

Constraint Parsing und psychologische Plausibilität

11. März 2005

Neuantrag auf Gewährung einer Sachbeihilfe

1 Allgemeine Angaben

Antragsteller:

Name: Wolfgang Menzel, Dr.-Ing.
Dienststellung: Professor für Informatik
Geburtsdatum: 1.2.1953
Nationalität: deutsch
Geschäftszeichen: Me 1472/1-2
Institution: Arbeitsbereich "Natürlichsprachliche Systeme"
Fachbereich Informatik, Universität Hamburg
dienstliche Adresse: Vogt-Kölln-Straße 30
22527 Hamburg
Telefon/Fax: (040) 42 883 - 24 35 / (040) 42 883 - 25 15
E-Mail menzel@informatik.uni-hamburg.de
Privatadresse: Goethestraße 16
25421 Pinneberg
Telefon: (04101) 69 19 45

Tabellarischer Lebenslauf:

1971	Abitur an der Frederic-Joliot-Curie-Oberschule Görlitz
1971 - 1978	Studium und Forschungsstudium Informationstechnik an der TU Dresden
1979	Promotion zum Dr.-Ing. an der Fakultät für Elektrotechnik der TU Dresden, Prädikat: magna cum laude
1978 - 1990	Wissenschaftlicher Mitarbeiter am Zentralinstitut für Sprachwissenschaft der Akademie der Wissenschaften der DDR, Berlin Durchführung von Forschungsarbeiten - zur textgesteuerten Sprachsynthese - zur morphologischen und syntaktischen Analyse des Deutschen - zur Maschinellen Übersetzung - zur Fehlerdiagnose für den Sprachunterricht
1991 - 1992	Mitarbeiter der Forschungsgruppe Computerlinguistik, Berlin Durchführung von Forschungsarbeiten - zur Fehlerdiagnose für den Sprachunterricht - zum Lexikon in integrierten Speech-Language-Systemen
seit 1992	Professor für Informatik, Universität Hamburg Forschungen - zur Architektur sprachverarbeitender Systeme - zur robusten Sprachanalyse - zu Lehrsystemen für den Fremdsprachenunterricht verheiratet, drei Kinder

1.1 Thema

Constraint Parsing und psychologische Plausibilität

1.2 Kennwort

CoPa

1.3 Fachgebiet und Arbeitsrichtung

Kognitionswissenschaft / Künstliche Intelligenz / Computerlinguistik / Psycholinguistik
Parsing, Constraint-Optimierung

1.4 Voraussichtliche Gesamtdauer

Das Forschungsvorhaben soll über eine Dauer von 3 Jahren durchgeführt werden.

1.5 Antragszeitraum

Der Antrag zur Unterstützung des Forschungsvorhabens umfasst einen Zeitraum von 36 Monaten.

1.6 Gewünschter Beginn der Förderung

Der Beginn der Förderung wird zum 1.10.2005 gewünscht.

1.7 Zusammenfassung

Psycholinguistische Forschung zielt auf die Aufklärung und das Verständnis der Mechanismen des menschlichen Sprachverstehens. Dabei ergänzen sich experimentelle Befunde und deren Interpretation im Rahmen mathematisch-formaler Modelle wechselseitig. Zu diesem Zweck werden zunehmend solche Modellklassen heangezogen, die auch eine Erfassung *gradueller* Abhängigkeiten gestatten.

Als Alternative zu den bereits etablierten probabilistischen Modellen soll hier ein Ansatz zu Satzanalyse auf der Basis von gewichteten Constraints im Hinblick auf sein Potenzial zur Erklärung psycholinguistischer Phänomene betrachtet werden. Über die Möglichkeiten probabilistischer Modelle hinaus gestattet er es, die verschiedenen Informationsbeiträge zum Sprachverstehen so zu isolieren, dass ihr Zusammenwirken einer experimentellen Untersuchung zugänglich wird. Zudem bietet er einen vielversprechenden Erklärungsansatz für das Zusammenspiel von Sprachsystem und Erfahrungswissen bei der Beurteilung gradueller Abweichungen von der sprachlichen Norm.

In dem Projekt sollen Erkenntnisse aus der Psycholinguistik, der Informatik und der maschinellen Sprachverarbeitung zusammengeführt werden, in der Erwartung, dass sich aus einer solchen Synthese eine wechselseitige Befruchtung der Arbeiten in den genannten Gebieten ergibt. Es besteht die begründete Aussicht, dass mit dem hier verfolgten Ansatz ein Erklärungsmodell für psycholinguistische Tatbestände gefunden werden kann, das es ermöglicht, den traditionell als antagonistisch betrachteten Gegensatz zwischen Modellen zur Reanalyse bzw. zum Constraint-Lösen zu überwinden.

2 Stand der Forschung, eigene Vorarbeiten

2.1 Stand der Forschung

2.1.1 Psycholinguistik

Die psycholinguistische Forschung ist in starkem Maße durch die erheblich eingeschränkten Beobachtungsmöglichkeiten am System der menschlichen Sprachverarbeitung geprägt. Unter den wenigen, überhaupt zur Verfügung stehenden “Beobachtungsfenstern” spielt sicherlich das Verfolgen von Augenbewegungen zur systematischen Untersuchung der zeitlichen Abläufe beim Lesen eine zentrale Rolle. Andere Analysetechniken, wie die funktionelle Magnetresonanztomografie, die Ableitung von ereigniskorrelierten Potenzialen oder aber die Probandenbefragung kommen wegen ihrer spezifischen Nachteile nur für spezielle Fragestellungen zum Einsatz.

Wegen der grundsätzlichen Wechselwirkung zwischen Forschungsmethodik und -gegenstand wirkt sich die beschränkte Zugänglichkeit zum Untersuchungsobjekt naturgemäß auch auf die Setzung von Forschungsprioritäten aus. Im Mittelpunkt stehen daher vorrangig Ursachen für die beobachtbaren zeitlichen Phänomene, die sich z.B. aus der Art des Zusammenwirkens der verschiedenen sprachlichen Ebenen ergeben können. Hier konkurrieren zwei Erklärungsansätze miteinander: Zum einen wird eine modulare Kapselung der syntaktischen Verarbeitung als ein autonomes Subsystem postuliert, wodurch diese von außersyntaktischem Einfluss weitgehend abgeschirmt ist (Frazier, 1984). Alternativ dazu wird von einer weitgehenden Interaktion zwischen syntaktischer und semantischer Verarbeitung ausgegangen, sodass auch außersyntaktische Information einen deutlichen Einfluss auf die Verarbeitung einer sprachlichen Äußerung ausüben kann (Trueswell et al., 1994).

Eine Entscheidung zugunsten eines dieser Modelle erweist sich als extrem schwierig, solange nur Beobachtungen über den Zeitverlauf der menschlichen Sprachverarbeitung zur Verfügung stehen. Diese Schwierigkeiten liegen vor allem darin begründet, dass ja durch den zeitlich inkrementellen Charakter der Satzanalyse auch bei der Annahme einer strikten Kapselung ein außersyntaktischer Einfluss durch die Verarbeitung und Bewertung von initialen Satzfragmenten wirksam werden kann. Ein starkes Argument zugunsten der Modularitätshypothese ergibt sich vor allem aus ereigniskorrelierten Potenzialen, in denen sich syntaktische und semantische Regelverletzungen in ganz charakteristischer Weise unterscheiden lassen (Friederici, 1995, 1999). Zudem hat sich gezeigt, dass sich die strukturell bedingten Lesezeitverzögerungen durch außersyntaktische Einflussfaktoren zwar neutralisieren, nicht jedoch invertieren lassen.

Eine zweite wesentliche Ursache für die beobachtbaren zeitlichen Phänomene wird in der Arbeitsweise des menschlichen Sprachverarbeitungssystems vermutet. Hier stehen sich serielle und parallele Erklärungsansätze gegenüber. Dabei beschränkt sich die Analyse in einer seriellen Architektur auf jeweils eine einzige Strukturhypothese. Lesezeitverzögerungen werden dementsprechend auf Reanalyseprozesse zurück geführt. Auf diese Weise ergibt sich eine nahe liegende Erklärung für die sogenannten “Gartenpfadphänomene”, bei denen wegen der starken lokalen Präferenz für eine global inkonsistente Lesart überhaupt keine Analyse mehr gefunden werden kann.

Im Gegensatz dazu gehen parallele Ansätze davon aus, dass im Falle von lokaler Mehrdeutigkeit auch mehrere Strukturhypothesen gleichzeitig verfolgt werden. Durch sukzessiven Abgleich der sprachlichen Evidenz mit den Constraints des Sprachsystems wird dann eine Entscheidung über die jeweils zu präferierende Lesart herbeigeführt. In Abhängigkeit von der

relativen Stärke der vorliegenden Evidenz wird hierfür mehr oder weniger viel Zeit benötigt. Im Gegensatz zu dem stark an formalen Grammatikmodellen orientierten seriellen Modell beruft sich der parallele Ansatz eher auf die Aktivierungsmechanismen des neuronalen Substrats.

