et al., 2000; Münte, Schiltz & Kutas, 1998; Schwarz, 2000). Other types of cohesive markers include word repetition, anaphors, and most importantly conjunctions and adverbs. A particle such as *because*, for instance, signals a connection between propositions, but also adds information about its type, in this case a causal connection. Coherence building must therefore involve the interplay of various processes, including lexical, semantic and pragmatic interpretations.

In order to investigate these coherence building processes, I devised a task appropriate for both patient studies and neuroimaging applications. The text materials were 120 coherent, minimal stories consisting of two sentences, the context and the target sentence. Care was taken that the connection was not purely based on associative overlap, but that it required the additional retrieval and integration of general world knowledge. For the incoherent condition, context sentences of two stories were switched. The participants' task was a simple YES/NO-decision to indicate whether the sentence had a pragmatic, content based connection. The second variable of interest, cohesion, was included to study how lexical information influences the knowledge based inference process. For each target sentence, two versions were used. The first version contained cohesive markers (Halliday & Hasan, 1976), whereas in the second version, these lexical cues to coherence were omitted. Examples are provided in Table 2.1 (for English translations, please refer to the examples in the original article). As can be seen in the example, the cohesive markers render the incoherent, unrelated sentence pairs slightly odd, because the lexical and pragmatic cues to coherence clash.

Table 2.1. Sentence examples for the four conditions of the experiment. The cohesive markers are printed in Italics. The incoherent sentence pairs are created by switching the context sentences of two coherent trials.

	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.

In the first experiment, a group of 24 students read the sentences one at a time, and subsequently made the coherence judgment. The goal of the experiment was to confirm the discriminability of coherent and incoherent sentence pairs. Moreover, we wanted to test whether the presence of a cohesive tie influenced the inference process as postulated.

> **Hypothesis 2.1**: Cohesive markers facilitate the inference process, whereas they render the detection of coherence gaps more difficult. This effect influences both error rates and reading times for the target sentences.

The results are displayed in Figure 2.1. The overall error rate was less than 5%. This performance level close to ceiling confirmed that the coherent and incoherent versions of the sentence pairs were easily discriminable. Slightly more errors were made on coherent than on incoherent trials. The expected interaction was visible in the patterns of the mean error rates, but it was not statistically reliable.

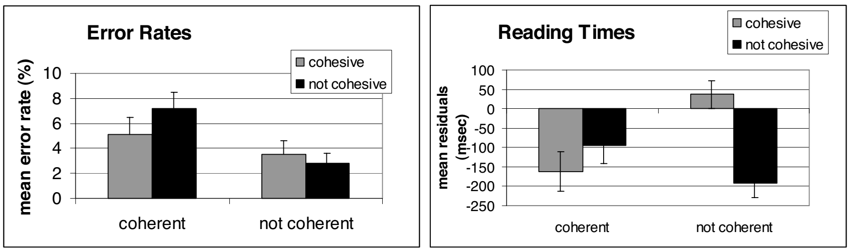


Figure 2.1. Performance data for the coherence judgment task. Both error rates and reading times (controlled for sentence length) show the expected interaction between pragmatic coherence and lexical cohesion.

The analysis of the reading times for the target sentence and of the judgment times (not shown) yielded significant interactions between Cohesion and Coherence, which were mostly due to the incoherent condition. When a lexical cohesive marker erroneously signalled a pragmatically implausible connection, the processing times increased.

These results confirm that the materials for the coherence judgment task were appropriately designed for subsequent use in neuroscientific studies. Three criteria were met. First, low error rates are important for use with a brain damaged population. Even without specific language or text comprehension deficits, an unspecific decrease in performance is usual. Decisive conclusions, in particular on differential effects of the experimental variables can only be drawn when the overall performance level remains above chance. The second criterion was the comparability of response times for coherent and incoherent sentence pairs. Usually, coherent sentences are read faster than incoherent ones (e.g., Gernsbacher, 1990; Keenan, Baillet & Brown, 1984; Myers, Shinjo & Duffy, 1987), so that in a subsequent fMRI study possible differences in the BOLD response to these two conditions might be caused by different processing times. And third, the interaction between cohesion and coherence confirms that both lexical and pragmatic properties were used for the coherence judgment. Thus, a description of the interplay between these two contributing factors becomes feasible.

A fourth criterion concerns the generalizability of the results to other non-student populations, an important requirement for applying the task in patient studies. Whether the coherence judgment task is equally easy for an older, non-student group is the topic of the next chapter.

Chapter 3
Inference processes across the life span

The study described here is going to appear in *Aging, Neuropsychology and Cognition* (Ferstl, 2006).

For applications of the paradigm in neuropsychological studies, an evaluation of possible age effects is necessary. Adult patients with brain damage are more heterogenous with respect to education and occupation and they are from a different age bracket as the student populations tested in experimental cognitive psychology. Before concluding that a possible deficit in the coherence judgment task might be due to a patient's brain lesion, it is therefore necessary to exclude the possibility of age-related changes in performance even in the absence of brain damage.

Language processes are considered relatively resistant to age-related decline (Kliegl, Mayer, Junker & Fanselow, 1999). More specifically, only those language comprehension tasks are more difficult for older people that use speeded conditions, require allocation of resources or working memory functions, inhibitory functions, or the verbatim use of the input (cf. Burke, 1997; Hasher & Zacks, 1988; Obler, Fein, Nicholas, & Albert, 1991; Salthouse, 1996; Zacks & Hasher, 1988). In contrast, when general world knowledge, plausibility evaluations or integrative processes are needed, older people do not show any decrements in performance, and sometimes even excel in meaning based tasks (Stine-Morrow & Miller, 1999; Wingfield & Stine-Morrow, 2000; Zacks, Hasher & Li, 2000; Baltes & Baltes, 1990). The same idea is reflected in distinctions between gist-based and item-based processing (Koutstaal & Schacter, Galluccio & Stofer, 1999), or cristalline vs. fluid intelligence (Horn, 1982). Common to all of these approaches is the observation that older people rely more on pragmatic information and general world knowledge. By slowly adapting their processing strategies to take optimal advantage of this rich store of experience, older people are able to compensate for the concurrent decline in "hardware", i.e., in functions such as working memory or rapid attention allocation (Raz, 2000).

In psycholinguistic terms, the literature suggests that the situation model level is intact in the elderly whereas the surface and textbase levels are subject to secondary effects of cognitive decline. Studies on the inference abilities of older comprehenders confirm this general pattern. Valencia-Laver and Light (2000) stress the importance of the task used for testing inference abilities. While data from off-line procedures hinted at age-related deficits, on-line measures often confirm that inference generation is mostly intact (Valencia-Laver & Light, 2000; Hess, 1995; Hamm & Hasher, 1992). Radvansky, Zwaan, Curiel & Copeland (2001) used a sentence recognition task to test the strength of the different text representation levels. When the statements were from the situation model level, however, older adults 60 to 90 years of age performed equally well as younger adults. However, older adults had difficulties with verbatim or paraphrase statements, indicating problems on the surface structure and text base levels.

An open question is when these text comprehension differences appear during the life span. In the Day Clinic for Cognitive Neurology at the University of Leipzig, a majority of patients either suffered closed-head injury, with a typical age range of about 25-40, or have vascular etiologies, with a typical age of about 45-60 years. These age groups are younger than the elderly population usually studied in gerontological research. Although there is evidence for age-related changes in syntactic and semantic processes long before retirement age (Gunter, Jackson & Mulder, 1992; 1998), but studies on text comprehension skills in middle aged adults are rare (Hess, 1995; van der Linden et al., 1999; Ulatowska, Hayashi, Cannito & Fleming, 1986).

For the coherence judgment paradigm, however, an analysis of the task requirements suggests little or no age-related decline. The coherence judgment, as described before, can easily be based on the situation model alone. Although the task is to create a bridging inference between two successive sentences, the incoherent trials are already characterized by a topic change, in which different situation models or scenarios are to be combined. Furthermore, the sentences are short and syntactically simple, so that working memory demands remain low. The sentences were auditorily presented at normal speaking rate and the responses were to be given without any time pressure.

> **Hypothesis 3.1:** The performance in the coherence judgment task remains stable across the life span. Observable decrements for older participants are restricted to general slowing, but they are not due to qualitatively different processes.

Using detailed analyses of reading time patterns, Stine-Morrow & Miller (1999) showed that older readers focus less on surface level features and concentrate more on the situation model level (see also Radvansky, Gerard, Zacks, & Hasher, 1990; cf. Baltes & Baltes, 1990). Such a strategy is expected to develop gradually, so that it might be apparent even in middle aged subjects. In our paradigm, a strong reliance on surface level features would predict a clear interaction between cohesion and coherence, whereas a focus on the pragmatic content would render cohesive ties less effective. Thus, an evaluation of the influence of cohesive ties, a feature on the surface level, can be used to see whether a similar shift in resource allocation takes place for the coherence judgment task.

> **Hypothesis 3.2:** If older comprehenders rely less on surface level features, the influence of cohesive ties is expected to diminish across the life span.

To test these hypotheses, I conducted an auditory version of the coherence judgment task with a group of 39 adult participants. Their age was uniformly distributed across the range of 26 – 64 years, with an average of 45 years. The group of participants resembled the typical patient population with respect to education and occupation. Because the sentences were played over head phones at normal reading speed, the response times of the coherence judgements were measured from the offset of the sentences. Thus, it was not necessary to correct for sentence length as in Experiment 1.

The pattern of results, displayed in Figure 3.1, is similar to that of the student group tested in Experiment 1. Most importantly, the accuracy was as high as that for the younger student participants, with less than 6% errors overall. This result confirmed that the task was appropriate for use with a middle aged, non-student population.

Qualitatively, the results also replicated the previous findings. For both errors and response times the expected interaction between Coherence and Cohesion obtained. The facilitation of coherent trials by cohesive markers was even more pronounced than in the previous experiment. Thus, there was no evidence for a reduced use of lexical cues in the older group of participants.

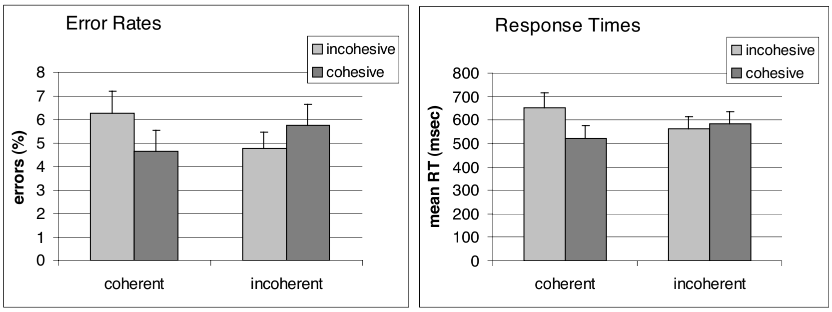


Figure 3.1. Error rates and response times in the coherence judgment task for a group of middle aged participants (26 – 64 years).

Scatterplots of the individual data are displayed in Figure 3.2. As can be seen, age had an impact neither on overall error rate, nor on overall response times. The only significant age effect was an interaction between coherence and age in the response times. A few of the oldest participants in our sample took somewhat longer for coherent than for incoherent trials.

As stated in hypothesis 3.1, these results are consistent with the expectation that the coherence judgment be based on conceptual or situation model processes, and that age effects would be minimal. Only three participants made more than 10% errors, and they were in an intermediate age range from 30 to 50, rather than among the oldest participants.

In contrast to hypothesis 3.2, there was no evidence for a differential use of the cohesive ties by older comprehenders. As in the previous reading version of the task, the interactions between Cohesion and Coherence in both error rates and response times confirmed the efficacy of cohesion for aiding coherence building. Importantly, these interactions did not depend on the participants' ages, but remained similar across the age range.

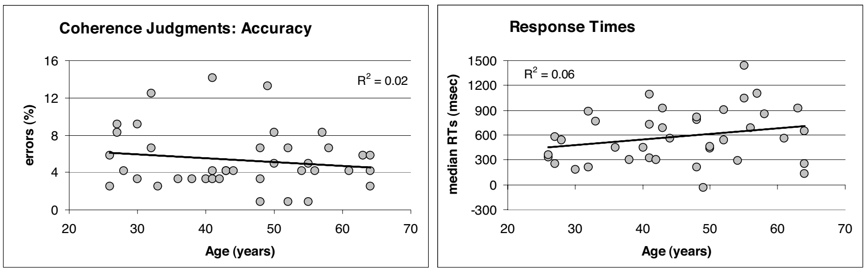


Figure 3.2. Scatterplots of the error rates and response times (measured from sentence offset) as a function of the participant's age. The task was equally easy across all age levels.

Taken together, the results from both behavioral studies show that the coherence judgment task is appropriate for use in neuropsychological research. Despite the simple YES/NO-decision and the corresponding ceiling effect in performance, qualitative effects due to the presence or absence of a cohesive marker were observed in both experiments. Importantly, these effects were comparable across a wide range of age levels. It remains an open question, however, whether these results generalize to elderly comprehenders as usually tested in studies on cognitive aging.

Chapter 4
Coherence building: Lesions of left prefrontal cortex

The results described in this chapter are published in *Neuropsychology* (Ferstl, Guthke & von Cramon, 2002).

As summarized in the introduction, inference deficits have been observed after RH damage. There is an extended literature using different materials and tasks confirming that RH patients have difficulties with inference processes (for reviews see Lehmann & Tompkins, 2000; Brownell & Martino, 1998; Beeman, 1993; Nicholas & Brookshire 1995). It has also been pointed out, however, that the qualitative descriptions of communication deficits after FL damage closely resemble that of the RH syndrome (McDonald, 1993; Nicholas & Brookshire, 1995). Moreover, inspecting the methodological details of many inference studies, Lehmann and Tompkins (2000) argued that the evidence for a RH inference deficit is less convincing when adopting usual experimental standards. In particular, most studies use off-line tasks which do not tap immediate, automatic inference processes but require memory and language production. Furthermore, in many studies the contents of the experimental materials prohibit a clean separation of text comprehension processes on one hand from "hot cognition" (Goel & Dolan, 2004), such as affective, emotional, social or pragmatic factors (Brownell & Martino, 1998).

A further reason for the overlap between frontal and RH communication deficits lies in patient selection procedures, and in particular, in the specificity of the brain damage considered. In many studies, RH patients are compared to a control group of healthy adults. Thus, eventual effects might be due unspecific consequences of brain damage, rather than specific RH pathology. A comparison with a matched control group of LH patients is seldom feasible. The aphasic deficits of this latter group are seen to prohibit them from taking part in experiments requiring complex language comprehension. And finally, the overlap in behavioral deficits of the two patient populations might be caused by an overlap of their lesion locations. Many studies on RH communication test patients predominantly with cerebro-vascular etiologies, and in particular with infarctions of the middle cerebral artery. Thus, the RH patients' lesions often encroach upon the frontal lobe (McDonald, 1993; Brownell & Martino, 1998). In contrast, many studies on frontal communication deficits study CHI patients. Although it is clear that CHI patients have a high likelihood of frontal pathology (e.g., Adams & Victor, 1993), and their focal lesions can be localized, diffuse axonal injury can lead to dysfunctions of a larger fronto-temporal network.

The literature on text comprehension skills after CHI or other causes of FL damage provides clear evidence for deficits similar to those of RH patients. In particular, FL or CHI patients have benn shown to have problems with inferencing and the weighing of information (Nicholas & Brookshire, 1995; Haut, Petros & Frank, 1990), the derivation of a story topic (Hough, 1990), memory for text (Novoa & Ardila, 1987; Kaczmarek, 1984, 1987), question answering (Channon & Crawford, 2000), and the goal-directed use of background knowledge (Ferstl, Guthke & von Cramon, 1999).

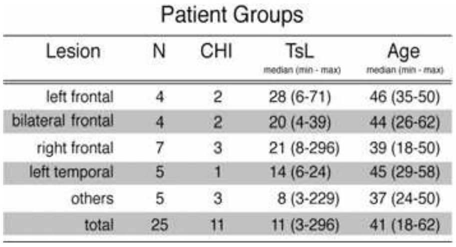
For the patient study described in this chapter, care was taken to select patients with circumscribed lesions, so that the respective influence of the brain lesions' location and lateralization could be evaluated. Based on the radiological diagnosis, frontal and temporal lesions were considered. The group of interest consisted of patients with FL lesions. According to the theories outlined in the introduction, the predictions for the experiment are as follows:

> **Hypothesis 4.1:** According to the RH hypothesis, patients with right-sided lesions are expected to have difficulties with the coherence judgment task, whereas the FL hypothesis predicts inference deficits for all patients with FL damage.

The participants were current or former patients of the Day Clinic of Cognitive Neurology at the University of Leipzig. The patients' therapists provided the clinical diagnosis, as well as the scores on a variety of neuropsychological tests. 25 non-aphasic patients were tested and their lesion locations were scored based on anatomical MRI scans. All patients were in the chronic stage of their illness, at least three months after suffering brain damage. The age range of 18 to 62 years corresponded well to that of the previous experiment (Chapter 3).

Based on the radiological diagnosis, five groups were then formed. Patients in the control group had no frontal or left temporal lesions (C). The patients with no frontal, but left temporal lesions formed the left temporal group (LT), and the patients with frontal lesions were split into three groups according to the lesion lateralization (RF: right frontal, LF: left frontal, BF: bilateral frontal). Table 4.1 provides information on the demographic characteristics of the participants.

Table 4.1. Characteristics of the participants. About half of the patients had closed-head injury (CHI). Time since lesion is provided in months, age in years.



Patient Groups

Lesion	N	CHI	TsL median (min - max)	Age median (min - max)
left frontal	4	2	28 (6-71)	46 (35-50)
bilateral frontal	4	2	20 (4-39)	44 (26-62)
right frontal	7	3	21 (8-296)	39 (18-50)
left temporal	5	1	14 (6-24)	45 (29-58)
others	5	3	8 (3-229)	37 (24-50)
total	25	11	11 (3-296)	41 (18-62)

Identical to the procedure of the experiment described in Chapter 2, the sentences were presented visually. To accomodate the patients' increased processing times and possible attentional problems, the experiment was shortened to 80 trials.

The accuracy data are shown in Figure 4.1. The patients in the control group (C), but also the RF and LT patients, had no difficulty with the coherence judgment task whatsoever. The qualitative pattern across the four conditions confirmed that all patient groups, except the LT group, used cohesive ties in the same way as healthy control subjects. Most importantly, the two groups whose lesions reached into the left-frontal lobe (LF and BF) had dramatically increased error rates for the coherent condition, but not for the incoherent condition. Thus, these patients exhibited a clear deficit which was due to finding a plausible connection between coherent sentences, rather than to detecting a break in coherence.

Post-hoc comparisons confirmed that this strong effect was independent of etiology and independent of the particular grouping criterion. In particular, for testing the RH hypothesis more stringently, all patients with RH damage (n =

13) were compared to the remaining participants. Even though this alternative RH group included the bifrontal patients whose performance was poor, there was no evidence for this results to generalize across the RH group (9.5% vs. 7.3% errors for the no-RH and RH groups, respectively).

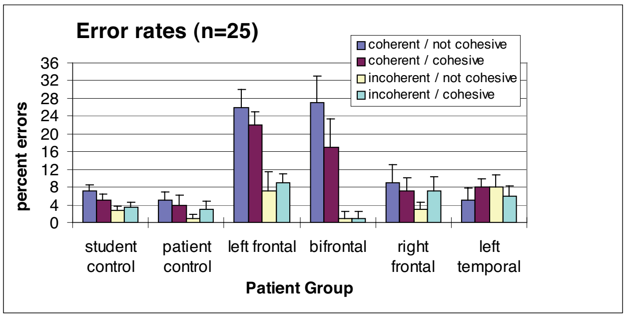


Figure 4.1. Error rates in the coherence judgment task for a group of 25 patients with brain lesions. For better comparison, the result from the student control group (see Chapter 2) are also shown.

A similar coherence effect was seen in the response time data. Only the LF patients took considerably longer for their YES responses than for their NO responses. To illustrate the homogeneity of this result across subjects, Figure 4.2 shows a scatterplot of individual data. The graph shows the coherence effect, i.e, the difference between coherent trials and incoherent trials in the response times as a function of the coherence effect for the error rates. For reducing the large interindividual variability, the response times had first been transformed into z-scores. It can be seen that the score for seven of the eight LF patients falls into the upper right quadrant, indicating that both measures showed increased difficulty for the coherent trials. In contrast, for most patients without LF lesions, the reaction times showed the opposite effect, consistent with the well-known psycholinguistic finding that coherence facilitates processing (Keenan, Baillet & Brown, 1984; Myers, Shinjo & Duffy, 1987).

To exclude the possibility that these results were secondary to accompanying neuropsychological deficits, the patients' clinical profiles were considered, as reflected in standard diagnostic tests (e.g., WMS-R, Wechsler, 1987; BADS, Wilson et al., 1996; TAP, Zimmermann & Fimm, 1992; Horn, 1983; CVLT, Delis, Kramer, Kaplan, & Obler, 1987; MWT, Lehrl et al., 1991; digit ordering test

(DOT), Werheid, et al, 2002). The groups were well matched with respect to memory and attention deficits. Correlations with the test results confirmed that the error rates on coherent and incoherent trials tapped qualitatively different processes. Long-term memory measures were related to errors on incoherent trials, whereas correlations of errors on coherent trials with measures tapping strategy use (e.g., verbal fluency) strengthened the conclusion that executive functions are necessary for successful inferencing.

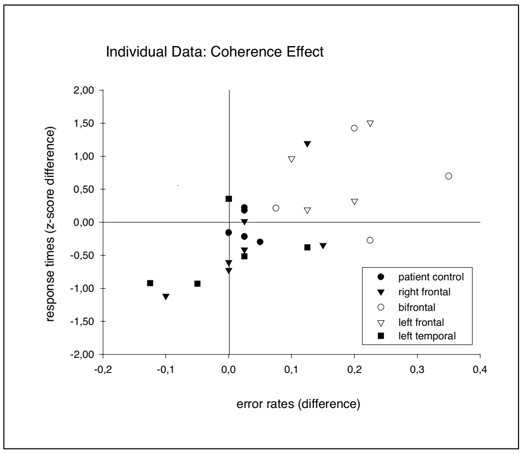


Figure 4.2. The individual data for all 25 patients. Shown is the coherence effect in the response times as a function of the coherence effect in the error rates. The two dependent measures converge for most patients. Most importantly, the data of patients with left- or bifrontal lesions fell into the upper left quadrant, indicating consistent difficulties with coherent trials compared to incoherent trials.

The data of the patient study confirmed the hypothesis of the left frontal lobe being important for inference processes. With this particular task and patient group, the RH hypothesis could not be supported. Neither right frontal nor left temporal lesions produced a comparable deficit. The increases in error rates and response times affected the coherent condition only. Thus, the LF patients had a problem with inferencing, rather than with erroneously activating possible but implausible connections.

Chapter 5
The neuroanatomy of inference processes: an fMRI study

The study summarized in this chapter is published in *Cognitive Brain Research* (Ferstl & von Cramon, 2001).

The patient study provided clear evidence for left frontal brain regions being more important for the coherence judgment than the right hemisphere. For a more specific delineation of the subregions involved during the coherence judgment task, we conducted a study using functional magnetic resonance imaging (fMRI; for methodological introductions see Cabeza & Kingstone, 2001; Jezzard, Matthews & Smith, 2001; Schwartz, Kischka, & Rihs, 1997). To derive more specific hypotheses, the results of previous imaging studies on coherence building are briefly summarized (cf. Mar, 2004; Ferstl, in press, b; Ferstl & von Cramon, 2005).

Fletcher et al. (1995) conducted a PET experiment to study the functional neuroanatomy of Theory-of-Mind (ToM) or mentalizing processes. ToM refers to the appreciation of other people's motivations, goals, and feelings as a driving force for their behavior (Frith & Frith, 1999; 2003). Using story materials developed for the assessment of ToM abilities (Happé, 1994), Fletcher et al. (1995) compared the reading of stories with or without a ToM component to that of unrelated sentences. The comparison of unrelated sentences to the stories yielded bilateral activation in the middle frontal gyrus and the dorsal precuneus, indicating increased demands on memory and attention (cf. Gruber, 2001; Brass, Zysset & von Cramon, 2001; Fink et al., 1997). Behavioral data confirmed that encoding for subsequent comprehension questions was more difficult when the sentences were unrelated. In the reverse comparison, stories elicited activation in the aTL bilaterally and the left temporo-parietal junction, as well as an area in the lower precuneus (PCC/prec; cf. Zysset et al., 2002, 2003). Partly consistent with the results of Mazoyer et al. (1993), the dmPFC (BA 8) was also active – but only in those stories that contained a ToM component.

Two studies on situation model building, to which I will return in Chapter 7, are also relevant here. Coherence was manipulated by providing or omitting necessary background knowledge (St. George, Kutas, Martinez & Sereno,1999; Maguire, Frith & Morris, 1999). Despite the similarities of the studies the results diverged. While St. George et al. (1999) reported temporal RH activation for incoherent stories, and interpreted this result as consistent with the RH hypothesis, Maguire et al. (1999) found the aTL, PCC/prec and the mPFC (BA 10/11) to be sensitive to the coherence manipulation.

Similar to the design of the coherence judgment paradigm, Robertson et al. (2000) investigated the influence of cohesion on coherence building. In their fMRI experiment lists of sentences with definite articles were compared to the same sentence lists with indefinite articles. The former version facilitates a coherent interpretation of the unrelated sentences, because definite noun phrases are assumed to refer to a part of the already established situation. In the comparison of the language trials with non-letter strings, the left-dominant extended language network (ELN) was reported, with the right-sided activations restricted to the temporal lobe. As a function of the articles' definiteness, the authors reported a lateralization difference in prefrontal cortex. The less coherent sentence lists activated the left IFG and a region in the fronto-medial cortex known to be sensitive to error processing and response competition (Ullsperger & von Cramon, 2001). The more coherent trials with definite articles, in contrast, elicited activation in the right IFG. These lateralization differences were interpreted as consistent with the RH hypothesis (cf. Beeman, 1993).

Taken together, these studies once again confirm the importance of the ELN regions for language processing in context. Coherence building was reflected either in activations in parts of this network, particularly the aTL, dmPFC and PCC/prec (Fletcher et al., 1995; Maguire et al., 1999), or in right hemisphere activations (St. George et al., 1999; Robertson et al., 2000). However, the discrepancy between these results could not yet be attributed to differences in methods or materials. Finally, cohesive ties as a cue to coherence engaged the lateral prefrontal cortex (Robertson et al., 2000; cf. Friederici, Opitz & von Cramon, 2000).

Based on these suggestions from the literature, three predictions for the coherence judgment paradigm follow:

> **Hypothesis 5.1**. The comparison of language comprehension to the non-word baseline is expected to activate an extended bilateral, but left-dominant network of fronto-temporal regions.
>
> **Hypothesis 5.2**. Subregions of the left language cortex, i.e. left lateral fronto-temporal regions, are expected to be sensitive to lexical cohesion.
>
> **Hypothesis 5.3**. With respect to inferencing, the RH hypothesis predicts right-lateralized fronto-temporal regions to be sensitive to coherence. The FL hypothesis, on the other hand, predicts FL activation.

To test these hypotheses, the written version of the experiment was administered while participants were scanned. In addition to the sentence pairs as described in Chapter 2, non-word letter strings were presented as a control condition. To match the decision and integration components of the coherence judgment task, the participants were asked to decide whether the appearance of the two letter-string sentences (either upper and lower case, or all upper case) were equal or different. Comparing the BOLD contrast elicited by the language trials to that elicited by the baseline was expected to control for perceptual and response components of the task, and thus to show a reflection of the brain activation during text comprehension.

The analysis of the BOLD contrast used the general linear model (cf. Lohmann et al., 2001; Friston et al., 1994), time-locked one second before the end of the target sentence. At this point, processing of the surface level has been completed, and the integration of the current sentence with the discourse context proceeds.

In the overall comparison of the language trials against the baseline, the general pattern as sketched in Chapter 1 was found. A left-dominant fronto-temporal network of brain regions was more engaged during language processing in context than during the processing of non-word letter strings. The activations from the group analysis, overlaid on an individual brain, are shown and described in Figure 5.1.

Sentences vs. Control Condition

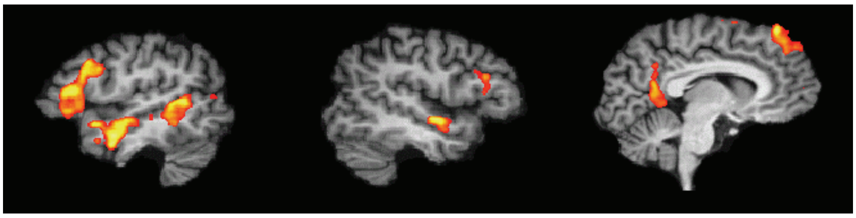
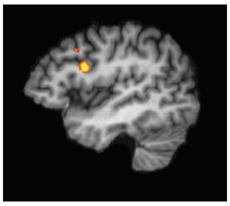


Figure 5.1. Patterns of activation for the comparison of all language trials, independent of their coherence, to the baseline consisting of non-word letter strings. The left-dominant network includes left lateral activations along the superior temporal sulcus, the inferior frontal gyrus, and the more dorsal junction of the inferior frontal with the precentral sulcus (left panel). In the right hemisphere (middle panel), the anterior temporal and dorsal prefrontal regions are significant as well. The left medial view shows retrosplenial and fronto-medial activations (right panel).

The results from the specific contrasts evaluating the influence of coherence and cohesion are shown in Figures 5.2 and 5.3. Illustrating the significant interaction between these two factors, Figure 5.2 displays the pairwise comparison of the cohesive with the incohesive incoherent trials. This comparison sheds light on how the language processing system deals with the mismatch between the lexical cohesion – which signals a connection – and the pragmatic unrelatedness. The result shows left-dominant activation in the area at the junction of the precentral and inferior frontal gyri. This region is also active during executive function tasks, such as dual task, Stroop, task switching, and many more (Dove, et al., 2000; Brass, Zysset & von Cramon, 2001; Brass, Derrfuss, Forstmann & von Cramon, 2005). In the context of the coherence judgment, this activation reflects the increase in task difficulty for the cohesive, incoherent condition. We cannot decide whether the posterior lateral PFC activation indicates the integration of inconsistent information, the inhibition of the irrelevant or erroneous cohesive markers, or the reactivation the current task set. All of these explanations have in common that they refer to a domain-general executive function.

An alternative interpretation, more consistent with hypothesis 5.2, would ensue if we anatomically allocated this IFG activation to Broca's area. And in fact, a recent meta-analysis has localized the inferior junction area more dorsally and posteriorly in the foot region of the middle frontal gyrus (Derrfuss, Brass & von Cramon, 2004). The functional attribution would then focus on the increase in lexical and/or syntactic processing (see Stowe, Haverkort & Zwarts, 2005) induced by the cohesive markers (cf. Friederici, Opitz & von Cramon, 2000). However, this explanation would predict an engagement of the left IFG for all cohesive sentences rather than for the mismatch sentences only.



Incoherent only: Cohesive - Incohesive

Figure 5.2. Activations of the group analysis superimposed on an individual brain. Shown is the lateral prefrontal activation indicating the pragmatic garden-path.

The contrast most important for an evaluation of hypothesis 5.3 was the comparison of coherent and incoherent trials. No brain region was significantly more activated when the sentences were unrelated as compared to when they were coherent. In the reverse contrast, two left medial regions emerged. As shown in Figure 5.3, there was activation in the fronto-medial wall (dmPFC: BA 9/10) and in the posterior cingulate cortex and lower precuneus (PCC/prec: BA 23/31).

These results are important for the evaluation of the RH hypothesis (cf. Chapter 1; Beeman, 1998; Brownell & Martino, 1998). Although the RH contributed to language processing in context, as shown in the overall comparison, the focussed contrasts did not uncover contributions of specific regions within the non-dominant hemisphere (cf. Mason & Just, 2004). Thus, rather than supporting the RH hypothesis, the findings once more confirmed the role of the left prefrontal cortex, and in particular the dmPFC for language processing in context.

Coherent vs. Incoherent

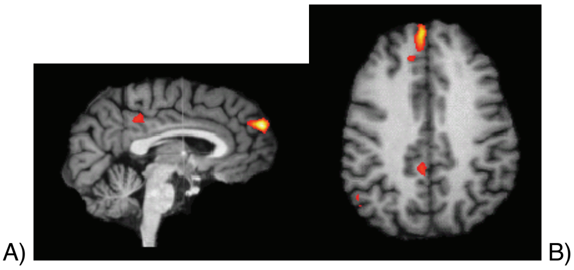


Figure 5.3. Activations of the group analysis superimposed on an individual brain. Shown is the sagittal (A) and the horizontal view (B). The result from the contrast comparing coherent to incoherent trials includes a fronto- and a parieto-medial region.

The functional attribution of the dmPFC rests on clinical descriptions of patients after infarctions of the anterior cerebral artery. These patients often exhibit problems in the realm of motivation and drive. They suffer from a reduction in self-guided activities, which in extreme cases might include akinetic mutism (Marin, 1991). With respect to language, a reduction in self-initiated language production has been observed. In the context of the coherence judgment task, we interpreted the fronto-medial activation as reflecting the self-guided initiation of a non-automatic, knowledge based inference process. The knowledge aspect of this process was reflected in the concurrent activation of the PCC/prec (cf. Zysset et al., 2002; 2003; Maguire, et al., 1999; Fletcher et al., 1995).

A closer look at the neuroimaging literature on the dmPFC suggested an alternative explanation, though. As summarized earlier, several studies using language stimuli also reported this area to be active (Fletcher et al., 1995; Vogeley et al., 2001). However, some of these studies investigated a very different process: Theory-of-Mind. Whether the dmPFC activation for coherent trials might have been due to concurrent mentalizing activities is the topic of the next chapter.

Chapter 6
Coherence or Theory-of-Mind?

The full account of the experiment reported in this chapter was published in *Neuroimage* (Ferstl & von Cramon, 2002).

The fronto-medial area found to be crucial for inferencing in Chapter 4 has received much attention recently. Rather than having been associated with language processing, the functional proposals reach from cognitive processes related to inferencing, such as inductive reasoning (Goel & Dolan, 2001) or evaluation (Zysset, Huber, Ferstl & von Cramon, 2002), to very different types of functions, such as emotional processing, the processing of self-relevant stimuli or a default state in the absence of stimulation (Greene et al., 2001; Gusnard, Akbudak, Shulman & Raichle, 2001; Raichle, et al, 2001; Simpson et al., 2001a, 2001b). The most important function attributed to this region are Theory-of-Mind processes (Frith & Frith, 1999; 2003; Fletcher, et al., 1995; Vogeley et al., 2001; Saxe, Carey & Kanwisher, 2004). Theory-of-Mind (ToM) or mentalizing refers to the ability to understand that other people "have a mind of their own", i. e., that they base their actions and decisions on their own beliefs, goals and motivations (cf. Premack & Woodruff, 1978; Brothers, 1990; Baron-Cohen, Leslie & Frith, 1985). Developmental studies show children acquire a full grasp of ToM, as reflected for instance in the understanding of false belief or deception tasks, only at an age of about 4 years (Wellman, 1993). Whether this development is primary or whether it hinges on the acquisition of complex language and executive functions is a matter of debate (cf. Baron-Cohen, Tager-Flusberg & Cohen, 1995, 2004; Baron-Cohen, 2003).

There is overwhelming evidence that the dmPFC plays a role during ToM processes. In addition to the story comprehension task described earlier (Happé, 1994; Fletcher et al., 1995; Gallagher et al., 2000, Vogeley, 2001), in which mental state inferences are compared to so-called "physical" inferences, a variety of other ToM type tasks have been used. These include the recognition of emotional facial expressions (Baron-Cohen, 2003), the attribution of intentions (Castelli, Happé, Frith & Frith, 2000; Brunet, Sarfati, Hardy-Baylé & Decety, 2000), identifiying with people from another historical period (Goel, Grafman, Sadato, & Hallett, 1995) or the comprehension of cartoon stories with or without a ToM component (Gallagher et al., 2000). The results show an engagement of the dmPFC during ToM processes, independent of the modality (verbal or non-verbal), the aspect of ToM (e.g., false belief, intention, gaze direction), or the specific task instructions.

When inspecting the example from Chapter 2 about Charlotte's birthday invitation, it is obvious that the scenario includes inferences about the protagonist's goals and motivations, and that connecting the sentences to each other necessitates inferences on the characters' mutual state of mind. Consequently, an alternative, plausible account for the findings of the frontomedial activation during inferencing is that the coherence of the sentence pairs could be evaluated based on an automatically elicited, concurrent ToM process. The comprehender decides whether the behavior of the protagonists makes sense, rather than whether a plausible briding inference was created. And in fact, following the linguistic relevance theory (Sperber & Wilson, 1995) Frith & Frith (2003) argue that the core of communication is understanding the motivations and intentions of other people, be it the protagonists of stories, or the communication partner. Thus, it might be that ToM and the use of language in context are inseparable (cf. Ferstl, in preparation).

Conversely, it has been suggested that autism, which can be conceptualized as an extreme ToM deficit which leads to impaired social behavior, might be explained by domain-general deficit of global, integrative processes (Frith, 1989; Happé, 2000) This so-called *central coherence* theory has also been tested using inferences in text comprehension (Jolliffe & Baron-Cohen, 1999, 2000). Thus, a close link between coherence building and ToM is likely.

The experiment described in this chapter was designed as an attempt to tear apart the relative contributions of ToM and inferencing to the activations observed in the coherence judgment paradigm. Using an auditory version, the goal was to minimize ToM processes by conducting the coherence judgment task on only those sentence pairs that did not mention humans. The first two examples for each of the conditions in Table 2.1 are animate, the last two inanimate sentence pairs. In the experiment, the same materials were used, but the order of presentation was controlled. In the first part of the experiment (Logic), only inanimate sentence pairs were presented. The task was to decide whether the sentences had a logical connection. The behavioral results confirmed that the responses were indistinguishable from those of the previous coherence judgment task. In the second part of the experiment (ToM), the animate sentence pairs were presented. In order to explicitly elicit ToM processes during this part of the experiment, the participants were instructed to decide whether they could understand the protagonists' feelings, motivations, intentions, and actions. The question of interest was which of the resulting four conditions would engage the dmPFC. The pattern of activation across the conditions provides information on the necessity and sufficiency of coherence building and ToM processes, respectively. Concerning the dmPFC involvement, the following hypothesis was derived:

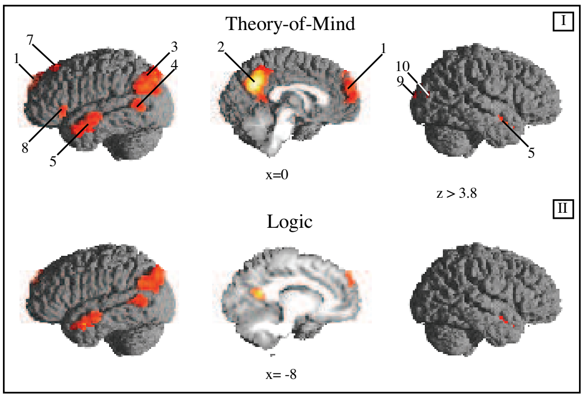
> **Hypothesis 6.1.** If the function of the dmPFC was ToM, it is expected to be engaged during all trials in the ToM block. If the function was coherence building, the dmPFC is expected to be active during coherent trials in both blocks. If both processes are necessary, only the coherent trials in the ToM block are expected to elicit dmPFC activation, and if both are sufficient (i.e., there is a domain-general overarching process), then the dmPFC should contribute in all conditions, except during the incoherent trials in the Logic block.

The results from scanning 9 participants during the two tasks are displayed in Figure 6.1. The upper two panels show the extended language network which in both blocks emerges from the comparison of all language trials to the perceptual baseline. The medial activations, both in dmPFC as well as in the PCC/prec were somewhat stronger and more extended in the ToM-block. Consistent with the hypothesis that coherence building is sufficient for engaging the dmPFC, there were significant contributions of the dmPFC to coherence building, even if ToM processes played little or no role (lower panel).

Consistent with the hypothesis of ToM processes being sufficient for engaging the dmPFC, both conditions in the second part of the experiment activated this region equally (Figure 6.1, upper panel (I)), A comparison of the time courses for the four conditions displayed in Figure 6.2 confirms that the coherent, non-ToM condition led to an increase in percent signal change, rather than to a modulation of deactivation (cf. Raichle et al., 2001).

These results are consistent with the third part of Hypothesis 6.1. Neither ToM nor coherence alone are necessary for dmPFC activation, but both are sufficient. Thus, we concluded that the functionality of the fronto-medial structures was related to a domain-independent general process, encompassing both ToM as well as coherence building. A proposal for what this process might be follows in the discussion in Chapter 8.

Language-Control



Logic: Coherent-Incoherent

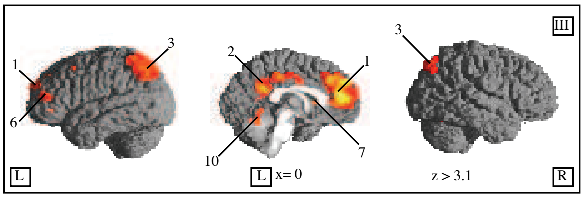


Figure 6.1. Activations of the group analysis overlaid on a mean brain of the participants. The comparison of language trials against baseline provides another replication of the general pattern described in Chapter 1. Besides apparently stronger activation in the medial structures in the ToM part (I), compared to the Logic part (II), the networks are remarkaby similar. The lower panel (III) shows the comparison of coherent to incoherent trials in the Logic part. In addition to the replication of fronto- and parieto-medial activations, this contrast revealed contributions of prefrontal and parietal areas.

