

# Zur Konturierung des Faches Computerlinguistik

Udo Hahn, Matthias Heyn, Ludwig Hitzenberger  
Bernhard Kelle, Hans-Dieter Lutz, Klaus Wothke

Die Zahl der Computerlinguistik-Studiengänge nimmt ständig zu. Die Unterschiedlichkeit der fachlichen und institutionellen Einbindung und die Verschiedenartigkeit der Studiengangstypen<sup>3)</sup> mit entsprechend verschiedenen Abschlüssen erschwert die Überschaubarkeit und Vergleichbarkeit der Studienangebote für

- Lehrende
- Studienanfänger und fortgeschrittene Studierende
- Studienberater
- eventuelle Arbeitgeber.

Auch über Ausbildungsziele, Berufsqualifikation und Berufsprofil von CL-Absolventen bestehen bisher keine festumrissenen, sondern allenfalls vage und meist differierende Vorstellungen.

Auf Anregung der Arbeitskreise "Ausbildung und Berufsperspektiven" sowie "LDV und Nachbarn" haben daher Beirat und Vorstand der GLDV im März 1986 Arbeitstreffen interessierter Mitglieder initiiert, deren Ziel es war, etablierte und im Aufbau begriffene Studiengänge, deren Inhalte, Ausbildungsprofile sowie die sich daraus ergebende Art von Berufsqualifikation zu analysieren. Langfristige Zielvorstellungen für diese Arbeit waren,

- das Profil des Faches CL genauer zu konturieren und damit die Berufsqualifikation eines CL-Absolventen beschreibbar zu machen,
- inhaltliche Minimalvorstellungen (je Studiengangstypus) zu entwickeln,
- Lehr- und Lernziele einzelner Lehrveranstaltungen zu beschreiben,
- die "Durchlässigkeit" (Studienort-Wechsel) zwischen einzelnen Studiengängen zu erhöhen oder überhaupt erst zu ermöglichen,
- die Überschaubarkeit für die Studienberatung zu verbessern.

Für die gestellte Aufgabe erwies es sich als zweckmäßig, einen Diplomstudiengang Computerlinguistik (ohne eine institutionelle Einbindung) als *"tertium comparationis"* (s. Kap. 3) anzusetzen, zu dem die existierenden oder geplanten Studiengänge dann in Beziehung gesetzt werden können. Einig war man sich auch über die Fachbezeichnung "Computerlinguistik" (CL).

## **Empfehlungen [der GLDV] zur Konturierung des Faches Computerlinguistik:**

- 1 Inhalte und Methoden des Faches Computerlinguistik
- 2 Ausbildungsprofil und Fähigkeiten von Computerlinguisten

### 2.1 Ausbildungsprofil und Berufsprofil

### 2.2 Fähigkeiten des Computerlinguisten

- 3 Computerlinguistik als Diplomstudiengang
- 4 Computerlinguistik als Hauptfach eines Magisterstudienganges
- 5 Computerlinguistik als Nebenfach
- 6 Computerlinguistik als Teil in einem anderen Fach
- 7 Weiterbildungsstudiengänge in CL

## 1 Inhalte und Methoden des Faches Computerlinguistik

### Grundlagen

Konstitutive Elemente der Computerlinguistik sind:

- a) die Beschreibung natürlicher Sprache,
- b) die Ausrichtung dieser Beschreibung an der Verarbeitung durch Computer,
- c) die methodische Fundierung dieser Beschreibung durch den Bezug auf theoretische Prinzipien der Linguistik und Informatik.

Demzufolge bezieht die CL einige ihrer Methoden und Inhalte aus ihren Nachbardisziplinen Linguistik und Informatik, die sie einer eigenen Zielsetzung<sup>4)</sup> folgend in ihren Methoden- und Inhaltekanon integriert. In einer ersten Näherung sind dies:

### "Vorfeld Linguistik"

- 1) Grundlegende Methoden empirischen Arbeitens und Kenntnis ihrer Ergebnisse im Bereich deskriptiver, struktureller, logischer oder quantitativer Linguistik
- 2) Modelle zur Phonologie, Morphologie, Syntax, Semantik, Pragmatik
- 3) Theoriegeleitete Beschreibung ausgewählter sprachlicher Phänomene (Nominalphrasen, Quantifikation, Koordination. . .)
- 4) Validierung von (Teil-)Theorien des Sprachsystems

