

Äquivokationen

In der Spracherkennung

Michael Baumann
532 – 0225336
michael.baumann@student.tuwien.ac.at

Seminar (mit Bachelorarbeit)

Inhalt

- Einführung
 - Äquivokation, Ambiguität, Prosodie
- Signale
 - Beispiele: Schüttelreim, Homograph und Homophon
- Spracherkennungssystem
 - ASR, Schematischer Aufbau
- Methoden
 - HMM, Neuronale Netze
- Anwendung

Äquivokationen

- Aus der Philosophie:
 - Als Äquivokation bezeichnet man eine Aussage, bei der ein äquivoker Begriff zu Missdeutungen führen kann.
Z.B.: Die Birne ist rot. → Kann eine Frucht aber auch eine Glühbirne gemeint sein.
 - Die Äquivokation kann die Grundlage eines logischen Fehlschlusses sein, wenn ein äquivoker Begriff in verschiedener Bedeutung verwendet wird.
- In welchem Bereich kommt dies in Computerlinguistik vor?

Äquivokationen

= Mehrdeutigkeit gleichlautender (äquivoker) Begriffe



Homonyme

Unterschiedliche
Bedeutung und auch oft
unterschiedliche Herkunft,
z.B.: Morphem „Tau“

Polysemie

Gemeinsame sprachliche
Wurzel und/oder
abgeleitete Bedeutung



Homographie

Dieselbe Schreibweise
bei unterschiedlicher
Aussprache

Homophonie

Unterschiedliche
Schreibweise
bei dieselber Aussprache

Ambiguität

- Von Ambiguität sprechen wir, wenn Äußerungen auf mehrere Weisen interpretiert werden können.
- Mehrdeutigkeiten oder Ambiguität (lat. ambiguus, zweifelhaft)
 - Wenn genau zwei: auch Doppeldeutig bzw. Zweideutig
- Hat nicht mit Wörtern bzgl. gleichlautender Begriffe zu tun (sondern in der Semantik)

Prosodie

- Sprachliche Eigenschaften hinsichtlich:
 - Wortakzent
 - Silbenbetonung
 - Satzmelodie
- Hilfreich für Auflösung von homographen Wörtern
- Beispiel (Satzmelodie):
 - Du gehst in dein Zimmer? (Frage)
 - Du gehst in dein Zimmer! (Befehl)

Signale

■ Spracherkennungssysteme

Für eingeschränkte Bereiche mit gutem Trainer-Modell wie z.B. Berichte/Protokolle diktieren sind sie ideal.