Independent of the exact interpretation of the dmPFC activation is the fact that we can use it as an indicator of knowledge based inferencing activities. In addition to the two experiments reported in Chapters 5 and 6, we have recently shown that the dmPFC activation can be modulated by task instructions and thus does not merely depend on the stimulus properties of the sentence pairs (Siebörger, Ferstl, & von Cramon, 2003). In this study, the peak latencies of the hemodynamic response were related to the decision component of the coherence judgment, rather than to the comprehension phase.

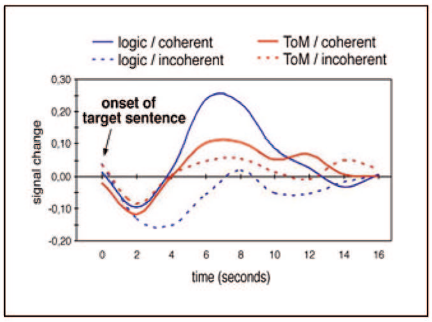


Figure 6.2. Time course of activation in the dmPFC (BA 9/10; [-6,45,12]). The largest signal increase was observed for the coherent condition in the non-ToM part of the experiment.

Zysset, Huber, Ferstl & von Cramon (2002; Zysset et al., 2003) provide further evidence for this functional interpretation. In their studies on evaluative decision making, a modulation of the relative contributions of the fronto- and parieto-medial cortices was observed. Comparing evaluative judgments with judgments based on episodic memory, the evaluative component drove the frontal activation, whereas the knowledge component drove the parieto-medial activation. Finally, there are studies implicating the dmPFC for inductive reasoning processes (Goel & Dolan, 2001). Taken together, these findings provide sufficient evidence to conclude that the dmPFC in concert with the PCC/prec region implements non-automatic, explanation-based inference processes (cf. Singer, Graesser & Trabasso, 1994; Gueraud & O'Brien, 2005).

shown that the dmPFC activation can be modulated by task instructions and thus does not merely depend on the stimulus properties of the sentence pairs (Siebörger, Ferstl, & von Cramon, 2003). In this study, the peak latencies of the hemodynamic response were related to the decision component of the coherence judgment, rather than to the comprehension phase.

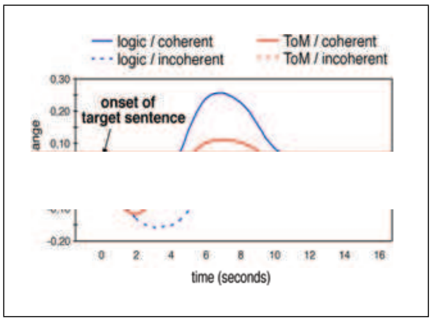


Figure 6.2. Time course of activation in the dmPFC (BA 9/10; [-6,45,12]). The largest signal increase was observed for the coherent condition in the non-ToM part of the experiment.

Zysset, Huber, Ferstl & von Cramon (2002; Zysset et al., 2003) provide further evidence for this functional interpretation. In their studies on evaluative decision making, a modulation of the relative contributions of the fronto- and parieto-medial cortices was observed. Comparing evaluative judgments with judgments based on episodic memory, the evaluative component drove the frontal activation, whereas the knowledge component drove the parieto-medial activation. Finally, there are studies implicating the dmPFC for inductive reasoning processes (Goel & Dolan, 2001). Taken together, these findings provide sufficient evidence to conclude that the dmPFC in concert with the PCC/prec region implements non-automatic, explanation-based inference processes (cf. Singer, Graesser & Trabasso, 1994; Gueraud & O'Brien, 2005).

Chapter 7
Situation model building: Integrating emotional and temporal information

The experiment summarized in this chapter appeared in the *Journal of Cognitive Neuroscience* (Ferstl, Rinck & von Cramon, 2005).

The coherence judgment paradigm provided evidence for a contribution of the left prefrontal cortex, and in particular, for the left dmPFC to inference processes. In contrast, there was neither evidence for a special role of the right hemisphere, nor for any additional activation for incoherent trials. One reason for this latter, somewhat counterintuitive result might have been the special task demands. In everyday language comprehension, coherence breaks are rare and detrimental to communicative success. Thus, they are particularly salient and likely to induce delibarate inference processes or metacognitive processes. In the coherence judgment task, on the other hand, incoherent trials are frequent and task relevant, so that participants are likely to just abandon deeper processing after detecting the coherence break.

The goal of the next study was to investigate inference processes in the context of locally coherent, natural texts. Rather than including as a control condition incomprehensible, unconnected language materials (cf. Maguire et al., 1999; St. George et al., 1999), situation model updating was studied using the inconsistency paradigm (Bower & Morrow, 1990; Rinck, Hähnel & Becker, 2001, Rinck, Gámez, Díaz & de Vega, 2003). In this paradigm, short stories are presented in which some information is locally coherent, but inconsistent with the global situation model. Based on the event-indexing model (Zwaan, Magliano & Graesser, 1995; Zwaan, Langston & Graesser, 1995; Rinck & Weber, 2003), several different information aspects are varied. The detection of the inconsistency, or an increase in reading times correspond to the salience of the target information in the on-going situation model representation. The adoption of this paradigm for an fMRI experiment enabled us to study situation model updating in a differentiated way within a rather natural comprehension task. Before deriving hypotheses and describing the study in more detail, two prior neuroimaging studies explicitly concerned with situation model building are briefly reviewed (see also Chapter 5).

Early convincing demonstrations of the psychological reality of the situation model were the studies by Bransford and colleagues (Bransford & Johnson, 1972; Bransford & Franks, 1972). These experiments established that the verbatim form of sentences could not be accurately remembered as long as they were consistent with the situation described. They also showed that comprehension fails when the reader or listener does not know what the text is about (cf. Bartlett, 1932). Lacking a situation model for loosely structured and semantically vague texts, comprehension failed and participants had severe problems with encoding and recalling text information. When either a title or a picture (Dooling & Lachman, 1971) provided the necessary background knowledge before reading, recall performance improved dramatically. In two imaging studies this paradigm was adopted for testing the influence of background knowledge on brain activation during comprehension.

St. George, Kutas, Martinez & Sereno (1999) used the version in which a title provides the necessary link for hooking up the incoming sentences to the necessary background knowledge. For instance, a description of doing laundry is formulated so vaguely (e.g., *you have to sort everything according to colors*) that the topic is usually not inferrable, and the sentences remain globally incoherent. In a previous ERP study, St. George, Mannes & Hoffmann (1994) had found reduced N400 amplitudes for content words when the title was provided, a result suggesting facilitated semantic integration during comprehensible passages (cf. Kutas & Hillyard, 1980). In their fMRI study, St. George et al. (1999) used the same materials. In the titled, coherent condition, the network of activation was left-lateralized, whereas in the untitled condition, the pattern of activation was clearly bilateral. Consequently, there were interactions between title condition and hemisphere, in particular in middle and inferior temporal regions. The data were interpreted as consistent with the RH hypothesis: the increase in inference demands in the incomprehensible, untitled condition led to increases in RH activation.

Maguire, et al. (1999) used the picture version of this paradigm to study three different factors influencing the ease of situation model building. Based on Dooling and Lachman (1971), descriptive texts were used that are globally incoherent unless a pictorial depiction of the scene provides the necessary background knowledge. Using a number of conditions, including different story types and repeated presentations, Maguire et al. (1999) identified four brain regions differentially sensitive to comprehensibility and coherence, closely resembling the pattern described by Fletcher et al. (1995). Activations in the anterior temporal lobe and the fronto-medial cortex (BA 11) were observed for comprehensible stories, i.e. when the language input could be integrated in a coherent representation. A fronto-polar region was related to recall and memory for previously presented stories. Finally, the PCC/prec was most sensitive to successful situation model building: it was more active during the second presentation as compared to the first, it was more active when a picture aided comprehension compared to an incoherent story, and its activation correlated with the comprehensibility ratings given by individual subjects.

Although these two experiments used very similar paradigms, the results do not even partially converge. A number of methodological differences render the studies incomparable. St. George et al. (1999) provided evidence for the RH hypothesis (cf. Robertson et al., 2000; Mason & Just, 2004), whereas the results of Maguire et al. (1999) confirm the importance of the regions in the extended language network (cf. Fletcher et al., 1995; Ferstl & von Cramon, 2001a, 2002). Common to both studies was the fact that the comparisons of interest were rather global, and the unrelated conditions did not enable the readers or listeners to set up any situation model at all.

As a more specific test of the text contents on situation model updating, we utilized the inconsistency paradigm and conducted an event-related fMRI study using coherent, comprehensible stories only. In all conditions, it was easy to set up a situation model. Two issues were of interest. First, we wanted to identify the linguistic processes needed for detecting and integrating a global inconsistency. And second, based on the event-indexing model (Zwaan, Langston & Graesser, 1995), we were curious about whether different information aspects would influence this updating process. Taking advantage of the previous research on situation model building using an inconsistency paradigm (Rinck, Hähnel & Becker, 2001; Rinck & Weber, 2003; Otero &

Kintsch, 1992), stories varying emotional and temporal or chronological situation model aspects were selected. These information aspects have been shown repeatedly to elicit strong behavioral effects, indicating that readers and listeners closely monitor these types of information. Emotional inconsistencies were created by contradicting the inferred affective status of the protagonist, and chronological inconsistencies were created, for instance by reversing the order of two events. In all cases, the consistent and inconsistent versions differed by one or a few words only. Examples for the two story types and the two versions are shown in Table 7.1.

Table 7.1. Sample stories for the two information types. 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.

Chronological	Emotional
Today, Markus and Claudia would finally meet again. Markus' 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.

The resulting 2 x 2 – design allowed us to independently investigate the effect of consistency and information type, respectively. Although some aspects of the study were exploratory, two rather general a-priori hypotheses were formulated.

Hypothesis 7.1. Emotional and chronological information aspects engage brain areas whose functions are related to these contents. In particular, emotional information is expected to elicit activation in the orbito-frontal or ventro-medial prefrontal cortices, or in parts of the limbic system (e.g., the amygdalae). Chronological information is expected to engage lateral prefrontal, premotor, or parietal regions.

Hypothesis 7.2. Inference processes in globally inconsistent texts resemble those required for local coherence building. Thus, the dmPFC is expected to play a role during situation model updating.

The results were analyzed in two time windows. First, we evaluated the immediate processing of the target word (inconsistent vs. consistent). Independent of the information type, detection of an inconsistent word elicited activation in the right anterior temporal lobe. The left-sided homologue was equally active in both conditions. This result is consistent with the general form of the RH hypothesis. Because the stories were rather similar with respect to their affective connotation (Brownell & Martino, 1998; Lehman & Tompkins, 2000), or their requirements for semantic activation (cf. Beeman, 1998), the most likely, but speculative explanation for this result draws on the resource account (e.g., Navon, 1984). The inconsistencies required more thorough processing of the input, so that the right aTL was engaged more than during the consistent trials.

The comparison between emotional and chronological stories, shown in Figure 7.1., confirmed that the information aspect elicited qualitative differences in processing. Emotional information led to activation in the ventro-medial prefrontal cortex (vmPFC) and the extended amygdala complex, a pattern clearly consistent with the interpretation that merely listening to a story with emotional content suffices for involving brain areas related to emotion processing (Davidson & Irwin, 1999; Luan Phan, Wager, Taylor & Liberzon, 2002; Kringelbach & Rolls, 2004). In contrast, temporal or chronological information activated an extended fronto-parietal network. We interpreted parts of this network, clearly related to attention and working memory, as reflecting the reinstatement search needed for reaccessing the relevant context information (Kintsch & van Dijk, 1978; cf. Rinck et al., 2003, for converging eye movement data). In addition, there was activation in the right temporo-parietal junction (TPJ), an area implicated for the processing of magnitude, spatial relations and - most importantly - temporal information (Walsh, 2003, cf. Dehaene et al., 1999). These results clearly confirm that the specific story information leads to qualitatively different patterns of brain activation.

In a more extended window starting at the target word, but also including the remainder of the stories, we found systematic differences related to the type of available integration processes. In the chronological stories, the target sentence needed to be checked against context information mentioned several sentences earlier. As shown in the left panel of Figure 7.2., inconsistent information, most prominently in the chronological stories, but also seen across all stories, activated the orbital part of the IFG bilaterally (BA 47/11). This area has been implicated for memory based decisions (Petrides, Alivisatos, & Frey, 2002; Nobre, Coull, Frith, & Mesulam, 1999) and for the processing of contextually unexpected information (Baumgärtner, Weiller & Büchel, 2002; Caplan & Dapretto, 2001).

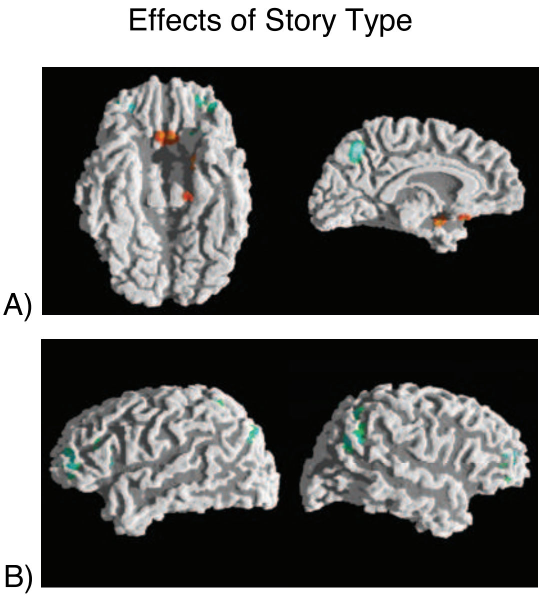
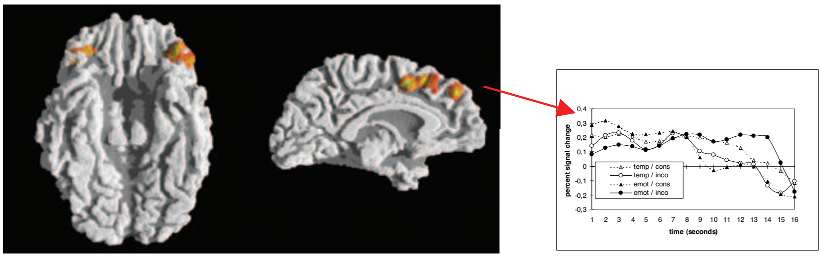


Figure 7.1. Comparison of the activation elicited by encountering the target words in the two story types. Emotional target words – shown in red -, independent of their consistency with the context, engaged the ventromedial cortex and the extended amygdala complex, as displayed in the bottom view and the left medial view (A). Temporal target information – shown in green - activated a widely distributed fronto-parietal network, displayed in the lateral views (B).

In contrast, in the emotional stories the relevant context information had to be integrated across a number of sentences throughout the previous story up to the target information. In addition, the inconsistencies were not all-or-none, but could possibly be reconciled by a further explanation-based inference. For example, when the description of a cheerful party is followed by the statement that the host feels sad, comprehenders can easily construct an elaborative explanation for why this might be the case. Consequently, processing of the information following the inconsistent target information was reflected in increased involvement of the left dmPFC. This result is displays in the right panel of Figure 7.2. The accompanying time course diagram shows that the dmPFC activation remains on a high level until the end of inconsistent emotional stories, whereas it falls off earlier in the remaining three conditions.

Information Integration



all stories *emotional stories only*

Figure 7.2. The integration of inconsistent information activated the lateral inferior/orbito-frontal cortex in both hemispheres, as shown in the left panel. The integration of emotional inconsistencies, in contrast, requires contributions of the medial prefrontal cortex, reaching from the pre-SMA anteriorly into BA 9. The time course diagram illustrates the signal change for the four conditions. In the inconsistent emotional stories, the dmPFC activation remains on a high level until the very end of the story (approx. 8 seconds after the target information).

The design of this study disentangled different subprocesses of situation model building. The finding of right aTL activation for inconsistent story information lends support to the interpretation of the aTL having a role for the immediate construction of the text base. This result is also consistent with the RH hypothesis. Furthermore, confirming hypothesis 7.1. the two distinct situation model aspects elicited qualitatively different patterns of activation, - both during detection and integration -, although in behavioral studies both types of inconsistencies lead to increases in reading times or question answering times. In particular, the vmPFC activation related to emotional processes clearly shows that story comprehension immediately engages the listeners affective system. These findings supports the idea of language comprehension eliciting an immediate immersion in the experiential world (Zwaan, 2004; Glenberg &

Kaschak, 2003). Converging evidence for a qualitative distinction was recently provided in an accompanying patient study. Jentzsch, Ferstl & Guthke (2005) showed that, when controlling memory demands, CHI patients with likely dysfunction of vmPFC had considerable difficulties with the emotional stories only, but not with the chronological stories.

Finally, the role of the dmPFC for inference processes was further elucidated. The integration of emotional inconsistencies into the unfolding story representation engaged this brain region. Although the experiment does not exclude the possibility of increases in emotional (Greene et al., 2001) or ToM processes (Frith & Frith, 2003), this finding is consistent with the idea of the dmPFC implementing underspecified, ideosyncratic, inductive inference processes (cf. Goel & Dolan, 2001; Zysset et al., 2002, 2003).

Chapter 8
Discussion

The results reported in this volume confirm that language processing in context is a highly complex endeavor that requires an intricate interplay of linguistic and cognitive processes. The prerequesite for conducting a number of neuroscientific studies on text comprehension was the development of a paradigm appropriate for both patient and imaging studies. Behavioral data confirmed that the coherence judgment task chosen for this purpose reflected interactions between lexical and pragmatic processes (Chapter 2) and that it was easily understandable for participants of all age and education levels (Chapter 3).

A patient study uncovered specific deficits in this task after left FL lesions. A more differentiated look at the functional neuroanatomy of text comprehension takes advantage of the results of three fMRI studies (Chapters 5-7). Besides the perisylvian language areas, reflecting word and sentence level processes (cf. Bookheimer, 2002; Brown & Hagoort, 1999; Caplan, 1987, 1992; Gernsbacher & Kaschak, 2003; Martin, 2003), there was evidence for a contribution of the ELN, and of brain regions related to working memory, emotion processing, attention, and many more. Several cortical areas emerged as consistently involved during a variety of text comprehension tasks. In the following discussion, the respective roles of the FL and the RH are reevaluated. Subsequently, I speculate on the role of the aTL and of dmPFC. These two regions appear to be most specific to language processing in context, and thus most relevant in light of the comprehension framework outlined in the introduction. In conclusion, I will return to the research questions posed at the outset and evaluate whether the patient study and the neuroimaging data provide hints for the further development of text comprehension theories.

The discussion also draws on recent reviews of neuroscientific studies of text comprehension (Ferstl, in press, b; Ferstl & von Cramon, 2005).

Text comprehension: A right hemisphere function?

The role of the right hemisphere during language processing in context is still a mystery. There is no doubt about a considerable contribution of the RH (Mason & Just, 2004; St. George et al., 1999; Vogeley et al., 2001; Bottini et al., 1994; Caplan & Dapretto, 2001; Xu et al., 2005); but many studies, even those testing complex language comprehension, report left-lateralized or left-dominant activation (Maguire et al., 1999; Fletcher et al., 1995; Baumgärtner, Weiller & Büchel, 2002; Rapp et al., 2004; Crinion et al., 2003). In our own studies, a similar picture emerged. The right aTL was sensitive to inconsistencies in short stories (Chapter 7), but neither in the patient study (Chapter 4) nor in the imgaging experiments on local coherence processing (Chapters 5 and 6), there was evidence for a specific role of the RH. Based on these mixed results, the global and somewhat simplifying claim that higher level text comprehension is a RH function seems premature (cf. Bookheimer, 2002; Mar, 2004).

The reason for these seemingly discrepant results might be in part methodological. It is always difficult to argue based on null results. A failure to find RH involvement might be due to the choice of regions of interest (e.g., Baumgärtner, Weiller & Büchel, 2002; cf. Kircher, et al., 2001), to interindividual lateralization differences, or merely to the lack of experimental power (cf. Ferstl & von Cramon, 2001b). Furthermore, there is more interindividual anatomical variability in the RH compared to the LH. Thus, a whole brain analysis in which activation is averaged across participants in each voxel might not yield an overlapping region of activation, whereas a region of interest analysis, in which larger anatomical regions are combined, is more likely to uncover distributed RH activation (e.g., Mason & Just, 2004; St. George et al., 1999).

A more productive explanation for the divergence of results is that the various experiments elicit qualitatively different processes. For instance, more RH involvement might be found in experiments with affective or social connotation or in experiments using longer texts, in which the integration demands increase with a more extended discourse context (Xu et al., 2005; cf Ferstl & von Cramon, 2001a). Thus, it is necessary to specify the cognitive processes thought to be realized in the right hemisphere, to specify subregions within the RH and to design experiments tailored to distinguishing between competing theories. Until we have a broader empirical data base, the role of the right hemisphere during text comprehension remains elusive.

Higher level comprehension: A function of the lateral pre-frontal cortex?

The lateral prefrontal cortex plays an important role during text comprehension. In line with the clinical descriptions of non-aphasic language deficits after CHI or FL damage, patients with left-frontal lesions had considerable difficulty with the coherence judgment task (Chapter 4), and prefrontal regions were consistently activated in the three imaging studies (Chapters 5-7). To understand the specific contributions, however, it is necessary to subdivide the frontal lobes, and to consider functional attributions for these regions.

What are the specific functions needed for comprehension in context? First, phonological, syntactic and semantic analysis is carried out. These processes are likely to engage the inferior frontal gyrus (Bookheimer, 2002). In particular when semantic selection or semantic integration with the context is required, the triangular part of the IFG is implicated (e.g., Baumgärtner, Weiller & Büchel, 2002; Rapp et al., 2004; Thompson-Schill, d'Esposito & Kan, 1999). In the studies reported here, the IFG (BA 45) was seen in the ELN when sentences were compared to a perceptual baseline, both in the coherence judgment (Chapter 5) and under ToM instructions (Chapter 6), and, particularly for coherent trials under the Logic instructions (Chapter 6). In all of these conditions, lexico-semantic processing is required.

When the information is unexpected or a decision needs to be based on recent memory, IFG activation reaches into the orbital part (BA 47/11; cf. Petrides, Alivisatos & Frey, 2002; Nobre, Coull, Frith & Mesulam, 1999, for studies using non-verbal materials). In line with this interpretation, this region was engaged during the integration of inconsistent information into the recently built situation model (Chapter 7), and this was more pronounced for the memory demanding chronological stories.

Furthermore, executive processes are required for the application of task rules (cf. Derrfuss, Brass & von Cramon, 2004; Brass, Derrfuss, Forstmann & von Cramon, 2005), for inhibition of task irrelevant information (Dove et al., 2000; Zysset, Müller, Lohmann & von Cramon, 2001), or for sequencing (Crozier et al., 1999; Sirigu et al., 1998; cf. Mar, 2004). These processes are likely to engage the posterior lateral prefrontal cortex (BA 9/6/44). Consequently, we interpret the activation close to the junction of the inferior frontal and precentral sulci found in the coherence judgment experiment (Chapter 5) as reflecting the task requirements, rather than an increase in phonological or syntactic processes (cf. Bookheimer, 2002; Saxe, Carey & Kanwisher, 2004).

Finally, language processing, in particular the comprehension of longer texts, always engages attention and working memory. These processes are likely to involve the anterior parts of the middle frontal gyrus (Gruber, 2001; Gruber & von Cramon, 2003). Once more, this region was part of the ELN in the experiments using the coherence judgment task. More specific activation was shown during situation model updating, when chronological information had to be integrated with the context information stated several sentences earlier. Whether this activation was due to the integration demands or to the distance within the discourse is the topic of a study currently carried out (Ferstl, Siebörger & von Cramon, in preparation).

Comprehension in a sentence context: The anterior temporal lobes

Closely linked to the inferiormost frontal lobe is the aTL. The most important proposals for aTL functionality include memory functions, in particular for autobiographical and episodic memory, and for semantic processes, in particular category specific retrieval of proper names or animate entities (e.g., Damasio, Grabowski, Tranel, Hichwa & Damasio, 1996; Leveroni et al., 2000; Maratos, Dolan, Morris, Henson, & Rugg, 2001; Martin & Chao, 2001).

In language comprehension studies, the aTLs are activated in studies both on the text and the sentence level when the integration of incoming words into a semantically based representation is needed (Stowe et al., 1998; Mazoyer et al., 1993). In our own studies (Chapters 5-7), the specific location was at the anteriormost end of the superior temporal sulcus, reaching into the superior and middle temporal gyri, but being slightly more posterior than the temporal pole. The aTLs are also part of the network repeatedly described to be active during ToM tasks (Frith & Frith, 2003; Gallagher & Frith, 2003). Frith & Frith (2003) attribute a process called "narrativization" to this brain region which is related to the episodic memory functions of the aTL.

Somewhat speculative, we postulated that the aTLs serve a function related to the propositionalization process (Kintsch & van Dijk, 1978). Such a view is consistent with the finding of aTL involvement during sentence processing (Bavelier et al., 1997; Müller et al., 1997; Stowe, et al., 1998) and its senstivity not only to semantic content (Baumgärtner, Weiller & Büchel, 2002) but also to differences in syntax or phrasal structure (Stowe, Haverkort & Zwarts, 2005; Friederici, Rüschemeyer, Hahne & Fiebach, 2003, Vandenberghe, Nobre & Price, 2002). Although there is evidence for the language specificity at least of the left aTL (Humphries et al., 2001; cf. Long & Baynes, 2002), this notion does not necessarily hinge on a verbal or abstract representation (cf. Zwaan, 2004). Rather, it might also refer to a chunking process in which phrases are identified and combined into content units.

The activation of the right aTL when inconsistent information was encountered in a story context (Chapter 7) could be explained as reflecting the increased demands on the text base construction. Inconsistent information requires the addition of novel propositions or content units not yet included in the discourse representation. In line with resource theory (cf. Chapter 1), this increased demand might lead to additional RH activation in the aTL homolog. Again, this explanation requires converging evidence and replications using other materials.

Making sense of what happens: The dorso-medial prefrontal cortex

One of the most interesting results of the studies reported here is the importance of the dmPFC (Ferstl & von Cramon, 2001a, 2001b, 2002; Ferstl, Rinck & von Cramon, 2005; Fletcher et al., 1995; Mazoyer et al., 1993; Schmalhofer et al., 2004; Siebörger, Ferstl & von Cramon, 2003; Vogeley et al., 2001). Dinstinct from the slightly more posterior region in BA6/8 which has recently been associated with processing under uncertainty and error related processing (Volz, Schubotz & von Cramon, 2003; Ullsperger & von Cramon, 2001), the dmPFC (BA 8/9/10) seems to be the most likely candidate as an inference region. Besides by our own results, this conclusion is supported by findings from reasoning experiments and studies on evaluative judgments (Goel, Gold, Kapur & Houle, 1997; Zysset, et al., 2002, 2003). However, there are a variety of other proposals for the function of the anterior dmPFC, including reflections of a default state in the absence of stimulation (Gusnard, Akbudak, Shulman, & Raichle, 2001), self-referential processes (Northoff & Bermpohl, 2004), emotion (Greene, Sommerville, Nystrom, Darley & Cohen, 2001), moral judgments (Moll, de Oliveira-Souza, Eslinger et al., 2002; Heekeren, et al., 2003), and, most prominently, ToM processes (Fletcher et al., 1995; Frith & Frith, 2003; Saxe, Carey & Kanwisher, 2004).

More specifically, Gallagher and Frith (2003) postulated that the dmPFC function during ToM processing concerns the decoupling of information from reality. In their view, the overlap of the results of language processing or reasoning experiments with ToM studies is a direct consequence of the inseparability of ToM and communication. When adopting this point of view, it is a challenge to make predictions concerning the circumstances under which the dmPFC should be particularly involved. Why is it more active during the processing of coherent as compared to incoherent trials (Chapter 5)? Based on the results of the study successfully separating inference processes from ToM processes (Chapter 6), we postulated a more general process encompassing both ToM components and inferencing. This account makes use of the observation of patients with lesions in this area to have problems with drive and motivation (Marin, 1991). Thus, we argued, the function of the dmPFC is related to an integration of the inner world with the external stimulation. This function is closely related to the comprehender's self (cf. Northoff & Bermpohl, 2004), and consequently involved in a wide variety of tasks, not only those using self-referential stimuli (cf. Siebörger, Ferstl, & von Cramon, 2003; Siebörger, Ferstl, Volkmann & von Cramon, 2004). What is important in the context of text comprehension studies, however, is that dmPFC activation is observed when the task's time frame allows for hypothesis formation, when there is no right or wrong answer, but a ideosyncratic response criterion needs to be established, and when prior knowledge or contextual information plays a role.

All these criteria were met when the consistency of emotional information had to be evaluated (Chapter 7). In contrast, the chronological information was either right or wrong, and its evaluation solely depended on an accurate memory representation of the discourse context. Thus, we interpreted the dmPFC activation as distinguishing qualitatively different inference processes rather than being due to the emotional content. As pointed out in Chapter 7, though, further studies using different information aspects are needed to separate the inference requirements from the information type.

In two experiments, we observed concurrent activation of the dmPFC and the PCC/prec (Chapters, 5-6). Given the close link to the dmPFC via pericallosal connections, we interpret the pairwise activation of the fronto- and parieto-medial regions to reflect different components of the same domain-independent process (Chapters 5 and 6; Siebörger, Ferstl & von Cramon, 2003). More specifically, the PCC has been interpreted as related to situation model updating (Maguire et al., 1999; Fletcher et al., 1995; Schmalhofer, et al., 2004; Siebörger, et al., in preparation; see also Kuperberg et al., 2000). Based on the comparison of evaluative judgments with those based on episodic memories, it has been argued that the posterior medial activation reflects the knowledge aspect of the inference process (Zysset, Huber, Ferstl, & von Cramon, 2002; Zysset, Huber, Samson, Ferstl & von Cramon, 2003; Fletcher et al., 1995), i.e., the integration of the novel information with prior discourse context or background knowledge. In the experiment specifically designed to study situation model updating (Chapter 7), this brain region did not play a particular role. While dorsal precuneus activation for chronological stories was related to increased attention to the global context, there was no differential activation in the lower precuneus (PCC/prec). Because the scenarios described in the stories remained constant and the type of background knowledge was comparable, the evaluative aspect of the inference process was more prominent than the knowledge aspect.

Implications for text comprehension theories

In addition to the increasing knowledge about the functional neuroanatomy of regions involved during text comprehension, neuroimaging provides experimental data which in turn might influence psycholinguistic theories. Three important issues were posed in the introduction in Chapter 1.

The first question was concerned with the interface between the surface level and the higher level representations. How do linguistic differences propagate to a conceptual or situational understanding? Relevant to this question are, of course, studies on syntactic parsing (see Kaan & Swaab, 2002; Friederici, 2002; Hagoort, Brown & Osterhout, 1999, for reviews; see also Ferstl & Flores-d'Arcais, 1999). The main interest of many of these studies, however, is how a correct reading is obtained, rather than how subtle differences in wording influence the quality or the content of the resulting conceptual representations. In the study reported in Chapter 4, a direct test of the interaction between lexical and pragmatic information was conducted. The difficulty of reconciling a cohesive marker with an implausible pragmatic interpretation led to activation in left prefrontal cortex (Ferstl & von Cramon, 2001a). Similarly, Robertson et al. (2000) showed that connecting sentences with definite articles activated the right prefrontal cortex, compared to a less coherent version using indefinite articles. These findings clearly show that neuroimaging methods are sufficiently sensitive for capturing subtle effects caused by linguistic variation.

The second issue concerned the status of qualitatively different inference types. The evidence seems strong that the fronto-medial cortex plays a role for inference processes (Ferstl & von Cramon, 2001a, 2002; Schmalhofer et al., 2004). The results reported here have been replicated using auditory presentation (Ferstl & von Cramon, 2001b), instruction-based deliberate inferences (Siebörger, Ferstl & von Cramon, 2003), and coherence building using other materials (Siebörger, in preparation). On the other hand, there is evidence for RH regions, in particular the right temporal lobe, to be also related to inferencing (Mason & Just, 2004; St. George et al., 1999; Caplan & Dapretto, 2001). A speculative account consistent with both findings is based on a qualitative distinction between different inference types. In line with the coarse coding hypothesis (Beeman, 1993), those inferences might be realized in the RH which can be drawn by associative activation of distant semantic fields and subsequent integration into a common representation. In contrast, knowledge based or explanation-based inferences requiring the use of background knowledge and discourse information as well as a deliberate plausibility judgment might engage the fronto- and parieto-medial cortices. The findings of dissociable integration processes in our study on situation model building (Chapter 7) are preliminary evidence for such an interpretation, but the materials were not sufficiently controlled to exclude alternative explanations. If a qualitative distinction of the neuroanatomical sequels of elaborative and bridging inferences could be confirmed in a direct within-subjects comparison, neuroimaging would provide a useful tool for separating inference types by analysing their realizations in fronto-medial or right temporal brain regions, resp.

The third issue concerned the representation of situation models. As shown in Chapter 7, qualitative differences in information aspects have a direct reflection in domain-specific brain activation, consistent with the view of an experiental grounding of language (Zwaan, 2004; Glenberg & Kaschak, 2002). Ventro-medial prefrontal activation was observed for emotional information (Chapter 7) and similar results have been obtained for the comprehension of verbal humour (Goel & Dolan, 2001; 2004; Siebörger, Ferstl, Volkmann & von Cramon, 2004). In contrast, fronto-parietal activation was observed during the processing of chronological information (Chapter 7). Thus, neuroimaging facilitates a qualitative dissociation of different situation model aspects (cf. Perfetti, 1999). Behavioral research has identified a multitude of variables influencing situation model building (Zwaan & Radvansky, 1998). Spatial or visual situation models have been studied most extensively, but so far, no neuroimaging data on this information aspect are available (but see Ruff, Knauff, Fangmeier, & Spreer, 2003, for a related reasoning experiment). The results reported here are promising because they prove the sensitivity of the imaging methods for capturing subtle differences caused by language variations.

Conclusions

The clinical diagnosis and treatment of non-aphasic language deficits requires a better understanding of the functional neuroanatomy of text comprehension processes. The research reported here adds to the still small but growing neuroimaging literature on this topic. The promise of neuroimaging methods lies in a more detailed description of the concert of cognitive processes contributing to comprehension in context. Rather than identifying single brain regions specific to subprocesses of comprehension, the task of the future will be a delineation of the relative contributions of components of the extended language network. The thorough understanding of text comprehension processes and the combination of lesion and neuroimaging studies provides the basis for this research endeavor.

PART III:
Bibliography

Bibliography

Adams, R. D., & Victor, M. (1993). *Principles of Neurology*. New York, NY: McGraw-Hill.

Albrecht J. E., & O'Brien, E. J. (1993). Updating a mental model: maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory and Cognitiion, 19*, 1061-1070.

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

American Speech-Language-Hearing Association (1988). The role of the speech-language pathologists in the identification, diagnosis, and treatment of individuals with cognitive-communicative impairments. ASHA, 30, 79.

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspective on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging* (pp. 1-24). New York: Cambridge University Press.

Baron-Cohen, S. (2003). The essential difference: Male and female brains and the truth about autism. NY: Basic books.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition, 21,* 37-46.

Baron-Cohen, S., Tager-Flusberg H. & Cohen, D.J. (Eds.) (2004). *Understanding other minds: perspectives from developmental cognititve neuroscience*. Oxford: Oxford University Press.

Baron-Cohen, S., Tager-Flusberg, H. & Cohen, D. J. (Eds.) (1995). *Understanding other minds: Perspectives from autism.* Oxford: Oxford University Press.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Baumgärtner, A., Weiller, C., & Büchel, C. (2002). Event-related fMRI reveals cortical sites involved in contextual sentence integration. *NeuroImage, 16*, 736-745.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V. P., Karni, A., Printer, A. , Braun, A., Lalwani, A., Rauschecker, J. P., Turner, R. & Neville, H. (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 and C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 255-284). Mahwah, NJ: Lawrence Erlbaum.

Benowitz, L. I., Moya, K. L, & Levine, D. N. (1990). Impaired verbal reasoning and constructional apraxia in subjects with right hemisphere damage. *Neuropsychologia, 28,* 231-241.

Benson, D. F., & Ardila, A. (1996). *Aphasia: A Clinical Perspective*. Oxford: Oxford University Press.

Beyer, R., Guthke, T., & Pekrul, U. (1996). Repräsentation von Textwissen im menschlichen Gedächtnis [Representation of text knowledge in human memory]. *Zeitschrift für Psychologie, 204*, 199-232.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M. & Cox, R.W. (1999). Conceptual processing during the conscious resting state: a functional MRI study. *Journal of Cognitive Neuroscience, 11*, 80-93.

Binder, J. R., Frost, J. A., Hammeke, T.A., Cox, R. W., Rao, S. M. & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience, 171*, 353-362.

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing*. Annual Review of Neuroscience, 25*, 151-188.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S. J. & Frith, C.D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: a positron emission tomography activation study. *Brain, 117,* 1241-1253.

Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science, 247*, 44-48.

Bransford, J. D., & Franks, J. J. (1972). The abstraction of linguistic ideas: A review. *Cognition, 1*, 211-249.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.

Brass, M., Derrfuss, J., Forstmann, B. & von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends in Cognitive Sciences, 9*, 314-316.

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *NeuroImage, 14*, 1416-1423.

Broca, P. (1865). Sur le siège de la faculté du langage articulé [On the localization of the articulatory language faculty]. *Bulletin d'Anthropologie, 6*, 377-393.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. [Comparative localization of the cortex based on its zytoarchitecture] Leipzig: J. A. Barth.

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.

Brookshire, R. H., & Nicholas, L. E. (1993). *The Discourse Comprehension Test.* Tucson, AZ: Communication Skill Builders, A Division of the Psychological Corporation.

Brothers, L. (1990). The social brain: A project for integrating primate behaviour and neurophysiology in a new domain. *Concepts in Neuroscience, 1*, 27-51.

Brown, C. M., & Hagoort, P. (Eds.) (1999). *The neurocognition of language*. Oxford: Oxford University Press.

Brownell, H. H., Gardner, H., Prather, P., & Martino, G. (1995). Language, communication and the right hemisphere. In: Kirshner H.S. (Ed.), *Handbook of neurological speech and language disorders* (pp. 325-349), New York: Marcel 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 and C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 309-328). Mahwah, NJ: Lawrence Erlbaum.

Brownell, H. H., Potter, H. H., Bihrle, A. M. & Gardner, H., (1986). Inference deficits in right brain-damaged patients. *Brain and Language, 27,* 310-321.

Brunet, E., Sarfati, Y., Hardy-Baylé, M.C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage, 11*, 157-166.

Burke, D. M. (1997). Language, aging, and inhibitory deficits: Evaluation of a theory. *Journal of Gerontology: Psychological Sciences, 52B,* P254-P264.

Cabeza, R., & Kingstone, A. (Eds.) (2001). *Handbook of functional neuroimaging of cognition*. Cambridge, MA: MIT Press.

Caplan, D. (1987). Neurolinguistics and linguistic aphasiology: An introduction. Cambridge: Cambridge University Press.

Caplan, D. (1992). Language: structure, processing, and disorders. Cambridge, MA: MIT Press.

Caplan, R. & Dapretto, M. (2001). Making sense during conversation: an fMRI study. *NeuroReport, 12*, 3625-3632.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12,* 314-325.

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

Chapman, S. B., Levin, H. S., & Culhane, K. A. (1995). Language impairment in closed head injury. In H. S. Kirshner (Ed.), *Handbook of neurological speech and language disorders* (pp. 387-414). New York: Marcel Dekker.

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

Chiarello, C., & Beeman, M. (1998). On codes of meaning and the meaning of codes: Semantic access and retrieval within and between hemispheres. In: M. Beeman and C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 141-160). Mahwah, NJ: Lawrence Erlbaum.

Corbalis, M. C. (1997). Mental rotation and the right hemisphere. *Brain and Language, 57*, 100-121.

Crinion, J.T., Lambon-Ralph, A., Warburton, E.A., Howard, D. & Wise, R.J.S. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain, 126,* 1193-1201.

Crozier, S., Sirigu, A., Lehéricy, S., Moortele, P.F. van de, Pillon, B., Grafman, J., Agid, Y., Dubois, B. & LeBihan, D. (1999). Distinct prefrontal activations in processing sequence at the sentence and script level: an fMRI study. *Neuropsychologia, 37,* 1469-1476.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature, 380*, 499-505.

Davidson, R. J, & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences, 3*, 11-21.

Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia, 41,* 127-138.

Dehaene, S., et al., (1999). Sources of mathematical thinking: behavioral and brain imaging evidence. *Science, 284*, 970-974.

Dehaene, S., Dupoux, E., Mehler, J., Cohen L., Paulesu, E., Perani, D., van de Moortele, P.-F., Lehericy, S. & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport, 8*, 3809-3815.

