

# Einführung in die Linguistik

## Butt & Co.

Do. 12:15 - 13:45

Fr. 12:15 - 13:45

Infos etc.

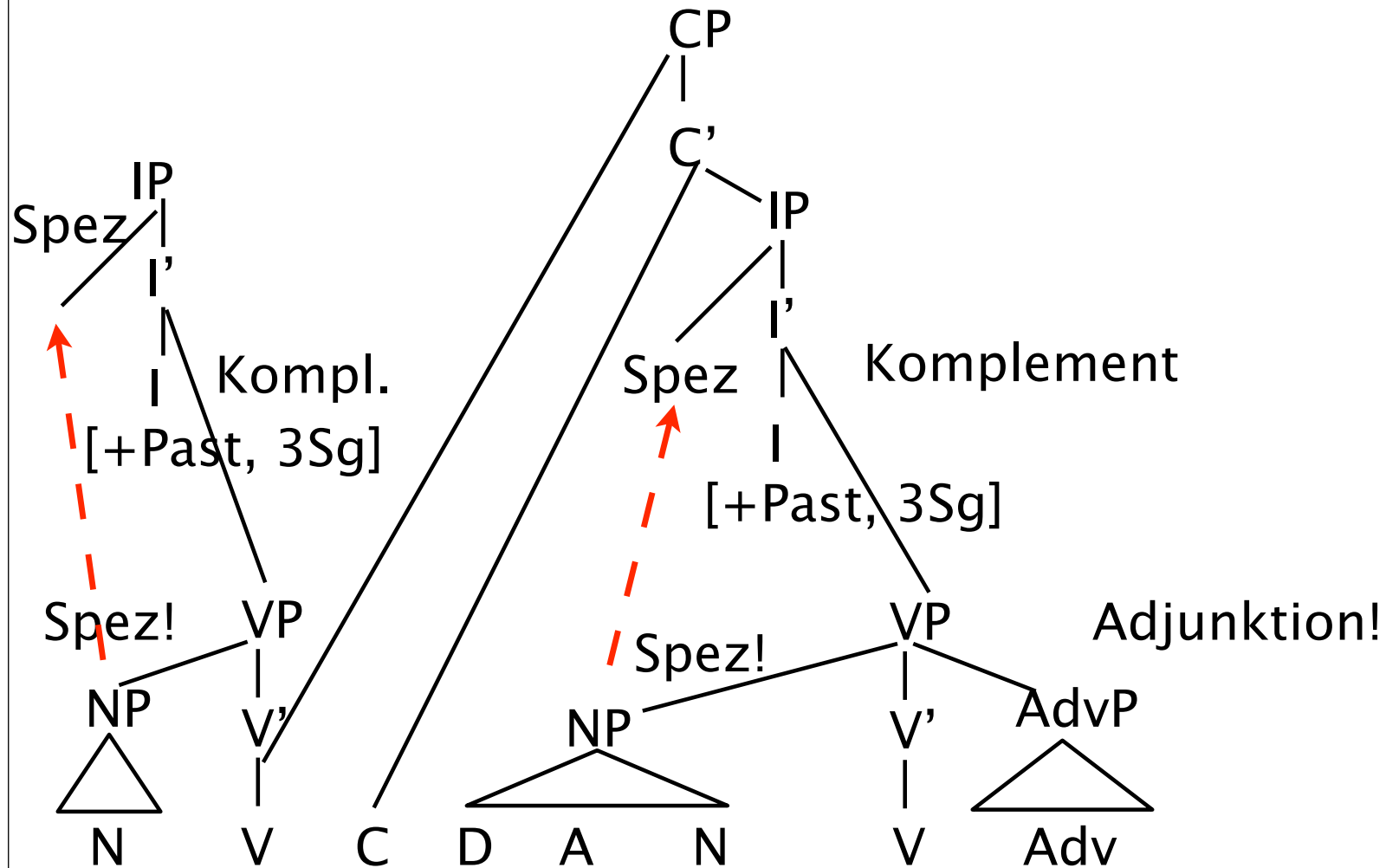