■ Aber der Vergleich von äquivoker Wörter und unpräzisen Sätze

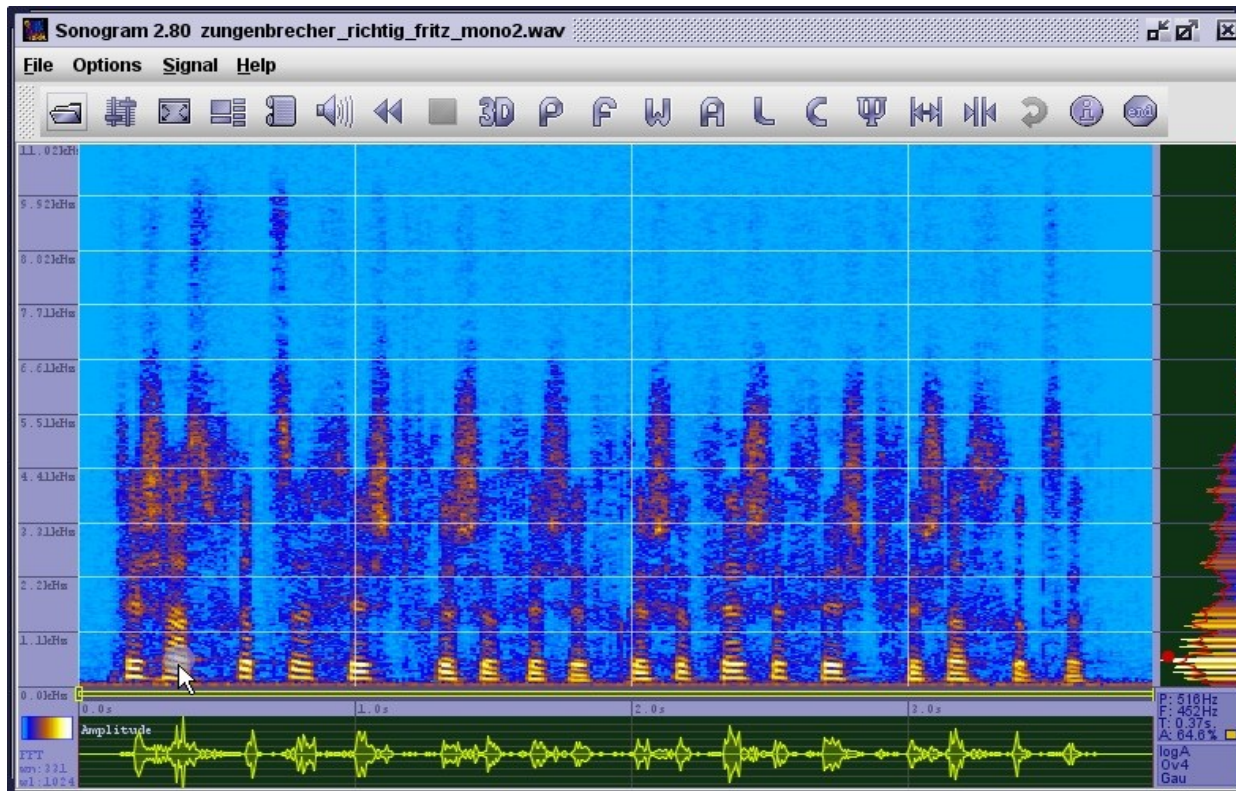
- bereiten schon die besten Systeme die Probleme.
(sowie Polysemie oder ihre Vagheit)

■ Beispiele:

- Zungenbrecher
- Lerche und Lärche (Homophon)
- modern und modern (Homograph)

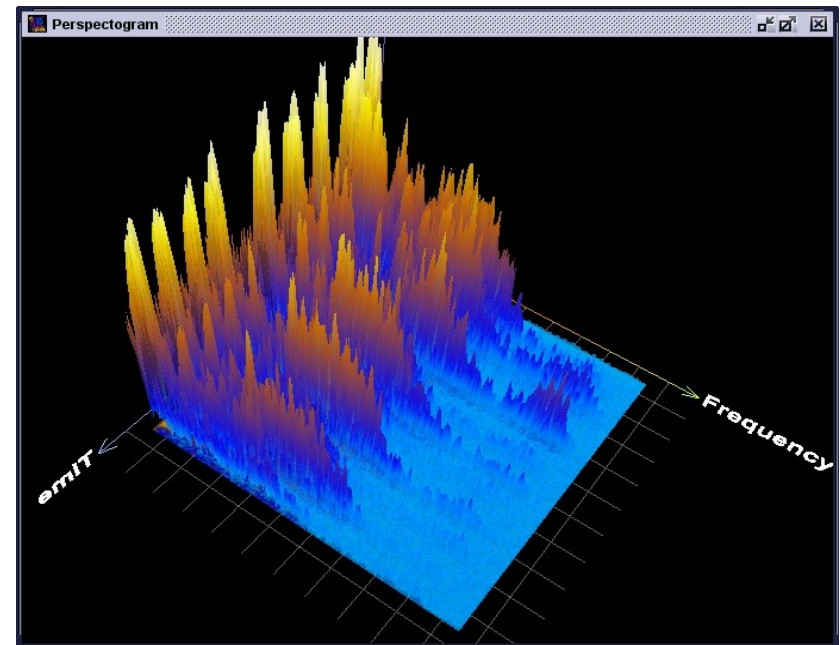
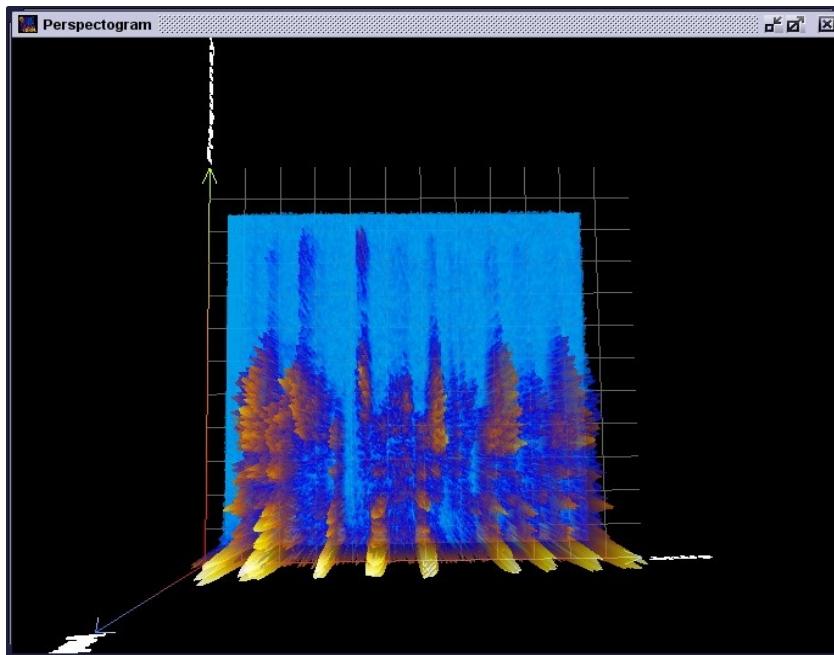
Signale

