

Sprachverstehen

Vorlesung an der TU Chemnitz

Wintersemester 2011/2012

Dr. Johannes Steinmüller

Johannes Steinmüller

- 1/B309
- Tel.: 531 35198
- stj@informatik.tu-chemnitz.de
- Seite zur Vorlesung:

<http://www.tu-chemnitz.de/informatik/KI/edu/spraver/>

Inhalt

- Einführung – Überblick
- Allgemeine Begriffe – Sprachliche Einheiten
- Aufbau von Sprachverarbeitungssystemen
- Syntax
 - Grammatiken
 - Übergangsnetze
 - Merkmalsstrukturen
 - Unifikationsgrammatiken
- Semantik
 - Aufgaben
 - Semantische Netze
 - Logikformalismen
 - Diskursrepräsentationstheorie
- Erkennung gesprochener Sprache
 - Nichtlineare Zeitanpassung
 - Hidden Markov Modelle
 - Statistische Spracherkennung
- Sprachmodellierung
- Zeitliches Schließen

Nicht in dieser Vorlesung

- Sprache bei Menschen
- Kognition
- Signalverarbeitung
- Neuronale Netze
- Sprachgenerierung

Literatur

St. Euler: **Grundkurs Spracherkennung**
Vieweg, 2006

B. Pfister, T. Kaufmann: **Sprachverarbeitung**
Springer, 2008

Kai-Uwe Carstensen u. a.: **Computerlinguistik und Sprachtechnologie**
Spektrum Akademischer Verlag, 2001, 2004, 2010

www.linguatec.net (Spracherkennung, Sprachausgabe, Übersetzung)

www.sympalog.de (Dialog)

www.lt-world.org/ (Language Technology World)

1 Einführung

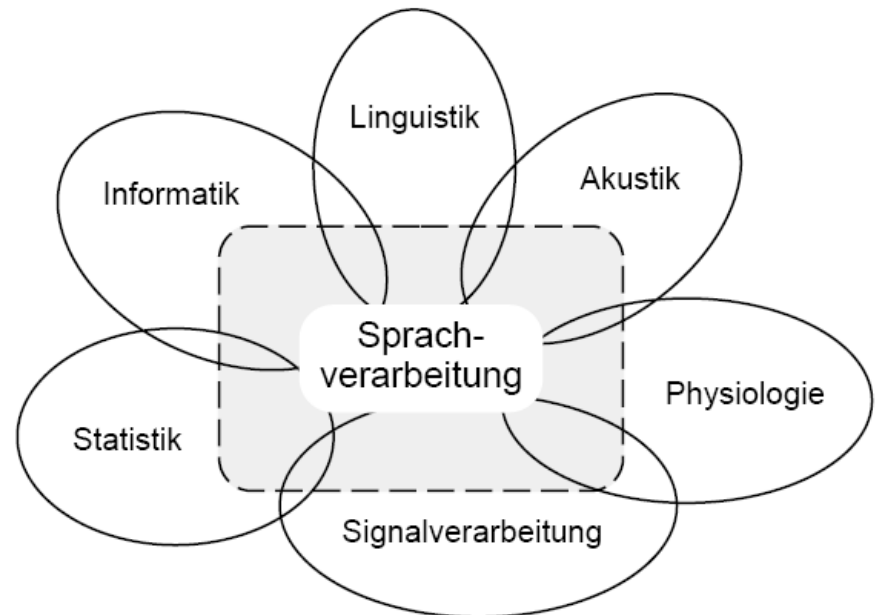
Inhalt

- Einordnung
- Sprache
- Was ist Sprachverarbeitung?
- Anwendungen der Sprachverarbeitung
- Geschichte

1.1 Einordnung

Interdisziplinärer Charakter

- Informatik (Künstliche Intelligenz)
- Linguistik (Computerlinguistik)
- Mathematik (Logik, Statistik)
- Kognitive Psychologie (Kognition)
- Neuroinformatik, Neurobiologie
- Informationswissenschaft
- Signalverarbeitung
- Philosophie