<http://ling.uni-konstanz.de>

=> Studium

=> Lehrveranstaltungen

=> Einführung in die Linguistik

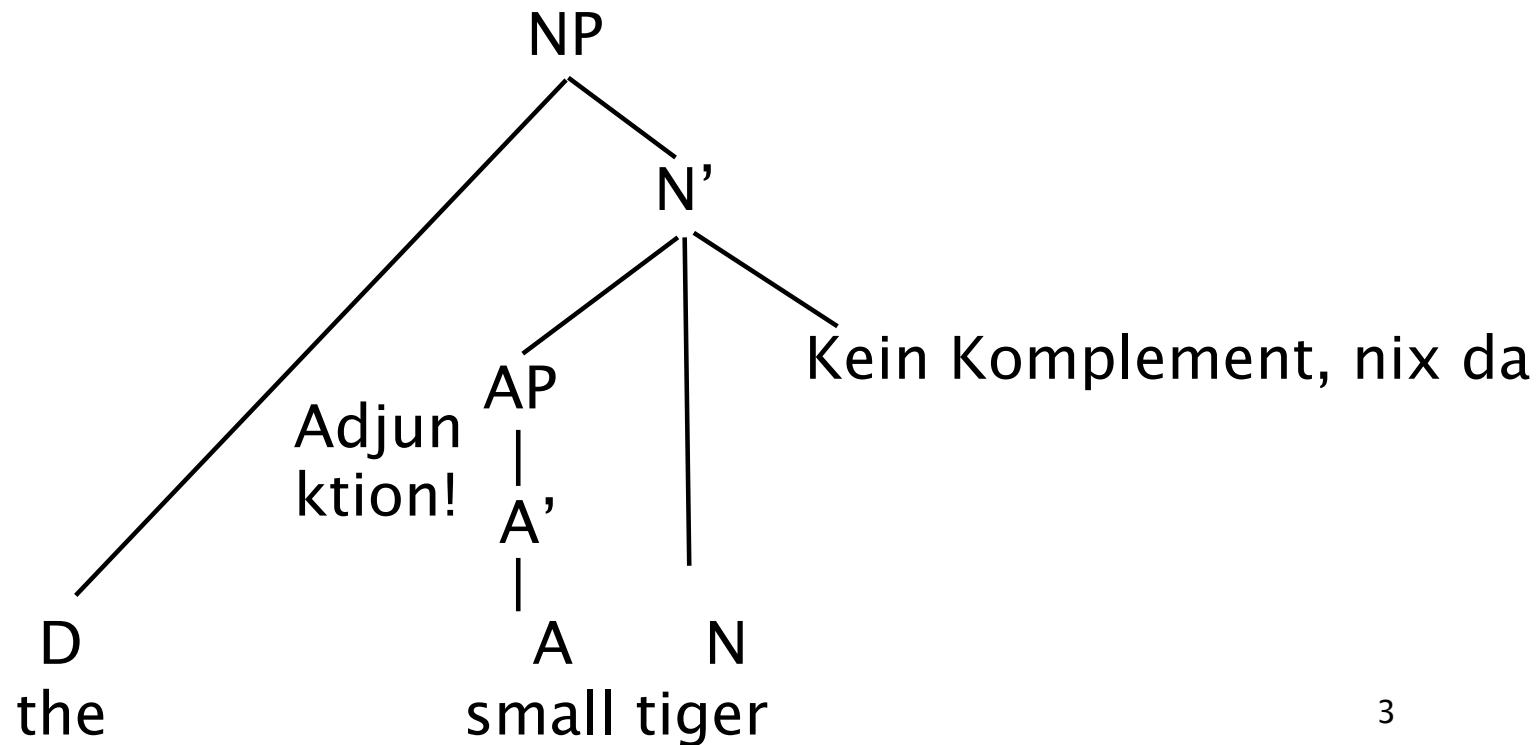
# Wiederholung X'

