

Interdisziplinäre Fachdidaktik: Natürliche Sprachen und Formale Sprachen

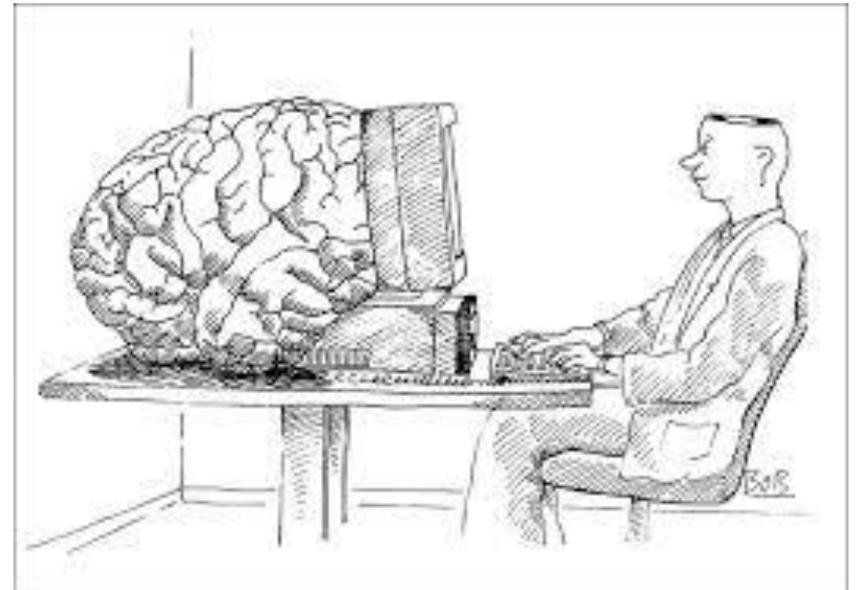
Wie verstehen Computer natürliche Sprachen?

LVA – Leiter: Elizabeta Jenko
Wilfried Grossmann

erstellt von

Egbert Althammer 9031924

Herbert Schindelka 8526917



Intention und Zielgruppe (fachdidaktische Sicht)

- Überblick über das Thema
- Neugierig machen auf das Thema
- Motivation zum Thema zu schaffen
- Zielgruppe
 - BHS 4. Klasse (Informatik)
 - AHS 7. Klasse (Wahlpflichtfach Informatik)

→ Anregung für VWA bzw. Diplomarbeit

Agenda

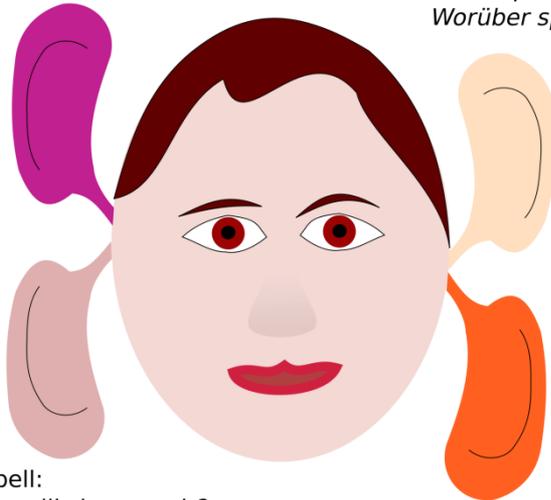
- Grundsätzliche Überlegungen
- Anatomie, Physik und Informatik
- Saarbrückener Pipeline Modell
- Problemstellungen
- Natürliche Sprache vs. künstliche Sprache
- Spracherkennung vs. Stimmerkennung
- Anwendungen

