The Journal of General Psychology, 2007, *134*(4), 405–413

The Relations Between Personality and Language Use

CHANG H. LEE
Sogang University and Pusan National University

KYUNGIL KIM
Ajou University

YOUNG SEOK SEO
Konkuk University

CINDY K. CHUNG
University of Texas

ABSTRACT. The authors of this study provided basic descriptive data on the correlation between personality tests and Korean language use. Native Korean-speaking students (*N* = 80) at Pusan National University completed 2 personality tests, the Myers-Briggs Type Indicator (MBTI; I. B. Myers, M. H. McCaulley, N. L. Quenk, & A. L. Hammer, 1998) and the 5-Factor Inventory (O. P. John, E. M. Donahue, & R. L. Kentle, 1991). They completed a stream-of-consciousness essay, which the authors analyzed using the Korean version of Linguistic Inquiry and Word Count (KLIWC; C. H. Lee, J. Shim, & A. Yoon, 2005). Personality traits were significantly correlated with linguistic variables. Furthermore, the observed correlations between the two types of variables in Korean were consistent with the results of previous LIWC studies conducted in English. The authors discuss language use as a marker of personality.

Keywords: 5-factor personality test, KLIWC, language use, MBTI, personality

IN THEIR DAILY LIVES, people often make judgments about a person's age, intentions, and emotional state based on how the person uses language. Over the past several decades, various software programs were developed to systematically capture the relations between language use and different aspects of personality (e.g., Foltz, 1996; Gottschalk, Stein, & Shapiro, 1997; Hart, 1984).

This research was supported by Grant No. R01-2005-000-10512-0 from the Ministry of Science and Technology, the Korea Science and Engineering Foundation, to Chang H. Lee and Ajou University Research Grant No. 20060790 to Kyungil Kim.

Address correspondence to Kyungil Kim, Ajou University–Psychology, San 5, Woncheon-dong, Yeongtong-gu Suwon 443-749, Korea; kyungilkim@ajou.ac.kr (e-mail).

Pennebaker, Francis, and Booth's (2001) Linguistic Inquiry and Word Count (LIWC) is one text analysis program used extensively for research in nonclinical populations. LIWC searches for more than 2,000 words and word stems and categorizes them into linguistic dimensions (e.g., words per sentence, particles, pronouns), psychological dimensions (e.g., cognitive words, emotion words, social orientation words), personal concerns (e.g., death, money, sex), and relativity (e.g., time, space, motion). The program reports the number of words per category as a percentage of the total number of words appearing in a text file. The LIWC program has been validated extensively and has been shown to have sound psychometric properties (see Pennebaker, Mehl, & Niederhoffer, 2003).

Lee, Shim, and Yoon (2005) developed the Korean version of the LIWC (KLIWC) using variables in the original English version and incorporating grammatical and cultural variables that are integral to Korean culture. For example, most Korean words are composed of multiple morphemes (Yoon & Kwon, 1997); each letter in a Korean word has some meaning. Thus, the program must count how many morphemes are in a word. In addition, Korean culture emphasizes respect of elders, allowing for respectful words not found in English. Thus, the program must calculate the ratio of respectful words.

We conducted the present study to validate the KLIWC as a measure of personality and investigate the relation between language use and personality using a Korean sample. Pennebaker and King (1999) provided data on the relation between the Big Five personality dimensions and English language use. A factor analysis of the 72 English LIWC variables provided four main factors: immediacy, distinctions, social past, and rationalization. Immediacy includes first-person singular pronouns, articles, long words, present tense, and discrepancies. Distinctions include exclusiveness, tentativeness, negations, and inclusive words. Social past includes past tense, positive emotion words, and social words. Last, rationalization includes insight, negative emotion, and causation words. Correlating these factors to the Big Five personality dimensions showed that immediacy is negatively correlated with openness. Also, distinctions are negatively correlated with extraversion and conscientiousness.

