

Theoretische Grundlagen der Computerlinguistik: eine Positionsbestimmung*

Burghard Rieger

Vorbemerkung zur Panel-Diskussion

Theoretische Grundlagen der Computerlinguistik

Das Programmkomitee hatte dieses Saarbrücker Symposium unter das Thema *Computerlinguistik und ihre theoretischen Grundlagen* gestellt. Dabei durfte man sicher sein, daß dem ersten Teil dieses thematischen Rahmens durch die *computerlinguistischen* Sektionsvorträge genügt werden würde; inwieweit diese auch die *theoretischen Grundlagen* der Disziplin würden abdecken können, war angesichts der unvermeidlichen Beschränkung der den Vortragenden zugestandenen Redezeiten weniger gewiß. Und da diese Ungewißheit in bezug auf die *theoretischen Grundlagen der Computerlinguistik* möglicherweise auch noch anders als durch den einengenden Zeitrahmen motiviert erschien, wurde eine Podiumsveranstaltung angesetzt, die von Burghard Rieger vorbereitet und organisiert und am 10. März 1988 in Saarbrücken auch moderiert wurde. Was in den Fachvorträgen also möglicherweise zu kurz kam, hier — so die Idee — sollte der programmtechnische Ort und die nötige Zeit bereitgestellt werden, die relevanten Fragen zu stellen und zu diskutieren.

Die folgenden Beiträge von Burghard Rieger (LDV/CL, Universität Trier), Manfred Bierwisch (z.Z. MPI-Psycholinguistik, Nijmegen), Christopher Habel (Informatik, Universität Hamburg), Hans Uszkoreit (LILOG, IBM-Stuttgart) und Wolfgang Wahlster (Informatik, Universität Saarbrücken) geben — auch in dieser Reihenfolge — die von jedem der Panelists zu Beginn der Veranstaltung vorgetragenen Positionen wieder. Die vorliegenden Texte wurden im Hinblick auf die Richtungen und Resultate der anschließenden Diskussion von den Autoren überarbeitet und so zum Teil auf wenige Punkte kondensiert. Sie stellen die persönlichen Sichtweisen der Teilnehmer zum Thema und seiner Diskussion in Saarbrücken dar. Deren quasiobjektivierende Zusammenfassung erschien deswegen entbehrlich, zumal sie — nach

*Erschienen in: Bátori, I.S./Hahn, U./Pinkal, M./Wahlster, W. (Hrsg): *Computerlinguistik und ihre theoretischen Grundlagen*. (KI-Informatik-Fachberichte Bd. 195) Berlin/Heidelberg/New York/London/Paris/Tokyo (Springer) 1988, S. 192–197.

Ansicht der Beiträger — einen nicht nur nicht erwünschten sondern darüber hinaus sicherlich auch ungerechtfertigten Gültigkeitsanspruch zu etablieren helfe.

Computerlinguistik: eine Positionsbestimmung

Das Thema dieser Panel-Veranstaltung könnte suggerieren, daß die *theoretischen Grundlagen* und deren Klärung auch eine Voraussetzung dafür darstellen, daß von *Computerlinguistik* als einer eigenen Disziplin sinnvoll überhaupt gesprochen werden dürfe. Es hätte sich von daher leicht in jene seit mehr als 15 Jahren andauernde Diskussion hinüberwechseln lassen, welche die *Computerlinguistik* von anderen Disziplinen abzugrenzen und unter den verschiedensten fachlichen, forschungs- und ausbildungs-praktischen, nationalen, gesellschaftlichen, etc. Gesichtspunkten zu definieren sucht. Dies sollte hier vermieden werden zugunsten wenigstens des Versuchs einer eher wissenschaftstheoretischen Positionsbestimmung.