Beide Fragestellungen (modular/interaktiv bzw. seriell/parallel) stehen in enger Beziehung zueinander. So wird für modulare Architekturen typischerweise eine serielle Verarbeitung angenommen, bei der innersyntaktische Präferenzen auf der Basis von strukturbezogenen Kriterien über die Reihenfolge entscheiden, in der die Strukturhypothesen aktiviert werden. Demgegenüber sind interaktive Architekturen gut mit dem parallelen Verarbeitungsmodell kombinierbar, wobei über die Interaktion außersyntaktische Evidenz bereitgestellt wird.

Keiner der beiden Erklärungsansätze kann für sich einen Bezug zu realen Berechnungsmodellen beanspruchen. Zur Vereinfachung erfolgt daher die Diskussion von Entscheidungsalternativen ausnahmslos so, als ob bei der Satzanalyse immer nur zwei alternative Hypothesen zur Auswahl stehen würden. Demgegenüber ist ein realer Parser stets mit massiver lokaler Mehrdeutigkeit konfrontiert und muss oftmals hunderte, wenn nicht gar tausende von faktisch gleichrangigen Strukturhypothesen verwalten. Bis heute existiert kein einziger rein symbolischer Parser, der ausgehend von der Konsistenzannahme einer binären Logik breite sprachliche Abdeckung mit der Fähigkeit zur vollständigen Disambiguierung verbinden kann. Hinzu kommt das grundsätzliche Problem, dass klassische logikbasierte Grammatikformalismen völlig ungeeignet sind, das Sprachverstehen im Falle von ungrammatischen Äußerungen zu erklären. Ihre Anwendung ist daher nur unter stark idealisierenden Bedingungen möglich, die mit den Möglichkeiten menschlicher Sprachverarbeitung nur noch wenig Gemeinsamkeiten aufweisen.

Wie stark die bisherige Denkweise in dieser Diskussion durch die Konsistenzannahme der formalen Logik geprägt ist, zeigt das Beispiel der Arbeit von (Meng und Bader, 2000), die die Entscheidung zugunsten einer seriellen Architektur an der unterschiedlichen Behandlung von echten Gartenpfadphänomenen einerseits und grammatisch inkorrekten Sätzen andererseits ausgehen. Dabei legen sie die Hypothese zugrunde, dass ein paralleler Parser nicht in der Lage sein soll, ungrammatische Strukturhypothesen zu reaktivieren. Diese Forderung lässt sich jedoch bei neueren Ansätzen zur Modellierung gradueller Wohlgeformtheit (vgl. Abschnitt 2.1.2 bzw. 2.1.3) nicht länger aufrecht halten.

Eine Alternative bietet sich hier etwa mit der Optimalitätstheorie (Prince und Smolensky, 1993), die grundsätzlich auch Constraintverletzungen zulässt, für die aber auch noch keine praktikablen Implementationen bekannt sind. Einen Schritt weiter gehen die parallelen Aktivierungsmechanismen, die den Grad der Akzeptabilität von Analysevarianten unmittelbar modellieren und diesen dadurch einen unterschiedlichen Aktivierungsgrad zuweisen. Diese Aktivierung dient dann als Grundlage für die erforderlichen Disambiguierungsentscheidungen. Eine Umsetzung (meist in Form von konnektionistischen Modellen) war bisher allerdings auch nur für isolierte Einzelbeispiele möglich, bzw. beschränkt sich von vornherein nur auf Simulationsmodelle (Anderson et al., 2004; Vasishth und Lewis, 2004). Derzeit ist noch weitgehend unklar, wie im konnektionistischen Paradigma (a) die inkrementelle Arbeitsweise der menschlichen Sprachverarbeitung nachgebildet werden kann, wie (b) strukturelle Zusammenhänge ohne Informationsverlust zu repräsentieren sind und wie (c) eine Übertragung der Prinzipskizzen auf unrestringierten Text erfolgen könnte.

2.1.2 Computer-Psycholinguistik

Ausgangspunkt für die Idee zur Erweiterung des Methodeninventars der Psycholinguistik um computergestützte Modellierungen und Simulationen ist sicherlich die Tatsache, dass mit den beachtlichen Erfolgen der probabilistischen Verfahren in der Computerlinguistik erstmalig Verarbeitungskomponenten zur Verfügung stehen, die in zentralen Parametern überhaupt erst einen Vergleich mit menschlichen Sprachverarbeitungsleistungen gestatten. Crocker und Keller (2005) rechnen hierzu:

- Effizienz: Erfahrungsbasierte Heuristiken gestatten auch bei massiver Mehrdeutigkeit eine schnelle Entscheidung über die plausibelste Strukturhypothese.
- Abdeckung: Im Gegensatz zu den derzeit in der Psycholinguistik dominierenden “Spielzeug”-Modellen, erfordert die Erklärung von menschlichen Sprachverarbeitungsprozessen auch die Fähigkeit zur Verarbeitung beliebiger Äußerungen.
- Robustheit: Analog zum menschlichen Vorbild sollte ein Computermodell auch robust gegenüber Versprechern, Häsitationen, Neuansätzen und kleineren grammatischen Abweichungen sein.
- Adaptierbarkeit: Ein Computermodell der Sprachverarbeitung sollte sich ebenfalls an die jeweilige Sprachumgebung anpassen können.
- Performanz: Zusätzlich zu Effizienz und breiter Abdeckung, sollte das Modell auch die Probleme der menschlichen Sprachverarbeitung erklären können, die traditionell den Schwerpunkt der experimentellen Untersuchungen in der Psycholinguistik bilden.

Hinzufügen muss man hier noch den Aspekt der vollständigen Disambiguierbarkeit, der eine elementare Voraussetzung für einen Vergleich mit der menschlichen Satzanalyse darstellt. Alle diese Forderungen werden von konsistenzbasierten Berechnungsmodellen nicht oder nur unzureichend erfüllt. Probabilistische Modelle hingegen bieten zumindest für die Aspekte der Effizienz, der Abdeckung, der vollständigen Disambiguierung und der Adaptierbarkeit befriedigende Lösungen an. Inwieweit auch eine Erklärung der menschlichen Performanz auf Sprachdaten möglich sein wird, ist Gegenstand der aktuellen Forschung. Der Unterschied zur “klassischen” Psycholinguistik besteht dabei vor allem darin, dass Aussagen über die Konsequenzen probabilistischer (bzw. allgemeiner, erfahrungsbasierter) Ansätze zur Sprachanalyse ohne umfangreiche Experimente mit einer lauffähigen Implementation überhaupt nicht mehr möglich sind. Die prinzipielle Verfügbarkeit eines reinen “Gedankenmodells” kann nicht mehr als ausreichend angesehen werden, da sich in Abhängigkeit von den erst im Training auf authentischen Daten zu ermittelnden Modellparametern durchaus ganz gegensätzliche Schlussfolgerungen ergeben können.

Die große Aufmerksamkeit, die probabilistische Modelle derzeit auf sich ziehen, ist vor allem dadurch zu erklären, dass unsymmetrische Häufigkeitsverteilungen das Zeitverhalten der menschlichen Sprachverarbeitung in vielfältiger Weise beeinflussen. Zuerst nachgewiesen wurde dies für die verschiedenen Aspekte des lexikalischen Zugriffs, die unmittelbar von der a priori Wahrscheinlichkeit des Wortes abhängen (Crocker und Corley, 2002; Trueswell et al., 1993; Garnsey et al., 1997). In (Mitchell et al., 1996) wird dieser Ansatz auch auf die Häufigkeit von Strukturvarianten erweitert. Derartige Abhängigkeiten stehen im Mittelpunkt der Arbeiten an stochastischen Parsern bzw. restringierten Fragmenten davon (Collins, 1999;

Charniak, 2000; Roark, 2001; Jurafsky, 1996). Probabilistische Parser arbeiten üblicherweise mit generativ erzeugten und daher hochgradig übergenerierenden Regelmengen und schränken den daraus resultierenden, extrem großen Suchraum durch häufigkeitsbasierte Präferenzen ein.

Zwei grundsätzliche methodische Probleme bestehen für alle Arten von erfahrungsbasierten Modellen: Zum einen betrifft dies die Wahl eines adäquaten Korrelats für menschliche Analysezeiten. Da ein quasi-linear Zeitbedarf auf seriellen Prozessoren derzeit wohl nicht erreichbar ist, probabilistische Modelle Häufigkeitsaspekte aber direkt in Präferenzen und Entscheidungskriterien für die Hypothesenselektion abbilden, kann hier ein Zusammenhang hergestellt werden, indem hohe Ableitungswahrscheinlichkeiten als Prädiktor für einen geringen kognitiven Verarbeitungsaufwand interpretiert werden. Ein solcher Zusammenhang konnte z.B. für einige lokale Disambiguierungsentscheidungen (Crocker und Brants, 2000), die syntaktisch-semantische Passfähigkeit von Komplementen (Jurafsky, 1996) und globale syntaktische Ableitungswahrscheinlichkeiten (Keller, 2003, 2004) nachgewiesen werden.