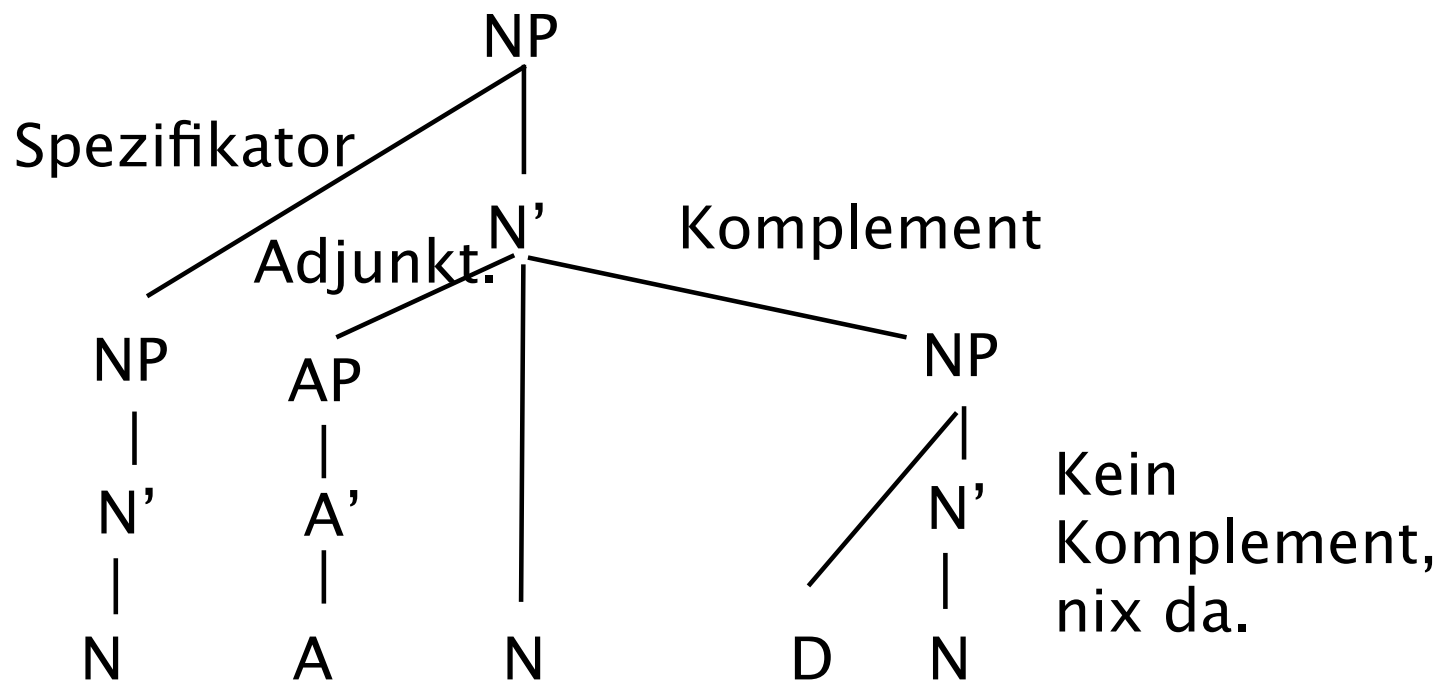


She say that the small tiger laugh loudly.

She said that the small tiger laughed.

