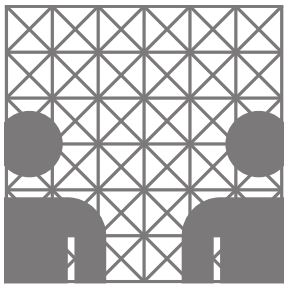
Specialization Module

Speech Technology

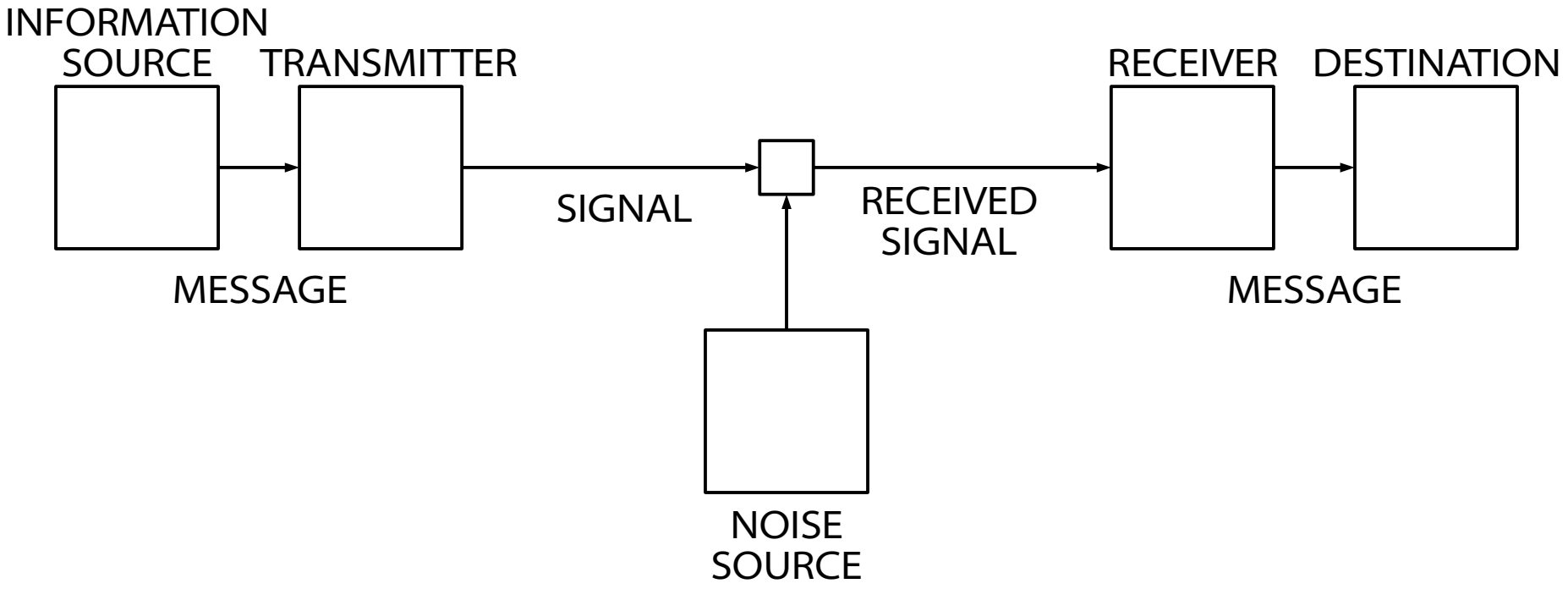
Timo Baumann
baumann@informatik.uni-hamburg.de