### "Vorfeld Informatik"

- 1) Mathematische Grundlagen (Algebra, Graphentheorie, formale Logik, Modelltheorie, Statistik); informatische Grundlagen (Rechnerarchitektur, Compilerbau, Betriebssysteme, Programmiersprachen, Software Engineering, Datenstrukturen, Algorithmentheorie)
- 2) Theorie formaler Grammatiken/Sprachen, Automatentheorie
- 3) DB-Systeme
- 4) Problemlösungsverfahren (Suchen, Planen, Lernen)
- 5) Testtheorie (Software- Validierung)

### Zentrale CL-Bereiche

- 1) Grundlagen der Mensch-Computer-Interaktion (MCI) (Interaktionstheorie, Kommunikationsforschung, Kognitionswissenschaft; Dialogmodelle, Argumentationsmodelle, Benutzermodelle)
- 2) Parsing-Verfahren, Parsing-Strategien und grammatische Spezifikationsprachen<sup>5)</sup> 3)  
Wissensrepräsentationsformalismen und -verarbeitungsprozeduren
- 4) Architekturen von natürlichsprachlichen Systemen
- 5) Systemevaluierung (Effektivität, Effizienz und empirische Adäquatheit implementierter natürlichsprachlicher Systeme, Technologiefolgenabschätzung)

## 2 Ausbildungsprofil und Fähigkeiten von Computerlinguisten<sup>6)</sup>

### 2.1 Ausbildungsprofil und Berufsprofil

Das Ausbildungsprofil eines Computerlinguisten setzt sich zusammen aus

- a) Grundwissen bezüglich der Computerlinguistik mit ihren Vorfeldern Linguistik und Informatik und
- b) Vertiefungswissen, das in der Regel durch einen Anwendungsbereich bestimmt wird, der wiederum die Hinzunahme von Anwendungswissen und Inhalten eines Nebengebietes erfordern kann. Dieses Vertiefungswissen ist auf mögliche Berufsfelder bezogen.

Grundwissen und Vertiefungswissen werden ergänzt durch praktisches Erfahrungswissen (interne und externe Praktika, Übungen, u.a.).

Vertiefungswissen kann in den folgenden Anwendungsbereichen (exemplarische, nicht vollständige Liste) erworben werden:

- Maschinelle Textverarbeitung / Bürokommunikation
- Maschinelle / Maschinengestützte Lexikographie
- Information und Dokumentation
- Maschinelle / Maschinengestützte Übersetzung
- Natürlichsprachliche Frage-Antwort-Systeme / Dialogschnittstellen
- Computerunterstützter Unterricht
- Entwicklung CL-spezifischer Tools

Entsprechend qualifiziert der hier vorgeschlagene Studiengang für Tätigkeiten im Bereich von Forschungseinrichtungen, von Fachinformations- und -dokumentationseinrichtungen, des Verlagswesens, in einschlägigen Softwarehäusern, im EDV-Ausbildungsbereich, u. a.

Grund- und Vertiefungswissen gemeinsam setzen den CL-Absolventen in den Stand, die Anforderungen des folgenden Fähigkeitenkataloges zu erfüllen.

### 2.2 Fähigkeiten des Computerlinguisten

Der Computerlinguist muß vorhandenes CL-Wissen adäquat einsetzen können. Das bedeutet:

2.2.1 Der Computerlinguist muß in der Lage sein zu entscheiden, ob ein sprachliches Problem auf dem gegenwärtigen Stand des Know-How mithilfe des Rechners sinnvoll zu lösen ist.

2.2.2 Der Computerlinguist muß die Fähigkeit haben, offene Problemfelder zu erkennen und neue Problemlösungen zu suchen.

2.2.3 Der Computerlinguist muß in der Lage sein, selbständig sprachliche Problemstellungen so aufzubereiten, daß sie sich mit dem Computer lösen lassen. Das bedeutet:

- Er muß mit fachspezifischen Verfahren eine Lösung finden und dabei gezielt das vorhandene sprachliche linguistische und informatische Wissen einsetzen (Problemanalyse).
- In einem zweiten Schritt muß der Lösungsweg in einer für die maschinelle Verarbeitung geeigneten Form dargestellt werden (Spezifikation).
- Er muß den Lösungsweg in einem dritten Schritt in ein Programm(-System) umsetzen können (Implementation).

