

Wegleitung für das Lizentiats-Studium der Computerlinguistik an der Universität Zürich¹

Inhaltsverzeichnis

ADRESSEN UND PERSONEN	1
WAS IST COMPUTERLINGUISTIK?	2
BERUFSAUSSICHTEN FÜR STUDIERENDE DER COMPUTERLINGUISTIK	4
WER SOLLTE COMPUTERLINGUISTIK STUDIEREN?	4
STUDIENMÖGLICHKEITEN	5
ZEITLICHE UND THEMATISCHE EINPASSUNG DER COMPUTERLINGUISTIK IN DAS STUDIUM	6
EMPFOHLENE STUDIENPLÄNE	7
STUDIENORDNUNG: LEHRE	8
STUDIENORDNUNG: PRÜFUNGEN	9
SCHWERPUNKTE DER FORSCHUNG AN DER UNIVERSITÄT ZÜRICH	13
DAS WEB-ANGEBOT DER COMPUTERLINGUISTIK	14
COMPUTERLINGUISTIK-LITERATUR	16

¹ Diese Studienordnung ist gültig für Studierende, die ab dem Wintersemester 2004/2005 Computerlinguistik studieren oder als Wahlgebiet belegen. Für Studierende, die ihr Studium früher begonnen haben, gilt die Studienordnung vom 10.6.2003 bzw. 5.3.2001. Studierende, die sich nach dem **Sommersemester 2006 immatrikulieren**, können nur den neuen Bachelor- bzw. Masterstudiengang belegen. Für diese Studiengänge gibt es eigene Wegleitungen auf unserer Homepage.

Adressen und Personen

Die Computerlinguistik an der Universität Zürich ist ein Fach der Philosophischen Fakultät. Sie ist aber technisch und administrativ dem Institut für Informatik angegliedert und auch dort angesiedelt (Universität Zürich-Irchel, Bau 27).

Postalische Adresse

Institut für Computerlinguistik
Binzmühlestrasse 14
CH-8050 Zürich (Oerlikon)

Web-Adresse

<http://www.cl.unizh.ch/>

Sekretariat

Corinne Maurer
Büro 2.C.07
Tel. 044 635 43 31
E-Mail: maurer@ifi.unizh.ch

Fachbibliothek

(eingegliedert in die Bibliothek des Instituts für Informatik)
Raum 0.C.09
Tel. 044 635 4357

Dozenten

Prof. Dr. Michael Hess
Büro 2.B.02
Tel. 044 635 43 77
E-Mail: hess@cl.unizh.ch

lic. phil. Alexandra Bünzli
Büro 2.B.04
Tel. 044 635 67 20
E-Mail: buenzli@cl.unizh.ch

Dr. Simon Clematide
Büro 2.B.03
Tel. 044 635 7121
E-Mail: siclemat@cl.unizh.ch

Dr. Manfred Klenner
Büro 2.B.03
Tel. 044 635 43 25
E-Mail: klenner@cl.unizh.ch
(zuständig für Studienfachberatung)

Was ist Computerlinguistik?

Die Computerlinguistik ist eine junge Disziplin im Überschneidungsbereich von Sprachforschung und Informatik. Sie untersucht,

1. **wie die menschliche Sprache als Mittel zur Übermittlung, Speicherung und Verarbeitung von Information verwendet wird, und**
2. **wie man diese Prozesse auf dem Computer modellieren kann.**

Dies geschieht einerseits aus theoretischem Interesse und andererseits unter praktischen Gesichtspunkten.

Praktische Anwendungen der Computerlinguistik

Es gibt eine wachsende Zahl von praktischen Anwendungen der Computerlinguistik; die Computerlinguistik wird sich in Zukunft zu einem wirtschaftlich wichtigen Zweig der Softwarebranche entwickeln. Zu diesen Anwendungen gehören Systeme wie

- **Maschinelle Übersetzungssysteme:** Eine ganze Reihe von Systemen für die Maschinelle Übersetzung zwischen verschiedenen natürlichen Sprachen sind heute schon kommerziell verfügbar. Die Qualität der erzeugten Übersetzungen ist allerdings noch beschränkt, und die Maschinelle

Übersetzung ist weiterhin eines der wichtigsten Forschungsgebiete innerhalb der Computerlinguistik.