Delis, D. C., Kramer, J. H., Kaplan, E., & Obler, B. A. (1987). *The California verbal learning test: Adult version*. San Antonio, TX: The Psychological Corporation.

Derrfuss, J., Brass, M., & von Cramon, D. Y. (2004). Cognitive control in the posterior frontolateral cortex: Evidence from common activations in task coordination, interference control, and working memory. *NeuroImage, 23,* 604-612.

Doane, S. M., McNamara, D. S., Kintsch, W., Polson, P.G. & Clawson, D. (1992). Prompt comprehension in UNIX command production. *Memory & Cognition, 20*, 327-343.

Dolan, R. J. (2002). Emotion, cognition and behavior. *Science, 298*, 1191-1194.

Dooling, D. J., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology, 88*, 216-222.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C.J. & von Cramon, D.Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Cognitive Brain Research, 9*, 103-109.

Ettlin, T. M., 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. (2001). Learning from Text. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*, Volume 3.13: Cognitive Psychology and Cognitive Science (Ed. W. Kintsch). Amsterdam: Elsevier.

Ferstl, E. C. (in press). The functional neuroanatomy of text comprehension: What's the story so far? In F. Schmalhofer & C. Perfetti (Hrsg.), *Higher Level Language Processes in the Brain: Inference and Comprehension Processes*. Mahwah: NJ: Lawrence Erlbaum.

Ferstl, E. C. (2006). Text comprehension in middle aged adults: Is there anything wrong? *Aging, Neuropschology and Cognition, 13,* 62-85.

Ferstl, E. C., & Flores d'Arcais, G. B. (1999). The reading of words and sentences. In A. D. Friederici (Hrsg.), *Language comprehension: A biological perspective*, 2nd edition (pp. 175-210). Berlin: Springer.

Ferstl, E. C., & Guthke, T. (1998). Diskursanalyse als Hilfsmittel zur klinischen Evaluation von nicht-aphasischen Sprachstörungen [Discourse analysis as a tool for the clinical evaluation of non-aphasic language deficits]. In Ohlendorf, I. M., Widdig, W., & Malin, J.-P. (Hrsg.), Bonn-Bochumer Beiträge zur Neuropsychologie und Neurolinguistik BBB. Band 5. *Arbeiten mit Texten in der Aphasietherapie* (pp. 113-143). Freiburg: Hochschul-Verlag.

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., Guthke, T., & von Cramon, D. Y. (2002). Text comprehension after brain injury: Left prefrontal lesions affect inference processes. *Neuropsychology*, *16*, 292-308.

Ferstl, E. C., & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In H. Oostendorp &. S. Goldman (Hrsg.), *The construction of mental models during reading*, (pp. 247 - 277). Mahwah, NJ: Lawrence Erlbaum.

Ferstl, E. C., Rinck, M., & von Cramon, D. Y. (2005). Emotional and temporal aspects of situation model processing during text comprehension: An event-related fMRI study. *Journal of Cognitive Neuroscience*, *17*, 724-739.

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

Ferstl, E. C., & von Cramon, D. Y. (2001b). Inference processes during text comprehension: Is it the left hemisphere after all? *Journal of Cognitive Neuroscience (Supplement), 128.*

Ferstl, E. C., & von Cramon, D. Y. (2002). What does the fronto-medial cortex contribute to language processing: Coherence or Theory of Mind? *NeuroImage, 17,* 1599-1612.

Ferstl, E. C., & von Cramon, D. Y. (2005). Sprachverstehen im Kontext: Bildgebende Studien zu Kohärenzbildung und Pragmatik [Language comprehension in context: neuroimaging studies on coherence building and pragmatics]. *Sprache – Stimme – Gehör*, *29*, 130-138.

Ferstl, E. C., Walther, K., Guthke, T., & Cramon, D. Y. (2005). Assessment of story comprehension deficits after brain damage. *Journal of Clinical and Experimental Neuropsychology, 27*, 367-384.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain, 120*, 1779-1791.

Fletcher, P. C., Happé, 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.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*, 78-84.

Friederici, A. D., Opitz, B., & von Cramon, D. Y. (2000). Segregating semantic and syntactic apsects of processing in the human brain: an fMRI investigation of different word types. *Cerebral Cortex, 10*, 698-705.

Friederici, A. D., Rüschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex, 13,* 170-177.

Friston, K. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2,* 189-210.

Frith, C. D., & Frith, U. (1999). Interacting minds – a biological basis. *Science, 286*, 1692-1695

Frith, U. (1989). *Autism. Explaining the enigma*. Blackwell, Oxford, GB.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, 358*, 459-473.

Fuster, J. M. (1997). The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe. Philadelphia, PA: Lippinkott-Raven.

Fuster, J. M. (1999). Cognitive functions of the frontal lobes. In B. L. Miller, & J. L. Cummings (Eds.), *The human frontal lobes: functions and disorders* (pp. 187-195). New York: Guilford.

Gallagher, H. L., & Frith, C.D. (2003). Functional imaging of `theory of mind`. *Trends in Cognititve Sciences, 7*, 77-83.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia, 38,* 11-21.

Germani, M. J., & Pierce, R. S. (1992). Contextual influences in reading comprehension in aphasia. *Brain and Language, 42*,308-319.

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

Gernsbacher, M. A. (Ed.) (1994). *Handbook of psycholinguistics*. San Diego, CA: Academic Press.

Gernsbacher, M. A., & Kaschak, M. P. (2003) Neuroimaging studies of language production and comprehension. *Annual Review of Psychology, 54*, 16.1-16.24.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent characters' emotional states? *Cognition & Emotion, 6*, 89-111.

Geschwind, N. (1979). Specializations of the human brain. *Scientific American, 241*,180-199.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558-565.

Glindemann, R. & von Cramon, D.Y. (1995). Kommunikationsstörungen bei Patienten mit Frontalhirnläsionen [Communciation deficits in patients with frontal lobe lesions]. *Sprache, Stimme, Gehör,* 19, 1-7.

Goel, V. & Dolan, R.J. (2001). The functional anatomy of humor: segregating cognitive and affective components. *Nature Neuroscience, 4*, 237-261.

Goel, V., & Dolan, R. J. (2004). Reciprocal neural response wihtin lateral and ventral medial prefrontal cortex during hot and cold reasoning. *NeuroImage, 20*, 2314-2321.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport, 8,* 1305-1310.

Goel, V., Grafman, J., Sadato, N., & Hallett M.(1995). Modeling other minds. *Neuroreport, 6*, 1741-1746.

Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (Eds.) (2003). *Handbook of discourse processes*. Mahwah, NJ: Lawrence Erlbaum.

Graesser, A. C., Millis K. K., & Zwaan R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163-189.

Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105- 2108.

Gruber, O. (2001). Effects of domain-specific interference on brain activation associated with verbal working memory task performance. *Cerebral Cortex, 2001*, 1047-1055.

Gruber, O., & von Cramon, D. Y. (2003). The functional neuroanatomy of human working memory revisited: Evidence from 3T fMRI studies using classical domain-specific interference tasks. *Neuroimage, 19*, 797-809.

Guéraud, S. & O'Brien, E. (2005). Components of comprehension: A convergence between memory-based processes and explanation-based processes. Special Issue: Discourse Processes, Mahwah, NJ: Lawrence Erlbaum.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1992). An electrophysiological study of semantic processing in young and middle aged academics. *Psychophysiology, 29,* 38-54.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1998). Priming and aging: An electrophysiological investigation of N400 and recall. *Brain & Language, 65,* 333-355.

Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of the Sciences, 98,* 4259-4264.

Guthke, T., Hauptmann, A., & Ferstl, E. C. (2001). Empirische Untersuchung zur Textverarbeitung bei hirngeschädigten Patienten [Empirical studies on text comprehension after brain damage]. *Aphasie und verwandte Gebiete, 15*, 29-41.

Gyselinck, V., & Tardieu, H. (1999). The role of illustrations in text comprehension: what, when , for whom , and why?. In H. Oostendorp &. S. Goldman (Hrsg.), *The construction of mental models during reading*, (pp. 195-218). Mahwah, NJ: Lawrence Erlbaum.

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

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. Psychology and Aging, 7, 56-64.

Happé, F. (1994). An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistics, mentally handicapped and normal children and adults. *Journal of Autism and Developmental Disorders, 24*, 129-154.

Happé, F. (2000). Parts and wholes, meaning and minds: Central coherence and its relation to theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 203-221). Oxford: Oxford University Press.

Hartley, L. L., & Jensen, P. J. (1992). Three discourse profiles of closed-head-injury speakers: theoretical and clinical implications. *Brain Injury, 6*, 271-282.

Hartley, L. L., & Levin, H. S. (1990). Linguistic deficits after closed head injury: A current appraisal. *Aphasiology, 4*, 353-370.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension and aging: A review and a new view. In G. H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 22, pp. 193-225). San Diego, CA: Academic.

Haut, M. W., Petros, T. V., & Frank, R. G. (1990). The recall of prose as a function of importance following closed head injury. *Brain Injury, 4,* 281-288.

Heekeren, H. R., Wartenburger, I., Schmidt, H., Schwintowski, H.-P., & Villringer, A. (2003). An fMRI study of simple ethical decision-making. *NeuroReport, 14,* 1215-1219.

Hess, T. M. (1995). Aging and the impact of causal connections on text comprehension and memory. *Aging and Cognition, 2*, 216-230.

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes*, Vol. 8 (pp. 237-278). New York: Plenum.

Horn, W. (1983). *Leistungsprüfsystem (LPS)* [Performance assessment system]. Göttingen: Verlag für Psychologie.

Hough, M. S. (1990). Narrative comprehension in adults with right and left hemisphere brain-damage: Theme organization. *Brain and Language*, 38, 253-277.

Humphries, C., Willard, K., Buchsbaum, B. & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *Neuroreport, 12*, 1749-1752.

Jentzsch, R.T., Ferstl, E.C., & Guthke, T. (2005). Situationsmodellbildung beim Textverstehen: Eine behaviorale Patientenstudie nach frontaler Hirnschädigung [Situation model building in text comprehension: A behavorial study of patients with frontal lobe lesions]. In K. W. Lange, K.-H. Bäuml, M. W. Greenlee, M. Hammerl, & A. Zimmer (Hrsg.), *Experimentelle Psychologie: Beiträge zur 47. Tagung experimentell arbeitender Psychologen*. Lengerich: Pabst Science Publishers, 94.

Jezzard, P., Matthews, P. M., & Smith, S. M. (Eds.) (2001). *Functional magnetic resonance imaging: an introduction to methods.* Oxford, GB: Oxford Unviersity Press.

Jolliffe, T. & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome: is local coherence impaired? *Cognition, 71,* 149-185.

Jolliffe, T. & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism of Asperger's syndrome. Is global coherence impaired? *Psychological Medicine, 30,* 1169-1187.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences, 6*, 350-356.

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 neurolinguistic perspective. In E. Perecman (Ed.), *The frontal lobes revisited*, (pp. 225-240). Hillsdale, NJ: Lawrence Erlbaum.

Keenan, J. M, Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior, 23*, 115-126.

Kent, R. D. (Ed.) (2004). *The MIT encyclopedia of communication disorders*. Cambridge, MA: MIT Press.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95,* 163-182.

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

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review, 7*, 257-266.

Kintsch, W., & van Dijk, T. A. (1978) Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Kircher, T. T. J., Brammer, M., Andreu, N. T., Williams, S. C. R., & McGuire, P. K. (2001). Engagement of right temporal cortex during processing of linguistic context. *Neuropsychologia, 39*, 798-809.

Kliegl, R., Mayr, U., Junker, M. & Fanselow, G. (1999). Testing age invariance in language processes. In S. Kemper & R. Kliegl, *Constraints on language: Aging, grammar, and memory* (pp. 137-168). Norwell, MA: Kluwer Academic.

Koutstaal, W., Schacter, D. L., Galluccio, L. & Stofer, K. A. (1999). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging, 14*, 220-237.

Krause, B. J., Schmidt, D., Mottaghy, F. M., Taylor, J., Halsband, U., Herzog, H., Tellmann, L., & Müller-Gärtner, H.-W. (1999). Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates: A PET-study. *Brain, 122*, 255-263.

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence form neuroimaging and neuropsychology. *Progress in Neurobiology, 72*, 341-372.

Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., Lythgoe, D. J., Williams, S. C. R., & David, A. S. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: An fMRI study. *Journal of Cognitive Neuroscience, 12*, 321-341.

Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

Larkins, B. M., Worrall, L. E., Hickson, L. M. H. (2000). Functional communication in cognitive communication disorders following traumatic brain injury. In L. E. Worrall, & C. M. Frattali (Eds.), *Neurogenic communication disorders: a functional approach* (pp. 206-219). Thieme.

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

Lehrl, S., Merz, J., Burkhard, G., & Fischer, B. (1991). Mehrfachwahl-Wortschatz-Intelligenztest (MWT-A). [Multiple choice vocabulary intelligence test]. Erlangen, D: Perimed.

Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *Journal of Neuroscience, 20*, 878-886.

Lichtheim, L. (1885). On aphasia. *Brain, 7*, 433-484.

Lohmann, G., Müller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L, & von Cramon, D. Y. (2001). Lipsia: A new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics, 25,* 449-457.

Long, D. L., & Baynes, K. (2002). Discourse representation in the two cerebral hemispheres. *Journal of Cognitive Neuroscience, 14,* 228-242.

Luan Phan, K., Wager, T., Taylor, S. F. & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage, 16,* 331 –348.

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.

Mannes, S., M., & Kintsch, W. (1991). Planning routine computing tasks: Understanding what to do. *Cognitive Science, 15,* 305-342.

Mar, R. A., (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neuropsychologia, 42*, 1414-1434.

Maratos, E. J., Dolan, R. J, Morris, J. S., Henson R. N A., & Rugg, M. D. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia, 29*, 910-920.

Marin, R. W. (1991). Apathy: A neuropsychiatric syndrome. *Journal of Neuropsychiatry and Clinical Neuroscience, 3*, 243-254.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology, 11*, 194-201.

Martin, R. C. (2003). Language processing: functional organization and neuro-anatomical basis. *Annual Review of Psychology, 54,* 15.1-15.35.

Mason, R. A., & Just, M. A. (2004). How the brain processes causal inferences in text: A theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychological Science, 15*, 1-7.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience, 5*, 467-479.

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

McDonald, S. (2000). Exploring the cognitive basis of right-hemisphere pragmatic language disorders. *Brain and Language, 75*, 82-107.

McDonald, S., Togher, L., & Code, C. (Eds.) (1999). *Communication disorders following traumatic brain injury*. Hove, GB: Psychology Press.

McDonald, S., & Wales, R.(1986). An investigation of the ability to process inferences in language following right hemisphere brain damage. *Brain and Language, 29*, 68-80.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*, 440-466.

Metzinger, T. (2000). The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. In T. Metzinger (Ed.), *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 285 – 306). Cambridge, MA: MIT Press.

Meyer, M., Friederici, A. D. & von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: Event related fMRI reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research, 9*, 19-33.

Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourao-Miranda, J., Andreiuolo, P. A. & Pessoa, L. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience, 22*, 2730-2736.

Monetta, L., & Joanette, Y. (2003). Specificity of the right hemsiphere's contribution to verbal communication: The cognitive resources hypothesis. *Journal of Medical Speech-Language Pathology, 11,* 203-211.

Müller, R.-A., Rothermel, R. D., Behen, M. E., Muzick, O., Mangner, T. J., & Chugani, H. T. (1997). Receptive and expressive language activations for sentences: A PET study. *Neuroreport, 8*, 3767-3770.

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

Murray, L. L. (1999). Attention and aphasia: Theory, research and clinical implications. *Aphasiology, 13*, 91-111.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*, 453-465.

Navon, D. (1984). Resources: A theoretical soup stone? *Psychological Review, 91*, 206-234.

Nichelli, P., Grafman, J., Pietrini, P., Clark, K., Lee, K. Y., Miletich, R. (1995). Where the brain appreciates the moral of a story. *NeuroReport, 6*, 2309-2313.

Nicholas, L. E., & Brookshire, R. H. (1995). Comprehension of spoken narrative discourse by adults with aphasia, right-hemisphere brain damage, or traumatic brain injury. *American Journal of Speech-Language Pathology, 4*, 69-81.

Nobre, A. C., Coull, J. T., Frith, C.D. & Mesulam, M. M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nature Neuroscience, 2*, 11-12.

Norris, D. G., Zysset, S., Mildner, T. & Wiggins, C. J. (2002). An investigation of the value of spin-echo based fMRI using a stroop colour-word matching task and EPI at 3 Tesla. *NeuroImage, 15*, 719-726.

Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences, 8*, 102-107.

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

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics, 12*, 433-452.

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in text: What readers believe versus what they read. *Psychological Science, 3,* 229-235.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production: A contribution to the definition of language epicenters with PET. *NeuroImage, 11*, 347-357.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, F., Dupoux, E., Fazio, F. & Mehler, J. (1996). Brain processing of native and foreign languages. *Neuroreport, 7,* 2439-2444.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167-208). Oxford: Oxford University Press.

Petrides, M., Alivisatos, B., & Frey, S. (2002). Differential activation of the human orbital, mid-ventrolateral, and mid-dorsolateral PFC during the processing of visual stimuli. *Proceedings of the National Academy of the Sciences, 99*, 5649-5654.

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

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 516-526.

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* (p. 18-28). Baltimore, MD: John Hopkins University Press.

Radvansky, G. A., Gerard, L. D., Zacks, R. T., & Hasher, L. (1990). Younger and older adults' use of mental models as representations for text materials. *Psychology and Aging, 2,* 209-214.

Radvansky, G. A., Zwaan, R. A., Curiel, J. M. & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging, 16*, 145-160.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, 98*, 676-682.

Rapp, A. M., Leube, D. T., Erb, M., Grodd, W., & Kircher, T. T. J. (2004). Neural correlates of metaphor processing. *Cognitive Brain Research, 20*, 395-402.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1-91). Mahwah, NJ: Lawrence Erlbaum.

Rinck, M. (2000). Situationsmodelle und das Verstehen von Erzähltexten: Befunde und Probleme [Situation models and the understanding of narrative texts: Findings and problems.] *Psychologische Rundschau, 51*, 115-122.

Rinck, M., Gámez, E., Díaz, J. M., & de Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory & Cognition, 31,* 77-86.

Rinck, M., Hähnel, A., & Becker, G. (2001). Using temporal information to construct, update, and retrieve situation models of narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 67-80.

Rinck, M., & Weber, U. (2003). Who, when, where: An experimental test of the event-indexing model. *Memory & Cognition, 31,*:1284-1292.

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.

Ruff, C.C., Knauff, M., Fangmeier, T. & Spreer, J. (2003). Reasoning and working memory: common and distinct neuronal processes. *Neuropsychologia, 41,* 1241-1253.

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 55,* 87-124.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals and understanding: An inquiry into human knowledge structures. Hillsdale, NJ: Lawrence Erlbaum.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes, 33*, 105-132.

Schmalhofer, F., & Perfetti, C. A. (Eds.) (in press). *Higher Level Language Processes in the Brain: Inference and Comprehension Processes*. Mahwah: NJ: Lawrence Erlbaum.

Schmalhofer, F., Raabe, M., Friese, U., Pietruska, K., & Rutschmann, R. (2004). Evidence from an fMRI experiment for the minimal encoding and subsequent substantiation of predictive inferences. Proceedings of the Meeting of the Cognitive Science Society.

Schwartz, A., Kischka, U. & Rihs, F. (1997). Funktionelle bildgebende Verfahren. In U. Kischka, C.-W. Wallesch, & G. Wolf (Eds.), *Methoden der Hirnforschung* (p. 295-318). Heidelberg: Spektrum Akademischer Verlag.

Schwarz, M. (2000). *Indirekte Anaphern in Texten* [Indirect anaphors in texts]. Tübingen: Niemeyer.

Scott, S.K., Leff, A. P., Wise, R. J. S. (2003). Going beyond the information given: a neural system supporting semantic interpretation. *NeuroImage, 19*, 870-876.

Siebörger, F.T., Ferstl, E.C., Volkmann, B. & von Cramon, D.Y. (2004). Spass beiseite! Eine fMRI-Studie und eine behaviorale Patientenstudie zu verbalem Humor und sprachlicher Revision. [All joking aside! An fMRI study and a behavioral patient study on verbal humor and linguistic revision] *Zeitschrift für Neuropsychologie, 15*.

Siebörger, F. T., Ferstl, E. C., & von Cramon, D. Y. (2003). Making sense of nonsense: An fMRI study of task induced coherence processes. *Journal of the International Neuropsychological Society, 9*, 506.

Simpson, J. R., Drevets, W. C., Snyder, A. Z., Gusnard, D. A., & Raichle, M. E. (2001a). Emotion-induced changes in human medial prefrontal cortex: II. During anticipatory anxiety. *Proceedings of the National Academy of the Sciences, 98*, 688-693.

Simpson, J. R., Snyder, A. Z., Gusnard, D. A., & Raichle, M. E. (2001b). Emotion-induced changes in human medial prefrontal cortex: I. During cognitive task performance. *Proceedings of the National Academy of the Sciences, 98*, 683-687.

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

Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory & Language, 33*, 421-441.

Sirigu, A., Cohen, L., Zalla, T., Pradat-Diehl, P., van Eeckhout, P., Grafman, J. & Agid, Y. (1998). Distinct frontal regions for processing sentence syntax and story grammar. *Cortex, 34,* 771-778.

Sperber, D. & Wilson, D. (1995). *Relevance: communication and cognition*, 2nd edition. Oxford: Blackwell.

Springer, S. P., & Deutsch, G. (1997). *Left brain, right brain*. New York: Freeman.

Stemmer, B., & Whitaker, H. A. (1998). *Handbook of neurolinguistics.* San Diego, CA: Academic Press.

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.

St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience, 6*, 70-83.

Stine-Morrow, E. A. L., & Miller, L. M. S. (1999). Discourse processing and aging. Resource allocation as a limiting factor. In S. Kemper & R. Kliegl (Ed.), *Constraints on language: Aging, grammar, and memory* (pp. 53-76). Norwell, MA: Kluwer Academic.

Stowe, L. A., Broere, C. A. J., Paans, A. M. J., Wijers, A. A., Mulder, G., Vaalburg, W., & Zwarts, F. (1998). Localizing components of a complex task – Sentence processing and working memory. *Neuroreport, 9,* 2995-2999.

Stowe, L. A., Haverkort, M., & Zwarts, F. (2005). Rethinking the neurological basis of language. *Lingua, 115*, 997-1042.

Thompson-Schill, S.L., D'Esposito, M. & Kan, I.P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron, 23,* 513-522.

Tzourio, N., Crivello, F., Mellet, E., Nkanga-Ngila, B. & Mazoyer, B. (1998). Functional anatomy of dominance for speech comprehension in left handers vs. right handers. *Neuroimaging, 8,* 1-16.

Ulatowska, H. K., & Chapman, S. B. (1994). Macrostructure in aphasia. In R. L. Bloom, L. K. Obler, S. De Santi, & J. S. Ehrlich (Eds.), *Discourse analysis and applications: Studies in adult clinical populations* (pp. 29-46). Hillsdale, NJ: Lawrence Erlbaum.

Ulatowska, H. K., Hayashi, M. M., Cannito, M. P., & Fleming, S. G. (1986). Disruption of reference in aging. *Brain & Language, 28*, 24-41.

Ullsperger, M. & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociatoin of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage, 14*, 1387-1401.

Valencia-Laver, D. L. & Light, L. L. (2000). The occurrence of causal bridging and predictive inferences in young and older adults. *Discourse Processes, 30*, 27-56.

van den Broek, P. W. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 536- 539-588. San Diego, CA: Academic Press.

van den Broek, P. W. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores- d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423-445). Hillsdale, NJ: Erlbaum.

van der Linden, M., Hupet, M., Feyereisen, P., Schlestraete, M.-A., Bestgen, Y., Bruyer, R., Lories, G., El Ahmadi, A., & Seron, X. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology and Cognition, 6*, 32-55.

van Dijk, T. A. & Kintsch, W. (1983) *Strategies of discourse comprehension.* New York: Academic Press.

Vandenberghe, R., Nobre, A. C. & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*, 550-560.

Vogeley, K., Bussfeld, P. Newen A., Herrmann S., Happé, F., Falkai, P., Maier, W., Shah, N.J., Fink, G.R., & Zilles, K. (2001). Mind reading: Neural mechanisms of Theory of Mind and self-perspective. *NeuroImage*, *14*, 170-181.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2003). Predicting events of varying probability: uncertainty investigated by fMRI. *NeuroImage, 19*, 271-280.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*, 483-488.

Wechsler, D. (1987). *Wechsler memory scale-revised.* New York: Psychological Corporation.

Weissman, D. & Banich, M. T. (2000). The cerebral hemispheres cooperate to perform complex but not simple tasks. *Neuropsychology, 14*, 41-59.

Wellman, H. M. (1993). Early understanding of mind: The normal case. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 10-39). Oxford: Oxford University Press.

Werheid, K., Hoppe, C., Thöne, A., Müller, U., Müngersdorf, M., & von Cramon, D. Y. (2002). The adaptive digit ordering test: Clinical application, reliability, and validity of a verbal working memory test. *Archives of Clinical Neuropsychology, 17*, 547-565.

Wernicke, D. (1874). The aphasic symptom complex: a psychological study on a neurological basis. Breslau: Kohn & Weigert.

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

Wingfield, A. & Stine-Morrow, E. A. L. (2000). Language and speech. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 359-416). Mahwah, NJ: Lawrence Erlbaum.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. *NeuroImage, 25*, 1002-1015.

Zacks, R. T. & Hasher, L. (1988). Capacity theory and the processing of inferences. In L. L. Light & D. M. Burke (Eds.), *Language, memory and aging* (pp. 154-170). New York: Cambridge University Press.

Zacks, R. T. & Hasher, L. & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 293-358). Mahwah, NJ: Lawrence Erlbaum.

Zalla, T., Phipps, M., & Grafman, J. (2002). Story processing in patients with damage to the prefrontal cortex. *Cortex, 38*, 215-231.

Zimmermann, P., & Fimm, B. (1992). Testbatterie zur Aufmerksamkeitsprüfung (TAP) [Test battery for the assessment of attention]. Würselen, Germany: Psychologische Testsysteme.emory. *Psychological Bulletin, 123,* 162-185.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation*, Vol. 44, (pp. 35-62). Amsterdam: Elsevier.

Zwaan, R. A., Langston, M. C. & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*, 292-297.

Zwaan, R. A., Magliano, J. P. & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 386-397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.

Zysset, S., Huber, O., Ferstl, E. C. & von Cramon, D. Y. (2002). The anterior fronto-medial cortex and evaluative judgment: An fMRI study. *NeuroImage, 15*, 983-991.

Zysset, S., Huber, O., Samson, A., Ferstl, E. C. & von Cramon, D. Y. (2003). Functional specialization wihtin the anterior medial prefrontal cortex: a functional magnetic resonance imaging study with human subjects. *Neuroscience Letters, 335*, 183-186.

Zysset, S., Müller K., Lohmann, G., & von Cramon D. Y. (2001). Color-word matching stroop task: Separating interference and response conflict. *Neuroimage,13,* 29-36.

PART IV:
Appendix

List of Tables and Figures
Verzeichnis der Tabellen und Abbildungen

Tables

Figures

Cumulative Publications
Kumulierte Publikationen

1. 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.

2. Ferstl, E. C. (2006).
   Text comprehension in middle aged adults: Is there anything wrong?
   Aging, Neuropschology and Cognition, 13, 62-85.

3. Ferstl, E. C., Guthke, T., & von Cramon, D. Y. (2002).
   Text comprehension after brain injury: Left prefrontal lesions affect inference processes.
   Neuropsychology, *16*, 292-308.

4. Ferstl, E. C., & von Cramon, D. Y. (2002).
   What does the fronto-median cortex contribute to language processing: Coherence or Theory of Mind?
   NeuroImage, 17, 1599-1612.

5. Ferstl, E. C., Rinck, M., & von Cramon, D. Y. (2005).
   Emotional and temporal aspects of situation model processing during text comprehension: An event-related fMRI study.
   Journal of Cognitive Neuroscience, 17, 724-739.

ELSEVIER

Cognitive Brain Research 11 (2001) 325–340

COGNITIVE
BRAIN
RESEARCH

www.elsevier.com/locate/bres

Research report

The role of coherence and cohesion in text comprehension: an event-related fMRI study

Evelyn C. Ferstl*, D. Yves von Cramon

Max-Planck-Institute of Cognitive Neuroscience, Stephanstrasse 1a, D-04103 Leipzig, Germany

Accepted 26 December 2000

Abstract

Text processing requires inferences for establishing coherence between successive sentences. In neuropsychological studies and brain imaging studies, these coherence-building processes have been ascribed to the right hemisphere. On the other hand, there is evidence for prefrontal brain damage causing non-aphasic language disorders, in which text level processes are impaired. In this study, we used an event-related, whole-head fMRI methodology to evaluate the contributions of prefrontal areas and the right hemisphere to coherence building. We scanned 12 participants while they read 120 sentence pairs and judged their coherence. Four conditions were used, resulting from crossing coherence and cohesion (i.e. the presence of a lexical connection). A behavioral pretest confirmed that cohesion aided establishing coherence, whereas it hindered the detection of coherence breaks. In the fMRI study, all language conditions yielded activation in left frontolateral and temporolateral regions, when compared to a physical control task. The differences due to coherence of the sentence pairs were most evident in larger activation for coherent as compared to incoherent sentence pairs in the left frontomedian wall, but also in posterior cingulate and precuneal regions. Finally, a left inferior prefrontal area was sensitive to the difficulty of the task, and in particular to the increase in processing costs when cohesion falsely indicated coherence. These results could not provide evidence for a special involvement of the right hemisphere during inferencing. Rather, they suggest that the left frontomedian cortex plays an important role in coherence building.

Theme: Neural basis of behavior

Topic: Cognition

Keywords: Text comprehension; Event-related functional magnetic resonance imaging; Frontal lobe

1. Introduction

An important process during natural language understanding is the establishment of coherence. In order to correctly interpret the current utterance, it is necessary for the comprehender to link its meaning to the prior context by using general knowledge of the world. In many cases, the specific relationship between the input just encountered and the communicative context is left implicit. However, the comprehender's assumption is that the speaker or writer follows an intention, and thus, that the utterance is relevant to the prior context.

The process needed for establishing coherence is called *inferencing*. In psycholinguistic research, a taxonomy of inferences has been attempted. The question of interest is which types of inferences are made automatically and on-line (for reviews see Refs. [21,47]). For instance, it has been proposed that only bridging inferences required to establish local coherence between subsequent sentences, and those inferences based on easily accessible knowledge are made automatically [37]. According to this so-called minimalist framework, all other types of inferences, such as causal, predictive, or elaborative inferences, are dependent on strategic processes determined by the comprehender's goals. In contrast, Trabasso and van den Broek [52] argue that causal and goal-related inferences are always included during comprehension, and even more, that finding out *why* certain things happen and *why* certain actions are carried out is one of the most fundamental functions of comprehension.

---

*Corresponding author. Tel.: +49-341-994-0116; fax: +49-341-994-0221.

E-mail address: ferstl@cns.mpg.de (E.C. Ferstl).

0926-6410/01/$ – see front matter
PII: S0926-6410(01)00007-6

In neuropsychological research, inference processes have been studied in the context of so-called non-aphasic language deficits [44]. While in right handers there is abundant evidence for a left-hemispheric dominance of language processes on the word and sentence level, the prevalent conclusion from the available empirical studies on inferences is that text level processes are located somewhere in the right hemisphere, both for automatic and controlled inferences. For instance, Beeman [2] found that bridging inferences based on associative relationships between content words are difficult after right-hemispheric lesions. Based on these results and converging results from hemifield priming studies (for a summary see Ref. [12]), Beeman [3] proposed the so-called *coarse coding hypothesis*. This hypothesis holds that the activation of a wide semantic field of related concepts is a process subserved by the right hemisphere, while the inhibition of contextually irrelevant associations takes place in the left hemisphere. Thus, Beeman argues, the right hemisphere patients' failure to draw bridging inferences is due to their failure to activate a sufficiently large field of world knowledge associations.

Regarding more controlled inferences, studies on right hemisphere patients' deficits have focused on pragmatic functions [9,10]. For instance, the comprehension of indirect requests, understanding jokes and metaphors, or the appreciation of an utterance's social implications, are processes said to be impaired after right hemisphere lesions. Inferences required for this type of non-literal language comprehension are, of course, much more dependent on the use of general world knowledge and interpretative functions than the bridging inferences studied by Beeman [2]. Underlying processes suggested to cause this type of deficit include difficulties with revising previous interpretations, deficits in monitoring, detection and repair of comprehension problems, the failure to go beyond the literal meaning and consider alternative readings, or the failure to take into account situational constraints.

As pointed out by McDonald ([36]; see also Ref. [48]), patients with frontal lobe lesions and patients with right-hemisphere damage exhibit remarkably similar non-aphasic language deficits. In empirical research, these two patient groups are often not clearly separable; however, when an attempt is made to differentiate lesion lateralization, the special role of the left prefrontal cortex for text level language processing becomes apparent [15,25,42]. An alternative approach to dissociating frontal and right-hemispheric language functions is the mapping of brain activity in healthy subjects.

In contrast to neurolinguistic research, neuroimaging studies on language processes have only recently begun to venture beyond the word level. Besides first investigations of syntactic processes [13,24,26,38], there are a few positron emission tomography (PET) or functional magnetic resonance imaging (fMRI) studies on the processing of language in a discourse context. Using PET, Mazoyer et al. found extensive, predominantly left-lateralized, temporolateral and inferior prefrontal activation for listening to stories compared to a resting baseline ([35]; see also Ref. [53]). The regions specifically implicated in processing of continuous speech, as compared to the processing of single words, were the polar region of the temporal lobes (BA 38) bilaterally. However, the only region activated solely during listening to stories, but not during processing of other meaningful language stimuli, was in the left superior frontal gyrus (BA 8).

Converging evidence for this finding was reported by Fletcher et al. [17] from a PET study. Their direct comparison of story processing with the comprehension of unrelated sentences yielded an increase of brain activity in the polar region of the temporal lobe (BA 38) bilaterally, but also in the posterior division of the left superior temporal gyrus (STG; BA22/39) and in the posterior cingulate region (BA 23/31). A left frontomedian region (BA 8) presumably close to that reported by Mazoyer et al. [35] was found to be active only during the processing of 'Theory of Mind' stories, i.e. stories that required the attribution of a character's motivations and mental state (see also Ref. [19] for a replication). While Mazoyer et al. did not use any comprehension task, Fletcher et al. explicitly instructed the participants to attempt a motivational attribution for the Theory of Mind stories, but not to do so for other stories. Using instructions to induce different levels of processing as well, Nichelli et al. [40] presented Aesop's fables during PET scanning. When comparing a task requiring semantic analysis of the input with a text level task requiring the deduction of the fable's moral or theme, the latter activated the orbital part of the inferior frontal gyrus (BA 47) and the anterior portion of the middle temporal gyrus (BA 21) in the right hemisphere. Because the derivation of the gist of a story requires the integration of disparate text information across sentence boundaries, the authors interpreted this result as evidence for the coarse coding hypothesis.

Two recent imaging studies explicitly targeted coherence processes [33,49]. Both are based on a classic paradigm in which comprehension processes are manipulated by providing or withholding theme information prior to reading short stories [8]. The wording of the stories is kept so vague that comprehenders cannot succeed in establishing global coherence, unless they are given a title or an organizing picture. When this theme information is available, it activates a mental model or *situation model* of the text, i.e. a global cognitive representation of what the text is about [55]. Each sentence in the story can now be linked to this situation model. Consider, for example, the incoherent sentence pair 'First you have to sort everything into piles. Powder is needed.'. When theme information is missing, these sentences cannot be connected into a unitary, coherent representation. When the topic 'Washing clothes' is given, however, the sorted laundry of the first sentence and the laundry detergent of the second can be associatively connected. Thereby, the sentences become easily comprehensible and locally coherent.

Testing the hypothesis that the right hemisphere is involved in coherence building, St. George et al. [49] presented stories visually with or without a title. The data from fMRI were then analyzed in an inferior prefrontal and three (superior, middle, inferior) temporal regions of interest. Consistent with the hypothesis, there were differential effects for the two hemispheres. In the left hemisphere, activation was comparable for the two conditions. In the right hemisphere, in contrast, the untitled, and thus incoherent condition, yielded more active voxels than the titled, coherent condition. This result was particularly pronounced in the temporal cortices. Once more, these findings of an involvement of right hemispheric brain areas during the processing of language in context, were interpreted as being consistent with the coarse coding hypothesis. Maguire et al. [33] used the pictorial version of the paradigm [8]. A picture showing an unusual scenario is presented before the presentation of its verbal description. Without the picture, the text information alone does not suffice for the comprehender to derive a situation model and thus, to render the text coherent. Maguire et al. [33] report results from PET scanning of participants while they listened to stories with or without having seen a relevant picture. Surprisingly, there was almost no overlap with the results of St. George et al. [49]. When the unusual stories presented with a relevant picture were compared to the same stories without pictures, a posterior cingulate area (BA 31) proved active. Moreover, when other, inherently coherent stories, were compared to the stories without a picture, activation was found in the polar region of the left temporal lobe (BA 38) and in a ventromedial orbital region (BA 11).

Another means to manipulate coherence, while holding semantic and pragmatic features of the language input constant, was employed by Robertson et al. [46]. They presented lists of sentences with unrelated content. When the articles in these sentences were all indefinite (e.g. 'A carpenter walked down the street. Some children were crying.'), the comprehenders were less likely to link the sentences into one situation model. In contrast, when the articles were replaced with definite articles ('The carpenter walked down the street. The children were crying.') it is more likely that the sentence lists are considered coherent, and that a common situational representation is formed. fMRI showed more activation in right-sided prefrontal areas for the coherent sentence lists (with definite articles) than for the less coherent sentence lists (with indefinite articles).

Taken together, the results from neuroimaging studies provide a somewhat inconsistent picture. While there was evidence for right hemispheric involvement during inferencing ([40,46,49]; see also Ref. [7]), the regions of activation reported in other studies were predominantly left lateralized or bilateral [17,33,35]. There is some evidence on the contribution of the frontomedian wall, but its specificity remains unclear.

Of course, there are important methodological differences between these studies, such as the comprehension task, the choice of a baseline and the reported comparison, the extent of the regions scanned or the regions of interest used for analysis. However, the most important difference causing the disparate results may concern the language materials used. In all studies, the coherent stimuli were narrative texts consisting of several sentences, and more than 100 words in length. During the comprehension of such stories, global processes, such as setting up a situation model, play a crucial role. These processes are greatly influenced by the macro-structure of the text [27,28], by the familiarity of the content, by its affective connotations, and by the comprehender's goals. For example, the stories in the Maguire et al. study differ from those used by St. George et al. with respect to the familiarity of the topic.

The present study was designed to map local coherence processes independent of global text factors or task induced strategic processes. Taking advantage of an event-related fMRI methodology we are able to scan the hemodynamic response for trials lasting a few seconds only. Thus, it is possible to use materials for which local and global coherence coincide. Instead of stories consisting of a number of sentences with differing degree of coherence to each other and to the global theme we used 'minimal stories' made up of pairs of sentences. Because we were interested in identifying prefrontal brain structures involved in text level processes, the sentence pairs were not linked via simple associative connections between content words, but required the active retrieval and use of general world knowledge. Examples are provided in Table 1.

To be able to contrast successful coherence building with the detection of incoherence, while holding lexical, syntactic and semantic features constant, we switched the first sentences of pairs of trials and thereby created unrelated sentence pairs. In order to avoid task artifacts, we used a simple coherence judgment task. Participants were asked to decide whether the two sentences were pragmatically related to each other or not. To manipulate the ease of the coherence judgment, we used a lexical feature. In half of the sentences, we added *cohesive ties*, i.e. lexical information explicitly signaling the relationship between sentences. Cohesive ties, such as pronouns or conjunctions, influence sentence processing (e.g. Refs. [39,46]), but their use is often difficult for brain-injured patients (e.g. Ref. [30]). Note, that cohesion does not render inference processes superfluous, it just facilitates them. For instance, in the first example in Table 1, the causal connective 'therefore' indicates the type of inference, but not its pragmatic content. The comprehender still has to infer the full causal chain: palms become sweaty when someone is nervous, and exams make people nervous. In the third example, on the other hand, the cohesive tie 'that's when' provides the temporal relationship be-

Table 1
Example materials for the four conditions of the experiment (translated from the original German); cohesive ties are printed in italics

[1] Coherent/Incohesive
Mary's exam was about to start. The palms were sweaty.
Laura got a lot of mail today. Some friends had remembered the birthday.
Sometimes a big truck drives by the house. The dishes start to rattle.
The lights have been on since last night. The car doesn't start.

[2] Coherent/Cohesive
Mary's exam was about to start. *Therefore*, *her* palms were sweaty.
Laura got a lot of mail today. *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 why* the car doesn't start.

[3] Incoherent/Incohesive
Laura got a lot of mail today. The palms were sweaty.
Mary's exam was about to start. Some friends had remembered the birthday.
The lights have been on since last night. The dishes start to rattle.
Sometimes a big truck drives by the house. The car doesn't start.