Grundüberlegungen zu „verstehen“ und „natürliche Sprache“

Schulz von Thun

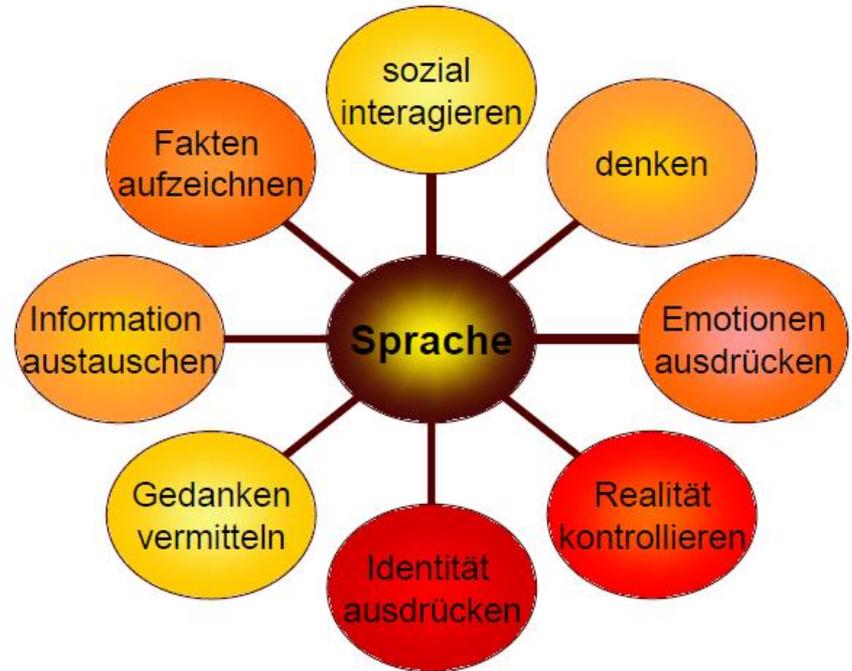
Selbstaussage:
Was offenbart sie über sich?

Sachaspekt:
Worüber spricht sie?



Appell:
Was will sie von mir?

Beziehungsaspekt:
Wie steht sie zu mir?



Kommunikation ist nicht das, was A sagt sondern B versteht (n. Watzlawik)

Wie verstehen Computer natürliche Sprache?

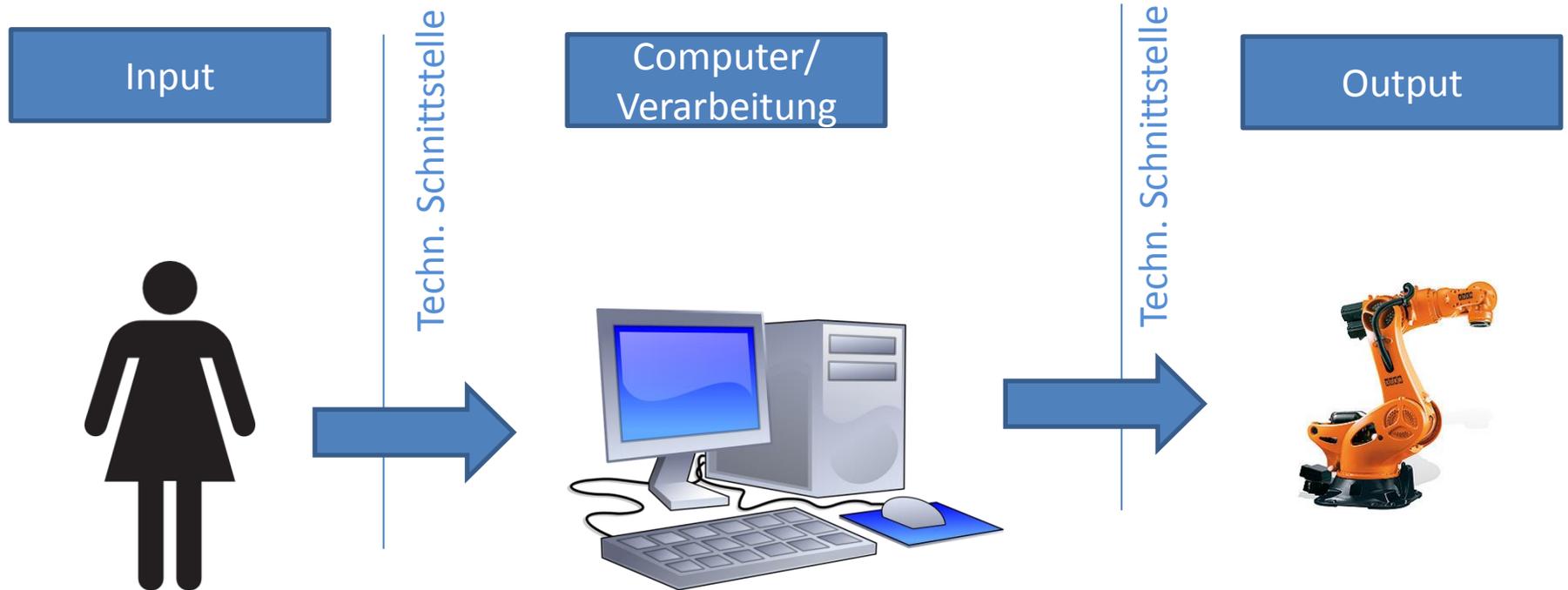
Was heißt verstehen?