Entgegen der Meinung mancher Wissenschaftstheoretiker ist die Genese einer wissenschaftlichen Disziplin ja in den seltensten Fällen — im Falle der CL gewiß nicht — als fortschreitende Entwicklung von den theoretischen Grundlagen, über die Konsolidierung fachlicher Forschung bis hin zur Anwendung von deren Ergebnissen in der Praxis nachzeichnenbar. Die theoretischen Grundlagen, ihre Reflexion und mögliche Revision ergeben sich häufig erst aus der Notwendigkeit, auf (vermeintliche oder tatsächliche) Erschütterungen zu reagieren, die ein Fach oder die in einem fachlichen Zusammenhang wissenschaftlich Tätigen betreffen. Gerade in der Entwicklung des Faches, das heute als *Computational Linguistics* etabliert (und mit *Computerlinguistik* im Deutschen nur ungenau wiedergegeben) ist, haben derartige Erschütterungen (z.B. Machine Translation und ALPAC-Report; Wissensrepräsentation und Declarative/Procedural-Controversy; Strong/Weak-AI-Positionen und ihre Diskussion aus sprachphilosophischer, kognitionspsychologischer und linguistischer Sicht) die in diesem Bereich tätigen Linguisten, Informatiker, Psychologen, etc. zu der Einsicht kommen lassen, daß sie — trotz aller im einzelnen divergierenden Positionen — dennoch auf bestimmten ihnen gemeinsamen theoretischen Grundlagen aufbauen. Eine neuerliche, gerade erst sich abzeichnende Beunruhigung dieser Art scheinen die jüngsten Weiterentwicklungen älterer Modelle zur Verarbeitung nicht-symbolisch repräsentierter Information auszulösen (vgl. etwa Rumelhart/McClelland 1986), die — wie zu Beginn der Informatik — wieder kybernetischen Vorstellungen näherstehen und unter der Bezeichnung *Neuer Konnektionismus* auf entsprechende Neuentwicklungen der Rechnerarchitektur zu *Massiver parallel Verarbeitung* treffen, welche ihrerseits schon zu zahlreichen neuartigen Verarbeitungsmodellen sog. *Neuronaler Netze* geführt haben.

1. In allen Wissenschaften sind Theorien, Modelle und Beschreibungen in der Regel die in eigenen (intersubjektiv oder doch überindividuell entwickelten) Aussagesystemen nachvollziehbar dargestellten Resultate von Bemühungen um das Verständnis und/oder die Erklärung von (beobachteten, erschlossenen oder auch nur vermutete-

ten) Zusammenhängen zwischen Entitäten, die — bei näherer Prüfung — ihrerseits Resultate von Bemühungen um das Verständnis und/oder die Erklärung von Zusammenhängen zwischen Entitäten darstellen, die ihrerseits Resultate von ... und so weiter, bis zu beliebiger Tiefe (oder auch umgekehrt: beliebiger Höhe) des — einzig vom jeweiligen Stand der Forschung abhängigen — Reflexionsniveaus einer beliebigen Disziplin.

Auch ohne Problematisierung der Unterscheidung von *Erklären* und *Verstehen* (als den vermeintlichen Aufgaben der Natur- und Geisteswissenschaften) läßt sich doch differenzieren zwischen *Theorien*, die allgemeine und umfassende Zusammenhänge formulieren, den daraus entwickelten *Modellen*, die kleinere und überschaubare Ausschnitte dieser Zusammenhänge abbilden, und der experimentellen *Erprobung* bzw. praktischen *Anwendung* dieser Modelle, welche als Erhebung und Vergleich von Daten, Überprüfung und Test von Hypothesen, Beschreibung und Analyse von Strukturen, Entwicklung und Simulation von Prozessen, etc. erst Rückschlüsse auf Adäquatheit und explikativen Wert einer Theorie zu ziehen erlauben (vgl. Stachowiak 1973).

