

TRANCE* – Ein System für maschinelle Satzanalyse: Syntax

Henning BRINKMANN**, Olaf HEISSENBÜTTEL,
Thorsten KOLLSTEDE, Christian LERCH, Michael SCHMEING,
Holger SCHNEIDER, Mario SCHOMAKERS, Jens WENDLER***,
Theodor KUJUMDJIEFF****.

Zusammenfassung

Im Rahmen des Projekts TRANCE wurde ein *Werkzeug für Linguisten* entwickelt, das ihnen die Möglichkeit bietet, *Sätze einer natürlichen Sprache mittels Computer zu analysieren*. Das Resultat der Analyse wird in Form eines syntaktischen Baums vom Computersystem dargestellt. Ein formalisiertes Sprachmodell, bestehend aus einer Phrasenstrukturgrammatik und einem Lexikon, dient als Grundlage für die Satzanalyse. Das System erlaubt dem Benutzer, die Grammatikregeln sowie die Lexikoneinträge selbst zu bestimmen. Ersetzbarkeit der Regeln der Phrasenstrukturgrammatik und Austauschbarkeit der Lexikoneinträge sind die beiden Hauptmerkmale dieses Systems.

Somit ist der Benutzer in der Lage, unterschiedliche Modelle der erforschten Sprache zu untersuchen und anhand der Baumdarstellung zu analysieren. Das System soll den Linguisten bei der Sprachforschung und bei der Lehre zugute kommen. Prinzipiell kann das Sprachmodell in seiner Funktionalität und in der Komplexität der zu erforschenden Sprachmodelle durch das Einführen von Merkmalen (*features*) erweitert werden.

* (TRansformational Sentence ANalysis for Computer Supported Education).

** Die Autoren dieses Projekts haben ihre Studien im Fachbereich Informatik der Carl von Ossietzky Universität Oldenburg absolviert.

*** Jens Wendler hat die DOS-Programmvariante des Projekts entwickelt. Er ist zur Zeit unter jens.wendler@netplus-oldenburg.de zu erreichen.

**** Theodor Kujumdjieff leitet das Projekt TRANCE, das aus seiner Vorlesung "Grundlagen der Verarbeitung natürlicher Sprache" entstanden ist. Er ist zur Zeit im Fachbereich Literatur- und Sprachwissenschaften der Carl von Ossietzky Universität Oldenburg tätig, erreichbar unter Kujumdjieff@Informatik.Uni-Oldenburg.de.

1. Zielsetzung

Das System TRANCE wurde als ein Werkzeug für Forschung und Lehre entwickelt. Bei der Systementwicklung wurden zweierlei Ziele angestrebt: Forschungsziele und didaktische Ziele. Das System kann praktisch für alle Indoeuropäischen Sprachen angewendet werden.

1.1 Forschungsziele

- Das unmittelbare Ziel des Projekts war es, ein linguistisches Werkzeug für die Erforschung natürlicher Sprache herzustellen. Mittels dieses Computerprogramms soll der Sprachwissenschaftler die Möglichkeit bekommen, unterschiedliche formale Beschreibungen verschiedener Sprachen zu untersuchen. Als Grundlage dient eine Generative Grammatik (Phrasenstrukturgrammatik), deren Regeln austauschbar sind, sowie ein Lexikon, dessen Einträge ebenso austauschbar sind. Die Visualisierungskomponente dient dazu, den Analyseprozeß in der meistverbreiteten Form – als syntaktischen Baum – darzustellen. Die Überlegungen, die zu der Entwicklung eines solchen Systems geführt haben, waren folgende:
 - a) *Disambiguierung*: Formale Sprachbeschreibungen (Phrasenstrukturgrammatiken) sind von Menschen schwer in die Praxis umzusetzen. Je nach Regelsatz und Kategorienzugehörigkeit ist zum Beispiel der englische Ausdruck *The English history teacher* entweder als “Ein Engländer, der Geschichte unterrichtet” oder als “Ein Lehrer der englischen Geschichte” zu interpretieren. Ähnliche Erscheinungen, verbunden mit verschiedenen Analysemöglichkeiten, sind häufig schwer nachzuvollziehen. Das TRANCE-Werkzeug soll dazu dienen, den Prozeß der Disambiguierung zu unterstützen.
 - b) *Prüfung der Regelmenge auf Akzeptabilität*: Oft führen Regeln zur formalen Akzeptanz von Ausdrücken, die von Sprechern als inakzeptabel angesehen werden. Dies wird erst nach einer mechanistischen Überprüfung der Regelmenge ersichtlich. Eine solche mechanistische Überprüfung fällt dem Forscher schwer, da er seine Sprachkenntnisse und sein Allgemeinwissen während des Analyseprozesses nicht völlig ignorieren kann und unbeabsichtigt auf sie zurückgreift. Das Compu-