2.2.4 In interdisziplinären Softwareentwicklungsprojekten muß der Computerlinguist zwischen den verschiedenen Ansätzen der an der Lösung eines computerlinguistischen Problems beteiligten Fachrichtungen vermitteln können. Kritische Bereiche sind dabei die verschiedenen Problemlösungsstile, Bewertungsmuster und nicht zuletzt die Fachterminologien der daran beteiligten Disziplinen.

- 2.2.5 Ein Computerlinguist muß in der Lage sein, Lösungen in der CL zu bewerten und diese Bewertung explizit zu begründen. Das setzt die Kenntnis von spezifischen Eigenschaften sowie von Vor- und Nachteilen unterschiedlicher Lösungswege voraus und die Kenntnis des innovativen Elementes bezüglich vorangegangener Lösungen.
- 2.2.6 Ein Computerlinguist muß weiterhin die Fähigkeit haben, die sprachlich fundierten Teilaspekte von Anwenderproblemen zu erkennen, in Kooperation mit dem Anwender zu beurteilen und gegebenenfalls einer Lösung zuzuführen.

Schließlich muß der Computerlinguist in der Lage sein, an der Etablierung eines in der Öffentlichkeit noch nicht geläufigen Berufsbildes mitzuarbeiten.

### 3 Computerlinguistik als Diplomstudiengang

Der Diplomstudiengang teilt sich in ein

- Grundstudium, das mit dem Vordiplom abgeschlossen wird, und
- ein Hauptstudium, das mit dem Hauptdiplom abgeschlossen wird.

Der Diplomstudiengang ist ein berufsqualifizierender Studiengang im üblichen Sinne.

Die Studiendauer beträgt 10 Semester + 1 Semester<sup>7)</sup> die Diplomarbeit.

Im Grundstudium und zu Beginn des Hauptstudiums liegen die Pflichtveranstaltungen; diese werden im Laufe des Hauptstudiums von Wahlpflichtveranstaltungen abgelöst.

Das Verhältnis zwischen Pflichtveranstaltungen, Wahlpflichtveranstaltungen und Wahlpflichtveranstaltungen aus dem Vertiefungsbereich (Anwendungssysteme) soll stets 3 : 2 : 1 betragen.

Der hohe Anteil von Wahlpflichtveranstaltungen soll die Möglichkeit eröffnen, örtliche und institutionelle Besonderheiten an den einzelnen Studienorten berücksichtigen zu können.

Der Diplomstudiengang umfaßt 120 SWS; dabei entfallen auf Pflichtveranstaltungen 60 SWS, auf Wahlpflichtveranstaltungen ebenfalls 60 SWS (dabei wiederum auf Wahlpflichtveranstaltungen aus dem Vertiefungsbereich 20 SWS).

Zum Fach Computerlinguistik ist ein Nebenfach mit 24 - 40 SWS zu studieren. Als Nebenfächer kommen in Frage alle Fächer, die starke inhaltliche Bezüge zur Computerlinguistik haben (Informatik, Linguistik, Mathematik, Psychologie) und darüber hinaus alle Fächer, in denen CL-Bedarf im weitesten Sinne besteht. Lokale Gremien entscheiden über die Zulassung der Nebenfächer.

Im folgenden wird ein Rahmen für einen Diplomstudiengang angegeben, der sich auf den in Kapitel 1 angegebenen Methoden- und Inhaltekanon bezieht.

Dabei wird besonderer Wert darauf gelegt, daß die beiden Vorfächer Linguistik (als rekonstruktive Wissenschaft) und Informatik (als konstruktive Wissenschaft) nicht unvermittelt nebeneinander stehen, sondern bereits im ersten Semester durch eine eigene Veranstaltung miteinander verklammert werden, die dem interdisziplinären Charakter der CL entspricht.

Vorangestellt werden einige Erläuterungen zu den im Diplomstudiengang angesprochenen Veranstaltungsformen und Studienleistungen und einige generelle veranstaltungsunabhängige Empfehlungen.

#### Veranstaltungsformen

Alle Veranstaltungen, die keinen Vorlesungscharakter haben, verfolgen den Zweck, theoretisches Wissen durch praktische Fähigkeiten zu ergänzen bzw. in praktische Fertigkeiten umzusetzen; sie sind aber auch dazu gedacht, den Studierenden Gelegenheit zu geben, Ergebnisse ihrer Studien und Erfahrungen zu formulieren und auf diesem Wege auch ihre Kooperations- und sprachliche Ausdrucksfähigkeit auszubilden.