- Beispiel: Zungenbrecher
„Fischers Fritze fischt frische Fische. Frische Fische fischt ...“



Signale

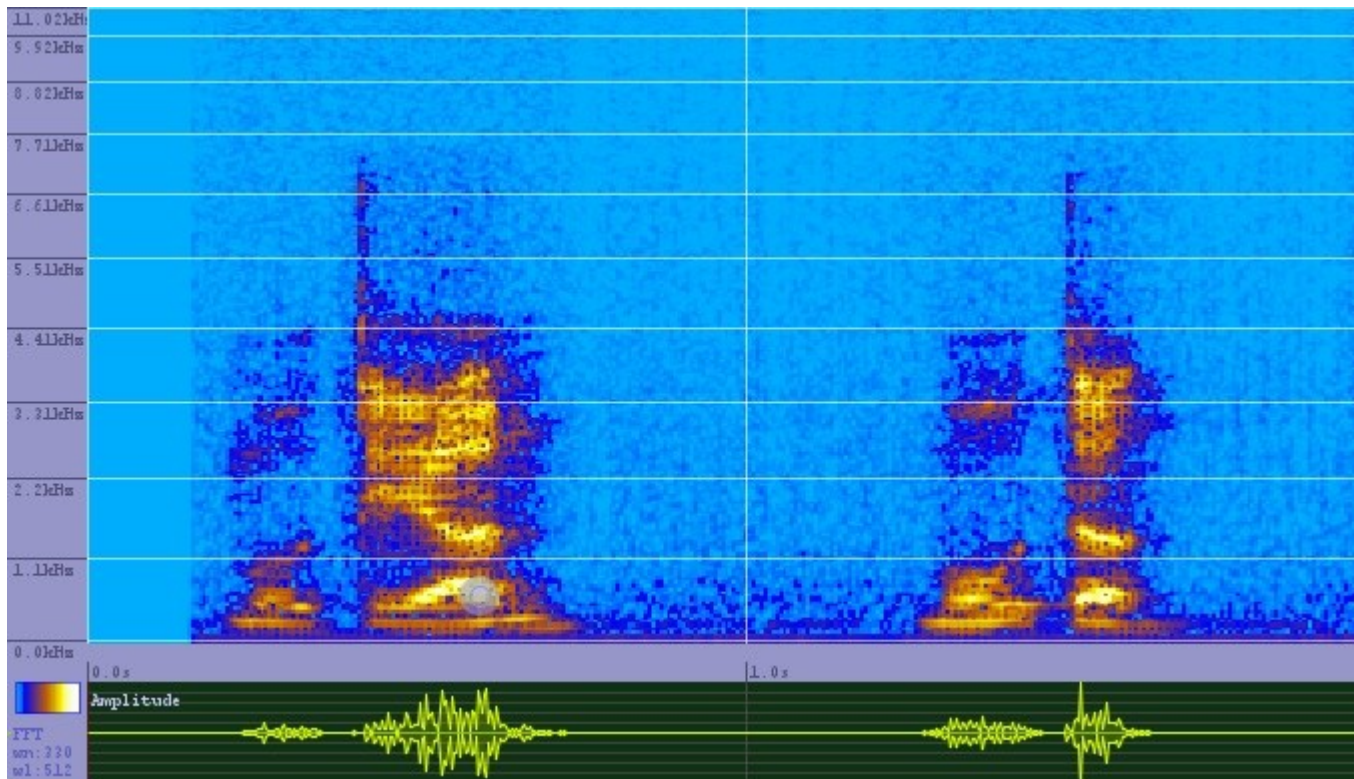
- Zungenbrecher:
„Fischers Fritze fischt frische Fische.“



