

COMPUTERLINGUISTIK UND INFORMATIK SELBSTVERSTÄNDNIS EINER DISZIPLIN*

Prof. Dr. István Bátori

Universität Koblenz-Landau, Abteilung Koblenz

Institut für Computerlinguistik

CL und Informatik sind untereinander auf mehreren Ebenen verbunden:

1. Für beide Disziplinen sind Sprachen, sprachliche Formulierungen, Sprachverstehen, sprachliche Repräsentationen, die Benutzung der Sprachen in Kommunikation, der sprachliche Verarbeitungsprozess von zentraler Bedeutung. Sprachen und Computer, genauer: die von-Neumann-Typ-Rechner sind symbolverarbeitende Einrichtungen: Wesentlich ist dabei nicht die physikalische Fluktuation der Zustände, sondern die vermittelten Inhalte. An dieser wesenhaften Ähnlichkeit ändert auch der Umstand nichts, daß die natürlichen Sprachen einen anderen (viel höheren) Komplexitätsgrad besitzen, als die (künstlichen) Sprachen, die im Umgang mit den Rechnern benutzt werden.

2. Informatik und CL sind auch methodisch eng verbunden. In der CL werden die Konzepte, Beschreibungstechniken und die Software-Instrumente der Informatik gezielt aufgegriffen

* Eröffnungspanel auf der Jahrestagung der GLDV, in Kiel, 3.3.1993, nach der Diskussion überarbeitete Fassung

und benutzt. Mit Hilfe der Informatik-Instrumente strebt die CL eine hohe Beschreibungspräzision an. Da die Komplexität der natürlichen Sprachen, die der Programmiersprachen übersteigt, reichen die entlehnten formalen Mittel für die adäquate Erfassung der natürlichen Sprachen nicht aus. Daher müssen die in der CL entwickelten Systeme über die Informatik-Vorbilder hinausgehen. Beispiel: Compiler-Compiler YACC und die split stack parser von Tomita oder die Entwicklung von relationalen Datenbanken zu Systemen mit nichtmonotonen Vererbungstechniken im lexikalischen Bereich.

3. Informatik und CL sind zweifelsohne auch durch die professionelle Computerbenutzung verbunden. Allerdings ist diese Verbindung nicht einmalig. Zeitgenössische Wissenschaft arbeitet generell mit professioneller Rechnerunterstützung. Computerunterstützung allein könnte die einmalig enge Verbindung zwischen Informatik und CL nicht erklären. Die instrumentale Nutzung der Computer für die CL als Wissenschaft ist nicht konstitutiv.

Die Beziehung zwischen CL und Informatik ist nicht symmetrisch: von einem Infor-

matiker wird nicht erwartet, daß er über linguistische oder CL-Kenntnisse verfügt; ein Computerlinguist muß dahingegen die Verarbeitungstechniken der Informatik für Sprachen kennen. Verarbeitungstechniken für Sprachen sind allerdings nicht - wie oben bereits angedeutet - eine isolierbare Teilmenge der Informatik, sondern Informatik schlechthin. Von einem (idealen) Computerlinguisten wird eine Doppelkompetenz (Informatik und Linguistik) erwartet. Es wäre unsinnig, dieses oder jenes Teilgebiet der Informatik für die CL als irrelevant oder nutzlos zu erklären. Je umfangreicher das Informatikwissen ist, umso größer ist die Katalisationsfläche für innovative Entwicklungen.

An dieser Stelle soll hervorgehoben werden, daß die CL eine forschungsorientierte Disziplin ist:

1. CL ist noch im Aufbau, wir stehen hier eher am Anfang einer Entwicklung und nicht mitten in einem saturierten Software-Feld. Die heutigen CL-Systeme sind Prototypen und funktionieren unbefriedigend.
2. Die CL-Absolventen finden typischerweise Stellen in Forschungs- und Entwicklungsprojekten.
3. Auch die CL-Ausbildung ist forschungsorientiert.

Meine Beobachtungen stützen sich auf die Daten unserer Absolventen in Koblenz. Im Prinzip kann ein Universtätsabsolvent nach dem Studium drei Wege einschlagen: 1. den Weg der Praxis, 2. den der Forschung und 3. den der Lehre. Die CL bietet z. Zt. zwei aus den obigen drei Möglichkeiten:

1. eine Stelle in der Forschung (genauer R&D) oder
2. eine Nachwuchsstelle in der Lehre.

Im Gegensatz zu Tiermedizin, BWL oder Jura, existiert die Praxis als dritte Karrieremöglichkeit nicht. Unsere Absolventen, insoweit sie computerlinguistisch "im Beruf" bleiben, haben typischerweise Stellen in R&D-Abteilungen, auch wenn sie in der Wirtschaft tätig sind.

Wenn unsere Koblenzer Informatikstudenten mit Anwendungsfach CL fertig sind, gibt es für sie drei Möglichkeiten:

1. Sie nehmen eine (lukrative) Stelle in der Wirtschaft oder in der Verwaltung an und hängen CL an den Nagel.
2. Sie bekommen eine Stelle in einem CL-Projekt und machen Forschung full time.
3. Sie bleiben in der Lehre (und machen Forschung nebenbei).

Schließlich möchte ich darauf hinweisen, daß Forschung in der CL und Forschung in der Linguistik stark konvergieren. Es gibt eine wachsende Anzahl von Modellen, die sowohl in der CL als auch in der Linguistik benutzt werden. LFG, GPSG und HPSG sind die prominentesten. Daher ist es nicht zweckmäßig, eine scharfe Trennlinie zwischen echter linguistischer Forschung und CL-orientierter Grundlagenforschung zu ziehen.

Computerlinguistik ist die moderne Form der linguistischen Forschung. Die Konvergenz der beiden Felder (CL und Linguistik) ist positiv zu bewerten und es ist auch angebracht, dieser Sachlage Rechnung zu tragen.