- **Natürlichsprachliche Zugangssysteme zu technischen Systemen:** Dies sind Programme, welche den Zugriff auf Datenbanksysteme, Expertensysteme u.a. in natürlicher Sprache statt in schwer erlernbaren formalen Abfragesprachen erlauben.
- **Systeme zum Verstehen und Generieren gesprochener Sprache:** Die Bedienung von Geräten etc. durch gesprochene Sprache; Diktiermaschine; Vorlesesysteme; telefonische Dialog- und Auskunftssysteme.
- **Textzugriffs- und Textverwaltungssysteme:** Dokumentenretrieval-Systeme (z.B. für oft sehr umfangreiche technische Dokumentationen, Gesetzestexte); Textfiltersysteme (Sortieren von Meldungen elektronischer Nachrichtendienste, massgeschneiderte elektronische Zeitungen); Textverstehenssysteme (teilweises Erfassen des Inhalts der Texte und Beantworten von Fragen dazu); automatische Abstractingsysteme (Erstellen von Zusammenfassungen von Texten im Hinblick auf ein allgemeines Profil oder eine spezifische Frage).

- **Systeme zum Generieren geschriebener Texte:** Erzeugen von technischer Dokumentation aus der formalen Beschreibung der Geräte; Erzeugen von Börsenberichten, Wetterberichten etc. aus den Rohdaten.
 - **Intelligente Textkritiksysteme:** Grammatik- und Stilkorrekturprogramme.
 - **Intelligente computergestützte Sprachlehrsysteme:** Die heutigen Systeme vergleichen oft einfach die Eingaben des Lernenden mit fest eingegebenen Mustern von Wortfolgen und bewerten sie danach als „korrekt“ oder „falsch“. Besser ist es, die grammatikalische Struktur eines vom Lernenden eingegebenen Satzes zu errechnen, seine Korrektheit zu ermitteln und eventuelle Fehler möglichst genau zu lokalisieren.
- auf Korrektheit und Vollständigkeit hin testen lassen.
- Der Computer bietet Möglichkeiten, Untersuchungen am sprachlichen Material vorzunehmen, welche sonst wegen ihres Aufwandes ausgeschlossen wären, und daraus Erkenntnisse für die Linguistik abzuleiten (z.B. statistische Analyse von Textsammlungen, Sammlung von wörtlich übereinstimmenden Textstellen (Konkordanzen), Lokalisieren spezifischer syntaktischer Konstruktionen in Textsammlungen).
 - Die in der Informatik allgemein und der Computerlinguistik speziell entwickelten Analysekonzepte erweisen sich oftmals als sehr interessant für die theoretische Linguistik (z.B. das Konzept der sog. Unifikation oder die klare Trennung zwischen deklarativem Wissen und prozeduraler Verarbeitung).

Beziehung der Computer-linguistik zur theoretischen Linguistik

Die Bedeutung der Computerlinguistik für die theoretische Linguistik liegt v.a. in den folgenden Punkten:

- Die Theorien der modernen Linguistik sind mittlerweile derart komplex geworden, dass sie sich zunehmend nur noch durch die Implementation auf dem Computer

Berufsaussichten für Studierende der Computerlinguistik

Für Computerlinguisten bieten sich vielfältige Einsatzfelder in allen oben angesprochenen Bereichen der Sprachtechnologie. Während bis in die 80er Jahre Computerlinguisten vor allem in der Forschung arbeiteten, finden sich in den letzten Jahren vermehrt Stellen bei Übersetzungsabteilungen grosser Unternehmen, Softwarehäusern (insbesondere im Bereich der Übersetzerhilfsmittel und Information Retrieval), Verlagen und Bibliotheken.