- Klare Intention des Users
- Klare Vorstellung von dem was der Computer tun soll

- Technische Schnittstelle

- befehlsorientiert
- und/oder „Freier Text“

Warum möchten wir (eigentlich), dass uns der Computer versteht?



Warum möchten wir (eigentlich), dass uns der Computer versteht?

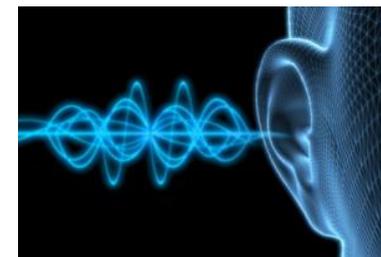


Techn. Schnittstelle

Mikrophon

Sprachaufzeichnung
Spracherkennung
Sprachsteuerung
Stimmerkennung

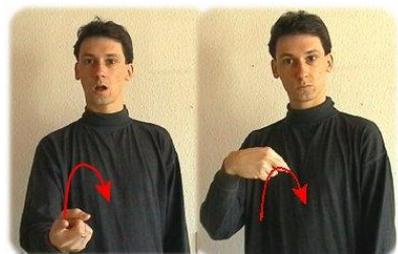
Sprachsynthese



Techn. Schnittstelle

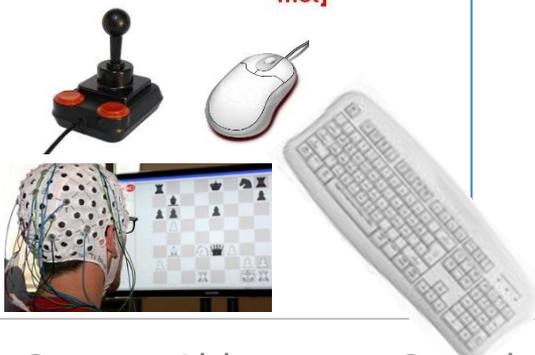
Video

Videoaufzeichnung
Mimik
Gestik (Gebärden)
Lippenlesen
Eye-Tracking



Du kommst zu mir.

kom mst]



