

Beschreibung des interdisziplinären Graduiertenkollegs

**Sprache, Mimik und Gestik
im Kontext technischer Informationssysteme**

**Ludwig–Maximilians–Universität München
Technische Universität München**

Präambel

Bei dem Thema »Sprache, Mimik und Gestik im Kontext technischer Informationssysteme« gehen wir davon aus, daß Sprache in ihrer spontan geäußerten Form in zukünftigen Informationssystemen eine in zunehmendem Maße wichtige Rolle spielt. Somit gibt es Gründe, das vorgeschlagene Thema einerseits interdisziplinär auf die hier beteiligten Fächer auszuweiten, andererseits aber eben auch ausdrücklich einzuschränken auf den besonderen Kontext, wie er mit technischen Informationssystemen vorgegeben ist.

Allein die Tatsache, daß es uns dank der technischen Entwicklung möglich wird, neben den Sprachsignalen jetzt auch die Mimik und Gestik von Sprechern einer digitalen Verarbeitung zugänglich zu machen, bringt Phonetiker und Ingenieure – sowohl der Informatik (Bildverarbeitung) als auch der Nachrichtentechnik (Sprachverarbeitung) – enger zusammen. Hinzu kommt das aktuelle Interesse der germanistischen Sprachwissenschaft, die sehr komplexen, aber digital verfügbaren neuen Daten mit linguistischen Methoden zu analysieren. Die Rolle des Kontextes, in welchem ein Sprecher seine Absichten (mit linguistischen und nicht-linguistischen Mitteln) auszudrücken versucht, bedarf in besonderem Maße einer Klärung, die nur aus der Perspektive einer Sprachphilosophie zu leisten ist, wie sie hier in München von Andreas Kemmerling vertreten wird. Schließlich kann über die Psycholinguistik nicht nur ein Großteil der umfangreichen Literatur über die nonverbalen Komponenten bei der natürlichsprachlichen Kommunikation verfügbar gemacht werden, vielmehr ergeben sich, im Kontext technischer Informationssysteme, völlig neue Aspekte der Effizienz von sprachlicher Kommunikation, und zwar in beiden Richtungen, der Ein- wie der Ausgabe von sprachlich formulierter Information.

Bei der Wahl des Themas und der Auswahl der beteiligten Disziplinen waren zwei weitere Gründe maßgebend, erstens im Hinblick auf die Lehre, zweitens im Blick auf die Forschung:

- Erstens existiert unseres Wissens für die anwendungsseitig sehr wichtigen sprachlichen Aspekte beim Einsatz von technischen Informationssystemen noch kein akademischer Studiengang, nicht einmal in Form eines Aufbaustudienganges. Hier möchten wir mit dem Graduiertenkolleg begabten Studenten die Möglichkeit eröffnen, sich für das neue Gebiet der zukünftigen Informationssysteme eine fundierte Qualifikation zu erwerben. Dies ist natürlich besonders interessant für die geisteswissenschaftlichen Fächer, deren Studierende jetzt verstärkt in berufliche Domänen vordringen, die bisher allein technisch ausgebildeten Akademikern vorbehalten waren.
- Zweitens möchten wir mit diesem Thema dazu beitragen, daß es auch im Bereich der aktuellen Grundlagenforschung zur Entwicklung von zukünftigen Informationstechnologien verstärkt zu einer Kooperation von natur- und geisteswissenschaftlichen Fächern kommen kann. Wir glauben, daß dafür gerade unter der Federführung der Phonetik eine besonders gute Chance besteht. Denn dieses Fach hat mit seinen beiden Komponenten, der Symbolphonetik und ihrer sprachwissenschaftlichen Kategorienbildung und der Signalphonetik mit an der Informationsübertragung durch Zeitfunktionen orientierten Theoriebildung, schon immer

eine Verbindung zwischen der naturwissenschaftlichen und geisteswissenschaftlichen Informations- und Sprachforschung gesucht und mit Erfolg praktiziert.

Für empirisch fundierte Untersuchungen können den Kollegiaten des Graduiertenkollegs schon im Laufe des Jahres 2000 umfangreiche digital aufbereitete multimediale Dialogdaten (MMDD) zur Verfügung gestellt werden. Denn nachdem das Institut für Phonetik und Sprachliche Kommunikation (IPSK), zusammen mit dem Bayerischen Archiv für Sprachsignale (BAS), für das **SmartKom**-Projekt des BMBF – hierbei handelt es sich um eines der neuen Leitprojekte der Bundesregierung zur "Mensch-Technik-Schnittstelle in der Wissensgesellschaft" – den Auftrag erhalten hat, eine intelligente mobile MMDD-Aufnahmestation zu entwickeln, bietet es sich geradezu an, die auf diese Weise verfügbar werdenden Daten im Rahmen des Graduiertenkollegs auch einer rein wissenschaftlichen Analyse zu unterziehen, wobei natürlich die Ergebnisse dann durchaus auch anwendungsseitig, insbesondere für die am SmartKom-Projekt beteiligten Firmen, von großem Interesse sein dürften.

1 Fachvertreter

Ludwig–Maximilians–Universität

Institut für Phonetik und Sprachliche Kommunikation
Schellingstr. 3/II VG
80799 München

Fachgebiet »Phonetik und Sprachliche Kommunikation«

Prof. Dr. H.G. Tillmann (Sprecher)

Tel. 089–2180–2758 (Sekretariat)

Tel. 089–2180–3546 (Büro)

Fax. 089–2800362

e–mail tillmann@phonetik.uni-muenchen.de

Institut für Phonetik und Sprachliche Kommunikation
Fach »Sprechwissenschaft und Psycholinguistik«
Oettingenstr. 67, Trakt F
80538 München

Fachgebiet »Sprechwissenschaft und Psycholinguistik«

Prof. Dr. G. Kegel

Tel. 089–2178–2650 (Büro)

Fax. 089–2178–2652

e–mail kegel@psycholinguistik.uni-muenchen.de

Institut für Deutsche Philologie
Schellingstr. 3 RG
80799 München

Fachgebiet »Germanistische Linguistik«

PD Dr. Th. Becker

Tel. 089–2180–2903

e–mail Thomas.Becker@lrz.uni-muenchen.de

Institut für Deutsch als Fremdsprache
Ludwigstr. 27/I
80539 München

Fachgebiet »Deutsch als Fremdsprache«

Prof. Dr. K. Ehlich

Tel. 089–2180–2116 (Sekretariat)

Fax. 089–2180–3999

e–mail Ehlich@DaF.Uni-Muenchen.De

Institut für Philosophie, Logik und Wissenschaftstheorie
Ludwigstr. 31/I
80539 München

Fachgebiet »Philosophie«

Prof. Dr. A. Kemmerling

Tel. 089–2180–3472

e–mail sak@lrz.uni-muenchen.de

Technische Universität München

Lehrstuhl für Mensch–Maschine–Kommunikation
Arcisstr. 21
80290 München

Fachgebiet »Elektro–und Informationstechnik«
Prof. Dr. M. K. Lang
Tel. 089 – 289 – 28541
Fax. 089 – 289 – 28535
e–mail lg@mmk.e–technik.tu–muenchen.de

PD Dr. G. Ruske
Tel. 089 – 289 – 28563
Fax. 089 – 289 – 28535
e–mail ruske@mmk.e–technik.tu–muenchen.de

Lehrstuhl Informatik IX (Bildverstehen und Wissensbasierte Systeme)
Orleansstr. 34
81667 München

Fachgebiet »Informatik«
Prof. Dr. Bernd Radig
Tel. 089–48095–121
Fax. 089–48095–203
e–mail radig@in.tum.de

Dr. Werner Meixner
Tel. 089–48095–281
Fax. 089–48095–203
e–mail meixner@in.tum.de

Fachgebiete:

- **Phonetik und Sprachliche Kommunikation:**
 - Adaptieren von sprachlichen Informationssystemen
 - Experimentelle Untersuchungen zur Natürlichkeit der Sprachausgabe (Sprach–synthese)
 - Detektion relevanter Informationen in Sprachsignalen (semantischer Fokus)
 - SONDERTHEMA: Die Interpretation von Gebärdensprache in informationstechnischen Systemen
- **Psycholinguistik und Sprechwissenschaft:**
 - Zeit– und Signalverarbeitung in der Mensch–Rechner–Interaktion
 - Einfluß von Orientierungsreizen auf Verstehen und Behalten gesprochener Sprache
 - Emotionale Einflußgrößen bei sprechsprachlicher Information
 - Struktur und Funktion von Sprache bei komplexen Problemen im problemlösenden Dialog

- **Germanistische Linguistik:**
 - Variation in der Standardaussprache
 - in Abhängigkeit von sozialen Faktoren (Bildung, Beruf etc.),
 - regionalen Faktoren (Dialekteinfluß, Mischung regionaler Varietäten, Interferenzen mit nicht-deutscher Muttersprache),
 - biologischen Faktoren (Geschlecht, Alter),
 - situativen Faktoren (Formalitätsgrad, Sprechtempo, Emotionen),
 - auch im Hinblick auf die Erfassung stilistischer Besonderheiten im
 - Umgang mit technischen Informationssystemen und deren mögliche Auswirkungen
 - auf Sprachwandeltendenzen in der Gegenwartssprache.

- **Deutsch als Fremdsprache:**
 - Nonverbale Kommunikation – Kontrastive Aspekte

- **Analytische Philosophie unter besonderer Berücksichtigung der Sprachphilosophie:**
 - Grundlegende Mehrdeutigkeiten des Kontextbegriffs in Semantik und Pragmatik
 - Die Intentionalität von Computern als begriffliches Problem

- **Mensch–Maschine–Kommunikation:**
 - Benutzermodellierung für gesprochene Eingaben auf akustisch–phonetischer Ebene, auf lexikalischer Ebene und auf Intentionsebene unter Berücksichtigung
 - der Anpassung des Benutzers an das System
 - und der Anpassung des Systems an den Benutzer.

- **Bildverstehen und wissensbasierte Systeme:**
 - Automatische Identifikation audiovisueller Sprachäußerungen auf der Grundlage von
 - Gesichtsdynamik
 - Sprechmotorik

2 Ziele, Programm und Struktur des Graduiertenkollegs

2.1 Zusammenfassung

An dem Graduiertenkolleg »Sprache, Mimik und Gestik im Kontext technischer Informationssysteme« wirken zwei Fakultäten der Ludwig–Maximilians–Universität und zwei Fakultäten der Technischen Universität München mit. Das Kolleg soll im Bereich der menschlichen Interaktion mit technischen Informationssystemen fehlende Erkenntnisse im Bereich der Grundlagenforschung gewinnen und innovative Anwendungskonzepte entwickeln. Derartiges kann heutzutage nur noch durch die interdisziplinäre Verknüpfung verschiedener Wissenschaftsdisziplinen ermöglicht werden. Neben wesentlichen Forschungsbeiträgen verspricht das Graduiertenkolleg auch hinsichtlich der Qualifikation der Kollegiaten nachhaltige Effekte. So können z.B. Studierende der Geisteswissenschaften technische Schlüsselqualifikationen erwerben und diese dann in Universitäten und Unternehmen einbringen oder selbst direkt in kleinen mittelständischen Unternehmen umsetzen.

Ein technisches Informationssystem ist eine Kombination aus Hard– und Software, die den Anwender mit Informationen bestimmter Domänen versorgt (zur Zeit z.B. Fahrplanauskünfte, Bibliotheksdienstleistungen, Adreßauskünfte, didaktische Programme zum Fremdspracherwerb etc.). Hierbei erfolgt die Eingabe der Anfrage und die Ausgabe der Informationen immer häufiger audiovisuell (z.B. (Internet–)Telefonie). Wesentliche Aspekte sind dabei technische Realisierbarkeit, Ergonomie und Ethik (z.B. im Sinne einer wertenden Technologiefolgenabschätzung)¹. Alle Aspekte setzen ein grundlegendes Verständnis der gesprochenen Sprache unter phonetischen, linguistischen, psycholinguistischen, ingenieurwissenschaftlichen und philosophischen Gesichtspunkten voraus – unter Berücksichtigung der verbalen und nonverbalen Ausdrucksformen von Sprecherintentionen.

Das Graduiertenkolleg wird eine effiziente Nutzung der vorhandenen Infrastruktur ermöglichen, dem massiven Interesse der beteiligten Institutionen an einer aktiven Kooperation Rechnung tragen und eine Intensivierung vorhandener bilateraler Kooperationen ermöglichen.

2.2 Forschungsprogramm

Die nachfolgend beschriebenen Forschungsthemen bilden Vorschläge für anzustrebende Dissertationen bzw. post–doktorale Arbeiten innerhalb des Graduiertenkollegs. Sie sind nach den Forschungsschwerpunkten der beteiligten Institute geordnet. Vorangestellt sind jeweils die themenspezifischen Vorarbeiten der Institute mit entsprechenden Literaturhinweisen.

¹ Die ethischen Aspekte der Interaktion mit technischen Informationssystemen werden in diesem Antrag noch weitgehend ausgeklammert, da diese nur nach einer vorherigen Bearbeitung der anderen, grundlegenden Aspekte theoretisch fundiert reflektiert werden können.