tersystem jedoch liefert **alle möglichen Analysen** – die “richtigen” sowie die “falschen”. Somit kann die Regelmenge objektiv – als eine in der Tat formal anzuwendende Methode – eingesetzt werden.

- Als globales Forschungsziel können die Bestrebungen betrachtet werden, das TRANCE-System so zu konzipieren, daß es als Grundlage für ein System, das für die Ziele des Maschinellen Übersetzens einzusetzen ist, dienen soll.

1.2 Didaktische Ziele

Die in den Forschungszielen aufgezählten Eigenschaften tragen dazu bei, daß das linguistische Werkzeug TRANCE ein breites Einsatzgebiet als ein *Werkzeug für den Unterricht* findet. Es könnte erfolgreich bei Veranstaltungen, die der Satzanalyse unterschiedlicher Sprachen gewidmet sind, eingesetzt werden.

2. Sprachwissenschaftliche Konzepte und Begriffe

Um eine formale Beschreibung natürlicher Sprache als Computerwerkzeug zu implementieren, wird eine Sprachtheorie benötigt, die als Grundlage für diese Implementation dienen soll. Dabei sind Teilgebiete der Sprachwissenschaft wie Lexikologie, Morphologie, Syntax und Semantik zu berücksichtigen. Notgedrungen mußte die Komplexität einzelner sprachlicher Phänomene außer Acht gelassen werden. Die Anwendung sprachwissenschaftlicher Konzepte wurde auf einige Grundbegriffe reduziert, die für die Ziele des Projekts eine ausreichende Sprachstrukturbeschreibung liefern. An dieser Stelle werden lediglich einige der wichtigsten Begriffe eingeführt, die für einen Einblick in die Funktionsweise des Systems erforderlich sind.

2.1 Grundkonzepte

Das System TRANCE dient primär der Satzanalyse. Aus diesem Grund, sowie wegen des verwendeten Algorithmus, wurde beschlossen, keine morphologische Analyse durchzuführen. Eine solche Analyse würde beispielsweise mögliche Wortformen ergründen, die aus einem Stamm ableitbar sind. Statt dessen wird ein Lexikon benutzt, das alle bekannten

Wortformen einschließt, die für die jeweiligen Satzbeispiele benötigt werden. Insofern sind ganze Lexeme – meist unvollständige Lexeme – die Daten, mit denen das System operiert.

Um Kongruenzübereinstimmungen bezüglich Genus, Numerus und Kasus zu ermöglichen, sind Merkmalsangaben (feature-Angaben) vorgesehen. Die Einführung des feature-Konzepts würde es erlauben, das System so zu erweitern, daß bei der Satzanalyse unterschiedliche Spracherscheinungen wie morphologische und analytische Kasusmarkierungen, semantische Einschränkungen und temporale Abhängigkeiten berücksichtigt werden.

Jede Wortform wird einer oder mehreren lexikalischen Kategorien zugeordnet – Verb (V), Nomen (N), Adjektiv (Adj) usw. Aus den lexikalischen Kategorien werden die entsprechenden Phrasen wie Nominalphrase (NP), Verbalphrase (VP) usw. gebildet. Die lexikalischen Kategorien sowie die syntaktischen Kategorien bilden die Satzkonstituenten.