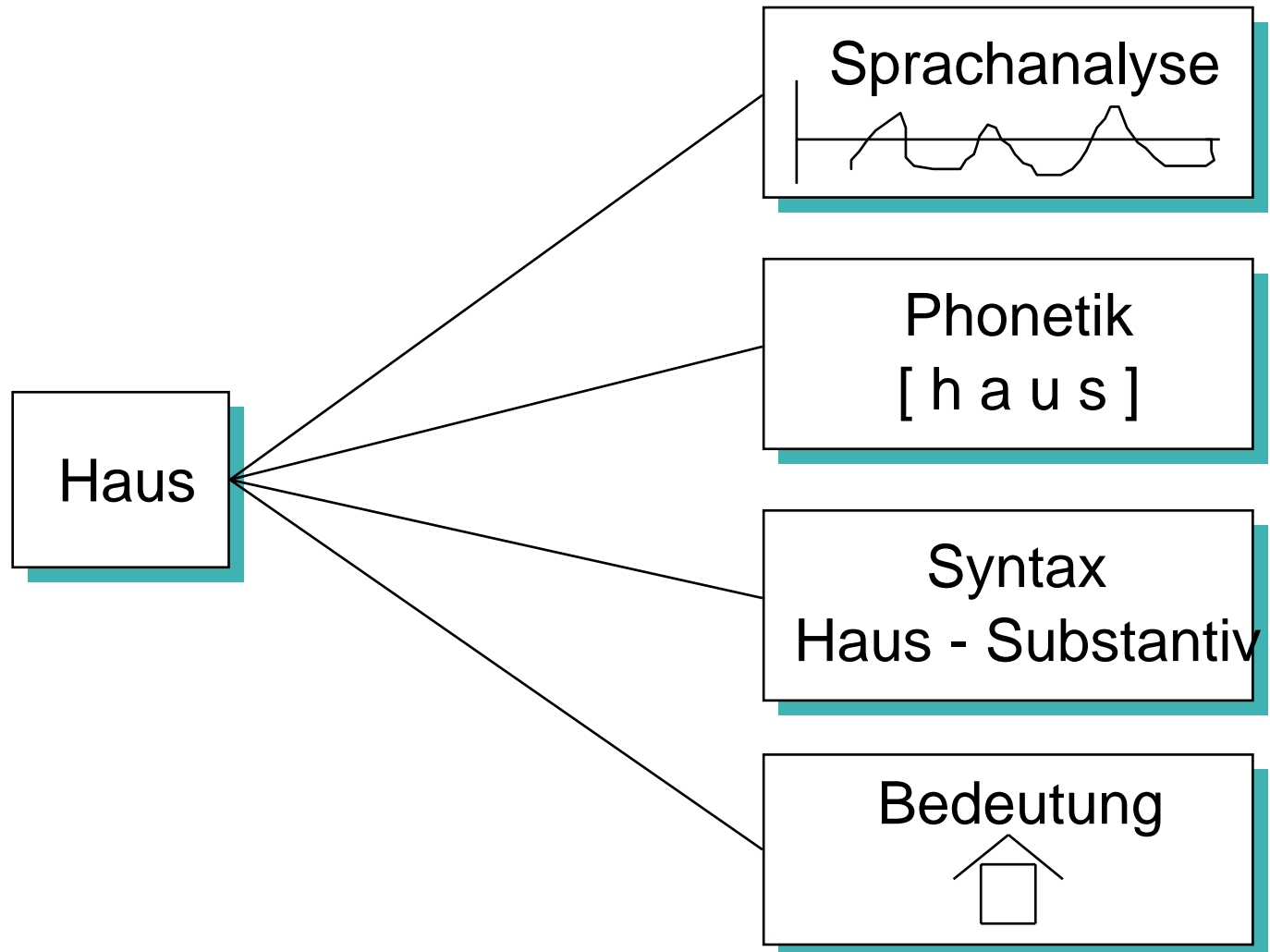


1.2 Sprache

Sprache

- dient zur Verständigung zwischen Menschen
- wichtig: Bedeutung einer sprachlichen Äußerung
- Sprecher und Hörer müssen eine Lautfolge in gleicher Weise verstehen
- geschriebene Sprache (Text)
- gesprochene Sprache (Sprachsignal)