Perspectogram (Achsen: x-y-z in Zeit - Frequenz - Amplitude)

Signale

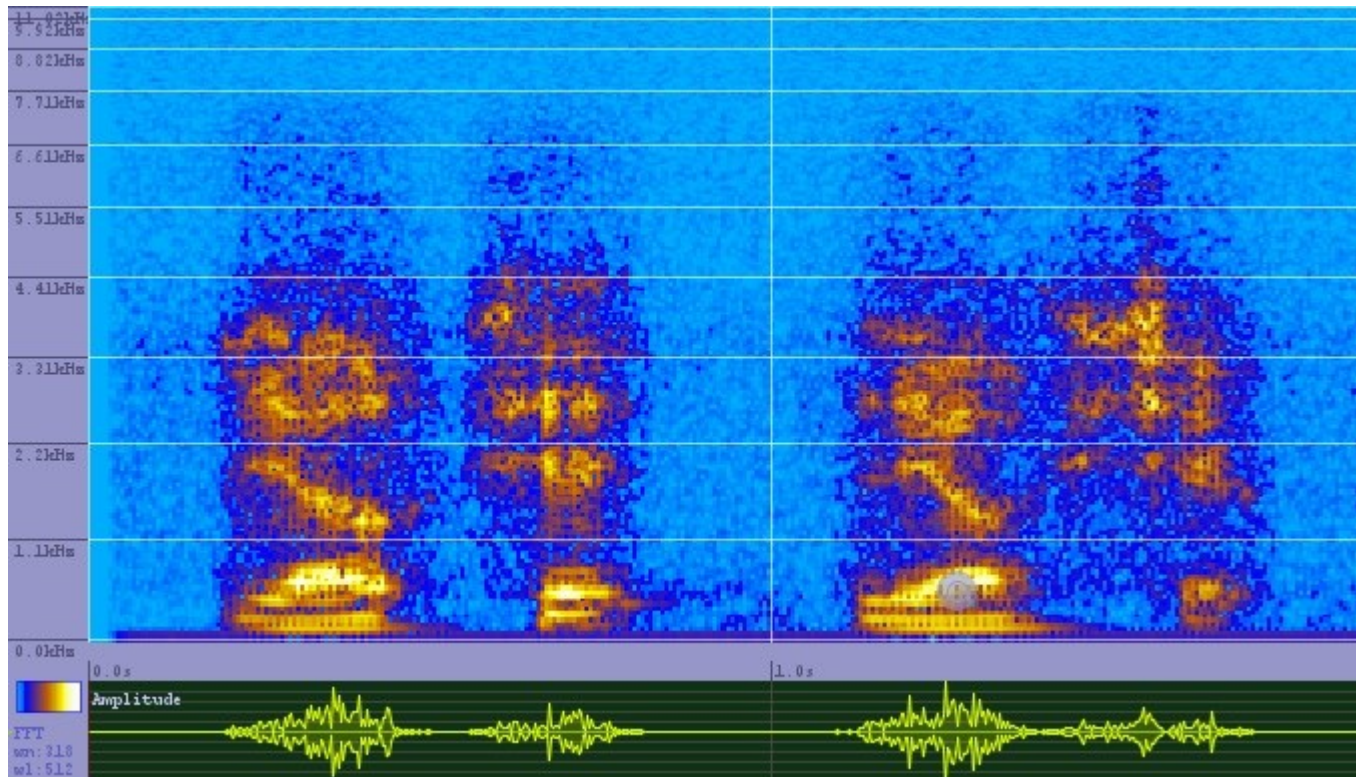
- Homograph: modern [mo'dɛɾn] (und) modern ['mo:dəɾn]