Bei der vorliegenden Systemversion wird eine syntaktische Analyse rein formal durchgeführt. Die kategoriebestimmenden Elemente sind Vertreter von Wortklassen wie Verb, Substantiv, Adjektiv, Adverb, Präposition usw. Wie bereits erwähnt, erlaubt jedoch das Systemkonzept, weitere Aspekte der Satzbauelemente zu berücksichtigen. Es ist zu erwarten, daß die rein syntaktische Satzanalyse derart erweitert wird, daß beispielsweise die Verteilung semantischer Rollen bei der Analyse eingeführt wird und Kriterien bezüglich der Übereinstimmung mit Allgemeinwissen überprüft werden können. Dies wäre der Übergang von einer syntaktischen zu einer semantischen Analyse. Bislang wird im Rahmen des Projekts die Semantik als rein lexikalische Semantik verstanden.

2.2 Generative Grammatik – Phrasenstrukturgrammatik

Der Schwerpunkt ist beim TRANCE-Projekt auf die syntaktische Satzanalyse gelegt. Als Grundlage für diese Art der Satzanalyse dient eine kontextfreie Phrasenstrukturgrammatik. An sich stellt eine generative Grammatik eine formalisierte Beschreibung – ein Modell – der Regeln und Objekte, die bei der Analyse bzw. der Generierung wohldefinierter Sätze unterschiedlicher Sprachen benutzt werden, dar. Diese formalisierte

Beschreibung bildet ein System von Regeln, die einen Ausdruck finden und eine semantische Interpretation erfahren, um die Satzbildung zu modellieren.

Die generative Grammatik besteht aus einem Alphabet und aus einer Menge von (Erzeugungs)Regeln. Das Alphabet seinerseits ist aus einem Eingangssymbol S (der Satz), einer Menge nicht-terminaler Symbole, auch Categoriesymbole genannt (NP – Nominalphrase, VP Verbalphrase, usw.) und einer Menge terminaler Symbole, den einzelnen Wortkategorien (V – Verb, N – Nomen, Art – Artikel, Adj – Adjektiv usw.), konstruiert. Den terminalen Symbolen werden letztlich Lexikoneinträge zugeordnet, die den lexikalischen Ausdruck der Satzstruktur bilden. Die Regeln verbinden terminale und nicht-terminale Symbole miteinander und ermöglichen die Ableitung wohldefinierter Sätze, ‘‘Satzteile’’ und Phrasen. In diesem Sinne kann eine generative Grammatik als Phrasenstrukturgrammatik oder Konstituentengrammatik bezeichnet werden. Folgendes Beispiel soll die eingeführten Begriffe erläutern:

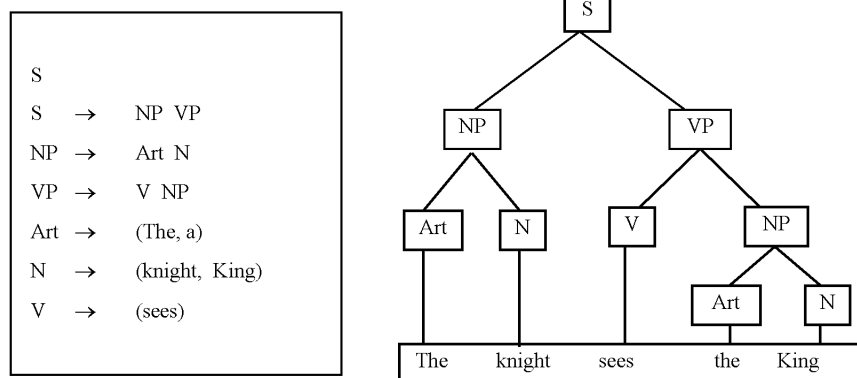


Abbildung 1

Die für das Projekt TRANCE als Grundlage benutzte Phrasenstrukturgrammatik wird als eine Regelmenge eingeführt, deren Hauptmerkmal die Austauschbarkeit der Regeln und der Lexikoneinträge ist. In diesem Sinne kann von einer frei definierbaren Phrasenstrukturgrammatik gespro-

chen werden. Eine Schlüsselfrage, die sich bei der Satzanalyse stellt, ist die Frage der Beziehung zwischen Syntax und Semantik.

