

## **COGS 523 Spring 2009**

### **Take-Home Final Exam**

#### **Due Date: 26th May, Class Hour**

*Please type your answers, and write clearly within the norms of academic writing, elaborating your points and justifying and exemplifying them where necessary.*

1. (40 pts) Read the Kreyer (2003) article provided as an additional file in this exam. Abstract away and indicate the steps of the corpus-based research methods used in the article. Evaluate the research presented with respect to corpus use related aspects (Note: I do not want you to go into the subject matter of the article any more than what is presented in the article; rather indicate the positive and negative aspects of corpus-related aspects of the research presented).
2. (20 pts) The case studies we reviewed in this course were mostly taking a corpus-based approach. A more radical stance in corpus linguistics is the “corpus-driven” approach as described by Tognini-Bonelli, trying to derive linguistic categories systematically from the recurrent patterns and frequency distributions that emerge from language in context. One distinguishing feature of this stance is to use such big corpora that there should be no serious effort for representativeness and balance; the sheer size of the big corpus will achieve the so-called cumulative representativeness. Another feature is the objections to any corpus annotation as that would bind the linguists to preconceived theories whereas they should aim of postulating linguistic categories entirely based on corpus data (McEnery and Wilson, 2001). Discuss the feasibility and possible objections to such an approach based on your experience of the material of this course.
3. (10 pts) What are the advantages and disadvantages of standard representations (such as TEI guidelines) for encoding information in corpora?
4. (Inspired from Forum Questions, 15 pts each)
  - a. How can we assure accuracy in analyzing a corpus?
  - b. One may feel that the idiomatic expressions in Turkish are more flexible than they are in English. How would you design a corpus-based research project based on this idea? Clarify and justify your steps.