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

tween the first and second sentence. However, the underlying causal connection of the heavy truck making the ground shake still needs to be inferred by the reader.

We assumed that crossing cohesion and coherence would lead to the following effects: When a cohesive tie is present in the coherent trial (Condition [2]), the inference becomes easier, because of additional, explicit lexical information. In contrast, when a cohesive tie falsely indicates a relationship between the sentences (Condition [4]), processing is rendered more difficult. This might be due to a competition between the pragmatic incoherence and the linguistic cohesion. Furthermore, in many cases, this condition yields a pragmatic garden-path effect, i.e, the usual communicative assumption of coherence leads to implausible or funny scenarios (as, for instance, in the last example in Table 1). These hypotheses were confirmed in a behavioral pretest (see Section 2.3).

With this type of material we conducted an event-related fMRI study. As a control condition, we used a baseline task in which non-word strings, similar in appearance to the experimental materials, were shown in the context of a purely physical task. The question of interest was whether we could dissociate processes related to language comprehension on the word and sentence level from those used for establishing coherence across sentence boundaries. In particular, we were interested in the lateralization of those latter processes and in the involvement of prefrontal brain regions indicating the use of executive functions.

2. Methods

2.1. Participants

Eight men and four women received reimbursement for participating in the experiment. Eleven of the 12 participants were students; and all were right-handed. None of the participants had any history of neurological disorder or other health problems preventing them from being exposed to the magnetic field. The average age was 23 years (SD=3.1, range 19–31). All participants had given informed consent.

2.2. Design and materials

The language trials were made up of 120 sentence pairs in which the second sentence (the target) was pragmatically related to the first (the context). Each target occurred in two versions: the cohesive version contained one or two lexical items (for instance, pronouns, 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. Thus, the experiment used a 2×2 within-subjects design with the factors Cohesion (yes/no) and Coherence (yes/no). Four different lists of 30 trials per condition were then created with the following constraints: in each list each target sentence appeared exactly once, and across lists, each target sentence appeared exactly once in each of the four conditions. The sentences were printed in upper and lower case. Following the rules of German orthography, the initial letters of words were printed in upper case when the word was at the beginning of a sentence or when the word was a noun. All other letters were printed in lower case.

Care was taken that the incoherent sentence pairs did not yield unintended pragmatic relationships. To confirm this and to test the hypothesis that Cohesion affects inference processes, the sentence pairs were pretested in a behavioral experiment (see Section 2.3).

For the control condition in the fMRI experiment, we used an additional 30 trials made up of non-word sentences. These stimuli were created by replacing letters in real words, so that the resulting displays resembled the language trials in appearance. The resulting non-words did not adhere to the orthographic and phonemic rules of German and were therefore not pronounceable. Half of the sentences were printed in all upper case, and half in upper and lower case, to closely resemble the experimental sentences in appearance. The non-word sentences were combined so that in 15 trials, the context and target had consistent case, and in 15 trials they were in different case.

Each of the four lists of experimental items was combined with the control condition trials. The order of the resulting 150 trials was then randomized, so that within each of three blocks of 50 trials, each condition appeared ten times.

The length of the target sentences was 6.9 words on average, with a range of 4–10 words. To avoid potential eye movement artifacts, we split the sentences into two displays to be presented successively. When possible, the line break coincided with phrase boundaries. The cohesive and incohesive versions of the target sentences were always split at the same position in the sentence. Since coherence was manipulated by switching the context sentences, rather than the target sentences, the line breaks and the presentation formats were identical in the coherent and the incoherent conditions. 60% of the context sentences were split into two displays as well, whereas the shorter context sentences were shown in only one display.

2.3. *Pretest*

The materials were pretested in a reading time experiment. The 120 trials were intermixed with 40 filler trials and presented one sentence at a time. Twenty-four students were instructed to read the sentences for comprehension and press a key on a button box whenever they had finished reading a sentence. After the presentation of the second sentence, a question mark appeared prompting the participants for a coherence judgment. The reading times for the two sentences, the proportion of correct responses and the judgment times were recorded. All dependent variables were then subjected to a 2×2 analysis of variance (ANOVA) with the factors Coherence and Cohesion.

The results confirmed the efficacy of the experimental manipulations. The error rates were very low overall, as can be seen in Table 2. There was a main effect of Coherence ($F(1,23)=5.6$, $P<0.05$), indicating that errors were more frequent for the coherent trials than for the incoherent trials.

For the reading times of the context sentences, there were no significant effects. Because the cohesive target sentences were slightly longer than the incohesive targets, it was necessary to factor out sentence length in the analysis of the target sentence reading times. For each participant separately, we first regressed the reading times on the number of words in the sentence. The residuals were then used in the statistical analysis. As shown in Table 2, the expected interaction between Coherence and Cohesion resulted ($F(1,23)=20.6$, $P<0.0001$). Cohesive ties facilitated comprehension of coherent sentence pairs, while they rendered the detection of coherence gaps more difficult. Planned pairwise comparisons confirmed that this interaction was mostly due to the inconsistent condition [4] in which cohesive ties falsely suggested a relationship between incoherent sentences ($F(1,23)=7.8$, $P<0.02$ for the main effect of Coherence in cohesive trials; $F(1,23)=2.2$, $P>0.15$, for the effect of Coherence in the incohesive trials). Despite the fast reaction times for the coherence judgments, the interaction between Cohesion and Coherence was significant for this dependent measure as well ($F(1,23)=5.7$, $P<0.05$).

The pretest confirmed that felicitous cohesive ties can facilitate coherence processes, and that misleading cohesive ties can render the detection of a coherence break more difficult. These effects were present most clearly in the sentence reading times, rather than in the accuracy or speed of the coherence judgments.

2.4. *Procedure*

The participants received written instructions before the scanning session started. Using examples for coherent and incoherent sentence pairs, the participants were told to indicate with a button press whether they considered the sentence pairs to be pragmatically related to each other or not. The participants were informed that trials with non-

Table 2
Results of the pretest of the sentence materials (mean (SD); $n=24$)

	Coherent		Incoherent	
	Incohesive [1]	Cohesive [2]	Incohesive [3]	Cohesive [4]
Correct responses (%)	92.8 (6.3)	94.9 (6.8)	97.2 (3.9)	96.5 (5.4)
Judgment times (ms)	461 (198)	459 (176)	451 (210)	475 (215)
Reading times for context sentence (ms)	2313 (717)	2300 (728)	2276 (595)	2274 (695)
Reading times for target sentence (ms)	2348 (539)	2359 (514)	2261 (527)	2550 (595)
Residuals of target sentence reading times (ms)	−95 (222)	−162 (254)	−193 (173)	37 (176)

word sentences would be interspersed. The task for these trials was to indicate whether the appearance (i.e. the typing in either all upper case or upper and lower case) was the same or different for the two non-word sentences. For this control task, examples were provided in which the letter case was either consistent across the two non-word sentences or not. For their YES-responses, half the participants were instructed to use the key on right side of the response box, and the other half to use the key on the left.

During the scanning session, the participants lay flat inside the magnet, with a response box placed into their right hands. The stimuli were projected onto a ground glass screen placed in the magnet bore above the subject's head. The subject viewed the screen using mirrored glasses with corrective lenses when necessary. The letters were printed in white on black background to avoid glaring, and a font size was chosen to allow for comfortable reading.

The functional measurement was conducted in three separate blocks, after each of which the participant could take a short break. Each block of 50 trials was preceded by one additional practice trial. The five conditions were presented in a pseudo-randomized order so that successive trials were from different conditions, and not more than three successive trials required the same response. Each trial lasted 20 s, with the following time course: After presentation of the fixation cross for 2 s, the one or two displays of the context sentence and the two displays of the target sentence were presented for 2 s each. If no response had been given before the offset of the last display, a question mark appeared as a reminder for the participant to provide the coherence judgment or the letter size consistency judgment. After the response, the screen was cleared and stayed blank for the remainder of the trial.

2.5. Data acquisition

A Bruker Medspec 30/100 system was used for magnetic resonance imaging at 3.0 Tesla. Prior to the functional scans, two anatomical scans were acquired for each participant using MDEFT sequences [54]. The first was a whole brain image acquired with a T1-weighted three-dimensional (3D)-segmented sequence ([29]; 128 sagittal, adjacent slices, 1.5 mm thick, 256×256 pixel matrix per slice). To enable alignment of the functional scans with this high-resolution image, anatomical 2D images were acquired, using the same number and orientation of slices as the functional scans ([41]; TE=6.1 ms, TR=1300 ms, 256×256 matrix).

During the functional scans, the BOLD response was measured using a single-shot gradient EPI-sequence (matrix 64×64, TE=30 ms, flip angle 90°, field of view 192 mm). Horizontal images were acquired for 16 slices (5 mm thickness, 2 mm spacing), parallel to the bicommissural plane (AC–PC). For most participants, six slices were below the AC–PC line, while ten slices were above, but care was taken that the temporal lobes, as well as prefrontal regions, were covered in full. In-plane resolution was 3×3 mm. We used a repetition time of 2 s (TR=2), and the presentation of the displays was triggered by the acquisition of the first slice of the current image.

2.6. Data analysis

Data analysis was conducted using the software package LIPSIA, developed as a tool for analyzing functional MRI data [32].

For each participant, the signal acquired during the functional scans was preprocessed as follows: First, a sinc-interpolation algorithm was applied to correct for the temporal spacing between the 16 slices of each image. Motion correction consisted of a global affinity linear transformation that optimized for each time step the linear correlation between the image at that time step and a predefined reference image. A baseline correction was then conducted using a temporal high pass filter with a cutoff frequency of 1/120 Hz. Furthermore, a spatial Gaussian filter was applied with a standard deviation of $\sigma=0.6$. With a voxel size of 3 mm, this standard deviation is equivalent to a full width at half maximum of 4.2.

For the statistical analysis we used a random effects model, with an algorithm identical to that used in SPM99 [18]. Specifically, we carried out the following sequence of processing steps: For each subject, the 2D-data were analyzed using the General Linear Model based on the aforementioned five conditions (four experimental conditions, one control condition). For the event-related model we time-locked the responses at the onset of the last display of the target sentence, and we assumed a lag of the hemodynamic response of 6 s from this event. Contrast codes were then used to detect significant activations by calculating t-statistics based on the parameter estimates of the full linear model.

We carried out the following six comparisons. The first contrast tested all four experimental conditions against the control condition, to see whether the resulting activations properly reflect language processing. For the language trials, the next three contrasts corresponded to testing the effects of the 2×2-design; specifically, one contrast tested the effect of Cohesion, one the effect of Coherence, and one contrast the interaction between these two factors. Finally, we used two planned pairwise comparisons to further specify the resulting interaction.

For the statistical analysis across participants, the six contrast images of each participant were fitted into a standard stereotaxic space [50] as follows: First, the 2D anatomical scan was rotated and shifted so that it mapped onto the 3D whole brain image. Then, linear scaling factors were calculated to transform the image to the standard size. The resulting co-registration matrices were finally applied to the six contrast images using trilinear interpolation [31]. For inferential statistics across participants, the standardized 3D-images for each contrast were

Table 3
Behavioral data for the coherence judgments during the fMRI scanning ($n=12$)

	Coherent		Incoherent	
	Incohesive [1]	Cohesive [2]	Incohesive [3]	Cohesive [4]
Correct responses: mean and SD in percent	92.8 (6.3)	94.2 (6.8)	95.3 (3.9)	93.1 (5.4)
Judgment times: mean and SD in ms	2041 (576)	2026 (590)	2091 (495)	2129 (506)

used to calculate voxelwise t-tests. The resulting t-values were then transformed into Z-scores. Only those voxels with $|Z|>3.09$ were considered significantly activated ($P<0.001$). Since this probability level is uncorrected for multiple comparisons, we defined an additional spatial extent threshold as follows: Areas of activation smaller than 125 mm^3 were neglected. This threshold roughly corresponds to the requirement that more than two adjacent voxels in the original image (of volume $3\times3\times5$ mm$=45$ mm^3) be significantly activated.

3. Results

3.1. Behavioral data

For each participant and each condition, the percentage of correct responses was calculated. In the baseline condition, participants answered 93% of the questions correctly (SD=7.1%). For the coherence judgments in the experimental trials, accuracy was just as high (mean=94%, SD=3.0). None of the effects in the 2×2-analysis with the factors Cohesion and Coherence were significant (see Table 3).

The reaction times were measured from the onset of the last display to the button press. Thus, times shorter than 2 s indicate that the response was given while the last display was still on the screen. Before averaging, all reaction times were first corrected by replacing high and low values with a cutoff value of 2 standard deviations above or below each participant's mean. For each participant and each condition, we then calculated mean reaction times.

In the baseline condition the mean reaction time was 1994 ms (SD=673). The means for the coherence judgments are shown in Table 3. Once more, none of the effects in the ANOVA with the within-subjects factors Cohesion and Coherence was significant ($F(1,11)=2.9$, $P=0.11$ for Coherence, $F(1,11)<1$ for Cohesion, and $F(1,11)=2.0$, $P=0.18$ for the interaction Cohesion$\times$Coherence). A similar result obtained when the reaction times were corrected for the length of the last display.

The reason for the failure to find significant results in the behavioral data was that, in contrast to the reading time pretest, the fMRI experiment required fixed presentation times. Although the response could be given as soon as the last display was shown, a majority of the participants waited almost always until its offset to provide their judgment. Thus, for these subjects, the reaction times did not properly reflect processing difficulty of the target sentence. Because of the clear results of the pretest, and the trends shown in the fMRI experiment, we can nevertheless be confident that the experimental conditions indeed required the different processes targeted.

3.2. fMRI results

Table 4 shows the areas of activation for the comparison of the language trials with the control task. The absence of activation in the visual cortex confirms that the control condition was appropriate for filtering out activation related to perceptual processing of the visually presented stimuli.

As expected, reading and comprehending sentences involved an extended frontotemporal area on the lateral surface of the left hemisphere. For better reference, we included in Table 4 the coordinates of local maxima within

Table 4
Brain regions significantly activated for the language conditions compared to the control condition

Language vs. control	Size (mm^3)	Z-Score	Side	Talairach–Fox coordinates		
				x	*y*	*z*
Superior frontal gyrus (BA 6) [1]	145	3.42	L	−10	15	52
Superior frontal gyrus (BA 8) [2]	2425	4.54	L	−8	44	42
Superior temporal sulcus (anterior) (BA 38/21) [3]	18434	4.90	L	−50	−12	−7
Superior temporal sulcus (posterior) (BA 21/37) [4]		4.84	L	−59	−49	13
Inferior frontal sulcus/precentral sulcus (BA 44/6) [5]		4.26	L	−46	20	19
Inferior frontal gyrus: pars triangularis (BA 45) [6]		4.23	L	−46	28	5
Posterior cingulate cortex/precuneus (BA 23/30) [7]	1188	3.93	L	−8	−54	23
Inferior frontal gyrus (BA 45) [8]	238	3.67	R	44	27	11
Superior temporal sulcus (anterior) (BA 22) [9]	1372	4.75	R	43	−10	−8
Cerebellum [10]	492	4.06	R	18	−79	−20

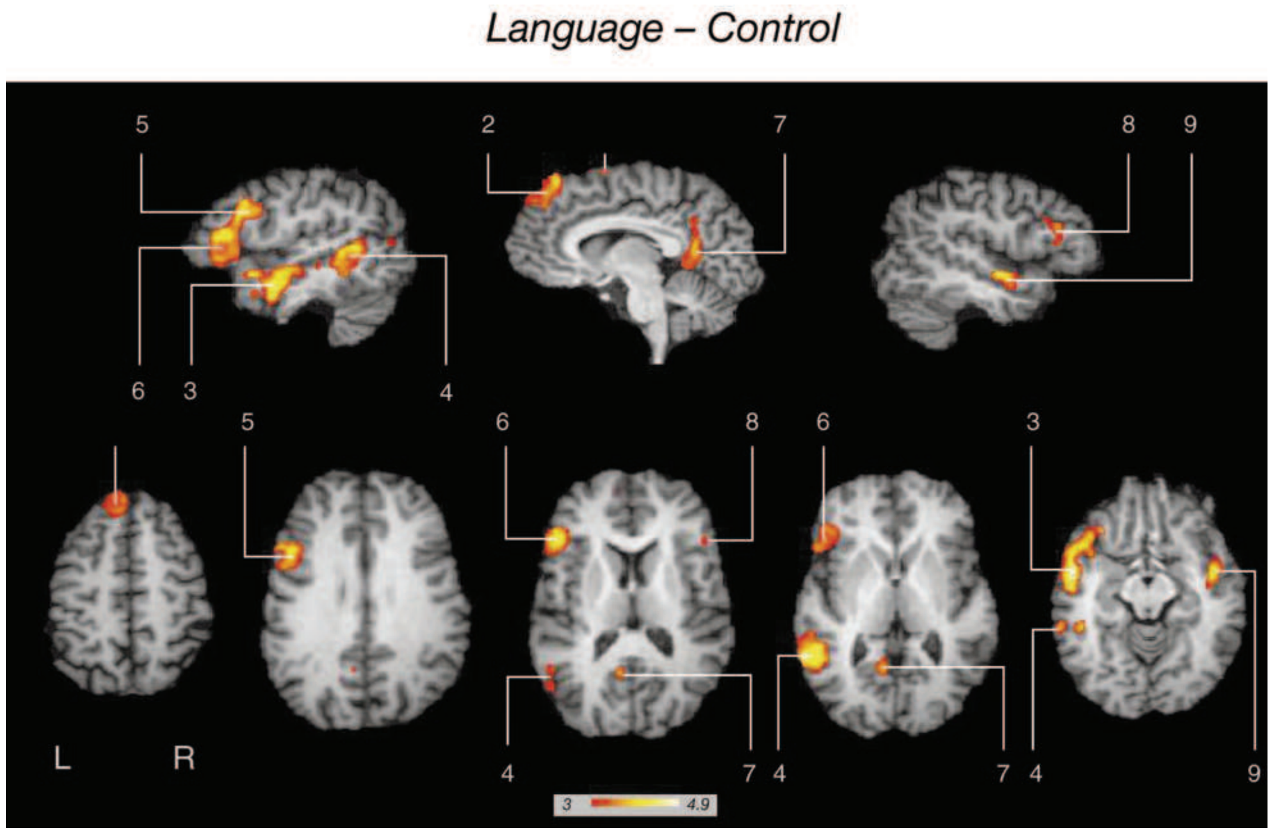


Fig. 1. The Z-map from the contrast comparing the language conditions to the physical control condition. The group data based on all 12 participants are superimposed on an individual brain. The upper row contains sagittal slices illustrating the left-lateral ($Y=-48$), the left-median ($Y=-8$) and the right-lateral ($Y=+44$) views. The second row shows corresponding horizontal slices in five planes ($Z=+38$, $+22$, $+10$, $+1$, -9), with the left side of the brain being on the left side of the images. The identification numbers for the activated regions are those provided in Table 4. Refer to the table for a more detailed description of the regions' characteristics.

this large region. As can be seen in Fig. 1, the temporal activation extended along the bank of the superior temporal sulcus (STS), reaching into the posterior division of the middle temporal gyrus (MTG; BA 21/37), but sparing most of the superior temporal gyrus (STG), and in particular, its posterior portion. The inferior temporal gyri were not active either. Besides a large activation peak covering the posterior part of the STS, there was a second peak in its anterior portion, extending into the polar region of the temporal lobe (BA 21/38).

The lateral prefrontal activation consisted of two adjacent regions in the ventrolateral frontal cortex. The more ventral peak centered on the pars triangularis of the inferior frontal gyrus (IFG; BA45), whereas there was almost no involvement of the neighboring pars opercularis (BA 44, Broca's area) during this language processing task. The second left frontolateral peak was located more superiorly, extending to the inferior frontal sulcus (IFS) and posteriorly to the inferior precentral sulcus (IPcS; BA 44/6). For both the anterior temporal and the inferior prefrontal activation we also found the right-hemispheric homologues to be somewhat involved, but clearly to a much smaller extent. Furthermore, the deep right cerebellum proved active.

Along the left superior frontal gyrus (SFG) we found two regions of activation, a small one in BA 6 and a more prominent one in BA 8. The latter extended from the dorsal surface of the SFG into its median division. Furthermore, we found an area of activation in the inferior precuneus (BA 23), extending downwards to the retrosplenial area (BA 30).

To shed light on the influence of the experimental variables, we now considered differences between the four conditions. In analogy to the 2×2-ANOVA reported for the behavioral data, we calculated the contrasts corresponding to the two main effects and their interaction. To further specify the interaction, we also conducted pairwise comparisons: for both the incoherent and the coherent conditions separately, we calculated the contrast testing the effect of Cohesion. The semantic and pragmatic content is identical for both cohesive and incohesive sentence pairs, as is the required response. Thus, these pairwise comparisons test the effect of lexical cohesion only, without possible confounds (e.g. retrieval of general world knowledge, response type or response success). Recall that we know from the pretest that, depending on coherence, cohesion has just the opposite effect: While cohesive ties aid inference processes in coherent trials, they hinder the detection of a coherence gap in the incoherent trials. The coordinates of the resulting activations are shown in Table 5.

The first contrast compared cohesive and incohesive trials. The most apparent difference between these conditions was the sentence length. Consistent with an impact of this feature on reading processes, independent of comprehension processes, we found activation in the vicinity of the frontal eye fields bilaterally (BA 6), which is displayed in Fig. 2. For no other brain regions were there increases in regional cerebral blood flow, and there were no significant activations in the comparison of incohesive vs. cohesive trials.

More central to the questions posed in this study is the comparison between coherent and incoherent trials. Unexpectedly, incoherent trials did not differentially activate any brain regions. The rejection of pragmatically unrelated sentence pairs did not yield measurable activation differ-

Table 5
Brain regions significantly activated for the contrasts testing the effects of Cohesion and Coherence

	Size (mm^3)	*Z*-Score	Side	Talairach–Fox coordinates		
				x	*y*	*z*
Main effect of Coherence:						
Coherent vs. Incoherent ([1],[2]–[3],[4])						
Posterior cingulate cortex/inferior precuneus (BA 23/31) [11]	178	3.42	L	−5	−34	39
Frontomedian wall/superior frontal gyrus (BA 9/10) [12]	1843	5.18	L	−4	58	13
Main effect of Cohesion:						
Cohesive vs. Incohesive ([2],[4]–[1],[3])						
Superior frontal sulcus/superior precentral sulcus (BA 6) [13]	133	4.03	L	−25	0	34
Superior frontal sulcus/superior precentral sulcus (BA 6) [14]	217	4.35	R	24	1	37
Interaction: Coherence×Cohesion:						
(a) difficult vs. easy ([1],[4]–[2],[3])						
Inferior frontal sulcus/inferior precentral sulcus (BA 44) [15]	143	3.77	L	−40	22	15
(b) Incoherent only:						
cohesive–incohesive ([4]–[3])						
Inferior precentral sulcus (BA 44/6) [16]	296	3.69	L	−44	9	19
Inferior frontal sulcus/inferior precentral sulcus (BA 44) [17]	328	4.02	L	−31	17	17
(c) Coherent only:						
cohesive–incohesive ([2]–[1])						
Intraparietal sulcus (BA 39/7) [18]	415	4.66	L	−28	−60	48
Supramarginal gyrus (BA 40) [19]	342	4.47	L	−45	−40	43

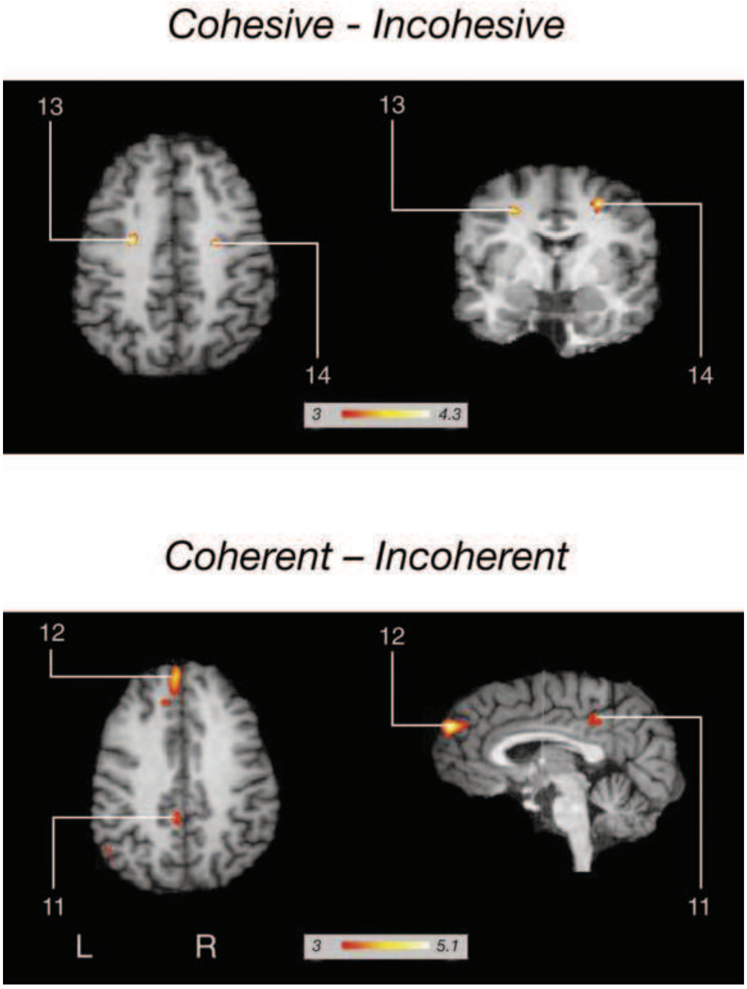


Fig. 2. Activations due to the experimental factors Coherence and Cohesion. The upper panel displays the *Z*-map for the main effect of Cohesion, independent of Coherence. The two significant brain regions in the subtraction of incohesive trials from cohesive trials are illustrated in a horizontal and a coronal view, centered in the Talairach–Fox coordinates of the left-sided region ($X=-25$, $Y=0$, $Z=34$). The lower panel illustrates the main effect of Coherence: displayed is the *Z*-map for the contrast subtracting the incoherent sentence pairs from the coherent sentence pairs, centered in the frontomedian region ($X=-4$, $Y=58$, $Z=13$). The identification numbers refer to the labels in Table 5, which contains further characteristics of the activated regions.

ences. In contrast, the successful establishment of a pragmatic connection, i.e. the derivation of an inference, activated two areas in the median wall of the left hemisphere, shown in the lower panel of Fig. 2. The posterior activation involved posterior cingulate (BA 31) and inferior precuneal (BA 23) areas. The much more prominent anterior activation, located in the medial portion of BA 10/9, seems to be the anteriormost extension of the frontomedian area that was equally activated in all language conditions.

When looking at the contrast reflecting the interaction between Cohesion and Coherence, we found one small region to be activated. The more difficult conditions elicited activation along the inferior bank of the IFS (BA 44), corresponding to the area activated in all language trials. Thus, this region was involved in the processing of all conditions, but more so when the task was more demanding (coherent/incohesive, incoherent/cohesive).

To further specify this interaction we calculated pairwise comparisons for each of the coherence conditions separately. In these comparisons, whose results are shown in Fig. 3, response, memory demands, and pragmatic knowledge are controlled. The sentences differ in one lexical feature only, namely the presence or absence of a cohesive tie. For the coherent condition, cohesion yielded small activations in the left supramarginal gyrus (BA 40) and along the intraparietal sulcus (BA 39/7). In contrast, for incoherent trials, for which the reaction time difference indicated a garden-path for the misleading cohesive sentences, cohesion yielded activation in the left inferior frontolateral cortex. One peak was close to the junction of the IFS and the IPcS, and the second peak lay within the

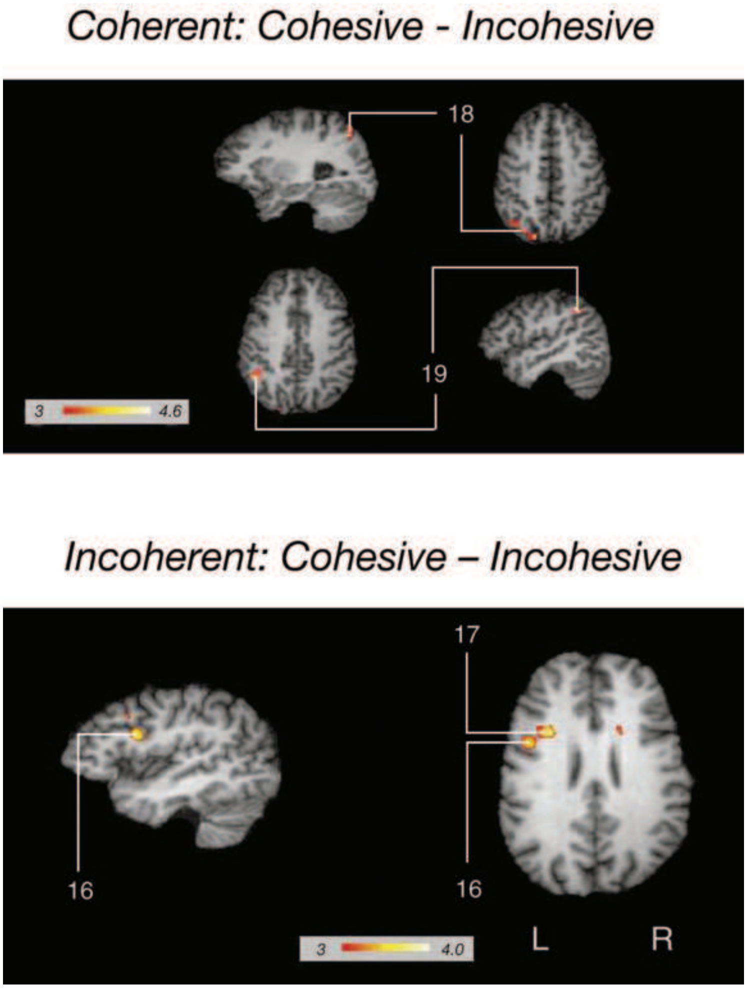


Fig. 3. Activations due to the interaction between Coherence and Cohesion. The upper panel displays the contrast testing the effect of cohesion for the coherent sentence pairs only. Images for the two regions that reached significance are shown separately. The first row displays the activation with the peak in ($X=-28$, $Y=-60$, $Z=48$), while the second row shows the activation with the peak in ($X=-45$, $Y=-40$, $Z=43$). The lower panel displays the effect of Cohesion on the coherent sentence pairs. Activation elicited by incohesive, coherent trials is subtracted from activation elicited by cohesive, coherent trials. Shown are a lateral and a horizontal view, centered in ($X=-44$, $Y=9$, $Z=19$). The identification numbers refer to the labels in Table 5 containing the characteristics of the areas of activation.

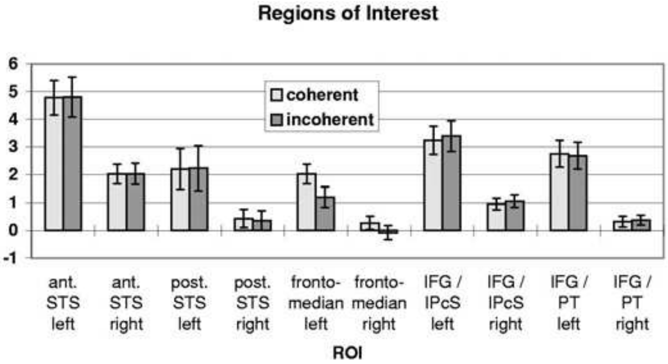


Fig. 4. Regions of interest analysis for lateral frontotemporal areas and the frontomedian wall. The means of the average Z-values in each of five regions of interest are displayed as a function of Coherence and Hemisphere. The error bars indicate one standard error above and below the mean. The left-hemispheric regions are defined as the spheres of radius 6 mm around the frontotemporal activation peaks provided in Table 4: ant.STS=[3], post. STS=[4], IFG/IPcS=[5], IFG-PT= [6]. The right hemispheric regions are defined as the homologue areas. The frontomedian ROI includes both the peak of area [2], as well as the peak for the contrast comparing coherent to incoherent trials [12].

lower branch of the left IPS. This result suggests that the significant interaction was mostly due to increases in processing demands for the incoherent, cohesive condition, rather than due to processing difficulties caused by missing cohesive ties in the coherent condition.

3.3. Regions of interest analysis

In order to further evaluate the apparent lateralization differences found in the group statistics, we conducted an additional analysis based on the activations in specifically defined regions of interest (ROIs). To each of four left-lateral regions of activation identified above, we defined a ROI of size 925 mm^3 as the sphere with radius 6 mm around the activation peak. To all four regions, we also considered the right-sided homologue. The activation levels in each of the resulting eight ROIs were obtained using the procedure suggested by Bosch [6]: for each participant separately, we first calculated individual Z-maps for both the contrast comparing coherent trials to the control task, as well as for the contrast comparing incoherent trials to the control task. After normalization and co-registration into a standard 3D-space, the average Z-values in the eight ROIs were determined for each participant. These means, whose averages are shown in Fig. 3, were then entered into a 2 (Hemisphere)×2 (Coherence)×4 (Region) ANOVA. There was a highly significant main effect of Region ($F(3,33)=10.7$, $P<0.0001$), a highly significant main effect of Hemisphere ($F(1,11)=44.1$; $P<0.0001$), but none of the effects involving the factor Coherence came close to significance (all Fs<1). Thus, this analysis confirms that the right-sided activations were considerably smaller than the left-sided ones, and suggests that Coherence did not modulate this left-hemispheric dominance.

To evaluate the frontomedian activation, we defined a larger ROI of size 2579 mm^3, including both the activation peak in the SFG found during all language trials, as well as the peak significant in the comparison of coherent vs. incoherent trials. Once more, the homologue area in the right-sided median wall was also considered. The centers were slightly shifted away from the median, so that there was no overlap between the left-sided and the right-sided ROIs. Average Z-values were then obtained for each participant as described above, and their means are also displayed in Fig. 4. In the 2 (Coherence)×2 (Hemisphere)-ANOVA, the clear main effect of Hemisphere confirmed the left-sided dominance ($F(1,11)=12.7$, $P<0.01$). Moreover, there was a highly significant main effect of Coherence ($F(1,11)=20.6$, $P<0.001$). The interaction of Coherence with Hemisphere shows that this effect was mostly driven by activation differences in the frontomedian wall of the left hemisphere ($F(1,11)=10.4$, $P<0.01$).[1]

---

[1] Because of well-known lateralization differences between men and women, we included in a second analysis the between-subjects factor Gender. Despite the small sample size of four women and eight men, the results showed marginally larger activation for women in the frontotemporal ROIs ($F(1,10)=4.6$, $P=0.07$), but Gender neither interacted with Hemisphere nor with Coherence. Contrary to the expectation of women's patterns of activation being less lateralized than men's, we obtained for the frontomedian ROI a significant interaction between Gender and Hemisphere, indicating that the difference between left and right-sided mean Z-values was even larger for women (mean=2.7, SD=1.4) than for men (mean=0.9, SD=1.1). This post hoc analysis suggests that gender differences might be an interesting topic for future research, but they do not seem to account for the lateralization effects found in this study.

4. Discussion

The present study was designed to investigate coherence processes during written language comprehension. Using whole-head fMRI with stimuli targeting local coherence processes rather than global text comprehension processes, we aimed at evaluating the respective contributions of prefrontal areas and the right hemisphere. First, we compared the language trials to a control task to confirm the consistency of our results with previously reported findings on language processing in the visual modality. Because in both language and control trials the stimuli consisted of letters and were similar in appearance, no additional activation in the striate and extra-striate visual cortices was seen for the language stimuli [22,23]. Thus, the judgment of a physical aspect of non-words properly controlled for visual perception and character recognition, without requiring lexical or semantic processing.

For the coherence judgment paradigm, we crossed the factors coherence, i.e. the pragmatic relatedness, and cohesion, i.e. the presence of a lexical connective. The accuracy of the coherence judgments was very high, showing that the conditions were easily separable. A reading time pretest, as well as the trends in the reaction times during the fMRI experiment, confirmed that lexical cohesion facilitates inference processes, while it renders the detection of incoherence more difficult. The factorial design enabled to separate those brain regions in the large network of activation that are involved in language processing on the sentence level from those activated by text level processes.

4.1. The left lateral convexity cortex

The frontotemporal activation found when comparing all language trials to the control trials, independent of condition, was predominantly left-lateralized. The left temporal lobe has been reported to be involved in language processing by numerous authors. In many studies, the activation for lexical or semantic tasks is located in the inferior lateral and basal temporal lobe (e.g. Refs. [16,51]). Other studies reported superior and middle temporal activation as well, not only for auditory but also for visual presentation [11,45]. Price [43] summarizes several findings and argues that the posterior temporal lobe is involved in the phonological and lexical aspects of word retrieval, while more anterior temporal activation seems to reflect the specific requirements of semantic decisions using verbal material. In contrast, Bavelier et al. [1] argue that anterior temporal areas are specifically involved during language comprehension in sentence contexts. Mazoyer et al. [35] report anterior temporal activation in all context conditions, even when the sentences were made up of pseudo-words or when they were semantically anomalous.

Consistent with the proposal of temporal activation being related to word and sentence level processes, we found activation both in anterior and posterior temporal regions along the banks of the left STS, and to a lesser extent, in the right anterior temporal region. These activations did not differ as a function of coherence. In contrast, Fletcher et al. [17] report posterior STG activation for the processing of coherent vs. incoherent language, i.e. for a comparison on the text level. However, more than posterior temporal areas, the polar temporal regions have been connected with text processing, in part based on the clinical observation that patients with anterior temporal lobectomies have specific difficulties with story recall. Neuroimaging results also indicate a role of the left or both temporal poles for processing language across sentence boundaries [17,19,33,35]. The temporal activation found in the present experiment extended into the polar region (BA 38), but the peak was located clearly more posteriorly, at the anterior end of the STS. One possible explanation for the result that this region was equally active in both coherence conditions is that the temporo-polar regions contribute to the construction of representations for larger units of text. The aforementioned studies used passages comprising narratives, so that text processing was guided by a story schema. In our experiment, the sentence pairs described minimal scenarios comprehensible without the need to develop or encode a global structure. Consistent with this account is the fact that Maguire et al. [33] found temporo-polar activation only for their standard stories that followed a canonic story schema. When the structurally unusual texts with and without the illustrating picture were compared, the left temporal pole was not significantly activated, despite the differences in global coherence.

The large left-sided, lateral prefrontal activation had two distinct foci. The peak of the activation in the inferior frontal gyrus was located in the pars triangularis (BA 45). In contrast to the pars opercularis (BA 44), an area shown to be involved during phonological and syntactic processing [13,26], the more anterior left IFG has been implicated in tasks requiring semantic activation, or more specifically, the goal-directed retrieval of semantic information [56] or the selection among competing semantic alternatives [51]. While the exact location within the IFG varies, it seems clear that inferior frontal regions become active in tasks requiring semantic processes (see Ref. [16]). In our experiment, the level of activation in the pars triangularis was comparable across conditions, so that it seems to reflect semantic processing on the sentence level.

In contrast, the second peak, located at the junction area of the IFS and the IPcS, was found to be specifically involved during processing of the more difficult conditions, and in particular, during processing of the incoherent condition with misleading cohesive ties. Therefore, activation in this region does not only reflect sentence level processing, but task-dependent processing influenced by the context condition. It is possible that this activation reflects the increased demand of semantic organization in these conditions (cf. Ref. [51]). However, because of the variety of tasks known to elicit activation in this area, we propose a functional role going beyond language-specific processes. Common to the tasks activating the junction area, including a Stroop paradigm [57] and a task-switching paradigm [14], is that they all require resolving the conflict between competing task sets. Similarly, responding correctly to our incoherent, but cohesive sentence pairs required resolving the conflict between a lexically based versus a pragmatically based response strategy.

4.2. The median wall

During all language trials, an area in the posterior cingulate cortex and the adjacent inferior precuneus (PCC) was active, and only the precuneal activation was modulated by coherence. Maguire et al. [33] report a similar region for their comparison of unusual stories with and without pictures, i.e. for the comparison of coherent vs. incoherent stories. They argue that it is involved in forming a mental model by integrating the current input with the background knowledge, in this case provided in the form of an illustrating cartoon. Fletcher et al. [17] report PCC activation for story comprehension as compared to the comprehension of unlinked sentences, whose exact localization depended on the type of stories presented. Once more, the authors explain this activation with the on-line attempt of integrating the current input with the previously established situation model of the story. A similar activation during a plausibility judgment task [7] suggests that these situational processes can take place on the sentence level, and several experiments using a complex categorization task on the word level also reported PCC activation [4,5]. The results of the present experiment also confirm the role of the PCC region in the encoding of a newly formed situation model representation. While for all sentences, a situation model is being constructed, an updating of the existing one is needed for coherent sentence pairs only.