Techn.
Interface

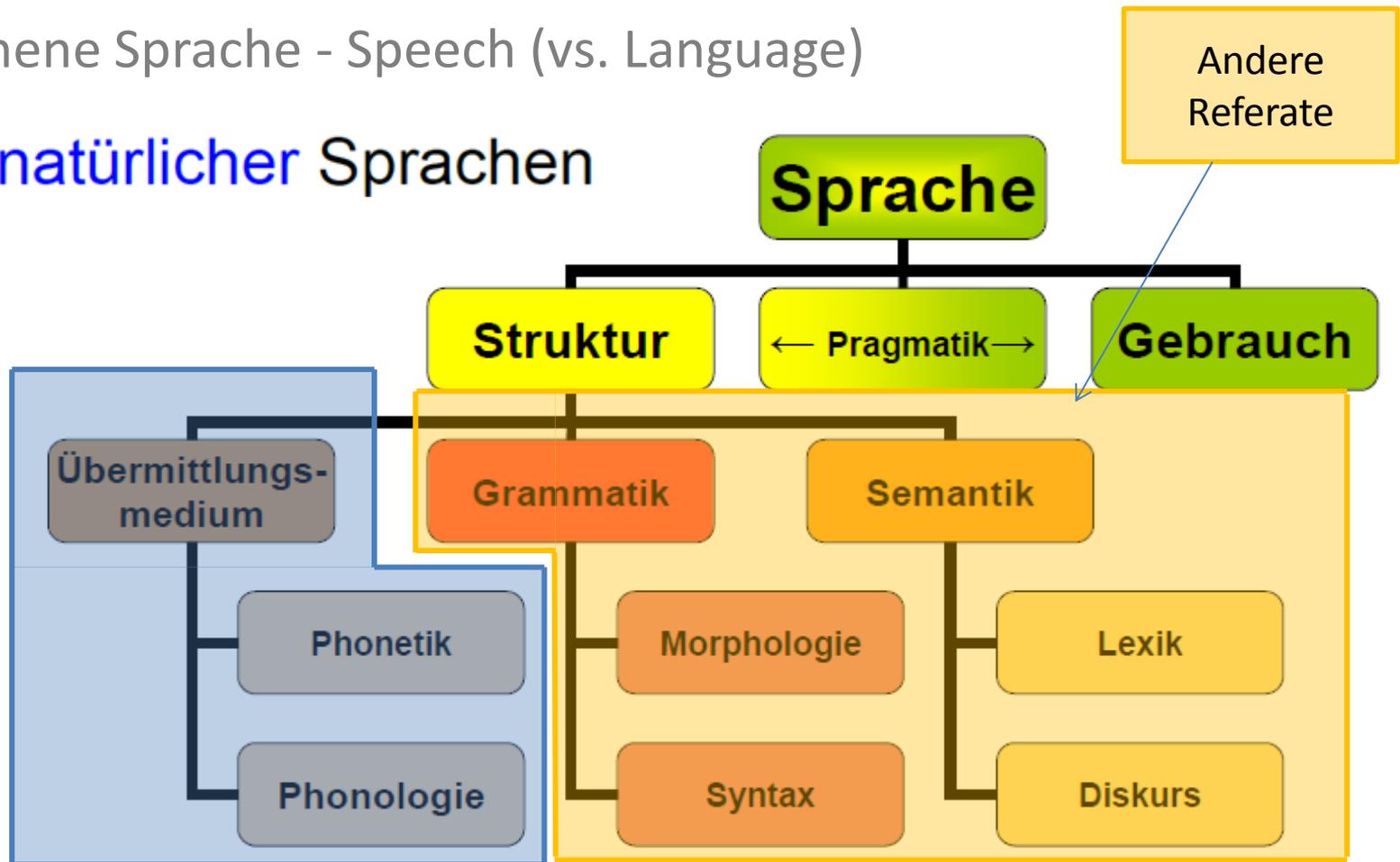
Maus, Joystick
Tastatur
„Steuerung über Gedanken“