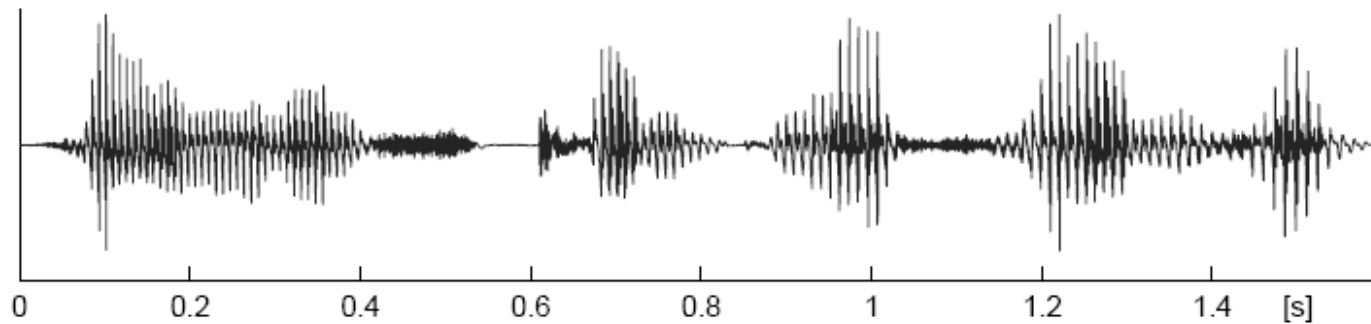
Verschiedene Betrachtungsarten der Sprache



Text und Sprachsignal

Text: *“Heinrich kommt nach Hause.”*

Sprachsignal:



Warum Sprache?

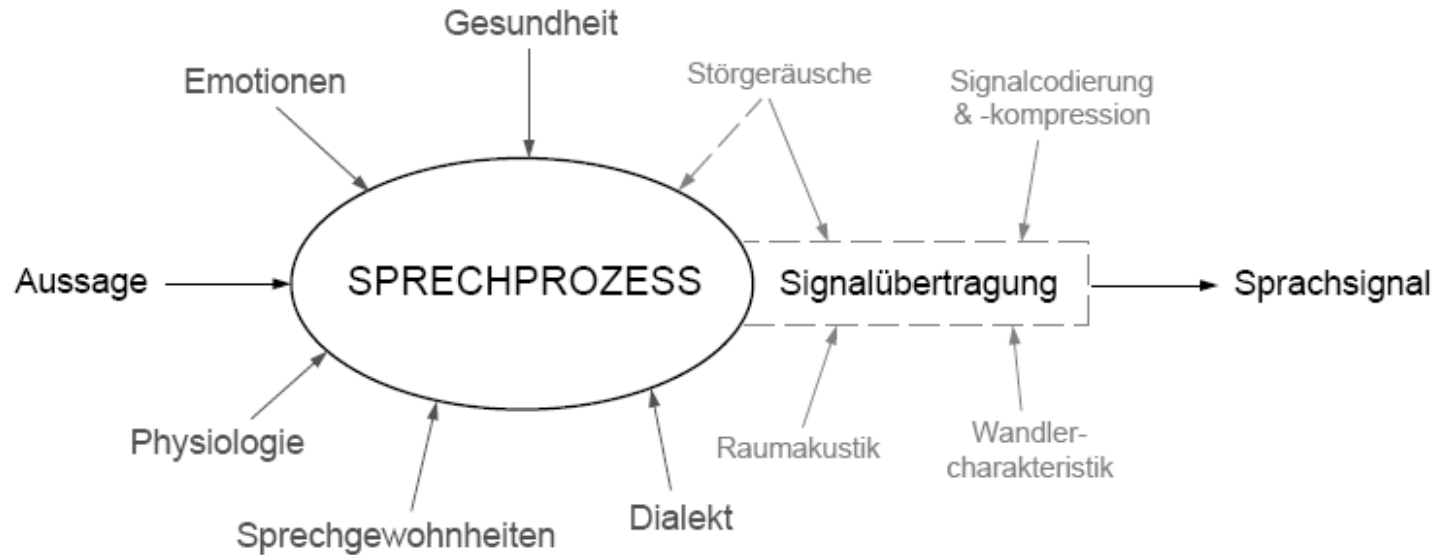
- natürlichste Kommunikationsform
- effizientes Mensch-Maschine-Interface
 - Tastatur: 100-150 Wörter pro Minute
(bei stark geschulter Bedienungsperson)
sonst nur 10-25 Wörter pro Minute
 - Sprache: 120-250 Wörter pro Minute
- kein physikalischer Mensch-Maschine-Kontakt
 - hoher Grad an Bewegungsfreiheit (Hände bleiben frei)
 - funktioniert über eine räumliche Distanz hinweg
 - keine visuellen Einschränkungen (Dunkelheit, schlechtes Licht)
- zusätzlicher Informationskanal
- kein wesentlicher Raumbedarf
- Zugang zum öffentlichen Telefonnetz