The most prominent region distinguishing coherent and incoherent trials was the medial portion of BA 9/10, to which the frontomedian activation during all language trials was posteriorly adjacent (BA 8). In the study by Mazoyer et al. [35] a left superior prefrontal focus was the only region exclusively activated by the story condition. Frontodorsal or frontomedian activations were also reported for inductive reasoning [20] and the aforementioned complex categorization and plausibility judgment tasks [4,5,7]. Common to these tasks is that an evaluation is required based on the integration of a verbally presented stimulus with both the experimental context, as well as the general background knowledge. Most importantly, Fletcher et al. [17] implicated a frontomedian area as the region where 'Theory of Mind' is located (see also Refs. [19,20]). In their comparison of stories requiring the attribution of the protagonist's mental state with stories without this component, the frontomedian region was active. Fletcher et al. discuss this similarity with the Mazoyer results and note that the stories used for the latter study also involved Theory-of-Mind attributions. In our experiment, frontomedian/frontodorsal structures were significant both for the comparison of all language trials to the control trials, and, even more clearly, for the comparison of coherent to incoherent sentence pairs. It is possible that, even without explicit instructions to consider the Theory of Mind, this frontomedian activation was mostly caused by those 50% of the trials that contained human protagonists and thus allowed to consider the thoughts and feelings of the people mentioned. To rule out this possibility we carried out a post hoc analysis, testing the contrast for the interaction between Animacy and Coherence. This comparison did not yield any significant areas of activation, indicating that the frontomedian activation was unlikely to depend on Theory-of-Mind attributions to animate protagonists.

An alternative account for the function of the frontomedian wall is related to clinical observations after, for instance, anterior cerebral artery infarctions. The symptoms described concern a reduction of motivation and volition, which in extreme cases can go as far as apathy [34]. In particular, the frontomedian wall is implicated for the internal generation of ideas and plans. Establishing coherence in the present experiment required a presumably non-automatic inference process for the integration of non-overlapping, externally presented information with general world knowledge. We believe that the frontomedian area has a function for the self-initiation of a cognitive process in the context of tasks that require the active utilization of the individual's background knowledge (cf. Ref. [20]). In the Theory-of-Mind studies [17,19], this cognitive process was caused by explicit instructions to identify with the feelings of the main character. In our experiment this process consisted of an inference process going beyond an associative, purely lexico-semantic elaboration.

4.3. Inference processes and the right hemisphere

Given our data, the role of the right hemisphere during inferencing remains unclear. In contrast to the good correspondence to the results of Fletcher et al. [17], and the partial overlap with the results reported by Maguire et al. [33], we could not replicate the involvement of right frontotemporal regions reported by St. George et al. [49] or Robertson et al. [46]. During language processing overall, the right hemispheric homologues of the left sided frontotemporal regions were less strongly activated and they were smaller in extent. Furthermore, there was no lateralization shift in fronto-temporal regions when comparing coherent, plausible trials to incoherent, unrelated sentence pairs.

One explanation for this discrepancy not only with these imaging studies, but also with the neuropsychological literature on inferencing, might be that resources in the non-dominant hemisphere do not become recruited unless the dominant hemisphere is lesioned, or unless unusual or erroneous stimuli are presented. Mazoyer et al. [35] found a reduction of their clear-cut left-hemispheric dominance for processing of syntactic prose and pseudo-word sentences (see also Refs. [38,40]). Following this hypothesis, the absence of right-hemispheric activation in the Maguire et al. study [33], in contrast to the St. George et al. study [49] might have been due to task differences. While there was no experimental task in the latter study — a setting in which the lack of a good situation model becomes apparent — the former study used memorization instructions, thereby drawing attention away from the insufficient situation model to the preserved textbase level. The left-hemispheric dominance found in the present study is also consistent with this explanation. There were no confusing or erroneous stimuli, because incoherence was task relevant. The participants did not get the feeling of being 'at loss', so that an additional recruitment of right hemispheric resources was unnecessary.

4.4. Conclusions

The experiment reported here enabled to identify the brain areas involved in distinct, but interacting processes during language comprehension in context. Consistent with previous results on word and sentence level processes we found extensive left fronto-temporal brain regions involved equally in all conditions. Neuropsychological theories proposing a particular role of the right hemisphere for inference processes could not be supported by the present data. Instead, we confirmed the role of posterior cingulate and adjacent precuneal areas for the successful construction of situation model representations. Most importantly, crossing the factors coherence and cohesion enabled us to attribute different components of language processing to distinct areas in the left prefrontal cortex. Besides the role of the left inferior frontal gyrus for lexico-semantic processes, and the function of the inferior frontal junction area for task set management, we found evidence for the involvement of the left frontomedian wall in the initiation of non-associative inference processes.

References

[1] D. Bavelier, D. Corina, P. Jezzard, S. Padmanabhan, V.P. Clark, A. Karni, A. Prinster, A. Braun, A. Lalwani, J.P. Rauschecker, R. Turner, H. Neville, Sentence reading — a functional MRI study at 4 Tesla, J. Cogn. Neurosci. 9 (1997) 664–686.

[2] M. Beeman, Semantic processing in the right hemisphere may contribute to drawing inferences from discourse, Brain Lang. 44 (1993) 80–120.

[3] M. Beeman, Coarse semantic coding and discourse comprehension, in: M. Beeman, C. Chiarello (Eds.), Right Hemisphere Language Comprehension: Perspectives From Cognitive Neuroscience, Lawrence Erlbaum, Mahwah, NJ, 1998, pp. 255–284.

[4] J.R. Binder, J.A. Frost, T.A. Hammeke, P.S.F. Bellgowan, S.M. Rao, R.W. Cox, Conceptual processing during the conscious resting state: A functional MRI study, J. Cogn. Neurosci. 11 (1999) 80–93.

[5] J.R. Binder, J.A. Frost, T.A. Hammeke, T.W. Cox, S.M. Rao, T. Prieto, Human brain language areas identified by functional magnetic resonance imaging, J. Neurosci. 17 (1997) 353–362.

[6] V. Bosch, Statistical analysis of multi-subject fMRI data: the assessment of focal activations, J. Magn. Reson. Imaging 11 (2000) 61–64.

[7] G. Bottini, R. Corcoran, R. Sterzi, E. Paulesu, P. Schenone, P. Scarpa, R.S.J. Frackowiak, C.D. Frith, The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study, Brain 117 (1994) 1241–1253.

[8] J.D. Bransford, M.K. Johnson, Contextual prerequisites for understanding: some investigations of comprehension and recall, J. Verb. Learn. Verb. Behav. 11 (1972) 717–726.

[9] H.H. Brownell, H. Gardner, P. Prather, G. Martino, Language, communication and the right hemisphere, in: H.S. Kirshner (Ed.), Handbook of Neurological Speech and Language Disorders, Marcel Dekker, New York, 1995, pp. 325–349.

[10] Y. Chantraine, Y. Joanette, D. Cardebat, Impairments of discourse-level representations and processes, in: B. Stemmer, H.A. Whitaker (Eds.), Handbook of Neurolinguistics, Academic Press, San Diego, CA, 1998, pp. 261–274.

[11] M.W.L. Chee, K. M O'Craven, R. Bergida, B.R. Rosen, R.L. Savoy, Auditory and visual word processing studied with fMRI, Hum. Brain Mapp. 7 (1999) 15–28.

[12] C. Chiarello, On codes of meaning and the meaning of codes: Semantic access and retrieval within and between hemispheres, in: M. Beeman, C. Chiarello (Eds.), Right Hemisphere Language Comprehension: Perspectives From Cognitive Neuroscience, Lawrence Erlbaum, Mahwah, NJ, 1998, pp. 141–160.

[13] M. Dapretto, S.Y. Bookheimer, Form and content: Dissociating syntax and semantics in sentence comprehension, Neuron 24 (1999) 427–432.

[14] A. Dove, S. Pollmann, T. Schubert, C. Wiggins, D.Y. von Cramon, Prefrontal cortex activation in task switching: An event-related fMRI study, Cogn. Brain Res. 9 (2000) 103–109.

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

[16] J.A. Fiez, Phonology, semantics, and the role of the left inferior prefrontal cortex, Hum. Brain Mapp. 5 (1997) 79–83.

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

[18] K.J. Friston, Statistical parametric maps in functional imaging: A general linear approach, Hum. Brain Mapp. 2 (1994) 189–210.

[19] H.L. Gallagher, F. Happé, N. Brunswick, P.C. Fletcher, U. Frith, C.D. Frith, Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks, Neuropsychologia 38 (2000) 11–21.

[20] V. Goel, B. Gold, S. Kapur, S. Houle, The seats of reason? An imaging study of deductive and inductive reasoning, Neuroreport 8 (1997) 1305–1310.

[21] A.C. Graesser, M. Singer, T. Trabasso, Constructing inferences during narrative text comprehension, Psychol. Rev. 101 (1994) 371–395.

[22] P. Hagoort, P. Indefrey, C. Brown, H. Herzog, H. Steinmetz, R.J. Seitz, The neural circuitry involved in the reading of German words and pseudowords: A PET study, J. Cogn. Neurosci. 11 (1999) 383–398.

[23] P.I. Indefrey, A. Kleinschmidt, K.-D. Merboldt, G. Krüger, C.

Brown, P. Hagoort, J. Frahm, Equivalent responses to lexical and non-lexical visual stimuli in occipital cortex: A functional magnetic resonance imaging study, Neuroimage 5 (1997) 78–81.

[24] M.A. Just, M. Rep, P.A. Carpenter, J.M. van Dijl, T.A. Keller, K. Suda, W.F. Eddy, G. Schatz et al., Brain activation modulated by sentence comprehension, Science 274 (1996) 114–116.

[25] B.L. J Kaczmarek, Neurolinguistic analysis of verbal utterances in patients with focal lesions of frontal lobes, Brain Lang. 21 (1984) 52–58.

[26] A.M. Kang, R. T Constable, J.C. Gore, S. Avrutin, An event-related fMRI study of implicit phrase-level syntactic and semantic processing, Neuroimage 10 (1999) 555–561.

[27] W. Kintsch, Comprehension: A Paradigm For Cognition, Cambridge University Press, Cambridge, 1998.

[28] W. Kintsch, T.A. van Dijk, Toward a model of text comprehension and production, Psychol. Rev. 85 (1978) 363–394.

[29] J.-H. Lee, M. Garwood, R. Menon, G. Adriany, P. Andersen, C.L. Truwit, K. Ugurbil, High contrast and fast three-dimensional magnetic resonance imaging at high fields, Magn. Reson. Med. 34 (1995) 308.

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

[31] G. Lohmann, Volumetric Image Analysis, Wiley, Chichester, UK, 1998.

[32] G. Lohmann, K. Müller, V. Bosch, LIPSIA: A Software Package For the Evaluation of Functional MRI Data, Technical Report, Max-Planck-Institute of Cognitive Neuroscience, Leipzig, 1999.

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

[34] R.S. Marin, Apathy: a neuropsychiatric syndrome, J. Neuropsychiatry Clin. Neurosci. 3 (1991) 243–254.

[35] B.M. Mazoyer, N. Tzourio, V. Frak, A. Syrota, N. Murayama, O. Levrier, G. Salamon, S. Dehaene, L. Cohen, J. Mehler, The cortical representation of speech, J. Cogn. Neurosci. 5 (1993) 467–479.

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

[37] G. McKoon, R. Ratcliff, Inferences during reading, Psychol. Rev. 99 (1992) 440–466.

[38] M. Meyer, A.D. Friederici, D.Y. von Cramon, Neurocognition of auditory sentence comprehension: Event related fMRI reveals sensitivity to syntactic violations and task demands, Cogn. Brain Res. 9 (2000) 19–33.

[39] T.F. Münte, K. Schiltz, M. Kutas, When temporal terms belie conceptual order, Nature 395 (1998) 71–73.

[40] P. Nichelli, J. Grafman, P. Pietrini, K. Clark, K.Y. Lee, R. Miletich, Where the brain appreciates the moral of a story, Neuroreport 6 (1995) 2309–2313.

[41] D.G. Norris, Reduced power multi-slice MDEFT imaging, J. Magn. Reson. Imaging 11 (2000) 445–451.

[42] O.P. Novoa, A. Ardila, Linguistic abilities in patients with prefrontal damage, Brain Lang. 30 (1987) 206–225.

[43] C.J. Price, 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, Academic Press, San Diego, CA, 1997, pp. 301–328.

[44] G.P. Prigatano, J.R. Roueche, D.J. Fordyce, 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, John Hopkins University Press, Baltimore, MD, 1986.

[45] K.R. Pugh, B.A. Shaywitz, S.E. Shaywitz, R.T. Constable, P. Skudlarski, R.K. Fulbright, R.A. Bronen, D.P. Shankweiler, L. Katz, J.M. Fletcher, J.C. Gore, Cerebral organization of component processes in reading, Brain 119 (1996) 1221–1238.

[46] D.A. Robertson, M.A. Gernsbacher, S.J. Guidotti, R.R.W. Robertson, W. Irwin, B.J. Mock, M.E. Campana, Functional neuroanatomy of the cognitive process of mapping during discourse comprehension, Psychol. Sci. 11 (2000) 255–260.

[47] M. Singer, Discourse inference processes, in: M.A. Gernsbacher (Ed.), Handbook of Psycholinguistics, Academic Press, San Diego, CA, 1994, pp. 479–515.

[48] B. Stemmer, Discourse studies in neurologically impaired populations: A quest for action, Brain Lang. 70 (1999) 402–418.

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

[50] J. Talairach, P. Tournoux, Coplanar Stereotaxic Atlas of the Human Brain, Thieme, New York, 1988.

[51] S.L. Thompson-Schill, M. D'Esposito, I.P. Kan, Effects of repetition and competition on activity in left prefrontal cortex during word generation, Neuron 23 (1999) 513–522.

[52] T. Trabasso, P. van den Broek, Causal thinking and the representation of narrative events, J. Mem. Lang. 24 (1985) 612–630.

[53] N. Tzourio, F. Crivello, E. Mellet, B. Nkanga-Ngila, B. Mazoyer, Functional anatomy of dominance for speech comprehension in left handers vs. right handers, Neuroimage 8 (1998) 1–16.

[54] K. Ugurbil, M. Garwood, J. Ellermann, K. Hendrich, R. Hinke, X. Hu, S.-G. Kim, R. Menon, H. Merkle, S. Ogawa, R. Salmi, Imaging at high magnetic fields: Initial experiences at 4T, Magn. Reson. Q. 9 (1993) 259.

[55] T.A. van Dijk, W. Kintsch, Strategies of Discourse Comprehension, Academic Press, New York, 1983.

[56] R. Wise, F. Chollet, U. Hadar, K. Friston, E. Hoffner, R. Frackowiak, Distribution of cortical neural networks involved in word comprehension and word retrieval, Brain 114 (1991) 1803–1817.

[57] S. Zysset, K. Müller, G. Lohmann, D.Y. von Cramon, Color word matching Stroop task: Separating interference and response conflict, Neuroimage 13 (2001) 29–36.

Aging, Neuropsychology, and Cognition, 13:62–85, 2006
Copyright © Taylor & Francis Group, LLC
ISSN: 1382-5585/05 print; 1744-4128 online
DOI: 10.1080/13825580490904237

Psychology Press
Taylor & Francis Group

Text Comprehension in Middle Aged Adults: Is There Anything Wrong?

EVELYN C. FERSTL
Max-Planck Institute of Human Cognitive and Brain Sciences,
Leipzig Day Clinic of Cognitive Neurology, University of Leipzig

ABSTRACT

Two experiments tested text comprehension skills in adults varying in age from 20 to 69 years. A coherence judgment task (n = 39), in which the pragmatic connection of two sentences had to be evaluated, did not yield any age effects. Both error rates and reaction times were largely independent of age. In a word recognition paradigm (n = 60), testing the strength of different text representation levels, there was a gradual, linear decrease in performance with age, accompanied by a similarly gradual increase in processing times. The results confirm that aging affects comprehension under high memory demands, but spares the conceptual, situation-based levels. Most importantly, it was shown that these changes become apparent during middle age already. Implications for education and neuropsychology are discussed.

Language comprehension in everyday life requires the interplay of a vast number of cognitive processes. Lexical, semantic, and syntactic processes on the word and sentence levels interact with structure-building and integration of general world knowledge on the text level—to name just a few examples for possible component players. Despite this complexity, language processes are so highly overlearned that—at least in standard situations—comprehension does not seem to require much conscious effort at all.

The empirical findings on text comprehension skills in aging populations mirror these seemingly contradictory observations. While there is no doubt that, for instance, working memory limitations in older comprehenders should adversely affect complex language processes, it seems that the potential deficits can be counteracted by a more efficient utilization of context information and background knowledge.

Address correspondence to: Evelyn C. Ferstl, Max-Planck Institute of Human Cognition and Brain Sciences, Stephanstr. 1 A, 04103 Leipzig, Germany. E-mail: ferstl@cbs.mpg.de

Psycholinguistic theory provides us with the conceptual framework for describing this in more detail. According to Kintsch (1998, 1988) and van Dijk & Kintsch (1983), three different levels of representations are built during text comprehension. The *surface level* encodes the exact wording of the utterances, and the *textbase* level represents the meaning, i.e., the semantic or propositional content. The *situation model* consists of an integration of the text information with the comprehender's background knowledge and encodes the overall gist of the text, i.e., "what the text is about" (Ferstl 2001; Ferstl & Kintsch, 1999; Graesser, et al., 1997; Zwaan et al., 1995; Zwaan & Radvansky, 1998). A very important subprocess needed for forming both textbase and situational representations is *inferencing* (Singer, 1994; Graesser, et al., 1994), i.e., using general world knowledge for appropriately supplementing the explicitly mentioned information. The classical example, "Annie invited Charly to her birthday party. He asked his mother for some money," illustrates how inferences are used for bridging coherence gaps between subsequent sentences and for establishing a causal structure in narratives. The reader is likely to infer that Charly wants to buy a present for Annie. Without the knowledge that it is usual to bring gifts to birthday parties, the sentences remain unrelated or *incoherent*. The situation model for this short text might also contain cakes, candles, and wrapping paper, i.e., pieces of information resulting from elaborative or predictive inferences (Singer, 1994).

In the literature on cognitive aging, language skills are considered relatively resistant to decline (Kliegl et al., 1999). If text comprehension deficits arise, they are consistently located on the surface and textbase levels. Age effects are most likely when the task is dependent on the verbatim form or when it relies crucially on working memory. On the other hand, it seems that older comprehenders show qualitatively similar patterns as younger ones when inference processes and situation model-building are investigated. Age effects are less likely when the task is dependent on conceptual integration using world knowledge and when sufficient processing time is available (for reviews, see Stine-Morrow & Miller (1999), Wingfield & Stine-Morrow (2000), and Zacks et al., (2000).

An open question is *when* these differences in text comprehension processes appear during the life span. Do they develop gradually with increasing age? Most aging studies compare a younger group, usually in their twenties, to a group of senior citizens in their sixties and older. However, there is evidence for age-related changes in semantic and syntactic processes long before retirement age (Gunter et al., 1992, 1998, 1999; Obler et al., 1991). Text comprehension studies including middle-aged adults are rare (e.g., Hess, 1995; van der Linden et al., 1999). And if a middle-aged group is included, it is sometimes considered the control group to which the elderly are compared, rather than being of interest in its own right (e.g., Ulatowska et al., 1986).

In the present study, we were interested in text comprehension processes across the adult life span. Specifically, the goal was the evaluation of two text comprehension tasks developed for research in clinical neuropsychology (cf. Ferstl et al., 1999, 2002; Ferstl & von Cramon, 2001, 2002) with respect to age-related changes. Participants in the age range of 20 years to 69 years performed two tasks tapping qualitatively different component processes. In the first experiment, we studied inference processes using a coherence judgment task, known to be dependent on left frontal lobe functions (Ferstl et al., 2002; Ferstl & von Cramon, 2001, 2002). The second experiment used a text recognition paradigm for dissociating different levels of text representation (cf. Beyer et al., 1996; Kintsch et al., 1990; Radvansky et al., 2001; Schmalhofer & Glavanov, 1986).

The questions of interest for both experiments were a) whether we could observe any age-related decline in performance, b) if so, whether this decline would be gradual across the age range or associated with a specific age group, and c) whether differential age effects would map on to dissociable subprocesses of text comprehension.

EXPERIMENT 1: COHERENCE BUILDING

Inference processes are a key component of text comprehension. They are required for establishing coherence, elaboration, prediction, and causal attribution (Graesser, et al., 1994; Singer, 1994; van den Broek, 1990). In general, on-line inference generation seems intact in older adults (see Valencia-Laver & Light (2000) for a recent review). Age differences arise mainly when the processing time is limited, when a revision is needed (Hamm & Hasher, 1992) or when memory demands are high. For instance, Hess (1995) found that older comprehenders were as sensitive to causal coherence as younger comprehenders, indicated by reading times to sentences varying in their causal relatedness. However, older participants did not spontaneously encode them sufficiently for the inferences to become effective retrieval cues during a subsequent memory test.

On a more general level, it has been argued that age-related decline of cognitive function is due to changes in frontal lobe function. For example, reductions in cortical volume and in glucose metabolism are most pronounced in the frontal lobes (Raz, 2000). While the evidence on corresponding age-related decline of executive functions is mixed (Boone, 1999), neuropsychological studies of inference processes confirm that the frontal lobes are crucial for language processing in context. Using a coherence judgment task, Ferstl and von Cramon (2001) studied the interaction between pragmatic coherence and lexical cohesion during inferencing. *Pragmatic coherence* refers to a connection between succesive sentences based on the sentence content. *Lexical cohesion*, in contrast, is present when a word, such as a conjunction (e.g., “therefore,” “before”) or a pronoun (e.g., “his”), explicitly signals a connection between the sentences. Functional MRI showed the involvement of lateral prefrontal cortex, particularly when cohesion erroneously suggested a connection between pragmatically incoherent sentences (Ferstl & von Cramon, 2001). In addition, a region in the left medial prefontal cortex was more active when an inference was successful, i.e., when a coherent sentence pair was read. Confirming the imaging results, Ferstl et al. (2002) reported deficits in coherence building for patients with left-frontal lesions.

The first experiment set out to test whether the same coherence judgment task would uncover any age-related differences. We tested a sample of participants ranging from 26 to 64 years of age. If inference processes were intact even in older participants, no age-related differences would be predicted for the error rates. However, the reaction times were expected to increase with age, either due to general cognitive slowing or to subtle differences in the cognitive processes involved in coherence building. Furthermore, if the functions attributed to the frontal lobes were already slightly affected in the older participants, they should rely more on lexical cohesion as a cue to coherence.

Methods

Participants

Thirty-nine adults volunteered for the experiment. The 19 women and 20 men received a small reimbursement for their expenses. The participants’ age varied from 26 years to 64 years. The age levels were uniformly distributed across the entire range (*M* = 45 years, *sd* = 11.9). Educational level was coded in two categories—college preparatory school, middle school—and occupation in three levels—academic/professional, clerical, worker/craftsmen. The mean ages of the two school levels were 43 years (n = 17; range 30–64, sd = 13.5) and 46 years (n = 22; range 26–64; sd = 10.6), respectively, and for the three occupational levels 42 (n = 17; range 26–64; sd = 13.6), 45 (n = 12; range 30–61; sd = 10.1) and 49 (n = 10; range 30–64; sd = 10.7) years. There was a slight tendency for the younger participants to be better educated, but this effect was not statistically significant (r = –.23 for occupation, r = –.14 for schooling).

Materials

One-hundred-twenty sentence pairs were written so that the first sentence (the context) was pragmatically related to the second sentence (the target). All of the syntactically simple sentences were about well-known everyday situations. By switching the context sentences of item pairs, an incoherent condition was created. For varying the lexical cohesion, each target was written in two versions: in the cohesive version a lexical item (e.g., a pronoun or a conjunction) signalled the pragmatic connection, whereas in the incohesive version, this lexical marker was omitted. Cohesion was fully crossed with coherence, so that four conditions resulted. Examples for these sentence stimuli, which were described in more detail elsewhere (Ferstl et al., 2002; Ferstl & von Cramon, 2001, 2002), are presented in Table 1.

All sentences were spoken by a female research assistant and digitally recorded using the software package CoolEdit. Because all context sentences were recorded in one session, and the target sentences in another session, there were no prosodic cues as to the coherence of the sentences. The presentation length was 2.5 s (range 1.3–4.6 s) for the context, and 2.8 s (range 1.6–4.0) for the target sentences. Cohesive targets were longer than incohesive targets (2.9 vs. 2.6 seconds, $t\ (119) = 15.5$, $p < .001$).

The 480 sentence pairs were put together in four counterbalancing lists of 120 trials each. Each list contained 30 trials of each condition, and each version of a particular sentence appeared in exactly one of the lists. The resulting lists were then subdivided into three blocks each. In each block of 40 trials, an additional 10 filler trials of pseudo-words were inserted, so that the procedure was equivalent to that of a previous study (cf. Ferstl & von Cramon, 2002). These trials are not of interest for the present study. In addition, one buffer trial at the beginning and at the end of each block was used. Thus, each block contained 52 trials, for a total of 156 trials per list. The order of the trials was pseudo-random within each block, so that not more than four adjacent trials required the same answer.

Table 1. Example materials for the four conditions of Experiment 1 (translated from the original German)

	Coherent	Incoherent
Cohesive	1. Mary's exam was about to start. *Therefore*, *her* palms were sweaty. Laura got a lot of mail today. *Her* friends had remembered *her* birthday.	1. Laura got a lot of mail today. *Therefore*, *her* palms were sweaty. Mary's exam was about to start. *Her* friends had remembered *her* birthday.
	2. 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 why* the car doesn't start.	2. The lights have been on since last night. *That's when* the dishes start to rattle. Sometimes a truck drives by the house. *That's why* the car doesn't start.
Incohesive	3. Mary's exam was about to start. The palms were sweaty. Laura got a lot of mail today. Some friends had remembered the birthday.	3. Laura got a lot of mail today. The palms were sweaty. Mary's exam was about to start. Some friends had remembered the birthday.
	4. Sometimes a truck drives by the house. The dishes start to rattle. The lights have been on since last night. The car doesn't start.	4. The lights have been on since last night. The dishes start to rattle. Sometimes a truck drives by the house. The car doesn't start.

Cohesive ties are printed in italics (from Ferstl & von Cramon, 2001).

Procedure

At arrival, subjects were randomly assigned one of the four counterbalancing versions and one of the keys on the response box as the YES key. Care was taken that these assignments were unconfounded with age. The participants received the instructions in writing. They stated that sentences would be played and that the task was to decide whether the sentences had "something to do with each other." Examples were given for coherent and incoherent sentence pairs. For a trial consisting of nonwords, the participants were instructed to always use the NO key.

After the experimenter had ensured that the instructions were understood, the experiment started. The sequence of trials was self-paced. Participants pressed the middle button on the response box whenever they were ready for the next trial. A trial started with the presentation of a fixation cross on the screen, accompanied by a signal tone. After 1000 ms, the auditory presentation of the context sentence started. The target sentence was played after an interstimulus interval of 500 ms. Simultaneous with the target presentation, a question mark appeared on the screen, indicating that a response could be made anytime. At the button press indicating the coherence judgment, the computer recorded the response key and the response time, measured from the onset of the target sentence.

The duration of the experiment was about 30 min, so that the entire session—including instructions and breaks—lasted up to 45 min.

Results

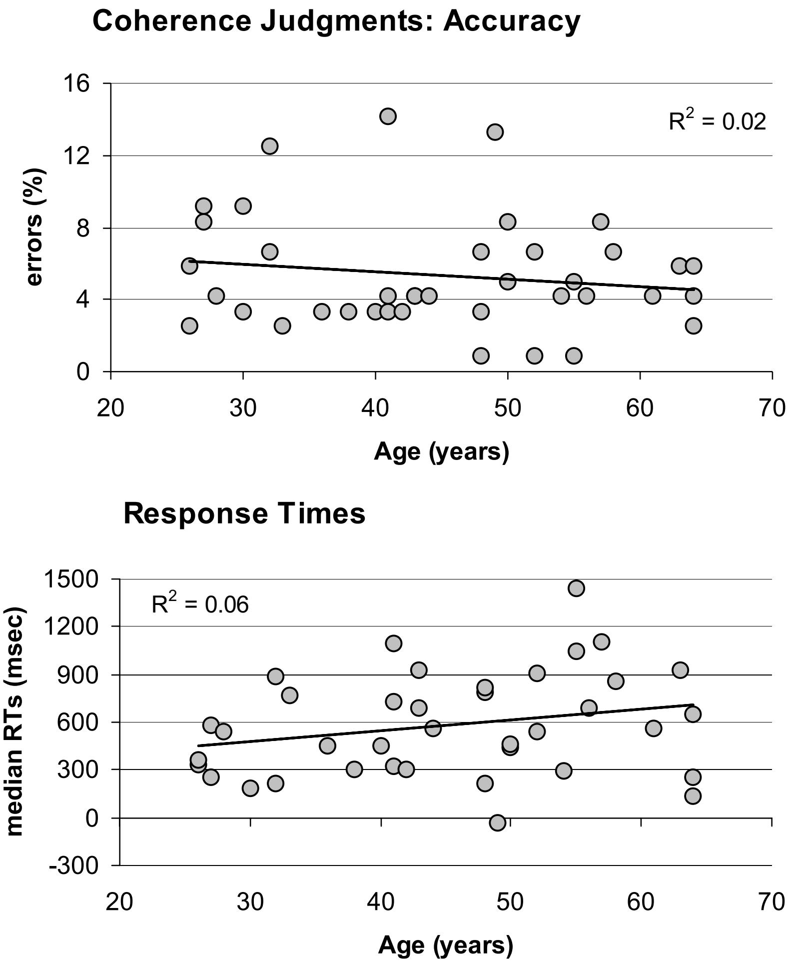
All analyses were based on the General Linear Model (GLM) including coherence and cohesion as within-subjects variables, and the continuous between-subjects variable age (cf. Judd & McClelland, 1989).

Errors

The overall performance level was excellent. On average, participants made 5.3 % errors (sd = 3.2). The error scores varied from .8% to 14.2%, but they did not systematically differ with age ($F(1,37) < 1$). To illustrate that this lack of an age effect was not just due to the variance, Figure 1 presents a scatterplot of the individual data. As can be seen, only three participants made more than 10% errors, and they were in the younger and middle age range.

The distribution of errors across the four conditions is shown in Figure 2. Whereas the main effects of coherence and cohesion did not even approach significance (F's < 1), the expected interaction was reliable ($F(1,37) = 5.5$, $p < .05$). For coherent trials, the cohesive marker facilitated the judgment, whereas for incoherent trials, the mismatch between the lexical and pragmatic information led to a very slight increase in error rates (cf. Ferstl & von Cramon, 2001). There was no interaction of the within-subjects variables with age (F's < 1).

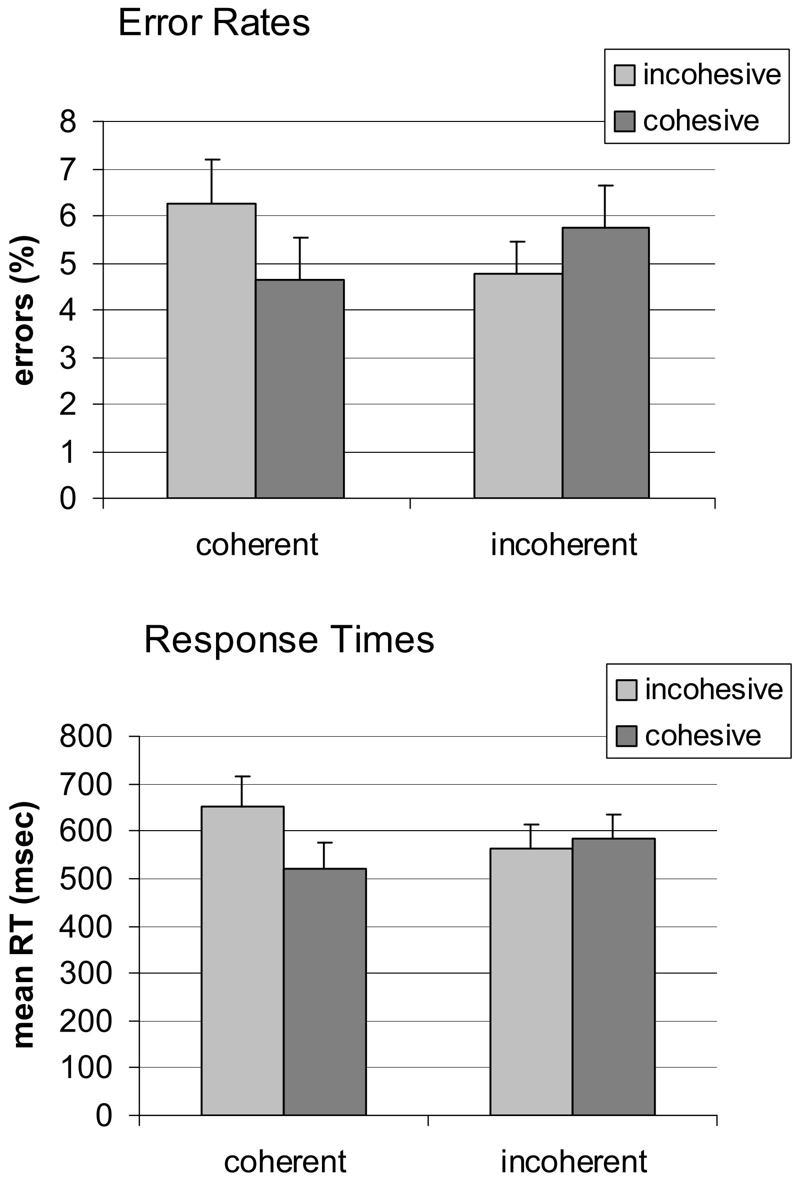
FIGURE 1. Individual data for the coherence judgment task. The upper panel shows a scatterplot of the error rates as a function of age. The lower panel shows the response times. The trend for an increase in response times was not significant (p > .13).



Response times

For the analysis of the response times, only correct trials were included. Although in some cases the coherence might have been detectable before the final word of the sentence, only a small proportion of the responses were given before the end of the sentence presentation. Inspection of the data showed that only one participant (age 30) consistently responded

FIGURE 2. Performance in the coherence judgment task as a function of coherence and cohesion of the target sentence. The upper panel shows the error rates, the lower panel the response times. The error bars indicate one standard error above the mean.



long before the end of the target sentence. The studentized residual from a regression of the overall response time on age confirmed that this observation was an outlier (stud-r < –3) (Judd & McClelland, 1989); in the subsequent analyses this participant's data was excluded.

For the remaining 38 participants, only 6.3% of the responses were given before the end of the target sentence presentation. The proportion of responses provided before sentence end was not correlated with age ($r = -.06$, $p = .7$). This result indicates that there were no age-related strategy differences in whether the participants waited with their judgment until the end of the sentence. To control for the variable presentation length and to obtain values that are more easily interpretable, the times from the offset of the target sentence were taken as the dependent variable. To obtain these values, the length of the sentence presentation time was subtracted from the recorded response time.

The median response time for each participant and each of the four conditions was taken as the basis for further analyses. The median was chosen because it is less affected by individual outliers, and therefore captures the central tendency of skewed reaction time data better than the mean (see Ratcliff, 1993).

Across the group of 38 participants, the mean response time was 580 ms (sd = 324), ranging from –31 ms to 1445 ms. The individuals' response times are also plotted in Figure 1 as a function of age. The very slight tendency for the oldest participants taking longer to respond was not statistically significant ($F(1,36) = 2.3$, $p > .13$).

The overwhelming proportion of the variance was accounted for by the factor cohesion and its interaction with coherence, as shown in the lower panel of Figure 2. The cohesive sentences were responded to faster ($F(1,36) = 10.0$; $p < .01$), and this effect was larger for the coherent trials ($F(1,36) = 40.5$; $p < .0001$; for the interaction cohesion × coherence). These strong effects were independent of the participant's age (F's < 1 for the interactions age × cohesion and age × cohesion × coherence).

There was no overall effect of coherence, but a significant interaction between age and coherence ($F(1,36) = 4.5$, $p < .05$). Whereas responses to incoherent trials did not vary with age ($r = .15$, ns), there was a trend for the responses to coherent trials to be longer for older participants ($r = .31$; $p = .06$). A few of the older participants tended to take especially long for establishing coherence, as compared to detecting coherence gaps. There was no such difference for younger participants.

Discussion

The accuracy data showed no age-related differences whatsoever. Only three participants made more than 10% errors, but they were in the young-to-middle age range. All participants could very easily distinguish the coherent and incoherent trials. This result is consistent with previous findings of mostly preserved inference processes in older people (cf. Valencia-Laver & Light, 2000). The presentation of two syntactically simple sentences did not put high demands on working memory, and the off-line coherence judgment task at the end of the target sentence gave sufficient time for integrating the text information with background knowledge.

Despite the low error rates, the expected interaction between coherence and cohesion was significant. Across the entire group, cohesion facilitated the inference process and rendered the detection of coherence gaps more difficult. This effect did not vary with age.

The pattern of the response times was very similar to the performance data. The clear influence of the factor cohesion on the ease of coherence building was confirmed, but—once more—these effects were independent of the participant's age. The interaction between age and coherence suggested that the coherent condition was more susceptible to age-related difficulties than the incoherent condition. In the previous patient study (Ferstl et al., 2002), it was the coherent condition as well that proved sensitive to frontal lobe damage. Thus, despite the lack of age effects on performance, the reading time patterns are at least suggestive of a connection between frontal lobe function and age-related decline in coherence building.

One explanation for the lack of age effects on performance and overall response times is that the age range of our sample of participants was not sufficently large. It is possible that a group of elderly (over 65 years of age) might have exhibited more difficulties with the coherence judgment task. An alternative explanation draws on the levels of text representation utilized. The coherence judgment task can be solved either by drawing a specific, text-based bridging inference, or by evalutating the plausibility of the situation model. This latter level is known to be preserved even in old age.

The second experiment was designed as a more focused test of text comprehension in middle age. In order to increase the likelihood of individual differences, the memory demands were elevated. Furthermore, the experiment included a direct test of the surface and text base levels. The recognition paradigm chosen for Experiment 2 differed from the coherence judgment paradigm of Experiment 1 because we wanted to take advantage of a well-established method sensitive to the variables of interest

EXPERIMENT 2: TEXT REPRESENTATION

The second experiment employed a recognition task testing the formation of different levels of text representation (cf. Kintsch et al., 1990; Schmalhofer & Glavanov, 1986). In contrast to the coherence judgment task of Experiment 1, this paradigm has been shown to be sensitive to age-related differences. Using sentence recognition, Radvansky et al. (2001) demonstrated for an elderly group (mean 74 years) that their response criteria was more liberal and that the discrimination performance for surface and textbase probes was lower than that of a younger control group.

We used a word recognition task to study these differential processes in adults of various ages, including middle-aged participants. Word recognition—instead of sentence recognition—was chosen because the comprehension demands at the time of test are lower. Thus, the recognition performance more directly reflects the encoding of the text information rather than integration processes at test (Beyer et al., 1996).

A story about two children skipping school and walking through the boy's home was presented visually. Memory for the narrative was indexed by the discrimination ability for text words against two different distractor types. The distractors had different levels of relation to the text (cf. Radvansky et al., 2001) Synonyms, or highly related words, were chosen as being very close in meaning to an explicitly mentioned word. Discrimination ability for synonyms requires the accurate representation of the surface structure of the text, i.e., of its exact wording (cf. Kintsch, 1988). Related words were chosen as being consistent with an elaborated text base or with the situation model of the text, i.e., these words are part of a text representation augmented by the comprehender's inferences.

Based on the previous studies on aging we expected the oldest participants in our sample to show a deficit with respect to surface level processes. In particular, the high memory demands were expected to lead to a decline in performance with increasing age. The question of interest was whether the performance during the middle age bracket would exhibit a qualitatively comparable, but smaller decline, or whether middle-aged participants would perform in a similar way as the younger participants.

Method

Materials

For the encoding task a story from a previous study was used (see Ferstl et al., 1999). The story was about two children skipping school and walking through the boy's house. The story structure is sequential, containing descriptions of the various rooms. The story was typed on three pages for presentation on three successive computer screens. The length of the story was 505 words.

From the story 48 "old" words were selected for the recognition task. They were evenly distributed across the length of the story. For the distractors, 24 words were selected as words that were very closely related to the surface structure of the text. Most of these words were synonyms of words mentioned in the text or were related via one associative link. Examples are German equivalents of *mom* for *mother* (*Mama – Mutter*), or *musty* for *moldy* (*modrig – muffig*). In a Thesaurus, there were entries for 22 of the 24 word pairs, with the remaining two words having no entries due to their low frequency (e.g., *Kurpark*). For some of the words, a superordinate was chosen for checking the relationship (e.g., *Eingangshalle – Halle*; *entry hall – hall*). These closely related words were used to test the surface representation. Another 24 words were chosen to be more distantly related to the story content. These distractors were needed to be inferred by the reader and were part of the story's elaborated textbase and/or situation model levels. Examples are *lawn* (*Wiese*) or *fence* (*Zaun*), both of which were not mentioned at all, but were likely to be part of the fictional garden described. Finally, eight words were selected to be entirely unrelated to the text, so that minimal comprehension could be ensured.

The mean word length was 8.4 letters, similar for old (range 4–15) and distractors (range 4–14). The occurrence frequency of the words was matched by using the CELEX database (Baayen et al., 1995). The median frequency (per million) of the targets was 9 (range 0–378), that of the distractors 4 (0–330). Table 2 contains an excerpt of the story, translated into English, together with examples for the recognition probe types.

The list of 104 words was first split into three blocks according to their position in the story (beginning, middle, end). Each of the three parts was then pseudo-randomized separately so that not more than three successive words were old or new, respectively. After the randomization, the beginning, middle, and end parts were combined in that order. Two differently ordered lists were created in this manner.