Die Dozenten der Computerlinguistik Zürich stehen in Kontakt mit Sprachtechnologie-Firmen, die z.T.

auch Praktika anbieten. Ein solches Praktikum stellt neben einem Auslandssemester eine wünschenswerte Ergänzung des Computerlinguistik-Studiums dar, ist aber keine studienrelevante Bedingung.

Studierende, die nicht beabsichtigen, in einem unmittelbar computerlinguistischen Bereich zu arbeiten, erhalten durch das Studium der Computerlinguistik eine wichtige Zusatzqualifikation in Richtung Sprachtechnologie, Neue Medien und Informationstechnologie.

Wer sollte Computerlinguistik studieren?

Computerlinguistik wird oft von Studierenden gewählt, die eine der Philologien (Germanistik, Anglistik, Romanistik o.a.) als Hauptfach haben. Wir empfehlen dann als weiteres Nebenfach Informatik, Allgemeine Sprachwissenschaft oder eine weitere Sprache (mit linguistischem Schwerpunkt). Überschneidungsbereiche gibt es aber auch mit der Pädagogik (Sprachlernsoftware), Psychologie (menschliche Sprachverarbeitungsprozesse) und mit der Philosophie (formale Logik, Epistemologie).

Von besonderem Interesse dürfte die Computerlinguistik daher sein

- für Studierende der an der Universität Zürich vertretenen (modernen und möglicherweise auch alten) Einzelsprachen sowie der Allgemeinen Sprachwissenschaft, gegebenenfalls auch der Philosophie, Psychologie und Pädagogik.
- für Studierende der Informatik und der Mathematik, an Universität und ETH.

Studienmöglichkeiten

Es gibt vier Möglichkeiten, Computerlinguistik an der Universität Zürich zu studieren.

1. Als **1. oder 2. Nebenfach in der Philosophischen Fakultät**. Die Lizentiatsarbeit kann auf dem Gebiet des ersten Nebenfachs geschrieben werden, sofern die Fakultät dem zustimmt (§10 Reglement über die Lizentiatsprüfung). Für die Computerlinguistik besteht keine Lateinplicht.
2. Als **Wahlgebiet innerhalb der Informatik für Studierende der Philosophischen Fakultät**, falls Informatik das 1. oder 2. Nebenfach ist. Informatik als Nebenfach bedarf, wie die meisten Fächer ausserhalb der Philosophischen Fakultät, der Genehmigung durch die Philosophische Fakultät (§8 Reglement über die Lizentiatsprüfung). Bei dieser Kombination kann Computerlinguistik nicht zusätzlich als weiteres Nebenfach gewählt werden.
3. Als **Wahlgebiet innerhalb der Informatik für Studierende mit Hauptfach Informatik**.
4. Als **Nebenfach für Studierende an anderen Fakultäten und Hochschulen** (z.B. als grosses oder kleines Nebenfach in der Mathematisch-naturwissenschaftlichen Fakultät). Hier gelten die Bestimmungen der entsprechenden Fakultäten und Hochschulen über zugelassene Nebenfächer.

Ungewohnte Zerlegungen deutscher Komposita

(als Nebenprodukt eines Systems zur automatischen Segmentierung deutscher Wörter)

Abtei#lungen	Red#aktion
Amerika#nische	Vorsitz#ende
Konsum#enten	Kau#fladen
Hau#sauf#gaben	Zug#riff
Wand#elan#leihe	Niko#laus
Zwei#fels#fall	Nacht#eile
Zwang#sauf#kauf	Dieb#stahl
Winterthur#erst#rasse	

Zeitliche und thematische Einpassung der Computerlinguistik in das Studium

Wie in allen Fächern der Philosophischen Fakultät gibt es auch in der Computerlinguistik keinen verpflichtend vorgeschriebenen Normallehrgang.

Im Interesse einer systematischen und effizienten Ausbildung wird aber ein strukturierter Studiengang mit weitgehend aufeinander abgestimmten Lehrveranstaltungen angeboten, den man normalerweise in 6 Semestern absolvieren kann.