**Vorlesung:**

Vorlesungen dienen zur Einführung in größere Teilgebiete der Computerlinguistik, Linguistik und Informatik. In zusammenhängender Darstellung werden wissenschaftliches Grund- und Spezialwissen vermittelt und die zugehörigen Arbeitstechniken und Methoden exemplarisch vorgestellt.

**Übung:**

Übungen begleiten Vorlesungen. Dabei ist ihre primäre Aufgabe nicht darin zu sehen, den Stoff der Vorlesung zu wiederholen, sondern darin, daß sich die Studierenden mit dem Stoff intensiv und selbsttätig auseinandersetzen. Im Mittelpunkt steht die Vermittlung und Einübung praktischer Fertigkeiten. Von dieser Ausrichtung her ist es notwendig, Übungen in Gruppen (mit max. 20 Teilnehmern) abzuhalten.

Es werden Hausaufgaben gestellt, die selbständig schriftlich zu bearbeiten sind. In den Übungsstunden besteht die Möglichkeit, Aufgabenlösungen vorzutragen, sachliche Probleme und Verständnisschwierigkeiten im Zusammenhang mit dem Stoff der Vorlesung und/oder Übung unter fachkundiger Leitung zu diskutieren. Übungen sind wesentliche und notwendige Ergänzung zu den Vorlesungen.

**Proseminar:**

Proseminare sind die Seminare des Grundstudiums. Dabei erhalten die Studierenden bereits im Grundstudium die Möglichkeit, selbständig ein eng begrenztes Thema zu erarbeiten, darzustellen und in einem kurzen Referat vorzutragen. Damit werden sowohl die in Übungen erworbenen methodischen Kenntnisse vertieft als auch die Fähigkeit geübt, Probleme und Sachverhalte darzustellen und zu vermitteln.

**Seminar:**

In den Seminaren werden Spezialbereiche bzw. spezielle Problemstellungen behandelt. Die Studierenden erhalten zur Aufgabe, ein komplexeres Thema anhand von Originalliteratur und weiterführender Literatur zu erarbeiten, kritisch-referierend darzustellen und in einem längeren Referat zur Diskussion zu stellen. Damit wird die Fähigkeit zur Darstellung und Vermittlung von Sachverhalten weiter vertieft.

**Projektübung:**

In Projektübungen soll eine umgrenzte, aber nicht-triviale Aufgabenstellung von kleinen Teams (3-4 Mitglieder) schrittweise und systematisch er- und bearbeitet werden: von der Definition der Problemstellung über die Zerlegung in Teilprobleme und die Abgrenzung und Bereitstellung des empirischen (sprachlichen) Materials und die Erarbeitung der notwendigen Beschreibungssprache(n)... bis hin zur Dokumentation von Stärken und Schwächen der vorgeschlagenen Problemlösung.

**Projektseminar:**

Die Projektseminare dienen sowohl dem Verständnis wichtiger neuerer Entwicklungen und aktueller Forschungsprobleme im Bereich der CL als auch der Vorbereitung auf die Berufspraxis. Hierbei arbeiten Gruppen von Studierenden an der Lösung umfangreicher CL-Probleme (von der linguistischen Analyse bis hin zur Implementierung) unter Anleitung von Dozenten. Neben dem Erwerb und der Festigung von CL-Kenntnissen soll die Arbeit in Projektseminaren vor allem folgende Themen berücksichtigen:

- psychologische und organisatorische Probleme der Teamarbeit,
- Rechnerunterstützung zur Organisation und Dokumentation der Arbeitsergebnisse,
- unterschiedliche Kommunikationsebenen und -formen bei der Software-Entwicklung und
- Konstruktion, wie Kommunikation innerhalb der Gruppe sowie mit Auftraggebern und Benutzern und Weiterentwicklern,
- Erfahrungen im Software-Entwurf, Abschätzung der Auswirkungen alternativer Entwurfsentscheidungen,

- Problematik der Aufspaltung in Teilaufgaben (Arbeitsteilung wie Modularisierung) und der Terminplanung,
- Test-, Evaluierungs- und Integrationsproblematik.

## **Studienleistungen**

### **Studienarbeit:**

Eine Studienarbeit steht bzgl. Komplexität und Zeitaufwand zwischen Seminararbeit und Diplomarbeit. Unter Betreuung eines Dozenten arbeiten die Studierenden ein Thema nach wissenschaftlichen Methoden auf, die im Studium bis zu diesem Zeitpunkt vermittelt worden sind. Dabei sollen nicht nur die theoretischen Kenntnisse, sondern auch Programmierkenntnisse und -fertigkeiten unter Beweis gestellt werden.