TABLE 2. Translated excerpt of the story used in Experiment 2. The words printed in italics are words used for recognition probes, either as targets, or as the basis for the synonym distractors. The examples for the probe types are illustrative only. In the original German version, the particular words were slightly different.

Text excerpt:

… She started walking towards the door, but Peter said: "No, come with me, I'll show you my *secret passage*!" He led her to the backside of the house. There Peter explained to Mary that he would always leave a window ajar, because he didn't like to carry a key. Both of the children squeezed through the opening. With a *jump* they landed on the humid ground. "It's always wet in here," Peter said, "particularly after a thunderstorm, and it smells all *moldy*." Then they climbed the staircase to the ground level. "This is like a fairy tale!" Mary thought when she entered the spacious lobby with at least seven doors leading into other rooms. Peter asked her to pick a door she wanted to open, and she selected the one in the middle. "This is my mother's *study*," Peter said while opening the door. By the window there was a desk, with a computer and a telephone on it. "My mother thinks I don't know what's in here, but I do!". He opened the first drawer and took out a small key….

Recognition probes:

Target words:	secret passage, moldy
Synonyms distractors	leap, office
Related distractors:	basement, rain
Unrelated distractors:	cigarettes, island

Procedure

The participants were seated in front of a computer and instructed that they would be tested on their memory for a story. They were told to read the story carefully to comprehend its meaning. They then proceeded at their own pace. After the story was presented on three successive screens, the recognition task started. The participants were told to decide after each word whether it had occurred in the text. Examples were given to illustrate that it was important to pay attention to whether the word had been presented in exactly the same form. Six practice trials with feedback followed. The recognition items were then presented one at a time in the middle of the screen until the participant pressed the key for indicating the YES/NO response. Response key and response times for the recognition items were recorded.

Participants

Thirty-one male and 29 female volunteers participated in the experiment, which was part of a larger study. They were paid a small reimbursement for their expenses. The sample of participants was different from that of Experiment 1, but selected from the same subject pool. The participants' ages were equally distributed across the entire age range from 20 to 69 years, with a mean of 44.6 years (sd = 14.8). Women and men were of the same age (mean(men) = 43.8; mean(women) = 45.4; $t(58) < 1$). The education level was also balanced for the different ages ($r = -.20$; $p > .12$, for the correlation between years of education and age). In addition, the participants' verbal short-term memory had been assessed using a digit span measure (cf. Wechsler, 1987). The number of correct trials, rather than the digit span, were used for a more differentiated range of scores. The short-term memory scores were negatively related to age ($r = -.31$; $p < .05$), indicating that—as expected—younger participants had a larger short-term memory capacity. In addition, a working memory measure taking into account executive components was available. The digit ordering test (DOT) requires participants to repeat a series of randomly presented digits in increasing order (Werheid et al., 2002). Once more, the number of correct trials was used. The DOT scores were neither closely related to age ($r = -.15$), nor to the digit span scores ($r = .15$). Thus, the two tests tapped different aspects of working memory.

Results

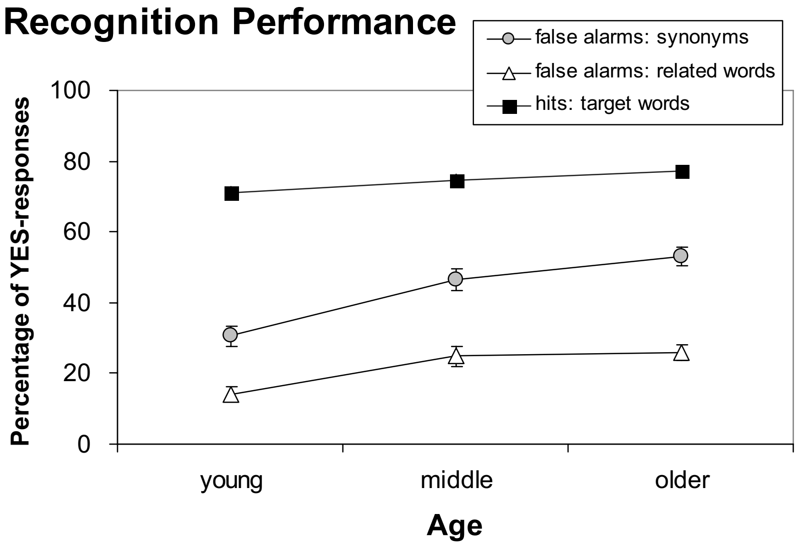
All analyses were based on the General Linear Model (GLM), including the within-subjects variable word type (target/synonym distractor/related distractor) and the continuous between-subjects variable age. For post-hoc comparisons as well as for the diagrams, the participants were grouped into three subgroups of equal size—young/middle/older (see section on post-hoc comparisons for a detailed description).

Performance

The unrelated words elicited only four errors in the entire sample of participants. Thus, these distractors are not included in the further analyses. The proportion of YES responses for targets and the two distractor types are shown in Figure 3 as a function of age. The hits were positively related to age. The older the participants, the better their performance on words mentioned in the text ($r = .41$, $p < .01$). The false alarms, averaged across distractor types, were also highly correlated with age ($r = -.63$, $p < .0001$). The older the participants, the fewer distractors they correctly identified as new.

Both of these results taken together indicate that younger participants were more conservative in their responses, i.e., that they were less likely to respond YES. To take into account these differences in response bias, a signal detection analysis was carried out (Swets, 1973). First, bias scores based on the overall performance were calculated for each participant. The larger the bias, the more conservative the decision criterion, i.e., the less likely are YES responses. The bias scores were negatively related to age, i.e., they confirmed that with age the decision criterion became more liberal ($r = -.65$; $p < .0001$).

Figure 3. The proportion of YES responses for targets, synonym, and related distractors in the recognition task. The error bars indicate one standard error above and below the mean. The data are displayed for three subgroups (see text). All analyses were based on the continuous variable age.



Second, d’ values were used as a measure of discrimination ability. The discrimination of related words was clearly better (M = 1.52, sd = .46) than that of synonyms (M = .85, sd = .41), reflected in a highly significant main effect of distractor type (F(1,58) = 211.5; p < .0001). Discrimination ability diminished with age (F(1,58) = 12.9; p < .001), but it did so equally for both distractor types (F(1,58) = 1.6; p > .2 for the interaction).

As expected, memory capacity contributed to the recognition performance. There was a strong positive correlation between digit span scores and overall d’ (r = .51; p < .0001). However—as pointed out in the method section—age and digit span were correlated as well. For an assessment of the relative contributions of memory and age, we calculated partial correlation coefficients. When partialling out digit span, the correlation of d’ for the synonyms was still highly dependent on age (r = –.40; p < .01). In contrast, the partial correlation between the d’ for related words and age failed to reach significance (r = –.20; p = .13). Thus, even when working memory was controlled, discrimination ability decreased with age, and this effect was more evident for the synonym distractors than for the related distractors.

Correlations of the discrimination scores with the working memory measure yielded a tendency for a relationship with the discrimination performance (r = .25, p = .05), which was also mostly due to the performance on synonyms (r = .25, p = .06) rather than that on related distractors (r = .17, p > .19).

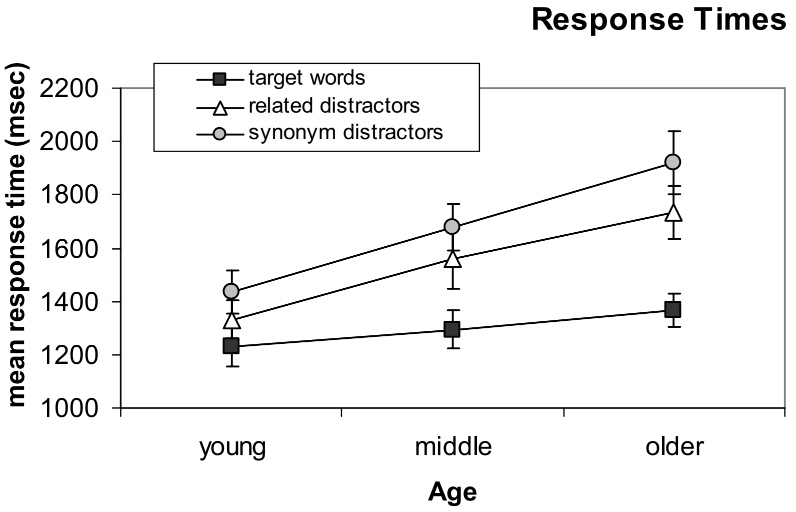
Response Times

The response times to unrelated distractors were not analyzed because of the small number of instances. Due to a technical error, reaction times from two participants were not available (ages 39 and 58). One participant’s response times (age = 38; RT = 3223 ms) were so far above the group mean (M = 1407 ms, sd = 389; range 872–2050 ms), that his data was excluded as well. The studentized residual (stud-r = 6.6) from the regression of overall response time on age confirmed that this observation was an outlier.

For the remaining 57 participants, the median response times for each of three conditions (hits, synonyms, and related distractors) were taken as the dependent variables in an ANCOVA. The factor word type was coded in two contrasts: the first analysis compared the YES responses (i.e., the hits) to the NO responses (i.e., correct rejections overall). The second contrast compared reaction times for correct responses to the two distractor types—synonym and related. In both analyses, age was entered as a continuous between-subjects variable.

The means are shown in Figure 4 as a function of age and word type. The response times increased with age, and targets were accepted faster than distractors were correctly rejected. Correspondingly, there was a clear effect of response (F(1, 56) = 73.6, p < .0001), as well as a reliable main effect of

FIGURE 4. Reaction times for correct responses (i.e., YES responses for old words and NO responses for distractors), as a function of age and word type. The error bars indicate one standard error above and below the mean. Once more, the analyses were based on the continuous variable age. The subgroups were combined to make the graph more easily readable.



age (F(1,55) = 14.1, $p < .001$). In addition, the difference between YES and NO responses became larger with increasing age, resulting in a significant interaction between response and age (F(1,55) = 16.6, $p < .0001$).

When separating the response times for correctly rejecting synonyms from those for rejecting related words, both main effects were again reliable. Older participants took longer to respond (F(1,55) = 22.2, $p < .0001$) and related words were rejected faster than synonyms (F(1,55) = 10.1; $p < .001$). However, there was no interaction between age and distractor type (F(1,55) < 1), indicating that the distractor effect was similar in magnitude for all age levels.

In contrast to the discrimination performance, the response times did not vary with verbal short-term memory. For none of the three word types was there a significant correlation of the response times with the digit span score or the DOT score (p's > .19). Neither was there a relationship between the response times and discrimination ability d' ($r = .06$; ns). However, the estimate of the response bias converged with the response times. Participants with a more liberal criterion took longer for correctly rejecting distractors than participants with a more conservative criterion ($r = -.41$, $p < .01$).

Post-Hoc Comparisons

To directly answer the question of whether the participants in the middle age range behaved qualitatively similar as the younger or as the older ones, the participants were split into three groups of 20 participants each (young: 20–37, M = 27, std = 5.9; middle: 38–53, M = 46, std = 5.0; older: 55–69, M = 61, std = 4.3). Note that for the response time analyses, two participants in the middle age group and one in the older were excluded. Bonferroni tests were calculated as post-hoc comparisons with a corrected alpha-level of .017. The performance variables considered were the hitrate, the correct rejections overall, and the d'-measures for both distractors. The reaction time variables were the response times to targets, related distractors, and synonyms, respectively. Pairwise comparisons yielded significant differences between the young and the older group for all measures ($|t|$'s > 2.52, $p < .017$), except for the response times to targets ($t(37) = -1.4$). In contrast, in the comparison of the middle and the older groups, there was no significant effect at all, neither for the performance measures, nor for the response times ($|t|$'s < 1.7; $p > .11$). Finally, the comparison of the young and middle age groups showed that the proportion of correct rejections was significantly higher for the young group ($t(38) = 3.8$, $p < .001$), as was the d' for synonyms ($t(38) = 2.8$, $p < .01$). No other variable reached the significance level (t's < 2.2, $p > .04$). These differences between the youngest and the middle group confirm that the decline in performance started in middle age already.

Discussion

The results of the recognition task differed greatly from those of the coherence judgment task. Performance measures as well as response times were sensitive to age-related differences in processing. Older participants applied more liberal response criteria, and they correctly recognized more items from the text. When taking into account these bias differences, the discrimination performance proved to gradually decline with age, and this effect was comparable for both synonym and related distractors.

The response times, mirroring the performance data, increased with age, and this effect was more pronounced on the distractors than on the targets. In addition, there was an expected correlation between verbal working memory and the recognition task. However, the age effects remained highly reliable even when the contribution of verbal working memory was factored out.

These results are in correspondence with those reported by Radvansky et al. (2001). In their study, a group of elderly participants was more liberal in their response criteria and their discrimination ability for surface and text-base recognition probes was lower than that of younger comprehenders. The synonym distractors used in the present experiment also tested the surface level of representation and clear age-related decline was found. The related distractors, in contrast, do not as readily allow the separation of the elaborated text base, i.e., a text base including inferences and the situation model representation. Finally, the hit rates—in this experiment equivalent to the situation model level (because of the low number of errors on the unrelated distractors)—also confirmed that on the situation model level the older participants' performance even increased.

The novel result of the experiment presented here is that the age effects appeared gradually, across the entire age range. Even for participants as "young" as in their forties a decline in recognition performance was apparent, and these differences could be confirmed in pairwise group comparisons.

GENERAL DISCUSSION

Two experiments were conducted to test text comprehension processes across the adult life span. The different tasks and different materials uncovered dissociable age-related changes in comprehension. While the coherence judgment accuracy was unaffected by age, except for slight effects on response times, the recognition task yielded a gradual decline. The most important result is that these age effects were observable in middle-aged participants long before retirement age.

The results converge with those of previous studies on text comprehension processes showing that older people rely more on the situation model representation than on surface and textbase representations (Radvansky & Curiel, 1998; Radvansky et al., 1990, 2001; Stine-Morrow & Miller, 1999) and that their performance is comparable or even superior, when processing is conceptually based (Hess & Flannagan, 1992). One interpretation of these findings is that they are caused by differential allocation of resources to the text representation levels. Older adults might focus more on the situation model level to optimally allocate their somewhat limited resources (Wingfield & Stine-Morrow, 2000; Baltes & Baltes, 1990).

In the coherence judgment task, the plausbility of the situation model suffices for making the decision about a pragmatic connection. This observation does not imply that the surface structure was neglected during situation model building. The additional factor cohesion, which was included for a test of surface level features, was used by adults of all age groups in a similar way. There was no evidence that older participants relied less—or more—on cohesive ties than younger participants.

In the recognition task, the use of two distractor types allowed for a direct test of the different representation levels. Performance on the situation model level, indicated by the discrimination ability for unrelated distractors, and in our case equivalent to the hitrate, became even slightly better with age. In contrast, both surface level and the elaborated text base representations, were equally compromised in older comprehenders. These results were mirrored in the response times. In addition, the bias measures suggested that the older participants slightly changed their response criterion to a more liberal one, a finding that is also consistent with a decision based on the situation model rather than on the lower representation levels.

Recognition memory—usually tested in the context of word list learning paradigms—is considered easier than recall and therefore less susceptible to age-related decline. In a review on memory functions in middle-aged adults, Lavigne & Finley (1990) concluded that recognition for visual stimuli is affected by age, with a steady increase in false alarm rates, but that recognition of verbal material does not show a corresponding age effect. The present results suggest that recognition for words shows similar age-related effects if the encoding is meaning-based rather than surface level-based. This result is consistent with the finding that older adults rely more on gist-based memory during recognition tasks (e.g., Tun et al., 1998; Koutstaal et al., 1999).

The main finding for the recognition experiment was that there was a gradual decline in performance across the entire age range. Radvansky et al. (2001), although not having included a middle-aged group, suggested that some age effects might be due to a selection bias, rather than being caused by deficits in the older populations. Radvansky et al. (2001) argue that the young groups usually consist of college students who apply different strategies to the experimental tasks rather than a population of older people not as used to taking tests and memorizing information. In the present experiment, however, this was not the case. All participants were recruited from the same subject pool, and the education levels were matched across different age groups. Most importantly, the middle age group (from 38 to 53 years of age) was equally well educated as the youngest group. Thus, the age-related differences must be due to a change in recognition memory for text information.

Of course there are other studies on changes of text comprehension skill during the life span. For instance, Hess (1995), in his second experiment on causal inferences, included a group of middle-aged participants (age 36–58). The stimuli consisted of sentence pairs, comparable to the materials of Experiment 1, and the variable of interest concerned the connection between these sentences. Reading time patterns and subsequent recall performance of the middle-aged participants were more similar to those of the younger ones than to those of the older. The results of Experiment 1, in which most participants were of an age falling into Hess' young- and middle-aged group, are consistent with this finding. It remains an open question whether a group of elderly participants, comparable to the older group studied by Hess (1995) would show an impairment in the coherence judgment task.

For their structural modelling attempt, van der Linden et al. (1999) tested 150 adults from age 30 to 80 on a number of cognitive tasks. Although not specifically discussed by the authors, the inspection of the means shows age effects throughout the entire range of age levels and for all of their measures, including the language tasks. However, these data are based on rather global text comprehension measures, and thus do not allow for a separation of specific subprocesses.

The present experiments were intended to evaluate middle-aged people's performance, rather than as a stringent test of theories on cognitive aging. To evaluate the differential contributions of other cognitive factors (cf. van der Linden et al, 1999), we will briefly discuss the most influential theoretical proposals (see Kliegl et al., 1999). In particular, the present results are inspected with respect to working memory capacity (Craik, 1986) and inhibition deficits (Hasher & Zacks, 1988; Zacks & Hasher, 1988; cf. Burke, 1997).

A memory capacity account does not readily make predictions for the coherence judgment task. Given the results, it seems likely that the short sentence contexts used could easily be held in working memory even if the capacity was reduced. In the recognition task, the digit ordering test that taps executive functions more than the storage component of working memory, was unrelated to performance. In contrast, short-term memory capacity, as measured by the digit span test, accounted for some of the variance, suggesting that storage more than executive manipulation played a role. However, even when the digit span score was factored out, a significant, additional effect of age remained, indicating that a capacity account alone is not sufficient. This conclusion is in line with the findings of van der Linden et al. (1999). In their structural equation model, age effects on language comprehension were mediated by working memory skills, but also by other contributing factors.

Changes in inhibitory control (Hasher & Zacks, 1988) account well for the increased false alarm rates in the recognition task. However, a failure to suppress irrelevant information would also have predicted an age effect for the processing of incoherent sentence pairs in the coherence judgment task. In the performance data, this result was not obtained, and the reaction times even showed the opposite pattern. Thus, we can conclude that eventual age effects concern the difficulty of providing the appropriate inference rather than suppressing a contextually irrelevant possible connection in pragmatically incoherent trials.

Finally, the evidence on a contribution of frontal lobe functions to age-related decline was mixed. As mentioned before, the increase in false alarm rates in the recognition task is consistent with a change of inhibitory control, which in turn is a function of the frontal lobes. The response times for the coherence judgment provided some evidence for inference processes becoming more difficult with age, a result that is consistent with the finding of frontal lobe patients' deficits in this task (Ferstl et al., 2002). On the other hand, the memory measures provided evidence for a larger contribution of short-term memory capacity than of the executive component of working memory. In addition, the performance in the coherence judgment task was independent of age.

Further empirical research is needed to evaluate whether and how these different factors influence inferencing and memory for discourse. The present data contribute to our knowledge of text comprehension processes during middle age, an age group that has not been sufficiently studied. In recent years, however, continuing education and life-long learning have become keywords describing the changing demands of the work place. Therefore, it is important to become aware of possible changes in cognitive abilities and/or strategies during adulthood (cf. Lavigne & Finley, 1990).

The present data are also relevant for applications in neuropsychology. When evaluating text comprehension processes in brain-injured patients, the age brackets differ greatly from those of the younger students usually participating in psycholinguistic studies. Thus, the identification of tasks that are sensitive to age-related decline is important for valid assessment (e.g., Brownell & Martino, 1998). The results of the two experiments described herein suggest that the off-line coherence judgment task is more appropriate for clinical assessment (cf. Ferstl et al., 2002) than the recognition task in which age alone accounts for severe performance decrements.

ACKNOWLEDGMENT

The contributions of many colleagues made this work possible. Melanie Wilke and Susanne Sinzig assisted with data collection; Florian Siebörger, Anja Meinke, and Laura Busse helped with the development of the materials. Katja Werheid made available the digit ordering scores for the second experiment, and Marcel Brass commented on the manuscript. Most importantly, I would like to thank D. Yves von Cramon and Thomas Guthke, not only for their comments on a first draft of the article, but also for the many stimulating discussions concerning our collaborative research.

REFERENCES

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). *The CELEX lexical database* Center for Lexical Information: Max-Planck-Institute of Psycholinguistics, Nijmegen [CD-ROM].

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspective on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging* (pp. 1–24). New York: Cambridge University Press.

Beyer, R., Guthke, T., & Pekrul, U. (1996). Repräsentation von Textwissen im menschlichen Gedächtnis [Representation of text knowledge in human memory]. *Zeitschrift für Psychologie*, *204*, 199–232.

Beyer, R., Guthke, T., & Thiele, B. (1995). Integration von Vorwissen beim Sprachverstehen in Abhängigkeit vom Lebensalter [Age dependent integration of background knowledge during language comprehension]. *Zeitschrift für Psychologie*, *203*, 361–378.

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

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: Lawrence Erlbaum.

Burke, D. M. (1997). Language, aging, and inhibitory deficits: Evaluation of a theory. *Journal of Gerontology: Psychological Sciences*, *52B*, P254–P264.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilites* (pp. 409–422). Amsterdam: Elsevier.

Ferstl, E. C. (2001). Learning from text. In N. J. Smelser, P. B. Baltes & W. Kintsch (Eds.), *International encyclopedia of the social & behavioral sciences Volume 3.13: Cognitive psychology and cognitive science* (pp. 8605–8609). Amsterdam: Elsevier.

Ferstl, E. C., Guthke, T., & von Cramon, D. Y. (2002). Text comprehension after brain injury: Left prefrontal lesions affect inference processes. *Neuropsychology*, *16*, 292–308 .

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., & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In H. Oostendorp & S. Goldman (Eds.), *The construction of mental models during reading* (pp. 247–277). Mahwah, NJ: Lawrence Erlbaum.

Ferstl, E. C., & von Cramon, D. Y. (2002). What does the fronto-median cortex contribute to language processing: Coherence or Theory of Mind? *NeuroImage*, *17*, 1599–1612.

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.

Graesser, A.C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*, 163–189.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1998). Priming and aging: An electrophysiological investigation of N400 and recall. *Brain & Language*, *65*, 333–355.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1992). An electrophysiological study of semantic processing in young and middle aged academics. *Psychophysiology*, *29*, 38–54.

Gunter, T. C., Vos, S. H., & Friederici, A. F. (1999). Memory or aging? That's the question: an electrophysiological perspective on language. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: Aging, grammar, and memory* (pp. 259–282). Norwell, MA: Kluwer Academic.

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. *Psychology and Aging*, *7*, 56–64.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation, 22,* (pp. 193–225). San Diego, CA: Academic Press.

Hess, T. M. (1995). Aging and the impact of causal connections on text comprehension and memory. *Aging and Cognition*, *2*, 216–230.

Hess, T. M., & Flannagan, D. A. (1992). Schema-based retrieval processes in young and older adults. *Journal of Gerontology: Psychological Sciences*, *47*, 52–58.

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

Kintsch, W. (1988). The construction-integration model. *Psychological Review*, *95*, 163–182.

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

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, *29*, 133–159 .

Kliegl, R., Mayr, U., Junker, M., & Fanselow, G. (1999). Testing age invariance in language processes. In S. Kemper, & R. Kliegl (Eds.), *Constraints on language: Aging, grammar, and memory* (pp. 137–168). Norwell, MA: Kluwer Academic.

Koutstaal, W., Schacter, D. L., Galluccio, L., & Stofer, K. A. (1999). Reducing gist-based false recognition in older adults: Encoding and retrieval manipulations. *Psychology and Aging*, *14*, 220–237.

Lavigne, V. D., & Finley, G. E. (1990). Memory in middle-aged adults. *Educational Gerontology*, *16*, 447–461.

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics*, *12*, 433–452.

Radvansky, G. A., & Curiel, J. M. (1998). Narrative comprehension and aging: The fate of completed goal information. *Psychology and Aging*, *13*, 69–79.

Radvansky, G. A., Gerard, L. D., Zacks, R. T., & Hasher, L. (1990). Younger and older adults' use of mental models as representations for text materials. *Psychology and Aging*, *2*, 209–214.

Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging*, *16*, 145–160.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510–532.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1–91). Mahwah, NJ: Lawrence Erlbaum.

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language*, *25*, 279–294.

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

Stine-Morrow, E. A. L., & Miller, L. M. S. (1999). Discourse processing and aging. Resource allocation as a limiting factor. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: Aging, grammar, and memory* (pp. 53–76). Norwell, MA: Kluwer Academic.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science*, *182*, 990–1000.

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13*, 230–241.

Ulatowska, H. K., Hayashi, M. M., Cannito, M. P., & Fleming, S. G. (1986). Disruption of reference in aging. *Brain & Language*, *28*, 24–41.

Valencia-Laver, D. L., & Light, L. L. (2000). The occurrence of causal bridging and predictive inferences in young and older adults. *Discourse Processes*, *30*, 27–56.

Van den Broek, P. W. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores-d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423–445). Hillsdale, NJ: Erlbaum.

Van der Linden, M., Hupet, M., Feyereisen, P., Schlestraete, M.-A., Bestgen, Y., Bruyer, R., Lories, G., El Ahmadi, A., & Seron, X. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology and Cognition*, *6*, 32–55.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wechsler, D. (1987). *Wechsler memory scale-revised.* New York: Psychological Corporation.

Werheid, K., Hoppe, C., Thöne, A., Müller, U., Müngersdorf, M., & von Cramon, D. Y. (2002). The adaptive digit ordering test: Clinical application, reliability, and validity of a verbal working memory test. *Archives of Clinical Neuropsychology*, *17*, 547–565.

Wingfield, A., & Stine-Morrow, E. A. L. (2000). Language and speech. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 359–416). Mahwah, NJ: Lawrence Erlbaum.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 293–358). Mahwah, NJ: Lawrence Erlbaum.

Zacks, R. T., & Hasher, L. (1988). Capacity theory and the processing of inferences. In L. L. Light & D. M. Burke (Eds.), *Language, memory and aging* (pp. 154–170). New York: Cambridge University Press.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, memory and cognition*, *21*, 386–397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

Neuropsychology
2002, Vol. 16, No. 3, 292–308

0894-4105/02/$5.00 DOI: 10.1037//0894-4105.16.3.292

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-

---

Evelyn C. Ferstl, Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany; Thomas Guthke, Tagesklinik für Kognitive Neurologie, University of Leipzig, Leipzig, Germany; D. Yves von Cramon, Max-Planck-Institute of Cognitive Neuroscience, and Tagesklinik für Kognitive Neurologie, University of Leipzig.

We thank Laura Busse and Anja Meinke for their assistance with preparing the materials and collecting data.

Correspondence concerning this article should be addressed to Evelyn C. Ferstl, Max-Planck-Institute of Cognitive Neuroscience, Stephanstrasse 1 A, 04103 Leipzig, Germany. E-mail: ferstl@cns.mpg.de

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 nonaphasic 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	—	Temporolateral (posterior T3)	C
113	33	M	229	TBI	—	—	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 × 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 coherent 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, $F(1, 20) = 14.3$, $p < .01$.

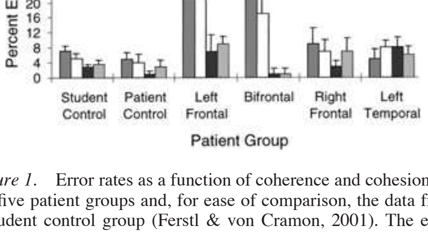


Figure 1. Error rates as a function of coherence and cohesion for the five patient groups and, for ease of comparison, the data from a student control group (Ferstl & von Cramon, 2001). The error bars indicate one standard error above the mean.

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,

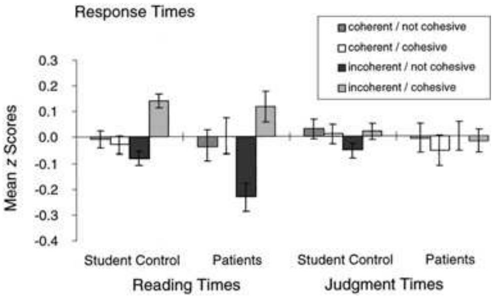


Figure 2. Response times as a function of coherence and cohesion for the entire patient group. Shown are the *z* scores for the reading times and the judgment times. The error bars are one standard error above the mean. The student control group's data from the pretest (Ferstl & von Cramon, 2001) are 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.

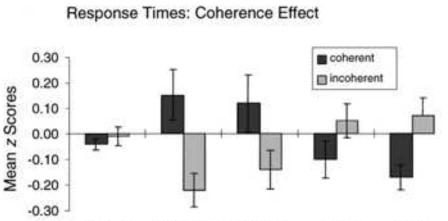


Figure 3. Response times as a function of coherence and patient group. Shown are the overall mean *z* scores (±1 *SE*), including both the reading and the judgment phases of the trials.

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

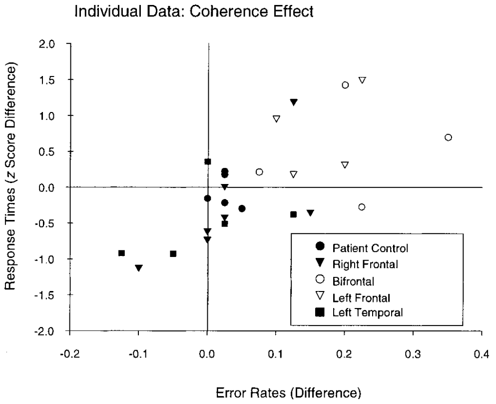


Figure 4. Scatterplot illustrating the relationship between the error rates and the response times. Shown is the coherence effect in the response time scores as a function of the coherence effect as reflected in the error rates. For both variables, the *coherence effect* is defined as the difference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

	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)
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 nonautomatic 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 relationship 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.

Received April 30, 2001
Revision received December 10, 2001
Accepted January 2, 2002 ■

NeuroImage **17,** 1599–1612 (2002)
doi:10.1006/nimg.2002.1247

What Does the Frontomedian Cortex Contribute to Language Processing: Coherence or Theory of Mind?

Evelyn C. Ferstl[1] and D. Yves von Cramon

Max Planck Institute of Cognitive Neuroscience, Stephanstrasse 1a, D-04103, Leipzig, Germany

Received December 20, 2001

The frontomedian cortex (FMC) has been shown to be important for coherence processes in language comprehension, i.e., for establishing the pragmatic connection between successively presented sentences. The same brain region has a role during theory-of-mind processes, i.e., during the attribution of other people's actions to their motivations, beliefs, or emotions. In this study, we used event-related functional magnetic resonance imaging at 3 T to disentangle the relative contributions of the FMC to theory-of-mind (ToM) and coherence processes, respectively. The BOLD response of nine participants was recorded while they listened to pragmatically coherent or unrelated sentence pairs. Using a logic instruction for inanimate sentence pairs, ToM processing was discouraged during the first part of the experiment. Using explicit ToM instructions for sentence pairs mentioning human protagonists, ToM processing was induced during the second part. In three of the resulting four conditions a significant increase in the BOLD response was observed in FMC: when ToM instructions were given, both coherent and incoherent trials elicited frontomedian activation, in replication of previous results showing involvement of the FMC during ToM tasks. When logic instructions were given, the coherent trials, but not the incoherent trials, activated the FMC. These results clearly show that the FMC plays a role in coherence processes even in the absence of concomitant ToM processes. The findings support the view of this cortex having a domain-independent functionality related to volitional aspects of the initiation and maintenance of nonautomatic cognitive processes. © 2002 Elsevier Science (USA)

Key Words: **coherence; theory of mind; functional magnetic resonance imaging; frontomedian cortex; text comprehension.**

INTRODUCTION

For social interaction, the ability to understand other people's beliefs, motivations, and goals is crucial. "Mentalizing" (Frith and Frith, 1999) or Theory of Mind (Premack and Woodruff, 1978), as this ability has been termed, is at the core of successful interaction with other human beings. Theory of Mind (ToM) is necessary for understanding and predicting other people's behavior and, thus, for reacting adaptively to the changing environment. There is a large body of knowledge about the development of this faculty in childhood (Wellman, 1993), and it has been shown that autistic children have specific deficits in ToM processing (Baron-Cohen *et al.,* 1985). Recently, a number of neuropsychological patient studies and neuroimaging studies have attempted to further our understanding of the neuroanatomical realizations of ToM abilities. A cortical region shown to be particularly important for ToM processes is the frontomedian cortex (FMC; also termed the dorsal medial prefrontal cortex), comprising mainly the medial aspects of Brodmann area (BAs) 9 and reaching into the adjacent portions of BAs 8, 10, and 32. In the following we first review the evidence for the FMC being involved during ToM processing, and then raise the question of whether other, more general processes might engage the FMC as well. Finally, we present an experiment designed to explicitly test the hypothesis that ToM processes are sufficient, but not necessary for FMC activation.

Patient studies on ToM deficits after acquired brain injury include investigations of right hemisphere involvement (Happé *et al.,* 1999; but see Siegal *et al.,* 1996) and the role of the amygdala (Fine *et al.,* 2001). Because most tasks employed for testing ToM abilities involve complex reasoning and working memory processes, on one hand, or the processing of affective information, on the other, a number of studies have focused on the effect of lateral prefrontal and orbitofrontal lesions on ToM abilities. Channon and Crawford (2000) found that left frontal lesions were most likely to cause a deficit in a story comprehension task involving ToM processes. Similarly, Rowe *et al.*

---

[1] To whom correspondence should be addressed. Fax: ++49 (341) 99 40 221. E-mail: ferstl@cns.mpg.de.

(2001) documented impairments in a false belief story comprehension task. Indpendent of laterality, most patients in their sample had lesions including medial or orbitofrontal regions. Additional evidence for the importance of medial, rather than dorsolateral frontal areas for ToM was provided by Stone *et al.* (1998), using a faux pas task, and by Stuss *et al.* (2001), using a deception task. Taken together, these few studies provide strong evidence for a frontal, and particularly a frontomedian, contribution to ToM abilities.

Additional evidence for the involvement of frontal brain areas is provided by a growing number of PET and fMRI studies on ToM processes. Whereas most patient studies used second- and third-order belief tasks or deception tasks, i.e., tasks whose properties are well known from the developmental literature on ToM processing, the imaging studies employed a set of different tasks. A most influential paradigm was introduced by Fletcher *et al.* (1995). In their PET study, they directly compared the processing of stories matched in length and complexity, but different with respect to an inherent ToM component. When ToM processing was required, the anterior and posterior cingulate cortices, a right inferior parietal region, and, most importantly, a region in the frontomedian wall (BA 8/9) were activated. This finding has been replicated with fMRI imaging, with visual stimuli (cartoon stories) instead of language stimuli (Gallagher *et al.,* 2000), and with stories worded in first person rather than in third person (Vogeley *et al.,* 2001).

Other studies using paradigms intended to induce different aspects of ToM processing yielded FMC activation as well. In a face perception study, a large, unspecific network was shown to be involved when the affective expression was to be judged, as compared with the person's gender (Baron-Cohen *et al.,* 1999); in an object perception study, more activation was found in FMC when an inference on other people's beliefs about the objects was required, compared with the retrieval of the objects' use from semantic memory. And finally, the perceived intentionality or goal orientation of visually presented stimuli predicted the FMC activation, both for meaningful cartoon stories (Brunet *et al.,* 2000) and for abstract movement patterns (Castelli *et al.,* 2001). Even for autistic or Asperger's patients, whose ToM abilities seem impaired, some FMC activation was reported during different ToM tasks (Baron-Cohen *et al.,* 1999; Happé *et al.,* 1996).

These imaging results clearly confirm that, among other brain regions, the FMC is part of the "social brain" (Brothers, 1990). However, evidence against the view of FMC being specific for ToM, i.e., against the view that ToM processes are necessary for activating the FMC, comes from imaging studies reporting FMC activation elicited by tasks other than ToM tasks. For instance, Zysset *et al.* (2002) reported FMC activation for the judgment of evaluative statements, Goel *et al.* (1997) for inductive as compared with deductive reasoning, and Gusnard *et al.* (2001) for self-referential judgments compared with neutral classifications. A recent study on moral judgment argued that the same pattern of frontomedian and retrosplenial activation is caused by emotional processing (Greene *et al.,* 2001). Another group of studies concerned higher-level language comprehension. FMC activation was found in the context of pragmatic comprehension (Bottini *et al.,* 1994; Nichelli *et al.,* 1995) and story comprehension (Maguire *et al.,* 1999).

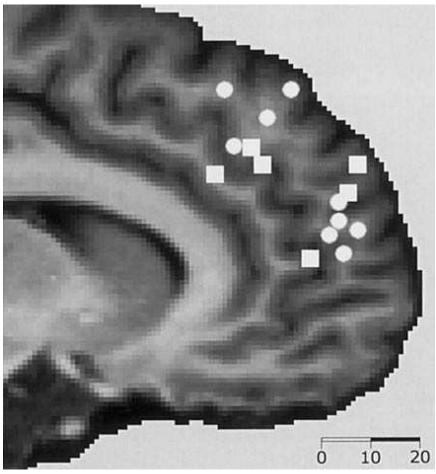


FIG. 1. Frontomedian peaks of activation from selected imaging studies (restricted to BA 8–10, irrespective of lateralization). The squares are from ToM tasks (Castelli *et al.,* 2000; Fletcher *et al.,* 1995; Gallagher *et al.,* 2000; Goel *et al.,* 1995; Vogeley *et al.,* 2001), the circles denote peaks from plausibility judgments, reasoning, evaluation, coherence judgment, and self-referential processing (Bottini *et al.,* 1994; Ferstl and von Cramon, 2001; Goel *et al.,* 1997; Greene *et al.,* 2001; Gusnard *et al.,* 2001; Zysset *et al.,* 2002). It is apparent that both types of tasks activate similar regions in FMC.

To illustrate how well the results of these studies overlap, despite the apparent variety of tasks employed, Fig. 1 displays the peaks of the activated regions for a selection of studies, projected onto the left median wall of an individual brain. As can be seen, there is no obvious anatomical division between studies investigating ToM processes, on one hand, and studies on plausibility, language, or evaluative judgments, on the other.

To illustrate two possible explanations for this overlap, let us consider a previous fMRI study on inference processes during text comprehension. Presented with sentence pairs (see Table 1), the participants' task was to evaluate their coherence (cf. Kintsch, 1998), i.e., to judge whether there was a pragmatic connection be-

TABLE 1

Sentence Examples

	Coherent	Incoherent
Part A: Logic	Sometimes a truck drives by the house. That's when the dishes start to rattle.	Sometimes a truck drives by the house. The car doesn't start.
	The lights have been on since last night. The car doesn't start.	The lights have been on since last night. That's when the dishes start to rattle.
Part B: Theory-of-Mind	Mary's exam was about to begin. Her palms were sweaty.	Mary's exam was about to begin. Some friends had remembered the birthday.
	Laura got a lot of mail today. Some friends had remembered the birthday.	Laura got a lot of mail today. Her palms were sweaty.

tween the two sentences or not.[2] During performance of this task, complex inference processes are carried out. It is necessary to comprehend both sentences, to store the context sentence's content in working memory, to retrieve relevant world knowledge, and finally to integrate all these information sources with the information provided in the second sentence. In contrast to the expectation that the perisylvian language areas, or their right hemispheric homologues, are active mainly during inferencing, we found medial cortical regions to be sensitive to coherence. Specifically, the FMC and a corresponding posterior cingulate region (PCC) were more active during the processing of coherent as compared with incoherent trials (Ferstl and von Cramon, 2001).

The first explanation for overlapping FMC activation is that in the studies on evaluative, reasoning or coherence processes, concurrent ToM processes were inadvertently induced, although the tasks were not explicitly designed for doing so. For instance, tasks in which the participants' opinions are asked for, such as those employed in the Gusnard and Zysset studies, can be linked to ToM processing under the assumption that understanding other people's minds is closely related to reflecting on one's own (Vogeley *et al.,* 2001). Any task in which the plausibility of statements needs to be evaluated, in particular in the context of human protagonists, might automatically elicit ToM processes. In our own materials for the coherence judgment task (Ferstl and von Cramon, 2001), sentence pairs with and without human protagonists were intermixed. For the animate sentence pairs (shown in the bottom rows of Table 1), the coherence judgments could have been based on a ToM evaluation. Rather than answering the question: "Can I find a pragmatic connection between these two sentences?" subjects might simply have answered the question: "Can I understand what these people are doing?"

A second explanation of why coherence and ToM processes both elicit FMC activation is that this cortex might subserve a more general function concomitant with both coherence and ToM processing. Gusnard *et al.* (2001), for instance, interpreted the BA 8/9 activation as being related to self-referential processes (cf. Vogeley *et al.,* 2001), and Greene *et al.* (2001), as related to emotional processes. In our own study, we interpreted the FMC activation as being related to the initiation and maintenance of nonautomatic cognitive processes (Ferstl and von Cramon, 2001), in particular when long-term memory and evaluative components are needed for task performance (Zysset *et al.,* 2002).