2.2.1 Phonetik und Sprachliche Kommunikation (Prof. Dr. Tillmann)

2.2.1.1 Themenspezifische Vorarbeiten

Die Arbeiten des Instituts für Phonetik und Sprachliche Kommunikation (IPSK) decken im wesentlichen die Bereiche Sprachcorpora, Spracherkennung und –synthese, empirische Dialektologie sowie die Elektromagnetische Artikulographie und Physiologie ab. Sämtliche dieser Bereiche sind wesentlich, um Fragen der Realisierbarkeit und Ergonomie der gesprochen-sprachlichen Interaktion mit technischen Informationssystemen zu beantworten. Damit deckt das Institut verschiedene Forschungsbereiche ab, die in ihrer sowohl technischen als auch sprachwissenschaftlichen Ausrichtung verdeutlichen, weshalb das Institut sinnvollerweise in der Rolle des Sprechers die Koordinierung des Kollegs übernehmen soll:

- Sprachcorpora

In enger Kooperation mit dem Bayerischen Archiv für Sprachsignale (BAS) entwickelt und realisiert das Institut Sprachcorpora für verschiedene Anwendungszwecke (vgl. z.B. Schiel, Draxler & Tillmann, 1997). Alle *data driven models* der Sprachtechnologie benötigen Sprachdaten, um

- (a) mit diesen Spracherkenner zu trainieren (vgl. z.B. Schiel, Kipp & Tillmann, 1998)
- (b) aus diesen Modelle für das Dialogverhalten menschlicher Anwender zu gewinnen (vgl. z.B. Tischer, 1995, zum Ausdruck von Emotionen)
- (c) »Bausteine« für die Sprachsynthese zur Verfügung zu haben (vgl. z.B. Zboril, Libossek & Metze, 1997).

Hierbei kommt der Annotation der Sprachsignale ebenso besondere Bedeutung zu (vgl. z.B. Schiel, Burger, Geumann & Weilhammer, 1998) wie der Aufnahme von Sprachdaten in für technische Informationssysteme realistischen Umgebungen (vgl. z.B. Pfitzinger, 1998, zur Sprachaufnahme in Automobilen)

- Spracherkennung und Sprachsynthese

Spracherkennung wird am IPSK in Verbindung mit dem Problem der automatischen Segmentierung (vgl. Kipp, Wesenick & Schiel, 1998) betrachtet, welches als das grundlegendere angesehen werden kann. Daß dieser Ansatz für die automatische Spracherkennung außerordentlich befruchtend sein kann, wurde durch die Dissertation von Kipp, 1998, und die Diplomarbeit von Kaleschke, 1997, nachgewiesen. Darüber hinaus liegen aus dem Institut verschiedene Arbeiten vor, die sich mit linguistischen (vgl. Tischer, 1997) und semantisch–logischen (Zboril & Metze, 1996) Grundlagenproblemen der Spracherkennung auseinandersetzen.

Im Rahmen von Industriekooperationen werden auch Sprachsynthesysteme entwickelt und evaluiert (zur Evaluation vgl. z.B. Zboril, Libossek & Metze, 1997)

- Empirische Dialektologie

Während dialektale Färbungen der Hochsprache einen wesentlichen Aspekt der Natürlichkeit einer gesprochen–sprachlichen Systemausgabe darstellen können, sind sie

durch ihren Variantenreichtum häufig ein Problem für die Spracherkennung. Ausgehend von entsprechenden Corpora (vgl. z.B. Burger & Schiel, 1998) wurden deshalb bereits verschiedene Aspekte der menschlichen Dialekterkennung untersucht (vgl. z.B. Burger & Draxler, 1998).

- Elektromagnetische Artikulographie und Physiologie

Um die Interaktion von (Laut-)Sprache, Mimik und Gestik verstehen zu können, werden grundlegende Kenntnisse über die neuronalen, neuro-muskulären und muskulären Prozesse der Sprachproduktion benötigt, wobei naturgemäß ein besonderer Fokus auf die Interaktion der verschiedenen am Sprechvorgang direkt beteiligten (Gesichts-)Muskeln mit anderen Gesichtsmuskeln zu legen ist. Wesentlich ist natürlich, daß zuvor die entsprechenden Prozesse des Sprechens selbst, insbesondere die muskulären, erforscht wurden. Hierzu wurden durch die Arbeiten der Gruppe um Phil Hoole wesentliche Grundlagen gelegt (vgl. z.B. Hoole & Kühnert, 1996, zur Zunge-Kiefer-Koordination, oder Hoole & Nguyen, im Druck, allgemeiner zu artikulographischen Methoden der Koartikulationsforschung).

Erwähnt seien auch die von der DFG geförderten Forschungsprojekte in diesem Bereich, nämlich zur Realisierung des sehr komplexen Vokalsystems des Deutschen, zur Entwicklung der drei-dimensionalen Artikulographie und, im Schwerpunkt Sprachproduktion, zur kompensatorischen Artikulation; ein weiteres DFG-Projekt ist dem Unterschied zwischen der normalen Sprechstimme und der Gesangsstimme gewidmet (mit dem Ziel, objektive Kriterien für die Bewertung guter Singstimmen zu entwickeln).

Schließlich liegen auch empirische Untersuchungen zur Gestaltung der Lehre im Fach Phonetik und Sprachlichen Kommunikation vor, die nachdrücklich die besondere Bedeutung der Sprachtechnologie als wohl wichtigstes Anwendungsgebiet phonetischen Wissens unterstreichen (vgl. z.B. Hazan et al., 1997).

Das IPSK ist in verschiedene Drittmittelprojekte (insbesondere Industriekooperationen) eingebunden, so daß neben dem theoretischen Wert der gewonnenen Erkenntnisse (für den die Literaturhinweise ein Indiz sein mögen) auch die hohe praktische Relevanz unterstrichen werden kann. Hierzu seien das VERBMOBIL-Projekt, das SPEECHDAT-Projekt (als Partner der Siemens AG), das AUTOSCOOUT-Projekt (ebenfalls gemeinsam mit der Siemens AG), die Entwicklung eines dreidimensionalen Artikulographen (gemeinsam mit der Firma Carstens sowie NTT, Tokyo), das RVG-Projekt (in Kooperation mit Bell Laboratories & Lucent Technologies) als Beispiele angeführt. Daß drei der jüngeren Mitarbeiter des Instituts unter den Preisträgern des Gründerwettbewerbs Multimedia des BMFT 1998 waren, zeigt auch, daß das Know-how des Instituts in innovativer Weise zur Existenzgründung umgesetzt werden kann.

Literaturhinweise²

Burger, S. & Schiel, F. (1998): RVG 1 – A Database for Regional Variants of Contemporary German; Proceedings of the FIRST INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION 1998, Granada, Spain.

² Bei den Literaturhinweisen der *Themenspezifischen Vorarbeiten* der beteiligten Institute kann es sich jeweils nur um eine stark fokussierte Auswahl handeln.

Burger, S. & Draxler, Ch. (1998): Identifying Dialects of German from Digit Strings; Proceedings of the FIRST INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION 1998, Granada, Spain.

Hazan, V., Dommelen, W. van, Cutugno, F., Fougeron, C., Frid, J., Köster, J.-P., Machuca, M., Mengel, A., Turk, A. & Zboril, D. (1997): Phonetics; in: Bloothof et al. (eds.): The Landscape of Future Education in Speech Communication Sciences – 1 Analysis; OTS; Utrecht.

Hoole, P. & Kühnert, B. (1996). Tongue–jaw coordination in German vowel production. Proceedings of the 1st ESCA tutorial and research workshop on Speech Production Modelling/4th Speech Production Seminar, Atrians, 1996, pp. 97–100.

Hoole, P. & Nguyen, N. Electromagnetic articulography in coarticulation research. In: W.H Hardcastle and N. Hewlett (eds.), Instrumental Studies of Coarticulation. Cambridge University Press (in press).

Kaleschke, Ch. (1997): Untersuchung von automatischen Spracherkennungssystemen mit Berücksichtigung von Aussprachevarianten, Diplomarbeit.

Kipp, A. (1998): Automatische Segmentierung und Etikettierung von Spontansprache, Dissertation.

Kipp, A., Wesenick, B. & Schiel, F. (1997): Pronunciation Modeling Applied to Automatic Segmentation of Spontaneous Speech; in: Proceedings of the EUROSPEECH 1997, Rhodos, Greece, pp. 1023–1026.

Pfitzinger, H.R. (1998): The Collection of Spoken Language Resources in Car Environments. Proc. of the First International Conference on Language Resources and Evaluation, vol. 2, pp. 1097–1100. Granada, Spain.

Schiel, F. (1993): Untersuchungen zur Sprecheradaption in Systemen zur automatischen Spracherkennung mit Hilfe stochastischer Modellierung. Verlag Shaker, Reihe Informatik, Aachen.

Schiel, F., Draxler, Ch. & Tillmann, H.G. (1997): The Bavarian Archive for Speech Signals: Resources for the Speech Community; in: Proceedings of the EUROSPEECH 1997, Rhodos, Greece, pp. 1687–1690.

Schiel, F., Burger, S., Geumann, A. & Weilhammer, K. (1998): The Partitur Format at BAS; Proceedings of the FIRST INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION 1998, Granada, Spain.

Schiel, F., Kipp, A. & Tillmann, H.G. (1998): Statistical Modeling of Pronunciation: It's not the Model, it's the data; Proceedings of the ESCA Tutorial and Research Workshop on 'Modeling Pronunciation Variation for Automatic Speech Recognition', May 1998, Kerkrade/Netherlands.

Tillmann, H.G. (1997): Die Mensch–Maschine–Kommunikation; in: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, Nr. 35, S. 3–16.

Tillmann, H.G. (1997): Eight Main Differences between Collections of Written and Spoken Language Data; in: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation, Nr. 35, S. 139–144.

Tischer, B. (1995): Acoustic correlates of perceived emotional stress. In: Proceedings of the ESCA–NATO Tutorial and Research Workshop on Speech under Stress (Lisbon, Portugal, 14–15 September 1995). Lisbon. S. 29–32.

Tischer, B. (1997): Selbstkorrekturen im Dialog. Regeln zur automatischen Syntaxverarbeitung. Linguistische Berichte 170, 312–344.

Zboril, D. & Metze, F. (1996): Indeterminateness in Qualitative and Quantitative Reasoning; in: Seventh International Workshop on Database and Expert Systems Applications; IEEE Computer Sciences Press, Los Alamitos.

Zboril, D., Libossek, M. & Metze, F. (1997): In–Development Assessment of Concatenation Synthesis by Nonnative Speakers; in: FIPKM 35: 115–119.

2.2.1.2 Skizzen möglicher Dissertationsprojekte

Die Phonetik und Sprachliche Kommunikation beschränkt sich im komplexen Gefüge sprachlicher Informationssysteme auf die lautsprachliche Übermittlung von Information, einerseits vom Menschen zum System (Spracheingabe), andererseits vom System zum Menschen (Sprachausgabe). Die dabei zu untersuchenden Aspekte sind:

- Adaptierung auf lautsprachlicher Ebene
- Experimentelle Untersuchungen zur Natürlichkeit der Sprachausgabe (Sprachsynthese)
- Detektion relevanter Informationen in Sprachsignalen (semantischer Fokus)
- Sonderthema: Gebärdensprache

Adaptierung auf lautsprachlicher Ebene

Strategien zur Sprecheradaption auf lautsprachlicher Ebene werden bereits seit ca. 15 Jahren im Gebiet der automatischen Spracherkennung verfolgt (z.B. Schiel, 1993). Der übliche Ansatz ist dabei eine statistische Modellierung der von einem neuen Sprecher benutzten Sprachmuster und eine schrittweise Anpassung der Modell–Parameter des Erkenners. Ein Problem ist durch die Tatsache gegeben, daß alle diese Ansätze keinerlei a priori vorhandenes Wissen über Sprecherpopulationen und die damit verbundenen Korrelationen von lautsprachlichen Effekten innerhalb solcher Populationen verwenden. Es wäre zu untersuchen, ob es mit Hilfe der heutzutage verfügbaren dialektspezifischen Sprachdatenbasen (Burger & Schiel, 1998) möglich ist, aus Einzelbeobachtungen am akustischen Sprachsignal des unbekanntem Sprechers auf andere regelmäßige lautsprachliche Effekte, die für die Sprecher–Population typisch sind, zu schließen (Burger & Draxler, 1998). Neben dialektspezifischen Sprechercharakteristika spielen auch soziolektspezifische sowie idiolektspezifische eine wesentliche Rolle für den Adaptionsprozeß, wobei sich insbesondere die Frage stellt, inwiefern visuelle Daten zur Modellierung der idiolektalen Sprechercharakteristika herangezogen werden können. Die

Beantwortung der hier aufgezeigten Fragestellungen wird wesentlich dazu beitragen, daß der Prozeß der Sprecheradaption eines sprachlichen Informationssystems – ähnlich wie beim Menschen – innerhalb von wenigen Sekunden erfolgen kann.

Experimentelle Untersuchungen zur Natürlichkeit der Sprachausgabe (Sprachsynthese)

In den vergangenen Jahren hat sich die konkatenative Synthese als die am natürlichsten klingende herausgestellt, so daß sie sich der klassisch-parametrischen Synthese zur Zeit als überlegen zu erweisen scheint. Dabei setzt sie das synthetische Sprachsignal aus Teilstücken von unter Studiobedingungen aufgezeichneten, sehr umfangreichen Äußerungen eines Sprechers, dessen Sprechercharakteristik weitgehend beibehalten bleibt, neu zusammen. Bei der konkatenativen Synthese (patentierte Verfahren: PSOLA, CHATR) werden die Probleme der natürlichen Anregung durch die Glottis sowie die Modellierung der Koartikulation verlagert in drei neue Problemkomplexe, die noch weiterer, sehr grundlegender experimenteller Untersuchungen bedürfen.

