This course is an introduction to computerized research methods, which are applied to large
data bases of language used in natural communicative settings to supplement more traditional
ways of linguistic analysis in all linguistic sub-disciplines.

**COURSE CREDIT REQUIREMENTS**

To meet the course credit requirements for this class, you are expected to

- attend class regularly
- come to class prepared, that is, do the required readings
- hand in a summary of a corpus-linguistic paper of your choice that appeared in a journal
  like, for instance, the *ICAME* journal, *Corpus Linguistics and Linguistic Theory, International
  Journal of Corpus Linguistics, Computers and the Humanities, Literary and Linguistic
  Computing, Corpora, or Computational Linguistics*, and that is not among the papers we
discuss in class (cf. below); the length of this summary should be no shorter than 5 pages,
double-spaced, 12 pt. font of running text (i.e., excluding a cover page, tables, figures, and
the like). The summary is due 1 December in class.
- pass a midterm exam that covers the contents of the class to that date;
- pass a final exam that covers the contents of the complete class.

For both the midterm exam and the final exam, you are allowed to bring all course materials,
readings, notes, etc.

**GRADING**

Your final grade will depend on the number of points you score. You can get the maximum of
100 points if you

- attend all classes for a maximum of 45 points (each unexcused absence reduces your score
  by 2.5 points)
- submit the article summary for a maximum of 15 points (each late-day reduces your score
  by 2.5 points)
- pass the midterm exam for a maximum of 15 points
- pass the final exam for a maximum of 25 points

Should you not be satisfied with your performance on either exam or the grade for the article
summary, you can score a maximum of 10 additional points by handing in a special written
assignment (feel free to talk to me about that anytime).

**CONTACT**

Office hours: Friday 2-3pm in SH 3523 or by appointment
Web: http://www.linguistics.ucsb.edu.faculty/swulff
Email: swulff@linguistics.ucsb.edu
### THE BASICS
1. **Sep 29** Introduction
2. **Oct 1** Online interfaces I: MICASE online
3. **Oct 6** Online interfaces II: BYU-CAE

### HOW TO COMPILE A CORPUS
4. **Oct 8** Compiling corpora / The web as corpus  
5. **Oct 13** Corpus transcription and annotation  
   Prep reading: Atkins et al. (1992)

### CORPUS METHODS: ROUGH AND READY
6. **Oct 15** Frequency lists (Introduction to AntConc)  
   Prep reading: Leech and Fallon (1992)
7. **Oct 20** Concordances  
8. **Oct 22** Collocations  
   Prep reading: Biber et al. (1996: 21-54)
9. **Oct 27** Patterns (Introduction to KfNgram)  
   Prep reading: Francis (1993)
10. **Oct 29** Collostructions  
11. **Nov 3** MIDTERM EXAM

### CORPUS METHODS: THE R-EAL DEAL
12. **Nov 5** R: Functions, arguments, vectors  
    Prep reading: Gries (2009: 72-84)
13. **Nov 10** R: Simple text processing  
    Prep reading: Gries (2009: 84-105)
14. **Nov 12** R: Advanced text processing  
15. **Nov 17** R: Simple frequency lists  
    Prep reading: Gries (2009: 120-123)
16. **Nov 19** R: Cultural word frequency comparisons  
    Prep reading: Gries (2009: 63-65; 131-137)
17. **Nov 24** R: Simple concordances
18. **Nov 26** R: Complex concordances  
    Prep reading: Gries (2009: 143-147)

### IT'S ALL DOWNHILL FROM HERE...
19. **Dec 1** Make a Wish! ARTICLE SUMMARY DUE TODAY
20. **Dec 3** Revision and course evaluation  
    Dec 12 12-3 pm: FINAL EXAM
REQUIRED READINGS


SUGGESTIONS FOR FURTHER READINGS (NOT OBLIGATORY)


