Downloaded By: [Newman, Matthew L.] At: 16:37 4 June 2008

TABLE 1
Main Effects of Gender on Language Use

		Female		Male		
LIWC Dimension	*Examples*	*M*	*SD*	*M*	*SD*	*Effect Size (d)*
Linguistic dimensions						
Word count		1,420	5,403	1,314	6,016	*ns*
Words per sentence		21.26	31.22	23.90	48.12	–0.07
Question marks		3.21	7.33	3.07	7.86	ns
Words ≥ six letters		13.99	4.42	15.25	5.91	–0.24
Numbers		1.37	1.31	1.59	1.55	–0.15
Negations	no, never, not	1.85	1.10	1.72	1.17	0.11
Articles	a, an, the	6.00	2.73	6.70	2.94	–0.24
Prepositions	on, to, from	12.46	2.44	12.88	2.64	–0.17
Inclusive words	with, and, include	6.42	1.88	6.34	2.03	*ns*
Exclusive words	but, except, without	3.82	1.54	3.77	1.64	*ns*
Psychological processes						
Emotions		4.57	1.99	4.35	2.07	0.11
Positive emotions	happy, pretty, good	2.49	1.34	2.41	1.40	*ns*
Optimism	certainty, pride, win	0.56	0.58	0.58	0.61	*ns*
Positive feelings	happy, joy	0.61	0.61	0.51	0.65	0.15
Negative emotions		2.05	1.65	1.89	1.56	0.10
Anxiety	nervous, afraid, tense	0.48	0.68	0.38	0.64	0.16
Sadness	grief, cry, sad	0.55	0.76	0.47	0.70	0.10
Anger	hate, kill	0.61	0.81	0.65	0.92	*ns*
Swear words	damn, ass, bitch	0.09	0.25	0.17	0.44	–0.22
Sensations		2.22	1.27	2.06	1.30	0.12
Feeling	touch, hold, feel	0.58	0.67	0.47	0.66	0.17
Hearing	heard, listen, sound	0.78	0.74	0.71	0.72	0.10
Seeing	view, saw, look	0.72	0.78	0.74	0.83	*ns*
Cognitive processes		7.35	2.57	7.17	2.82	0.07
Causation	effect, hence	1.02	0.76	1.02	0.88	*ns*
Insight	think, know	2.40	1.28	2.28	1.38	0.09
Discrepancy	should, would, could	2.32	1.31	2.23	1.46	0.07
Tentative	perhaps, guess	2.54	1.43	2.54	1.57	*ns*
Certaintyns	always, never	1.35	0.94	1.21	0.96	0.14
Hedge verb phrases	I + guess, I + reckon	0.57	0.67	0.50	0.67	0.11
Social processes						
Social words		9.54	4.92	8.51	4.72	0.21
Communication	talk, share, converse	1.26	0.95	1.20	0.95	*ns*
Friends	pal, buddy, coworker	0.37	0.51	0.33	0.53	0.09
Family	mom, brother, cousin	0.77	1.04	0.64	1.01	0.12
Humans	boy, woman, group	1.22	1.33	1.15	1.33	*ns*

(continued)

Downloaded By: [Newman, Matthew L.] At: 16:37 4 June 2008

TABLE 1 (*Continued*)

		Female		Male		
LIWC Dimension	*Examples*	*M*	*SD*	*M*	*SD*	*Effect Size (d)*
Pronouns		14.24	4.06	12.69	4.63	0.36
First-person singular	I, me, my	7.15	4.66	6.37	4.66	0.17
First-person plural	we, us, our	1.17	2.15	1.07	2.12	*ns*
Second person	you, you're	0.59	1.05	0.65	1.15	–0.06
Third person	she, their, them	3.41	3.45	2.74	3.01	0.20
Time and space						
Time	till, started, hour	4.09	1.94	4.03	2.14	*ns*
Past-tense verb	walked, were, had	4.36	2.97	4.02	2.84	0.12
Present-tense verb	walk, is, be	11.71	4.00	10.98	4.10	0.18
Future-tense verb	will, might, shall	1.27	1.03	1.33	1.18	*ns*
Space	here, up, around	2.40	1.18	2.47	1.31	*ns*
Motion verbs	walk, move, go	1.22	0.89	1.15	0.93	0.07
Current concerns						
Occupation	work, class, boss	2.34	1.88	2.59	2.10	–0.12
Money	cash, taxes, income	0.25	0.39	0.29	0.49	–0.10
Metaphysical	death, god	0.41	0.88	0.47	0.97	–0.06
Leisure	house, TV, music	1.11	0.92	1.07	1.08	*ns*
Home	house, kitchen, lawn	0.80	0.76	0.68	0.79	0.15
Sports	football, play, game	0.13	0.30	0.19	0.50	–0.15
Physical functions	ache, breast, sleep	1.33	1.22	1.28	1.29	*ns*
Sex	lust, pregnant, gay	0.30	0.53	0.27	0.60	*ns*

Note. Means (except for word count and words per sentence) refer to percentages of the total words in a sample. Effect size (Cohen's *d*) was calculated by dividing the mean difference by the pooled standard deviation. Positive effect sizes mean women used the category more; negative effect sizes mean men used it more. All mean differences except those labeled "*ns*" were significant at $p < .001$, based on univariate statistics from a multivariate analysis of variance.

1991). Based on this literature, we also predicted that the size of the gender difference would be largest in spoken language (i.e., the conversation category) because it is more natural and spontaneous (e.g., Biber, 1991).

Method

Text corpus.[1] Our archive of electronic text samples represented 70 studies from 22 laboratories. These laboratories included 14 universities in the United States (63 studies), 1 university in New Zealand (4 studies), and 3 universities in England (3 studies). Forty-four of the studies (63%) were conducted by at least one of the authors. The studies were conducted over a 22-year period (1980–2002), and included samples of fiction going back as far as the 17th century. All the files

[1]We are happy to make this corpus available to other researchers. Interested parties can contact James W. Pennebaker for more information at pennebaker@mail.utexas.edu