- Zum ersten klingt die aus Teilstücken zusammengesetzte Synthesestimme zwar sehr natürlich, doch wie läßt sich die Synthese auf die Stimme und den Idiolekt eines anderen Sprechers verändern?
- Zum zweiten stellen Wörter in ihrer isolierten Aussprache (also in einer akzeptablen Zitierform) auch bei dieser Art der Synthese ein großes Problem dar, insbesondere dann, wenn die dazu erforderlichen Teilstücke im Ausgangssprachkorpus fehlen.
- Zum dritten stellt sich die Frage nach den am besten geeigneten elementaren Äußerungsfragmenten. Die Größe dieser Signalstücke und die jeweilige Strategie ihrer Konkatenation sind eng miteinander verknüpft. Während bei größeren Einheiten (z.B. Wörtern, Phrasen) in vielen Fällen eine prosodische Manipulation, die bezüglich der Intonation gut erforscht ist, als hinreichend betrachtet werden kann, gibt es bei kleineren Einheiten (z.B. Diphonen, Halbsilben, Silben) noch viele ungelöste Probleme, die wir untersuchen wollen. Genügt es z.B. (um die hörbaren Diskontinuitäten der konkatenativen Synthese zu vermeiden), die Methoden zur Auffindung der jeweils für den Signalkontext benötigten Signalstücke in einer entsprechend großen Signaldatenbank eines gegebenen Sprechers zu verbessern? Oder müssen sie einer Reihe von Signalmanipulationen (u.a. Veränderung von Formanttransitionen und des Grades der spektralen Ausprägung von Zielpositionen) unterzogen werden? Hierbei kommen wieder Aspekte der parametrischen Sprachsynthese ins Spiel, wenn auch in der Form der gezielten Manipulation vorgegebener, natürlich erzeugter Ausgangsdaten.

Die Beantwortung der hier aufgeworfenen Fragen setzt ein grundlegendes Verständnis der Regelmäßigkeiten der Sprachproduktion (einschließlich der akustischen Artikulation) und der Sprachperzeption voraus. Für experimentelle Untersuchungen auf diesem Gebiet steht durch die am Lehrstuhl aufgebaute Synthesystem-Entwicklungsumgebung ein effektives Hilfsmittel im Analysis-by-Synthesis-Paradigma zur Verfügung, mit dem sehr gezielt bestimmte Hypothesen und Theorien überprüft werden können.

An eine solche erste (und damit grundlegende) experimentelle Untersuchung können sich dann weitere Dissertationen anschließen, in denen z.B. untersucht wird, inwieweit die für die gezielte Modifikation von Signalstücken benötigten Parameter vom »lokalen Sprechtempo« abhängen, das sich ja mit dem prosodischen Ablauf einer Äußerung ändert. Außerdem kann dann besser untersucht werden, wie die mimische Dynamik von

Gesichtsbewegungen mit der akustischen Sprachsynthese in Übereinstimmung gebracht werden kann (vgl. dazu auch die Projekte in 2.2.4 und 2.2.6).

Detektion relevanter Informationen in Sprachsignalen (semantischer Fokus)

Die Entwicklung auf dem Gebiet der automatischen Spracherkennung der letzten Dekade hat gezeigt, daß bei spontansprachlicher Eingabe von Information keinesfalls damit gerechnet werden kann, daß ein technisches System den tatsächlich gesprochenen Text wortwörtlich, d.h. mit allen spontansprachlichen Komponenten (wie Häsitationen und Abbrüchen) sowie metasprachlichen Einschüben und Überlagerungen (wie Lachen, Räuspern etc.) zu verarbeiten und dem Sinn nach zu interpretieren in der Lage ist. Ähnlich wie auch der Mensch bei der Aufnahme von Sprache nicht das Hauptaugenmerk auf die Details der sprachlichen Wiedergabe legt, sondern sich auf den wesentlichen Inhalt fokussiert, müßten sprachliche Informationssysteme eigentlich selbsttätig die relevanten, d.h. informationstragenden, Teile des akustischen Sprachsignals identifizieren. Zur Bewältigung der jeweiligen Aufgabe müssen der Kontext erfaßt und durch diesen den lokalisierten Signalstücken mögliche bzw. wahrscheinliche Sprecherintentionen zugeordnet werden können. Hier dürfte, wie in einer gemeinsamen Lehrveranstaltung mit Dr. Harald Höge (dem Leiter der Entwicklungsabteilung für Sprachtechnologien der Siemens AG) gezeigt wurde, der klassische Entropiebegriff der Informationstheorie eine neue Bewertung erfahren. Interessante Fragestellungen in diesem Zusammenhang sind:

- Welche Strategien verfolgt der Mensch bei dieser Aufgabe?
- Sind akustische Korrelate zum Informationsgehalt im Sprachsignal nachweisbar?
- Wie können diese technisch verwertet werden?

Sonderthema:

Seit dem Wintersemester 1999 gehören einführende Lehrveranstaltungen zur Gebärdensprache der Gehörlosen zum Studienprogramm des IPSK. Deshalb bietet es sich in einem Graduiertenkolleg zum Thema Sprache, Mimik und Gestik im Kontext technischer Informationssysteme geradezu an, Fragestellungen wie die der automatischen Übersetzung von gebärdensprachlichen Äußerungen in orthographische Texte und deren lautsprachliche Realisierung (in enger Kooperation mit der digitalen Bildverarbeitung) zu untersuchen und im Rahmen einer Dissertation in den grundlegenden Aspekten zu behandeln, und dies gilt auch umgekehrt für die Überführung von textsprachlich formulierter Information in gebärdensprachliche Äußerungen auf den Terminal eines Rechners.

2.2.2 Sprechwissenschaft und Psycholinguistik (Prof. Dr. Kegel)

2.2.2.1 Themenspezifische Vorarbeiten

Die vorgeschlagenen Projekte psycholinguistischer Forschung können den Bereichen »Menschliche Zeit- und Sprachverarbeitung«, »Sprache und Emotion« und »Handlungsbegleitende Sprache« zugeordnet werden.

Die Frage nach dem Zusammenhang zwischen Zeit- und Sprachverarbeitung beschäftigt unser Fach schon seit einigen Jahren. Neben einem DFG-Projekt zur Untersuchung der subjektiven Zeitverarbeitung bei sprachauffälligen und sprachunauffälligen Kindern (Kegel

et al., 1988) sowie Forschungsarbeiten zur Modellbildung zum Zusammenhang von Sprach- und Zeitverarbeitung (Steffen & Werani, 1996; Kegel & Tramitz, 1991, Dames & Lautenbacher, 1989), entstehen regelmäßig Magisterarbeiten (Dames, 1985; Steffen, 1995; Moser, 1995) und derzeit auch Promotionen zu diesem Thema (Moser, Steffen). Neben den zeitlichen Aspekten sollen zudem psychophysiologische Daten als objektivierendes Kriterium erhoben werden. Erkenntnisse aus der langjährigen Erfahrung mit der Erhebung, Auswertung und Interpretation psychophysiologischer Daten im Rahmen psycholinguistischer Fragestellungen, insbesondere im Gebiet der Sprach- und Medienforschung, sollen dazu herangezogen werden (Kegel et al., 1985; Arnhold, 1984; Dahlmeier, 1984). Die vorgeschlagenen Projekte sollen die bisher gewonnenen Erkenntnisse auf komplexerer Ebene ergänzen (s. Projekt 1 und Projekt 2).

Der Bereich »Emotion und Sprache« ist seit langem Schwerpunkt unserer psycholinguistischen Forschung (Tischer, 1987, 1988a, 1988b, 1993). Dabei wurden psycholinguistische und phonetische Ansätze erfolgreich miteinander verknüpft. Auch hier entstanden mehrere Magisterarbeiten. Diese Kenntnisse sollen in einem weiteren Projekt erhalten und vertieft werden (s. Projekt 3).

Der Bereich »Handlungsbegleitende Sprache« beim Problemlösen ist momentan der aktuelle Fokus des Forschungsinteresses. Aus diesem Bereich entsteht ein Projekt, welches das psycholinguistische Frageinteresse zur handlungsbegleitenden Sprache mit Aspekten aus der Aphasie verbindet (Werani und Kegel, 1998). Ferner ist gerade eine experimentelle Untersuchung der handlungsbegleitenden Sprache in der Interaktion mit dem Computer im Abschluß begriffen, die sich speziell mit interaktionsspezifischen Problemlösungsstrategien und nutzererwünschten »sprachlichen Fähigkeiten« des Computers befaßt (Kachelrieß, Dissertation in der Abschlußphase). Die Ausweitung dieses Bereichs ist hinsichtlich pathologisch-therapeutischer Aspekte sowie der Mensch-Maschine-Kommunikation überaus erstrebenswert (s. Projekt 4).

Weitere einschlägige Aktivitäten des Faches betreffen fach- und institutionenübergreifende Kongresse und Workshops u.a. zum Bereich Grundschule und Computer mit spezieller Thematisierung der Förderung des Schriftspracherwerbs (Huber et al., 1998).

Literaturhinweise

Arnhold, T. (1984): Psychophysiologische Reaktionen auf verbale Stimuli: Die experimentelle Erfassung von Aufmerksamkeitsveränderungen während der auditiven Sprachwahrnehmung. Magisterarbeit. LMU München.

Dahlmeier, K. (1984): Reaktionen des autonomen Nervensystems auf sprachliche Reize. Die Messung der elektrodermalen Aktivität zur Erfassung der Aufmerksamkeit beim Lesen eines Tests. Magisterarbeit. LMU München.

Dames, K. & Lautenbacher, S. (1989): Die Zeitstruktur von Nachsprechleistungen bei sprachentwicklungsgestörten Kindern. In: Kegel, G.; Arnhold, T.; Dahlmeier, K.; Schmid, G.; Tischer, B. (Hrsg.): Sprechwissenschaft und Psycholinguistik 3. Opladen.

Dames, K. (1985): Einfluß der Syntax auf die Zeitstruktur der Nachsprechleistungen sprachentwicklungsgestörter und sprachunauffälliger Kinder. Magisterarbeit. LMU München.

Huber, L., Kegel, G. & Speck–Hamdan, A. (1998): Einblicke in den Schriftspracherwerb. Braunschweig. Westermann Schulbuchverlag GmbH.

Kegel, G. (1985): Kind ohne Sprache. In: Füssenich, I.; Gläß, B. (Hrsg.): Dysgrammatismus. Theoretische und praktische Probleme bei der interdisziplinären Beschreibung gestörter Kindersprache. Heidelberg.

Kegel, G. (1990): Sprach- und Zeitverarbeitung bei sprachauffälligen und sprachunauffälligen Kindern. In: Kegel, G.; Arnhold, T.; Dahlmeier, K.; Schmid, G.; Tischer, B. (Hrsg.): Sprechwissenschaft und Psycholinguistik 4. Opladen.

Kegel, G., Arnhold, T. & Dahlmeier, K. (1985): Sprachwirkung. Opladen. Beiträge zur psychologischen Forschung, Bd. 6.

Kegel, G., Dames, K. & Veit, S. (1988): Die zeitliche Organisation sprachlicher Strukturen als Sprachentwicklungsfaktor. In: Kegel, G.; Arnhold, T.; Dahlmeier, K.; Schmid, G.; Tischer, B. (Hrsg.): Sprechwissenschaft und Psycholinguistik 2. Opladen.

Kegel, G. & Tramitz, C. (1991): Olaf – Kind ohne Sprache. Opladen.

Moser, S. (1995): Zeitverarbeitung und Sprachstörung. Zur Entwicklung eines diagnostischen Verfahrens. Magisterarbeit. München.

Steffen, A. (1995): Zur Bestimmung der Ordnungsschwelle. Ein experimenteller Vergleich.

Steffen, A. & Werani, A. (1996): Ein Experiment zur Zeitverarbeitung bei der Sprachwahrnehmung. In: Kegel, G.; Arnhold, T.; Dahlmeier, K.; Schmid, G.; Tischer, B. (Hrsg.): Sprechwissenschaft und Psycholinguistik 6. Opladen.

Tischer, B. (1987): Sprache und Emotion. Theoriebildende Ansätze und ihre Bedeutung für Sprechwissenschaft und Psycholinguistik. In: Kegel, G.; Arnhold, T.; Dahlmeier, K.; Schmid, G.; Tischer, B. (Hrsg.): Sprechwissenschaft und Psycholinguistik 2. Opladen.

Tischer, B. (1988a): Kein Spaß im Wortfeld der Gefühlsbegriffe? Zehn Methoden zur Kennzeichnung der Wortfeldzugehörigkeit. *Archiv für Psychologie*, 140, 15–31.

Tischer, B. (1988b): Ein Verfahren zur Messung der vokalen Kommunikation von Gefühlen. *Sprache & Kognition*, 7, 205–216.

Tischer, B. (1993): Die vokale Kommunikation von Gefühlen. Weinheim. Psychologie Verlags Union.

Werani, A. & Kegel, G. (1998): Die Relevanz handlungsbegleitender Sprache für Aphasiker beim schlußfolgernden Denken – Ein Beitrag zum Zusammenhang von Sprache und Denken, eingereichter DFG-Antrag).

2.2.2.2 Skizzen möglicher Dissertationsprojekte

Zeit- und Signalverarbeitung in der Mensch-Maschine-Interaktion

Die menschliche Verarbeitung gesprochener Sprache wird durch drei interne interagierende Zeitverarbeitungssysteme (subjektive „Uhren“) gesteuert. Im Einzelnen unterstützen diese Systeme die Segmentierung, Klassifizierung und Integration von Signalstrecken auf der Laut-, Silben- und Satzebene. Die Systeme regeln neben der auditiven ebenfalls die visuelle und die taktile Wahrnehmung. Ihre Metrik scheint sich aber zum großen Teil über die zeitimmanenten Anforderungen des Sprachsignals herauszubilden. Einerseits korrelieren Störungen der internen Zeitverarbeitung, wie sie bei Kindern und Aphasikern beobachtet werden, mit teilweise massiven Störungen der Sprachverarbeitung. Andererseits führen Signale, die der subjektiven Zeitverarbeitung nicht angepaßt sind, zu Problemen bei ihrer Wahrnehmung und Verarbeitung. Das gilt auch für die akustische und optische (nicht unbedingt sprech- oder schriftsprachliche) Signalgebung des Computers. Das Projekt konzentriert sich auf die Anpassung dieser Signalgebung an die menschliche Zeitverarbeitung mit dem Ziel, einen Beitrag zur Optimierung der Interaktion zu leisten.

