

Einführung in die Korpuslinguistik

Beispielsätze:

„Der Junge besteigt den Zug, denn er fährt nach Hause.“

„Ihre Ideen sind lachsrot, isolierverglast und ungefähr 3 cm lang.“

„Peter bekommt von seiner Mutter das Lügen verboten.“

„Frieder läßt Eva Matthias einen Kuss geben.“

„Adam läßt Eva Monika Peter Paul ein Schwein schlachten lassen lassen lassen.“

(Erben, Johannes (1998): Zum Erscheinen der IDS-Grammatik. In: Sprachwissenschaft 23, 367-380, 369f.)

Was ist eigentlich ein Beispiel / ein Beispielsatz und was soll es / er leisten?

Beispiel: " ein beliebig herausgegriffener, typischer Einzelfall (als Erklärung für eine bestimmte Erscheinung od. einen bestimmten Einzelfall" (Duden 1999, 2, 510).

Sind selbstkonstruierte Beispiele sinnvoll?

Woher bekommen Linguisten/Linguistinnen eigentlich die Daten, die sie zur Überprüfung ihrer Theorien/Hypothesen brauchen?

- Introspektion
- Informantenbefragung
- Datenerhebungen (Belegsammlung)
- Korpora

Korpus: feste Sammlung von Texten, die als Ganzes untersucht und mit statistischen Methoden befragt wird.

- kann nach bestimmten Kriterien gesammelt werden oder auch opportunistisch (man nimmt, was man bekommt)
- kann riesig groß sein (>200 m Wörter) oder ganz klein
- kann auf allen linguistischen Ebenen annotiert sein (Wortart, Semantik, Syntax, ...)

Ziel einer Korpusanalyse:

- Was wird tatsächlich verwendet?
- Was ist in der ‚Norm‘ zu finden?
- Wie häufig wird etwas benutzt, d.h. wie stark ist die funktionale Belastung?
- Wie kann man die Daten analysieren?

Linguistische Daten

Introspektion	Korpusdaten
Kompetenz: was ist grammatisch?	Performanz: was kommt vor?
Produktionssystem, das alle grammatischen Äußerungen einer Sprache hervorbringt	Modell, das die Phänomene und Verteilungen innerhalb eines bestimmten Korpus beschreibt
qualitativ (kategorial)	qualitativ + quantitativ (probabilistisch)

Welche Daten geeignet sind, folgt aus einer möglichst präzise gestellten Forschungsfrage!