Projizieren können: N, V, I, A, Adv, P, C  
Folgendes projiziert (erstmal) nicht: D





Lisas gutes Verständnis der Syntax

# Computerlinguistik

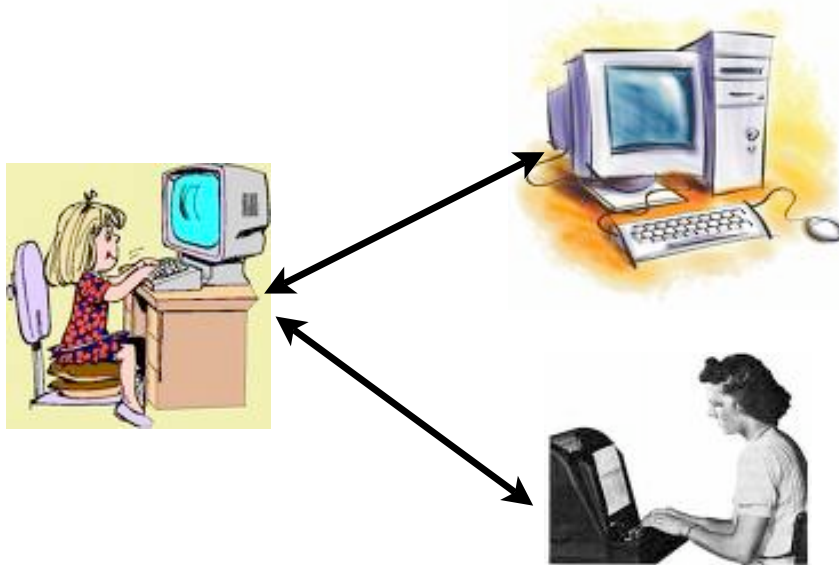
# Intelligente Maschinen?

## **Turing Test:**

Maschine bauen, die im Kommunikationsverhalten nicht vom Menschen zu unterscheiden ist.



(Alan Turing 1950)