1.3 Was ist Sprachverarbeitung?

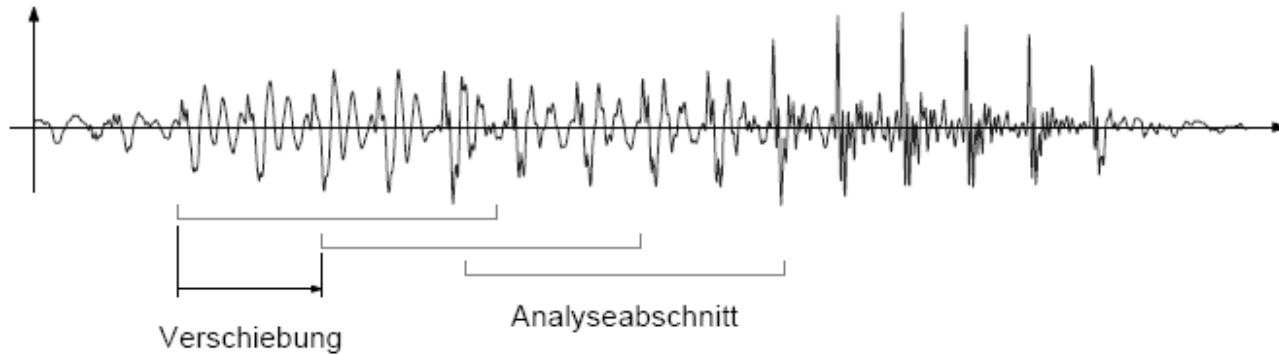
Sprachverarbeitung

- Informationsquelle (Mensch)
- Schallwellen, elektrisches Signal
- Digitalisierung
- Signaltransformation (Spektrum, Merkmale)
 - Berechnung von ca. 30-50 Merkmalen alle 10 ms
 - dazu betrachtet man einen kurzen Abschnitt (20 – 40 ms) des Sprachsignals
- Spracherkennung

Sprechprozess – Sprachsignal



Berechnung von Merkmalen



System zur Sprachverarbeitung

- ***Sprachanalyse***: Eine sprachliche Eingabe wird verstanden, syntaktisch und semantisch analysiert, eine Bedeutung zugeordnet und ausgewertet.
 - ***automatische Spracherkennung***
 - ***Sprachverstehen***
- ***Sprachsynthese*** (Sprachgenerierung): Eine Antwort wird wieder in Sätze (ev. akustische Form) zur Ausgabe an den Benutzer übersetzt.