The present study was designed to provide evidence for the second explanation. The goal was to prove ToM processing to be sufficient, but not necessary for activating the FMC. The study had two event-related parts (A and B) in each of which coherent and incoherent trials were intermixed. Part A was designed to show that coherence alone could activate the FMC, even in the absence of ToM processes: we first selected the inanimate sentence pairs from our previous study (Ferstl and von Cramon, 2001) for separate presentation (Part A). To discourage ToM processes, the instructions focused on the internal logic of the sentence pairs. Part B was designed to show that ToM alone could activate the FMC, in the absence of coherence. During Part B, the animate sentence pairs were presented. To induce ToM processing, the participants were asked to try to identify with the protagonists mentioned, and to attempt to understand their motivations for and feelings during the actions described. Since the crucial comparison concerned the replication of the coherence effect in the *absence* of concurrent ToM processes, Part A was always presented first.

The predictions are clear-cut. If in FMC a specific ToM process was realized, its activation would be ex-

---

[2] Note that we use the term *coherence* in its linguistic sense only. The related concept of *central coherence* used in autism research is considerably more general (Frith, 1989; see Happé, 2000, for a review), although it encompasses language processes (cf. Jolliffe and Baron-Cohen, 1999, 2000).

pected during Part B only, but not during Part A. If, in contrast, the frontomedian activation was due to coherence processes, FMC activation would be expected for coherent trials, during Part A as well as during Part B, but not for incoherent trials. If, as we believe, both processes share a more general, domain-independent component related to the initiation and maintenance of nonautomatic cognitive processes, we expect a replication of the coherence effect during Part A, as well as FMC activation for both coherent and incoherent trials during Part B.

METHODS

Participants

Five women and four men, all right-handers, received reimbursement for participating in the experiment. None of the participants had any history of neurological disorder or other health problems preventing them from being exposed to the magnetic field. The median age was 24 years (range 22–27). All participants had given informed consent.

Design and Materials

The experiment consisted of two independent parts (A and B) with identical procedure. In both parts, the within-subjects factor Coherence (coherent vs incoherent) was manipulated, but the types of materials as well as the instructions were varied. In Part A (Logic) ToM processes were minimized, whereas in Part B (ToM) ToM processes were explicitly induced. The materials were taken from the previous fMRI study (Ferstl and von Cramon, 2001). The language trials were made up of 120 coherent sentence pairs in which the second sentence (the target) was pragmatically related to the first (the context). For Part A, those 60 sentence pairs were selected that referred to objects only and did not mention human protagonists. For Part B, those 60 sentence pairs were used that contained references to people. The incoherent conditions were created by switching the context sentences of two coherent trials. Examples for the resulting four types of trials are provided in Table 1.[3]

All sentences were read by a female speaker and tape recorded. By separating the recording of the context sentences and the target sentences, prosodic cues as to the coherence of the trials were avoided. For the computer presentation the recordings were digitalized and their loudness was adjusted.

The presentation length for the sentence pairs was 5.3 s on average (SD = 0.7, range = 3.0–8.1 s), and there were no differences as a function of Part or Coherence (all F's < 1).

For the control condition in the fMRI experiment, we used an additional 32 trials made up of pseudo-word sentences. Because it was necessary to control for both the auditory perception and the task demands (listening to the two "sentences," relating them to each other, and providing a motor response), we used two different types of pseudo-sentences. Half of the pseudo-sentences contained some German function words or function morphemes (e.g., *Der* Miefensalm *ist noch* kolmut *ge*klubet), and half did not (Molsa erkau lanschdal ettenbul giller). Thirty-two pseudo-sentence pairs were then created so that in 16 of the trials, the two pseudo-sentences were from the same item type (consistent), and in the other 16 they were from different item types (inconsistent). The length of the control trials was comparable to that of the experimental trials ($M = 5.4$ s, range = 3.6–6.7 s, SD = 0.8).

For the event-related experiment, we intermixed stimuli from the three conditions (coherent, incoherent, control) as follows: For both Part A and Part B separately, four counterbalancing lists of the 60 experimental items were created. Within each list, 30 target sentences were paired with their coherent context sentence; the remaining 30 target sentences were paired with the incoherent context. Across the four lists, each sentence appeared twice in each of the two conditions. The 16 control trials were intermixed with the experimental trials. Once more, half of the control trials were in each of the two consistency conditions. The order of the trials in each of the resulting lists was randomized separately, with the constraint of not more than three consecutive trials being in the same condition. In each of the two experiment parts, a warmup trial at the beginning and a buffer trial at the end were added. Thus, the two parts of the experiment consisted of a total of 78 trials each.

Procedure

The participants received written instructions for Part A (Logic) before the scanning session started. The participants were told to carefully listen to the sentence pairs and to indicate after the second sentence whether there was a logical connection to the first sentence or not. Two examples each were given for coherent and incoherent sentence pairs and the appropriate response (YES or NO) was indicated. Furthermore, the participants were informed that trials with nonword sentences would be interspersed. Examples were given for the two types of stimuli. The participants were told that they were from different artificial languages, one of which sounded somewhat like German while the other had no resemblance to German. The task for these trials was to indicate whether the pseudo-language was the same or different for the two

---

[3] As in our previous experiment (Ferstl and von Cramon, 2001), half of the target sentences in each condition contained cohesive ties, i.e., words explicitly signaling a connection between the sentences, such as pronouns and conjunctions.

nonword sentences. For their YES responses, half of the participants were instructed to use the key on the right side of the response box, and the other half to use the key on the left.

During the scanning session, the participants lay flat inside the magnet, with a response box placed into their right hands. They were instructed to keep their eyes closed throughout the experiment. The stimuli were presented to both ears through headphones. To protect from the scanner noise, participants wore ear plugs that muffled the noise while allowing the stimuli to be heard.

The functional measurement for Part A involved the presentation of one of the four lists of inanimate sentence pairs. Each trial lasted 20 s, with the following time course: A short beep alerted the subject to the beginning of the presentation 500 ms before the first sentence was played. Then the two sentences were presented with an interstimulus interval of 500 ms. The timing of the trial was such that the onset of the second sentence always occurred 6200 ms after the trigger signaling the beginning of the trial. This timing yielded variable intertrial intervals, whose length depended on the response time of the previous trial and the duration of the context sentence of the current trial. The reaction times were measured from the onset of the target sentence up to the participant's button press for the YES or NO response. There was an unfilled pause until the next presentation started.

The instructions for the second part of the experiment (Part B: ToM) were read to the participants after the first part had been completed. Remaining in the scanner, the participants were informed that the subsequent sentences would refer to human protagonists. The instructions further stated (translated from German): "Your task is to identify with the people mentioned. You should try to put yourself into their shoes, i.e., to understand their motivations, feelings, and actions. After the second sentence, please press the YES key if you succeeded, and the NO key if you did not. It is especially important to take into account both sentences." Furthermore, the instructions stated that the control task was continued in the same way as in the first part of the experiment. After the experimenter had ensured that these instructions were understood, functional measurement for Part B started. The 78 animate trials were presented with a time course identical to that of Part A.

The presentation of the two parts of the experiment lasted 52 min, and the entire scanning session had a duration of about 65–70 min.

Data Acquisition

A Bruker Medspec 30/100 system was used for magnetic resonance imaging at 3.0 T. Prior to the functional scans, two anatomical scans were acquired for each participant using MDEFT sequences (Ugurbil *et al.,* 1993). The first was a whole-brain image acquired with a T1-weighted 3D segmented sequence (Norris, 2000; 128 sagittal, adjacent slices, 1.5 mm thick, 256 × 256-pixel matrix per slice; TR 1.3 s, TE 10 ms). To enable alignment of the functional scans with this high-resolution image, anatomical T1-weighted 2D images were acquired, using the same number and orientation of slices as the functional scans.

During the functional scans, the BOLD response was measured using a single-shot gradient EPI sequence (matrix 64 × 64, TE = 30 ms, flip angle 90°, field of view 192 mm, acquisition bandwidth 100 kHz). Horizontal images were acquired for 16 slices (5-mm thickness, 2-mm spacing), parallel to the bicommissural plane (AC-PC). For most participants, 6 slices were below the AC–PC line, while 10 slices were above, but care was taken that the temporal lobes, as well as prefrontal regions, were covered in full. In-plane resolution was 3 × 3 mm. We used a repetition time of 2 s (TR = 2), and the presentation of the displays was triggered by the acquisition of the first slice of the current image.

Data Analysis

Data analysis was conducted using the software package LIPSIA, developed as a tool for analyzing functional MRI data (Lohmann *et al.,* 2001).

For each participant, the signal acquired during the functional scans was preprocessed as follows: First, a sinc-interpolation algorithm was applied to correct for the temporal spacing between the 16 slices of each image. Motion correction consisted of a global affine linear transformation that optimized for each time step the linear correlation between the image at that time step and a predefined reference image. A baseline correction was then conducted using a temporal high-pass filter with a cutoff frequency of 1/40 Hz. Furthermore, a spatial Gaussian filter was applied with a standard deviation of $\sigma = 0.8$. With a voxel size of 3 mm, this standard deviation is equivalent to a FWHM of 5.65.

For the statistical analysis we carried out the following sequence of processing steps (Lohmann *et al.,* 2001): For each participant separately, the 2D data were analyzed using the General Linear Model (cf. Friston, 1994) based on a 3 (Trial: coherent, incoherent, control) × 2 (Part A, Part B) within-subject design. For the event-related model we time-locked the BOLD responses 1000 ms before the offset of the target sentence. A synthetic hemodynamic response function with a lag of 6 s was assumed. The model equation, i.e., the observation data, the design matrix, and the error term, was convolved with a Gaussian kernel with a dispersion of 4 s FWHM. Contrast codes were then used to detect significant activations by calculating t statistics based on the parameter estimates of the full linear model, and subsequently transforming the t values into Z scores. For each participant, the following five comparisons were made: For each part separately all language trials were compared with the control trials, and then the coherent with the incoherent trials. Because of the fixed order of the two parts of the experiment, it was not feasible to directly compare the activations in the two parts. However, to statistically strengthen the resulting differences, some results from the interaction contrasts between Language and Part, as well as Coherence and Part, are also reported.

For the statistical analysis across participants, all Z maps of each participant were fitted into a standard stereotaxic space (Talairach and Tournoux, 1988) as follows: First, the 2D anatomical scan was rotated and shifted so that it mapped onto the 3D whole-brain image. Then, linear scaling factors were calculated to transform the image to the standard size. The resulting coregistration matrices were finally applied to the contrast images using trilinear interpolation (Lohmann, 1998). The resulting standardized 3D images were finally transformed using deformation fields obtained from a nonlinear adjustment of the participants' individual anatomies (Thirion, 1998). The resulting Z maps were then averaged across participants (Bosch, 2000). Only those voxels with $|Z| > 3.09$ were considered significantly activated ($P < 0.001$). Since this probability level is uncorrected for multiple comparisons, we defined an additional spatial extent threshold. Areas of activation smaller than 225 mm^3 were neglected. This threshold roughly corresponds to the requirement that at least 5 adjacent voxels in the original image (of volume $3 \times 3 \times 5$ mm $= 45$ mm^3) be significantly activated.

For the displays of the results, the deformation fields from the nonlinear scaling were applied to the individual participants' anatomical images, and a mean brain was obtained by averaging these normalized images. In all figures, the Z maps are overlaid on this mean brain for our group of participants.

RESULTS

Behavioral Data

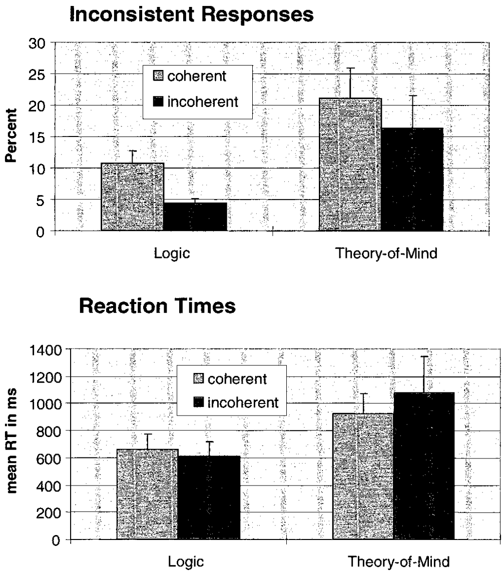


FIG. 2. Behavioral data for the two parts of the experiment. Top: Mean percentage of the responses coinciding with the coherence of the sentence pair. Bottom: Mean response times. The error bars are one standard error above the mean.

For a measure of performance, the proportions of answers coinciding with the coherence of the sentence pair were calculated. A 2×2 ANOVA with the factors Part and Coherence yielded a main effect of Part ($F(1,8) = 10.1$, $P < 0.01$). During the first part, only 7.6% (SD = 3.3) of the responses deviated from those predicted by Coherence, whereas in the second part 18.7% (SD = 11.4) of the responses deviated. Although neither the main effect of Coherence nor the interaction between Coherence and Part reached significance ($F(1,8) = 2.2$ and $F(1,8) < 1.0$, respectively), inspection of the data showed that this was due to large intersubject variability during Part B. Planned pairwise comparisons confirmed that Coherence had an effect during Part A ($F(1,8) = 9.6$, $P < 0.02$). Slightly more answers deviated for the coherent trials than for the incoherent trials. No such effect was observed during Part B ($F(1,8) < 1$).

For an analysis of the response times, measured from the onset of the target sentences, the presentation duration of the target sentences was subtracted. The resulting reaction times correspond to the latency of the response after the offset of the target sentence. The means of the response times are shown in Fig. 2 for both parts as a function of coherence. The overall mean reaction time was 818 ms (SD = 408) and varied from 400 ms up to 1700 ms. To control for this large variance, the reaction times were first standardized for each subject separately. The statistical analysis, based on the resulting Z scores, yielded a main effect of Part ($F(1,8) = 7.4$, $P < 0.05$). The reaction times were longer during Part B than during Part A. Neither the main effect of Coherence ($F(1,8) < 1$), nor the interaction with Part ($F(1,8) = 2.4$, $P = 0.15$) reached significance.

These results confirm that the participants heeded the instructions. In both performance and reaction times, there were clear effects of the experimental task. Furthermore, the response pattern during Part A

TABLE 2

Characteristics of the Cortical Regions of Activation for the Contrast Comparing All Language Trials with the Control Task[a]

			Part A: Logic					Part B: Theory of Mind				
Language–Control	BA	Side	Size	$Z_{\max}$	X	Y	Z	Size	$Z_{\max}$	X	Y	Z
1. Frontomedian cortex	9/10/32	L	960	5.32	−16	51	30	7,548	6.57	−19	49	30
2. Retrosplenial cortex/precuneus	29/30/23/31	L	1,745	6.66	−11	−59	15	17,946	9.10	−10	−58	21
	29/30	R	608	5.39	8	−57	14	—	-	6	−55	12
3. Temporoparietal cortex	39/22	L	8,541	7.71	−43	−63	27	10,624	9.11	−50	−65	25
4. Posterior STS	22/21	L	—	—	−60	−58	7	869	5.66	−59	−58	8
5. Anterior STS	21	L	4,314	8.26	−53	−6	−13	5,806	8.56	−54	−13	−8
	21	R	1,682	6.47	45	−2	−16	3,767	8.27	46	−2	−16
6. Fusiform/parahippocampal gyrus	35/36	L	1,491	5.88	−33	−34	−9	1,002	5.49	−31	−33	−9
	35/36	R						296	4.41	28	−25	−12
7. Superior frontal sulcus	8	L						1,179	5.24	−19	29	48
8. Inferior frontal gyrus	45	L						349	4.79	−48	17	0
9. Lateral occipital gyri	19	R						793	4.33	29	−93	13
10. Parieto-occipital cortex	39/19	R						537	4.65	41	−77	12

[a] Shown are the approximate Brodmann areas, Z values, and Talairach coordinates for the peaks and the size of the region of activation (in mm^3). The activations are displayed in Fig. 3, using the same labels. The thresholds were set to $Z > 3.8$ and a minimal size of the regions of 225 mm^3.

closely resembled that of the previous experiments using a coherence judgment task.

fMRI Results

For Part A and Part B separately, Table 2 provides a list of the significant regions of activation from the contrast comparing all language trials with the control task. To focus the discussion, only the most prominent areas are listed ($Z > 3.8$, $P < .0001$; extent $> 225\ \text{mm}^3$).

Figure 3 illustrates that remarkably similar networks were engaged in both parts of the experiment, with Part B yielding stronger activations than Part A. As expected, language processing in context involved the anterior portion of the superior temporal sulcus bilaterally (BA 21), more pronounced in the left hemisphere. In both parts, there was activation along the collateral sulcus (fusiform/parahippocampal gyri) on the left side, with the corresponding region on the right being significant during the ToM part only. In both parts, there was strong involvement of a temporoparietal region (BA 39/22), including the angular gyrus. Three additional, small regions reached the threshold during Part B only: the tip of the triangular part of the inferior frontal gyrus (BA 45), a region in the parieto-occipital transition area (BA 39/19), and one in the lateral occipital gyri (BA 19).

In the median wall, there were two large regions of activation. The center of gravity of the bilaterally active posterior region lay in the retrosplenial area (BA 29/30). In Part B, this activation extended further into the left inferior precuneus (BA 23/31). The anterior border was delineated by the marginal branch of the cingulate sulcus, the posterior border by the “common stem” and the parieto-occipital sulcus. The frontal region, clearly left lateralized, fell into the median part of the superior frontal gyrus, dorsal to the paracingulate sulcus (posterior BA 9). In Part B, this activation extended ventrally into BA 10 and dorsally into the superior margin (BA 9/10/32), and it was accompanied by a more lateral, small area in the superior frontal sulcus (BA 8).

The most apparent difference between the activation patterns for the two parts of the experiment was that the median regions were more strongly activated during Part B than during Part A. To statistically strengthen the different involvement of median structures during the two parts, we calculated the contrast coding the interaction between the effects of language and the effects of part. In confirmation of the descriptive differences, both the posterior ($Z_{\max} = 5.73$; size = 18,595 mm^3; peak [1 −66 38]), as well as the anterior median region ($Z_{\max} = 4.51$; size = 2299 mm^3; peak [−4 52 10]) proved to be more active during Part B than during Part A.

The results for the contrast comparing coherent with incoherent trials is shown in Table 3. For Part B, under the ToM instructions, this contrast did not yield any significant regions of activation, indicating that coherent and incoherent trials were processed in a similar fashion. During Part A, using the Logic instructions, the results resembled those of our previous study (Ferstl and von Cramon, 2001): None of the regions were more active during incoherent trials than during coherent trials. For the reverse comparison, however,

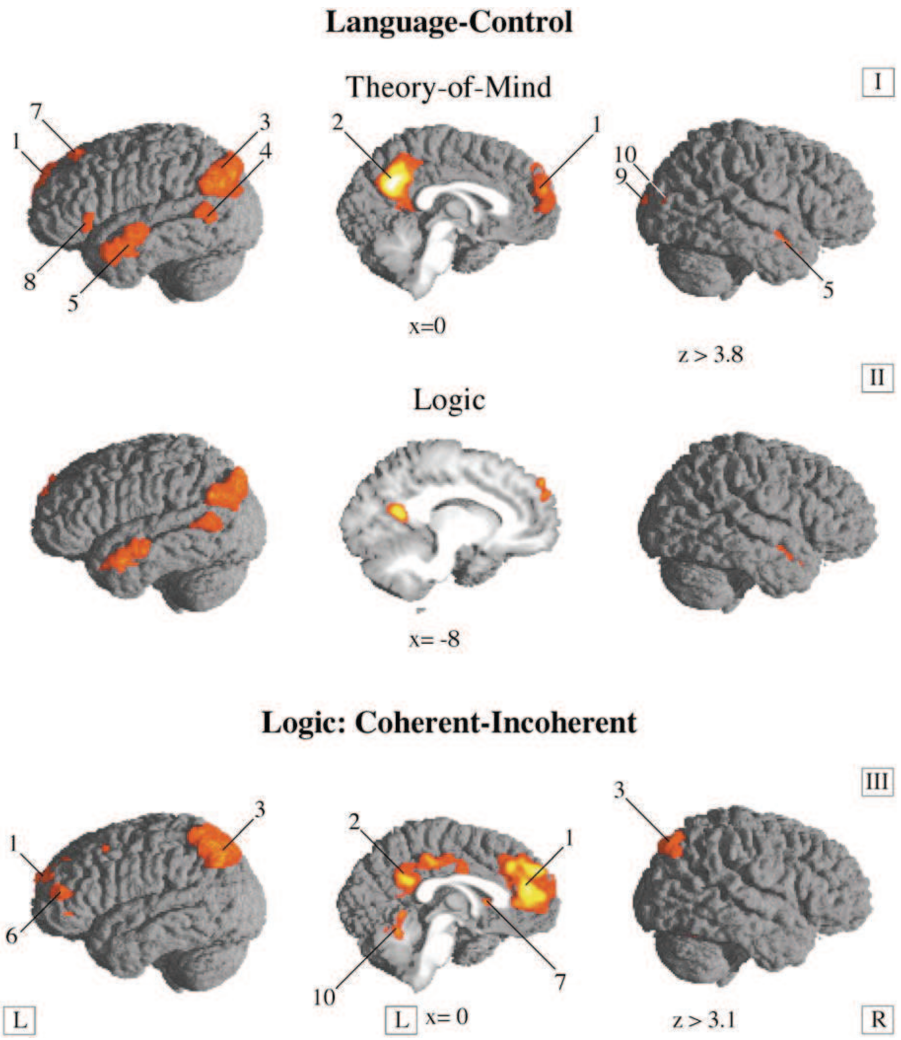


FIG. 3. Significant regions of activation are projected onto the cortical surface of an average brain, obtained by nonlinear transformation of the participants' individual anatomies. In all three panels, sagittal views of the left lateral, left medial, and right lateral cortices are shown from left to right. (I, II) Activations from the contrast Language–Control task, with Panel I showing regions of activation during Part B (ToM), and Panel II during Part A (Logic). For both contrasts, a threshold of $Z > 3.8$ was used. The numbers labeling the regions are those used in Table 2. (III) Significant activations from the contrast Coherent–Incoherent during Part A (Logic). For this contrast, we used a threshold of $Z > 3.1$ and the labels from Table 3.

we replicated the activations in the retrosplenial area (BA 29/30) on the left, extending into the posterior cingulate sulcus (BA 23). Similarly, there was a large left lateralized region of activation in the FMC, spreading from the anterior cingulate gyrus over the cingulate and paracingulate sulci (BA 24/32) into the median part of the superior frontal gyrus (BA 9/10). In addition, bilateral parietal structures, including the supramarginal gyri, were sensitive to coherence. Two left prefrontal regions, one in the superior frontal sulcus (BA 9), the other in the orbital part of the inferior frontal gyrus (BA 45/47) and, finally, an area in the

TABLE 3

Cortical Regions of Activation for the Contrast Comparing Coherent with Incoherent Trials ($Z > 3.09$, Size > 225 mm^3)[a]

			Part A: Logic				
Coherent–Incoherent	BA	Side	Size	$Z_{\max}$	X	Y	Z
1. Frontomedian cortex/ACC	9/10/24/32	L	24005	5.57	−6	26	35
2. Retrosplenial area/PCC	29/30/23	L	6281	4.51	−4	−28	35
3. Posterior parietal cortex	39	L	9350	4.98	−54	−62	41
	39	R	1705	4.05	43	−70	44
4. Inferior precentral sulcus	6/44	L	753	4.24	−36	−2	25
5. Superior frontal sulcus	10	L	374	4.15	−27	34	34
6. Inferior frontal gyrus	45/46/47	L	1393	4.10	−45	33	1
7. Caudate nucleus	—	L	7106	4.66	−13	3	18
8. Midbrain (VTA)	—	L	873	4.90	−8	−20	−6
9. Hippocampal formation	—	L	691	4.19	−25	−26	−6
10. Cerebellum	—	L	1433	3.72	−3	−53	−15
11. Cerebellum	—	R	978	3.85	32	−58	−20

[a] There were no significant results for the comparison of incoherent with coherent trials. During Part B (Theory of Mind), the two conditions did not yield any different regions of activation either.

fundus of the left inferior precentral sulcus (BA 44/6), proved to be more active during coherent trials than during incoherent trials.[4]

For further confirmation of the differential effect of Coherence in the median wall of the left hemisphere during the two parts of the experiment, the contrast coding the interaction between Coherence and Part was calculated. And indeed, this contrast yielded two regions in the cingulate sulcus that were significantly more sensitive to coherence during Part A than during Part B (BA 8/9/32, [−6 25 35], $Z_{\max} = 4.59$, size = 4498 mm^3; and BA 9/10/32 [−4 39 16], $Z_{\max} = 3.8$, size = 668 mm^3). In contrast to these FMC activations, the coherence effect in the PCC/precuneus did not significantly differ for the two parts of the experiment.

Region-of-Interest Analysis

To evaluate which of the observed differences provided by the statistical comparisons was due to activation of the relevant areas, rather than to deactivation in one of the comparison conditions, the frontomedian changes in the BOLD contrast were analyzed in more detail. The signal change in the raw data provides information about task-induced effects independent of the particular choice of control task. Six locations were chosen to cover the extent of the region of activation as shown in Fig. 3. The six locations and their coordinates are displayed in Fig. 4. For each peak, the percentage signal change in the corresponding voxel of the original image (size 27 mm^3) was separately averaged for each participant and each condition. Figure 5 displays the means of the resulting values for each of the four conditions, across a window of 11 s, starting 1 s before the offset of the target sentences. The graphs clearly show that for all but the most ventral region (BA 10), the sentences elicited a positive-going curve. Moreover, for these five peaks, the condition with the largest increase in the BOLD response was the coherent condition during Part A. In contrast, the coherence effect in BA 10 (Talairach coordinates [−4 52 10]) was apparently due to a negative-going curve in the incoherent condition.

To further ensure that that the activations were statistically generalizable and stable across subjects, an additional region-of-interest (ROI) analysis was

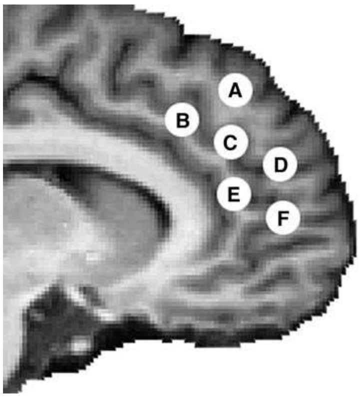


FIG. 4. Six regions of interest in FMC.

---

[4] The activation of components of the mesofrontal dopaminergic pathway on the left side, i.e., the ventral tegmental and retrorubral area together with the caudate nucleus/nucleus accumbens, is consistent with the frontomedian cortical activation (Williams and Goldman-Rakic, 1998). However, a more thorough discussion of this issue is beyond the scope of this article.

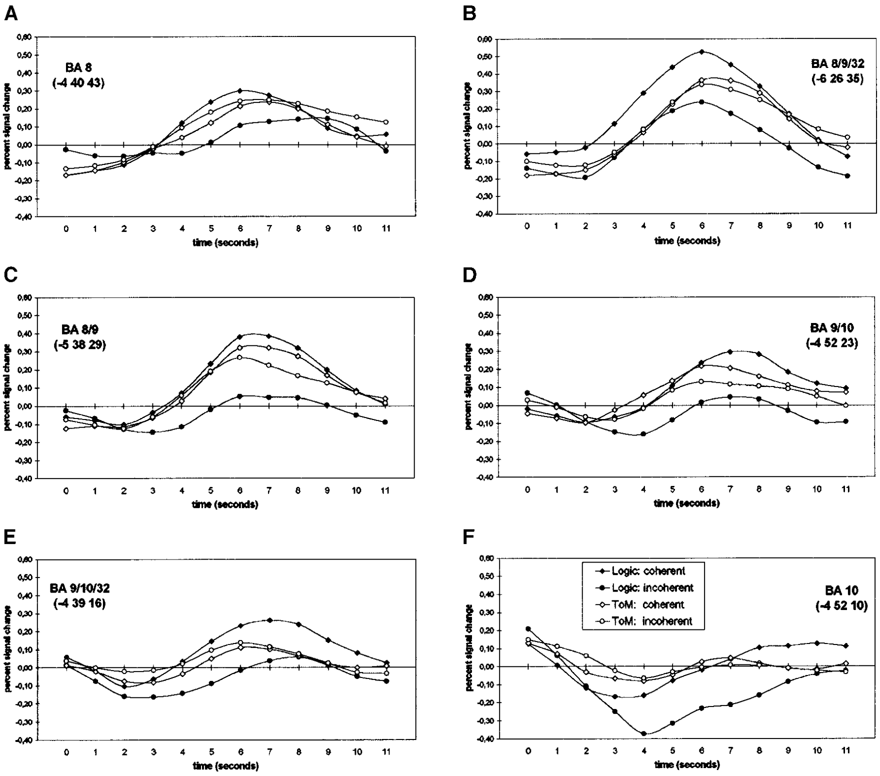


FIG. 5. Time course of activation in six frontomedian regions of interest (cf. Fig. 4). Shown is the percentage signal change for both parts of the experiment as a function of condition. The plots are time-locked at 1 s before the offset of the target sentences, which corresponds to the time locking used in the modeling of the event-related fMRI analysis.

conducted. Following the method proposed by Bosch (2000), Z maps were calculated for each condition and each participant, and nonlinearly normalized as described above. The peaks for six ROIs considered were identical to those taken for the time course diagrams. In a spherical region with a radius of 6 mm around the peak, the Z values were averaged. Figure 6 shows the resulting means across participants, as a function of ROI and condition.

In the overall ANOVA with the within-subject factor ROI, Coherence, and Part, the main effects of ROI ($F(5,40) = 12.0$, $P < 0.0001$) and Coherence ($F(1,8) = 5.5$, $p < 0.05$) were reliable. In contrast, there was no difference between the two parts of the experiment ($F(1,8) < 1$). Interactions between Coherence and Part ($F(1,8) = 9.4$, $P < 0.05$) and between ROI and Part ($F(5,40) = 2.7$, $P < 0.05$) once more confirm that the FMC activations varied systematically with the instructions and the materials. Most importantly, the

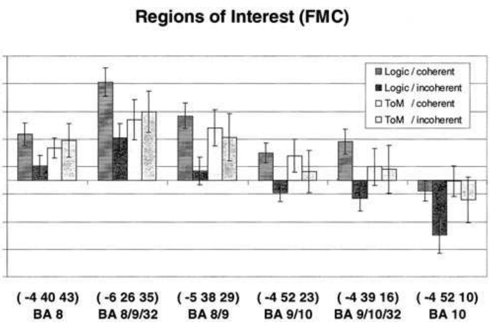


FIG. 6. Means and standard errors of the z values of the six regions of interest, shown for both parts of the experiment as a function of coherence. The regions shown are A–F (cf. Fig. 4) from left to right.

interaction between Coherence and ROI ($F(5,40) = 2.7$, $P < 0.05$) and the three-way interaction between Coherence, Part, and ROI indicated that the effects of coherence were reflected in different regions of the FMC, dependent on the type of processing ($F(3,24) = 7.4$, $P < 0.01$). Planned pairwise comparisons, evaluating the effect of Coherence for each part and each ROI separately, yielded highly significant effects for all six ROIs during Part A (F's > 13.0, P's < 0.01). During Part B, coherence had no corresponding effect (F's < 1). This analysis confirms that across the group of subjects, the effects of coherence during Part A were stable and generalizable.

DISCUSSION

In the present study we replicated the finding of the frontomedian cortex to be involved during Theory-of-Mind processes. Incoherent as well as coherent trials elicited FMC activation during Part B, confirming that ToM processing was sufficient for engaging the FMC. This result was also reflected in stronger FMC activation during Part B than during Part A. Supporting our hypotheses, the data also provided evidence for the claim that ToM processes are not necessary for activating this region, but that they are an instantiation of a more general cognitive process subserved by the FMC. During Part A the comprehension of coherent, plausible sentence pairs elicited frontomedian activation even though ToM processes were minimized. This influence of coherence during Part A only, but not during Part B, was also confirmed in a reliable interaction.

The behavioral data confirmed that the participants heeded the instructions; i.e., they used different processing strategies for the two parts of the experiment. Most importantly, the responses mostly coincided with the coherence of the sentence pairs during the Logic instructions. Thus, the comparison between coherent and incoherent trials tapped similar processes as in the previous study (Ferstl and von Cramon, 2001). Furthermore, it took longer to make a decision about whether the sentence pair described a comprehensible behavior (Part B) than about whether there was a logical connection (Part A). During the debriefing, most participants' self-reports confirmed that they attempted to identify with the protagonists. These results show that the different instructions for the sentence pairs with and without human protagonists had the intended effect of excluding ToM processes in Part A and inducing them in Part B.

The activations found in the contrast of all language trials against the control task replicate many previous studies on language processing in context (e.g., Mazoyer *et al.,* 1993; Ferstl and von Cramon, 2001). In particular, the bilateral temporal activation, including anterior and posterior parts of the superior temporal sulcus, as well as the temporoparietal junction area, was expected. The absence of a larger left-sided lateral prefrontal involvement in this contrast was probably due to the auditory presentation, in which both the language and the pseudo-word trials shared phonological processes. In the following, we focus on an interpretation of the two regions in the median wall of the left hemisphere.

It has been proposed that frontomedian regions are active during rest or during less attention-demanding tasks, and that more taxing cognitive processes yield deactivation or attenuation of these regions (Gusnard *et al.,* 2001, Raichle *et al.,* 2001). Gusnard *et al.* (2001) provided evidence for the argument that more ventral regions follow this pattern indeed (BA11), but that in their study, the more dorsal regions (BA 8/9/32) were positively activated by self-referential processing (i.e., a pleasantness judgment). In our study, we could show for six peaks of activation—differentially sensitive to the experimental condition—that a similar dissociation holds. In five of the six regions of interest, the BOLD contrast increased for the coherent conditions, whereas in the most ventral area, it decreased for the incoherent condition. As Binder *et al.* (1999) argued, the FMC activation found during rest conditions possibly reflects self-guided thought processes in the absence of a task. The fact that, in our study, activation increased and that these increases were dependent on the experimental condition confirms that the FMC plays a role not only as an area taking part in the representation of a "default state" (Raichle *et al.,* 2001), but also during the performance of tasks requiring a self-guided, nonautomatic cognitive process, such as ToM and coherence tasks.

Other influential proposals for a specific function of the FMC, besides ToM, include affective processing (Greene *et al.,* 2001; Lane *et al.,* 1997; see also Maddock, 2001, for the proposal of the retrosplenial cortex/PCC being involved in emotion processing). Raichle *et al.* (2001; see also Simpson *et al.,* 2001a,b) see the precuneus as a region constantly monitoring the perceptual input, and the FMC as assessing the emotional significance of the external stimuli. The evidence for these proposals comes from the observation of FMC activation during tasks presumably eliciting emotions. The interpretations of the results are less than conclusive, though. For instance, Greene *et al.* (2001) used a moral judgment on complex moral dilemmas. Thus, the decision task required a variety of subprocesses, including the comprehension of the different texts and the decision taking into account the persons' value systems. It is not obvious that the emotional content was the crucial factor leading to FMC activation, rather than, let's say, inferences based on episodic memory, the ambiguity and uncertainty of the decision, or the more complex problem-solving process. Moreover, the data of the present study show that FMC activation was elicited by a condition in which emotional processes might play a supporting role at best. Of course, it is always possible for one experimental condition to be accompanied by stronger emotional involvement than another (see Simpson *et al.,* 2001 a,b, for a similar argument), but it is not clear why this should be the case for the processing of coherent as compared with incoherent sentence pairs with emotionally neutral content.

A second proposal links the observed activations to self-referential processing (Gusnard *et al.,* 2001). Although this concept is closely related, and not easily separable from the affect account (for instance, Lane *et al.,* 1997, studying emotional processes, used an almost identical pleasantness judgment task as Gusnard *et al.,* 2001, studying self-referential processing), it focuses more on the personal relevance of the to-be-processed stimuli than on their emotional content. Vogeley *et al.* (2001) recently reported an experiment in which the self-relevance or self-perspective was systematically varied. The results showed that the FMC was engaged during both ToM and self-relevant processing. Zysset *et al.* (2002) indirectly manipulated self-reference in a study concerned with evaluative judgments. When the verification of factual statements, related to self-referential, episodic memories, was compared with that of semantic, general world knowledge statements, a very similar pattern of FMC/PCC activation was found as in the present study. Interestingly, the FMC activation was even stronger when an evaluative component was added, whereas the reverse was true for the PCC activation. These results indicate that the PCC activation is modulated by retrieval processes (cf. Krause *et al.,* 1999) whereas the FMC activation increases with the inference demands of the task, over and above the self-referential nature of the stimuli. The data reported here confirm that self-reference, as specifically induced by the ToM instructions during Part B, leads to strong involvement of the FMC/PCC regions. However, as pointed out before, self-referentiality of the stimulus content is not necessary for eliciting this activation, because it was also observed for coherent trials during Part A.

To delineate the role of the FMC, not only for coherence and ToM processes, we postulate a more general common component shared by the multitude of other tasks shown to engage this brain region. How can we describe this underlying function? Neuropsychological descriptions of patients with ischemic infarctions in the cortical territory of the anterior cerebral artery focus on changes in drive and volition (Marin, 1991). Patients with bilateral lesions suffer from akinetic mutism, characterized by a complete lack of mental animation (Damasio, 1999). Even in unilateral cases, patients often show a lack of ideas, they fail to initiate actions, or they neglect to carry out plans. With respect to language, patients with left-sided lesions have a reduction of spontaneous speech. Based on these observations was our previous interpretation of the FMC being involved whenever the initiation and maintenance of nonautomatic cognitive processes are required, independent of the domain and independent of the content (Ferstl and von Cramon, 2001). This assumption fits very well with the adaptive coding model (Duncan, 2001), postulating that a fundamental principle of prefrontal cortex might be its potential to be driven by many different kinds of input.

We now go one step further and link this domain-independent component to a concept from philosophy. Metzinger (2000) described a component of subjective experience, the so-called *self-model,* as a “transient computational module, episodically activated by the system in order to regulate its interaction with the environment” (p. 290). Not to be intermixed with the self-concept, encompassing the knowledge about the self, the self-model is continuously updated by integrating external and internal information in a dynamic way. Thus, the engagement of the self-model is dependent on both the stimulus features and the cognitive, motivational, and volitional status of the system. Speculative at this point, the self-model seems a promising account for the domain-general functionality of the FMC (see also Vogeley *et al.,* 1999).

Researchers are just beginning to collect data directly targeting FMC functionality and to develop theoretical accounts thereof (e.g., Gusnard *et al.,* 2001). Thus, a falsification of our still underspecified proposal seems difficult. Any experimental task requires some integration of the internal world with the external stimulus, with the focus being shifted from one to the other by both task requirements and stimulus properties. And indeed, FMC activation has been shown to be a matter of degree, rather than all-or-none (e.g., Castelli *et al.,* 2000; Greene *et al.,* 2001; Zysset *et al.,* 2002), and it cannot directly be predicted by stimulus complexity or task difficulty. Even during rest conditions FMC activation has been observed—similar to that elicited by knowledge-based semantic processing—indicating the initiation and maintenance of self-guided thought (Binder *et al.,* 1999). Consequently, conclusions about which tasks will and which tasks will not engage the FMC are highly dependent on the control task. For instance, the rather demanding memory task used by Fletcher *et al.* (1995; Gallagher *et al.,* 2000) as a control task (“unrelated sentences”) requires self-guided encoding strategies and is thus predicted to lead to comparable FMC activation as the comprehension of the “physical stories”—a condition whose inference requirements make FMC engagement likely. Further research is needed for directly investigating these issues. We are confident that the proposal sketched above will prove useful for generating testable hypotheses which will, in turn, aid in refining the theoretical understanding of the functional neuroanatomy of the FMC.

CONCLUSIONS

In this study we presented evidence for a domain-independent, general-purpose functionality of the FMC. Based on previous observations of involvement of FMC during coherence processes as well as during ToM, emotional, or self-referential processing, we designed a study in which coherence and ToM were independently varied. The result of activation of both the FMC and an accompanying region in the PCC ToM processing, but also whenever sentence pairs were coherently related, suggests that inference processes are sufficient for engaging median cortical structures. We believe that the function of the FMC includes the initiation and maintenance of nonautomatic cognitive processes. However, further research is needed to obtain a better understanding of this brain area's functionality. A future theory must include a more thorough delineation of the subdivisions of FMC, as well as a consideration of volitional and motivational aspects of cognition.

REFERENCES

Baron-Cohen, S., Leslie, A. M., and Frith, U. 1985. Does the autistic child have a 'theory of mind'? *Cognition* **21:** 37–46.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., and Williams, S. C. R. 1999. Social intelligence in the normal and autistic brain: An fMRI study. *Eur. J. Neurosci.* **11:** 1891–1898.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., and Cox, R. W. 1999. Conceptual processing during the conscious resting state: A functional MRI study. *J. Cogn. Neurosci.* **11:** 80–93.