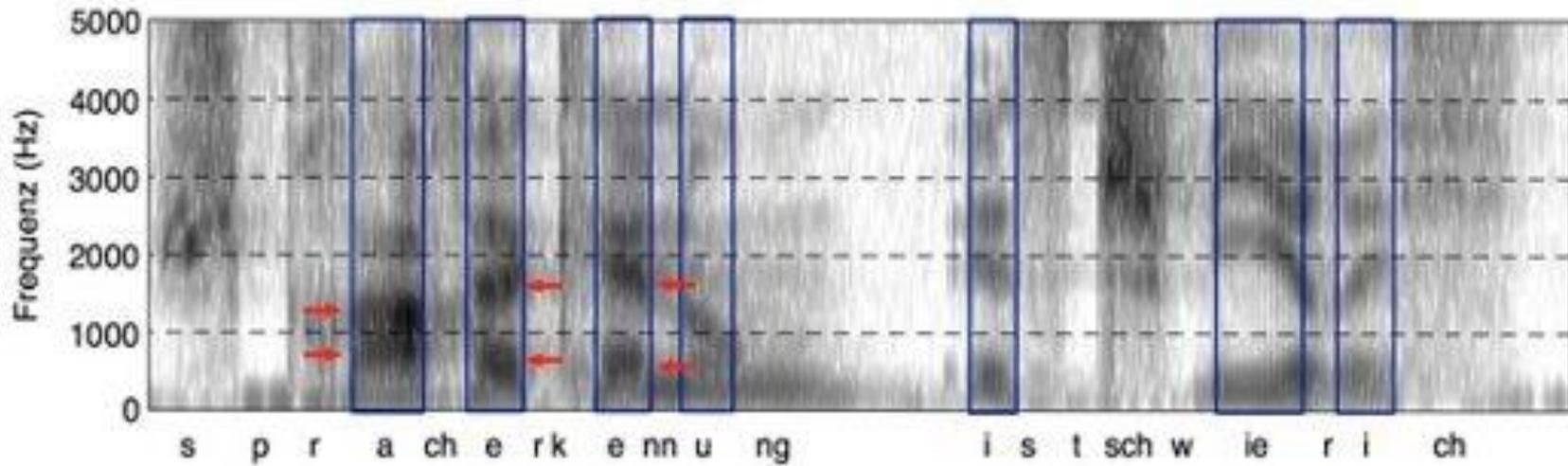
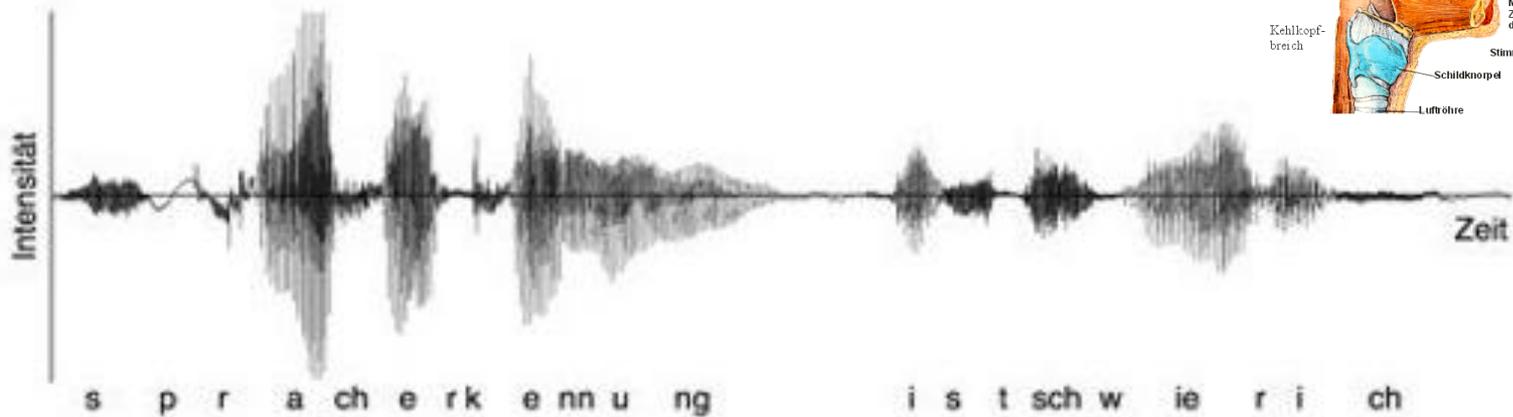
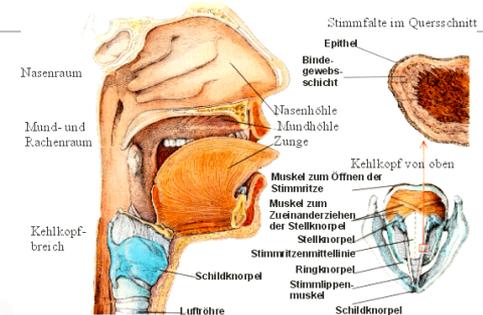
Fokus für heute

Gesprochene Sprache - Speech (vs. Language)

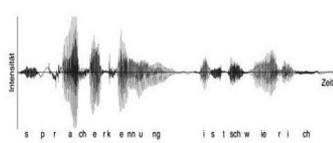
Modell natürlicher Sprachen



Zur Anatomie, Physik und Informatik



„Techn.“ Umsetzung



Signal –
aufbereitung
(Filterung)

Umweltgeräusche
Rauschen
Lautstärke

(Sprecher-)Training

Aussprache-
wörterbücher

Sprachsynthese



Eigentl.
Signalanalyse

Mustererkennung

Phoneme

„Wörterbücher“

Silben/Wörter

Phonem: kleinste lautliche Einheit auf der Ebene des Sprachsystems, die selbst keine sprachliche Bedeutung trägt, sondern nur bedeutungsunterscheidend ist