Das zweite zu lösende Problem betrifft den Arbeitsmodus des probabilistischen Parsers, der für einen direkten Vergleich mit den inkrementellen Prozessen der menschlichen Sprachverarbeitung natürlich ebenfalls strikt inkrementell angelegt sein muss. Leseschwierigkeiten, wie sie beim Menschen beobachtet werden, sind lokale und daher flüchtige Phänomene. Wenn etwa (Keller, 2003, 2004) hierbei auf globale Ableitungswahrscheinlichkeiten für ganze Sätze zurückgreift, so muss davon ausgegangen werden, dass die lokalen Phänomene teilweise im stochastischen Rauschen untergehen.

Trotz der deutlichen Vorteile probabilistischer Modelle im Hinblick auf die oben genannten Aspekte, ergibt eine genauere Analyse durchaus auch eine Reihe von Problemen im Hinblick auf ihre kognitive Plausibilität. Hierzu sind zu rechnen:

- Probabilistische Modelle liefern ohne weiteres keine Erklärung für die (dynamischen) Einflüsse aus dem extralinguistischen Kontext. Da die Modelle auf umfangreichen, speziell annotierten Textsammlungen trainiert werden, ist die Anpassung an die Bedingungen individueller Kommunikationssituationen nicht möglich. Eine diesbezügliche Erhebung von Daten muss als nicht praktikabel betrachtet werden.
- In mehreren Fällen sind systematische Abweichungen zwischen der Anbindungspräferenzen bei der Syntaxanalyse und den in Korpusdaten beobachtbaren Häufigkeitsverteilungen gefunden worden (Hindle und Rooth, 1993; Pickering et al., 2000).
- Stochastische Modelle verbergen den individuellen Beitrag verschiedener Wissensquellen und linguistische Ebenen in einer monolithischen Gesamtarchitektur, die einer detaillierten Analyse nur sehr schwer zugänglich ist. Dadurch lassen sich Phänomene, die durch das Zusammenwirken unterschiedlicher Wissensquellen (z.B. an der Syntax-Semantik-Schnittstelle) hervorgerufen werden, nur eingeschränkt untersuchen.

Diese Probleme zeigen sich besonders deutlich, wenn über die Erklärung der temporalen Effekte hinaus auch das graduelle Grammatikalitätsempfinden des Menschen mit einbezogen werden soll. Trotz seiner erheblichen Bedeutung für die Beurteilung der Wohlgeformtheit natürlichsprachlicher Äußerungen wurde dieses Gebiet bisher aber erst sehr wenig bearbeitet. Verschiedene Autoren argumentieren in dieser Frage sehr deutlich zugunsten einer strikten Separation der graduellen Grammatikalität von Häufigkeitsparametern (z.B. Keller (2000); Abney (1996); Riezler (1997)).

Sollte sich tatsächlich herausstellen, dass hier eine Unverträglichkeit mit einem stochastischen Erklärungsansatz vorliegt, stellt sich nicht nur die Frage, ob probabilistische Modelle hinreichend gut geeignet sind, die temporalen Charakteristika menschlicher Sprachverarbeitung zu erklären, sondern auch, wie diese zeitlichen Effekte in allgemeinere Modellvorstellungen zur Behandlung gradueller Abstufungen eingebettet werden können. Dies erfordert letztendlich Berechnungsmechanismen, die in der Lage sind, häufigkeitsbasierte Evidenz mit Präferenzen aus dem Sprachsystem selbst zu kombinieren.

2.1.3 Constraint-basiertes Parsing

Alle in der Tradition der generativen Grammatiken verwurzelten Formalismen verstehen Grammatikregeln als elementare Bausteine, aus denen sich nach bestimmten Kombinationsregeln komplexe Strukturbeschreibungen zusammensetzen lassen. Das Fehlen einer Regel in der Grammatik führt unter dieser Annahme oftmals zu einem Zusammenbruch der Analyse, der nur durch externe Mechanismen der Regelaufweichung im Parser aufgefangen werden kann. Ursache für diese extreme Empfindlichkeit gegenüber ungesehenen bzw. unerwarteten Daten ist die starre Kopplung von Wohlgeformtheitsbedingungen an die strukturbildenden Elemente der Grammatik. Dadurch führt eine Verletzung dieser Bedingungen zwangsläufig zum Ausschluss der betreffenden Interpretation.

Im Gegensatz dazu beschreiben constraint-basierte Ansätze den Raum möglicher Strukturbeschreibungen durch Bedingungen, die völlig unabhängig vom eigentlichen strukturellen Grundgerüst der Grammatik definiert sind. Constraints bewirken dann ein sukzessives Ver ringern der Hypothesenvielfalt, bis im Idealfall nur noch eine Strukturhypothese übrig bleibt. Dadurch lässt sich Parsing durchgängig als Prozess der strukturellen Disambiguierung auffassen. Im Falle einer zu stark einschränkenden Grammatik kann problemlos auf Standardannahmen zurückgegriffen werden und eine wichtige Ursache für den Abbruch der Analyse entfällt.

Bereits (Karlsson, 1995) hat auf die Attraktivität eines solchen Modells aus kognitiver Sicht hingewiesen, da in diesem Denkansatz der kognitive Aufwand stark mit dem Disambiguierungsfortschritt korreliert, während klassische Parser eher zum entgegengesetzten Zusammenhang tendieren: Zusätzlicher Analyseaufwand schlägt sich üblicherweise in einer erhöhten Mehrdeutigkeit und damit einer zunehmenden Unsicherheit über das Parsingresultat nieder. Dieser Vorteil kommt jedoch nur zum Tragen, wenn die oben genannte Trennung von Strukturbeschreibungen und Constraints tatsächlich realisiert werden kann. Während dies für die ebenfalls constraint-basierte Head-driven Phrase Structure Grammar (Pollard und Sag, 1994) wohl nur schwer möglich ist, bieten sich dependenz-basierte Modelle hierfür geradezu an, da bei ihnen der Raum möglicher Strukturbeschreibungen sehr regulär aufgebaut und (bei geeigneter Repräsentation) vollständig rekursionsfrei ist.

(Karlsson et al., 1995) realisieren die Idee des constraint-basierten Parsings auf der Basis von Dependenzrelationen, die hinsichtlich ihrer genauen Anbindungspunkte teilweise unter spezifiziert sind. Demgegenüber verwendet die Constraint Dependency Grammar (Maruyama, 1990) vollständig spezifizierte Dependenzbäume und besitzt damit gute Voraussetzungen für einen direkten Vergleich mit alternativen Parsingansätzen. Der Ansatz wurde auch auf probabilistische Modelle erweitert und mit Erfolg zur Sprachmodellierung und zum stochastischen Parsing eingesetzt (Wang und Harper, 2002, 2004). Eine Alternative hierzu besteht in der Verwendung verletzbarer Constraints auf die im folgenden Abschnitt 2.2 genauer eingegangen wird.

2.2 Eigene Vorarbeiten

Die eigenen Vorarbeiten beziehen sich vor allem auf Arbeiten zum Parsing mit gewichteten Constraints im Rahmen der Weighted Constraint Dependency Grammar (WCDG) (Menzel, 1994, 1995, 1998; Heinecke et al., 1998; Schröder et al., 2000; Daum et al., 2003; Foth et al., 2005b,a, submitted), die von der DFG in zwei Vorgängerprojekten gefördert wurden. Das hierbei verwendete robuste Parsingverfahren ist als Prozedur zur Constraint-Optimierung über relationalen Strukturbeschreibungen realisiert. Constraints werden zu diesem Zweck mit numerischen Gewichtungen annotiert, auf deren Grundlage eine *optimale* Strukturbeschreibung für eine gegebene Äußerung ermittelt werden kann. Da die relationalen Strukturen als Dependenzbeziehungen zwischen Wortformen interpretiert werden und die Optimierungsprozedur eine *eindeutige* Unterordnung jeder Wortform erzwingt, kann das Verfahren auch als Disambiguierungsprozedur für Abhängigkeitsbäume aufgefasst werden.

Verschiedene Verfahren zur (ggf. approximativen) Berechnung der optimalen Strukturbeschreibung wurden implementiert und auf ihre spezifischen Eigenschaften hinsichtlich Vollständigkeit, Korrektheit und Ressourcenbedarf untersucht. Dies betrifft u.a. Algorithmen zur Best-First-Suche, zum heuristisch gesteuerten Pruning sowie Techniken der fehlergesteuerten Strukturtransformation (Menzel und Schröder, 1998; Foth, 1999; Foth et al., 2000b; Daum und Menzel, 2002).