Sonogram: modérn und módern

Signale

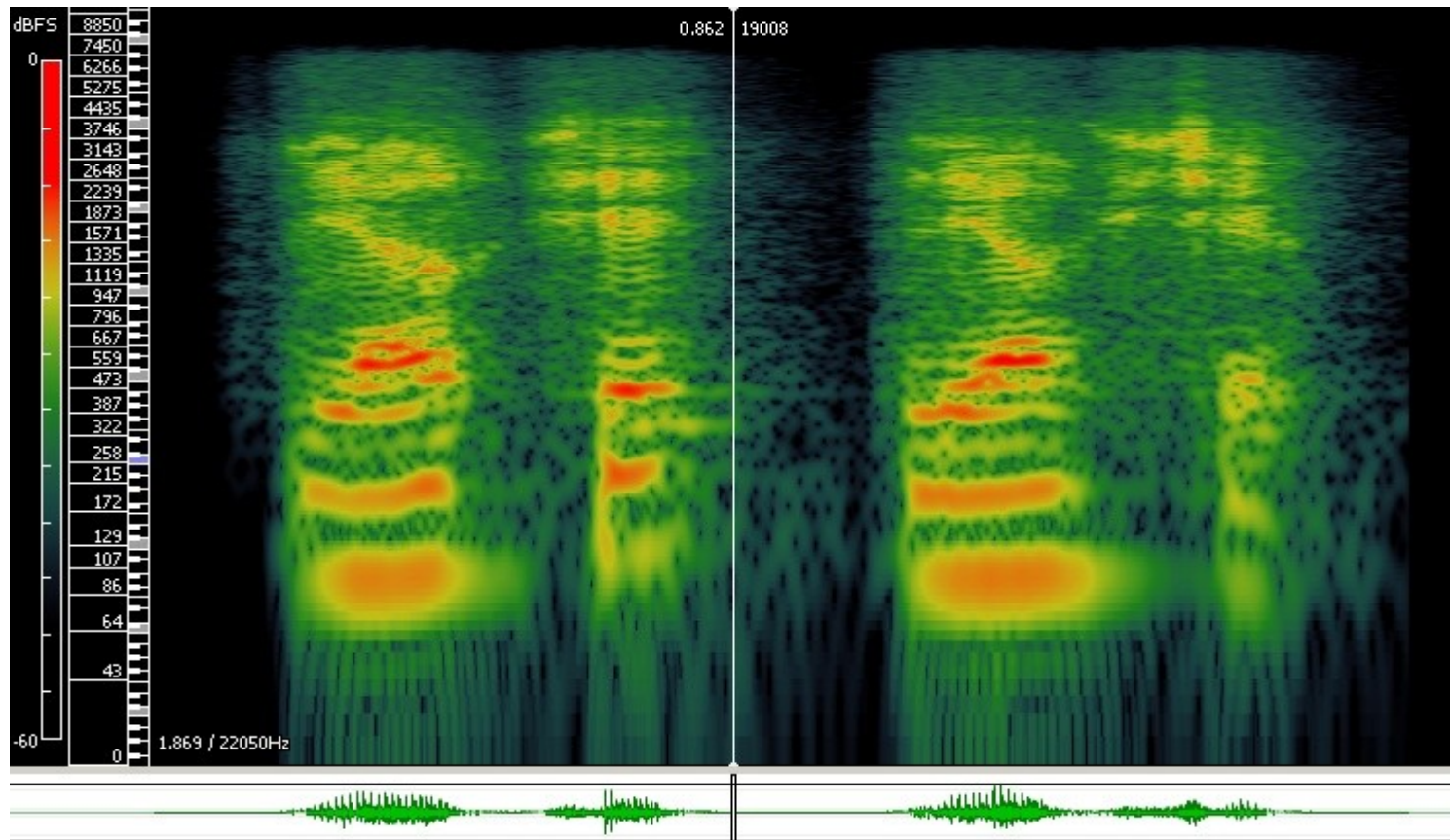
- Homophon (1): Lärche [ˈlɛrçə] (und) Lerche



