Zur Geschichte der Computerlinguistik

Wolfgang Menzel
Universität Hamburg, Vogt-Kölln-Straße 30, 22527 Hamburg
menzel@informatik.uni-hamburg.de

1 Die Ursprünge

Die frühen Entwicklungen zur Computertechnologie in den dreißiger und vierziger Jahren waren sehr stark durch die Hinwendung zu numerischen Problemstellungen geprägt. Dieser Umstand spiegelt sich recht deutlich in den ursprünglichen Namensgebungen wider: computational machinery, machine à calculer, ordinateur, вычислительная машина, Elektronenrechner usw. Allerdings wurde auch damals schon das enorme Potential der neuen Technologie für die Behandlung rein symbolischer Verarbeitungsaufgaben erkannt. Ausschlaggebend hierfür war wohl nicht zuletzt der erfolgreiche Einsatz zur Dechiffrierung verschlüsselter Nachrichtentexte, der letztendlich auch die maschinelle Übersetzung der natürlichen Sprache als Spezialfall einer Dekodierungsaufgabe realisierbar erscheinen ließ [Wea49]. Zugleich wurden erste Überlegungen zu den prinzipiellen Möglichkeiten maschinellen Informationsverarbeitung angestellt [Tur50]. Auch wenn es sich dabei angangs noch um reine Gedankenexperimente handelte, so bezogen sie sich doch ebenfalls auf ein Szenario, dass dem Bereich der maschinellen Sprachverarbeitung zuzuordnen ist, und setzten damit die prinzipielle Realisierbarkeit eines natürlichsprachlichen Dialogs zwischen Mensch und Maschine indirekt schon einmal voraus.

In diesen frühen Überlegungen weisen die sich abzeichnenden Lösungsansätze zur maschinellen Sprachverarbeitung durchaus noch eine gemeinsame Wurzel auf, die stochastische Informationstheorie [SW49]. Aus deren Perspektive erscheint ein fremdsprachlicher Text als das Ergebnis der Übertragung einer Nachricht über einen gestörten Kanal. Die Aufgabe etwa der maschinellen Übersetzung besteht dann darin, den ursprünglichen Nachrichtentext unter Verwendung der sprachspezifischen Symbolwahrscheinlichkeiten und der Kanalcharakteristika beim Empfänger zu rekonstruieren.

War zu diesem Zeitpunkt die Einheit des methodischen Inventariums noch weitgehend gewahrt, so konnte man schon bald darauf eine stärkere Aufspaltung in stochastische Verfahren einerseits und symbolische Ansätze andererseits beobachten. Während erstere vor allem im Bereich der Informationswissenschaft, aber auch zur Verifizierung der Autorenschaft eines Textes zum Einsatz kamen, wurden letztere geradezu zum Synonym der späteren Computerlinguistik und dominierten die Entwicklung des Gebiets über einen erstaunlich langen Zeitraum.

Für diese recht einseitige Entwicklung lassen sich sicherlich mehrere Gründe identifizieren. Zum einen war da Chomsky's Diktum [Cho57], dass prinzipiell kein statistischer Ansatz in der Lage sein kann, den fundamentalen Unterschied zwischen den beiden Sätzen

(1) Colorless green ideas sleep furiously.

(2) Furiously sleep ideas green colorless.

zu erfassen, da man mit einiger Sicherheit davon ausgehen darf, dass keiner von beiden jemals in einem englischen Diskurs auftreten würde, und somit einer stochastischen Beobachtung *per se* nicht zugänglich ist. Es hat letztendlich mehr als vier Jahrzehnte intensiver Forschung benötigt, um erkennen zu können, dass diese Annahme grundfalsch war, und dass sich unter Zuhilfenahme versteckter Variablen durchaus stochastische Modelle auf ganz gewöhnlichen englischen Korpusdaten trainieren lassen, die tatsächlich einen Unterschied von mehr als fünf Größenordnungen zwischen den Wahrscheinlichkeiten für diese beiden Sätze vorhersagen [Per00].