In this article, we investigated the relation between KLIWC variables and several personality tests taken by Korean participants. We observed whether language variables of the KLIWC reflect personality factors as in previous LIWC studies (e.g., Pennebaker & King, 1999). Personality factors and their markers are culture common, but individual levels of them may differ among cultures. It is possible that in a culture, the level of a specific personality variable is relatively higher or lower than it is in the other cultures. However, researchers generally believe that personality-trait structure and the manifestations of its factors are universal (McCrae & Costa, 1997). Therefore, we predicted that the relations between language factors of the KLIWC and personality tests would be consistent with those observed in previous LIWC studies, subsequently validating the use of the current KLIWC.

Method

Participants

Eighty undergraduate students enrolled in a psychological statistics class at Pusan National University participated in the study for course credits. All of the students were native Korean speakers. The mean age of participants was 21.7 years (*SD* = 2.31 years). Of the sample, the majority were female (53 women, 27 men).

Procedure

Each participant sat alone in a room while completing two personality tests in Korean: the 5-Factor Inventory (John, Donahue, & Kentle, 1991) and the Myers-Briggs Type Indicator (MBTI; Myers, McCaulley, Quenk, & Hammer, 1998). The 5-Factor Inventory was available online and lasted about 20 min, whereas the MBTI paper test lasted about 30 min. Next, participants hand wrote a stream of consciousness essay for 20 min. Subjects were told that they could write about any topic that came to their mind as they were writing.

As briefly discussed in the previous section, the KLIWC includes two major parts: linguistic and nonlinguistic (e.g., psychological or personal concern-related) dimensions. The linguistic dimension consists of standard linguistic variables (e.g., word count, words per sentence, ratio of long words to short words) and variables concerning parts of speech such as adverb, auxiliary verb, verb, suffix, or specific types of nouns (e.g., proper nouns). In the nonlinguistic part, the psychological dimension includes negative emotion words (e.g., sad, angry), positive emotion words (e.g., happy, laugh), and cognitive words (e.g., because, hence), whereas the personal-concern dimension includes variables such as occupational (e.g., job, school), financial (e.g., money), or physical-state-related (e.g., sleep, sex) words. For each writing sample, the KLIWC computed the percentage of total words represented by each of these variables. We then correlated the scores of each of the linguistic output variables with the personality dimensions of the MBTI and 5-Factor Inventory.

Results

We analyzed the participants' writing to ensure that the individuals did not differ significantly in language ability. We measured the length of each participant's essay. The average word count of the essays was 341.6 words (*SD* = 41.8 words). None of the participants used fewer than 295 words.

We noticed significant correlations between the KLIWC variables and the 5-Factor Inventory and between the KLIWC variables and personality dimensions of the MBTI. As shown in Table 1, 13 of 86 KLIWC variables were correlated with at least one of the 5-Factor personality dimensions. Specifically, the KLIWC variable verb positively correlated with the 5-Factor personality variable extroversion.

TABLE 1. Correlations Between KLIWC Variables and 5-Factor Personality Test (O. P. John, E. M. Donahue, & R. L. Keuthe, 1991)

	Big Five factor				
KLIWC variable	Extroversion	Agreeableness	Conscientiousness	Stability	Openness
Number of sentences	.14	–.02	–.24*	.27*	.28*
Verb	.24*	.01	.07	.06	.07
Exclamation	.05	–.23*	.03	.11	.19
Proper noun	.08	.12	.08	–.27*	–.31*
Personal pronoun	.13	.01	.11	.04	.24*
Adverb	–.12	.14	–.17	.09	–.38*
Emotion word	–.09	.11	–.19	–.31*	–.03
Cognitive word	.09	.06	.33*	–.15	.12
Swear word	.09	.08	–.24*	.12	.08
Anger	.10	–.02	–.01	–.24*	–.02
Expectation	.04	.25*	.30*	–.15	.12
Confidence	.03	.11	.27*	.01	.01
Sleep	–.12	–.08	–.16	.17	–.31*

Note. KLIWC = Korean version of the Linguistic Inquiry and Word Count (C. H. Lee, J. Shin, & A. Yoon, 2005).
*$p < .05$.

According to the manual of the 5-Factor Personality Test (John et al., 1991), the facets of extroversion are sociability, gregariousness, assertiveness, and activity. This correlation indicates that the higher a participant scored on extroversion, the more he or she used verbs in natural language.

