

Proseminar, wozu?

Verarbeitung gesprochener Sprache

Proseminar

Wolfgang Menzel

menzel@informatik.uni-hamburg.de

Arbeitsbereich Natürlichsprachliche Systeme

Fachbereich Informatik

Universität Hamburg



Proseminar Verarbeitung gesprochener Sprache 1

“Proseminare dienen zur Einführung
in das wissenschaftliche Arbeiten im Grundstudium.”

Studienordnung Diplom Informatik



Proseminar Verarbeitung gesprochener Sprache 2

Proseminar, wozu?

- Schlüsselqualifikationen erwerben
 - Recherchieren
 - Strukturieren
 - Präsentieren
 - Moderieren
- Wechselwirkungen zwischen Informatiksystemen und ihrem Einsatzfeld
 - Verknüpfung von fachinhaltlichen Schwerpunkten mit Wirkungs- und gesellschaftlichen Fragestellungen
 - Technische Systeme im Kontext betrachten

Studienordnung BSc Informatik



Proseminar Verarbeitung gesprochener Sprache 3

Proseminar, wozu?

primäre Fragestellungen:

- Themenanalyse
- Themenaufbereitung, Systematisierung
- Ideenfindung
- Literaturarbeit: Recherche, Selektion, Auswertung
- Systemevaluation
- Präsentation
- Diskussion



Proseminar Verarbeitung gesprochener Sprache 4

Proseminar, wozu?

sekundäre Fragestellungen:

- Problemlösung
- Entwurf
- Realisierung



Proseminar Verarbeitung gesprochener Sprache 5

A. Sprachverarbeitung, wozu?

1. Gesprochene Sprache

- Welche Informationen enthält ein Sprachsignal?
- Wie werden diese Informationen im Signal kodiert?
- Welche Bestandteile sind für technische Anwendungen interessant?



Proseminar Verarbeitung gesprochener Sprache 7

Verarbeitung gesprochener Sprache

A. Sprachverarbeitung, wozu?

- Analyse des Phänomenbereichs
 - Brainstorming: Potenzielle Anwendungsfelder
 - Analyse realer Anwendungsfälle
 - Diskussion
- ## B. Sprachverarbeitung, wie?
- Seminarvorträge zu Einzelthemen
- ## C. Sprachverarbeitung: Historische Wurzeln
- Exkursion: Historische Sammlung des Phonetischen Instituts



Proseminar Verarbeitung gesprochener Sprache 6

A. Sprachverarbeitung, wozu?

2. Sprachverarbeitung: Einsatzmöglichkeiten und Leistungsparameter

- Brainstorming: Anwendungsszenarien
- Was sind Störfaktoren? Wie lässt sich ihr Einfluss verringern?

Literatur: Tageszeitungen, populärwissenschaftliche Zeitschriften und Bücher



Proseminar Verarbeitung gesprochener Sprache 8

A. Sprachverarbeitung, wozu?

3. Anwendungsbereich I: Diktiersysteme
4. Anwendungsbereich II: Dialogsysteme
5. Anwendungsbereich III: Sprachlehrsysteme
 - Evaluation
 - Erfahrungsbericht
 - Bewertung



B. Sprachverarbeitung, wie?

7. symbolische Signalbeschreibung
 - Beschreibungsebenen für gesprochene Sprache: Signal, Phonetik, Phonologie, Morphologie, Prosodie, ...
 - Verhältnis zur geschriebenen Sprache
 - Beschreibungseinheiten: akustische / artikulatorische Merkmale, Phone, Allophone, Phoneme, Silben, Morphe, Wörter, ...
 - Visualisierung von Sprachsignalen

Literatur: Fellbaum (1984), Holmes (1991), jedes beliebige Handbuch der Sprachwissenschaft