Auf der anderen Seite hatte die einseitige Bevorzugung symbolischer Verfahren aber sicherlich auch ganz praktische Gründe, die vor allem in der mangelnden Leistungsfähigkeit der damals verfügbaren Hardware zu suchen sind. Derartige Beschränkungen bevorzugen in der Tat symbolische Ansätze in ganz entscheidender Weise: So lässt sich etwa die prinzipielle Idee eines symbolischen Verfahrens immer auch anhand eines extrem stark vereinfachten Modells (wenige Regeln, geringer Abdeckungsgrad usw.) demonstrieren, wobei sich die eigentlichen Schwierigkeiten dann natürlich bei der Verallgemeinerung auf größere Sprachausschnitte einstellen. Dagegen muss bei einem vergleichbaren stochastischen Ansatz bereits für das allererste Experiment ein ganz erheblicher Aufwand im Bereich der Datensammlung und der sehr ressourcenintensiven Schätzverfahren (Training) geleistet werden.

2 Symbolische Sprachverarbeitung

Die frühen Arbeiten zur symbolischen Sprachverarbeitung orientierten sich einerseits sehr stark an den vorhandenen linguistischen Beschreibungsebenen (Morphologie, Syntax, Semantik), zum anderen aber auch an den unmittelbaren Bedürfnissen praktischer Anwendungen, wie Maschinelle Übersetzung und Informationsrecherche. Im Mittelpunkt standen daher Untersuchungen zur lexikalischen Repräsentation und morphosyntaktischen Analyse von Wortformen, sowie zur syntaktischen Analyse von Sätzen.

Auf der Ebene der Morphotaktik lässt sich ein starker Trend hin zu elementaren Techniken aus dem Bereich der Endlichen Automaten bereits seit den frühesten Ansätzen nachweisen. Hinsichtlich der lexikalischen Beschreibungen konzentrierten sich die Bemühungen stark auf die syntaktischen Auswirkungen von Wortbildungs- und Flexionsprozessen, während die semantischen Aspekte lange Zeit eher ausgeklammert wurden. Seit den achtziger Jahren wurden verstärkt Anstrengungen unternommen, die Redundanz im Lexikon zu reduzieren. Ein erster Schritt hierzu stellte die systematische Nutzung von Transducern zur Modellierung der phonologischen Variation [Kos83] dar. Durch geeignete Vererbungsmechanismen konnte auch auf der Seite der Lexikoninformation eine kompaktere Beschreibung erreicht werden. Um dabei dem Spannungsverhältnis zwischen Regel und Ausnahme angemessen Rechnung zu tragen, kamen dabei zumehmend auch Techniken der nichtmonotonen Vererbung zum Einsatz [EG89].

Wichtigster Motor für die Aktivitäten zur syntaktischen Analyse waren sicherlich die Bedürfnisse der Maschinellen Übersetzung, wo man sich von dem Rückgriff auf syntaktische Repräsentationen einen deutlichen Forschritt gegenüber den rein wortformbasierten Ansätzen versprach. Zum anderen lag hier ein enger Berührungspunkt mit parallelen Entwicklungen im Bereich der Programmiersprachen vor, wo beim Compilerbau durchaus vergleichbare Techniken zum Einsatz kamen. Dadurch gabe es insbesondere in den sechziger und siebziger Jahren eine starke wechselseitige Befruchtung.

Kontrovers wurde vor allem die Frage nach dem jeweils geeignetsten Grammatiktyp diskutiert, wobei im wesentlichen Ansätze zur Modellierung der Phrasenstruktur [Cho57] bzw. der Abhängigkeitsbeziehungen [Tes59], aber auch Kategorialgrammatiken [BH54] verwendet wurden. Besonders einflussreich war hierbei die Schule der Transformationsgrammatik [Cho57, Cho65], obwohl diese wegen der zugrundeliegenden generativen Sicht letztendlich keinerlei praktikable Sprachanalysesysteme hervorgebracht hat. Breiten Raum nahmen Untersuchungen zur effizienten Realisierung der syntaktischen Analyse (Parsing) ein. Wichtige Meilensteine stellen der Nachweis eines polynomialen Algorithmus für beliebige kontextfreie Grammatiken [Ear70], sowie die Idee der Wiederverwendung partieller Analyseergebnisse beim Chart-Parsing [Kap73, Kay73] dar.