Auskunfts-system für InterCity-Verbindungen

S: *Hier ist die automatische InterCity-Auskunft. Was kann ich für Sie tun?*

B: *Ich will morgen abend nach Frankfurt.*

S: *Sie können ab Bonn fahren um [...]*

B: *Gibt es auch noch einen früheren Zug?*

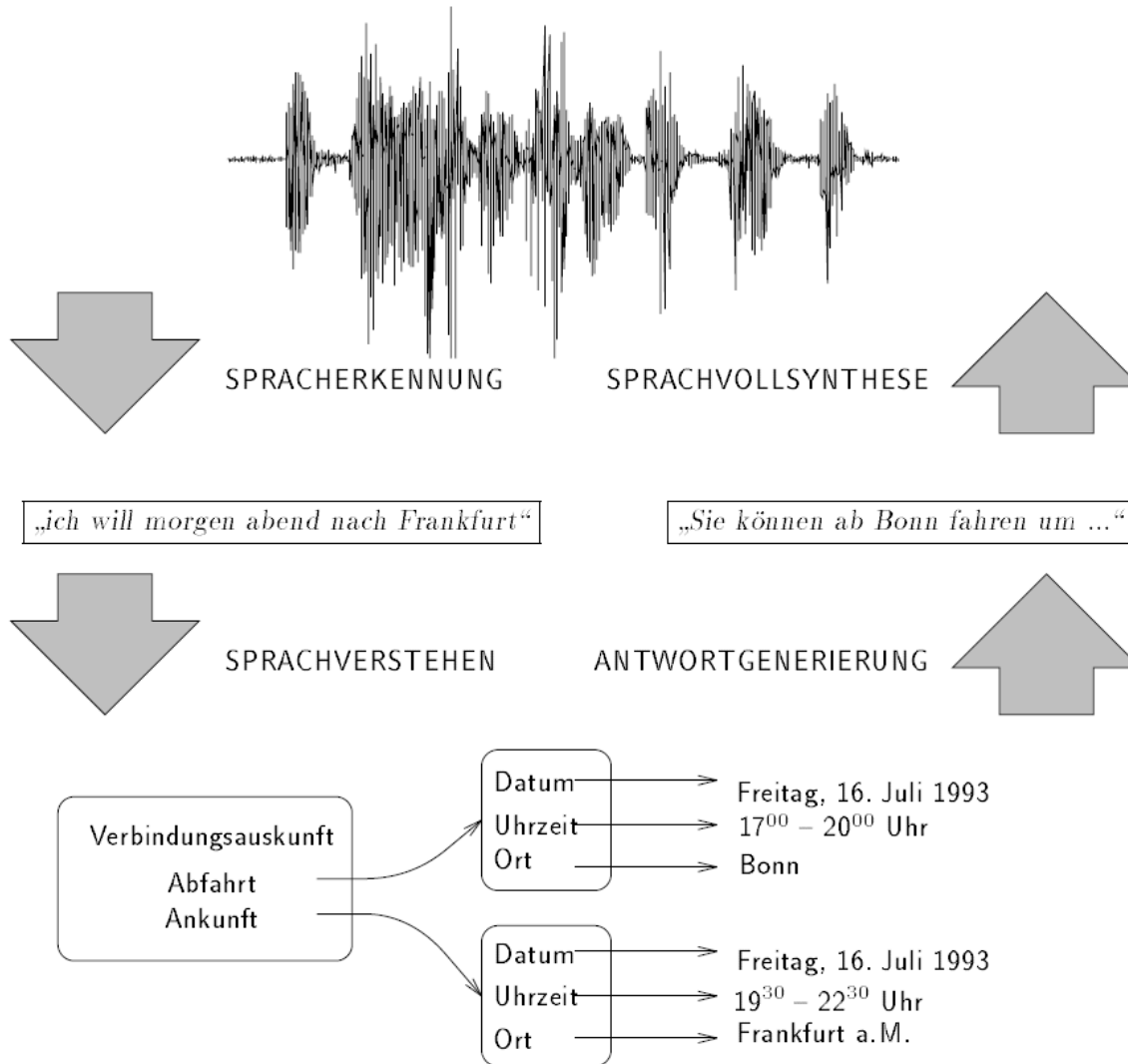
S: *Bis wann möchten Sie spätestens in Frankfurt ankommen?*

B: *Bis einundzwanzig Uhr.*

S: *Sie können ab Bonn fahren um [...]*

B: *Vielen Dank. Auf Wiedersehen.*

Auskunfts-system für InterCity-Verbindungen



Fähigkeiten und Eigenschaften

- Einzelworterkennung
 - Sprachpause zwischen 2 Wörtern
 - Mustererkennung beruht auf einem Vergleich ganzer Wörter
- Wortkettenerkennung
 - sind in der Lage, Ausdrücke, die mehrere Wörter enthalten, zu verarbeiten und die einzelnen Wörter zu zerteilen
 - nur innerhalb der Kette fließend
 - z.B.: Ziffern einer Telefonnummer
- Erkennung fließender Sprache
 - spontane Sprache