2.3 Autonomie der Syntax und der Semantik

Beim Entwerfen eines Computersystems, das Sätze analysiert, müßten alle Aspekte der Sprachbeschreibung berücksichtigt werden. Durch die Entwicklung der Sprachforschung bedingt können einzelne Gebiete wie Lexik, Morphologie, Syntax und Semantik unterschieden werden, die jedoch nicht streng voneinander zu trennen sind. Die verschiedenen Auffassungen bezüglich der Aufgaben einzelner Wissenschaftsgebiete, sowie die unterschiedlichen Ansätze beim Modellieren der Sprachphänomene innerhalb eines sprachtheoretischen Systems sind teilweise widersprüchlich, und die Möglichkeiten ihrer Vereinbarung sind eingeschränkt. Infolgedessen sind weder Syntax noch Semantik als völlig autonom zu betrachten. Inhalt und Ausdruck eines Satzes bedingen sich gegenseitig, was einen Zusammenhang zwischen Syntax und Semantik zur Folge hat.

Als Beispiel für diesen Zusammenhang soll die Wortstellung im Deutschen und im Englischen dienen. Im Englischen ist die Wortstellung bedeutungsunterscheidend. Im folgenden Beispiel bestimmt allein die Wortstellung den Sinn des einzelnen Satzes:

The knight sees the King. vgl. *The King sees the knight.*

Im Deutschen hingegen kann die Wortstellung freier variieren. Beispiele:

Der Ritter sieht den König. vgl. *Den König sieht der Ritter.*

Häufig bestimmt die Bedeutung lexikalischer Einheiten die Akzeptabilität paradigmatischer Beziehungen im Satz. Folgendes Beispiel soll dies erläutern:

Er flog durch die Luft. vgl. **Er flog durch den Stein.*

Ein weiteres Beispiel wird die Wechselwirkung zwischen Bedeutung und syntagmatischen Beziehungen im Satz veranschaulichen. In diesem Fall wird die Opposition Aktiv-Passiv in Betracht gezogen, wobei die Bedeutung nicht als reiner Informationsgehalt aufgefaßt wird, sondern vielmehr als diskursbezogene Erscheinung.

Unexpectedly the police arrested Earl Grey.

Unexpectedly Earl Grey was arrested by the police.

Bei den unterschiedlichen Satzkonstrukten findet eine Promotion bzw. eine Demotion des Subjekts statt, die diskursmotiviert ist. Inhaltlich betrachtet sind beide Sätze gleich. Diskurspragmatisch gesehen besitzen sie unterschiedliche Bedeutung. Bei dem Passivsatz findet eine Promotion des Patiens statt, das syntaktisch als Subjekt kodiert wird. Dadurch, daß es als Objekt kodiert wird, erfährt das Agens eine Demotion. Die Bedeutung der syntagmatischen Relationen betrifft nicht den ausgedrückten Sachverhalt, sondern bezieht sich auf diskurs-pragmatische Phänomene.

Es bedarf keinerlei weiterer Beispiele, um zu behaupten, daß Syntax und Semantik als keine vollständig autonomen Gebiete bei der Satzanalyse zu betrachten sind. Daraus folgt, daß bei einer formalen Sprachbeschreibung mehrere Standpunkte miteinander zu vereinigen sind, um verschiedene Aspekte der Sprachinterpretation zu berücksichtigen. Bei einer Erweiterung des Umfangs und der Ansatzkomplexität eines Sprachmodells sind im allgemeinen genauere Analysen zu erwarten.

3. Systemkonzept

3.1 Allgemeine Beschreibung und Funktion einzelner Module

Das TRANCE-Programmsystem wurde entsprechend den Forschungs- und Lehrzielen konzipiert. Aus der Sicht des Benutzers – des Sprachwissenschaftlers – bietet das System folgende Funktionen an:

- Eingabe einer Menge von Wortformen – ein Lexikon.
- Ergänzung bzw. Austausch von Lexikoneinträgen.
- Eintragung oder Austausch von Syntaxregeln (Grammatik), nach denen die Satzanalyse durchgeführt werden soll.
- Eingabe des zu analysierenden Satzes.
- Graphische Darstellung der Satzanalyse in Form eines Syntaxbaums.