„Wissenschaftlicher“ Ansatz

- Disziplin: Computerlinguistik
- Saarbrücker Pipeline Modell
 - **Spracherkennung:** Schall in Phoneme
 - **Tokenisierung :** Die Buchstabenkette wird in Wörter, Sätze etc. segmentiert.
 - **Morphologische Analyse:** Personalformen oder Fallmarkierungen werden analysiert, um die grammatische Information zu extrahieren und die Wörter im Text auf Grundformen zurückzuführen, wie sie z. B. im Lexikon stehen.
 - **Syntaktische Analyse** Die Wörter jedes Satzes werden auf ihre strukturelle Funktion im Satz hin analysiert (z. B. Subjekt, Objekt, Modifikator, Artikel, etc.).
 - **Semantische Analyse:** Den Sätzen bzw. ihren Teilen wird Bedeutung zugeordnet. Dieser Schritt umfasst potentiell eine Vielzahl verschiedener Einzelschritte, da Bedeutung schwer fassbar ist.
 - **Dialog- und Diskursanalyse:** Die Beziehungen zwischen aufeinander folgenden Sätzen werden erkannt

Warum ist Spracherkennung schwierig?

- **Gewählte Sprache an sich**
- **Variabilität**
 - Signalqualität
 - Phonetisch: Unterschiedliche Aussprache (verschiedene Sprecher, Dialekte, Ansprache, Plauderei, W/M/Kind, ...)
 - Linguistisch: Homonym, Homophon (Ferse vs. Verse)
- **Computer besitzt kein Weltwissen**
 - Pragmatisches Wissen was kann jetzt überhaupt gesagt werden
 - Keine multimodale Erfassung (vgl. aber Mensch Experiment McGurk)
- **Segmentierung** (Buchstaben, Wörter, Sätze, ...)
- **Spontane Wörter**
 - Abgebrochene Wörter, Ähhh, Mmmhhh
- **Unbekannte Wörter**
- **Erkennen von Emotionen, Erkennen von Witz, Satire**

Vor- und Nachteile

Vorteile:

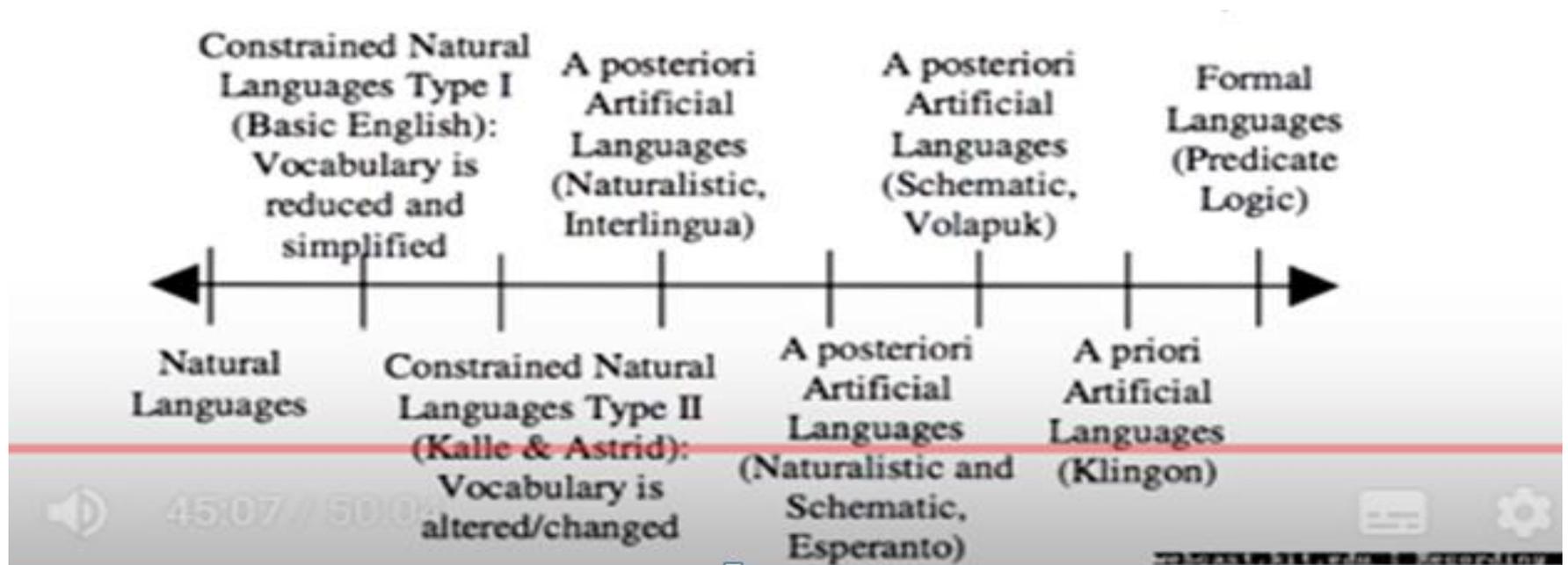
- Hände und Augen frei
- Natürliche Eingabe
 - Bedienung durch Laien
- Hohe Bandbreite
 - 1000 - 4000 Zeichen pro Minute
(Tatstatur 200 – 1000)

