

Language and Personality, Fall 2004
A Semi-Syllabus, more a vague roadmap really

Aug 30	LIWC: How it works. The nature of dictionaries and arithmetic. Defining personality. Text archives.	
Sept 13	Class projects with LIWC	LIWC manual; Pennebaker & King
Sept 20	Psychometrics of words. Linking individual differences to language use – sex and age	Pennebaker & Stone; Newman, Groom et al
Sept 27	Benjamin Whorff	Whorff; Slobin
Oct 4	Culture and friends	Wierzbicka
Oct 11	Culture and class projects	Wierzbicka
Oct 18	Nouns, verbs, parsing and tagging	Nisbett
Oct 29	Text analyses in the social sciences: An overview. What is the right level of analysis?	Pennebaker, Mehl, & Niederhoffer
Nov 1	LSA, factor analysis, and other ways to think about words	Campbell & Pennebaker
Nov 8	Brains, bodies, hormones	Heberlein et al Schervitz Pennebaker, Groom et al
Nov 15	Pronouns I	Brown
Nov 22	Pronouns II	tba
Nov 29	More function words: Articles and Prepositions	tba

Campbell, R.S. & Pennebaker, J.W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*, 60-65.

Cohn, M.A., Mehl, M.R., & Pennebaker, J.W. (in press). Linguistic Markers of Psychological Change Surrounding September 11, 2001. *Psychological Science*.

Heberlein, A.S., Adolphs, R., Pennebaker, J.W., & Tranel, D. (2003). Effects of damage to right-hemisphere brain structures on spontaneous emotional and social judgments. *Political Psychology, 24*, 705-726.

Newman, M., Groom, C., Stone, L.D., & Pennebaker, J.W. (2004). Sex differences in language use: An analysis of text samples from 70 studies. Manuscript submitted for publication.
CD Filename: SexDiff.doc

Nisbett, R. (2003). *The geography of thought* (or the 4-page newspaper summary).

Pennebaker, J.W., Groom, C.J., Loew, D., & Dabbs, J. (2004). Testosterone as a social inhibitor: Two case studies of the effect of testosterone treatment on language. *Journal of Abnormal Psychology, 113*, 172-175.

Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Erlbaum Publishers (www.erlbaum.com).

- Pennebaker, J.W. & King, L.A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77, 1296-1312.
- Pennebaker, J.W., Mehl, M.R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54, 547-577.
- Pennebaker, J.W., & Stone, L.D. (2003). Words of wisdom: Language use over the lifespan. *Journal of Personality and Social Psychology*, 85, 291-301.
- Slobin, D. (1996). From “thought and language” to “thinking for speaking.” In J.J. Gumpers and C. Levinson (Eds.), *Rethinking linguistic relativity* (pps, 70-96). Cambridge, MA: Cambridge University Press.
- Wierzbicka, Anna (1997). *Understanding cultures through their key words*. New York: Oxford University Press
- Whorff (1942) reading: (From Language, Mind, and Reality)
<http://sloan.stanford.edu/mousesite/Secondary/ThoughtReality.htm>

Articles and Data Sets for the Class

This is an odd class in many ways. In addition to mastering LIWC and learning about other text analytic procedures, this course provides indepth training in data mining. Over the years, I have amassed a remarkable text archive which now has over 300,000 text files. In this class, you are invited to analyze whatever text files that I have at my disposal. Should you get some interesting results, I hope you publish them.

A complicated matter with data sets like this involves ownership. Most of the data sets have been collected by me or people in my lab. Others involved collaborations with others. Based on previous experience, I doubt that most other authors would have problems with your publishing results based on their data. However, you will need their permission. Some may ask to be an author which should not be a problem.

A related issue concerns confidentiality. I have tried to remove identifying information from the text files. However, this may not always have been possible. Please consider all text files (not in the public domain) as confidential research documents. Do not share the raw text files, LIWC program files, or unpublished papers with anyone not in this class without my permission.

Authors	Nickname	Sample info	Other info
Cohn, Mehl, & Pennebaker	livejournal	About 71,000 diary entries by 1,081 people in 3 mo prior and post 9/11	age, sex, region of U.S.
Davison, Pennebaker, & Dickerson	support group	About 4,000 bulletin board posts among people in 6-7 illness groups	illness group, sex, person with disease or friend
Gortner & Pennebaker	A&M	about 500 newspaper articles in UT and A&M papers surrounding Bonfire disaster	About 800 bulletin board memorial posts. Includes some sex and age info
Groom & Pennebaker	personals	about 1600 online personals entries: males, females,	age, sex, ses, & misc info (preferences in mate

		gays, straights	selection)
Newman, Pennebaker, Berry, & Richards	deception	5 studies; people are asked to lie or tell the truth. Lab study. students	sex, truth/lie, position on topic
Niederhoffer & Pennebaker	Linguistic style matching	2 college student IM sessions. About 100. Also 16 Watergate tapes.	For students: age, sex, some attitude items. Some judge ratings
Pennebaker, Groom, Loew, & Dabbs	testosterone	2 case studies: FTM and older male. Diary or outgoing emails for a year	time and date
Pennebaker & King	Linguistic style	5 Intro psych samples, 100-500/year. Writing 4 times: SOC, college, TAT, and other topics	Pretesting, Big 5, SAT, grades in class, health records, other personality info
		5 years of summer school students (n=34) who wrote 10 times on different topics	sex, writing topic
		18 days of inpatient diaries for about 12 people in substance abuse clinic in England	sex, maybe age and diagnosis
		40 SESP members' with 15 journal abstracts	sex, year of publication
Pennebaker & Lay	Giuliani	35 press conferences	date, life circumstances of Giuliani
Pennebaker, Mayne, & Francis	Roshoman	2 interviews of 30 gay men who lost a lover to AIDS; 2 weeks and 1 year after	depression, age, maybe some other info
Pennebaker & Niederhoffer	Life	about 650 people who filled out a huge questionnaire online and described a picture	see LIFE questionnaire for items.
Pennebaker & Stone	Aging	Author data: complete published works of 10 famous novelists, poets, and playwrights	age when articles were published, nationality, sex, genre
Stirman & Pennebaker	Suicide	at least 15 poems from 14 poets who committed suicide or didn't	age, sex, nationality

Other data sets:

Disclosure studies: Dozens of data sets where people write about trauma or about superficial topics. Most studies include multiple other measures

Published work: several projects involving books, song lyrics, chapters of the bible, movie scripts

Natural conversations: real world interactions (EAR), "natural" lab interactions, live TV and radio transcripts.

Interviews: people being interviewed about childhood (AAI), issues concerning welfare and education