

Computerlinguistik am Institut für Kommunikationsforschung und Phonetik (IKP) der Universität Bonn

*Winfried Lenders Institut für Kommunikationsforschung und
Phonetik (IKP) der Universität Bonn*

Das Institut für Kommunikationsforschung und Phonetik der Universität Bonn befaßt sich schon seit über 30 Jahren mit Fragen der Anwendung des Computers auf Sprache. Es kann als Pionier auf den Gebieten der Verbindung datenverarbeitender Methoden mit geisteswissenschaftlichen Fragestellungen gelten. Das Institut und die dort betriebene Lehre soll im folgenden mit seiner Entstehungsgeschichte und Ausrichtung vorgestellt werden.

Schon Mitte der 60er Jahre wendete das IKP die damals verfügbaren Großrechner der Universität Bonn zur Bearbeitung von Texten an, besonders zur lexikographischen Erschließung des Gesamtwerks Immanuel Kants und der Werke einiger mittelhochdeutscher Autoren. Aus diesen Forschungsarbeiten hat sich das Lehrprogramm des IKP im Bereich LDV /CL entwickelt. 1972 wurde in Bonn die erste Professur für *Linguistische Datenverarbeitung* in Deutschland eingerichtet. Gleichzeitig wurde Linguistische Datenverarbeitung als Lehrfach in das Curriculum des Faches Kommunikationsforschung und Phonetik integriert. Sichtbarer Ausdruck dieser Integration war ein Lehrplan „Linguistische Datenverarbeitung“, der erste seiner Art in Deutschland. Seitdem hat sich dieses Fach in Bonn wie an vielen Universitäten in Forschung und Lehre zu einer eigenständigen Disziplin entwickelt.

Neben dem starken Bezug zu philologischen und sprachwissenschaftlichen Themen spielen im Bonner Lehrangebot Themen der maschinellen Sprachübersetzung, der natürlichsprachlichen Systeme und der maschinellen Lexikographie von je her eine wesentlichen Rolle. Dabei geht es nicht so sehr um die in diesen Systemen dargestellten Formalismen, sondern um ihren allgemeinen Aufbau und um ihren Charakter als Modelle oder Teilmodelle zwischenmenschlicher Kommunikation. Es geht um die Komponenten dieser Systeme, die insgesamt die

Darstellung des Computers als Sprachbenutzer ermöglichen. In diesem Zusammenhang ist es im Hinblick auf maschinelle Sprachübersetzung und Sprachverstehenssysteme ein Anliegen, die Grenzen der maschinellen Sprachverarbeitung sichtbar zu machen. Leitgedanke ist dabei die Erkenntnis, daß Sprache und Text nicht als isolierte Objekte untersucht und verarbeitet werden können, sondern daß sie immer in ihren kommunikativen Zusammenhängen gesehen werden müssen. Wichtige Themen der Computerlinguistik sind daher auch die zeichen- und handlungstheoretischen Grundlagen der sprachlichen Kommunikation sowie die kognitionswissenschaftlichen und philosophischen Fragen, bei denen es um Bedeutung, Verstehen und Verständigung geht.

Auf eine Besonderheit der Bonner Computerlinguistik ist aufmerksam zu machen, auf ihre enge Verbindung zur Phonetik. Diese ist darauf zurückzuführen, daß das Institut im Jahre 1921 als Institut für Phonetik gegründet worden ist. In den 50er Jahren dieses Jahrhundert wurden Methoden der Informationstheorie eingeführt, die dann zur formalen Sprachbetrachtung und maschinellen Sprachverarbeitung führten. Infolgedessen ist das Studium der Computerlinguistik in Bonn nicht auf die geschriebene Sprache begrenzt, sondern erstreckt sich auch auf Sprache in ihrer gesprochenen Form. Jeder Student, der sich für Computerlinguistik als Studienschwerpunkt entscheidet, muß einen Teil seiner Lehrveranstaltungen im Bereich Phonetik und Sprachsignalverarbeitung absolvieren. Auf diesem Wege werden zumindest die Grundlagen der Phonetik, der Sprachsignalverarbeitung und der akustischen Mensch-Maschine-Kommunikation vermittelt.

Linguistische Datenverarbeitung bzw. Computerlinguistik kann in Bonn nach wie vor mit dem Ziel des Magisterexamens als *Schwerpunkt* des Haupt- oder Nebenfaches Kommunikationsforschung und Phonetik studiert werden. Voraussetzungen ab dem WS 1997/98 kann dann das Fach darüber hinaus als eigenständiges Nebenfach gewählt werden, und zwar sowohl zum Hauptfach Kommunikationsforschung und Phonetik, Schwerpunkt Phonetik, als auch als Nebenfach zu einem beliebigen anderen Hauptfach.