Schwierigkeitsgrad: leicht



A. Sprachverarbeitung, wozu?

6. Gesprochene Sprache, die alternative Mensch-Maschine-Schnittstelle?
 - Diskussion: Pro und Kontra



B. Sprachverarbeitung, wie?

8. parametrische Signalbeschreibung
 - Maximalwertdetektion, Energie
 - Autokorrelation, Nulldurchgangsdichte
 - Spektrum, Fourier-Transformation, Spektrogramm, Formant
 - Filter, Filterbänke
 - Cepstrum
 - Visualisierung von Signalparametern

Literatur: Vary et al. (1998), Schafer und Rabiner (1975), Eppinger und Herter (1993), Euler (2006)

Schwierigkeitsgrad: schwer



B. Sprachverarbeitung, wie?

9. Sprachsynthese

- Historische Vorläufer
- Formantsynthese
- Diphonsynthese
- PSOLA
- Korpus-basierte Synthese

Literatur: Eppinger und Herter (1993), Vary et al. (1998), Hess (1996), Portele (1996), Balestri et al. (1999), Stöber et al. (1999)

Schwierigkeitsgrad: mittel



B. Sprachverarbeitung, wie?

10. Worterkennung I: Linearer Mustervergleich

- Abstandsmessung
- Dynamic Time Warping

Literatur: Vaissière (1985), Silverman und Morgan (1990), Furui (1989), Holmes (1991), Eppinger und Herter (1993), Ruske (1994), Euler (2006)

Schwierigkeitsgrad: schwer



B. Sprachverarbeitung, wie?

11. Worterkennung II: Hidden-Markov-Modelle

- Markov-Modelle
- Hidden-Markov-Modelle
- Vektorquantisierung
- Training

Literatur: Rabiner und Juang (1986), Eppinger und Herter (1993), Picone (1990), Jurafsky und Martin (2000), Wendemuth (2004), Fink (2003) Euler (2006)

Schwierigkeitsgrad: sehr schwer



B. Sprachverarbeitung, wie?

12. Sprachmodellierung

- reguläre Grammatiken
- Bigramme, Trigramme
- stochastische Grammatiken
- Perplexität

Literatur: Jelinek (1976), Jelinek (1990), Ney (1992), Lee (1999), Jelinek (1998), Charniak (1993), Jurafsky und Martin (2000), Euler (2006)

Schwierigkeitsgrad: schwer



B. Sprachverarbeitung, wie?

13. Prosodie in der Spracherkennung

- Betonung
- Phrasengrenzen
- Satzmodus

Literatur: Vaissière (1988), Nöth (1989), Nöth et al. (1997), Kompe (1997)

Schwierigkeitsgrad: mittel



Proseminar Verarbeitung gesprochener Sprache 17

B. Sprachverarbeitung, wie?

15. VoiceXML

- XML wozu?
- Grundidee, Systemarchitektur
- Standardisierungsbemühungen
- Produkte

Literatur: WWW

Schwierigkeitsgrad: leicht



Proseminar Verarbeitung gesprochener Sprache 19

B. Sprachverarbeitung, wie?

14. Dialogmodellierung

- Anwendungen
- Dialogtypen
- Formalismen

Literatur: Bernsen et al. (1998), Smith und Hipp (1994), Pieraccini et al. (1997), Harris (2005), Dahl (2004), citetmctear:2004

Schwierigkeitsgrad: mittel



Proseminar Verarbeitung gesprochener Sprache 18

C. Sprachverarbeitung: Historische Wurzeln

16. Exkursion

- Schausammlung des Phonetischen Instituts

Literatur: Grieger (1989)



Proseminar Verarbeitung gesprochener Sprache 20

Ablaufplan

23.10.2006	Einführung	11.12.2006	Vorträge 3
30.10.2006	Diskussion: Aufg.anal.	18.12.2006	Vorträge 4
6.11.2006	Brainstorming: Anwend.	9.1.2007	Vorträge 5
13.11.2006	Erfahrungsberichte 1	16.1.2007	Vorträge 6
20.11.2006	Erfahrungsberichte 2	23.1.2007	Vorträge 7
27.11.2006	Vorträge 1	30.1.2007	Vorträge 8
4.12.2006	Vorträge 2	6.2.2007	Abschlussdiskussion