Einen Überblick über diesen Studiengang und die empfohlene Abfolge der Veranstaltungen gibt der tabellarische Studienplan. Man beachte, dass das Curriculum je nach den gemachten Erfahrungen in der Zukunft modifiziert werden kann.

Es empfiehlt sich (v.a. für Studierende mit Computerlinguistik als 1. Nebenfach), mit dem Besuch der Computerlinguistik-Veranstaltungen frühzeitig im Studium zu beginnen. Man kann den Studiengang jedes Wintersemester beginnen.

Das Studienangebot umfasst pro Semester normalerweise 12 Stunden, von denen aber beim Beachten der empfohlenen Abfolge von Veranstaltungen i.d.R. nicht mehr als 3

Vorlesungen/Kurse im gleichen Semester zu besuchen sind.

Im 1. Jahr lernen die Studierenden im Kurs „Programmiertechniken in der CL“ die logische Programmiersprache PROLOG und wenden sie an, um verschiedene Probleme der Computerlinguistik zu lösen.

Da sich Programmieren nur lernen lässt, wenn man auch selbst programmiert, werden regelmässig Übungen angeboten, die selbstständig oder im Rahmen einer betreuten Übungsstunde (2h pro Woche) von den Studierenden gelöst werden.

Computerlinguistik ist inhaltlich eng verzahnt mit der Informatik und der Linguistik. An der Universität Zürich werden folgende Verbindungen besonders betont:

1. bezüglich der Informatik: die Logik-Programmierung und die Künstliche Intelligenz (KI)
2. bezüglich der Sprachfächer in der Philosophischen Fakultät: die linguistischen Aspekte der einzelnen vor allem europäischen Sprachen sowie Vergleiche zwischen verschiedenen Sprachen

Empfohlene Studienpläne

Computerlinguistik als 1. Nebenfach (11 Vorlesungen/Kurse + 2 Seminare)

1. Sem.	Einführung in die CL I	Programmiertechniken in der CL I + Übungen		
	<i>Teilakzessprüfung</i>	<i>Teilakzessprüfung</i>		
2. Sem.	Einführung in die CL II	Programmiertechniken in der CL II + Übungen	Lexikonaufbau und Morphologieanalyse	
	<i>Teilakzessprüfung</i>	<i>Teilakzessprüfung</i>		
3. Sem	Methoden der KI in der CL*		Formale Grammatiken und Syntaxanalyse	Quantitative Methoden der CL*
bis	Semantikanalyseverfahren*	Spezialvorlesung*	Corpuslinguistik	Seminar
6. Sem.	Diskursanalyse*	Spezialvorlesung*		Seminar
	<i>Lizentiatsprüfung</i>			

Computerlinguistik als 2. Nebenfach (8 Vorlesungen/Kurse + 1,5 Seminare)

1. Sem.	Einführung in die CL I	Programmiertechniken in der CL I + Übungen		
	<i>Teilakzessprüfung</i>	<i>Teilakzessprüfung</i>		
2. Sem.	Einführung in die CL II	Programmiertechniken in der CL II + Übungen	Lexikonaufbau und Morphologieanalyse	
	<i>Teilakzessprüfung</i>	<i>Teilakzessprüfung</i>		
3. Sem.	Methoden der KI in der CL [§]		Formale Grammatiken und Syntaxanalyse	Quantitative Methoden der CL [§]
bis	Semantikanalyseverfahren [§]	Spezialvorlesung [§]	Corpuslinguistik	Seminar
6. Sem.	Diskursanalyse [§]			kl. Seminar
	<i>Lizentiatsprüfung</i>			

* Es müssen vier dieser sechs Veranstaltungen ausgewählt werden.

§ Es muss eine dieser fünf Veranstaltungen ausgewählt werden.