UNIVERSITÄT HAMBURG, DEPARTMENT OF INFORMATICS
NATURAL LANGUAGE SYSTEMS GROUP



Speech as a Communication System

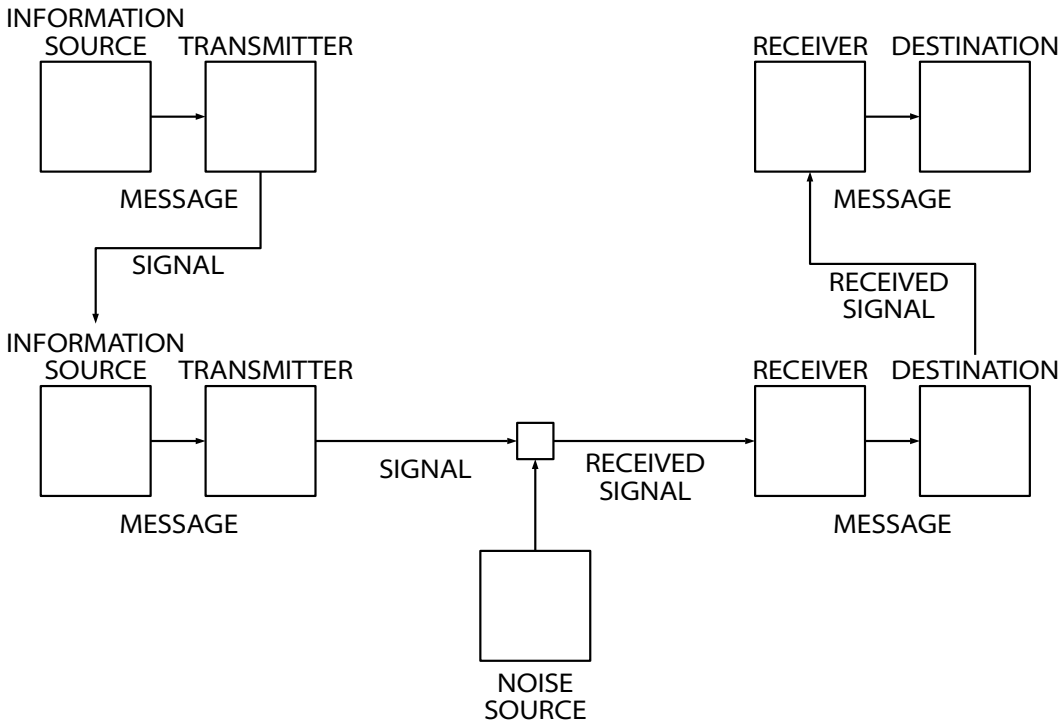
A technical view at communication



- Shannon-Weaver (1949) model of communication
  - requirements: model noise source, add redundancy
  - assumptions: corruption of messages is detected, transmitter and receiver are matched

Stratification of the Language System

- communication is rarely direct
- e.g. telephoning:



- the higher-level signal becomes a lower-level message
- layering applies the foundational scientific principle of analysis (Descartes 1637; repr. 1824)

Conventional division of linguistics into subdisciplines

area	objective
pragmatics	the study of meaning in context
semantics	the study of meaning
syntax	the study of sentence structure
lexicology	the study of words
morphology	the study of forming words
phonology	the study of a language's sound system
phonetics	the study of speech sounds

- is this conventional division correct?
  That is, does it accord to the best possible analysis?
- what are the interfaces between the areas?
- is everything covered by these areas?

Systems theory research on linguistics

- Serébrennikov et al. (1975) give a detailed account of why the aforementioned layers (strata) are reasonable
  - however, often strong interdependence between layers and the layers' elements (e.g., words are represented by phonemes *in a complex way*, e.g. based on the words' context)
- Levelt (1989): psycholinguistic proof for layering
  - (for speaking: conceptualization, formulation, articulation)

- we'll often resort to simple interfaces between layers, when we build technical systems
- what's better: many simple modules, or fewer complex modules? What are the dimensions of „better“?