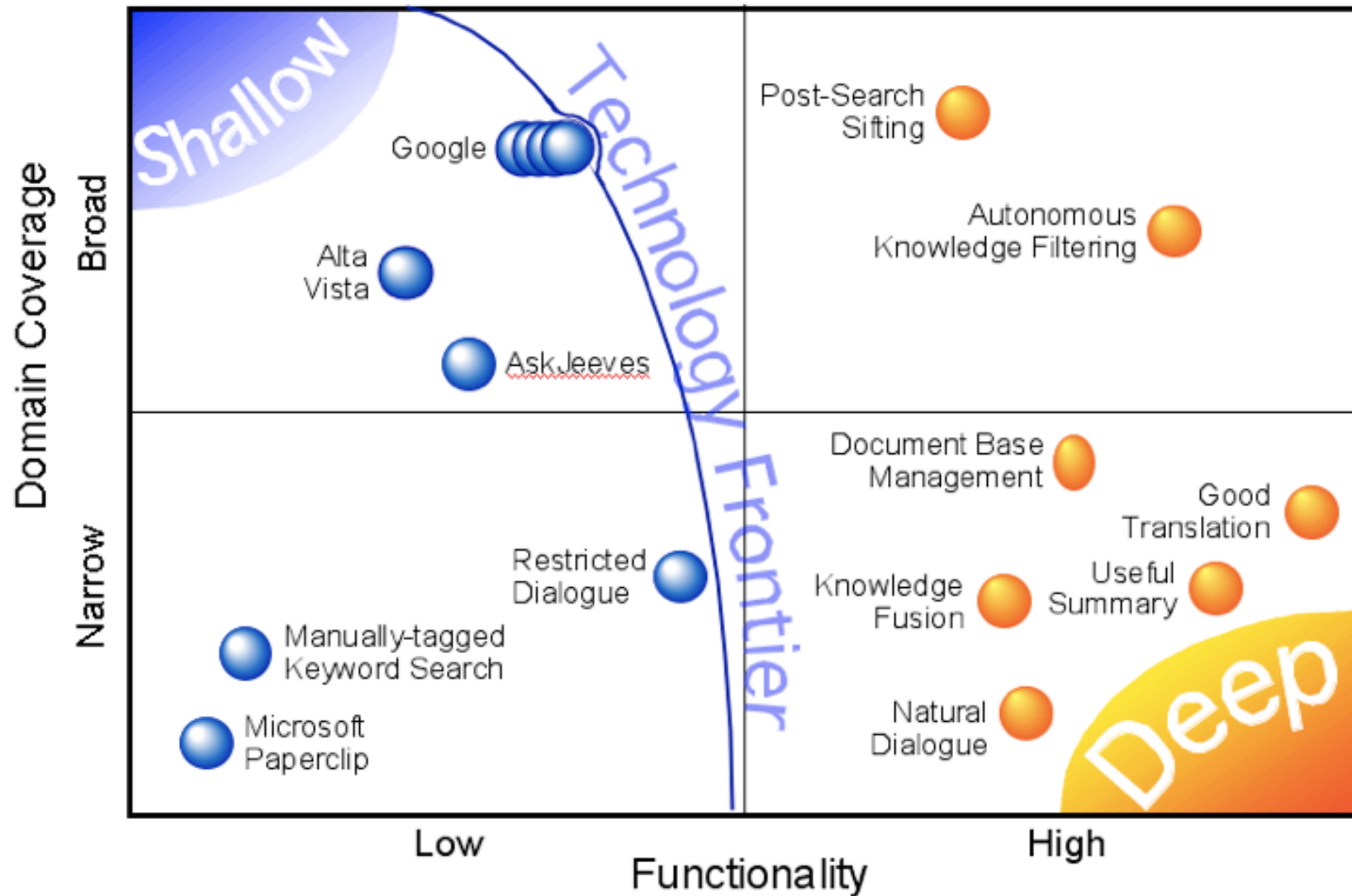



# Intelligente Maschinen?

## **Code knacken = Maschinelle Übersetzung???**

- Computer wurden im 2. Weltkrieg erfolgreich zum Code knacken eingesetzt.
- Alan Turing war einer der sehr bekannten Code Knacker.
- Idee/Erwartung: Sprachen zu übersetzen ist wie Code Knacken, müsste also schnell machbar sein.

# Applications of Language Engineering



 [http://www.google.de/language\\_tools?hl=en](http://www.google.de/language_tools?hl=en)[Web](#) [Images](#) [Videos](#) [Maps](#) [Shopping](#) [Groups](#) [Mail](#) [more](#) ▼

# Google

## Language Tools

### Search across languages

Type a search phrase in your own language to easily find pages in another language. We'll translate the results for you to read.

Search for:

My language:  Search pages written in:

Tip: Use [advanced search](#) to restrict your search by language and country without translating your search phrase.

### Translate text

Hans fliegt über Paris.

### Translate a web page

# Typische Übersetzung

## Rund und gesund

Ein gemütliches kleines Bäuchlein ist gesund. Wer seinen Rettungsring und das Hüftgold pflegt, kann sich entspannt zurücklehnen und auch weiterhin herzhaft zubeißen. Die Botschaft von Ärzten und Forschern ist schließlich eindeutig: Menschen mit leichtem bis mittlerem Übergewicht leben länger und werden seltener krank als ihre dünnen Zeitgenossen.

## Approximately and healthy

A cosy small belly flax is healthy. Who maintains its Rettungsring and the Hüftgold, can back-lean relaxed and also further heartfelt zubeissen. The message of physicians and researchers is finally clear: Humans with easy to middle predominance live longer and become more rarely ill than their dry contemporaries.

# Typische Übersetzung

## Executives Check In At Cheaper Hotels

As the recession unfolds, a growing number of budget-minded business travellers are shifting to lower price hotels, whether by choice or because their bosses are telling them to.

## Hauptleiter überprüfen sich innen in den preiswerteren Hotels

Während die Rezession ausbreitet, verschieben sich eine wachsende Anzahl von Etat-gekümmerten Geschäftsreisenden auf Hotels des niedrigeren Preises, ob durch Wahl oder, weil ihre Chefs ihnen zu erklären.