Zur Grundlagenausbildung gehört das Erlernen von Programmiersprachen ebenso wie die Fähigkeit, Computerprogramme für die Verarbeitung von Sprache und Text zu schreiben und anzuwenden. Bevorzugte Programmiersprachen sind Pascal, C, C++ und Prolog. Arbeitsmaterialien stehen in Form von Textkorpora, der Bonner Wortdatenbank, selbstentwickelten Tools für die maschinelle Verarbeitung gesprochener und geschriebener Sprache zur Verfügung. Zur praktischen Ausbildung dienen Programmierübungen, experimentelle Praktika, Transkriptionsübungen etc. Anwendungsgebiete, die in Vorlesungen behandelt, für die es aber auch Praktika und Übungen gibt, sind die nach wie vor expandierenden Bereiche der maschinellen Sprachübersetzung, der maschinellen Lexikographie, der natürlichsprachlichen Systeme und der Verfahren der Mensch-Maschine-Kom-

munikation. Für die praktische Ausbildung stehen ein CIP-Pool und Workstations sowie - für die phonetische Ausbildung - zahlreiche Geräte zur akustischen Analyse zur Verfügung. Die Computerlinguistik ist dem ACO RUM angeschlossen.

Aufgrund des Anspruchs, nicht nur die technischen Aspekte der Sprachverarbeitung zu lehren, sondern Sprache auch als Teil eines kommunikativen Prozesses zu sehen, den es allerdings möglichst genau zu beschreiben gilt, wird sehr viel Wert auf die Vermittlung von kommunikations- und sprachtheoretischen Themen gelegt. So werden die Studierenden mit den Grundbegriffen der Semiotik ebenso vertraut gemacht, wie mit den Grundlagen der Grammatiktheorie und der Sprachphilosophie. Autoren wie Karl Bühler, Charles Morris, Hans Georg Gadamer gehören zum ständigen Repertoire. Und schließlich ist zu erwähnen, daß auch die zahlreichen Bezüge zu anderen Fächern, z.B. zur kognitiven Psychologie und zur Logik, vor allem in Seminaren des Hauptstudium behandelt werden. Diese eher *kommunikationsbezogenen* Themen werden in praktischer Hinsicht durch Ausbildung in Rhetorik, Argumentation und Sprechkunde ergänzt.

Seit etwa einem Jahr nimmt das Fach gemeinsam mit einigen anderen Fächern der Philosophischen Fakultät an Überlegungen und Planungen teil, deren Ziel die Einrichtung eines Nebenfachstudiengangs "Medienwissenschaft" an der Universität Bonn ist. Ein solches Fach stellt - auch unter den Perspektiven der Berufsaussichten - eine ideale Ergänzung zur Computerlinguistik dar.

Immer wieder verwendete Literatur (Auswahl):

1. Grundlegende Primärliteratur:

Ferdinand de Saussure: Cours de Linguistique Generale. 1916.

Charles Morris: Foundations of the Theory of Signs. 1938

Karl Bühler: Sprachtheorie, 1934.

Noam Chomsky: Aspects of a theory of syntax. 1965.

Charles Fillmore: The case for case. In: Bach/Harms (eds.): Universals in Linguistic Theory.

New York 1969.

John Searle: Speech Acts. 1969.

2. Lehrbücher

:

John Lyons: Introduction to theoretical linguistics. Cambridge: Cambridge University Press
1968.

John Lyons: Semantic. Vol. I, II. Cambridge: Cambridge University Press 1977.

Winfried Lenders/Gerd Willee: Linguistische Datenverarbeitung. Ein Lehrbuch. Opladen: Westdeutscher Verlag. 2. Aufl. 1997.

- George W. Smith: Computers and Human Language. New York, Oxford: Oxford University Press, 1991.
- W. John Hutchins/Harold L. Somers: An Introduction to Machine Translation. London etc.: Academic Press, 1992.
- Terry Winograd: Language as a Cognitive Process. 1983.
- Mark G. Sobell: A practical guide to the UNIX system.
- Leon Sterling and E. Shapiro: The Art of PROLOG. Cambridge, Mass.: The MIT Press, 1986.
- Michael A. Covington: Natural Language Processing for Prolog Programmers. Englewood Cliff: Prentice Hall, 1994.
3. *Handbücher:*
- Istv3.n Batori/Winfried Lenders/Wolfgang Putschke (Hrsg.): Handbuch Computerlinguistik. Berlin: De Gruyter, 1989.
- Günther Görz (Hrsg.): Einführung in die künstliche Intelligenz. 1. Aufl. Bonn etc.: Addison Wesley 1993.