Chain model of Communication

find message that describes idea

determine structure to convey meaning

sequentialize structure to word stream

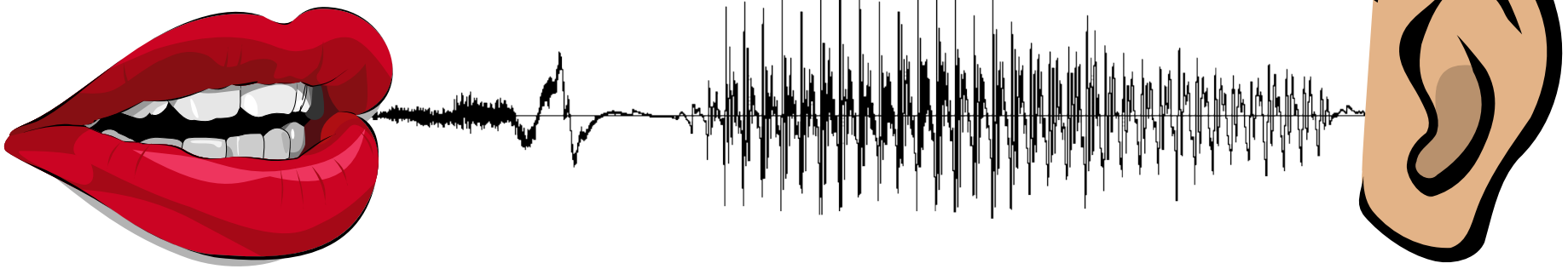
represent words through sounds

recover idea described by message

determine meaning of structure

recover structure of sequence

recombine souds to words



Chain model of Communication

find message that describes idea

determine structure to convey meaning

sequentialize structure to word stream

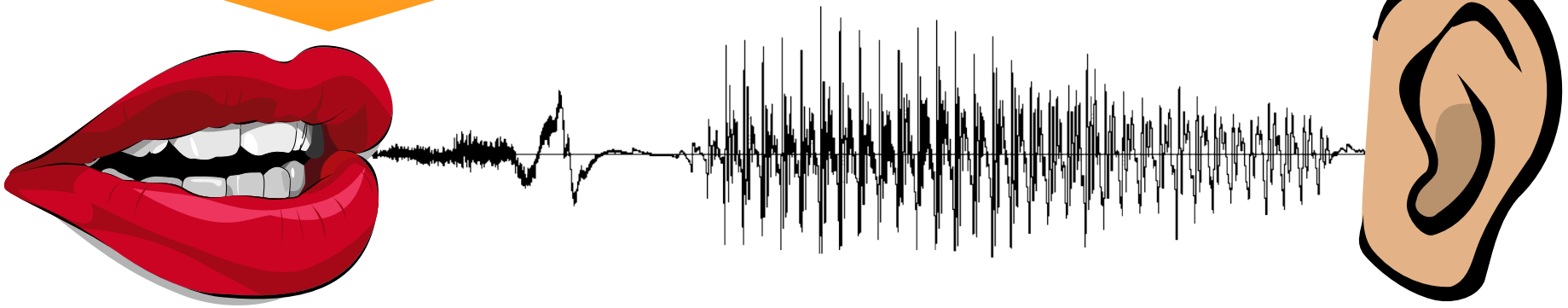
represent words through sounds

recover idea described by message

determine meaning of structure

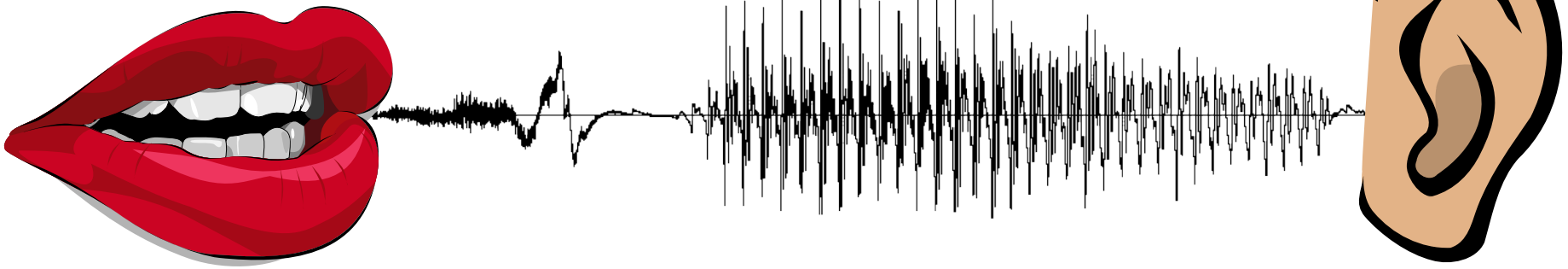
recover structure of sequence

recombine souds to words



Chain model of Communication

find message that describes idea	recover idea described by message
determine structure to convey meaning	determine meaning of structure
sequentialize structure to word stream	recover structure of sequence
represent words through sounds	recombine souds to words