### **Diplomarbeit:**

Die Diplomarbeit ist die unter Anleitung eines Hochschullehrers anzufertigende schriftliche Abschlußarbeit. Sie stellt den wesentlichen Studienschwerpunkt dar. Mit ihr sollen die Studierenden zeigen, daß sie in der Lage sind, eine nicht zu eng gewählte Aufgabenstellung innerhalb einer vorgegebenen Frist nach wissenschaftlichen Methoden selbständig zu bearbeiten und adäquat darzustellen.

## **Empfehlungen**

1. Es wird empfohlen, ein Betriebspraktikum vorzusehen. Dieses Praktikum, das nicht zu den von der Hochschule anzubietenden Veranstaltungen gehört, soll dem/der Studierenden Einblicke in die Berufspraxis ermöglichen, die die Hochschule weder theoretisch noch praktisch vermitteln kann.
2. Die Durchführung von Übungen sollte durch Tutoren unterstützt werden.

## Rahmen für einen Diplomstudiengang Computerlinguistik

### Das Grundstudium umfaßt im:

1. Sem.	Einführung in die Computerlinguistik I, (V + Ü) Einführung in die Grundlage der Linguistik, (V + Ü) Mathematische und informatische Grundlagen I, (V+Ü) Programmierpraxis (V+Ü)	4SWS 4SWS 4SWS 4 SWS = 16 SWS
2. Sem.	Einführung in die Computerlinguistik II (V+Ü) Grundlegende Methoden empirischen Arbeitens (V + Ü) Mathematische und informatische Grundlagen II (V+Ü)	4SWS 4SWS 4 SWS = 12 SWS
3. Sem.	Theoriegeleitete Beschreibung ausgewählter sprachlicher Phänomene I (Projektübung) .. Programmiersprachen und Programmierkonzepte (V + U)	8 SWS 4 SWS = 12 SWS
4. Sem.	Theoriegeleitete Beschreibung ausgewählter sprachlicher Phänomene II (Proseminar) Datenbank-Systeme (V+Ü) Problemlösungsverfahren (V + Ü)	2 SWS 4SWS 4 SWS = 10 SWS
		: 50SWS

### Das Hauptstudium umfaßt 70 SWS:

- 34 SWS für Lehrveranstaltungen u.a. zu folgenden thematischen Bereichen:
  - Modellierung von Mensch-Computer-Interaktion
    - Parsingverfahren und -strategien
    - Wissensrepräsentationsverfahren
    - Systemarchitekturen
    - Evaluierung
- 20 SWS für Lehrveranstaltungen zu Vertiefungswissen aus einem Anwendungsbe-  
reich,
- 6 SWS für die Studienarbeit,
- 10 SWS für Projektseminar.

gesamt: 120 SWS

## Inhalte der Veranstaltungen (Wesentliches in Stichworten)

### **EINFÜHRUNG IN DIE COMPUTERLINGUISTIK I+ II**

Einordnung der CL; CL und ihre Anwendungsbereiche; Basisalgorithmen der Sprachverarbeitung, Grundlagen des Parsing von natürlichen Sprachen; Grundlagen der Interaktionstheorie; zwischenmenschliche Kommunikation; Mensch-Computer- Interaktion (MCI); gesellschaftliche Auswirkungen.

Grundlagen von Informationssystemen: Informationsbegriff; Funktionsklassen von Informationssystemen (Referenzretrieval, Datenretrieval, inferenzbasierte Systeme), natürlichsprachliche Komponenten (Indexing, Klassifikation, Abstracting, Datenbanksynthese, Faktenextraktion, Maschinelle Übersetzung, Textgenerierung), Interaktionsklassen von Informationssystemen (Frage-Antwort-Systeme, Dialogsysteme); ...

### **EINFÜHRUNG IN DIE GRUNDLAGEN DER LINGUISTIK**

Linguistik als empirische Wissenschaft; deskriptive Linguistik, formale Linguistik, kontrastive Linguistik, ...; Sprache und Welt; sprachliche Einheiten; linguistische Subsysteme (Phonologie, Morphologie, Syntax, Semantik, Pragmatik), ...