Der Einfluß von Orientierungsreizen auf Verstehen und Behalten gesprochener Sprache

Wahrnehmung und Verstehen gesprochener Sprache wird durch sprachliche und parasprachliche Aufmerksamkeitsimpulse gesteuert respektive teilweise erst ermöglicht. Die vom Sprecher gesetzten Aufmerksamkeitsimpulse führen beim Hörer zu Orientierungsreaktionen mit begleitenden Reaktionen des autonomen Systems, z.B. kurzfristigen Veränderungen der Hautleitfähigkeit, die sich psychophysiographisch erfassen lassen. Die Orientierungsreaktionen gewährleisten aufgrund der Senkung von Wahrnehmungsschwellen die bessere Aufnahme der informationstragenden Passagen der Rede. Für die Erfassung der kurz- und langfristigen Verstehens- und Behaltensleistung stehen bewährte Techniken zur Verfügung. Das Projekt konzentriert sich auf die eigentlich sprachlichen Aufmerksamkeitsimpulse, die im Prinzip ja bereits in der klassischen Rhetorik erörtert wurden. Grundsätzlich zu klären sind Form und Wirkungsausmaß der Aufmerksamkeitsimpulse sowie ihr Timing im Redefluß. Speziell zu klären ist die Abhängigkeit dieser Größen von Redesituation und Textsorte.

Emotionale Einflußgrößen bei sprechsprachlicher Information

Das Projekt konzentriert sich auf den emotionalen Status von Personen während der Sprachproduktion und dessen Auswirkung auf die Sprechleistung. Zu unterscheiden ist zwischen „konstitutiven emotionalen und situationsspezifischen emotionalen Größen. Zum ersten Bereich sind Größen wie Intro-/Extrovertiertheit, Stressempfänglichkeit, Redeangst, Selbstbewußtsein etc. zu erfassen. Hierzu stehen Testverfahren zur Verfügung. Der zweite Bereich ist zu weit, als daß er annähernd repräsentativ untersucht werden könnte. Daher wird eine Beschränkung auf Belastungsgrößen wie Videoaufzeichnung, Hörerreaktion und Störungen vorgenommen. Neben einer Sprecherbefragung sollten zur Auswirkung dieser Größen psychophysiographische Erhebungen in Erwägung gezogen werden. Zentral für die Auswertung sind Veränderungen im gestischen und phonetischen Verhalten (z. B. Reduzierung der Gestik, gefüllte Pausen, Sprechrate). Hinsichtlich Sprache, Körpersprache und Emotionalität werden geschlechtsspezifische Differenzen behauptet, die zum Teil in der Forschung auch nachgewiesen wurden. Die Untersuchung ist unter dem Gesichtspunkt dieser Fragestellung zu planen.

Struktur und Funktion von Sprache bei komplexen Problemen im problemlösenden Dialog

Die Erforschung handlungsbegleitender Sprache ist in der Psycholinguistik ein zentrales Thema. Nach Wygotski geht die egozentrische Sprache des Kindes in die innere Sprache über und übernimmt Funktionen der Handlungsplanung, Handlungssteuerung und Handlungskontrolle. Mit der Entstehung der inneren Sprache und somit eines inneren Dialogs vollzieht sich in gewissen Belangen eine Entkoppelung von der Umwelt. Die Außenlenkung der Bezugsperson nimmt ab, und der innere Dialog übernimmt selbststeuernde Funktionen. Die Unterteilung in eine kommunikative und eine kognitive Funktion der Sprache wird offensichtlich. In diesem Projekt geht es um die Betrachtung dieser verschiedenen sprachlichen Funktionen. Durch das Lösen komplexer Probleme soll die innere Sprache nach außen treten, ferner soll durch eine dialoge Situation mit dem Versuchsleiter in den Problemlöseprozeß eingegriffen werden. Somit können verschiedene sprachliche Analysen durchgeführt werden:

1. die handlungsbegleitende Sprache des Problemlösenden,
2. die korrigierende Sprache des Versuchsleiters,
3. die sprachlichen Reaktionen des Problemlösenden. Die Analyse trägt zur Klärung von Struktur, Funktion und Stabilität handlungsbegleitender Sprache bei.

2.2.3 Philosophie (Prof. Dr. Kemmerling)

2.2.3.1 Themenspezifische Vorarbeiten

Eine Frage, die im Zentrum meiner Forschung der vergangenen 20 Jahre stand, war: Welche Rolle spielen Absichten und ihre Erkenntnis in der sprachlichen Verständigung und zwar insbesondere in derjenigen Art sprachlicher Verständigung, die konstitutiv ist für die soziolektale Bedeutung von Wörtern und Sätzen? Der Ausgangspunkt meiner Überlegungen dabei ist ein gebrauchstheoretischer: Die soziolektale Bedeutung verdankt sich der Art und Weise, wie sprachliche Zeichen verwendet werden. Aber welchem Aspekt der Verwendung? Auf diese Frage gibt es zwei einander entgegengesetzte idealtypische Antworten. Die intentionalistische, die insbesondere auf das Werk von Paul Grice zurückgeht, besagt: Es hängt davon ab, welche Absichten der Sprecher mit seiner Äußerung bekundet. Die konventionalistische Antwort, die mit sehr unterschiedlicher Tendenz ausgearbeitet wurde (Wittgenstein, Austin, Searle, v. Savigny wären hier zu nennen), besagt: Es hängt davon ab, welche geregelten Verschiebungen der Rechte und Pflichten durch die Äußerung zustande kommen. In beiden Antworten wird eine Konzeption davon vorausgesetzt, was ein Äußerungskontext ist und wie seine Merkmale in der sprachlichen Verständigung wirksam werden. Für den intentionalistischen Gebrauchstheoretiker ist der Kontext vornehmlich ein Ort, an dem sich das individuelle Wissen von Sprecher und Adressat zu einem "gemeinsamen" (oder "wechselseitigen") Wissen zusammenfügen, von dem beide profitieren: Der Sprecher braucht nur wenig zu äußern, um viel zu sagen und noch mehr zu verstehen zu geben; der Adressat braucht nur wenig zu hören, um mit Leichtigkeit viel zu verstehen. Für den konventionalistischen Gebrauchstheoretiker hingegen hat der Kontext weniger mit kognitiven Kategorien wie Wissen und Verstehen zu tun als mit normativen wie: zulässige Äußerung und korrekte Reaktion. In diesem Arbeitsbereich betreue ich derzeit eine Dissertation mit dem Arbeitstitel "Nicht-sprachliche Mittel in sprachlicher Verständigung", in der es u.a. darum gehen wird, welche besondere

Rolle ikonische und mimetische Zeichen in sprachlicher Verständigung spielen – wobei die Besonderheit dieser Rolle u.a. darin zu sehen ist, daß diese Zeichen nicht als Repräsentationen, sondern als bezugnehmende Präsentationen funktionieren. Der Begriff des Bezugnehmens steht auch im Zentrum eines von mir geleiteten DFG-Projekts ("Der Begriff des Bezugnehmens im Lichte des Vergleichs von sprachlichem und bildlichem Bezugnehmen").

Ein weiterer Forschungsschwerpunkt betrifft die Frage, ob überhaupt und falls ja, in welchem Ausmaß Computersimulationen kognitiver Leistungen darüber Aufschluß geben können, wie Menschen diese Leistungen vollbringen. Ein grundlegendes Problem dabei ist, ob dem programmierten Computer unter Umständen ein besonderer Status zwischen bloß instrumentellem Artefakt und intelligentem Subjekt zukommt, und welche Merkmale eine Begrifflichkeit haben müßte, mit der dieser Sonderstatus angemessen wiedergegeben werden könnte. (Beispielsweise scheint ja klar, daß ein Computer Absichten nicht in dem Sinne ausdrücken kann wie ein Mensch, der sie sprachlich bekundet, ja nicht einmal wie ein Tier, das sie im Verhalten deutlich werden lassen kann. Dennoch gibt es vielerlei Verrichtungen von Computern, angesichts deren es einfach zu wenig ist, wenn nur darauf insistiert wird, daß sie aus begrifflichen Gründen keine Absichten haben können.)

Literaturhinweise

Kemmerling, A. (1976a): Bedeutung und Sprachverhalten, in: E. von Savigny (Hrsg.), Probleme der sprachlichen Bedeutung, Kronberg, 73–99

Kemmerling, A. (1976b): Konvention und sprachliche Kommunikation, Münchner Dissertation.

Kemmerling, A. (1977): Was Grice mit "Meinen" meint – Eine Rekonstruktion der Griceschen Analyse rationaler Kommunikation, Forschungsberichte des IPK der Universität München 8, 1977, 121 – 166

Kemmerling, A. (1979a): Gebrauchstheorie ohne konstitutive Regeln, in: Zwischenstation, Universität Bielefeld, 124– 127

Kemmerling, A. (1979b): Was Grice mit "Meinen" meint, in G. Grewendorf (Hrsg.), Sprechakttheorie und Semantik, Frankfurt, 67–118, (erweiterte und überarbeitete Fassung von Kemmerling, 1977)

Kemmerling, A. (1979): Zur Empirie der Semantik, Forschungsberichte des IPK der Universität München 12, 65–10

Kemmerling, A. (1980): How Many Things Must a Speaker Intend, Before He is Said to Have Meant?, Erkenntnis 15 (1980), 333–41

Kemmerling, A. (1983): Starre Designation und kontingente Wahrheiten a priori, Forschungsberichte des IPK der Universität München 18.

Kemmerling, A. (1986): Utterer's Meaning Revisited, in: R. Grandy/R. Warner (eds.), Philosophical Grounds of Rationality: Intentions. Categories. Ends, Oxford, 131–155

Kemmerling, A. (1988): Philosophischer Kognitivismus und die Repräsentation sprachlichen Wissens, in: G. Heyer/J. Krems/G.Görtz (Hrsg.), Wissensarten und ihre Darstellung, Heidelberg/Berlin/New York, 21–6

Kemmerling, A. (1987): Die Maschine spricht Deutsch. (Oder nicht?), in: Vorträge des Germanistentages Berlin 1987, Bd. 4, hrsg. Von N. Oellers, Tübingen 1988, 63–71

Kemmerling, A. (1989): Kognitive Wissenschaft und umgangssprachlicher Überzeugungsbegriff, in: B. Becker (Hrsg.), Zur Terminologie in der Kognitionsforschung, Arbeitspapiere der GMD 385, St. Augustin, 150–1

Kemmerling, A. (1991): Grice, in: J. Nidaümelin (Hrsg.), Philosophie der Gegenwart in Einzeldarstellungen. Stuttgart, 199–204

Kemmerling, A. (1991): Implikatur, in: A. von Stechow/D. Wunderlich (Hrsg.), Semantik – Ein internationales Handbuch der zeitgenössischen Forschung, Berlin/New York, 319–333

Kemmerling, A. (1992): Bedeutung und der Zweck der Sprache, in: W. Vossenkuhl (Hrsg.), Von Wittgenstein lernen, Berlin, 99–120

Kemmerling, A. (1992): Der Geist als Computer, in: A.B. Cremers/R. Haberbeck/J. Seetzen/I. Wachsmuth (Hrsg.), Künstliche Intelligenz – Leitvorstellungen und Verantwortbarkeit. VDI Report 17, 16–27

Kemmerling, A. (1993): The Philosophical Significance of a Shared Language, in: R. Stoecker (Hrsg.), Reflecting Davidson – Donald Davidson Responding to an International Forum of Philosophers, Berlin/New York, 85–116

Kemmerling, A. (1997): Die (sei's auch metaphorische) These vom Geist als Computer, in: W. Lenzen (Hrsg.), Das weite Spektrum der Analytischen Philosophie. Festschrift für Franz von Kutschera, Berlin/New York, 112–134

Kemmerling, A. (1997): Der bedeutungstheoretisch springende Punkt sprachlicher Verständigung, in: Kommunikationsversuche. Theorien der Kommunikation, hrsg. von G.–L. Lueken, Leipzig, 60–106

Kemmerling, A. (1997): Zur sog. Naturalisierung von Intentionalität, in: A. Burri (Hrsg.), Sprache und Denken, Berlin/New York, 237–258 (eine frühere Fassung erschien in: Wittgenstein Studies 1997)

Kemmerling, A. (1997): Zwei philosophische Konzeptionen des Sprachverstehens, 24 S., erscheint in: U. Metschl/U. Winko (Hrsg.), Analyse und Verstehen, Paderborn

Kemmerling, A. (1999): Gricy Actions, 20 S., erscheint in: G. Cosenza (ed.), Paul Grice's Heritage, Oxford

Kemmerling, A. (o.D.): Sagen können, was man denkt und glaubt, 47 S., Manuskript

2.2.3.2 Skizzen möglicher Dissertationsprojekte

Grundlegende Mehrdeutigkeiten des Kontextbegriffs in Semantik und Pragmatik

Der Kontextbegriff spielt eine Schlüsselrolle in der zeitgenössischen Semantik und Pragmatik. Seit Frege ("Der Gedanke", 1918) wird der Begriff des Äußerungskontexts herangezogen, um das semantische Funktionieren von Indexikalia und Demonstrativa zu erklären. Dieser Aspekt ist in den vergangenen 25 Jahren insofern immer wichtiger geworden, als eine Reihe von neueren bedeutungstheoretischen Argumenten darauf hinzudeuten scheinen, daß es neben den offenkundig indexikalischen bzw. demonstrativen Ausdrücken (wie "ich", "hier", "heute" oder "der da", "dort" usw.) auch eine unübersichtlich große Vielzahl von sog. versteckt indexikalischen Ausdrücken gibt (vieldiskutierte Beispiele sind "wissen" und "Wasser"). Der Kontext fungiert hier als Komponente der Bezugsbestimmung. Eine andere explanatorische Aufgabe des Kontextbegriffs liegt darin, zu verschiedenen Arten der Desambiguierung im weitesten Sinn beizutragen. Zur Desambiguierung ist dabei zu rechnen: die Auflösung lexikalischer und struktureller Mehrdeutigkeiten, die Beseitigung von Vagheitsphänomenen aller Art, die Ergänzung von elliptischen Sätzen, die Anreicherung von inhaltsarmen Äußerungen, die "Relevantierung" von unthematischen Gesprächsbeiträgen und so weiter.

