The Development of the Chinese Linguistic Inquiry and Word Count Dictionary

Chin-Lan Huang,[1] Cindy K. Chung,[2] Natalie Hui,[3] Yi-Cheng Lin,[4] Yi-Tai Seih,[2] Ben C. P. Lam,[3] Wei-Chuan Chen,[4] Michael H. Bond,[3] and James W. Pennebaker[2]

[1]Department of Humanities and Social Sciences, National Taiwan University of Science and Technology

[2]Department of Psychology, University of Texas at Austin

[3]Department of Psychology, The Chinese University of Hong Kong

[4]Department of Psychology, National Taiwan University

The analysis of natural language opens a window to the exploration of thoughts, feelings, and personalities. To analyze word use, Pennebaker and his colleagues (Pennebaker, 2011) developed a computer software, Linguistic Inquiry and Word Count (LIWC). LIWC reports on the percentage of words in a text file that are in the grammatical, psychological, and content categories of its dictionary, and was created by judges who rated whether each entry belonged in a category. This paper describes the development of the Chinese LIWC dictionary and establishes its reliability and validity. Study 1 involved the translation of terms from the English LIWC dictionary to Chinese, and judges' ratings of whether each entry belonged in a category. Furthermore, some categories unique to the Chinese language were added to the Chinese LIWC dictionary. These resulted in a total of approximately 6,800 words across 30 linguistic categories and 42 psychological categories. In Study 2, we analyzed one hundred texts and their translations using both the English LIWC dictionary and Chinese LIWC dictionary. Fifty texts of varied genres and authors were written in English and translated to Chinese; others were written in Chinese and translated to English. Reliable correlations were found between the English LIWC and Chinese LIWC categories, indicating acceptable equivalence between the two dictionaries. In Study 3, we analyzed 30 bulletin board messages from a site for people suffering from depression, and 30 messages from a site on a control topic (personal experiences with part-time jobs). Similar to what past researchers have found in English, depressed people used more first person singular pronouns, fewer first person plural pronouns, and more negative emotion words than non-depressed people, confirming the validity of the Chinese LIWC. Just as the English LIWC has led to discoveries in the social sciences, the Chinese LIWC now opens new windows to the psychology of Chinese speakers and authors by providing an efficient means by which to analyze open-ended texts. Future research, applications, and revisions of the Chinese LIWC dictionary are discussed.

Keywords: *expressive writing, LIWC (Linguistic Inquiry and Word Count), text analysis*