### **MATHEMATISCHE UND INFORMATISCHE GRUNDLAGEN I + II**

Algebra; Graphentheorie; formale Logik (incl. Theorem-Beweis-Verfahren); Modelltheorie; Statistik; Algorithmen und Datenstrukturen; Theorie formaler Sprachen und Automatentheorie; ...

### **PROGRAMMIERPRAXIS**

Grundlagen der strukturierten Programmierung anhand einer ersten Programmiersprache; praktischer Umgang mit Rechnern und Betriebssystemen;...

### **GRUNDLEGENDE METHODEN EMPIRISCHEN ARBEITENS**

Beobachtung, Beschreibung, Erklärung; Datenerhebung, Transkription; Aufbau von Corpora, Corpusanalyse (Distributionsanalyse, IC-Analyse, H'); Klassifizierung, Hypothesenbildung, Abstraktion, Generalisierung, Theorienbildung, Theorienüberprüfung;

### **THEORIEGELEITETE BESCHREIBUNG**

#### **AUSGEWÄHLTER SPRACHLICHER PHÄNOMENE I+II**

z.B.: Typisierung von Sprechakten nach AUSTIN, HABERMAS, SEARLE, ...;

z.B.: syntaktisch/semantische Beschreibung von Verben im Rahmen der Valenztheorie, der TG, der Kasusgrammatik, ...;

z.B.: Kennzeichnung und ihre sprachliche Realisierungsmöglichkeiten (Referenzsemantik, Syntax, Wortbildung, ...); ...

### **DB-SYSTEME**

Grundlegende Konzepte von DB-Systemen; conceptual modelling; logische Datenorganisation; Datenmodelle; Sprachen für Datenmodelle (DDL,DML); Relationentheorie; physikalische Datenorganisation; Integrität; praktischer Umgang mit DB-Systemen; ...



**PROGRAMMIERSPRACHEN UND PROGRAMMIERKONZEPTE**

Über die erste Programmiersprache hinausgehende Konzepte, z.B. funktionaler Programmiersprachen (wie LISP), objektorientierter Programmiersprachen (wie SMALLTALK), logikorientierter Programmiersprachen (wie PROLOG), ...; Programmierumgebungen; Prototyping;...

**PROBLEMLÖSUNGSVERFAHREN**

Pattern-Matching; generate-and-test; means-end-Analysis; heuristische Suche ...; formale Schlußverfahren (Deduktion, Induktion, Abduktion); Lernverfahren; Planungsverfahren; ...

**MODELLIERUNG VON MENSCH-COMPUTER-INTERAKTION**

Kooperationsprinzipien; adaptive Systeme: Benutzermodellierung, Überzeugungssysteme; Sprechakt-Erkennung, Handlungsplangenerierung, Sprechaktgenerierung; Argumentationsmodelle, Dialogmodellierung; adaptierbare Systeme; ...

**WISSENSREPRÄSENTATIONSVERFAHREN**

Wissensrepräsentationsformalismen: Logiken; Semantische Netze, Frames, Scripts; Produktionssysteme; Wissensrepräsentationssprachen (FRL, KL-ONE, KRYPTON, OPS5, ...);

Beschreibungsformalismen für natürliche Sprachen: Dependenz, TG, GPSG, LFG, G & B, u.a.; grammatische Spezifikationssprachen, Unifikation;

**PARSING-VERFAHREN UND -STRATEGIEN**

Prinzipien des Compilerbaus; deterministisches, probabilistisches Parsing; top-down, bottom-up, island, chart, ...; unifikationsbasiertes Parsing; Parser-Generatoren; ...

**SYSTEMARCHITEKTUREN**

sequentielle, parallele, hierarchische, heterarchische, blackboard-Architektur; ...;

Architekturen von natürlichsprachlichen Systemen; ...

**EVALUIERUNG**

Leistungstests (Effizienz, Effektivität, Funktionalität) von natürlichsprachlichen Systemen; Prinzipien von Experimentaufbau und -durchführung bei der Entwicklung von natürlichsprachlichen Systemen...

**4 Computerlinguistik als Hauptfach eines Magisterstudiengangs**

Der Magisterstudiengang teilt sich in ein

- Grundstudium, das mit der Zwischenprüfung abgeschlossen wird, und
- ein Hauptstudium, das mit der Magisterprüfung abgeschlossen wird.

Der Magisterstudiengang mit CL als Hauptfach ist ein dem Diplomstudiengang entsprechend berufsqualifizierender Studiengang. Die Studiendauer beträgt mindestens 8 Semester + 1 Semester für die Magisterarbeit.