Sonogram: Lärche und Lerche

Signale

- Homophon (2): Lärche / Lerche

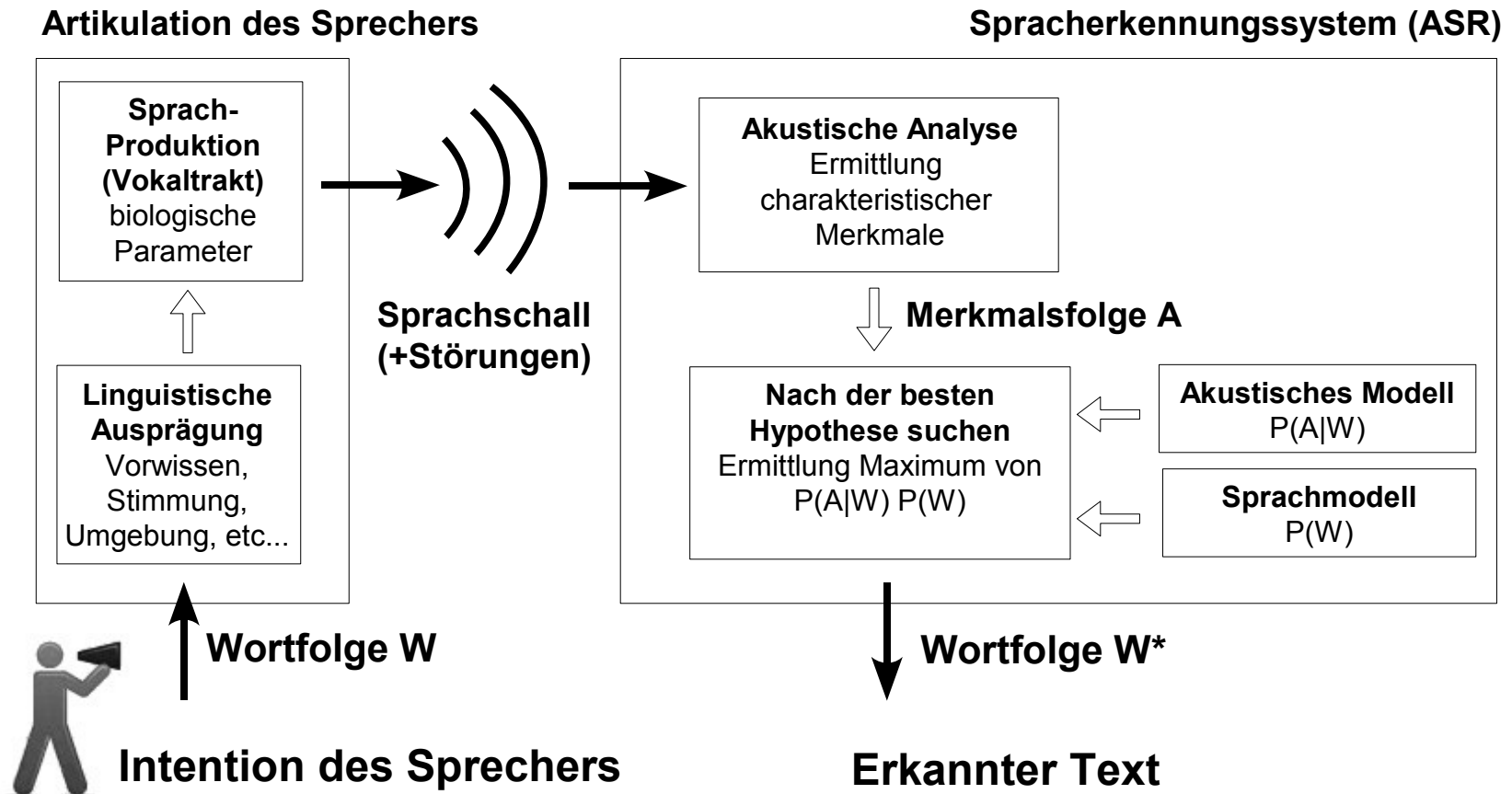


Sonogram: Lärche und Lerche (Sonic Visualiser)

Spracherkennungssysteme

- Automatic Speech Recognitions – ASR
- Vorverarbeitung, Normierung, Training
- Klassifikationen
- Nichtlineare Zeitanpassung
- Mustererkennung
 - HMM – Hidden Markov Model
 - Neuronale Netze