Unter Berücksichtigung der prozeduralen Eigenschaften der einzelnen Verfahren sind im Hinblick auf eine möglichst große Nähe zur menschlichen Sprachverarbeitung insbesondere die transformations-basierten Methoden von Interesse. Bei ihnen wird versucht, eine initiale Strukturhypothese in einer Folge von Reparaturschritten sukzessiv in eine besser bewertete Strukturbeschreibung zu überführen. Die Auswahl der aussichtsreichsten Transformations-schritte erfolgt dabei auf der Grundlage von Informationen über aktuell beobachtete Constraintverletzungen. Sie gestatten es, diejenigen Elemente der Struktur zu identifizieren, die den größten Reparaturbedarf haben.

Auf dieser Grundlage wurde ein Parser für das Deutsche entwickelt, der aus der Sicht der psycholinguistischen Adäquatheit über einige sehr attraktive Eigenschaften verfügt (Menzel, 1994, 1999, 2002):

- Vollständige Disambiguierung: Analog zu probabilistischen Ansätzen ist der Parser dank der Constraintgewichte und der darüber erfolgenden Optimierung immer in der Lage, die von ihm präferierte Strukturhypothese eindeutig zu bestimmen.
- Robustheit: Im Gegensatz zu konsistenzbasierten Modellen stellen Constraintverletzungen bei der Arbeit mit einer typischen Constraintgrammatik keine Problemfälle dar, sondern sind vielmehr Grundlage aller Disambiguierungsentscheidungen und damit eine wichtige Quelle für die Steuerung der Analyse. Abweichungen von der grammatikalischen Norm werden daher in großem Umfang toleriert, so dass sich der Parser als äußerst robust gegenüber extra- bzw. ungrammatischem Eingabedaten erweist. Analyseabbrüche treten praktisch nie auf.
- strukturelle Echtzeitfähigkeit: Transformationsbasierte Optimierungsverfahren verfügen über eine inhärente Robustheit gegenüber Ressourcenbeschränkungen. Dadurch, dass (nach Ablauf einer initialen Latenzzeit) eine vollständige Analyse immer zur Verfügung steht, kann das Verfahren unter Verzicht auf eine möglicherweise noch zu erwartende Qualitätsverbesserung jederzeit unterbrochen werden. Analysequalität und Rechenzeit sind somit gegeneinander austauschbar.

- Diagnosefähigkeit: Constraintverletzungen können als Diagnosen für beobachtete Abweichungen von der sprachlichen Norm interpretiert werden. Neben ihrer zentralen Rolle bei der Steuerung des Parsing-Prozesses stehen sie auch als Quelle für graduell abgestufte Grammatikalitätsurteile zur Verfügung. Wichtig ist dabei, dass sich diese Beurteilung primär auf das Sprachsystem bezieht, d.h. unabhängig von der Auftretenswahrscheinlichkeit einer sprachlichen Konstruktion erfolgen kann.

Im Gegensatz zu zahlreichen *ad hoc*-Mechanismen für die Fehlerbehandlung stellt die Arbeit mit verletzbaaren Constraints zudem einen wichtigen Aspekt kognitiver Plausibilität sicher: Das Ignorieren einer Abweichung erfordert weniger (Berechnungs-) Aufwand, als deren Erkennung bzw. gar Diagnose oder Korrektur (Menzel, 1995).

- breite Abdeckung: Es konnte gezeigt werden, dass der Parsing-Ansatz auf der Basis gewichteter Constraints in der Lage ist, auch unrestringierten Text mit vergleichsweise hoher Sicherheit zu analysieren. So wurde auf einem Standardtestset von 1000 Sätzen aus dem NEGRA-Korpus ein f-score für die Kantengenauigkeit (Anbindung und Unterordnungsrelation korrekt) von 87.0% erreicht, während konkurrierende (probabilistische) Parser hier bei 76.1% (Dubey und Keller, 2003) bzw. 81.7% (Schiehlen, 2004) liegen. Eine Kontrollexperiment auf einem Testset von ca. 13.700 Sätzen aus einem breiten Spektrum von alternativen Textsorten zeigt, dass dieses Resultat auch auf neu hinzukommenden Daten reproduzierbar ist (Foth et al., 2005b). Inzwischen konnte auf der Basis von Supertag-Informationen eine weitere Steigerung der Genauigkeit auf dem Referenz-Testset bis auf 91.0% erreicht werden (Foth et al., submitted). Damit steht für das Projekt der derzeit zuverlässigste Parser für das Deutsche zur Verfügung.
- Unabhängigkeit von der Implementationsbasis: Für die Lösung des Optimierungsproblems stehen auch bei Beschränkung auf das transformations-basierte Paradigma noch verschiedene Algorithmen zur Verfügung. Bisher wurden hierzu ein eher technisch angelegtes Verfahren zur Tabu-Suche (Glover, 1989; Foth et al., 2000b), bzw. ein stärker aktivierungs-basierter Ansatz zur gesteuerten lokalen Suche (guided local search) (Voudouris, 1997; Daum und Menzel, 2002) erfolgreich genutzt. Es kann erwartet werden, dass auch eine Übertragung auf quasi-parallele Lösungsverfahren möglich sein wird.

Die bisher durchgeführten Experimente haben bestätigt, dass sich gewichtete Constraints hervorragend für eine Kombination mit *probabilistischen* Komponenten eignen. Dies wurde durch Experimente zur Einbindung eines Wortarten-Taggers, eines Chunkers und eines Supertaggers in den WCDG-Parser nachgewiesen (Daum et al., 2003; Foth et al., submitted)). Dadurch, dass es die gewichteten Constraints erlauben, eine inhärent unsichere Information als zusätzliche Evidenz in den Entscheidungsprozess einzubringen, ohne sich von dieser Information abhängig zu machen, kann der Beitrag der einzelnen Informationsquellen recht präzise ermittelt werden. Dabei erweist sich der Wortarten-Tagger als mit Abstand wichtigste Informationsquelle, die die Genauigkeit um einen Faktor 1.7 steigert. Von diesem Niveau ausgehend erbringen Chunker und Super-Tagger immerhin noch einen Beitrag von jeweils ca. 2% absolute Steigerung der Genauigkeit.

Literatur

Abney, Steven, 1996. Statistical methods and linguistics. In J. Klavans und P. Resnik (Hg.) *The Balancing Act: Combining Symbolic and Statistical Approaches to Language*, S. 1–26.

- MIT Press, Cambridge, MA.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C. und Qin, Y., 2004. An integrated theory of the mind. *Psychological Review*, Bd. 111(4), 1036–1060.
- Charniak, Eugene, 2000. A maximum-entropy-inspired parser. In *Proc. of the North American Chapter of the ACL*, S. 132–139. Morgan Kaufmann, Seattle, Washington.
- Collins, Michael, 1999. *Head-Driven Statistical Models for Natural Language Parsing*. Phd dissertation, University of Pennsylvania.
- Crocker, Matthew W. und Brants, Thorsten, 2000. Wide coverage probabilistic sentence processing. *Journal of Psycholinguistic Research*, Bd. 29(6), 647–669.
- Crocker, Matthew W. und Corley, Steffan, 2002. Modular architectures and statistical mechanisms: The case from lexical category disambiguation. In Suzanne Stevenson und Paola Merlo (Hg.) *The Lexical Basis of Sentence Processing*. John Benjamins, Amsterdam.
- Crocker, Matthew W. und Keller, Frank, 2005. Probabilistic grammars as models of gradience in language processing. In Gisbert Fanselow, Caroline Féry, Ralph Vogel und Matthias Schlesewsky (Hg.) *Gradience in Grammar: Generative Perspectives*, S. forthcoming. Oxford University Press, Oxford.
- Daum, Michael, 2004. Dynamic dependency parsing. In Frank Keller, Stephen Clark, Matthew W. Crocker und Mark Steedman (Hg.) *Proc. of ACL 2004 Workshop on Incremental Parsing*, S. 67–73. Association for Computational Linguistics, Barcelona, Spain.
- Daum, Michael, Foth, Kilian A. und Menzel, Wolfgang, 2003. Constraint based integration of deep and shallow parsing techniques. In *Proc. 11th Conference of the European Chapter of the ACL*, S. 99–106. Budapest, Hungary.
- Daum, Michael und Menzel, Wolfgang, 2002. Parsing natural language using guided local search. In F. van Harmelen (Hg.) *Proc. 15th European Conference on Artificial Intelligence*, S. 435–439. IOS Press, Amsterdam.
- Debusmann, Ralph, Duchier, Denys, Koller, Alexander, Kuhlmann, Marco, Smolka, Gert und Thater, Stefan, 2004a. A relational syntax-semantics interface based on dependency grammar. In *Proceedings of the COLING 2004 Conference*. Geneva/SUI.
- Debusmann, Ralph, Duchier, Denys und Kuhlmann, Marco, 2004b. Multi-dimensional graph configuration for natural language processing. In *Proceedings of the International Workshop on Constraint Solving and Language Processing*. Springer, Roskilde/DEN.
- Dubey, Amit und Keller, Frank, 2003. Probabilistic parsing for German using sister-head dependencies. In *Proc. 41st Annual Meeting of the Association of Computational Linguistics, ACL-2003*, S. 96–103. Sapporo, Japan.
- Ferreira, F., Christianson, K. und Hollingworth, A., 2001. Misinterpretations of garden-path sentences: Implications for models of reanalysis. *Journal of Psycholinguistic Research*, Bd. 30, 3–20.