Die Ein- und Ausgabemodule stellen die *Benutzungsoberfläche* dar. Die Kommunikation zwischen Benutzer und System findet über die Benutzungsoberfläche statt. Der restliche Datenaustausch zwischen den Modulen

ist nur indirekt von den Benutzern zu beeinflussen. Schematisch ist das System in Abbildung 2 dargestellt. Zweckbedingt stellt das Schema lediglich die Hauptmodule dar.

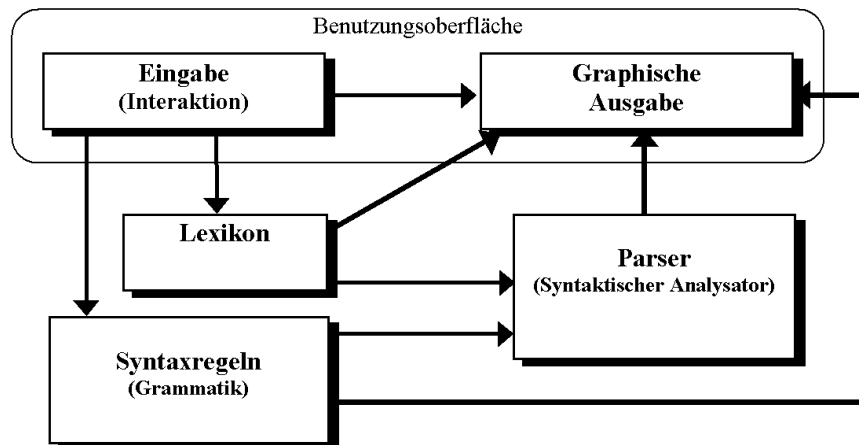


Abbildung 2

Das Eingabemodul

Das Eingabemodul (Interaktionsmodul) gibt dem Benutzer die Möglichkeit, Lexikoneinträge neu einzugeben und zu ändern, Syntaxregeln einzugeben und zu ändern, Sätze für den Analysevorgang einzugeben und die Ausgabe zu beeinflussen.

Modul der Graphischen Ausgabe

Das Modul der Graphischen Ausgabe dient dazu, die eingegebenen (vorhandenen) Lexikoneinträge und Syntaxregeln wiederzugeben, die zur Analyse eingegebenen Sätze zu zeigen und letztlich das Resultat der Satzanalyse graphisch – in Form von Syntaxbäumen – darzustellen.

Lexikonmodul

Das Lexikonmodul empfängt Daten (Lexikoneinträge) vom Eingabemodul und speichert sie. Auf diese Weise entsteht das Lexikon des Systems. Es ist vorgesehen, daß auch bereits existierende Lexika benutzt werden können. Das Lexikon liefert bei Bedarf die Daten über eine in ihm

existierende Wortform an das Parsermodul und/oder die Graphische Ausgabe weiter.

Grammatikmodul

Das Grammatikmodul dient dazu, Syntaxregeln von der Eingabe zu übernehmen, sie zu speichern und an den Parser weiterzuleiten. Regeln werden auch der Graphischen Darstellung zur Verfügung gestellt, wenn der Benutzer die Regelmenge visualisiert haben möchte.

Parsermodul

Das Parsermodul enthält den Parsingalgorithmus. Dieses Modul nutzt die Daten, die vom Lexikon und von der Grammatik bereitgestellt werden (Lexikoneinträge und Syntaxregeln) dazu, um die tatsächliche Analyse durchzuführen. Das Analyse-Ergebnis wird an die Graphische Darstellung weitergeleitet, um dort als Syntaxbaum visualisiert zu werden.