Fähigkeiten und Eigenschaften

- sprecherabhängig
 - auf einen Sprecher zugeschnitten
 - System muss mit Sprachproben trainiert werden
- sprecheradaptiv
 - während des Erkennungsvorganges passt sich das System an die besonderen Eigenschaften des Sprechers an
- sprecherunabhängig

Wortschatz

Alarmstopschalter	1						
Menü-Steuerung (J/N)	2						
Zahlen/Ziffern	$10+x$						
Gerätebedienung	20-200						
Auskunftsdialog	500-2000						
Alltagssprache	8 000-20 000						
Diktiermaschine						20 000-50 000	
Deutsch ohne Fremdwörter							ca. 300 000

1.4 Anwendungen

Anwendungen der Sprachverarbeitung

- Sprachsignalcodierung
 - Übertragung, Speicherung und Chiffrierung von Sprache
- Verbesserung der Sprachqualität
- Sprechertransformation
 - ein Sprachsignal, das von einem Sprecher X gesprochen worden ist, so zu verändern, dass es als die Stimme eines Sprechers Y wahrgenommen wird
- Sprechererkennung
 - Identifikation: Erkennen einer Person anhand ihrer Stimme (bekannter Personenkreis)
 - Verifikation: entscheiden, ob ein Sprecher zu einem bekannten Personenkreis gehört (Zugangskontrollen)
- Sprachidentifikation
- Keyword-Spotting
 - Wurden bestimmte Schlüsselwörter gesagt?
- Spracherkennung
- Sprachsynthese
- Übersetzung, Dolmetschen
- Dialoganwendungen
- Google
- Geschäftsprozesse
 - Briefe klassifizieren

Übersetzung

Beispiel:

Der Geist ist willig, aber das Fleisch ist schwach.

Dieser Satz wurde ins Russische übersetzt und dann wieder zurück. Es entstand:

Der Wodka ist gut, aber das Essen verdorben.

Anwendungsklassen in der Spracherkennung

- Kommandoworterkenner
 - kleiner Wortschatz
 - Steuerung von Geräten
 - zuverlässig unter extrem schwierigen Bedingungen
- Diktiersysteme
 - Sprecheradaptiv
 - großer Wortschatz (100000 Wörter)
 - Bedeutung des gesprochenen Textes wird nicht verstanden
- interaktive Dialogsysteme
 - sprecherunabhängig
 - spontane zielgerichtete Äußerungen, Telefon
 - z.B.: Kinoreservierung, Hotelvermittlung, Zugverbindungen

1.5 Geschichte

Geschichte

- 1791: Sprechmaschine (KEMPELEN)
Buch "Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine,,
baute auch den Schachautomaten "Der Türke"
- 1956: Theorie der formalen Sprachen (CHOMSKY)
- 1966: Verarbeitung natürlicher Sprachen mit einfachen
Mustertransformationen (ELIZA)
<http://www.med-ai.com/models/eliza.html.de>
- 1968: semantische Netze
- 1970: effektive Grammatikformalismen (ATN-Grammatiken)

Geschichte

- 1972: **LUNAR** -- Auskunftssystem über die von den Mondlandungen mitgebrachten Gesteinsproben
- 1977: **GUS** -- Beratung und Reservierung bei Flugreisen
- 1990: **SPICOS II** -- Dialog mit einer Datenbank (verarbeitet fließend gesprochene Sprache und generiert akustische Antwort)
- 1994: **TANGORA** -- sprecherabhängiges Diktiersystem (20000 Wortformen, Pausen zwischen den Wörtern)
- 1995: **EVAR** -- IC Zugauskunft
- 1991 - 2001: **VERBMOBIL** -- Simultanübersetzung von Verhandlungsdialogen

aktueller Stand

- Kommandoworterkenner
- Diktiersystem, Spracherkennung
www.linguatec.net
- Dialogsysteme (Kino, Sport, Wetter, Reservierung, Börsenauskunft) – SymRec / SymComponents
www.sympalog.de
- Keyword-Spotter – SymSpot
www.sympalog.de
- Erkennung von Emotionen aus Sprachsignalen – SymEmotion
www.sympalog.de