A worked example

- A: <I want the soup to be saltier>

  A: <There is a saltshaker on the other end of the table>

  A: <Someone is sitting between me and the saltshaker>
- A: <I could ask her to hand it to me>
- A: „Could you pass the saltshaker, please?“
- channel: [kUd ju: pa:s T@ sAltSEk@6 pli:z]
- B receives: „Could you pass the saltshaker, please?“
- B: <Oh, he needs the salt>
- B: <I'll hand it to him>
- B: hands the saltshaker to A.

Small group exercise:

Develop and sketch out ways of how
the worked example can go wrong.
Describe in what ways it's going wrong.
What parts of the system are failing?

(Mis-)Alignment within layers

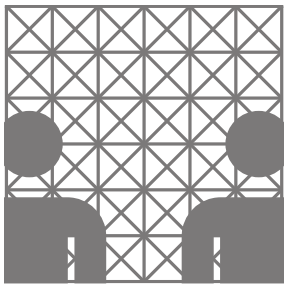
- e.g. you say „Please hand me the salt.“
- what's the implication if I respond „Here's the saltshaker.“

Thank you.

baumann@informatik.uni-hamburg.de

https://nats-www.informatik.uni-hamburg.de/SLP16

UNIVERSITÄT HAMBURG, DEPARTMENT OF INFORMATICS
NATURAL LANGUAGE SYSTEMS GROUP



Further Reading

- Noisy-channel model, chain model of communication:
  - Shannon and Weaver (1969): *The Mathematical Theory of Communication*, paperback edition, reprinted from the Bell System Technical Journal, July and October 1948. The University of Illinois Press. InfBib: T SHA 148 (also in StaBi and on the Web).
  - M. Pétursson & J. Neppert (1996): Elementarbuch der Phonetik. Buske. StaBi: F Ling 062/6.
- Introduction to (German) Linguistics:
  - Grewendorf, Hamm and Sternefeld (1989): *Sprachliches Wissen. Eine Einführung in moderne Theorien der grammatischen Beschreibung*, Suhrkamp. InfBib: A GRE 49762.
  - ... or any other introduction to linguistics; or the relevant chapter in: Jurafsky and Martin (2009): *Speech and Language Processing*. Pearson International. InfBib: A JUR 4204x.
- Systems theoretic views on complex systems in general and on language in particular:
  - Serébrennikov et al. (1975): *Allgemeine Sprachwissenschaft* (translated and edited by Zikmund and Feudel), Volume 2. Akademie-Verlag Berlin. StaBi: A 1974/563: 2.
  - Bertalanffy (1972): „The History and Status of General Systems Theory". In: *The Academy of Management Journal* 15(4), pp. 407-426. via Google Scholar.
  - Larsen-Freeman and Cameron (2008): *Complex Systems and Applied Linguistics*, Oxford University Press. StaBi: A 2009 / 7836.

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

- ~100 minutes
- should prosody be in the slide on linguistic subdisciplines?
- further parts (

Desired Learning Outcomes

- students understand the technical model of communication and its limitations
- students understand Descartes' scientific principle of analysis which leads to stratification as a principle for describing communication systems
- students know the linguistic subdisciplines that accord to stratification and that don't fit well (e.g. prosody)
- students understand the problem of modularization and can discuss advantages/disadvantages of particular models
- students are aware of the many ways that communication can break down (or have smaller problems) and are able to categorize problems and discuss such a categorization