3.2 Algorithmusauswahl

Als syntaktischer Analysator (Parser) wurde ein *Chart-Parser* ausgewählt und weitgehend in Übereinstimmung mit dem *Earley Algorithmus* implementiert (NAUMANN & LANGER 1994). Bei der Auswahl des Algorithmus wurden die Eigenschaften von Top-down und Bottom-up Parsing-Algorithmus berücksichtigt. Bei dem im TRANCE-System benutzten *Earley Chart Parser* existieren keinerlei Einschränkungen, die Zyklen oder Tilgungsregeln betreffen. Außerdem ist es nicht erforderlich, eine Chomsky-Normalform zu befolgen.

3.3 Software-ergonomische Richtlinien

Ausgehend von den unmittelbaren Zielen, die den Systembenutzer betreffen, wurde Wert darauf gelegt, einige Grundrichtlinien der Softwareergonomie zu befolgen. Das System wurde unter Berücksichtigung von *Konsistenz, Robustheit, Fehlertoleranz und Rückmeldung, Transparenz, Selbsterklärbarkeit* und *Menschengerechtigkeit* entwickelt. Ohne diese Grundrichtlinien im einzelnen zu erörtern, sei an dieser Stelle lediglich erwähnt, daß die meisten davon möglicherweise mit den Eigenschaften des menschlichen Kurzzeitgedächtnisses in Verbindung stehen (SLAVOVA & KUJUMDJIEFF 1998).

3.4 Implementation

Eines der Ziele bei der Ausarbeitung des Systemkonzepts bestand darin, eine Plattformunabhängigkeit und die damit verbundene Portabilität zu erreichen. Darüber hinaus wurden Überlegungen angestellt, eine zukünftige Version des Systems für einen Internet-Benutzungsmodus vorzusehen. Ausgehend von diesen Voraussetzungen wurden mehrere Varianten in Angriff genommen. Die einfachste Variante wurde für eine MS-DOS-Umgebung entwickelt (WENDLER 1997). Sie verfügt über die gesamte Funktionalität unseres Systems, ist im Hinblick auf Benutzerfreundlichkeit und Erweiterbarkeit aber um einiges eingeschränkt. Die Zielplattformen für das Projekt sind die Rechnersysteme DEC-Stations und PCs.

Das Systemkonzept setzt eine Windows-ähnliche Umgebung voraus und ist für das Public-Domain-System wxWindows realisiert worden. Bei Abschluß der Java-Variante könnte das System weltweit über das Internet benutzt werden.

4. Einschränkungen und Erweiterungsmöglichkeiten

In der vorliegenden Version des Systems sind einige Einschränkungen vorhanden, die den Analyseprozeß beeinträchtigen. Als erstes ist die lexikalisch-morphologische Einschränkung bezüglich der Kongruenzen von Genus, Numerus und Kasus anzuführen. Eine Systemerweiterung durch das Einführen von Merkmalen (*features*) wird jedoch diese Einschränkung weitgehend aufheben. Eine weitere modelltheoretische Einschränkung ist die Erscheinung, die "diskontinuierliche Konstituenten" genannt wird. Als Beispiel dient folgender Satz:

Einige Verbesserungen sind im Laufe der Besprechung vorgeschlagen worden.

Als Hauptproblem, das in einer zukünftigen Variante des Systems in Angriff genommen werden soll, ist das Problem der fehlenden semantischen Analyse zu sehen. Um mittels semantischer Merkmale und entsprechender Regeln eine Übereinstimmung mit dem allgemein Bekannten – mit dem "gesunden Menschenverstand" – zu erreichen, wäre es erforderlich, den *Earley Algorithmus* erheblich zu modifizieren. Dies würde erlauben, semantische Rollen wie Agens, Patiens u.ä., sowie diskursrele-

vante Beziehungen wie Thema/Rhema, Fokus usw. in einem Sinnkontext auf ihre Gültigkeit zu überprüfen.

Die Erweiterungsmöglichkeiten lassen sich in den folgenden Punkten zusammenfassen:

- Einführung von Merkmalen (*features*), die Kongruenztests und die Vergabe von semantischen Rollen erlauben würden.
- Überarbeitung des Parsingalgorithmus im Hinblick auf eine parallele Verarbeitung. Die parallele Verarbeitung würde die gleichzeitige Überprüfung mehrerer Regelsätze (Kongruenz, semantische Konsistenz, usw.) erlauben.
- Implementierung des Systems in einer Java-Umgebung, was einen uneingeschränkten Zugriff ermöglichen würde.