- Foth, Kilian, Menzel, Wolfgang und Schröder, Ingo, 2005a. Robust parsing with weighted constraints. *Erscheint in: Natural Language Engineering*.
- Foth, Kilian A., 1999. *Transformationsbasiertes Constraint-Parsing*. Diplomarbeit, Universität Hamburg, Fachbereich Informatik.
- Foth, Kilian A., By, Tomas und Menzel, Wolfgang, submitted. Guiding a constraint dependency parser with supertags. In Srinivas Bangalore und Aravind K. Joshi (Hg.) *Complexity of Lexical Descriptions and its Relevance to Natural Language Processing: A Supertagging Approach*. MIT-Press, Cambridge, MA.
- Foth, Kilian A., Daum, Michael und Menzel, Wolfgang, 2005b. Parsing unrestricted german text with defeasible constraints. In H. Christiansen, P. R. Skadhauge und J. Villadsen (Hg.) *Constraint Solving and Language Processing*, Bd. 3438 von *Lecture Notes in Artificial Intelligence*, S. 88–101. Springer-Verlag, Berlin.
- Foth, Kilian A., Menzel, Wolfgang, Pop, Horia F. und Schröder, Ingo, 2000a. An experiment in incremental parsing using weighted constraints. In *Proc. 18th International Conference on Computational Linguistics, Coling 2000*, S. 1026–1030. Saarbrücken.
- Foth, Kilian A., Menzel, Wolfgang und Schröder, Ingo, 2000b. A transformation-based parsing technique with anytime properties. In *Proc. 4th Int. Workshop on Parsing Technologies*, S. 89–100. Trento, Italy.
- Frazier, Lyn, 1984. Modularity and the representational hypothesis. In *Proceedings NELS 15*, S. 131–144. Providence, Rhode Island.
- Friederici, Angela D., 1995. The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, Bd. 50, 259–281.
- Friederici, Angela D., 1999. The neurobiology of language comprehension. In Angela D. Friederici (Hg.) *Language comprehension: A biological perspective*, S. 263–301. Springer-Verlag, Berlin, Heidelberg, New York.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E. M. und Lotocky, M. A., 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, Bd. 37, 58–93.
- Gibson, E., 1991. *A Computational Theory of Human Linguistic Processing: Memory Limitations and Processing Breakdown*. Phd thesis, Carnegie Mellon University, Pittsburgh, PA.
- Glover, F., 1989. Tabu search I. *ORSA Journal on Computing*, Bd. 1(2), 4–32.
- Heinecke, Johannes, Kunze, Jürgen, Menzel, Wolfgang und Schröder, Ingo, 1998. Eliminative parsing with graded constraints. In *Proc. 17th International Conference on Computational Linguistics, 36th Annual Meeting of the ACL, Coling-ACL '98*, S. 526–530. Montreal, Canada.
- Hindle, D. und Rooth, M., 1993. Structural ambiguity and lexical relations. *Computational Linguistics*, Bd. 19, 103–120.

- Jurafsky, Daniel, 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, Bd. 20, 137–194.
- Karlsson, Fred, 1995. Designing a parser for unrestricted text. In Fred Karlsson, Atro Voutilainen, Juha Heikkilä und Arto Anttila (Hg.) *Constraint Grammar – A Language-Independent System for Parsing Unrestricted Text*, Kap. 1, S. 1–40. Mouton de Gruyter, Berlin, New York.
- Karlsson, Fred, Voutilainen, Atro, Heikkilä, Juha und Anttila, Arto (Hg.) 1995. *Constraint Grammar – A Language-Independent System for Parsing Unrestricted Text*. Mouton de Gruyter, Berlin, New York.
- Keller, Frank, 2000. *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. Ph.d. thesis, University of Edinburgh.
- Keller, Frank, 2003. A probabilistic parser as a model of global processing difficulty. In *Proc. 25th Annual Conference of the Cognitive Science Society*, S. 646–651. Erlbaum, Mahawah, NJ.
- Keller, Frank, 2004. Robust models of human parsing. In *Proc. 3rd Coling-Workshop on Robust Methods in Analysis of Natural Language Data*, S. 11–13. Geneva.
- Maruyama, H., 1990. Structural disambiguation with constraint propagation. In *Proc. 28th Annual Meeting of the ACL*, S. 31–38.
- Meng, M. und Bader, M., 2000. Ungrammaticality detection and garden path strength: evidence for serial parsing. *Language and Cognitive Processes*, S. 615–666.
- Menzel, Wolfgang, 1994. Parsing of spoken language under time constraints. In T. Cohn (Hg.) *Proc. 11th European Conference on Artificial Intelligence*, S. 560–564. Amsterdam.
- Menzel, Wolfgang, 1995. Robust parsing of natural language. In *KI-95: Advances in Artificial Intelligence*, S. 19–34. Springer-Verlag, Berlin.
- Menzel, Wolfgang, 1998. Constraint satisfaction for robust parsing of natural language. *Theoretical and Experimental Artificial Intelligence*, Bd. 10(1), 77–89.
- Menzel, Wolfgang, 1999. A strongly interactive but symbolic parsing architecture. In *Abstracts Int. Conf. Architectures and Mechanisms for Language Processing, AMLaP-99*, S. 69. Edinburgh.
- Menzel, Wolfgang, 2002. System architecture as a problem of information fusion. In *Proc. Int. Symposium Natural Language Processing between Linguistic Inquiry and Systems Engineering*, S. 74–84. Hamburg.
- Menzel, Wolfgang und Schröder, Ingo, 1998. Decision procedures for dependency parsing using graded constraints. In Sylvain Kahane und Alain Polguère (Hg.) *Proc. Coling-ACL Workshop on Processing of Dependency-based Grammars*, S. 78–87. Montreal, Canada.
- Mitchell, D. C., Cuetos, F., Corley, M. M. B. und Brysbaert, M., 1996. Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, Bd. 24, 469–488.

- Pickering, M. J., Traxler, M. J. und Crocker, Matthew W., 2000. Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language*, Bd. 43, 447–475.
- Pollard, Carl und Sag, Ivan A., 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago.
- Prince, Alan und Smolensky, Paul, 1993. Optimality theory: Constraint interaction in generative grammar. Technical Report 2, Rutgers University, Center for Cognitive Science.
- Riezler, Stefan, 1997. Quantitative constraint logic programming for weighted grammar applications. In Christian Retoré (Hg.) *Logical Aspects of Computational Linguistics (LACL'96)*, S. 346–365. Springer, Berlin.
- Roark, Brian, 2001. Probabilistic top-down parsing and language modeling. *Computational Linguistics*, Bd. 27(2), 249 – 276.
- Schiehlen, Michael, 2004. Annotation strategies for probabilistic parsing in german. In *Proc. 20th International Conference on Computational Linguistics, COLING 2004*, S. 390–396. Geneva, Switzerland.
- Schröder, Ingo, Menzel, Wolfgang, Foth, Kilian A. und Schulz, Michael, 2000. Modeling dependency grammar with restricted constraints. *Traitement Automatique des Langues (T.A.L.)*, Bd. 41(1), 97–126.
- Trueswell, J., Tanenhaus, M. und Garnsey, S. M., 1994. Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, Bd. 39, 558–593.
- Trueswell, J. C., Tanenhaus, M. K. und Kello, C., 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Bd. 19, 528 – 553.
- Vasishth, Shravan und Lewis, Richard L., 2004. Modeling sentence processing in act-r. In Frank Keller, Stephen Clark, Matthew Crocker und Mark Steedman (Hg.) *Proc. of ACL 2004 Workshop on Incremental Parsing*, S. 82–87. Association for Computational Linguistics, Barcelona, Spain.
- Voudouris, Christos, 1997. *Guided Local Search for Combinatorial Optimisation Problems*. Dissertation, Department of Computer Science, University of Essex, Colchester, UK.
- Wang, Wen und Harper, Mary P., 2002. The SuperARV language model: Investigating the effectiveness of tightly integrating multiple knowledge sources. In *Proc. Conf. on Empirical Methods in Natural Language Processing, EMNLP-2002*, S. 238–247. Association for Computational Linguistics, Philadelphia, PA.
- Wang, Wen und Harper, Mary P., 2004. A statistical constraint dependency grammar (cdg) parser. In Frank Keller, Stephen Clark, Matthew W. Crocker und Mark Steedman (Hg.) *Proc. of ACL 2004 Workshop on Incremental Parsing*, S. 42–49. Association for Computational Linguistics, Barcelona, Spain.