2. Das in dieser Skizze wissenschaftlichen Arbeitens angedeutete Prinzip, wonach (fortschreitender oder revidierender) Erkenntnisgewinn als Leistung und Resultat jener — im weitesten Sinne beschreibenden — Aktivitäten erscheint, durch die zunächst chaotische Regellosigkeiten versuchsweise in regelhafte Zusammenhänge und Strukturen überführt werden oder umgekehrt auch zunächst für fundamental gehaltene Einheiten sich möglicherweise als komplexere System- oder Funktionsgefüge darstellen können, erweist sich in seinem kognitiven Kern als *rekursiv*. Denn indem einerseits eine als vorhanden zunächst akzeptierte und als erfahrbar analysierte Wirklichkeit (oder Ausschnitte davon) in ihren Zusammenhängen erkannt und in zunehmend verfeinerten Repräsentationssystemen abgebildet wird, werden andererseits durch eben diese Abbildungen spezifizierte, neue Gegebenheiten allererst konstituiert (vgl. Rieger 1985). Sie können so als (zumindest in Ausschnitten) veränderte Realität zu erneuten Aktivitäten herausfordern, deren kognitive Leistungen und Resultate — im Falle der Wissenschaften — deren *Kontinuität* und *Dynamik* ausmachen.

Dieses Konstitutionsprinzip charakterisiert aber nicht nur die Dynamik, mit der konkurrierende *wissenschaftliche* Theorien einander ablösen oder ihre nach unterschiedlichen Paradigmen konzipierten (formalen, theoretischen, deskriptiven, strukturalen, prozessualen, prozeduralen, etc.) Modelle modifizieren lassen. Dasselbe Prinzip (vgl. Suppes 1982) kann vielmehr allgemein als eine Art Grundmuster zur Kennzeichnung auch solcher *kommunikativer* Prozesse gelten, die in Situationen bestimmter verbaler (diskursiv-dialogischer) Interaktion durch regelgeleiteten Gebrauch von (natürlich-sprachlichen) Zeichen(-ketten) *Bedeutungen* entstehen lassen, welche von den daran beteiligten (Zeichen-)Verwendern auch *verstanden* werden (können): also *Sprache* als kommunikativer *Prozeß*, der auf der *Verwendung* sehr komplexen *Wissens* beruht, das er selbst verändert (vgl. Winograd 1983).

3. Daraus läßt sich in erster Näherung eine Bestimmung auch der *Computerlinguistik* ableiten, welche sie durch die spezifische Kombination von *Forschungsgegenstand*, *Erkenntnisinteresse* und *Untersuchungsmethode* von benachbarten Disziplinen zu unterscheiden erlaubt. Danach ist für die *CL* konstitutiv,

- daß ihr *Forschungsgegenstand* die natürliche Sprache als einen auf komplexem Wissen beruhenden kommunikativen Prozeß (möglicherweise vielfältiger kognitiver Teilprozesse) bestimmt und sie als Performanzphänomene ausweist;
- daß ihr *Erkenntnisinteresse* auf die Strukturen dieses Wissens und die Organisation dieser Prozesse gerichtet ist, die analysiert, repräsentiert und in (Teil-)Modellen (re-)konstruiert werden müssen, um so zu einer simulativen Modellierung des Zusammenwirkens einzelner Komponenten im Prozeßverlauf zu kommen;
- daß ihre *Untersuchungsmethode* eine spezifische Modellierung betrifft, die als operable Darstellungen hypothetischer Entwürfe dessen, was der Fall sein könnte, nicht nur eine Formulierung sondern zugleich die berechenbare Überprüfung dieser Hypothesen bereitstellt.