Studienordnung: Lehre

1. Hauptvorlesungen

(regelmässig und in fester Reihenfolge angeboten)

Einführende Vorlesungen

- Einführung in die Computerlinguistik I
- Einführung in die Computerlinguistik II
- Programmiertechniken in der Computerlinguistik I
- Programmiertechniken in der Computerlinguistik II

Die Programmiertechniken I und II werden durch eine wöchentliche Übungsstunde (2h), d.h. mit praktischer Arbeit am Computer, ergänzt.

Grundlegende Vorlesungen

- Corpuslinguistik
- Lexikonaufbau und Morphologieanalyseverfahren
- Formale Grammatiken und Syntaxanalyse

Weiterführende Vorlesungen

- Methoden der Künstlichen Intelligenz in der Computerlinguistik
- Diskursanalyse
- Semantikanalyseverfahren
- Quantitative Methoden in der Computerlinguistik

2. Spezialvorlesungen

(zu wechselnden Themen; nicht in regelmässigem Turnus angeboten; oft von auswärtigen Dozenten gelehrt)

- Ausgewählte Fragen der Computerlinguistik (wechselnde Schwerpunkte)
- Syntaxtheorien für die Sprachverarbeitung
- Textverstehen
- Intelligente Verfahren der Informationssuche
- Maschinenunterstützte Übersetzung
- Computer in der Sprachwissenschaft

Nach Rücksprache mit den Dozenten der Computerlinguistik können auch Veranstaltungen der Informatik als Spezialvorlesung anerkannt werden. So z.B.

- Logik-Programmierung
- Mensch-Rechner-Interaktion
- Einführung in die Künstliche Intelligenz

Externe Kurse

Kurse/Veranstaltungen, die an Sommer-/Herbstschulen erfolgreich besucht wurden (Bescheinigung), werden ebenfalls als Spezialvorlesung angerechnet.

3. Seminare

Zum erfolgreichen Besuch eines Seminars gehört das Halten eines Referats, die Annahme einer schriftlichen Seminararbeit und die Teilnahme am Seminar an mindestens 80% der Sitzungen – ausgenommen vorgängig bewilligter Urlaub oder Abwesenheiten infolge von Krankheit.

Kleines Seminar

Beim sogenannten ‚kleinen Seminar‘ entfällt die schriftliche Seminararbeit.

Programmierprojekt

Mit Bewilligung des Fachvertreters Computerlinguistik kann eine der verlangten Seminararbeiten ersetzt

werden durch eine Semesterarbeit, die aus der Entwicklung einer substantiellen Implementation besteht. Voraussetzung zur Annahme sind die Abgabe des kommentierten Quellcodes in elektronischer und gedruckter Form, ein kurzer technischer Bericht, das Halten einer mündlichen Präsentation vor Publikum und die Übertragung des Copyrights am Programmcode an die Universität Zürich.

Seminarbescheinigung

Der erfolgreiche Besuch von Seminaren resp. die Annahme einer Semesterarbeit wird auf der Studienkarte schriftlich bestätigt (Seminarbescheinigung). Bestätigungen müssen bei der Anmeldung zur Lizentiatsprüfung vorliegen.

Studienordnung: Prüfungen

Akzessprüfung

Die Akzessprüfung setzt sich aus 4 schriftlichen Teilakzessprüfungen zusammen. Geprüft wird der Stoff aus den Vorlesungen: Einführung in die Computerlinguistik I und II, Programmiertechniken in der Computerlinguistik I und II.

Die Prüfungen finden immer zur gleichen Uhrzeit wie die entsprechende Vorlesung statt und zwar entweder in der letzten Semesterwoche oder in den beiden ersten Wochen der vorlesungsfreien Zeit.

Die Akzessprüfung gilt als bestanden, wenn alle vier Teilbereiche erfolgreich absolviert worden sind. Die Teilbereiche können zu beliebigen Zeitpunkten abgelegt werden.

Die Akzessprüfung ist obligatorisch für alle Studierenden der Computerlinguistik ausser für jene Studierende, die CL als Wahlgebiet im 2. Nebenfach Informatik wählen. Das Bestehen der Akzessprüfung ist Bedingung für den Besuch von Seminaren.