3 Ziele und Arbeitsprogramm

3.1 Ziele

Ziel des Projektes ist eine systematische Untersuchung des Potentials, das Grammatiken mit gewichteten Constraints zur Erklärung psycholinguistischer Phänomene bieten. Obwohl bereits seit einigen Jahren wichtige Erkenntnisse in dieser Richtung vorliegen (vgl. 2.2), sind erst jetzt mit dem erfolgreichen Abschluss der Arbeiten an einer robusten und qualitativ hochwertigen Grammatik für das Deutsche die Bedingung gegeben, um erste Experimente zur Überprüfung von konkreten Adäquatheitshypothesen in Angriff nehmen zu können.

Die bisherigen Erfahrungen bei der Entwicklung von und der Arbeit mit dem WCDG-System lassen erwarten, dass es mit Verfahren auf der Basis gewichtete Constraints gelingen könnte, die Vorteile der probabilistischen Ansätze mit Möglichkeiten zur direkten Beeinflussung des Modells auf der symbolischen Ebene zu kombinieren. Damit bietet dieser Ansatz auch weitaus umfangreichere Möglichkeiten zum Überprüfen psycholinguistischer Hypothesen, als dies auf der Basis rein probabilistischer Modelle möglich wäre:

- Durch gezielte Manipulation von Constraints und Constraintgewichten können im Modell Präferenzen individuell eingestellt werden und lassen sich so auch zur experimentellen Überprüfung von Gegenthesen heranziehen. Probabilistische Modelle hingegen sind durch das Training auf eine direkte Abbildung der Regularitäten aus den Korpusdaten festgelegt. Dies ist ein Vorteil im Hinblick auf eine effiziente und robuste Verarbeitung, erweist sich aber aus methodischer Sicht als Nachteil, wenn es nicht nur um hinreichende Adäquatheitskriterien, sondern auch um den Nachweis ihrer Notwendigkeit geht.
- Der constraint-basierte Ansatz gestattet es, die verschiedenen linguistischen Ebenen durch Modularisierung voneinander zu isolieren. Dabei wird jedoch nicht eine Gliederung des Parsers in autonome Prozesse erzwungen. Statt dessen werden verschiedene linguistische Repräsentationen in modularer Weise durch Constraints beschrieben. Ebenso werden die Schnittstellen zwischen den linguistischen Ebenen untereinander durch gewichtete Constraints realisiert. Somit kann der Einfluss und das Zusammenwirken an den Schnittstellen zwischen den Ebenen (insbesondere das Verhältnis zwischen syntaktischer und semantischer Information) im Detail studiert werden. Voraussetzung dafür ist allerdings die Erweiterung der derzeit verwendeten Grammatik um zusätzliche Repräsentationsebenen.
- Dank der inhärenten Robustheit ist constraint-basiertes Parsing weitgehend unabhängig von der Verfügbarkeit einzelner Wissensquellen. Deren Beitrag zum Analyseerfolg und zur Performanz des Gesamtsystems kann daher durch selektive Hinzunehmen bzw. Abschalten von Constraintmengen präzise gemessen werden.
- Gewichtete Constraints erlauben es, die Auswirkungen widersprüchlicher Informationsbeiträge bei der Satzanalyse gezielt zu untersuchen.

Gegen die Verwendung von manuell erstellten Grammatikregeln zur Erklärung psycholinguistischer Phänomene spricht natürlich, dass diese stark durch die subjektive Wahrnehmung des Grammatikautors geprägt sein können, während probabilistische Modelle aufgrund ihrer strikt empirischen Ausrichtung eher objektivierbare Aspekte des Sprachsystems widerspiegeln. Angesichts der Tatsache, dass der Parser seine hohe Analysequalität auf unrestringiertem Text in einer Vielzahl von Textsorten erreicht, besteht jedoch Grund zu der Annahme,

dass der subjektive Anteil an den Modellinformationen keinen substantiellen Einfluss auf die Belastbarkeit von Aussagen über das Verhältnis zum menschlichen Vorbild haben wird.

3.2 Arbeitsprogramm

Um mit dem WCDG-System Experimenten zur Überprüfung von Hypothesen zur psycholinguistischen Adäquatheit durchführen zu können, müssen drei wesentliche Voraussetzungen geschaffen werden:

- die Erweiterung von Parser und Grammatik um Mechanismen zur inkrementellen Analyse (Arbeitspaket 1, und 2),
- die Modellierung des Syntax-Semantik-Interface in der Grammatik zumindest in exemplarischer Form (Arbeitspaket 7), sowie
- die Entwicklung von Werkzeugen und -methoden zur Evaluation von Zwischenergebnissen der inkrementellen Analyse (Arbeitspaket 3).

Bei den Arbeiten zur inkrementellen Analyse kann von früheren Untersuchungen ausgegangen werden (Foth et al., 2000a), mit denen die prinzipielle Realisierbarkeit inkrementeller Verarbeitung bereits gezeigt werden konnte. Von der damals verfolgten Konzeption unterscheiden sich die jetzigen Planungen vor allem in der Modellierung der Erwartungen an den (noch unbekannt) rechten Kontext eines Wortes. Im Gegensatz zu den Untersuchungen in (Foth et al., 2000a) soll hierfür jetzt ein (temporär) unterspezifizierter Knoten verwendet werden, auf den etwa die Erwartungen über noch nicht gesättigte Valenzforderungen projiziert werden können (Daum, 2004). Eine der zentralen Fragestellungen für die Untersuchungen wird sein, inwieweit es gelingt den Hypothesenraum aufgrund der Informationen aus dem linken Kontext so weit zu einzuschränken, dass ein quasi-lineares Analysezeitverhalten erreicht werden kann. Es kann davon ausgegangen werden, dass die Kriterien hierfür in engem Zusammenhang mit den Bedingungen für das Auftreten von harten Gartenpfadphänomenen stehen.

Auch bei der Modellierung des Syntax-Semantik-Interface kann auf entsprechende Vorarbeiten zurückgegriffen werden, da bereits in den frühen Arbeiten zum Parsing mit gewichteten Constraints eine Mehrebenenrepräsentation vorgesehen war (Menzel, 1994, 1995; Schröder et al., 2000). Diese wurde jedoch aus Effizienzgründen später (vorübergehend) zurückgestellt, da bereits bei Sätzen aus der relativ kleinen Verbmobil-Domäne kein unmittelbarer Nutzen im Hinblick auf die Analysequalität nachweisbar war (Foth et al., 2005a). Wegen der prinzipiellen Schwierigkeiten mit einer domänenunabhängigen Modellierung der semantischer Einflüsse, kann ein solcher Nutzen erst recht nicht für die Verarbeitung von unrestringiertem Text erwartet werden. Für eine Überprüfung psycholinguistischer Adäquatheitshypothesen tritt dieser Aspekt jedoch in den Hintergrund, da die für die Experimente verwendeten Stimulusdaten ohnehin sorgfältig ausgewählt und ausbalanciert werden müssen. Somit erscheint eine *ad hoc*-Modellierung der einschlägigen semantischen Präferenzen durchaus praktikabel.

In der Zwischenzeit liegt mit den Arbeiten von (Debusmann et al., 2004a,b) im Rahmen der XDG (Extensible Dependency Grammar) ein vielversprechender Ansatz für eine orthogonale, dependenzbasierte Mehrebenenrepräsentation vor, der zwischen fünf systematisch aufeinander bezogenen Ebenen unterscheidet: unmittelbare Dominanz, lineare Anordnung, Tiefensyntax, Prädikats-Argument-Struktur und Skopus. Durch entsprechende Untersuchungen wird herauszufinden sein, ob sich hier ggf. eine Synthese mit der Idee der gewichteten Constraints

herstellen lässt.

Aufbauend auf diesen Erweiterungen von Parser und Grammatik sind dann eine Reihe von Experimenten zu psycholinguistischen Fragestellungen vorgesehen, mit denen die psycholinguistische Adäquatheit des constraint-basierten Ansatzes getestet werden soll. Hierbei stehen vier Themenkomplexe im Mittelpunkt:

- Welche menschlichen Verarbeitungspräferenzen lassen sich direkt aus den Eigenschaften des Parsers ableiten, auch wenn dieser bei der inkrementellen Analyse das grammatikalische Wissen über vollständige Sätze benutzt? Welche spezifischen Constraints für Satzpräfixe sind ggf. erforderlich? (Arbeitspaket 4)
- Wie wichtig ist die zentrale Grundannahme der Computerpsycholinguistik, dass für einen adäquaten Erklärungsansatz nur Modelle mit breiter Abdeckung geeignet sind? (Arbeitspaket 5)
Hierzu sollen eine Reihe von Experimenten mit stark überbeschränkten Grammatiken durchgeführt werden.
- Wie lassen sich menschliche Verarbeitungspräferenzen mit Constraints modellieren? Auf welcher Grundlage lassen sich Verarbeitungsschwierigkeiten vorhersagen? Wie wirken sich Annahmen über die Größe des verfügbaren Arbeitsspeichers auf die Verarbeitung aus? (Arbeitspakete 6 und 7)
Analog zu den diesbezüglichen Experimenten in der Psycholinguistik sollen
 - strukturbasierte Präferenzen (minimale Anbindung, später Abschluss usw.),
 - syntax-externe Einflüsse (thematische Passfähigkeit, Referenzauflösung, usw.),
sowie
 - informationsbezogene Präferenzen (Pickering et al., 2000)

betrachtet werden.