In diesem Projekt soll die Frage untersucht werden, ob es sich in diesen beiden Falltypen Bezugsbestimmung und Desambiguierung nicht um vollständig unterschiedliche Begriffe von Kontext ("konversationaler" versus "konkreter" Kontext) handeln muß, die in der zeitgenössischen Fachdiskussion in der Linguistik und Sprachphilosophie nicht klar genug von einander getrennt werden, woraus sich insbesondere auch Fehlcharakterisierungen des eigentlichen Problems der Kontextabhängigkeit ergeben.

Die Intentionalität von Computern als begriffliches Problem

Ein wesentliches Merkmal menschlicher Kognition ist die Fähigkeit, im Denken und Sprechen auf externe Gegenstände bezug zu nehmen. Ob Computer diese Fähigkeit ebenfalls besitzen oder jedenfalls besitzen könnten, ist seit geraumer Zeit heftig umstritten. Ausgangspunkt des Projekts sollte die (offenbar festgefahrene) Debatte zwischen Dennett und Searle um dessen berühmt-berüchtigtes Argument mit dem Chinesisch-Zimmer sein. In dieser Debatte werden die Schwierigkeiten besonders deutlich, die eine fruchtbare Erörterung dieser Frage als fast aussichtslos erscheinen lassen. Es besteht nämlich keine Übereinstimmung darüber, ob in diesem Zusammenhang gewisse wesentliche Voraussetzungen begrifflicher Art zulässig sind oder nicht.

Gibt es, wie von der einen Seite unterstellt wird, eine klare Grenzlinie zwischen (a) Systemen, IN DENEN Repräsentationen, die sich auf externe Sachverhalte beziehen, kausal aktiv sind, und (b) Systemen, DIE SICH in ihrem Wahrnehmen, Denken und Handeln auf externe Sachverhalte beziehen? Und ist jene "echte", "ursprüngliche", "intrinsische", "nicht-abgeleitete" Intentionalität vom Typ (b) nur bei einem System gegeben, das auch bewußtseinsfähig ist? Und kann als begriffliche Voraussetzung geltend gemacht werden, daß Computer – soweit sie nichts weiter sind als maschinelle Rechner – prinzipiell nicht bewußtseinsfähig sind?

In diesem Projekt ist herauszuarbeiten, welche Aussichten darauf bestehen, solche und

andere Fragen, die der festgefahrenen Diskussion um Computer-Intentionalität zugrundeliegen, in einer intersubjektiv verbindlichen Weise zu beantworten. Das Ziel sollte sein, einerseits deutlich zu machen, an welchen Stellen dieser Debatte es sich um echte Probleme und inhaltliche Streitpunkte handelt, die Gegenstand wissenschaftlicher Forschung sein können; andererseits aber auch darauf hinzuweisen, an welchen Stellen der Streit bloße Intuitionen über Begriffe wie "Bewußtsein" betrifft, für (und gegen) die es keine rationalen Argumente gibt.

2.2.4 Germanistische Linguistik (PD Dr. Becker)

2.2.4.1 Themenspezifische Vorarbeiten

Das Teilprojekt der Germanistischen Linguistik schließt sich eng an das der Phonetik und Sprachlichen Kommunikation an, da das Münchener Germanistische Institut nur zum Teil über die für dieses Projekt notwendige technische Ausstattung verfügt. Das seit einigen Jahren regelmäßig abgehaltene gemeinsame Oberseminar des germanistischen und des phonetischen Instituts zum Thema »Lautstrukturen des Deutschen« hat gezeigt, daß sich die beiden Disziplinen zwar weit auseinanderentwickelt und nur wenige aktuelle Fragestellungen gemein haben, sie aber doch in sehr fruchtbarer Weise zusammenarbeiten können. Aus der Zusammenarbeit in diesem Oberseminar sind bereits eine Dissertation (Restle, 1998) und eine Habilitationsschrift hervorgegangen (Becker, 1998, vgl. auch 1996a und 1996b), weitere Dissertationen und kleinere Publikationen sind im Entstehen. Andererseits sind schon früh germanistische Fragestellungen, wie die des artikulatorischen Korrelats des Silbenschnittgegensatzes im Deutschen (Vennemann, 1991), in die Forschung des IPSK eingegangen (Hoole et al. 1994).

Literaturhinweise

Becker, Th. (1996a): Zur Repräsentation der Vokallänge in der deutschen Standardsprache. *Zeitschrift für Sprachwissenschaft* 15. 3–21.

Becker, Th. (1996b): Die Aufhebung des Vokallängengegensatzes in unbetonten Silben der deutschen Standardsprache. *Deutsche Sprache* 24. 268–282.

Becker, Th. (1998): Das Vokalsystem der deutschen Standardsprache. Frankfurt am Main: Peter Lang (= *Arbeiten zur Sprachanalyse* 32, hg. v. Konrad Ehlich).

Hoole, Ph., Mooshammer, Chr. & Tillmann, H. G. (1994): Kinematic analysis of vowel production in German. In: *Proceedings of the International Congress of Spoken Language Processing (ICSLP)*, Yokohama 1994. Yokohama: Acoustical Society of Japan. Bd. I. 53–56.

Restle, D. (1998): Silbenschnitt – Quantität – Kopplung. Zur Geschichte, typologischen Einordnung und Repräsentation der Anschlußprosodie unter dem Blickwinkel einer Oszillationssilbentheorie. Diss. München.

Vennemann, Th. (1991) Syllable structure and syllable cut prosodies in modern Standard German. In: Pier Marco Bertinetto/Michael

Kenstowicz & Loporcaro, M. (Hgg.). Certamen Phonologicum II. Papers from the 1990 Cortona Phonology Meeting. Turin: Rosenberg & Sellier. 211–243.

2.2.4.2 Skizzen möglicher Dissertationsprojekte

Bei den zu verfolgenden Dissertationsprojekten geht es in erster Linie darum, Impulse, die sich aus dem genannten Oberseminar ergeben haben, für die Durchführung von germanistischen Promotionen nutzbar zu machen, bei denen es um eine Verbindung der germanistischen Theorien zum heute gesprochenen Deutsch und der technischen Verarbeitung von gesprochener Sprache geht.

Fast kein Sprecher der deutschen Standardsprache spricht die reine Norm des Aussprachewörterbuchs, sondern Abweichungen davon, die sich allerdings systematisieren lassen. Z. B. sind fast alle Varietäten der Standardsprache regional geprägt, wobei die Sprecher meist über mehrere Varietäten verfügen, die sich im Formalitätsgrad und im Grad der dialektalen Beeinflussung unterscheiden. Spracherkennung sollte mit diesen Abweichungen möglichst geschickt umgehen können, und die Berücksichtigung von Regionalismen und unterschiedlichen Formalitätsgraden kann die Akzeptanz synthetischer Sprache erhöhen. Außerdem seien die großen Fortschritte hervorgehoben, die sich aus dem Einsatz von intelligent programmierten Rechnern als digitalen, automatisch interviewenden Aufnahmestationen für die empirischen Untersuchungen im Bereich der Germanistischen Linguistik ergeben: Nicht nur läßt sich die Versuchsanordnung auf diese Weise sehr gut kontrollieren, auch die Daten werden in einer Form erhoben und gespeichert, die dann vom Rechner unmittelbar weiterverarbeitet werden kann. Mit den zu bearbeitenden Dissertationsthemen sollten seitens der Germanistik in erster Linie die stilistischen Besonderheiten des Sprechens mit technischen Informationssystemen geklärt werden, sowohl bei spontanem, als auch bei (durch spezielle Instruktion) geschultem Sprechen.

Regionale Varietäten der Standardausssprache

Welche Merkmale charakterisieren die einzelnen regionalen Varietäten der Standardausssprache? Sind alle diese Merkmale dialektbasiert? Lassen sie sich der Stärke nach ordnen, so daß sich implikative Skalen ergeben? (Jeder Sprecher der Standardsprache in Bayern, der in *klein* [g_0lo_~a_~_^]³ den Lenisplosiv [g_0] durch [k] ersetzt und die Nasalisierung der Vokale durch ein folgendes [n], der ersetzt auch das auffälligere und leichter zu kontrollierende Dialektmerkmal des Diphthongs [oa_^] durch [ai_^], so daß Formen wie [kloa_^n] nicht vorkommen, oder nur bei Sprechern, die den Dialekt zu imitieren versuchen; daher gilt: [oa_^] impliziert [g_0]). Wie korrelieren Dialekteinfluß und Formalitätsgrad? (In diesem Zusammenhang sind auch die Varietäten im Hinblick auf Lexikon, Syntax und Morphologie interessant.) Gibt es vom Formalitätsgrad abhängige oder regionale Unterschiede in Mimik und Gestik oder Unterschiede in Mimik und Gestik abhängig von der Bewertung der regionalen oder formalen Varietät des Gesprächspartners? Lassen sich aus alters- oder bildungsabhängigen Unterschieden im Dialekteinfluß und Formalitätsgrad Tendenzen des Sprachwandels in der Standardsprache der Gegenwart ableiten?

3 Wir verwenden hier das im Rahmen des SAM-Projektes der EU entwickelte Transkriptionssystem, das mit entsprechenden Erweiterungen in »Computer-coding the IPA: a proposed extension of SAMPA Summary version, not requiring an IPA character set« von John Wells beschrieben wird (<http://www.phon.ucl.ac.uk/home/sampa/x-sampa.htm>).

Adressatenorientiertes Sprechen mit dem Rechner

Was charakterisiert den situativen Kontext "Gespräch mit dem Rechner"? Ist es tatsächlich eine besondere Gesprächssituation (gibt es "Computer-Talk"?) oder läßt sie sich auf Besonderheiten der Mensch-zu-Mensch-Kommunikation reduzieren? Welchen Formalitätsgrad streben die Sprecher an, und drückt sich darin die Einstellung zu Rechnern aus? Gibt es hier geschlechtsspezifische Unterschiede? Solche Fragen lassen sich auf allen Beschreibungsebenen der Linguistik untersuchen (Syntax, Semantik, Pragmatik), insbesondere aber auch auf der phonologischen Ebene (Lautstruktur, Prosodie, Intonation) sowie der Ebene der nonverbalen Kommunikation.

Einrichtung des »talking-head«-Syntheseprogramms von Massaro für gesprochenes Deutsch

Dominic Massaro (UCSC) hat angeboten, sein für amerikanisches Englisch entwickeltes (sehr komplexes) Syntheseprogramm zur dreidimensionalen Generierung von sichtbaren Sprechbewegungen uns hier in München für systematische Untersuchungen zum Deutschen zur Verfügung zu stellen. Für einen Studierenden der Germanistik mit Nebenfach Informatik oder Phonetik, der sich für eine berufliche Tätigkeit im Bereich der Sprachtechnologie des Deutschen qualifizieren möchte, wäre es – speziell im Rahmen des Graduiertenkollegs – eine höchst reizvolle Aufgabe, den künstlich sprechenden Kopf von Massaro auf einen Sprecher des Deutschen zu adaptieren und die sich dabei ergebenden Unterschiede zum englisch-amerikanischen Original zugleich einer kontrastiven Analyse zu unterziehen, und zwar sowohl auf der segmentellen als auch auf der suprasegmentalen Ebene.

2.2.5 Mensch-Maschine-Kommunikation (Prof. Dr. Lang, PD Dr. Ruske)

2.2.5.1 Themenspezifische Vorarbeiten

Grundlegende eigene Vorarbeiten beziehen sich auf das Gebiet der automatischen Spracherkennung, des Trainings von Klassifikatoren und der automatischen Sprecheradaption. Die akustisch-phonetische Erkennung der gesprochenen Äußerung basiert auf den klassischen Hidden-Markov-Modellen (HMM) für Sprachlaute (Phoneme).

Bei der reinen Sprecheradaption standen zuerst die klassischen Transformationsverfahren im Vordergrund, die die spektralen Eigenschaften des Sprechers an das Erkennungssystem anpassen. Diese Verfahren können sehr effektiv eingesetzt werden. Darüberhinaus wurden verschiedene Methoden für die Ähnlichkeitsmessung zwischen den Verteilungen in den Zuständen aller HMMs untersucht, um die HMM-Parameter zu adaptieren. Es wurde hierbei davon ausgegangen, daß ähnliche Verteilungen bzw. ein ähnliches Lernverhalten in zwei Zuständen oft auf das Vorhandensein ähnlicher akustisch-phonetischer Ereignisse zurückzuführen ist. Bei Vorliegen eines bestimmten, aktuellen Zustands auf dem Viterbi-Pfad können auch alle diejenigen Zustände nachadaptiert werden, deren Verteilungen mit der aktuellen Verteilung verwandt sind.