Spracherkennungssystem



Spracherkennungssystem

■ ASR

- Analyse → Merkmalsfolge (Vektoren)
- + Sprachmodell
kann statistische, grammatikalische oder auch syntaktische Regeln berücksichtigen, sortiert (unpassende) Wortpaare aus
- + Akustisches Modell
hoch, tief, schnell, ambiente, bzw. Sprecher-Verfassung, etc.
- Beste Hypothese

Spracherkennungssystem

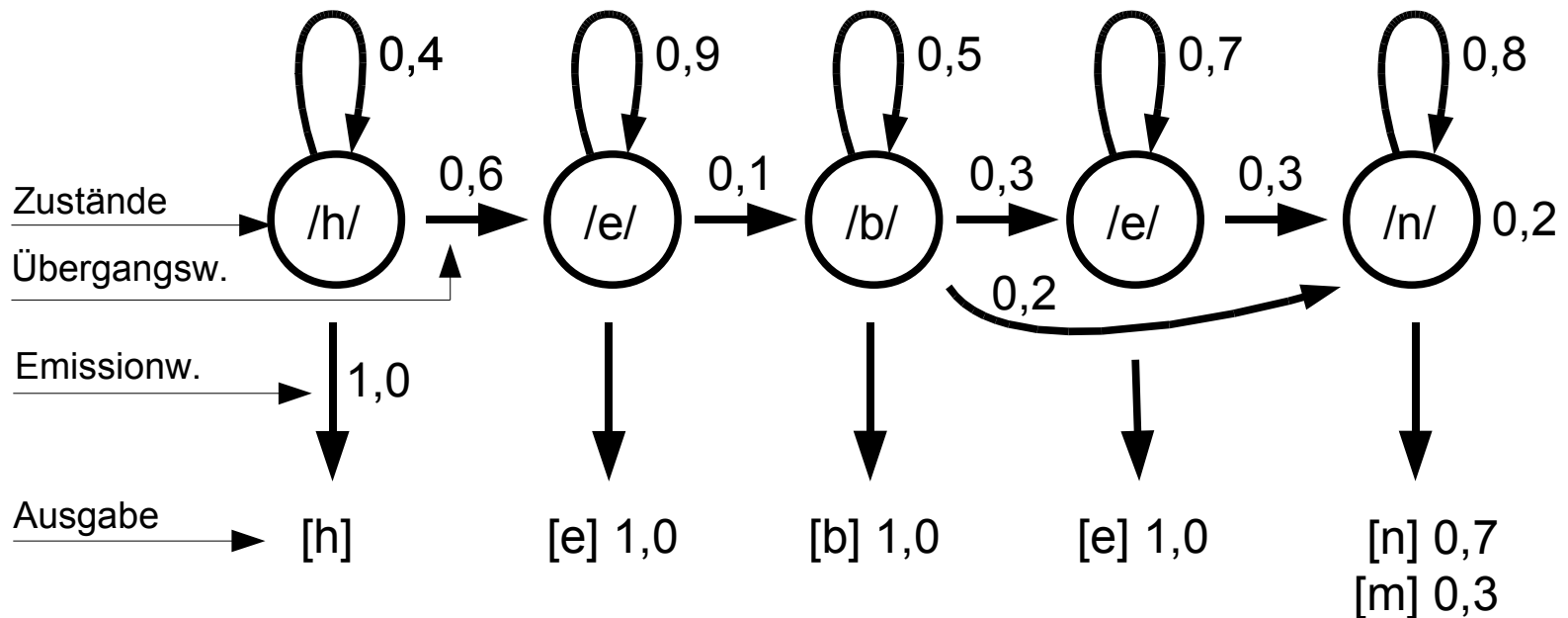
- Einzelworterkenner
 - einfach
 - Vokabular: zwischen einem und ca. hundert Wörter
 - Anwendung: Steuerung
- Kontinuierliche Erkennung
 - komplex
 - Zwei Anwendungsfälle:
 - Kleines Vokabular, sprecherabhängig, kein Sprachmodell
 - Großes Vokabular, sprecherunabhängig, mit Sprachmodell

Methoden

- Eine Methode zu finden um die identischen „Klang“-Wörter (vom Intention des Sprechers) richtig deuten zu können, ist aufwändig und kostenintensiv und zum teil noch weit entfernt.
- Methoden mit gute (phonetische) Annäherungsergebnisse sind eingeschränkt
- Häufigste Anwendungen:
 - HMM – Hidden Markov Model
 - Neuronale Netze