5. Schluß

Das vorgestellte System, das als Resultat des Projekts TRANCE entstanden ist, sowie die bislang gesammelte Benutzererfahrung zeigen, daß es gelungen ist, ein linguistikorientiertes Werkzeug zu schaffen, das für eine praktische Forschungsanwendung geeignet ist. Die Benutzungsoberfläche ist übersichtlich gestaltet, was die Bedienung erleichtert. Die Einarbeitung in den Umgang mit dem System ist benutzergerecht. Die Einarbeitungszeit kann als gering bezeichnet werden.

Das Werkzeug kann sowohl zu Forschungszwecken dienen als auch als lehrunterstützendes Mittel verwendet werden. Mögliche Erweiterungen des modelltheoretischen Anwendungsbereichs würden es erlauben, komplexere Sprachmodelle mittels dieses Systems zu erforschen.

Literatur

- AARTS, Jan and CALBERT, Joseph (1979): *Metaphor and Non-Metaphor*. Tübingen: Max Niemeyer.
- ALLEN, James (1987): *Natural Language Understanding*. Menlo Park, California: Benjamin & Cummings Publishing.
- CHOMSKY, Noam and HALLE, Morris (1968): *The Sound Pattern of English*. New York: Harper & Row.

- CHOMSKY, Noam (1972): *Syntactic Structures*. The Hague: Mouton.
- EBERLEH, Edmund et al. (1994): *Einführung in die Software Ergonomie*. Berlin: de Gruyter.
- FELTKAMP, Hendrik (1977): *SINNTAX: Interpretierende Transformationsgrammatik des Deutschen*. München: Schöningh.
- FILLMORE, C.H. (1972): The case for case. In: E. BACH and R. HARMS (eds.), *Universals in Linguistic Theory*. London: Holt, Rinehart and Winston.
- LEHFELDT, Werner (1995): *Einführung in die Sprachwissenschaft für Slavisten*. München: Otto Sagner.
- KUJUMDJIEFF, Theodor (1992): Features in CAD-Systemen – Überblick und linguistische Ansätze. Bericht Nr. AIS-8 – November, 1992, Universität Oldenburg.
- (1994): *Transformational Sentence Analysis for Computer Supported Education*. Fachinterne Arbeit, Fachbereich Literatur und Sprachwissenschaften, Universität Oldenburg.
- (1996): *Grundlagen der Verarbeitung natürlicher Sprache: Maschinelles Übersetzen, Satzanalyse und Generierung*. Vorlesung an der Universität Oldenburg, Wintersemester 1995/1996.
- NAUMANN, Sven und LANGER, Hagen (1994): *Parsing: Eine Einführung in die maschinelle Analyse natürlicher Sprache*. Stuttgart: Teubner.
- RADFORD, Andrew (1988): *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.
- SAMPSON, Geoffrey (1992): Machine Translation. In: Peter ROACH (ed.), *Computing in Linguistics and Phonetics*. London: Academic Press.
- SLAVOVA, V. and KUJUMDJIEFF, T. (1998): The Association-Cup Model of Natural Language Comprehension. In: *Proceedings of the 5th International Summer School in Cognitive Science*, Sofia (im Druck).
- WENDLER, Jens (1997): *Entwicklung von linguistischer Lernsoftware für den Satzanalyse-Unterricht*. Studienarbeit an der Universität Oldenburg, Oldenburg.
- WINOGRAD, Terry (1983): *Language as a Cognitive Process: 1. Syntax*. Reading, Massachusetts: Addison-Wesley.
- WIRTH, Niklaus (1995): *Algorithmen und Datenstrukturen*. Stuttgart: Teubner.
- WURZEL, W.-U. 1984: *Flexionsmorphologie und Natürlichkeit (Studia grammatica, 21)*. Berlin: Akademie Verlag.
- ZALIZNJAK, A.A. (1967): *Russkoe imennoe slovoizmenenie*. Moskva: Nauka.