

**Barbara Schlücker & Maik Walter**

Freie Universität Berlin

Institut für Deutsche und Niederländische Philologie

CBS Kopenhagen

Sokrates-Blockseminar

Oktober 2007

## HS Korpuslinguistik

Unter Korpora versteht man in der Linguistik Zusammenstellungen von sprachlichen Daten, die bezüglich eines Auswahlkriteriums gesammelt wurden. Durch die Auswertung der Korpora will man Aufschlüsse über den tatsächlichen Sprachgebrauch in bestimmten Domänen erhalten. Beispielsweise unterscheidet sich die Sprache der Politik ganz erheblich von der Sprache der Wirtschaft: Nicht nur die verwendete Lexik differiert hier, sondern auch die lexiko-grammatischen Muster.

Im ersten Teil des Seminars werden ausgewählte Korpora der deutschen Sprache vorgestellt. In Übungssequenzen wird der Umgang mit diesen Ressourcen an linguistischen Fragen eingeübt. Im zweiten Teil wird die Kompilation eigener Korpora (so genannter small corpora) thematisiert. Dazu werden Tools vorgestellt, die es erlauben, auf effiziente Weise Korpora zusammenzustellen und auszuwerten.

### Literatur:

Lothar Lemnitzer & Heike Zinsmeister (2006): Korpuslinguistik. Eine Einführung. Tübingen: Narr (Narr Studienbücher)

Mohsen Ghadessy, Alex Henry & Robert L. Roseberry (eds., 2001): Small corpus studies: theory and practice. Amsterdam: Benjamins (Studies in corpus linguistics)

## Seminarplanung

Tag	Thema	Literatur zur Vor- und Nachbereitung
Dienstag, 02.10.07 11.40-15.10 (inkl. 2 UE)	1. Korpuslinguistik und Korpora: Eine neue Perspektive auf die (deutsche) Sprache [+ Sprachstandsfeststellung]	Lemnitzer/Zinsmeister 2006: 5-59
Donnerstag, 04.10.07 11.40-15.10 (inkl. 2 UE)	2. Lexikalische Analysen (am Beispiel des Wortschatzportals der Universität Leipzig) 3. Grammatische Analysen (am Beispiel des DWDS)	Lemnitzer/Zinsmeister 2006: 60-100
Montag, 08.10.07 13.30-15.10	4. Lernerkorpora & mehrsprachige Korpora	Lemnitzer/Zinsmeister 2006: 127-174
Dienstag, 09.10.07 9.50-10.35 11.40-12.25	5. Die Kompilation von (eigenen) Korpora	Lemnitzer/Zinsmeister 2006: 101-126
Donnerstag, 11.10.07 11.40-15.10 (inkl. 2 UE)	6. Die Auswertung eigener Korpora (am Beispiel von TextStat) 7. Deutschsprachige Korpora: Ein Fazit	Lemnitzer/Zinsmeister 2006: 40-59, 101-126

### Literatur:

Die Basisliteratur ist Lothar Lemnitzer & Heike Zinsmeister (2006): Korpuslinguistik. Eine Einführung. Tübingen: Narr (Narr Studienbücher). Dieses Buch wird auch zur Anschaffung empfohlen. Dort findet sich eine große Anzahl von weiteren Literaturhinweisen.

### Weitere Literatur:

- Albert, R. / Koster, C.J. (2002): Empirie in Linguistik und Sprachlehrforschung. Ein methodologisches Arbeitsbuch. Tübingen: Narr
- Biber, D. / Conrad, S. / Reppen, R. (1998): Corpus linguistics: investigating language structure and use. Cambridge: CUP
- Breindl, E. (1997): "DaF Goes Internet! Neue Entwicklungen in Deutsch als Fremdsprache". In: Deutsche Sprache, 25, 289-342
- Burnard, L./ McEnery, T. (eds.) (2000): Rethinking language pedagogy from a corpus perspective. Frankfurt am Main: Peter Lang

- Carstensen, K.-U. et al. (eds., 2004): Computerlinguistik und Sprachtechnologie. Eine Einführung. Heidelberg: Spektrum Akademischer Verlag
- Donath, R. (Ed.): Kreative Textarbeit mit dem PC. Stuttgart u.a. : Klett
- Fillmore, C. (1992): Corpus linguistics vs. computer-aided armchair linguistics. In: Svartvik, J. (ed., 1992), Directions in Corpus Linguistics. Berlin: Mouton de Gruyter, 35 - 60
- Ghadessy, M, / Henry, A. / Roseberry, R. (eds.) (2000): Small Corpus Studies and ELT. Theory and prac-tice. Amsterdam/Philadelphia: Benjamins
- Helbig, G. / Buscha, J. (2003): Leitfaden der deutschen Grammatik. Berlin u. a.: Langenscheidt
- Hunston, S. (2002): Corpora in Applied Linguistics. Cambridge: CUP
- Kennedy, G. (1998): An introduction to corpus linguistics. London / New York: Longman
- Kettemann, B. / Marko, G. (eds.) (2002): Teaching and Learning by Doing Corpus Analysis, Amsterdam: Rodopi
- Leech, G. & Eyes, E. (1997): Syntactic Annotation: Treebanks. In: Garside et al. (eds.): Corpus Annotation: Linguistic Information from Computer Text Corpora. New York: Addison Wesley Longman, 34 - 52
- Lemnitzer, L. / Zinsmeister, H. (2006): Korpuslinguistik: Eine Einführung. Tübingen: Narr
- Linke, A. et al. (2004): Studienbuch Linguistik. 5., erw. Auflage. Tübingen: Niemeyer
- McEnery, T. / Wilson, A. (1996): Corpus linguistics. Edinburgh: Edinburgh University Press
- Meibauer, J. et al. (2002): Einführung in die germanistische Linguistik. Stuttgart: Metzler
- Mukherjee, J. (2002): Korpuslinguistik und Englischunterricht: Eine Einführung. Frankfurt a.M.: Peter Lang
- Partington, A. (1998): Patterns and Meanings. Using Corpora for English Language Research and Teaching. Amsterdam / Philadelphia: Benjamins
- Porte, G. K. (2002): Appraising Research in Second Language Learning. A practical approach to critical analysis of quantitative research. Amsterdam /New York: Benjamins
- Scherer, C. (2006): Korpuslinguistik. Heidelberg: Winter

## Zwei mögliche Ausgangspunkte für weitergehende WWW-Recherchen

- <http://www2.rz.hu-berlin.de/korpling/links/>
- <http://www.lemnitzer.de/lothar/KoLi>