Nachteile:

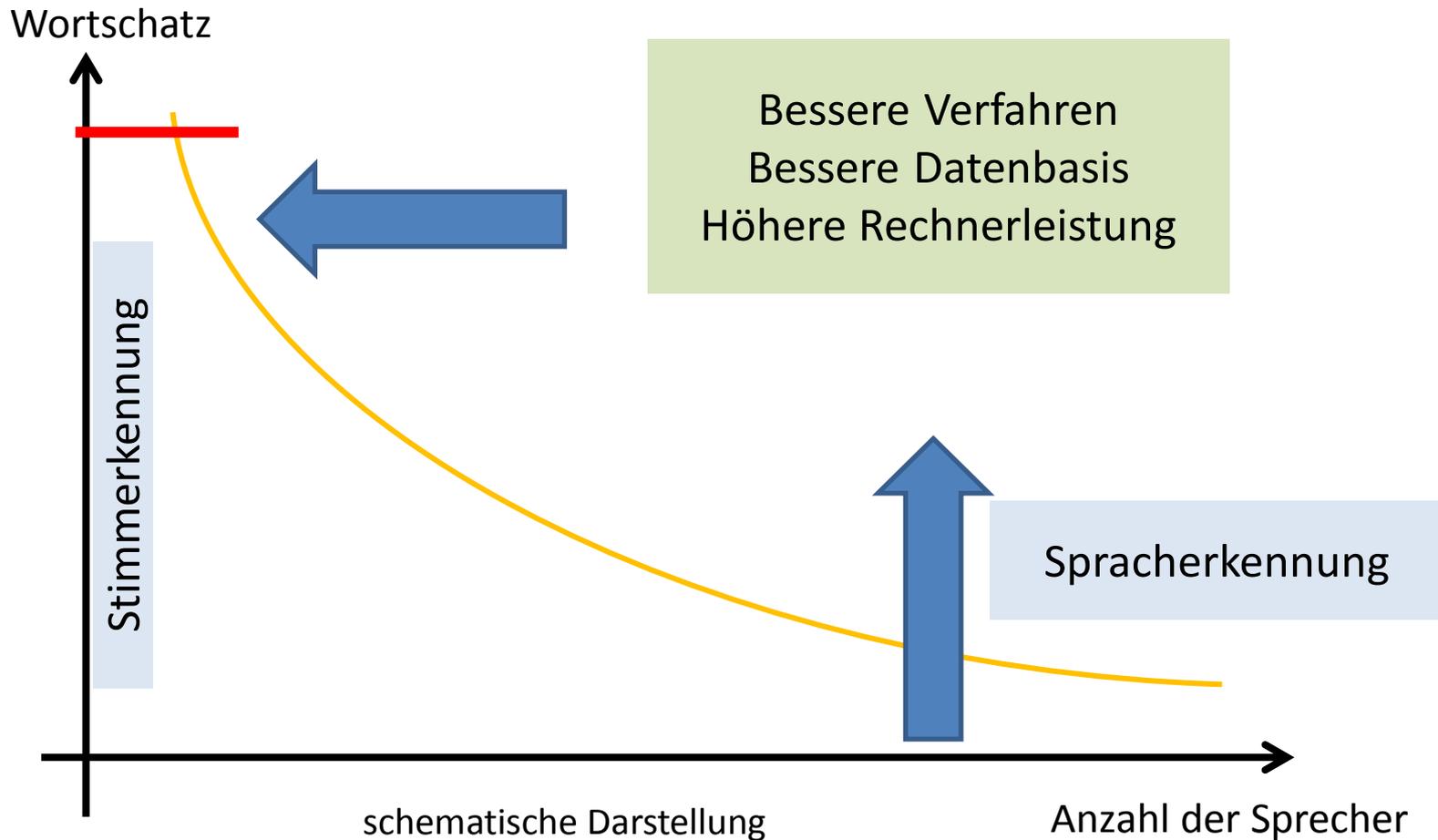
- Akustische Umgebung
 - Mehrere Sprecher gleichzeitig
 - Störgeräusche
- Technik / Mikrophone
 - Bedienung durch Laien
- Ungenügende Erkennungsleistung
 - Qualität wird besser
 - Fehlerbehandlung mit einplanen