The KLIWC variables emotion and anger were negatively correlated with the 5-Factor personality variable stability, indicating that people who score lower on stability are more likely to use emotional and anger-related words in their writing than are those who score higher on stability. According to the manual of 5-Factor personality test, stability is related to anxiety, depression, anger, emotion, insecurity, and embarrassment. Thus, we expected the significant correlations between this personality factor and emotion and anger words.

As shown in Table 1, the variable conscientious also showed significant correlations with KLIWC variables. Conscientiousness was positively correlated with cognitive, expectation, and confidence words and negatively correlated with swear words. Highly conscientious people are characterized as dependable, careful, thorough, responsible, and hardworking (John et al., 1991). Last, people who scored high on openness tended to use fewer words related to sleep.

As shown in Table 2, 8 of the 86 KLIWC variables were correlated with at least one MBTI personality variable. Specifically, the MBTI sensing was positively correlated with the ratio of phrases, the ratio of morphemes, and the use of suffixes. According to Myers and McCaulley (1985), sensing is related to characteristics such as realism, acute powers of observation, and memory for details. Therefore, we expected that people who scored highly on this variable would be more likely to use words and suffixes to increase verbosity and expression of relations. Suffix use was negatively correlated with the intuition factor on the MBTI, which is the opposite factor of sensing.

Introversion on the MBTI was positively related to the use of auxiliary words and negatively related to verb use. The correlation with auxiliary words is somewhat unclear. However, its negative correlation to verb use is in line with the positive correlation found between extroversion on the 5-Factor Inventory and verb use.

Last, the judgment and perception factors of the MBTI showed opposite patterns of correlation with cognitive words, expectation words, and swear words. According to Myers and McCaulley (1985), the judgment factor is related to self-control, a stronger superego, rule boundedness, and dependability. In contrast, perception is related to impulsivity, rebelliousness, changeability, and restlessness. The judgment factor was positively correlated with cognitive words and expectation words and negatively correlated with swear words. This is similar to the pattern of correlations of conscientiousness of the 5-Factor personality test.

In sum, the results of this study suggest that the variables of the KLIWC reflect variations on the many factors of personality tests, as in the previous LIWC studies (e.g., Pennebaker & King, 1999). Furthermore, those relations were consistent with the relations observed between the personality factors and their manifestations.

TABLE 2. Correlations Between KLIWC Variables and MBTI Factors

	MBTI factor							
KLIWC variable	Extroversion	Introversion	Sensing	Intuition	Thinking	Feeling	Judgment	Perception
Ratio of phrases	–.17	.15	.27*	–.18	.06	.14	.11	.12
Ratio of morphemes	–.02	.19	.28*	–.07	.05	–.14	.15	–.14
Auxiliary word	–.21	.33*	.01	.03	.50	–.19	.11	–.12
Verb	.21	–.27*	.18	–.27*	.05	–.11	–.01	–.02
Suffix	–.05	.01	.28*	–.27*	–.03	.04	.18	–.16
Cognitive word	.11	–.08	.11	–.03	.12	–.18	.25*	–.27*
Swear word	–.01	.02	.05	–.20	.10	–.14	.27*	–.28*
Expectation	.02	–.04	–.19	.11	.12	.09	–.26*	.25*

Note. KLIWC = Korean version of the Linguistic Inquiry and Word Count (C. H. Lee, J. Shin, & A. Yoon, 2005). MBTI = Myers-Briggs Type Indicator (I. B. Myers, M. H. McCaulley, N. L. Quenk, & A. L. Hammer, 1998).
*$p < .05$.

Discussion

We conducted this study to investigate how the use of the Korean language reflects various personality dimensions. Several strong correlations emerged. Extroversion of the 5-Factor Inventory and of the MBTI was reflected by a high rate of verb use. Stability of the 5-Factor Inventory was reflected by the use of emotion and anger words. In addition, the sensing factor of the MBTI was reflected by a high ratio of phrases and morphemes and the use of suffixes. Last, judgment of the MBTI and conscientiousness of the 5-Factor personality test were related to the use of cognitive words and expectation words, and low use of swear words.