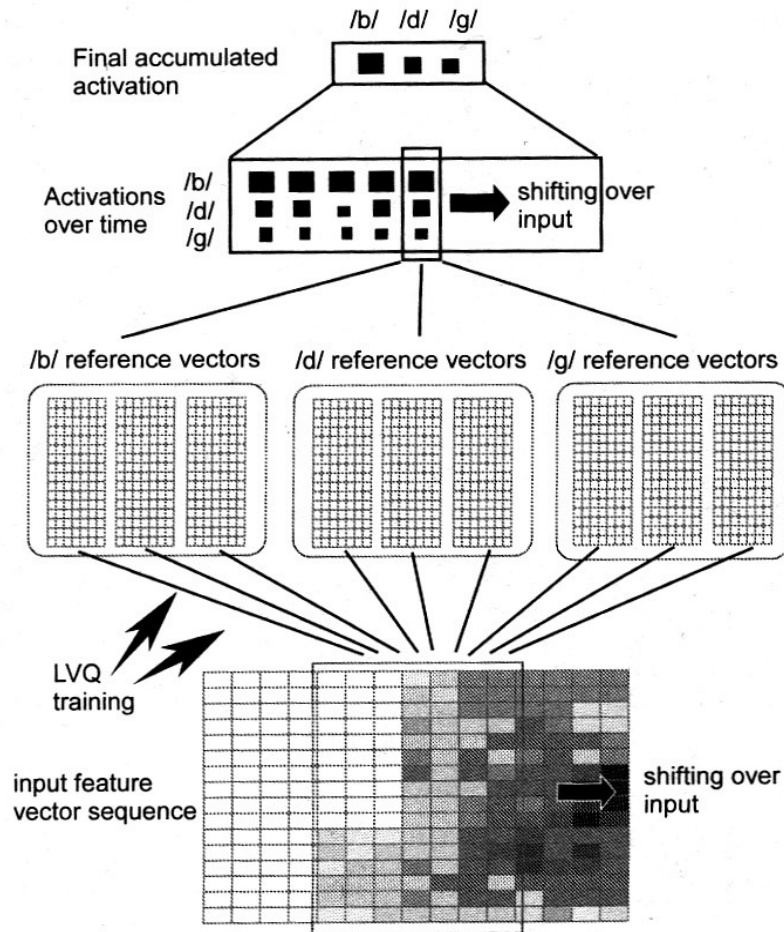
Spracherkennungssystem

- HMM – hidden markov model
 - (verborgene) Zustände, Übergangswahrscheinlichkeiten, Emissionswahrscheinlichkeiten, (sichtbare) Ausgabesymbole
 - Beispiel: „heben“



Spracherkennungssystem

- Neuronale Netze



Anwendung

- Sprachverstehen in Dialog-Systemen
 - Auskunft
 - Feedbacksystem

- Nebeneffekt:
 - Verfassung des Sprechers erkennen
 - Müde (langsam), Munter/Fit (schnell)

Zusammenfassung

- Erkennen von Mehrdeutigkeiten nach wie vor:
sehr Schwer (aber nicht unmöglich ;-)
- Vorwissen und Gedächtnis des Spracherkenners vorhanden?
- Abhängig vom aktuellen Rechnerleistung im Consumer-Bereich
(→ Rentabilität der Software / Hardware Anwendung)

Quellen

- Wendemuth A., Grundlagen d. Stochastischen Sprachverarbeitung, 2004
- Eppinger/Herter, Sprachverarbeitung, 1993
- Carstensen, Computerlinguistik – Eine Einführung, 2004
- Wu Chou, Pattern Recognition in Speech and Language Processing, 2003
- Schukat-Talamazzini, Automatische Spracherkennung, 1995
- Daniel v. Grünigen, Digitale Signalverarbeitung, 2001

- Verwendete Software (alle Freeware):
 - Sonogram v2.80 (sonogram, perspectogram)
 - WavePad Sound Editor v4.02 (editor, noise reduction)
 - Wavosaur v1.0.3.0 (record, PCM 22.050kHz, mono)
 - Sonic Visualiser v1.3 (sonogram)

- Weitere Zungenbrecher siehe:
<http://de.wikiquote.org/wiki/Zungenbrecher>