Lizentiatsprüfung

Voraussetzungen für die Anmeldung zur Lizentiatsprüfung

1. Bescheinigung über die bestandene Akzessprüfung in Computerlinguistik (ausser für Studierende, die CL als Wahlgebiet im 2. Nebenfach Informatik wählen).
2. Bescheinigungen über Seminarbesuche in Computerlinguistik (siehe Tabelle Seminarbesuche).
3. Bescheinigung über angenommene Lizentiatsarbeit

Tabelle: Seminarbesuche

<i>CL als 1. Nebenfach</i>	Bescheinigung über 2 Seminarbesuche in Computerlinguistik (je mit Seminararbeit)
<i>CL als 2. Nebenfach</i>	Bescheinigung über 2 Seminarbesuche in Computerlinguistik (ein Seminar mit Seminararbeit und eines nur mit Referat)
<i>CL als Wahlgebiet im 1. Nebenfach Informatik</i>	Bescheinigung über einen Seminarbesuch in Computerlinguistik (mit Seminararbeit)
<i>CL als Wahlgebiet im 2. Nebenfach Informatik</i>	keine Seminarbesuche in Computerlinguistik verlangt
<i>CL als Nebenfach für andere Studierende</i>	nach den Regeln der jeweiligen Institution

Art und Umfang der Lizentiatsprüfung in Computerlinguistik

<i>CL als 1. Nebenfach</i>	dreitägige Hausarbeit und mündliche Prüfung von 30 Minuten (Falls die Lizentiatsarbeit im 1. Nebenfach gemacht wurde, muss die Hausarbeit im Hauptfach gemacht werden.)
<i>CL als 2. Nebenfach</i>	mündliche Prüfung von 30 Minuten
<i>CL als Wahlgebiet im 1. oder 2. Nebenfach Informatik</i>	mündliche Prüfung von 30 Minuten (wovon ca. 50% zur Computerlinguistik gehören, der Rest zur Informatik allgemein)
<i>CL als Nebenfach für andere Studierende</i>	nach Absprache; nach Möglichkeit in Anlehnung an die Regeln der jeweiligen Fakultät

Prüfungsstoff für die mündliche Lizentiatsprüfung

	<i>CL als 1. NF</i>	<i>CL als 2. NF</i>	<i>CL im 1. NF Infor- matik</i>	<i>CL im 2. NF Informatik</i>
Einführung in die CL I	ja	ja	ja	ja
Einführung in die CL II	ja	ja	ja	ja
Programmiertechniken in der CL I	ja	ja	ja	ja
Programmiertechniken in der CL II	ja	ja	ja	
Lexikonaufbau und Morphologieanalyse	ja	ja	1 Veranstaltung aus diesen 5	
Formale Grammatiken und Syntaxanalyse	ja	ja		
Corpuslinguistik	ja	ja		
Semantikanalyseverfahren	4 Veranstaltungen aus	1 Veranstaltung aus		
Quantitative Methoden der CL				
Methoden der KI in der CL	diesen 6	diesen 5		
Diskursanalyse				
Spezialvorlesung				
Spezialvorlesung				

- Studierende mit Computerlinguistik als Wahlgebiet innerhalb der Informatik (CL innerhalb des Nebenfachs Informatik) beachten bitte auch das spezielle Merkblatt der Abteilung für Informatik.
- Für Studierende an anderen Fakultäten/Institutionen als der Philosophischen Fakultät wird die Lizentiatsprüfung nach Absprache und in Anlehnung an die Regeln der jeweiligen Fakultät/Institution durchgeführt.

Schwerpunkte der Forschung an der Universität Zürich

Die Computerlinguistik-Gruppe der Universität Zürich beschäftigt sich vor allem mit Informationsgewinnung aus Texten.

Das Ziel dieser Forschung ist die Entwicklung von Informationssystemen, die natürlichsprachliche Fragen auf Grund von natürlichsprachlichen Texten beantworten können.