Waren die frühen Systeme zur Sprachverarbeitung im wesentlichen *ad hoc*-Implementierungen bestimmter algorithmischer Ideen, so ist seit den siebziger Jahren eine zunehmende Tendenz hin zu generischen Formalismen zu verzeichnen, die dank ihres hohen Abstraktionsgrades dann auch für ganz unterschiedliche Verarbeitungsaufgaben eingesetzt werden können. Diese Entwicklung vollzog sich über spezielle Programmiersprachen mit teilweise noch stark prozedural orientierter Semantik (z.B. der durch gezielte Erweiterung aus den Endlichen Automaten entstandene Formalismus der Augmented Transition Networks, ATN [Woo70]), über stärker deklarativ angelegte Formalismen zur Darstellung linguistischen Wissens (z.B. die Baum- und Graphtransformationssprachen ROBRA [CB78] bzw. Systèmes-Q [Col70]), bis hin zu den rein deklarativen Formalismen auf der Basis der Unifikation (z.B. die unifikationsbasierten Grammatikformalismen mit kontextfreiem Grundgerüst, wie PATR-II [Shi86]). Mit den constraintbasierten Unifikationsformalismen [Shi92] liegt nunmehr auch ein rein deklaratives und dennoch berechnungsuniverselles Modell vor, das einerseits hohen Ansprüchen im Hinblick auf eine prinzipienorientierte und damit erkärungsadäquate Modellierung der Grammatik gerecht wird [Cho81, PS94], andererseits aber auch die Brücke zum Paradigma der Logikprogrammierung in der Informatik schlägt.

Generell sind durch die verstärkte Hinwendung zu universell verwendbaren

Formalismen, auch deren formale Eigenschaften verstärkt ins Blickfeld geraten. Ziel dieser Untersuchungen ist es vor allem, diejenigen Modellklassen zu identifizieren, die es gestatten, eine gegebene Problemstellung mit minimaler Mächtigkeit und größtmöglicher Effizienz zu lösen.

Universell verwendbare Formalismen eröffnen darüberhinaus auch die Möglichkeit zur Realisierung ebenenübergreifender Modelle, die sehr unterschiedliche Aspekte des sprachlichen Wissens integrieren können. Ein Beispiel hierfür ist die Konstruktion einer semantischen Repräsentation auf der Grundlage der Montague-Grammatik [Mon74], die dann mit den Mitteln der Unifikation in einem constraint-basierten Formalismus emuliert werden kann [BKU88]. Vergleichbare Erweiterungen sind auch zur Einbeziehung satzübergreifender Phänomene auf der Grundlage der Diskursrepräsentationstheorie (DRT) [KR93] möglich.

3 Korpusstatistische Verfahren

Das Wiedererwachen des Interesses an stochastischen Verfahren steht in engem Zusammenhang mit den deutlichen Fortschritten bei der Erkennung gesprochener Sprache seit Anfang der achtziger Jahre. Gerade in diesem Gebiet hat sich gezeigt, dass die automatische Ermittlung von Modellparametern aus einem speziell aufbereiteten Korpus von Sprachdaten (oftmals als Training bezeichnet), einen entscheidenden Schritt zur Lösung des Wissensakquisitionsproblems darstellt. Letztendlich wurde erst durch den konsequenten Einsatz solcher Trainingsverfahren die Erkennung mit großen Wortschätzen und mehreren Sprechern überhaupt ermöglicht [Jel76].

Für die erfolgreiche Anwendung stochastischer Techniken müssen mehrere, teils widersprüchliche Forderungen erfüllt sein:

- Zum einen muss die Struktur des Modells so gewählt werden, dass die Zahl der zu schätzenden Modellparameter und die verfügbaren Trainingsdaten in einem ausgewogenen Verhältnis stehen.
- Zum anderen sollte das Modell über genügend Freiheitsgrade verfügen, um die Struktur der Daten angemessen widerspiegeln zu können, gleichzeitig aber beschränkt genug sein, um eine Generalisierung über den Trainingsdaten zu erzwingen und ein "Auswendiglernen" der Einzelbeispiele zu vermeiden.

Ausgangspunkt des Modellentwurfs ist hierbei also nicht ein extern vorgegeber Adäquatheitsanspruch, wie dies für die symbolischen Verfahren charakteristisch ist, sondern vor allem die Frage der wirksamen Trainierbarkeit eines Modells auf einem vorgegebenen Datensatz.

