

# Proseminar

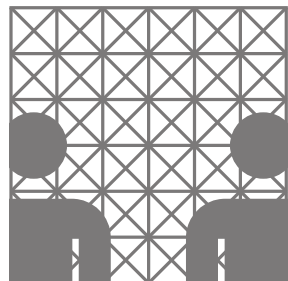
## Verarbeitung gesprochener Sprache

Timo Baumann

[baumann@informatik.uni-hamburg.de](mailto:baumann@informatik.uni-hamburg.de)



Universität Hamburg,  
Department of Informatics



# Heute

1. Vorstellung
2. Proseminar – was ist das, was soll das?
3. Aufbau des Seminars, Formalia
4. Brainstorming: Verarbeitung gesprochener Sprache

# Wer bin ich?

- Timo Baumann (baumann@inf..., F-406)
  - 2001-2007: Studium in Hamburg (Informatik, Phonetik, allerlei)
    - Auslandsaufenthalte –  
**integriert Auslandsaufenthalte in euer Studium!!**
  - 2007-2013: Promotion in Computerlinguistik (Potsdam, Stockholm Bielefeld, Hamburg)  
Thema: Inkrementelle Sprachdialogverarbeitung
  - seit 2011: Mitarbeiter/PostDoc in Hamburg  
im Arbeitsbereich Natürlichsprachliche Systeme
  - 2017-2018: Systems Scientist Carnegie Mellon University  
in Pittsburgh

wer seid ihr?

Was ist der Unterschied zwischen Schule und Universität?

Was macht Wissenschaft und Forschung aus?

# Aufbau (je)des (Pro-)Seminars

**[nats-www.informatik.uni-hamburg.de/VGS20/WebHome](http://nats-www.informatik.uni-hamburg.de/VGS20/WebHome)**

- Fachvortragstermine:
  - durch euch (fünf Stück)
  - durch mich (1-2 Stück)
- wissenschaftliche Methodenkompetenz:
  - Durchdringung und Gliederung (der Sprachverarbeitung)
  - Präsentation, Diskussion
  - Erstellen wiss. Arbeiten
  - Technikfolgenabschätzung
- Befähigung zur eigenständigen wissenschaftlichen Arbeit und zum aktiven Lernen

Brainstorming Anwendungsbereiche:  
Wo steckt überall Sprachtechnologie?





Proseminar  
Verarbeitung  
Gesprochener Sprache



Proseminar **Verarbeitung  
Gesprochener Sprache**  
SoSe 2015 Mi 10-12  
Kontakt: baumann@inf...



ALWAYS READY, CONNECTED,  
AND FAST. **JUST ASK.**

- Information, music, news, weather, and more—instantly
- Controlled by your voice for hands-free convenience
- Voice recognition hears you from across the room
- Connected to the cloud so it's always getting smarter

Proseminar **Verarbeitung Gesprochener Sprache**  
SoSe 2015 Mi 10-12

Anmeldung via Stine

Kontakt: baumann@inf...

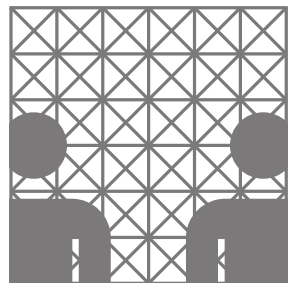


Kleingruppenarbeit (3-4 Personen):  
Strukturieren und ergänzen Sie  
die gesammelten Beispiele!



Vielen Dank.

[baumann@informatik.uni-hamburg.de](mailto:baumann@informatik.uni-hamburg.de)



# Further Reading

- Speech Signal Representation:
  - P. Taylor (2009): *Text-to-Speech Synthesis*. Cambridge Univ Press. ISBN: 978-0521899277. InfBib: A TAY 43070
  - D. Jurafsky & J. Martin (2009): *Speech and Language Processing*. Pearson International. InfBib: A JUR 4204x
- Phonetics:
  - M. Pétursson & J. Neppert (1996): *Elementarbuch der Phonetik*. Buske.
  - J. Neppert (1999): *Elemente einer akustischen Phonetik*. Buske.
- Phonology/Phonotactics/Phonological Systems:
  - E. Ternes (1999): *Einführung in die Phonologie*. Wiss. Buchgesellschaft. ISBN: 978-3534138708.

# Further Reading

- Speech Recognition in General:
  - D. Jurafsky & J. Martin (2009): *Speech and Language Processing*. Pearson International. InfBib: A JUR 4204x
- Token-Pass Algorithm:
  - Young, Russel, Thornton (1989): “Token Passing: A Simple Conceptual Model for Connected Speech Recognition Systems”, *Tech.Rep. CUED/F-INFENG/TR*, Cambridge University.
- The Sphinx-4 Speech Recognizer:
  - Walker et al. (2004): “Sphinx-4: A Flexible Open Source Framework for Speech Recognition”, *Tech.Rep. SMLI TR2004-0811*, Sun Microsystems.

# Notizen

# Lernziele

- Speech is the continuous signal of a symbolic language
- Speech is ambiguous on multiple layers of analysis
  - elements of speech
  - quasistationary signal, Phone/Laute, Phoneme, Allophone
- Speech recognition is the search for a most probable interpretation
  - allophone model, pronunciation model, language model (i.e., word sequence model)