- Antwortextraktionssysteme lokalisieren anhand der Frage relevante Textstellen.
- Als Fernziel gelten Fragebeantwortungs-Systeme, die Antworten generieren durch Schlussfolgern über den in den Texten enthaltenen Informationen.

Zur Zeit wird einerseits an Systemen gearbeitet, die englischsprachige Texte wie das Online-Handbuch des Unix-Betriebssystems oder das „Aircraft Maintenance Manual“ des Airbus 320 verarbeiten können und darin Antworten auf Benutzerfragen finden.

Alle Sätze dieser Texte werden morphologisch, syntaktisch und semantisch analysiert und schliesslich als logische Formeln repräsentiert. Der so verarbeitete Text bildet die Wissensbasis des Systems. Wenn eine Frage gestellt wird, so wird diese auf

die gleiche Weise analysiert und ebenfalls in eine logische Formel übersetzt. Durch logische Schlüsse über Wissensbasis und Frage kann die Antwort im Ausgangstext lokalisiert und angezeigt werden.

Andererseits arbeiten wir an einem deutschsprachigen Universitäts-Informationssystem, das natürlichsprachliche Fragen zu den administrativen Vorgängen an der Universität Zürich beantwortet. Das sind Fragen, die typischerweise an die Kanzlei gestellt werden, wie *Wo bekommt man das Vorlesungsverzeichnis?* Als Antwort soll jeweils ein möglichst angemessener, vollständiger Satz geliefert werden.

Auch um eine solche Frage beantworten zu können, benötigt das System die entsprechenden Informationen in Form von natürlichsprachlichen Texten. Ein Satz des Textes könnte z.B. lauten: *Das Vorlesungsverzeichnis kann in der Universitätskanzlei, im Studentenladen oder in verschiedenen Buchhandlungen gekauft werden.*

Zusätzlich beschäftigt sich die Gruppe mit dem Entwickeln von Grammatiken (*Grammar Engineering*), insbesondere für die deutsche Sprache.

Das Web-Angebot der Computerlinguistik

Die Computerlinguistik der Universität Zürich hat früh begonnen, Informationen zu Lehre und Forschung über das Internet verfügbar zu machen. Die Homepage der Computerlinguistik liegt unter:

<http://www.cl.unizh.ch/>

Dort finden sich unter anderem:

- Stundenpläne mit den Veranstaltungen der Computerlinguistik
- Lehrveranstaltungs-Unterlagen zu den meisten Computerlinguistik-Vorlesungen (z.T. mit interaktivem Übungsmaterial)
- eine Liste mit häufig gestellten Fragen (FAQ) zur Computerlinguistik
- computerlinguistische Beispielprogramme, die man über das Internet ausprobieren kann (z.B. ein sog. Tagger, der zu den Wörtern

eines eingegebenen deutschen Satzes die Wortarten ermittelt)

- ein aktueller Überblick über die Forschungsaktivitäten inklusive Zugriff auf die neuesten wissenschaftlichen Veröffentlichungen der Zürcher Computerlinguisten
- eine Liste aller Computerlinguistik-relevanten Zeitschriften in Zürcher Bibliotheken

Mailing-Liste

Für die Ankündigung von speziellen Veranstaltungen, Jobangeboten usw. unterhalten wir eine Mailing-Liste:

cl-list@ifi.unizh.ch

Für Anmeldung bitte eine Mail ohne weiteren Inhalt an:

cl-list-subscribe@ifi.unizh.ch

für Abmeldung an:

cl-list-unsubscribe@ifi.unizh.ch

Eigenwillige maschinelle Übersetzungen

Fehler durch falsche Wörterbucheinstellung:

„Abbau“	<i>shut down</i> statt: <i>reduction</i>
„Band“	<i>volume</i> statt: <i>tape</i>
„Die Leitung wurde verlegt.“	<i>The management was killed.</i>
„einführen“	<i>import</i> statt: <i>introduce</i>

Fehler bei der Übersetzung von Komposita:

„Armbanduhr“	<i>poor tape clock - bracelet clock</i>
„Barzahlung“	<i>nightclub payment</i>
„Vorschlaghammer“	<i>proposal hammer</i>

Die Fachschaft CLinZ/CH

Die Fachschaft CLinZ/CH ist der studentische Fachverein der Computerlinguistik an der Universität Zürich. CLinZ/CH steht für CL in Zürich / Schweiz (CH). Der Fachverein wurde 1996 gegründet und versteht sich als Bindeglied zwischen Lehrkörper und Studierenden. Zudem will er den Kontakt unter den Studierenden fördern. Um diese Ziele umzusetzen, führt CLinZ/CH vielfältige Aktivitäten durch.

- Unterstützung der CL-Studierenden in Zürich durch Archivierung von Prüfungsfragen und finanzielle Beiträge für Teilnahme an CL-Tagungen, CL-Sommerschulen etc.
- Herausgabe der semesterlich erscheinenden Broschüre CLIP (ComputerLinguistikInformationsPlattform) mit Interviews, Fachbeiträgen oder auch weniger ernst Gemeintem
- Sporadische Orientierung der CL-Studierenden via E-Mail: z.B. Veranstaltungsempfehlungen
- Jährliche Teilnahme an der TaCoS (Tagung der Computerlinguistik-Studierenden im deutschsprachigen Raum). Die TaCoS '98 wurde in Zürich organisiert.
- Studentische Workshops, z.B. zur theoretischen Linguistik, Computerlinguistik
- Organisation eines jährlich stattfindenden Schlitteltags und eines Sommerfests
- Teilnahme einer CL-Mannschaft an der SOLA-Stafette
- Design und Vertrieb eines eigenen T-Shirts

Alle Studierenden sind herzlich eingeladen dem Fachverein beizutreten. Wie das geht, erfährt man unter:

<http://www.cl.unizh.ch/CLinZ.CH>

Computerlinguistik-Literatur

Wer Computerlinguistik studiert, macht sich mit Vorteil schon von Anfang an mit englischen Fachbegriffen und -texten vertraut. Je eine deutschsprachige Referenz ist dennoch angegeben...

Einführende Literatur

Carstensen, Kai-Uwe et al. 2004. *Computerlinguistik und Sprachtechnologie - Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag. 2. Auflage.

Guter und aktueller Überblick! Zur Anschaffung empfohlen!

McTear, Michael. 1987. *The articulate computer*. Oxford: Blackwell.
Vergriffen. Verständlich und eingängig geschriebenes Buch, das der Frage nachgeht, wie Computer in natürlicher Sprache kommunizieren könn(t)en.

Smith, George. 1991. *Computers and Human Language*. New York: Oxford University Press.

Verständlich geschriebenes Lehrbuch, z.T. recht informatisch orientiert.

Cole, Ronald C. et al. 1997. *Survey of the State of the Art in Human Language Technology*: Cambridge University Press. (Elektronische Versionen im WWW verfügbar unter: <http://cslu.cse.ogi.edu/HLTsurvey>)

Sehr breite Thematik wird abgedeckt. Z.T. sehr summarisch und mit wenig Beispielen geschrieben.

Standardliteratur

Jurafsky, Daniel et al. 2000. *Speech and Language Processing*. Prentice-Hall.

Exzellentes und aktuelles Lehr- und Referenzwerk.

Allen, James. 1995. *Natural Language Understanding*. 2nd Edition. Addison Wesley.

Lehrbuch mit ausgebautem Semantik- und Pragmatikteil.

Dale, Robert et al. 2000. *Handbook of natural language processing*. New York: Dekker.

Verständliche und kurze Übersichten

Manning, Christopher D. et al. 2003. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Umfassendes Lehrwerk für statistische Methoden und Anwendungen

Hausser, Roland. 2000. *Grundlagen der Computerlinguistik: Mensch-Maschine-Kommunikation in natürlicher Sprache*. Berlin: Springer.

Gutes Lehrbuch. Gewisse Ansätze etwas übervertreten