Diese grundlegende Besonderheit teilen die stochastischen Verfahren mit einer ganzen Klasse von trainierbaren Modellen, zu denen auch Systeme zum konnektionistischen und symbolischen Lernen gehören. Wesentliches Charakteristikum ist also nicht so sehr die wahrscheinlichkeitstheoretische Fundierung des Ansatzes, sondern vielmehr die Tatsache, dass in der Trainingsphase die für die jeweilige Aufgabe relevanten statistischen Eigenschaften der Daten zur Modelladaption ausgenutzt werden.

Die wohl erste computerlinguistische Aufgabe, die Ende der achtziger Jahre mit korpusstatistischen Methoden erfolgreich bearbeitet wurde, war die Wortartendisambiguierung (Tagging) [DeR88]. Angespornt von diesen Anfangserfolgen wurden dann zunehmend anspruchsvollere Zielstellungen verfolgt und Erfahrungen mit komplexeren Modellstrukturen gesammelt. Zu diesen Aufgaben gehören

- die syntaktische Analyse (Parsing) unter Verwendung unterschiedlich stark strukturierter Repräsentationen, z.B. [BW92]
- die strukturelle syntaktische Diambiguierung, z.B. PP-Attachment [HR93],
- die semantische Lesartendisambiguierung,
- die automatischen Ermittlung lexikalischer Information und
- die bilinguale Übersetzung [BCP$^+$90].

Auch wenn bei den vielfältigen Experimenten zur Entwicklung korpusstatistischer Verfahren oftmals die klassischen Modellvorstellungen der strukturellen Linguistik Pate gestanden haben, so hat sich jedoch bald gezeigt, dass die elementaren Modellstrukturen der traditionellen Ansätze (z.B. kontextfreie Regeln) für eine direkte Übernahme in das neue Paradigma nur bedingt geeignet sind. Dies hat zu einer deutlichen Akzentverlagerung in zweierlei Hinsicht geführt:

- In vielen Fällen kann eine stochastische bzw. konnektionistische Modellierung besser über die elementaren Operationen des zugrundeliegenden Entscheidungsprozesses (z.B. Transformation von Symbolsequenzen, Parseraktionen, ...) erfolgen, als auf der Ebene der Modellstrukturen selbst [Mag95]. Somit rückt die Perspektive der Performanz wieder stärker in den Mittelpunkt.
- Das klassische Ideal einer redundanzarmen Beschreibung bringt gleichzeitig eine massive Verletzung der stochastischen Unabhängigkeitsannahme mit sich, so dass sich für eine erfolgreiche Modellierung vielfach sehr komplexe und hochgradig redundante Modellstrukturen besser eignen [Bod95].

Zunehmende Aufmerksamkeit wird nunmehr auch der Frage nach möglichen Synergieeffekten durch die Integration symbolischer, stochastischer und konnektionistischer Verfahren in hybriden Systemlösungen gewidmet [WW97].

4 Anwendungen der Computerlinguistik

Obwohl das anwendungsbezogene Problem der Maschinellen Übersetzung bereits am Anfang der Arbeiten zur Computerlinguistik stand, zieht es auch ein halbes Jahrhunder später noch ein unvermindert starkes Forschungsinteresse auf sich, das nur gegen Ende der sechziger Jahre durch die recht pessimistischen Prognosen des ALPAC-Reports (siehe [Hut86]) für kurze Zeit abgeschwächt worden war.

Dass trotz einer jahrzehntelangen und intensiven Forschungsarbeit auf diesem Gebiet noch immer wesentliche Fragen der Übersetzungsqualität, sowie der Portierbarkeit auf neue Anwendungsbereiche und Sprachpaare offen sind, zeigt zum einen, dass es sich bei der Maschinellen Übersetzung um ein überaus schwieriges Sprachverarbeitungsproblem handelt. Zum anderen wird aber auch deutlich, dass wir es hier mit einer typischen technologischen Fragestellung zu tun haben, die immer durch einen Kompromiss zwischen Anspruch und Wirklichkeit gekennzeichnet ist, und dass damit so etwas wie eine *endgültige* Lösung des gegebenen Problems auch gar nicht erwartet werden darf. In diesem Sinne steht die Maschinelle Übersetzung gleichberechtigt in einer Reihe mit anderen technologischen Aufgabenbereichen, die sich in einer ganz ähnlichen Situation befinden: Zwar existieren nach nunmehr schon mehreren Jahrhunderten intensiver Entwicklungsarbeiten zahlreiche brauchbare Lösungsansätze für das Problem des Transports von Personen und Gütern, dennoch sind auch hier keinerlei Aussichten auf eine abschließende Behandlung dieser Aufgabenstellung zu erkennen.