Für Magisterstudiengänge gibt es zwei verschiedene Typen:

- den Zwei-Fächer- Typ, in dem das 1. Hauptfach und das 2. Hauptfach jeweils ca. 60 SWS umfassen,
- den Drei-Fächer- Typ, in dem das Hauptfach ca. 60 SWS umfaßt und jedes der beiden Nebenfächer ca. 30 SWS beansprucht.

Unter den folgenden Voraussetzungen ist ein Magisterstudium CL einem Diplomstudium CL inhaltlich vergleichbar:

Für den Zwei-Fächer-Typ werden die dem jeweiligen "Vorfeld" zuschreibbaren Veranstaltungen vom 2. Hauptfach übernommen. Als 2. Hauptfächer kommen in Frage Linguistik, Informatik und sprachwissenschaftlich ausgerichtete Einzelphilologien. Wie die Verteilung im einzelnen aussieht, wird vom jeweiligen Standort abhängen, und sie sollte ihren Niederschlag finden in einer von den beteiligten Fächern zu erarbeitenden Studien- und Prüfungsordnung.

Für den Drei-Fächer-Typ in seiner Ausprägung CL und zwei der Nebenfächer Informatik, Linguistik, sowie sprachwissenschaftlich ausgerichtete Einzelphilologie werden Teile der dem jeweiligen "Vorfeld" zuschreibbaren Veranstaltungen vom 1. bzw. 2. Nebenfach übernommen. Um aber eine vergleichbare Qualifikation zumindest prinzipiell sicherzustellen, sind an die Inhalte, die in diesen Nebenfächern vermittelt werden, bestimmte Forderungen von seiten des Hauptfaches zu stellen. Diese Inhalte sind in eine von den beteiligten Fächern zu erstellende Studien- und Prüfungsordnung einzuarbeiten.

In jedem Falle wird das Hauptfach CL die Aufgabe zu erfüllen haben, die Inhalte des/der jeweils anderen Faches/Fächer auf die Veranstaltungsinhalte der CL zu beziehen.

Weitere sinnvolle Kombinationen mit CL als Hauptfach sind denkbar, sind aber mit dem im Kapitel 3 skizzierten Diplomstudiengang nicht mehr vergleichbar. Dazu würden die Kombinationen von CL mit Psychologie, Informationswissenschaft, Formale Logik, u.a. zählen. Kombinationen mit Fächern ohne jeglichen inhaltlichen Bezug zur CL werden nicht empfohlen, können aber für Studierende im Einzelfall sinnvoll sein.

## **5 Computerlinguistik als Nebenfach**

Computerlinguistik als Nebenfach sollte im Rahmen eines Diplomstudienganges oder im Rahmen eines Magisterstudienganges vom Drei- Fächer- Typ nur an solchen Universitäten angeboten werden, an denen CL auch als Hauptfach studiert werden kann oder aber an denen eine CL vorhanden ist, die personell, apparativ und mit Sachmitteln so ausgestattet ist, daß eine Forschung durchgeführt wird, die eine im Niveau akzeptable Lehre gewährleistet.

Welche Inhalte in welcher Form aus welchen Veranstaltungen des Rahmens für einen Diplomstudiengang (vgl. Kap. 3) gewählt werden, hängt einerseits ab von den örtlichen Gegebenheiten (vor allem von den Forschungsschwerpunkten) und andererseits von dem gewählten Hauptfach.

CL ist als Nebenfach zu einem Hauptfach Linguistik bzw. sprachwissenschaftlich ausgerichtete Einzelphilologie anders zu strukturieren als als Nebenfach zu einem Hauptfach Informatik.

Auf Grund des interdisziplinären Charakters des Faches CL erscheint es - von Einzelfällen abgesehen - wenig sinnvoll, CL mit Fächern zu kombinieren, die keinen Bezug zu CL haben.

## **6 Computerlinguistik als Teil in einem anderen Fach**

Wenn Computerlinguistik nur als Teil (Studienrichtung oder Teilfach) eines Nebenfaches in einem Magisterstudiengang vom Drei-Fächer- Typ auftritt, dann ist zu bezweifeln, ob diese Kombination sinnvoll ist. Sie hätte nämlich bei den heutigen Gegebenheiten zur Konsequenz, daß CL mit nur ca. 15 SWS vertreten wäre. Es ist fraglich, was ein derart reduziertes Lehrangebot enthalten soll und welche Qualifikation damit vermittelt werden soll.