Eine Reihe grundlegender Untersuchungen hatte zum Ziel, die Verarbeitung sowohl von Aussprachevarianten innerhalb von Wörtern als auch von Aussprachevarianten über

Wortgrenzen hinweg während des Erkennungsvorgangs zu ermöglichen. Unsere Arbeiten haben gezeigt, daß mögliche Abweichungen von der Standardtranskription stets regelhaft durch die davor und danach auftretenden Phoneme beschrieben werden können. Es wurde ein statistisches Verfahren entwickelt und implementiert, das anhand einer Sprachdatenbasis automatisch die grundlegenden Verschleifungen und zusätzlich auch sprecherspezifische Eigenheiten sowie dialektale Einflüsse erfaßt.

Eingehende Untersuchungen zum adaptiven Mensch–Maschine–Dialog wurden bereits im Rahmen eines intelligenten, adaptiven Tutorsystems durchgeführt. Diese Arbeiten zeigten neue Prinzipien zum maschinell unterstützten Dialog zwischen Mensch und Computer auf. Als Versuchsumgebung wurde dazu ein Intelligentes Tutorielles System (ITS) konzipiert und prototypisch realisiert. Anhand von Tests mit Versuchspersonen konnte demonstriert werden, wie sich durch dynamische Bewertung mithilfe von Lernzielen eine permanente Benutzereinschätzung durchführen läßt. Das Dialogverhalten und das Niveau der Tutorsitzung können damit automatisch an die individuellen Bedürfnisse des Benutzers angepaßt werden.

Literaturhinweise

Beham, M. (1995): Merkmalsextraktion und Regelgewinnung für die automatische Spracherkennung. Dissertation am Lehrstuhl für Datenverarbeitung, TU München, 1995.

Müller, J. & Stahl, H. (1998): Speech understanding and speech translation in various domains by maximum a-posteriori semantic decoding. Tagungsband EIS 98, La Laguna/Spanien, Band 2, S. 256–267.

Obermaier, A. (1995): Benutzereinschätzung in einem adaptiven Tutorsystem mit Hilfe von Lernzielen. In: Tagungsbd. ABIS'94–GI Workshop, Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen. Hrsg.: C.G. Thomas, GMD–Studien Nr. 244, Sankt Augustin, S. 63 – 68.

Obermaier, A. (1996): Modellierung eines adaptiven Mensch–Maschine–Dialoges und experimentelle Untersuchungen an einem Prototypen. Dissertation am Lehrstuhl für Mensch–Maschine–Kommunikation, TU München.

Pfau, T. & Ruske, G. (1998): Estimating the Speaking Rate by Vowel Detection. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Seattle/USA, 12. – 15. Mai 1998, II–945 – 948.

Pfau, T., Beham, M., Reichl, W. & Ruske, G. (1997): Creating Large Subword Units for Speech Recognition. Europ. Conf. on Speech Communication and Technology, Rhodos, 22.–25. 9. 1997, Vol. 3, 1191 – 1194.

Plannerer, B. & Ruske, G. (1993): A continuous speech recognition system using phonotactic constraints. EUROSPEECH, Berlin, 21.–23.9. 1993, 859 – 862.

Plannerer, B., Einsele, T., Beham, M. & Ruske, G. (1994): A continuous speech recognition system integrating additional knowledge sources in a data-driven beam search algorithm. Int. Conference on Spoken Language Processing ICSLP–94, Yokohama/Japan, 18.–22.

Sept. 1994, S01–5.1 – S01–5.4.

Plannerer, B. (1995): Erkennung fließender Sprache mit integrierten Suchmethoden. Dissertation am Lehrstuhl für Mensch–Maschine–Kommunikation, TU München.

Reichl, W. & Ruske, G. (1995): A hybrid RBF–HMM system for continuous speech recognition. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Detroit, 9.–12. Mai 1995, 3335 – 3338.

Reichl, W. & Ruske, G., Discriminative training for continuous speech recognition. Europ. Conf. on Speech Communication and Technology, Madrid, 18.–21. 9. 1995, 537 – 540.

Reichl, W., Caspary, P. & Ruske, G. (1994): A new model–discriminant training algorithm for hybrid NN–HMM systems. In: IEEE Proc. 1994 Int. Conf. Acoust. Speech and Signal Process., Adelaide/Austral., April 19–22, 1994, II–677 – II–680.

Reichl, W. (1996): Diskriminative Lernverfahren für die automatische Spracherkennung. Dissertation am Lehrstuhl für Mensch–Maschine–Kommunikation, TU München, 19.6.1996. Im Buchhandel: Verlag Dr. Shaker, Aachen.

Ruske, G. (1994a): Automatische Spracherkennung: Methoden der Klassifikation und Merkmalsextraktion. Zweite, erweiterte Auflage, Oldenbourg–Verlag, München Wien.

Ruske, G. (1994b): Methoden der Automatischen Spracherkennung. Informationstechnik und Technische Informatik (it + ti), Heft 2, Oldenbourg–Verlag, München Wien, 31 – 39.

Ruske, G., Falthausen, R. & Pfau, T. (1998): Extended linear discriminant analysis (ELDA) for speech recognition. Int. Conference on Spoken Language Processing ICSLP–98, Sydney, 30. Nov. –4. Dez. 1998, im Druck.

Schiel, F. (1993): Untersuchungen zur Sprecheradaptation in Systemen zur automatischen Spracherkennung mit Hilfe stochastischer Modellierung. Dissertation am Lehrstuhl für Datenverarbeitung, TU München.

Stahl, H., Müller, J. & Lang, M. (1996): An efficient top–down parsing algorithm for understanding speech by using stochastic syntactic and semantic models. ICASSP 96, Atlanta/USA, Vol. 1, S. 397–400.

Stahl, H., Müller, J. & Lang, M. (1997): Controlling Limited–Domain Applications by Probabilistic Semantic Decoding of Natural Speech. ICASSP 97, München, 1997, S. 1163–1166.

Wolfertstetter, F., Kipp A. & Ruske, G. (1994): Statistische Generierung von typischen Verschleifungsregeln für den Lexikonvergleich. In: Tagungsband KONVENS 1994, Wien, Sept. 28–30, 1994. Hrsg.: H. Trost, S. 397–406.

Wolfertstetter, F. & Ruske, G. (1994): Discriminative state–weighting in Hidden–Markov models. Int. Conference on Spoken Language Processing ICSLP–94, Yokohama/ Japan, 18.–22. Sept. 1994, S07–9.1 – S07–9.4.

Wolfertstetter, F. & Ruske, G. (1995): Structured Markov models for speech recognition. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Detroit, 9. –12. Mai 1995, 544 – 547.

Wolfertstetter, F. & Ruske, G. (1996): Discriminative training of stochastic Markov graphs for speech recognition. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Atlanta, 7. – 10. Mai 1996, 581 – 584.

2.2.5.2 Skizzen möglicher Dissertationsprojekte

Benutzermodellierung für gesprochene Eingaben

(Untertitel: „Untersuchungen zum adaptiven Verhalten eines informationsverarbeitenden Systems und seines Benutzers im sprachlichen Dialog)

Die zentrale Stellung in einem sprachlichen Informationssystem nimmt der Benutzer ein. Eine wirklich komfortable Nutzung eines solchen Systems erfordert eine weitreichende Einbeziehung des Benutzers, d.h. seiner Eigenschaften, seiner Gewohnheiten und Präferenzen. Das vorliegende Projekt hat daher zum Ziel, ein Benutzermodell aufzubauen, anhand dessen eine optimal angepaßte Handhabung für den Benutzer möglich wird. Das Benutzermodell enthält 3 Ebenen:

- a) die akustisch–phonetische Ebene (gesprochene Äußerungen),
- b) die lexikalische Ebene (Wortgebrauch),
- c) die Intentionsebene (Absicht).

Die Anpassung wird durch Strategien erreicht, die den Benutzer und das System besser aufeinander abstimmen. Bewußt soll diese Problematik von diesen beiden Seiten angegangen werden, denn beide Seiten können der optimalen Lösung entgegenkommen. Die Anpassungsstrategien umfassen daher 2 Ansätze:

- 1) das System paßt sich an den Benutzer an,
- 2) der Benutzer paßt sich an das System an.

Untersucht werden sollen diese Effekte am Beispiel eines Auskunftssystems, das mit gesprochener Sprache abgefragt wird.

Zu 1): Hier stehen die typischen Sprechereigenschaften im Vordergrund. Das System adaptiert sich an die akustisch–phonetischen Eigenschaften, an die Wortwahl und an die typischen Intentionen dieses Benutzers. Verfahren zur automatischen Spracherkennung und Sprecheradaptation sind bereits vorhanden. Untersucht werden sollen daher vor allem die Ebenen b) und c). Hier lassen sich sinnvoll Wizard–of–Oz Experimente einsetzen, wobei das Auskunftssystem simuliert wird.

Zu 2): Es ist sehr wichtig, auch die Möglichkeiten zu untersuchen, wie der Benutzer dem System entgegenkommen kann und will. Er kann z.B. durch Beispielsätze („Demos) etwas über den vorhandenen Wortschatz und die gewünschte Art des Sprechens erfahren. Das System kann ihn auf nicht verstandene Wörter aufmerksam machen und Alternativen anbieten („Meinen Sie etwa ..'). Das heißt, der Benutzer lernt in einem Klärungsdialog das System besser kennen. Hier sollen ebenfalls wieder Wizard–of–Oz Experimente zum

Einsatz kommen. Ein entsprechendes Usability Labor mit allen Einrichtungen ist vorhanden. Wissenschaftliche Fragestellungen sind die Akzeptanz des Benutzers und seine subjektive Beurteilung der Güte der Bedienbarkeit des Systems.

Diese Fragestellungen sind in diesem Zusammenhang größtenteils neu und von grundlegender Bedeutung für die Kooperation zwischen Mensch und Maschine.

2.2.6 Lehrstuhl für Informatik IX – Bildverstehen und Wissensbasierte Systeme (Prof. Dr. Radig, Dr. Meixner)

2.2.6.1 Themenspezifische Vorarbeiten

In München ist am Lehrstuhl von Professor Radig die Forschungsgruppe Bildverstehen angesiedelt. Durch die thematischen Schwerpunkte steht die Gruppe in kooperativer Verbindung mit der Forschungsgruppe Kognitive Systeme des Bayerischen Forschungszentrums für Wissensbasierte Systeme (FORWISS), das auf Initiative der bayerischen Staatsregierung 1988 als Universitäts-Institut an den drei Universitäten Erlangen–Nürnberg, München (TU) und Passau gegründet wurde. Die Arbeiten der Gruppe Kognitive Systeme sind auf digitale Bildverarbeitung spezialisiert und stehen unter der Leitung von Prof. Radig.

Aus der breiten Basis von Forschungsaktivitäten am Lehrstuhl bzw. der Arbeitsgruppe Kognitive Systeme seien hier nur zwei Beispiele genannt. Von 1994 bis 1996 arbeitete das FORWISS mit dem Projekt STABIL an der Entwicklung eines Systems zur videobasierten Detektion von Personen im 3D–Raum. Ziel dieses Systems war es, die Position von im Videobild erkennbaren Personen zu bestimmen.

Die mit dem System STABIL gesammelten Erfahrungen führten im Herbst 1996 zu der Entscheidung, ein von grundauf neues System mit einem strikt modellgetriebenen Ansatz zu entwickeln. Unter dem Namen STABIL++ wird seither eine C++–Klassenstruktur zur Repräsentation eines dreistufigen Modells entwickelt.

Die Merkmalsextraktion erfolgt mit dem am Lehrstuhl entwickelten Bildverarbeitungssystem HALCON, das die Extraktion von Bildmerkmalen auf der Basis von über 600 Bildverarbeitungsoperatoren ermöglicht.

Am Lehrstuhl wurden Bayessche Methoden bei der Erkennung von 3D–Objekten eingesetzt (Büsching 1998). Bei der Bewertung von Korrespondenzen zwischen Bildmerkmalen wurde dabei die a priori Wahrscheinlichkeit von Bildmerkmalen in Abhängigkeit von ihren Typen und Attributwerten berücksichtigt. Die Eignung von Bayesschen Netzen für Aufgaben im Bereich der Spracherkennung wird auch dadurch deutlich, daß ein enger Zusammenhang besteht zwischen Bayesschen Netzen, die temporale Abhängigkeiten zwischen Zufallsvariablen darstellen (dynamische probabilistische Netze, DPN), und Hidden–Markov–Modellen (HMM). HMM lassen sich danach als besonders einfache Form von DPN interpretieren. Andererseits haben sich HMM bereits bei der akustischen Spracherkennung bewährt.

Literaturhinweise

Azebajani, A., Wren, C. & Pentland, A. (1996): Real-Time 3-D Tracking of the Human Body. Technical Report 374, M.I.T. Media Laboratory, Perceptual Computing Section.

D. Büsching (1998): Verteilte Generierung von Objektmodellen. Als Dissertation eingereicht, Technische Universität München.

Bregler, (1997): Learning and Recognizing Human Dynamics in Video Sequences. In: Computer Vision and Pattern Recognition, pages 568–574. IEEE Computer Society Press.

Eckstein, W. & Steger, C. (1997): An Architecture for Computer Vision Application Development within the HORUS System. Journal of Electronic Imaging, 6(2), Apr. 1997.

Huggle, K. & Eckstein, W. (1995): Extraktion von Personen in Videobildern; in: G. Sagerer, S. Posch, F. Kummert (Hrsg.) Mustererkennung 1995, Deutsche Arbeitsgemeinschaft für Mustererkennung, Informatik aktuell, Springer-Verlag, 134–144.