Analog hierzu haben seit den achtziger Jahren einige Übersetzungssysteme durchaus auch die Reife zum Einsatz in speziellen Anwendungsszenarien erlangt. Ein Weg hierzu führte über die Beschränkung auf sehr spezielle Textsorten (z.B. Wetterberichte [Tho82]). Alternative Ansätze setzen stärker auf eine manuelle Nachbereitung der Übersetzungsresultate. Andere Entwicklungen wiederum zielen vor allem auf eine optimale Unterstützung des Humanübersetzers, dem eine Reihe von Werkzeugen zur Sicherung der terminologischen Konsistenz, zur Wiederverwendung bisheriger Übersetzungsresultate, sowie zur partiellen (Roh-) Übersetzung bei Routineaufgaben an die Hand gegeben werden soll.

Parallel zu den Arbeiten an der Maschinellen Übersetzung ist den letzten drei Jahrzehnten eine erstaunliche Vielfalt von Anwendungssystemen auf der Grundlage computerlinguistischer Verfahren entwickelt und teilweise auch schon zur Einsatzreife gebracht worden. In vielen Fällen sind diese Arbeiten erst durch die bedeutenden Fortschritte auf anderen Gebieten der Informationstechnologie initiiert bzw. vorangetrieben worden. So wurde die wohl erste erfolgreiche Anwendung morphologischer Analysetechniken zur automatischen Silbentrennung ganz wesentlich durch den umfassenden Übergang zum Photosatz im Druckereigewerbe Anfang der sechziger Jahre forciert. Erst mit der flächendeckenden Verbreitung der Mikrorechner seit den achtziger Jahren, steht diese Technologie als standardmäßiger Bestandteil aller Textverarbeitungssysteme auch einem Massenpublikum zur Verfügung. Vergleichbare Entwicklungen waren auch im Bereich der Hilfsmittel zur Rechtschreibprüfung und -korrektur zu verzeichnen [Pet80].

Recht deutlich lässt sich der Einfluss externer Faktoren auch auf dem Gebiet der Informationssuche nachvollziehen, wo durch die zunehmende Verbreitung des WWW eine deutliche Belebung der diesbezüglichen Forschungsaktivitäten zu verzeichnen ist [BYRN99]. Durch die explosionsartig anwachsende Menge der digital verfügbaren Information sind in diesem Zusammenhang eine Reihe von

2000		
		MÜ für gesprochene Sprache
		Informationsextraktion
	stochastisches Parsing	
1990	stochastisches Tagging	stochastische MÜ, Diktiersysteme
	Constraint-basierte Grammatiken	
	Vererbung im Lexikon	
	Unifikationsgrammatiken, Zweiebenenmorphologie	
1980	Diskursrepräsentationstheorie	
		MÜ im Routineeinsatz
	Semantikkonstruktion	Rechtschreibfehlerkorrektur
	Chart-Parsing	natürlichsprachliche Datenbankabfrage
1970	ATN-Grammatiken	Automatische Silbentrennung
	Morphologische Analyse	
1960	syntaktisches Parsing mit CFG	
		experimentelle MÜ
	Sprachverarbeitung als Zeichenkettenmanipulation	
1950	Erste Gedankenexperimente	

Abbildung 1: Zeittafel

Anwendungsszenarien mit zum Teil ganz neuartigen Anforderungen entstanden:

- die Online-Recherche, die sich insbesondere durch extreme Effizienzerwartungen auszeichnet.

- die Informationsfilterung und -klassifikation zur Zuordnung relevanter Dokumente z.B. bei der E-Mail-Sortierung bzw. als Grundlage hochgradig individualisierter Informationsangebote,

- die Informationsextraktion zur inhaltlichen Erschließung von Textdokumenten im Hinblick auf stark spezialisierte Informationsbedürfnisse.