Bosch, V. 2000. Statistical analysis of multi-subject fMRI data: The assessment of focal activations. *J. Magn. Reson. Imaging* **11:** 61–64.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S. J., and Frith, C. D. 1994. The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain* **117:** 1241–1253.

Brothers, L. 1990. The social brain: A project for integrating primate behaviour and neurophysiology in a new domain. *Concepts Neurosci.* **1:** 27–51.

Brunet, E., Sarfati, Y., Hardy-Baylé, M. C., and Decety, J. 2000. A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage* **11:** 157–166.

Castelli, F., Happé, F., Frith, U., and Frith, C. 2000. Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* **12:** 314–325.

Channon, S., and 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.

Damasio, A. R. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* Harcourt Brace, New York.

Duncan, J. 2001. An adaptive coding model of neural funciton in prefrontal cortex. *Nat. Rev. Neurosci.* **2:** 820–829.

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

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

Fine, C., Lumsden, J., and Blair, R. J. R. 2001. Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. *Brain* **124:** 287–298.

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

Friston, K. J. 1994. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2:** 189–210.

Frith, C. D., and Frith, U. 1999. Interacting minds, a biological basis. *Science* **286:** 1692–1695.

Frith, U. 1989. *Autism: Explaining the Enigma.* Blackwell, Oxford.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. 2000. Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia* **38:** 11–21.

Goel, V., Gold, B., Kapur, S., and Houle, S. 1997. The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport* **8:** 1305–1310.

Goel, V., Grafman, J., Sadato, N., and Hallett, M. 1995. Modeling other minds. *NeuroReport* **6:** 1741–1746.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. 2001. An fMRI investigation of emotional engagement in moral judgment. *Science* **293:** 2105–2108.

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. 2001. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc. Natl. Acad. Sci. USA* **98:** 4259–4264.

Happé, F. 2000. Parts and wholes, meaning and minds: Central coherence and its relation to theory of mind. In *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience* (S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen, Eds.), 2nd ed., pp. 203–221. Oxford Univ. Press, Oxford.

Happé, F., Brownell, H. H., and Winner, E. 1999. Acquired 'theory of mind' impairments following stroke. *Cognition* **70:** 211–240.

Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C. Dolan, R., Frackowiak, R., and Frith, C. 1996. 'Theory of mind' in the brain: Evidence from a PET scan study of Asperger syndrome. *NeuroReport* **8:** 197–201.

Jolliffe, T., and Baron-Cohen, S. 1999. A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition* **71:** 149–185.

Jolliffe, T., and Baron-Cohen, S. 2000. Linguistic processing in high-functioning adults with autism of Asperger's syndrome: Is global coherence impaired? *Psychol. Med.* **30:** 1169–1187.

Kintsch, W. 1998. *Comprehension: A Paradigm for Cognition.* Cambridge University Press, Cambridge.

Krause, B. J., Schmidt, D., Mottaghy, F. M., Taylor, J., Halsband, U., Herzog, H., Tellmann, L., and Müller-Gärtner, H.-W. 1999. Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates: A PET-study. *Brain* **122:** 255–263.

Lane, R. D., Fink, G. R., Chau, P. M. L., and Dolan, R. J. 1997. Neural activation during selective attention to subjective emotional responses. *NeuroReport* **8:** 3969–3972.

Lohmann, G. 1998. *Volumetric Image Analysis.* Wiley, Chichester.

Lohmann, G., Müller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., and von Cramon, D. Y. 2001. Lipsia: A new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput. Med. Imaging Graphics* **25:** 449–457.

Maddock, R. J. 2001. The retrosplenial cortex and emotion: New insights from functional neuroimaging of the human brain. *Trends Neurosci.* **22:** 310–316.

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

Marin, R. W. 1991. Apathy: A neuropsychiatric syndrome. *J. Neuropsychiatry Clin. Neurosci.* **3:** 243–254.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., and Mehler, J. 1993. The cortical representation of speech. *J. Cogn. Neurosci.* **5:** 467–479.

Metzinger, T. 2000. The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. In *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (T. Metzinger, Ed.), pp. 285–306. MIT Press, Cambridge, MA.

Nichelli, P., Grafman, J., Pietrini, P., Clark, K., Lee, K. Y., and Miletich, R. 1995. Where the brain appreciates the moral of a story. *NeuroReport* **6:** 2309–2313.

Norris, D. G. 2000. Reduced power multi-slice MDEFT imaging. *J. Magn. Reson. Imaging* **11:** 445–451.

Premack, D., and Woodruff, G. 1978. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* **1:** 516–526.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. 2001. A default mode of brain function. *Proc. Nat. Acad. Sci. USA* **98:** 676–682.

Rowe, A. D., Bullock, P. R., Polkey, C. E., and Morris, R. G. 2001. 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain* **124:** 600–616.

Siegal, M., Carrington, J., and Radel, M. 1996. Theory of mind and pragmatic understanding following right hemisphere damage. *Brain Lang.* **53:** 40–50.

Simpson, J. R., Drevets, W. C., Snyder, A. Z., Gusnard, D. A., and Raichle, M. E. 2001a. Emotion-induced changes in human medial prefrontal cortex: II. During anticipatory anxiety. *Proc. Natl. Acad. Sci. USA* **98:** 688–693.

Simpson, J. R., Snyder, A. Z., Gusnard, D. A., and Raichle, M. E. 2001b. Emotion-induced changes in human medial prefrontal cortex: I. During cognitive task performance. *Proc. Natl. Acad. Sci.* **98:** 683–687.

Stone, V. E., Baron-Cohen, S., and Knight, R. T. 1998. Frontal lobe contributions to theory of mind. *J. Cogni. Neurosci.* **10:** 640–656.

Stuss, D. T., Gallup, G. G., and Alexander, M. P. 2001. The frontal lobes are necessary for 'theory of mind'. *Brain* **124:** 279–286.

Talairach, J., and Tournoux, P. 1988. *Coplanar Stereotaxic Atlas of the Human Brain*. Thieme, New York.

Thirion, J. P. 1998. Image matching as a diffusion process: An analogy with Maxwell's demons. *Med. Image Analy.* **2:** 243–260.

Ugurbil, K., Garwood, M., Ellermann, J., Hendrich, K., Hinke, R., Hu, X., Kim, S.-G., Menon, R., Merkle, H., Ogawa, S., and Salmi, R. 1993. Imaging at high magnetic fields: Initial experiences at 4T. *Magn. Reson. Q.* **9:** 259.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. 2001. Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage* **14:** 170–181.

Vogeley, K., Kurthen, M., Falkai, P., and Maier, W. 1999. Essential functions of the human sellf model are implemented in the prefrontal cortex. *Consciousness Cognition* **8:** 343–363.

Wellman, H. M. 1993. Early understanding of mind: The normal case. In *Understanding Other Minds: Perspectives from Autism* (S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen, Eds.), pp. 10–39. Oxford Univ. Press, Oxford.

Williams, S. M., and Goldman-Rakic, P. S. 1998. Widespread origin of the primate mesofrontal dopamine system. *Cereb. Cortex* **8:** 321–345.

Zysset, S., Huber, O., Ferstl, E. C., and von Cramon, D. Y. 2002. The anterior fronto-median cortex and evaluative judgment: An fMRI study. *NeuroImage* **15:** 983–991.

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

Evelyn C. Ferstl[1,2], **Mike Rinck**[3], **and D. Yves von Cramon**[1,2]

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

Journal of Cognitive Neuroscience 17:5, pp. 724–739

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, Lambon-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, Diaz, & 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 ones 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-

Table 1. Sample Stories for the Two Story Types

Chronological	*Emotional*
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.

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.

tion model, a topic that currently receives much attention in studies on discourse processing (e.g., Zwaan & Singer, 2003). In addition, the salience of the inconsistencies and the information aspects were also selected for their neuroanatomical separability. Processing of temporal, sequential, or chronological information usually engages frontoparietal regions, including the premotor (Schubotz & von Cramon, 2001), prefrontal (Sirigu et al., 1998), or parietal (Walsh, 2003) cortices. In contrast, the processing of emotional information is known to engage medial prefrontal, orbitofrontal, and limbic structures, including the amygdalae (Dolan, 2003). Whether listening to stories suffices for eliciting activation in either of these networks was an open question.

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, Ferstl, & von Cramon, 2002; Ferstl & von Cramon, 2001, 2002).

Taken together, the design of the experiment included two factors: Consistency and Story Type. To ensure participants' attention to the story information, a consistency judgment after auditory presentation of the stories was required.

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 spin–echo 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 ($p > .11$).

Table 2. Behavioral Data

	Temporal Stories		*Emotional Stories*	
	Consistent	*Inconsistent*	*Consistent*	*Inconsistent*
Errors (%)				
Mean	18.8	34.4	10.6	37.5
SD	16.0	25.9	16.3	19.4
Reaction times (msec)				
Mean	990	799	837	906
SD	430	267	222	350

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

Event-Related Inconsistency Detection	*Area*	*Size*	*Max*	*Side*	*X*	*Y*	*Z*	*F*
Emotional > chronological	Ventromedial prefrontal cortex/supraorbital sulcus	945	4.19	R	2	20	−12	15.9**
	Extended amygdala	371	3.94	L	−22	2	−12	11.8*
	Pons	823	4.63	L	−8	−25	−27	20.9**
Chronological > emotional	Precuneus	2,927	4.81	L	−8	−61	44	21.9**
	Precuneus	354	4.01	R	4	−49	41	27.5***
	Intraparietal sulcus	474	3.72	L	−40	−55	50	21.4**
	Intraparietal sulcus/temporoparietal junction area	1,804	4.11/3.93	R/R	35/43	−55/−53	44/26	31.7***
	Parieto-occipital	790	4.10	L	−34	−82	35	17.5**
	Frontopolar	1,677	4.22	L	−23	47	12	32.8***
	Frontopolar	1,179	4.09	R	23	41	15	26.5***
	Anterior superior insula	636	3.89	L	−29	14	15	17.6**
Inconsistent > consistent	Anterior temporal/anterior inferior insula/posterior orbital gyrus	523	3.79	R	40	−1	−12	18.2**

$^*p < .01$; $^{**}p < .001$; $^{***}p < .0001$

$Z > 3.09$; $p < .001$, uncorrected; extent threshold, 324 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).

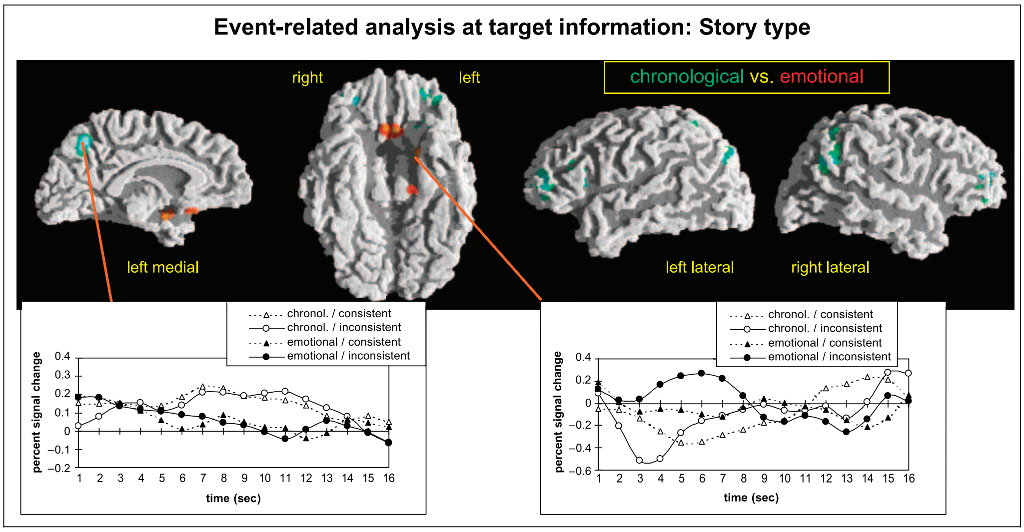


Figure 1. fMRI results for the main effect of story type in the event-related analysis, time locked at the target information. All activations are superimposed on a white-matter rendition of the participants' average anatomy. For descriptive details see Table 3. Areas in red/yellow were more active during 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.

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

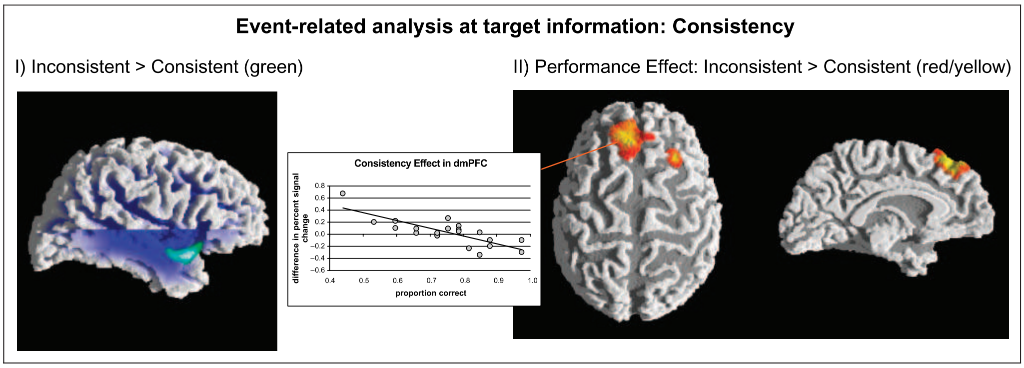


Figure 2. fMRI results for the main effect of consistency in the event-related analysis, time locked at the target information (cf. Table 3). (I) The area in green was more active when an inconsistency was heard. The temporal lobe is cut at $x = 43$. (II) Parametric second-level analysis of the consistency 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 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 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 ($r = -.80$). Even when an extreme value is removed, the correlation is still $r = -.69, p < .001$. 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 = 8.33 sec, E = 8.37 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 gyri 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-

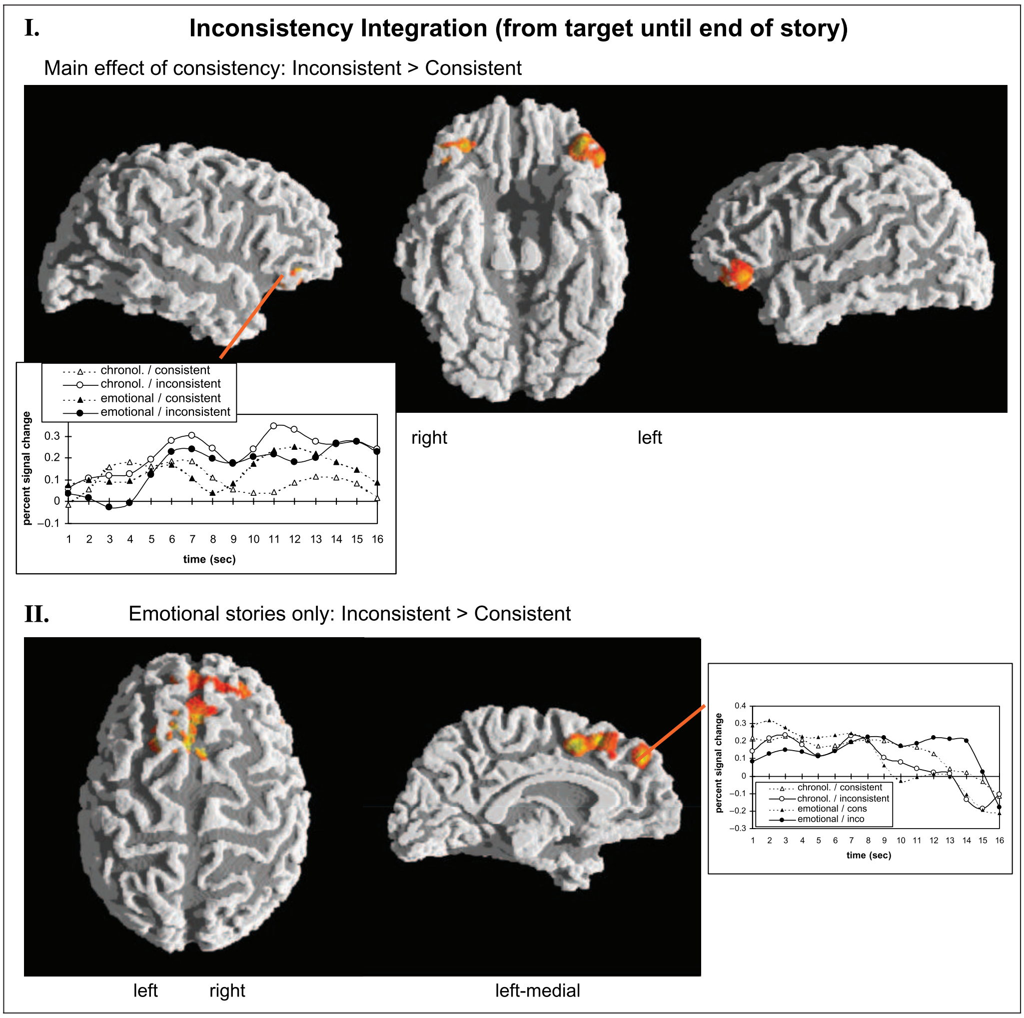


Figure 3. fMRI results for the epoch-related analysis, including the period from the target information until the end of the story (cf. Table 4). (I) In the upper panel, the main effect of consistency is shown. The time course diagram shows an increase in percent signal change for both inconsistent conditions. (II) As an illustration of the interaction between Consistency and Story Type, this panel displays the contrast comparing inconsistent to consistent trials for the emotional stories only. The diagram shows the time course of activation in the anteriormost peak. In all four conditions, the dmPFC is highly activated during the last few seconds of the story. The differences concern the point at which this activation levels off. The earliest point is observed for the consistent emotional stories, the latest for the inconsistent emotional stories, with the two chronological conditions being in between.

trast. 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

Table 4. Regions of Activation for the Epoch-Related Analysis Based on the Period from the Presentation of the Target Information until the End of the Story

Epoch-related Inconsistency Integration	*Area*	*Size*	Z_{max}	*Side*	*X*	*Y*	*Z*	*F*
Chronological > emotional	Precuneus	497	4.10	L	−8	−52	59	18.2**
	Precuneus	2,457	4.71	R	4	−64	53	34.3***
	Cuneus	327	4.04	R	11	−88	−9	19.2**
Inconsistent > consistent	Orbital part of IFG/lateral orbital gyrus	1,677	4.32	L	−50	32	−6	18.1**
	Orbital part of IFG/lateral orbital gyrus	539	4.75	R	35	35	−3	14.3*
Consistency × Story Type interaction	dmPFC	326	3.86	L	−1	44	38	15.1**
Chronological stories only: inconsistent > consistent	Orbital part of IFG/lateral orbital gyrus	554	3.84	L	−32	35	−6	12.6*
	Orbital part of IFG/lateral orbital gyrus	425	3.54	R	35	35	−3	10.8*
Emotional stories only: inconsistent > consistent	dmPFC, BA 6/8, rostral cingulate zone	2,890	4.69	L	−4	5	44	26.1***
	dmPFC, BA 8/9	1,285	3.97	L	−2	44	38	28.3***
	Junction area: IFG/Precentral gyrus	418	3.94	R	44	20	21	15.3**

IFG = inferior frontal gyrus; dmPFC = dorso-medial prefrontal cortex; *$p < .01$; **$p < .001$; ***$p < .0001$.

$Z > 3.09$; $p < .001$, uncorrected; extent threshold, 324 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. In the last column, the *F* values from the second-level analysis of the region are presented [$df = (1,19)$].

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 precentral 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.

Right Anterior Temporal Lobe

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-hemisphere 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, this intermediate step between the ongoing input and building a situation model representation becomes more difficult and requires increased right-hemispheric involvement.

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 (IPS), 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 (vlPFC; BA 47/11) during the integration of inconsistent information clearly differed from the frontopolar activations found in the event-related analysis. Left vlPFC engagement, close to the most lateral left-sided peak found here, was also reported in two recent studies on language processing in context. Baumgärtner, 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 vlPFC 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 vlPFC 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 × 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* = .9) 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 × 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 × 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 × 3 × 3 mm. The data were then analyzed using the general linear model (Friston, 1994), based on the within-subjects design Consistency × 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 mm^3, 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 mm^3 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 mm^3. 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 2-2004-117T2.

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

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience, 2,* 1032–1037.

Baumgärtner, A., Weiller, C., & Büchel, C. (2002). Event-related fMRI reveals cortical sites involved in contextual sentence integration. *Neuroimage, 16,* 736–745.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V. P., Karni, A., Printer, A., Braun, A., Lalwani, A., Rauschecker, J. P., Turner, R., & Neville, H. (1997). Sentence reading—A functional MRI study at 4 Tesla. *Journal of Cognitive Neuroscience, 9,* 664–686.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123,* 2189–2202.

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.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25,* 151–188.

Bosch, V. (2000). Statistical analysis of multi-subject fMRI data: The assessment of focal activations. *Journal of Magnetic Resonance Imaging, 11,* 61–64.

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *Neuroimage, 14,* 1416–1423.

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: Marcel Dekker.

Caplan, R., & Dapretto, M. (2001). Making sense during conversation: An fMRI study. *NeuroReport, 12,* 3625–3632.

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. S. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain, 126,* 1193–1201.

Damasio, A. (1995). *Descartes' error.* London: Picador.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature, 380,* 499–505.

Davidson, R. J, & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences, 3,* 11–21.

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation—A possible prelude to violence. *Science, 289,* 591–594.

Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia, 41,* 127–138.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain imaging evidence. *Science, 284,* 970–974.

Dolan, R. J. (2002). Emotion, cognition and behavior. *Science, 298,* 1191–1194.

Elliott, R., Rubinsztein, J. S., Sahakian, B. J., & Dolan, R. J. (2000). Selective attention to emotional stimuli in a verbal go/no-go task: An fMRI study. *NeuroReport, 11,* 1739–1744.

Ferstl, E. C. (2001). Learning from text. In N. J. Smelser, P. B. Baltes (Series Eds.) & W. Kintsch (Vol. Ed.), *International encyclopedia of the social and behavioral sciences: Vol. 3.13, Cognitive psychology and cognitive science.* Amsterdam: Elsevier.

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., Guthke, T., & von Cramon, D. Y. (2002). Text comprehension after brain injury: Left prefrontal lesions affect inference processes. *Neuropsychology, 16,* 292–308.

Ferstl, E. C., & Kintsch, W. (1999). Learning from text: Structural knowldege assessment in the study of discourse comprehension. In H. Oostendorp & S. Goldman (Eds.), *The construction of mental models during reading,* (pp. 247–277). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Ferstl, E. C., & von Cramon, D. Y. (2002). What does the fronto-medial cortex contribute to language processing: Coherence or Theory of Mind? *Neuroimage, 17,* 1599–1612.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain, 120,* 1779–1791.

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.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33,* 636–647.

Friston, K. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2,* 189–210.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, 358,* 459–473.

Gaddy, M. L., van den Broek, P., &, Sung, Y.-C. (2001). The influence of text cues on the allocation of attention during reading. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholingusitic aspects* (pp. 89–110). Amsterdam: John Benjamins.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of ''theory of mind'' in verbal and nonverbal tasks. *Neuropsychologia, 38,* 11–21.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent characters' emotional states? *Cognition & Emotion, 6,* 89–111.

Gernsbacher, M. A., & Kaschak, M. P. (2003) Neuroimaging studies of language production and comprehension. *Annual Review of Psychology, 54,* 16.1–16.24.

Glenberg, A. M., Havas, D., Becker, R., & Rinck, M. (2005). Grounding language in bodily states: The case for emotion. In R. Zwaan & D. Pecher (Eds.), *The grounding of cognition: The role of perception and action in memory, language, and thinking* (pp. 115–118). Cambridge: Cambridge University Press.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review, 9,* 558–565.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport, 8,* 1305–1310.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48,* 163–189.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293,* 2105–2108.

Gruber, O. (2001). Effects of domain-specific interference on brain activation associated with verbal working memory task performance. *Cerebral Cortex, 2001,* 1047–1055.

Gruber, O., & von Cramon, D. Y. (2003). The functional neuroanatomy of human working memory revisited: Evidence from 3T fMRI studies using classical domain-specific interference tasks. *Neuroimage, 19,* 797–809.

Hough, M. S. (1990). Narrative comprehension in adults with right and left hemisphere brain-damage: Theme organization. *Brain and Language, 38,* 253–277.

Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *NeuroReport, 12,* 1749–1752.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malvade, K., Beattie, B., Leon, A. C., & Stern, E. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Sciences, U.S.A., 96,* 10456–10459.

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

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

Kintsch, W., & van Dijk, T. A. (1978) Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Kircher, T. T. J., Brammer, M., Andreu, N. T., Williams, S. C. R., & McGuire, P. K. (2001). Engagement of right temporal cortex during processing of linguistic context. *Neuropsychologia, 39,* 798–809.

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence form neuroimaging and neuropsychology. *Progress in Neurobiology, 72,* 341–372.

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

Lohmann, G. (1998). *Volumetric image analysis.* Chichester, UK: Wiley.

Lohmann, G., Müller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L, & von Cramon, D. Y. (2001). Lipsia: A new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics, 25,* 449–457.

Long, D. L., & Baynes, K. (2002). Discourse representation in the two cerebral hemispheres. *Journal of Cognitive Neuroscience, 14,* 228–242.

Luan Phan, K., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16,* 331–348.

Maddock, R. J., & Buonocore, M. H. (1997). Activation of left posterior cingulate gyrus by the auditory presentation of threat-related words—An fMRI study. *Psychiatry Research: Neuroimaging, 75,* 1–14.

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.

Maratos, E. J., Dolan, R. J, Morris, J. S., Henson, R. N A., & Rugg, M. D. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia, 29,* 910–920.

Mason, R. A., & Just, M. A. (2004). How the brain processes causal inferences in text: A theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychological Science, 15,* 1–7.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience, 5,* 467–479.

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

Meyer, M., Friederici, A. D., & von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: Event related fMRI reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research, 9,* 19–33.

Nathaniel-James, D. A., & Frith, C. D. (2002). The role of the dorsolateral prefrontal cortex: Evidence from the effects of contextual constraint in a sentence completion task. *Neuroimage, 16,* 1094–1102.

Nobre, A. C., Coull, J. T., Frith, C. D., & Mesulam, M. M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nature Neuroscience, 2,* 11–12.

Norris, D. G. (2000). Reduced power multi-slice MDEFT imaging, *Journal of Magnetic Resonance Imaging, 11,* 445–451.

Norris, D. G., Zysset, S., Mildner, T., & Wiggins, C. J. (2002). An investigation of the value of spin–echo based fMRI using a Stroop colour–word matching task and EPI at 3 Tesla. *Neuroimage, 15,* 719–726.

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

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in text: What readers believe versus what they read. *Psychological Science, 3,* 229–235.

Petrides, M., Alivisatos, B., & Frey, S. (2002). Differential activation of the human orbital, mid-vetrolateral, and mid-dorsolateral prefrontal cortex during the processing of visual stimuli. *Proceedings of the National Academy of Sciences, U.S.A., 99,* 5649–5654.

Rao, S. M. Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience, 4,* 317–323.

Rinck, M., Gámez, E., Díaz, J. M., & de Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory and Cognition, 31,* 77–86.

Rinck, M., Hähnel, A., & Becker, G. (2001). Using temporal information to construct, update, and retrieve situation models of narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 67–80.

Rinck, M., & Weber, U. (2003). Who when where: An experimental test of the event-indexing model. *Memory and Cognition, 31,* 1284–1292.

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.

Schubotz, R. I. & von Cramon, D. Y. (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cerebral Cortex, 11,* 210–222.

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

Sirigu, A., Cohen, L., Zalla, T., Pradat-Diehl, P., van Eeckhout, P., Grafman, J., & Agid, Y. (1998). Distinct frontal regions for processing sentence syntax and story grammar. *Cortex, 34,* 771–778.

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.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 121,* 153–156.

Tabert, M. H., Borod, J. C., Tang, C. Y., Lange, G., Wie, T. C., Johnson, R., Nusbaum, A. O., & Buchsbaum, M. S. (2001). Differential amygdala activation during emotional decision and recognition memory tasks using unpleasant words: An fMRI study. *Neuropsychologia, 39,* 556–573.

Talairach, J., & Tournoux, P. (1988). *Coplanar stereotaxic atlas of the human brain.* New York: Thieme.

Ugurbil, K., Garwood, M., Ellermann, J., Hendrich, K., Hinke, R., Hu, X., Kim, S.-G., Menon, R., Merkle, H., Ogawa, S., & Salmi, R. (1993). Imaging at high magnetic fields: Initial experiences at 4T. *Magnetic Resonance Quarterly, 9,* 259.

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage, 14,* 1387–1401.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2003). Predicting events of varying probability: Uncertainty investigated by fMRI. *Neuroimage, 19,* 271–280.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7,* 483–488.

Whalen, P. J., Bush, G., McNally, R. J., Wilhelm, S., McInerney, S. C., Jenike, M. A., & Rauch, S. L. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry, 44,* 1219–1228.

Worsley, K. J., & Friston K. J. (1995). Analysis of fMRI time-series revisited–again. *Neuroimage, 2,* 173–181.

Zalla, T., Phipps, M., & Grafman, J. (2002). Story processing in patients with damage to the prefrontal cortex. *Cortex, 38,* 215–231.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6,* 292–297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 386–397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

Zwaan, R., A., & Singer, M. (2003). Text comprehension. In Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (Eds.), *Handbook of discourse processes,* (pp. 83–122). Mahwah, NJ: Lawrence Erlbaum.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin and Review, 10,* 954–958.

Zysset, S., Huber, O., Ferstl, E. C., & von Cramon, D. Y. (2002). The anterior fronto-medial cortex and evaluative judgment: An fMRI study. *Neuroimage, 15,* 983–991.

Zysset, S., Huber, O., Samson, A., Ferstl, E. C., & von Cramon, D. Y. (2003). Functional specialization within the anterior medial prefrontal cortex: A functional magnetic resonance imaging study with human subjects. *Neuroscience Letters, 335,* 183–186.

Declaration of Authorship
Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Habilitationsschrift selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie die wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Leipzig, 15. September 2005 Dr. Evelyn C. Ferstl

Curriculum Vitae

Dr. Evelyn C. Ferstl

Ausbildung

1994	Promotion (Ph. D.) in kognitiver Psychologie
1991	M. A. (Masters of Arts) in kognitiver Psychologie
1988 - 1994	Studium der Kognitionspsychologie an der University of Colorado at Boulder, USA
1987	Hauptdiplom im Fach Mathematik
1980 - 1987	Studium der Mathematik an der Ludwig-Maximilians-Universität, München

Wissenschaftlicher Werdegang

seit 2003	Wissenschaftliche Mitarbeiterin an der Tagesklinik für kognitive Neurologie, Universität Leipzig (DFG-Projekt)
1999 - 2003	Wissenschaftliche Mitarbeiterin am Max-Planck-Institut für neuropsychologische Forschung in Leipzig Abteilung Neurologie
1998 - 1999	Habilitationsstipendiatin des Sächsischen Ministeriums für Wissenschaft und Kunst
1995 –1997	Wissenschaftliche Mitarbeiterin am Max-Planck-Institut für neuropsychologische Forschung in Leipzig Abteilung Neuropsychologie
1994 – 1995	Post-doktoraler Forschungsaufenthalt an der Northwestern University, Evanston, Illinois, USA
1988 - 1994	Lehr- und Forschungsassistentin an der Fakultät für Psychologie an der University of Colorado at Boulder, Colorado, USA
1987 – 1988	Wissenschaftliche Assistentin an der mathematischen Fakultät der Ludwig-Maximilians-Universität, München
1985	Studentische Hilfskraft am Max-Planck-Institut für psychologische Forschung, München
1983 - 1987	Tutorin und Übungsleiterin an der mathematischen Fakultät der Ludwig-Maximilians-Universität, München

MPI Series in Human Cognitive and Brain Sciences:

1 Anja Hahne

Charakteristika syntaktischer und semantischer Prozesse bei der auditiven Sprachverarbeitung: Evidenz aus ereigniskorrelierten Potentialstudien

2 Ricarda Schubotz

Erinnern kurzer Zeitdauern: Behaviorale und neurophysiologische Korrelate einer Arbeitsgedächtnisfunktion

3 Volker Bosch

Das Halten von Information im Arbeitsgedächtnis: Dissoziationen langsamer corticaler Potentiale

4 Jorge Jovicich

An investigation of the use of Gradient- and Spin-Echo (GRASE) imaging for functional MRI of the human brain

5 Rosemary C. Dymond

Spatial Specificity and Temporal Accuracy in Functional Magnetic Resonance Investigations

6 Stefan Zysset

Eine experimentalpsychologische Studie zu Gedächtnisabrufprozessen unter Verwendung der funktionellen Magnetresonanztomographie

7 Ulrich Hartmann

Ein mechanisches Finite-Elemente-Modell des menschlichen Kopfes

8 Bertram Opitz

Funktionelle Neuroanatomie der Verarbeitung einfacher und komplexer akustischer Reize: Integration haemodynamischer und elektrophysiologischer Maße

9 Gisela Müller-Plath

Formale Modellierung visueller Suchstrategien mit Anwendungen bei der Lokalisation von Hirnfunktionen und in der Diagnostik von Aufmerksamkeitsstörungen

10 Thomas Jacobsen

Characteristics of processing morphological structural and inherent case in language comprehension

11 Stefan Kölsch

Brain and Music
A contribution to the investigation of central auditory processing with a new electrophysiological approach

12 Stefan Frisch

Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen

13 Markus Ullsperger

The role of retrieval inhibition in directed forgetting – an event-related brain potential analysis

14 Martin Koch

Measurement of the Self-Diffusion Tensor of Water in the Human Brain

15 Axel Hutt

Methoden zur Untersuchung der Dynamik raumzeitlicher Signale

16 Frithjof Kruggel

Detektion und Quantifizierung von Hirnaktivität mit der funktionellen Magnetresonanztomographie

17 Anja Dove

Lokalisierung an internen Kontrollprozessen beteiligter Hirngebiete mithilfe des Aufgabenwechselparadigmas und der ereigniskorrelierten funktionellen Magnetresonanztomographie

18 Karsten Steinhauer

Hirnphysiologische Korrelate prosodischer Satzverarbeitung bei gesprochener und geschriebener Sprache

19 Silke Urban

Verbinformationen im Satzverstehen

20 Katja Werheid

Implizites Sequenzlernen bei Morbus Parkinson

21 Doreen Nessler

Is it Memory or Illusion? Electrophysiological Characteristics of True and False Recognition

22 Christoph Herrmann

Die Bedeutung von 40-Hz-Oszillationen für kognitive Prozesse

23 Christian Fiebach

Working Memory and Syntax during Sentence Processing.
A neurocognitive investigation with event-related brain potentials and functional magnetic resonance imaging

24 Grit Hein

Lokalisation von Doppelaufgabendefiziten bei gesunden älteren Personen und neurologischen Patienten

25 Monica de Filippis

Die visuelle Verarbeitung unbeachteter Wörter.
Ein elektrophysiologischer Ansatz

26 Ulrich Müller

Die katecholaminerge Modulation präfrontaler kognitiver Funktionen
beim Menschen

27 Kristina Uhl

Kontrollfunktion des Arbeitsgedächtnisses über interferierende Information

28 Ina Bornkessel

The Argument Dependency Model: A Neurocognitive Approach to Incremental
Interpretation

29 Sonja Lattner

Neurophysiologische Untersuchungen zur auditorischen Verarbeitung
von Stimminformationen

30 Christin Grünewald

Die Rolle motorischer Schemata bei der Objektrepräsentation:
Untersuchungen mit funktioneller Magnetresonanztomographie

31 Annett Schirmer

Emotional Speech Perception: Electrophysiological Insights into the
Processing of Emotional Prosody and Word Valence in Men and
Women

32 André J. Szameitat

Die Funktionalität des lateral-präfrontalen Cortex für die Verarbeitung
von Doppelaufgaben

33 Susanne Wagner

Verbales Arbeitsgedächtnis und die Verarbeitung ambiger Wörter in
Wort- und Satzkontexten

34 Sophie Manthey

Hirn und Handlung: Untersuchung der Handlungsrepräsentation im ventralen
prämotorischen Cortex mit Hilfe der funktionellen Magnet-Resonanz-Tomographie

35 Stefan Heim

Towards a Common Neural Network Model of Language Production and
Comprehension: fMRI Evidence for the Processing of Phonological and
Syntactic Information in Single Words

36 Claudia Friedrich

Prosody and spoken word recognition: Behavioral and ERP correlates

37 Ulrike Lex

Sprachlateralisierung bei Rechts- und Linkshändern mit funktioneller Magnetresonanztomographie

38 Thomas Arnold

Computergestützte Befundung klinischer Elektroenzephalogramme

39 Carsten H. Wolters

Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain

40 Ansgar Hantsch

Fisch oder Karpfen? Lexikale Aktivierung von Benennungsalternativen bei der Objektbenennung

41 Peggy Bungert

Zentralnervöse Verarbeitung akustischer Informationen Signalidentifikation, Signallateralisation und zeitgebundene Informationsverarbeitung bei Patienten mit erworbenen Hirnschädigungen

42 Daniel Senkowski

Neuronal correlates of selective attention: An investigation of electrophysiological brain responses in the EEG and MEG

43 Gert Wollny

Analysis of Changes in Temporal Series of Medical Images

44 Angelika Wolf

Sprachverstehen mit Cochlea-Implantat: EKP-Studien mit postlingual ertaubten erwachsenen CI-Trägern

45 Kirsten G. Volz

Brain correlates of uncertain decisions: Types and degrees of uncertainty

46 Hagen Huttner

Magnetresonanztomographische Untersuchungen über die anatomische Variabilität des Frontallappens des menschlichen Großhirns

47 Dirk Köster

Morphology and Spoken Word Comprehension: Electrophysiological Investigations of Internal Compound Structure

48 Claudia A. Hruska

Einflüsse kontextueller und prosodischer Informationen in der auditorischen Satzverarbeitung: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

49 Hannes Ruge

Eine Analyse des raum-zeitlichen Musters neuronaler Aktivierung im Aufgabenwechselparadigma zur Untersuchung handlungssteuernder Prozesse

50 Ricarda I. Schubotz

Human premotor cortex: Beyond motor performance

51 Clemens von Zerssen

Bewusstes Erinnern und falsches Wiedererkennen:
Eine funktionelle MRT Studie neuroanatomischer Gedächtniskorrelate

52 Christiane Weber

Rhythm is Gonna Get you.
Electrophysiological Markers of Rhythmic Processing in Infants with and without Risk for Specific Language Impairment (SLI)

53 Marc Schönwiesner

Functional Mapping of Basic Acoustic Parameters in the Human Central Auditory System

54 Katja Fiehler

Temporospatial characteristics of error correction

55 Britta Stolterfoht

Processing Word Order Variations and Ellipses: The Interplay of Syntax and Information Structure during Sentence Comprehension

56 Claudia Danielmeier

Neuronale Grundlagen der Interferenz zwischen Handlung und visueller Wahrnehmung

57 Margret Hund-Georgiadis

Die Organisation von Sprache und ihre Reorganisation bei ausgewählten, neurologischen Erkrankungen gemessen mit funktioneller Magnetresonanztomographie – Einflüsse von Händigkeit, Läsion, Performanz und Perfusion

58 Jutta L. Mueller

Mechanisms of auditory sentence comprehension in first and second language: An electrophysiological miniature grammar study

59 Franziska Biedermann

Auditorische Diskriminationsleistungen nach unilateralen Läsionen im Di- und Telenzephalon

60 Shirley-Ann Rüschemeyer

The Processing of Lexical Semantic and Syntactic Information in Spoken Sentences: Neuroimaging and Behavioral Studies of Native and Non-Native Speakers

61 Kerstin Leuckefeld

The Development of Argument Processing Mechanisms in German.
An Electrophysiological Investigation with School-Aged Children and Adults

62 Axel Christian Kühn

Bestimmung der Lateralisierung von Sprachprozessen unter besonderer Berücksichtigung des temporalen Cortex, gemessen mit fMRT

63 Ann Pannekamp

Prosodische Informationsverarbeitung bei normalsprachlichem und deviantem Satzmaterial: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

64 Jan Derrfuß

Functional specialization in the lateral frontal cortex: The role of the inferior frontal junction in cognitive control

65 Andrea Mona Philipp

The cognitive representation of tasks
Exploring the role of response modalities using the task-switching paradigm

66 Ulrike Toepel

Contrastive Topic and Focus Information in Discourse – Prosodic Realisation and Electrophysiological Brain Correlates

67 Karsten Müller

Die Anwendung von Spektral- und Waveletanalyse zur Untersuchung der Dynamik von BOLD-Zeitreihen verschiedener Hirnareale

68 Sonja A.Kotz

The role of the basal ganglia in auditory language processing: Evidence from ERP lesion studies and functional neuroimaging

69 Sonja Rossi

The role of proficiency in syntactic second language processing: Evidence from event-related brain potentials in German and Italian

70 Birte U. Forstmann

Behavioral and neural correlates of endogenous control processes in task switching

71 Silke Paulmann

Electrophysiological Evidence on the Processing of Emotional Prosody: Insights from Healthy and Patient Population

72 Matthias L. Schroeter

Enlightening the Brain –Optical Imaging in Cognitive Neuroscience

73 Julia Reinholz

Interhemispheric interaction in object- and word-related visual areas