Klupsch, M. (1996): Ein objektorientiertes Bildanalysesystem in Smalltalk in Informatik aktuell: Mustererkennung 1996 – 18. DAGM-Symposium, B. Jähne, P. Geißler, H. Haußecker, F. Hering (Hrsg.), 561–571, Heidelberg, 11.–13.9. 1996.

Lanser, S. & Zierl, Chr. (1996a): On the Use of Topological Constraints within Object Recognition Tasks in 13th International Conference on Pattern Recognition, Track A: Computer Vision, Vienna, Austria, pp. 580–584, IEEE Computer Society Press, 1996.

Lanser, S. & Zierl, Chr. (1996b): MORAL: Ein System zur videobasierten 3D-Objekterkennung im Kontext autonomer, mobiler Systeme in: Autonome Mobile Systeme 1996, G. Schmidt und F. Freyberger (Hrsg.), Informatik aktuell, S. 88 – 98, Springer-Verlag, 1996.

Lanser, S., Zierl, Chr., Munkelt, O. & Radig, B.: MORAL – A Vision-based Object Recognition System for Autonomous Mobile Systems in: 7th International Conference on Computer Analysis of Images and Patterns, Kiel, September 10–12, pp. 33–41, Springer-Verlag, 1997.

Lückenhaus, M. (1998): A Multi-Agent System for Parallelizing Image Analysis Tasks accepted for: International Conference on Intelligent Autonomous Systems 5 (IAS-5), Sapporo, Japan, June 1–4, 1998.

Munkelt, O. & Kirchner, H. (1996). STABIL: A System For Monitoring Persons in Image Sequences. In: Image and Video Processing IV, Vol. 2666 of SPIE Proceedings, pages 163–179. SPIE – The Intern. Soc. for Optical Engineering.

Steger, C. (1996): Extraction of Curved Lines from Images in 13th International Conference on Pattern Recognition, Track B: Pattern Recognition, Vienna, Austria, 25.8.–30.8.1996.

Vatikiotis-Bateson, E., Eigsti, I.-M., Yano, S. & Munhall, K.G. (1997): Perceiver Eye Motion During Audiovisual Speech Perception; in: Forschungsberichte des Instituts für

Phonetik und Sprachliche Kommunikation der LMU, Nr. 35, S. 39–60.

Vatikiotis–Bateson, E., et al. (1996): The Dynamics of Audivisual Behavior in Speech. In: Speechreading by Humans and Machines. D.G. Stork, M.E.Hennecke (Herausg.). NATO ASI Series, Series F, Vol. 150, Springer Verlag.

2.2.6.2 Skizze möglicher Dissertationsprojekte

Bei den hier skizzierten Projekten geht es um die Anwendung der am Lehrstuhl Radig entwickelten Methoden der Bildverarbeitung auf audiovisuell gegebene Sprachäußerungen, wobei die dynamische Analyse und das automatische Verfolgen der mit der Sprachäußerung verbundenen Gesichtsmimik und Körpergestik im Vordergrund des Interesses steht. Auf der rein technischen Seite besteht das zentrale Problem in einer schritthaltenden dynamischen Bildverarbeitung mit einer adaptionsfähigen Architektur. Die folgenden Projekte sollen in enger Kooperation mit der Abteilung für Artikulographie des Instituts für Phonetik und Sprachliche Kommunikation und ihrem Leiter, Dr. Phil Hoole, sowie Dr. Eric Vatikiotis–Bateson vom ATR in Kyoto realisiert werden.

Der prosodische Zusammenhang von Mimik, Gestik und akustischer Artikulation

Die automatisch zu erfassende Gesichtsmimik (und gegebenenfalls zu beobachtende Veränderungen in der Körperhaltung) des Sprechers sollen daraufhin untersucht werden, auf welche Weise und in welchem Maße sie mit der rein akustisch zu beobachtenden Prosodie der lautsprachlichen Äußerung korrelieren. Dabei werden die von Vatikiotis–Bateson entwickelten Techniken zur generalisierten Erfassung der Gesichtsmimik herangezogen. Die prosodische Kennzeichnung soll auf der Basis einer automatischen Segmentation (mit dem am Institut für Phonetik und Sprachliche Kommunikation entwickelten MAUS–System) sowie einer (ggf. auch am Glottogramm zu überprüfenden) automatischen F0–Extraktion erfolgen.

Analyse der Unterschiede von artikulationsbedingter und sprachfreier Gesichtsmimik

Vatikiotis–Bateson und Yehia haben mit ihrem am ATR in Kyoto entwickelten System gezeigt, daß sich ein Sprachsynthesesystem, das auf der Technologie der neuronalen Netze beruht, mit zunächst audiovisuellem Material (also akustisches Sprachsignal plus Gesichtsmimik) so trainieren läßt, daß es lernt, allein aus der Beobachtung der Gesichtsmimik unter Ausblendung der sichtbaren Mund– und Lippenbewegungen die zugrundeliegende lautsprachliche Äußerung in einer einigermaßen verständlichen Form zu regenerieren. Bei dem hier skizzierten Projekt sollte davon ausgegangen werden, daß die optisch erfassbare Gesichtsdynamik bei einer Sprachäußerung nicht vollständig auf die zugrundeliegende Sprachmotorik zurückgeführt werden kann. Es soll vielmehr herausgefunden werden, welche Komponenten der Gesichtsmimik von der rein artikulatorischen Form einer lautsprachlichen Äußerung abhängen und welche sich als davon unabhängig extrahieren lassen. Dies gilt z.B. für emotional gesteuerte Gesichtsveränderungen. Mit dieser Untersuchung soll ein Beitrag zur Klärung der Frage geleistet werden, inwieweit sich die bei spontanem Sprechen zu beobachtende Gesichtsdynamik als Überlagerung der Wirkung unabhängiger motorischer Steuerungsmechanismen erklären lassen (wie z.B. beim Sprechen mit gleichzeitigem Lächeln).

2.3 Studienprogramm

2.3.1 Generelle Überlegungen

Das Studienprogramm legt besonderen Wert darauf, daß Kollegiaten, die bereits profunde Kenntnisse und die herausragende Fähigkeit zu wissenschaftlicher Arbeit in einem der Fachgebiete durch eine Magister- oder Diplomarbeit nachgewiesen haben, ihre Kenntnisse in den anderen Fachgebieten so weit ausbauen, daß sie im Rahmen von anwendungsorientierten Forschungsprojekten den Erkenntnisstand eigenständig erweitern können. Dies gewährleistet ein spezifisch auf das Kolleg abgestimmtes Studienprogramm, welches das durch die Promotionsordnungen der beteiligten Universitäten vorgeschriebene Studienprogramm einerseits inhaltlich und formal voll abdeckt, andererseits aber wesentlich darüber hinausgeht. Bei der folgenden Darstellung der verschiedenen Studienphasen ist zu beachten, daß die Realisierung des Forschungsprojektes parallel erfolgt.

1. Integration und Harmonisierung (Monat 1–12 der Förderung):

Die Kollegiaten (insbesondere, wenn sie von Hochschulen kommen, die nicht an dem Kolleg beteiligt sind) erwerben die im Rahmen des Promotionsstudiums durch Prüfungsordnungen vorgeschriebenen Leistungsnachweise. Darüber hinaus vertiefen Sie ihre Kenntnisse in forschungsthemenspezifischen Bereichen. Dies wird durch die an den Instituten regulär angebotenen Vorlesungen und Seminare gewährleistet.

2. Vernetzung und Präsentation (Monat 6–24 der Förderung):

Die Kollegiaten besuchen Veranstaltungen der nicht betreuenden Institutionen und präsentieren ihre (Zwischen)ergebnisse regelmäßig in einem eigens für das Kolleg eingerichteten Oberseminar aller Kooperationspartner. In diesem Zusammenhang sind auch Klausurtagungen in der Venice International University geplant, an deren Gründung die Ludwig-Maximilians Universität München beteiligt war und deren erklärtes Ziel die Internationalisierung der wissenschaftlichen Ausbildung ist.

3. Transfer (Monat 18 – Ende der Förderung):

In dieser Phase präsentieren die Kollegiaten ihre Ergebnisse auf internationalen Konferenzen, Tagungen und Messen sowie auf Workshops, die vom Kolleg selbst veranstaltet werden. Im Rahmen von kurzfristigen Aufenthalten bei Partnerinstituten im In- und Ausland sowie in der Industrie bereiten sie ihre berufliche Zukunft vor. Die Institute unterstützen auch die Einreichung der abgeschlossenen Arbeiten für ausgeschriebene Wissenschaftspreise.

In allen Ausbildungsphasen wird besonderer Wert auf den (durch Lehrveranstaltungen oder individuelle Betreuung gestützten) Erwerb von Schlüsselqualifikationen (wie Fremdsprachenkenntnisse und Präsentationstechniken) gelegt. Wo dies nicht durch Dozenten der beteiligten Institute gewährleistet werden kann, werden hierfür externe Dozenten eingeladen, wie dies z.B. am Institut für Phonetik und Sprachliche Kommunikation schon seit Jahren durch die enge Kooperation mit der Firma Siemens AG üblich ist. Derartige – bereits bestehende – Industriekooperationen sollen auch zur Vermittlung von außeruniversitären Praktika genutzt werden.

Für jeden Promotionsstipendiaten wird es neben dem betreuenden Hochschullehrer in jedem Partnerinstitut einen bestimmten Ansprechpartner geben (Patentprinzip). Auf diese Weise

wird gewährleistet, daß für alle Fachbereiche jedem Kollegiaten zu jeder Zeit ein kompetenter und hilfsbereiter Ansprechpartner zur Verfügung steht. Positive Seiteneffekte auf die Interdisziplinarität des Projekts sind zu erwarten.

Abschließend ist zu bemerken, daß das Studienprogramm insbesondere die berufliche Qualifikation der Kollegiaten im universitären und außeruniversitären Kontext verbessern soll. Hierzu gehört auch die straffe Abwicklung der Promotionsprojekte. Nicht zuletzt um diese zu gewährleisten, wird auf ein starres Studienprogramm verzichtet, an dessen Stelle im Projekt die individuelle Betreuung der Stipendiaten tritt, die auch Intensivprogramme zum Erwerb fachlicher und fachübergreifender Qualifikationen sowie von Schlüsselqualifikationen einschließt.

In diesem Zusammenhang ist auch ein jährlicher Workshop auf dem Gebiet des Graduiertenkollegs unter Beteiligung führender Vertreter der Forschungs- und Entwicklungsabteilungen von Siemens, Daimler-Benz, IBM, Sony und Philips geplant.

2.3.2 Konkretisierung des Programms

Die interdisziplinäre Zielrichtung des Studienprogramms zum GK "Sprache, Mimik und Gestik im Kontext technischer Informationssysteme" kann sehr gut im Hinblick auf die von der Bundesregierung geförderten neuen Leitprojekte zur "Mensch-Technik-Interaktion in der Wissensgesellschaft" erläutert werden. Die im Graduiertenkolleg studierenden Doktoranden sollen durch die Organisation eines dieser Thematik gewidmeten Studienprogramms mit grundlegenden Kenntnissen und Fertigkeiten vertraut gemacht werden, für die es auf eine fruchtbare Verknüpfung von Wissen ankommt, wie es traditionellerweise in die hier beteiligten natur- und geisteswissenschaftlichen Einzelfächer verteilt ist. Dieses Wissen soll zum einen an die Kollegiaten in der Form von speziellen Kursen vermittelt, gleichzeitig aber auch von den Kollegiaten auf einer gemeinsamen Basis aktiv miterarbeitet werden; so soll dieses fachübergreifende Wissen einerseits in die jeweils eigenen, ganz individuell ausgerichteten Promotionsvorhaben eingebracht werden können, doch andererseits soll dieses Wissen – über die in den Dissertationen erzielten Ergebnisse – wiederum für die gemeinsame Basis fruchtbar werden. Diesen interdisziplinären Kreislauf von gemeinsamer Lehre und Forschung in Gang zu setzen, ist das eigentliche Ziel des die Arbeit an den Dissertationen begleitenden Studienprogramms. Es hat drei wesentliche Komponenten:

1. Ein wöchentliches gemeinsames "Graduiertenkolloquium"

Diese zwei, bei Bedarf auch dreistündige Semesterveranstaltung hat sich im SIL Kolleg sehr bewährt. An ihr nehmen nicht nur alle Kollegiaten, sondern auch alle Lehrenden teil. Natürlich finden auch die einzuladenden Gäste hier eine Plattform für ihre Anregungen und deren Diskussion. Hauptsächlich aber werden hier die Dissertationsprojekte und deren jeweiliger Entwicklungsstand vorgestellt und diskutiert. Schließlich soll hier aber auch die Struktur (d.h. Form und Inhalte) der unter 2. genannten Lehrveranstaltungen reflektiert und für eine zukünftige Präsentation verbessert werden.