- Wie wirken sich Constraintverletzungen im Analyseverlauf eines constraint-basierten Parsers aus? (Arbeitspaket 8)
Auf der Grundlage der inhärenten Diagnosefähigkeit des Parsers soll dabei untersucht werden, ob sich der zeitliche Ablauf der menschlichen Fehlerdetektion, wie er sich in EEG-Ableitungen beim Auftreten von syntaktischen oder semantischen Abweichungen darstellt, auch mit den Mitteln der Constraint-Optimierung nachvollziehen lässt.
- Wie wirken sich Konfliktsituationen zwischen den verschiedenen linguistischen Repräsentationsebenen aus? (Arbeitspaket 9)
Hier sind vor allem inkonsistente Residualstrukturen bei der Reanalyse von besonderem Interesse (vgl. Ferreira et al. (2001)). Es soll überprüft werden, ob sich solche Divergenzen auch im Reanalyseverhalten eines transformations-basierten Parsers niederschlagen.

Diese Liste von Themenstellungen ist sicherlich nicht endgültig und wird im Projektverlauf in Abhängigkeit von aktuellen Entwicklungen in der psycholinguistischen Forschung zu modifizieren sein.

Grundsätzlich muss sich die Gestaltung der Experimente eng an die Methodik psycholinguistischer Untersuchungen anlehnen, um überhaupt eine Vergleichbarkeit der Resultate zu ermöglichen. Im Hinblick auf die Lesezeiten ist dabei die Wahl eines geeigneten Korrelats entscheidend. Nach dem Vorbild der Interpretation probabilistischer Daten können hierfür die Verarbeitungsschwierigkeiten mit (lokal) niedrigen Bewertungen in Beziehung gesetzt werden. Alternativ sollen auch konkrete Aufwandsmaße, wie die Anzahl von erforderlichen Transformationszyklen (als abstraktes Maß für den Reparaturaufwand), bzw. tatsächliche Rechenzeiten zum Vergleich herangezogen werden.

Als zusätzliche Beobachtungsdaten kommen in Frage:

- Präferenzen bei der Satzvervollständigung (vgl. Meng und Bader (2000))
- Inkonsistenzen in den Ergebnisstrukturen einer Mehrebenenrepräsentation (vgl. Ferreira et al. (2001))
- der Zeitpunkt der Fehlerdetektion durch ein lokales Absinken der Bewertung, auch unter Zeitdruck (speeded grammaticality judgment task, vgl. (Meng und Bader, 2000))
- Inspektion des Hypothesenraums auf mögliche Analysealternativen. Dadurch soll etwa herausgefunden werden, ob die typischerweise als Anbindungsalternativen diskutierten Strukturvarianten wirklich mit den konkurrierenden Hypothesen eines constraint-basierenden Parsers übereinstimmen.

Bei all den geplanten Experimenten soll ausschließlich von publizierten psycholinguistischen Daten und Modellen ausgegangen werden. Eigene psycholinguistische Experimente sind nicht vorgesehen.

Insgesamt zielen die Arbeiten auf eine Neubewertung grundlegender psycholinguistischer Erklärungsansätze im Lichte der Möglichkeiten des constraint-basierten Ansatzes ab. Besonders interessant ist dabei die Frage, ob sich der bisher als unumstößlich postulierte Antagonismus zwischen Reanalyse und parallelem Constraint-Satisfaction nicht aus der Perspektive des transformationsbasierten Parsings relativiert: obwohl einerseits der Parser die Menge aller möglichen Unterordnungsbeziehungen in gewichteter Form zur Verfügung hat, sorgt andererseits die Constraint-Optimierung dafür, dass jederzeit auch die momentan als optimal betrachtete Lesart bekannt ist. Hierdurch könnte sich ein typisches Reanalyseverhalten ergeben. Der Unterschied zu den klassischen Annahmen für eine parallele Architektur besteht dann nur noch darin, dass der Parser nicht mehr *vollständige* Lösungen verwaltet, sondern ein Vielzahl von bewerteten Baumfragmenten, aus denen der Optimierer die mit der besten Bewertung zusammenstellt. In diesem Zusammenhang wird auch das Verhältnis zu parallelen Modellen mit mehreren gewichteten Strukturhypothesen zu untersuchen sein, die im Gegensatz zum hier verfolgten Ansatz aber stets komplette Analysen für das gegebene Satzpräfix verwalten (Gibson, 1991).

Analog zur Serialitätsdiskussion stellt sich die Situation im Hinblick auf die Modularitätshypothese dar, wo sich eine Perspektivenverschiebung dadurch ergibt, dass ein Mehrebenenparser auf der Repräsentationsebene einer strikte Modularisierung verpflichtet ist, obwohl er auf der prozeduralen Ebene durchaus recht stark von interaktiven Elementen Gebrauch machen kann (Menzel, 1999).

Arbeitspakete

1. Umstellung des Parsers auf inkrementelle Analyse

6 PM, Abschluss 6. Projektmonat, Deliverable: Implementation

Das derzeit vorliegende Softwaresystem ist auf dynamische Constraint-Optimierung für Satzpräfixe unzustellen und um einem Mechanismus für die unterpezifizierte Modellierung des rechten Kontexts zu ergänzen.

2. Umstellung der Grammatik auf inkrementelle Analyse

6 PM, Abschluss 9. Projektmonat, Deliverable: Grammatik

Zusätzlich zu einer Umstellung auf die Behandlung noch unbekannter Wortformen im rechten Kontext muss überprüft werden, ob die Grammatik noch durch spezifische Constraints für die Präfixmodellierung ergänzt werden muss.

3. Evaluationsmethoden und -werkzeuge für partielle Analysen

3 PM, Abschluss 12. Projektmonat, Deliverable: Implementation, Report

Für die Auswertung von Experimentierdaten aus den Arbeitspaketen 4 bis 9 sind geeignete Werkzeuge zu entwickeln, die im Gegensatz zu den etablierten Methoden der Tatsache Rechnung tragen, dass nunmehr auch Messwerte für Teilanalysen berücksichtigt werden müssen.

4. Experimente zum Zeitverhalten der inkrementellen Analyse

3 PM, Abschluss 12. Projektmonat, Deliverable: Report

In diesem Arbeitspaket ist zu untersuchen, unter welchen Bedingungen durch heuristische Manipulation der Hypothesenmenge ein quasi-lineares Analysezeitverhalten erreicht werden kann. Hierfür ist insbesondere die prinzipielle Austauschbarkeit von Qualität und Rechenaufwand wesentlich.

5. Untersuchungen zur Aussagekraft von eingeschränkten Grammatikmodellen

6 PM, Abschluss 16. Projektmonat, Deliverable: Report

Mit diesen Untersuchungen soll geklärt werden, inwieweit eine breite sprachliche Abdeckung der Grammatik tatsächlich Vorbedingung für die Ermittlung psycholinguistisch relevanter Resultate ist. Hierzu sollen auch punktuelle Modelle für psycholinguistisch relevante Phänomene aus der englischen Sprache entwickelt werden.

6. Untersuchungen zur Vorhersage von Verarbeitungsschwierigkeiten

9 PM, Abschluss 22. Projektmonat, Deliverable: Report

Es soll geklärt werden, welche Parameter des Verfahrens zur Constraint-Optimierung für eine Vorhersage von graduell abgestuften Verarbeitungsschwierigkeiten bei der menschlichen Sprachverarbeitung herangezogen werden können. Kandidaten hierfür sind: Das Bewertungsprofil der optimalen Lösung, verfügbare Alternativen im Hypothesenraum und deren Bewertungen, sowie direkte Aufwandsmaße (z.B. Transformationszyklen oder

Rechenzeit). Die Untersuchungen greifen in starken Maße auf die Resultate von Arbeitspaket 4 und 5 zurück.