## Winnie The Pooh auf Babelfisch und Google

Seventy years ago a certain chubby **stuffed** bear came downstairs "**Bump, Bump, Bump,**" on the back of his head, **behind** Christopher Robin.

Vor siebzig Jahren kam ein bestimmter **chubby** **angefüllter** Bär unten "**Stoß, Stoß, Stoß,**" auf der Rückseite seines Kopfes, **hinter Christopher vor Robin.**

Vor siebzig Jahren eine gewisse **chubby** ausgestopften Bären kam die Treppe herunter "Bump, Bump, Bump," auf der Rückseite des Kopfes, hinter Christopher Robin.

## Einfache Beispiele: Google

Hans fliegt über Paris.

Google™



Hans flew over Paris.

Hans is flying via Paris.

The dog is jumping over the fence.

Der Hund springt über den Zaun.

Google™



The dog jumps over the fence.

Die Lampe hängt über dem Tisch.

Google™



The lamp hangs over the table.

The lamp hangs above the table.

## Einfache Beispiele: Google

Hans fährt gerade nach Paris.

Hans drives straight to Paris  
(Sommer 2009).

Hans just going to Paris. (Jetzt)

Hans is driving to Paris.

John likes to swim. → John mag schwimmen

John may swim. → John kann schwimmen.

John can swim.

John schwimmt gerne.

# Code Knacken = MÜ???

- Es gab einen “Code” der im 2. Weltkrieg nicht geknackt wurde.
- Das war Navajo, eine natürliche Sprache.
- Sprachen scheinen also ein ganz anderes Problem darzustellen, als ein künstliches erfundenes Verschlüsselungssystem.

# Intelligente Maschinen?

## **Turing Test:**

Maschine bauen, die im Kommunikationsverhalten nicht vom Menschen zu unterscheiden ist.

Bis jetzt haben wir es noch nicht geschafft.

# Sprache

Was ist an Sprache so schwierig?

**Kontext:** Worte haben oft unterschiedliche Bedeutungen, sie sind *ambig*.

Die Ambiguität kann nur durch Wissen über den Kontext aufgelöst werden.

Um den Kontext eines Satzes zu verstehen, brauchen wir **Weltwissen**.

# Weltwissen

Weltwissen wird durch Ontologien oder einer Ansammlung von Definitionen modelliert.

Bis jetzt: mühsames manuelles Zusammentragen und Kodierung von Information, Z.B WordNet.  
(<http://wordnet.princeton.edu/>)

# Sprache

Was ist an Sprache so schwierig?

**Struktur:** Sprachen unterliegen bestimmten Regeln, aber Sprachen unterscheiden sich systematisch von einander.

Zum Beispiel Wortstellung:

English: SVO

Deutsch: SOV (in Nebensätzen)

Irish Gaelisch: VSO

# Morphology und Syntax

- Um einen Satz maschinell zu bearbeiten muss er auf mehreren Ebenen analysiert werden.
- Bevor ein System ihn “verstehen” kann, muss klar sein, was die Morphologie des Satzes ist.
- Dann muss klar sein, was die Syntax eines Satzes ist.

# Wie funktionieren Babelfisch und Co?

- Hauptsächlich Wort für Wort Übersetzung
- Kein/Wenig Wissen über Kontext.
- Wenig Wissen über die Struktur der Sprachen.
- Riesige* Datenbank zu **Übersetzungspaaren.**

# Sprachmodellierung

Gute Sprachmodellierung kann zu besseren Applikationen führen.

Im Augenblick versucht man bessere **Frage-Antwort** Systeme zu bauen.

Die Universität Konstanz kooperiert mit der Firma *PowerSet*.

A dark blue rounded rectangular button with a thin black border, containing the text "QA Demo" in white.

QA Demo