Taxonomie von Sprachen

- Einzelne Kommandos
- Künstliche Sprachen
- Natürliche Sprachen



Spracherkennung vs. Stimmerkennung



Anwendungen (1)

Mensch-Maschine-Schnittstelle

- Spracherkennung
 - Diktiersysteme
 - Call-Center-Anwendungen (z.B. Verkauf von Kinokarten)
 - Auskunftssysteme (z.B. Fahrplanauskunft)
 - E-Commerce
 - Geschäfte über Mobil-Telefone
- Sprachsynthese
 - Je nach Komplexität der Interaktion
- Maschinelle Übersetzung geschriebener und gesprochener Sprache
 - Unterstützung von Human-Übersetzern
 - Domänenspezifische Übersetzungen
- Stimmerkennung in Kriminologie und Forensik

Hilfsmittel für Behinderte

- Vorlesegeräte für Blinde
- Fernsehen für Hörgeschädigte

Quelle: <https://hpsg.fu-berlin.de/~stefan/PS/cl-slides.pdf>

Anwendungen (2)

- Unterstützung bei der Textverarbeitung
(z.B. automatische Korrektur von Tipp- und Rechtschreibfehlern, Prüfung auf grammatische Richtigkeit)
- Unterstützung von Autoren beim Verfassen von Texten
(z.B. Finden des treffenden Ausdrucks oder der richtigen Terminologie, etwa bei der Verwendung eines kontrollierten Vokabulars in der technischen Dokumentation)
- Sprachsteuerung technischer Geräte oder Computer
- Automatisierte Messung von persönlichen Stärken anhand natürlicher Gespräche wie offene Interviews, Bewerbungsgespräche, Talkshows, Podiumsdiskussionen oder Gruppendiskussionen

Quelle: <https://de.wikipedia.org/wiki/Computerlinguistik>

Software (1)

- Große Auswahl an kostenpflichtiger sowie gebührenfreier Spracherkennungssoftware
(z.B. Dragon Professional, Testsieger 2016 laut www.netzsieger.de)
- Kaum Unterschiede bei der Qualität der Spracherkennung
- Lernfähig – erfordert intensives Training
- Wichtig für hohe Erkennungsrate:
 - Akzentfreies Sprechen des Users
 - Leistungsfähiger PC
 - Gutes Headset/Mikrofon
- Unterschiede in den Zusatzfeatures

Quelle: <https://www.netzsieger.de/k/spracherkennungssoftware>

Software (2)

- Siri (iOS 9):
 - Verwendung nur online möglich
 - „Werbung“: <https://www.apple.com/de/ios/siri>
- Cortana (Windows 10)
 - In deutscher Sprache, aber nicht mit Ländereinstellung Österreich verfügbar
 - <http://windows.microsoft.com/de-de/windows-10/getstarted-what-is-cortana>

Links

- Sprachsignale
<http://culturmag.de/litmag/siri-ously-neues-aus-der-computerlinguistik/55432>
- Wenn Computer Emotionen verstehen lernen
<http://www.welt.de/wirtschaft/webwelt/article3905112/Wenn-Computer-Emotionen-verstehen-lernen.html>
- McGurk - Effekt
<https://www.youtube.com/watch?v=jtsfidRq2tw>

Danke für die Aufmerksamkeit

Fragen?