Für die computerlinguistische Forschung scheint mir dabei nicht nur die Erweiterung des *Forschungsgegenstands* (Sprache als Prozeß) und die für das *Erkenntnisinteresse* leitende Hypothese (Wissensbasiertheit kommunikativer Prozesse) wichtig zu sein. Als entscheidender darf vielmehr eine epistemologische Dimension gelten, welche die *Untersuchungsmethode* und die sie auszeichnende neue Modellbildung (Berechenbarkeits-Postulat) betrifft. Sie erlaubt es,

- von der Vielzahl vorliegender *Daten* aufgrund beobachtbarer, regelhafter Zusammenhänge über deren Klassifikation zur formalen Beschreibung von *Strukturen* zu gelangen,
- durch die Deutung solcher *Strukturen* als Resultate von kognitiven *Prozessen* zu deren Analyse beizutragen, welche ein Regelwissen voraussetzt, dessen Anwendung (und Veränderung) gerade diese Prozesse auszeichnet,
- derartige *Prozesse* unabhängig von ihrer zeitlichen Dauer in Form von *Prozeduren* abstrakt zu repräsentieren, was deren (modularen) Aufbau zu studieren, zu erproben und zu modifizieren erlaubt,
- solche *Prozeduren* — in geeigneten formalen Sprachen und auf entsprechenden Maschinen realisiert — wieder zu in der Zeit ablaufenden *Prozessen* werden zu lassen, und
- anhand solcher auf geeigneten *Daten* operierenden *Prozesse*, jene beobachtbaren Zusammenhänge als Prozeßresultate intersubjektiv überprüfbar zu reproduzieren, deren *Regelmäßigkeit* beim Erkennen (Analysieren/Repräsentieren/Verstehen/etc.) von *Strukturen* generell vorausgesetzt werden muß.

Unsere Skizze des sogenannten *kognitiven Paradigmas* einer wissensbasierten Sprachverarbeitung hat eine quasi-empirische Modellbildung zu entwickeln versucht, die nicht gedeutet zu werden braucht als allmähliche Annäherung an das zu modellierende Original, sondern besser faßbar ist als schrittweise Entfernung von dem, was — nachweisbar in und überprüfbar durch Modellierungen — *nicht* der Fall ist. Diese gegenüber sowohl *symbolisch* als auch *verteilt* repräsentierter Information neutrale Position, wonach kognitive Leistungen allgemein als prozedurales Resultat fortschreitender Strukturierungen aufgefaßt werden (Winograd/Flores 1986), scheint derzeit noch am ehesten jene wissenschaftstheoretische Basis liefern zu können, auf die sowohl ältere wie neuere computerlinguistische Ansätze regelgrammatischer Analyse- und Erkennungs-Prozesse natürlich-sprachlicher Strukturen sich beziehen lassen, auf der ebenso aber auch symbolinterpretierende Ansätze zur Verarbeitung natürlicher Sprache als Verstehenssysteme fußen, die im weniger grundlagenorientierten Bereich der älteren und neueren Forschungen zur Künstlichen Intelligenz entstanden. Für beide ist zu hoffen, daß die Herausforderung, die mit der Emulation *paralleler* Verarbeitungsprozesse von *verteilt* repräsentierter Information gerade im *kognitiven* Bereich verbunden ist, nicht übersehen sondern angenommen wird.

Literatur

- Rieger, B. (1985): *Einleitung* zu Rieger (Hrsg): Dynamik in der Bedeutungskonstitution, (Papiere zur Textlinguistik Bd. 46). Hamburg (Buske), S. 1–17
- Rumelhart, D.E./McClelland, J.A./PDP-Research Group (1986): Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Cambridge, MA (MIT Press)
- Stachowiak, H. (1973): Allgemeine Modelltheorie. Wien/New York (Springer)
- Suppes, P. (1982): *Variable-Free Semantics with Remarks on Procedural Extensions* in: T.W. Simon/R.J. Scholes (Hrsg): Language, Mind, and Brain. Hillsdale, NJ (Lawrence Erlbaum), S. 21–34
- Winograd, T. (1983): Language as a Cognitive Process, Vol. 1: Syntax. Reading, MA (Addison-Wesley)
- Winograd, T./Flores, F. (1986): Understanding Computers and Cognition. A New Foundation for Design. Norwood, NJ (Ablex)