7. Untersuchungen zum semantischen Bias in der syntaktischen Verarbeitung

9 PM, Abschluss 28. Projektmonat, Deliverable: Report

Auf der Basis einer punktuellen Modellierung extra-syntaktischer Ebenen (lexikalische Subkategorisierungspräferenzen, thematische Passfähigkeit, oder aber auch die Verfügbarkeit von geeigneten Referenzobjekten im extralinguistischer Kontext usw.) soll untersucht werden, wie sich durch die zusätzlich verfügbare Evidenz das Parsingverhalten verändert und welche Parallelen sich zu experimentellen Daten aus psycholinguistischen Experimenten ergeben.

8. Untersuchungen zum Diagnoseverhalten

6 PM, Abschluss 32. Projektmonat, Deliverable: Report

Der Schwerpunkt liegt hier auf der grundsätzlichen Möglichkeit zur Unterscheidung unvollständiger Satzpräfixe von fehlerhaften, sowie dem Zeitpunkt zu dem sich Syntaxfehler bzw. semantische Abweichungen bemerkbar machen.

9. Untersuchungen zu Konflikten zwischen den linguistischen Ebenen

6 PM, Abschluss 36. Projektmonat, Deliverable: Report

Es soll untersucht werden, ob das unterschiedliche Disambiguierungsverhalten auf der syntaktischen und semantischen Ebene unter Umständen zu Inkonsistenzen in den Ergebnisstrukturen führen kann, wie sie auch beim Menschen beobachtet werden. Da die entsprechenden Daten nur für das Englische zur Verfügung stehen, sind aufbauend auf den Resultaten von Arbeitspaket 5 auch spezifische Vorbereitungsarbeiten zur Grammatikmodellierung erforderlich.

4 Beantragte Mittel

4.1 Personalbedarf

	Lfd. Nr.	Anzahl	Dauer	Tarif/Stunden	Aufgabenpakete
Mitarbeiter	1	1	36 Monate	BAT IIa	1,3,5,6,7,8,9
	2	1	36 Monate	BAT IIa/2	2,4,5,6,7,8,9
Hilfskräfte	3	2	36 Monate	73 h/Monat	

Wegen des hohen Anteils an softwarebezogenen Arbeiten in den Arbeitspaketen 1, 3, 5, 6, 7, 8 und 9 ist für die erste Mitarbeiterstelle eine Informatikerin mit einer ganzen Stelle BAT IIa vorgesehen. Die Notwendigkeit für eine ganze Stelle ergibt sich vor allem durch den Umfang und die Komplexität, die das WCDG-System inzwischen erreicht hat. Von der Stelleninhaberin soll auch die ständig erforderliche Pflege und Aktualisierung der Software übernommen werden.

Für die zweite beantragte Stelle ist eine Mitarbeiterin mit psycholinguistischem Hintergrund vorgesehen.

4.2 Wissenschaftliche Geräte

Für die Arbeiten steht ein im Rahmen von Vorgängerprojekten durch die DFG finanziertes Compute-Cluster mit derzeit 15 Dual-Prozessormaschinen zur Verfügung. Dieses Cluster wurde in zwei Ausbaustufen realisiert:

Ausbaustufe	Prozessoren	Inbetriebnahme
1	6 x 2	02/2002
2	9 x 2	12/2003

Seine Verfügbarkeit war eine wesentlichen Voraussetzung für die bisherigen Arbeiten im Bereich korpusbasierte Grammatikentwicklung und Evaluation. Da damit zu rechnen ist, dass hier in absehbarer Zeit Ersatzbedarf besteht, wird für das zweite Projektjahr die Erneuerung der 1. Ausbaustufe durch 10 Dual-Prozessormaschinen beantragt. Hierfür entsteht (inklusive der Hardware für Installation und Vernetzung) ein Aufwand von ca. 19.700,00 €.

Pos.	Beschreibung	Stückpreis	Anz.	Summe
1	Rack Kabinett	1.200,00 €	1	1.200,00 €
2	Dual Prozessor Rechner, Intel Xeon 2.8 Ghz	1.700,00 €	10	17.000,00 €
3	Gigabit Ethernet Switch, 16 Port	400,00 €	1	400,00 €
3	Tastatur/Konsole-Umschalter (KVM)	800,00€	1	800,00 €
4	Kabel (Strom, Netzwerk, KVM)	30,00 €	10	300,00 €
Gesamt:				19.700,00 €

Entsprechende Angebote wurden angefordert, von den betreffenden Firmen aber noch nicht unterbreitet. Sie werden nachgereicht, sobald sie vorliegen.

4.3 Verbrauchsmaterial

Für Verbrauchsmaterialien, wie z.B. Druckerpapier, Fotokopien, Toner und Datenträger muss mit einem Aufwand in Höhe von jährlich 400,00 € d.h. insgesamt

1.200,00 €

gerechnet werden.

4.4 Reisen

Projektjahr	Anzahl	Reiseziel	Kosten	Summe
1	1	Konferenzbesuch	1.900,00 €	3.900,00 €
	1	Workshopteilnahme	1.000,00 €	
	1	Arbeitsbesuch Saarbrücken	500,00 €	
	1	Arbeitsbesuch Leipzig	500,00 €	
2	1	Konferenzbesuch	1.900,00 €	3.900,00 €
	1	Workshopteilnahme	1.000,00 €	
	1	Arbeitsbesuch Potsdam	500,00 €	
	1	Arbeitsbesuch Saarbrücken	500,00 €	
3	1	Konferenzbesuch	1.900,00 €	2.900,00 €
	1	Workshopteilnahme	1.000,00 €	
Gesamt:				10.700,00 €

Begründungen:

- Die Arbeitsbesuche in Saarbrücken, Potsdam und Leipzig dienen dem Kennenlernen aktueller Entwicklungen der psycholinguistischen Forschung im deutschsprachigen Raum. Dabei spielt Saarbrücken als Zentrum der Computer-Psycholinguistik (Prof. Matthew Crocker), Leipzig wegen seiner umfassenden Forschungsarbeiten zur neuro-physiologischen Fundierung der Psycholinguistik (Prof. Angela Friederici) und Potsdam wegen der starken Bezüge zur linguistischen Grundlagenforschung insbesondere zur Optimalitätstheorie (Prof. Gisberg Fanselow) eine wichtige Rolle.
- Die Konferenz- bzw. Workshopteilnahmen sollen dazu genutzt werden, die Ergebnisse des beantragten Projekts vorzustellen und aktuelle Entwicklungen der psycho- und computerlinguistischen Forschung aufzunehmen. Hierfür sind folgende Konferenzserien einschlägig:

Cognitive Science	jährliche Konferenz zur Kognitionswissenschaft (einschließlich psycholinguistischer Fragestellungen)
AMLAP	jährliche Konferenz zur Psycholinguistik
Coling/ACL/EACL	jährlich im Turnus wechselnde Veranstaltungen, die sich in den letzten Jahren zunehmend auch psycholinguistischen Fragestellungen widmen

4.5 sonstige Kosten

entfällt

5 Voraussetzungen für die Durchführung des Vorhabens

5.1 Zusammensetzung der Arbeitsgruppe

Die Arbeitsgruppe wird aus dem Antragsteller, sowie den hier beantragten Mitarbeitern und Hilfskräften bestehen.

5.2 Zusammenarbeit mit anderen Wissenschaftlern

Im Rahmen des Projektes sollen Arbeitsbeziehungen zu den o.g. Arbeitsgruppen in Leipzig, Potsdam und Saarbrücken etabliert werden.

5.3 Auslandsbezug

entfällt

5.4 Apparative Ausstattung

Für die geplanten Implementierungs- und Experimentierarbeiten wird das Workstation-Cluster am Arbeitsbereich des Antragsteller genutzt (vgl. hierzu auch Abschnitt 4.2). Dabei kommt zum überwiegenden Teil selbstentwickelte Software aus Vorgängerprojekten zum Einsatz, die durch das hier beantragte Projekt weiterentwickelt werden soll. Die dafür erforderliche Basissoftware liegt frei verfügbar bzw. durch Campus-Verträge lizenziert vor.

5.5 Laufende Mittel für Sachausgaben

Die laufenden Mittel für sonstige Sachausgaben werden vom Fachbereich Informatik getragen.

6 Erklärung

Ein Antrag auf Finanzierung dieses Vorhabens wurde bei keiner anderen Stelle eingereicht. Wenn ich einen solchen Antrag stelle, werde ich die Deutsche Forschungsgemeinschaft unverzüglich benachrichtigen.

Den Vertrauensdozent der DFG an der Universität Hamburg, Herrn Prof. Dr. Reinhard Lieberei werde ich über diesen Antrag in Kenntnis setzen.

7 Unterschrift

(Prof. Dr.-Ing. Wolfgang Menzel)

8 Verzeichnis der Anlagen

Schriftenverzeichnis Menzel

Personalfragebogen Michael Daum

Sonderdrucke:

(Foth et al., 2000a)

(Schröder et al., 2000)

(Daum et al., 2003)

(Daum, 2004)

(Foth et al., 2005a)

(Foth et al., 2005b)

(Foth et al., submitted)