Ein Bereich, der vor allem von der gewaltigen Steigerung der Hardwareleistungsfähigkeit seit Beginn der neunziger Jahre profitiert hat, ist die automatische Spracherkennung, die insbesondere in Form von Diktieranwendungen zunehmende Verbreitung findet. Ein wesentlicher Berührungspunkt mit computerlinguistischen Forschungen ergibt sich hierbei durch die Notwendigkeit, Prädiktionen über Wortformsequenzen (Sprachmodellierung) in die Ermittlung des Erkennungsergebnisses einfließen zu lassen. Benötigt werden hierzu vor allem

Verfahren zur leichteren Modelladaption an neue Nutzer und unbekannte Textsorten, sowie Techniken zur besseren Einbeziehung nichtlokaler Abhängigkeiten auf den verschiedenen sprachlichen Ebenen.

Dass sich die fundamentalen Trends der Informationstechnologie durchaus nicht immer förderlich auf die Entwicklung computerlingustischer Anwendungen auswirken müssen, lässt sich etwa am Beispiel des natürlichsprachlichen Zugriffs zu Datenbanken beobachten, an den Mitte der achtziger Jahre erhebliche kommerzielle Hoffnungen geknüpft waren. Hier wurde die Entwicklung jedoch durch das Aufkommen graphischer Nutzerschnittstellen vollständig überholt. Für spezielle, aber typische Anwendungskontexte, wie Fahrplan- und Produktauskünfte, konnte alternativ zur geschriebenen Sprache ein Kommunikationskanal bereitgestellt werden, der eine bequemere und zugleich robustere Mensch-Maschine-Interaktion ermöglicht. Wichtige Aspekte dieser Technologie erfahren allerdings bereits heute eine Neuauflage in Dialogsystemen zur automatischen Telefonauskunft bzw. durch aktuelle Entwicklungsarbeiten zur automatischen Beantwortung von E-Mail im Servicebereich.

Literatur

[BCP+90] Brown, Peter F., John Cocke, Stephen A Della Pietra, Vincent J. Della Pietra, Frederick Jelinek, John D. Lafferty, Rober L. Mercer und Paul S. Roossin: *A Statistical Approach to Machine Translation.* Computational Linguistics, 16(2):79–85, 1990.

[BH54] Bar-Hillel, Yoshua: *Logical syntax and semantics.* Language, 30:230–237, 1954.

[BKU88] Bouma, Gosse, Esther König und Hans Uszkoreit: *A flexible graph-unification formalism and its application to natural-language processing.* IBM Journal of Reserach and Development, 32(2), 1988.

[Bod95] Bod, Rens: *Enriching Linguistics with Statistics: Performance Models of Natural Language.* Doktorarbeit, Department of Computational Linguistics, Universiteit van Amsterdam, 1995.

[BW92] Briscoe, Ted und Nick Waegner: *Robust Stochastic Parsing Using the Inside-Outside Algorithm.* In: *AAAI Workshop on probabilistically-based NLP Techniques*, 1992.

[BYRN99] Baeza-Yates, Ricardo und Berthier Ribeiro-Neto: *Modern Information Retrieval.* Addison Wesley, Wokingham, UK, 1999.

[CB78] Christian Boitet, Guillaume Pierre, Maurice Quèzel-Ambrunaz: *Manipulation d'arborescences et parallélisme: le système ROBRA.* In: *Proceedings 7th International Conference on Computational Linguistics, Coling '78*, 1978.

[Cho57] CHOMSKY, NOAM: *Syntactic Structures.* Mouton, Hague, 1957.

[Cho65] CHOMSKY, NOAM: *Aspects of the Theory of Syntax.* MIT Press, Cambridge, MA, 1965.

[Cho81] CHOMSKY, NOAM: *Lectures on Government and Binding.* Foris, Dordrecht, 1981.

[Col70] COLMERAUER, ALAIN: *Les systèmes-Q, un formalisme pur analyser et synthétiser des phrases sur ordinateur.* Technischer Bericht, TAUM, Université Montréal, 1970.