Dieser Zweifel gilt auch für den Fall, daß CL Teil eines Nebenfaches für einen Diplomstudiengang Informatik ist (zumal die Anzahl der SWS für Nebenfächer in einem Diplomstudiengang Informatik zwischen ca. 18 und 36 SWS differiert).

Tritt CL als Teil eines Hauptfaches auf, gilt das im ersten Abschnitt von Kap. 5 Gesagte.

## 7 Weiterbildungsstudiengänge in CL

Unter Weiterbildungsstudiengängen werden hier Studiengänge verstanden, die in den einzelnen Bundesländern als 'Aufbaustudiengang', 'Ergänzungsstudiengang' und 'Zusatzstudiengang' bezeichnet werden.

Unberücksichtigt bleiben in diesem Zusammenhang Zweitstudium und Weiterbildung im Rahmen von Einzelveranstaltungen (Kontaktstudium).

Studienvoraussetzung für einen Weiterbildungsstudiengang CL ist ein berufsqualifizierender Abschluß eines Hochschulstudiums (Mag., Dipl., Staatsex.); hierbei muß es sich um Fächer handeln, die in einem engeren inhaltlichen Zusammenhang zu CL stehen (wie z.B. Linguistik, Informatik, Einzelphilologien, Phonetik), oder um CL selbst.

Die Studiendauer beträgt in der Regel 4 Semester.

Der Studiengang wird durch eine Prüfungsordnung **und** eine Studienordnung bestimmt. Der Weiterbildungsstudiengang CL sollte nur an einer Hochschule angeboten werden, an der ein grundständiger CL-Studiengang existiert. Das Lehrangebot des Weiterbildungsstudiengangs ist institutionell und personell an das grundständige Lehrangebot zu binden.

Die Lehrveranstaltungen müssen sich inhaltlich nach den als Eingangsqualifikation geforderten Studienvoraussetzungen richten; ein Weiterbildungsstudiengang CL muß also bei vorangegangenem Studium der Informatik, Linguistik, Einzelphilologie, Phonetik oder CL jeweils verschiedene Lehrinhalte anbieten.

Eine Abstimmung zwischen den einzelnen Hochschulen bei der Einrichtung von Weiterbildungsstudiengängen CL ist unumgänglich.

## Anmerkungen

- (1) Mitgearbeitet und zur Entstehung dieses Papiers beigetragen haben:  
I.S. Batori, R. Drewek, B. Endres-Niggemeyer, A. Franzke, T. Gardner, U. Hahn, P. Hellwig, M. Heyn, L. Hitzberger, K. Jakob, B. Kelle, G. Knorz, W. Kreitmeier, W. Löser, H.-D. Lutz, M. Lutz-Hensel, P. Ovenhausen, H.P. Pütz, W. Putschke, U. Reyle, B. Rieger, B. Schaefer, K.-D. Schmitz, H. Schnelle, G. Schweisthal, G. Willee, K. Wothke
- (2) Die Bezeichnung Computerlinguistik (CL) wird hier stellvertretend für die Vielzahl der zur Zeit gebräuchlichen Bezeichnungen des Faches verwendet.
- (3) s. Studienführer Linguistische Datenverarbeitung (LDV) für die wissenschaftlichen Hochschulen der Bundesrepublik Deutschland, ermittelt und bearbeitet von Magdalene Lutz-Hensel, Stand Januar 1985.
- (4) Diese Zielsetzung ist zur Zeit nicht einheitlich. V gl. **LDV** Forum 5 (1987), 76-78 mit einer Vorversion dieses Papiers, oder Batori, I.; Krause, J.; Lutz, H.-D. (Hg.), Linguistische Datenverarbeitung. Versuch einer Standortbestimmung im Umfeld von Informationslinguistik und künstlicher Intelligenz. (= Sprache und Information. 4.) Tübingen: Niemeyer 1982.
- (5) Dem liegt ein nicht allein syntaktisch geprägtes Verständnis von Parsing und Grammatik zugrunde.
- (6) Die Bezeichnung "Computerlinguist" steht stellvertretend für "Computerlinguist" und "Computerlinguistin".
- (7) Bedingt durch die Inhalte des Studiengangs wird von einer Studienzeit von 10 Semestern ausgegangen. Die formalen Bedingungen für die Studiendauer mögen aufgrund von Länderverordnungen o.ä. anders aussehen, d.h. ein achtsemestriges Studium fordern.