2. Spezielle Lehrveranstaltungen der beteiligten Disziplinen

Jede beteiligte Disziplin bereitet sehr gezielt spezielle Beiträge zum gemeinsamen Studienprogramm vor. Hierbei kann es sich entweder sehr wohl um eine Auswahl von

normalen Seminaren handeln, die für diesen Zweck auch für die Kollegiaten mitangeboten, d.h. geöffnet werden. Oder es werden regelrechte Spezialkurse zu grundlegenden Themen in der Form von kompakten Blockveranstaltungen angeboten und durchgeführt. Was auf diesem Hintergrund die konkrete Gestaltung des Studienprogramms betrifft, so soll pro Semester wenigstens eine generelle Lehrveranstaltung aus einer der beteiligten Disziplinen für die Kollegiaten zum Generalthema erklärt werden, mit einer separaten Abschlussveranstaltung zu der sich der Lehrende exklusiv nur mit den Kollegiaten trifft. Und gegenüber diesen wöchentlich, in der Regel 2– bis 4–stündig zu belegenden Kursen, bietet es sich an, bestimmte grundlegende Themen durch kompakte Blockveranstaltungen abzudecken. Angefangen werden soll hier mit folgenden Themen, wobei die Reihenfolge noch offen ist:

- a) Becker: Lautstrukturen des Deutschen (Phonologie, Morphologie und ihre regionalen Ausprägungen),
- b) Ruske: Statistische Grundlagen von Sprachverarbeitungstechniken (HMM, neuronale Netze),
- c) Kemmerling: Nichtlinguistische Sprachtheorien,
- d) Tillmann: Allgemeinverständliche Einführung in die mathematischen Grundlagen der digitalen Signalverarbeitung (die z–Transformation und ihre Anwendungen bei linearen Systemen),
- e) Kegel: Psycholinguistische Grundlagen des Spracherwerbs,
- f) Radig: Methoden der Bildverarbeitung durch Automaten,
- g) Ehlich: Grundlagen der Analyse von nonverbaler Kommunikation (und später:) Die Maschine und das deiktische System.

Dies sind Beispiele. Die endgültige Form und Abfolge solcher Kompaktkurse soll im einzelnen noch genauer festgelegt werden. Was die aktive Komponente der Kollegiaten bei diesen Blockkursen betrifft, so denken wir daran, daß jeder Doktorand am Ende des Kurses ein Protokoll vorlegen muß, das bei der Anfertigung eines Skripts verwertet werden kann. Damit der beträchtliche Aufwand, der mit solchen einmaligen Spezialkursen verbunden ist, auf das Lehrdeputat des jeweiligen Kollegen angerechnet werden kann, müssen diese Lehrveranstaltungen natürlich auch den anderen Studierenden der TU bzw. LMU mitangeboten werden.

3. Interdisziplinäres Kolloquium "Beherrschung der Komplexität bei der Mensch–Maschine–Kommunikation".

Dieses im Begleitschreiben genannte Kolloquium, das von einer Forschergruppe aus Instituten der TU und LMU München während des Semesters in wechselnden Instituten durchgeführt wird, steht den Kollegiaten ebenfalls offen. Eine regelmäßige Teilnahme wird empfohlen. Auch ein Besuch von Veranstaltungen am Centrum für Informations– und Sprachverarbeitung (CIS) (einer zentralen universitären Einrichtung an der LMU, die nicht direkt am Graduiertenkolleg beteiligt ist) zu Natural Language Processing (NLP), bzw. Bereichen der Computerlinguistik wird empfohlen. Abschließend sei für die Gutachter noch einmal hervorgehoben, daß der Bezug des Studienprogramms zu den eingangs erwähnten Leitprojekten der Bundesregierung nicht nur ganz abstrakt – also eben allein schon durch die Thematik des GK – auf eine sehr grundsätzliche Weise gegeben ist, sondern jetzt auch durch die Tatsache, daß das IPSK mit seinem Vorschlag, eine intelligente MMDD–Aufnahmestation mit virtueller Wizard–of–Oz Technologie für die Versorgung der geplanten Leitprojekte mit den benötigten multimedialen Dialogdaten zu entwickeln, Erfolg

hatte, einen sehr konkreten unmittelbaren Kontakt zu den Leitprojekten selbst erhält. Die Entwicklung einer mobilen und intelligent programmierten Aufnahmestation zur gezielten Evokation multimedialer Dialogdaten dürfte nicht nur selbst zu einem aktuellen Thema des Studienprogramms werden, sondern zugleich auch eine Versorgung der einzelnen Dissertationsvorhaben mit den benötigten Daten sicherstellen, was zu einer deutlichen Entlastung führen dürfte, die nicht zuletzt eben auch dem Studienprogramm der einzelnen Kollegiaten sehr gut tun dürfte.

4. Für das Fach Germanistik wird empfohlen, daß die Kollegiaten (unter Anleitung der Betreuer) ein Hauptseminar für Studenten des Fachs Germanistische Linguistik, wie z.B. "Soziolinguistik und Sprachgeographie: Regionale Varietäten der Standardausprache und ihre Verwendung" anbieten; wenn es sich als sinnvoll erweist, sollten die Seminare zu einem späteren Zeitpunkt wiederholt werden. Außerdem wird der gemeinsame Besuch von Tagungen und Forschungseinrichtungen, sowie die Kooperation mit anderen Forschungseinrichtungen empfohlen.

2.4 Darstellung der bisherigen Bemühungen und Leistungen beteiligten Einrichtungen in der Graduiertenausbildung

Eine Graduiertenförderung auf dem hier anvisierten Gebiet existiert weder an der Technischen Universität München noch an der Ludwig–Maximilians–Universität München. Einzelaspekte des Themas wurden und werden an mehreren Instituten in Dissertationen, Diplom– und Magisterarbeiten untersucht. Hinweise zu abgeschlossenen und noch laufenden Arbeiten finden sich in den Beschreibungen der Vorarbeiten der beteiligten Institute (vgl. Kapitel 2.2).

2.5 Erwarteter Ertrag des Graduiertenkollegs und Auswirkungen auf die Entwicklung der betroffenen Fachgebiete und den Studienverlauf

Wie auch schon an anderer Stelle ausführlich beschrieben, soll das Graduiertenkolleg wesentliche Beiträge zur Grundlagenforschung wie auch zu anwendungsorientierten Forschungsgebieten im Bereich der gesprochensprachlichen Interaktion mit technischen Informationssystemen leisten.

In enger Verbindung mit dem Kolleg sollen diese thematischen Schwerpunkte auch für das reguläre Studienangebot der Institute weiter erschlossen werden. Dies gewährleistet letztlich auch die Nachhaltigkeit fachlich neuer Perspektiven in der Lehre. So werden einerseits den humanwissenschaftlichen und linguistischen Fächern anwendungsorientierte Fachaspekte erschlossen, was letztlich durch die Möglichkeit der empirischen Verifikation auch positive Rückwirkungen auf die Theoriebildung einschließt, andererseits wird die Vernetzung der ingenieurwissenschaftlichen Fächer mit den im Bereich der Sprachverarbeitung für sie äußerst wichtigen sprachwissenschaftlichen Fächern institutionalisiert und damit vertieft. So ist es auch denkbar, daß sich im Anschluß an das Kolleg der bereits in der Präambel angesprochene Mangel einer universitären Ausbildung zum Thema des Graduiertenkollegs durch Einrichtung eines (internationalen) Aufbaustudienganges realisieren läßt, für den die Erfahrungen im Rahmen des Graduiertenkollegs eine wesentliche Voraussetzung wären.

3. Strukturelle und organisatorische Voraussetzungen

3.1 Stellung des Graduiertenkollegs zu anderen wissenschaftlichen Einrichtungen

Aus den Darstellungen möglicher Promotionsthemen wird bereits deutlich, daß einige in enger Kooperation mit internationalen Partnern geplant und durchgeführt werden sollen. Dies ist neben der uneingeschränkten Kooperation der beteiligten Institute einschließlich besonders enger Partner (wie z.B. BAS und FORWISS) für das Kolleg von ganz besonderer Bedeutung.

Da das Kolleg sich auf die gesprochene Sprache (einschließlich ihrer nonverbalen Komponenten) konzentrieren wird, ist das Centrum für Informations- und Sprachverarbeitung, das seinen Schwerpunkt eher in der Verarbeitung der natürlichen (nicht-gesprochenen) Sprache sieht, nicht beteiligt.⁴

Bereits jetzt stehen dem Kolleg u.a. mit Prof. Dr. Günthner (CIS), Prof. Dr. Ehlich (als Inhaber des Lehrstuhls für Deutsch als Fremdsprache insbesondere für die Interaktion bestimmter Anwendergruppen mit technischen Informationssystemen) und Prof. Dr. Vennemann (als Inhaber des Lehrstuhls für Germanistische und Theoretische Linguistik) international renommierte Ansprechpartner beratend zur Seite.

Erkenntnisse, die im Rahmen des Graduiertenkollegs »Sprache, Information, Logik«, das ebenfalls an der Ludwig-Maximilians-Universität angesiedelt ist, gesammelt werden konnten, werden direkt in das Kolleg eingehen und damit eine anwendungsspezifische Umsetzung erfahren.

3.2 Kompatibilität des Graduiertenkollegs mit der bestehenden Studienstruktur der betreffenden Hochschule

Die beteiligten Institutionen sehen keinerlei Schwierigkeiten, das geplante Graduiertenkolleg innerhalb der bestehenden Promotions-, Prüfungs- und Studienordnungen der betroffenen Universitäten durchzuführen. Da nur eingeschriebene (LMU) bzw. in die Promotionsrolle eingetragene (TU München) Stipendiaten die Voraussetzungen für eine Förderung erbringen können, sind keinerlei Sonderregelungen notwendig.

3.3 Zugangsvoraussetzungen

Die Kollegiaten werden im Einklang mit den Promotionsordnungen der beteiligten Universitäten ausgewählt, wobei jeweils die Promotionsordnung des betreuenden Institutes bzw. Faches ausschlaggebend ist. Darüberhinaus finden folgende Merkmale von Bewerbern besondere Beachtung:

⁴ Sollte sich jedoch nach Ablauf der ersten Antragsphase eine diesbezügliche Ergänzung der Themenstellungen als wünschenswert erweisen, so ist dies ebenso möglich wie eine Ergänzung durch Institute der Psychologie, Deutsch als Fremdsprache, Didaktik oder Ethik. Wie schon an anderer Stelle bemerkt, sind jedoch vor der Bearbeitung von Themen der Technologiefolgenabschätzung oder der Zugänglichkeit der technischen Informationssysteme für breite Bevölkerungsschichten solche zu bearbeiten, die sich mit der technischen und ergonomischen Realisierbarkeit auseinandersetzen.

- besondere fachliche Leistungen (Prädiaktsexamen, Auszeichnungen, Preise)
- Auslands- und Industrieerfahrung (z.B. durch Praktika)
- Abschluß des grundständigen Studiums in der Regelstudienzeit
- einschlägige Publikationen
- EDV-Kenntnisse⁵
- Fremdsprachenkenntnisse (insbes. Englisch und andere Amtssprachen der EU sowie Japanisch)

Nach Bewilligung der Mittel werden die zu vergebenden Stipendien öffentlich ausgeschrieben. Auf der Basis ausführlicher Bewerbungsunterlagen einschließlich einer Skizze der geplanten Dissertation entscheiden dann die betreuenden Dozenten aller beteiligten Institute über die Eignung eines Bewerbers für das Kolleg. Die endgültige Auswahl wird nach einem Auswahlgespräch getroffen.

3.4 Erfolgskontrollen – Art der Abschlußprüfungen

Durch die im Rahmen des Promotionsstudiums zu erwerbenden Leistungsnachweise ist über die Auswahl hinaus eine Kontrolle des Studienerfolgs möglich. Von den Stipendiaten werden dabei mindestens gute, in der Regel sehr gute, Studienleistungen erwartet.

Der Erfolg der wissenschaftlichen Arbeiten wird durch regelmäßige Gespräche mit dem Betreuer und anderen Mitarbeitern der beteiligten Institute festgestellt. Wesentliche Bedeutung kommt dabei den Präsentationen im Rahmen des einzurichtenden Oberseminars zu. Wichtig ist, daß die Kollegiaten auch durch Selbstevaluation die Qualität ihrer Arbeit einzuschätzen lernen und regelmäßig kontrollieren. Hierzu können die Kontakte mit der Industrie und den einzuladenden Gastwissenschaftlern beitragen.

Die Abschlußprüfungen richten sich nach den Promotionsordnungen der betreuenden Institute bzw. Fächer, wobei für das Kolleg neben den Prüfungsleistungen insbesondere der weitere berufliche und wissenschaftliche Werdegang der Kollegiaten erfolgsbestimmend ist.

3.5 Von Seiten der Hochschule eingebrachte Grundausrüstung

Die an dem Graduiertenkolleg beteiligten Institute verfügen in jeder Hinsicht über die für ein solch anspruchsvolles Projekt notwendige Infrastruktur:

- Die Dozenten sind durch Publikationen, Drittmittelprojekte und Industriekooperationen auf dem Themengebiet des Graduiertenkollegs ausgewiesen.⁶
- Die beteiligten Institutionen verfügen über entsprechend ausgestattete, vernetzte DV-Anlagen.
- Die Lehrveranstaltungen der Institute werden den Graduierten offenstehen und diesen somit profunde themenspezifische Studien ermöglichen.

Die Ausstattung der Bibliotheken enthält alle wesentlichen der themenspezifisch relevanten Bücher und Zeitschriften.

⁵ Dies gilt für die Studierenden der Humanwissenschaften nur eingeschränkt, da diese ja gerade durch das Kolleg die Möglichkeit erhalten sollen, sich diese Kenntnisse zu erschließen.

⁶ Hier ist auch darauf hinzuweisen, daß sich nicht direkt am Kolleg beteiligte Wissenschaftler der beteiligten Institute in das Kolleg einbringen werden, wann immer dies für das Kolleg als solches oder für einzelnen Kollegiaten gewinnversprechend ist.

Hinweis:

Aus stilistischen Gründen wird nicht stets auf beide Geschlechter bezug genommen. Es sind jedoch – wo nicht explizit auf bestimmte Personen oder Personengruppen referiert wird – stets Angehörige beider Geschlechter gemeint. So wird die Aufnahme von Frauen in den Kreis der Kollegiaten ausdrücklich gewünscht und gefördert.

München, den 15.07.1999

Prof. Dr. H.G. Tillmann

Sprecher des Graduiertenkollegs
»Sprache, Mimik und Gestik im Kontext technischer Informationssysteme«

Vorstand des Instituts für Phonetik und Sprachliche Kommunikation