[DeR88] DEROSE, S. J.: *Grammatical category disambiguation by statistical optimization.* Computational Linguistics, 14:31–39, 1988.

[Ear70] EARLEY, J.: *An efficient context-free parsing algorithm.* Communications of the ACM, 13(2):94–102, 1970.

[EG89] EVANS, ROGER und GERALD GAZDAR: *Inference in DATR.* In: *Proceedings 4th Conference of the European Chapter of the ACL*, Seiten 66–71, Kopenhagen, 1989.

[HR93] HINDLE, D. und M. ROOTH: *Structural Ambiguity and Lexical Relations.* Computational Linguistics, 19:103–120, 1993.

[Hut86] HUTCHINS, W. J.: *Machine Translation, Past, Present, Future.* Ellis Horwood, Chichester, 1986.

[Jel76] JELINEK, F.: *Continuous Speech Recognition by Statistical Methods.* Proceedings of the IEEE, 64(4):532–556, 1976.

[Kap73] KAPLAN, RONALD M.: *A general syntactic processor.* In: RUSTIN, RANDALL (Herausgeber): *Natural Language Processing*, Seiten 193–241. Algorithmics Press, New York, 1973.

[Kay73] KAY, MARTIN: *The MIND system.* In: RUSTIN, RANDALL (Herausgeber): *Natural Language Processing*, Seiten 155–188. Algorithmics Press, New York, 1973.

[Kos83] KOSKENNIEMI, KIMMO: *Two-Level Model for Morphological Analysis.* In: *Proceedings 8th International Joint Conference on Artificial Intelligence*, Seiten 683–685, Karlsruhe, FRG, 1983.

[KR93] KAMP, H. und U. REYLE: *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Kluwer, Dordrecht, 1993.

[Mag95] MAGERMAN, DAVID M.: *Statistical Decision-Tree Models for Parsing.* In: *Proceedings 33rd Annual Meeting of the ACL*, Cambridge, MA, 1995.

[Mon74] MONTAGUE, RICHARD: *Formal Philosophy. Selected Papers of Richard Montague.* New Haven and London, 1974.

[Per00] PEREIRA, FERNANDO: *Formal grammar and information theory: Together again?* Philosophical Transactions of the Royal Society, 358(1769):1239–1253, 2000.

[Pet80] PETERSON, JAMES L.: *Computer Programs for Spelling Correction*, Band 96 der Reihe *Lecture Notes in Computer Science.* Springer, Berlin, 1980.

[PS94] POLLARD, CARL und IVAN A. SAG: *Head-Driven Phrase Structure Grammar.* The University of Chicago Press, Chicago, 1994.

[Shi86] SHIEBER, STUART M.: *An Introduction to Unification-Based Approaches to Grammar.* CSLI Lecture Notes. Stanford, CA, 1986.

[Shi92] SHIEBER, STUART M.: *Constraint-Based Grammar Formalisms.* Bradford, Cambridge, MA, 1992.

[SW49] SHANNON, CLAUDE und WARREN WEAVER: *The mathematical theory of communication.* University of Illinois Press, Urbana, 1949.

[Tes59] TESNIÈRE, LUCIEN: *Éléments de syntaxe structurale.* Paris, dt. Grundzüge der strukturellen Syntax, Stuttgart 1980 1959.

[Tho82] THOUIN, B.: *The METEO system.* In: LAWSON, V. (Herausgeber): *Practical experience of machine translation*, Seiten 39–44. North-Holland, Amsterdam, 1982.

[Tur50] TURING, ALAN M.: *Computing Machinery and Intelligence.* Mind, 59:433–560, 1950.

[Wea49] WEAVER, WARREN: *Translation.* In: LOCKE, W. N. und A. D. BOOTH (1955) (Herausgeber): *reprinted in: Machine Translation of Languages*, Seiten 15–23. MIT Press, Cambridge, MA, 1949.

[Woo70] WOODS, WILLIAM A.: *Transition Network Grammars for Natural Language Analysis.* Communications of the ACM, 13(10):591–606, 1970.

[WW97] WERMTER, STEFAN und VOLKER WEBER: *SCREEN: Learning a Flat Syntactic and Semantic Spoken Language Analysis Using Artificial Neural Networks.* Journal of Artificial Intelligence Research, 6(1):35–85, 1997.