The Korean language has some important differences compared with other languages such as English. Nonetheless, the observed relations between factors of personality tests and language variables of the KLIWC in the current study are consistent with the observations of the previous clinical and LIWC studies. For example, the current result that people who scored lower on stability were more likely to use emotion and anger-related words in their essays than were those who scored higher is consistent with general knowledge about neuroticism, emotional instability and anger being its most typical symptoms (Nevid & Rathus, 2005, p. 62). Furthermore, our study replicated Pennebaker and King's (1999) observations that conscientiousness had a negative relation with negation and negative emotion words and a positive relation with positive emotion words. Also, their finding that openness was positively related to the use of articles and long words is consistent with the positive correlation we found between the number of words used and openness. These two studies both suggest that writing complexity predicts the level of openness.

The finding that personality is revealed in language use has important theoretical implications. Previously, language use was not thought to reveal personality dimensions (Pennebaker & King, 1999). However, language use at the word level is an observable and measurable behavior shown to be relatively stable across time and contexts (Pennebaker & King), fitting well with current definitions of personality (e.g., Funder, 2004). Studies of the LIWC and the current study using the KLIWC have shown that word use does reflect personality and that the linguistic cues to some personality traits are similar across languages. As Pennebaker et al. (2001) wrote, text analysis programs can be used to examine personality and psychological states through word use without effortful interviews, providing practical benefits such as increased efficiency, reliability, and external validity.

Nonetheless, the present study should not be compared directly with Pennebaker and King's (1999) study, as they correlated personality dimensions to linguistic indexes derived by factor analysis. Our own preliminary factor analysis revealed more than 10 factors, with the first factor explaining less than 9% of the variance of all correlations. Pennebaker and King's Factor 1 explained about

22% of the variance. Thus, it would be inappropriate to use linguistic factors on the Korean sample in this study and to compare our simple correlations with correlations with factors. However, there are some important similarities between the two studies. For example, stability of the 5-Factor personality test was correlated with negative emotion words, such as anger words, in Pennebaker and King's study, as in the current study.

Given the 86 KLIWC variables in this study, the total number of eight differences that were significant at the .05 level is somewhat close to the number of differences that can be expected to occur by chance (see Table 2; i.e., 8/86 = .09). However, we note that the KLIWC and the LIWC do not capture all aspects of human language use. There are many personality-irrelevant variables in the KLIWC and the LIWC.

The strength of the observed correlations in the current study is another important issue. The factors of the 5-Factor personality test and the MBTI measure complex personality dimensions made of multiple attributes or facets. Individual KLIWC variables explain very specific aspects of individual language use. In general, such inconsistencies can lower the possibility of discovering potential relations.

If personality manifests in subtle linguistic variables, researchers should include smaller but more concrete personality factors in future studies and observe their individual relations with the various language variables. We suggest that future research be conducted with a concrete theory, which would constrain the hypothesis. Such efforts would provide clearer insight into the relation between language use and personality.

AUTHOR NOTES

Chang H. Lee is an associate professor in the departments of psychology at Sogang University and Pusan National University in Korea. His research interests include word recognition, sentence processing, and text analysis. **Kyungil Kim** is an assistant professor in the Department of Psychology at Ajou University in Suwon, Korea, and his research interests are decision making, reasoning, and knowledge representation. **Young Seok Seo** is an associate professor in the College of Education at Konkuk University in Seoul, Korea, and researches cultural difference, personality, and emotional expression. **Cindy K. Chung** is a PhD candidate in the Department of Psychology at the University of Texas at Austin. Her research interests are social psychology and psychosomatics, language and health, and trauma.

REFERENCES

Foltz, P. W. (1996). Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments, & Computers, 28,* 197–202.

Funder, D. C. (2004). *The personality puzzle* (3rd ed.). New York: Norton.

Gottschalk, L. A., Stein, M. K., & Shapiro, D. H. (1997). The application of computerized content analysis of speech to the diagnostic process in a psychiatric outpatient clinic. *Journal of Clinical Psychology, 53,* 427–441.

Hart, R. P. (1984). *Verbal style and the presidency: A computer-based analysis.* Orlando, FL: Academic Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The "Big Five" inventory: Versions 4a and 5b (technical report).* Berkeley: University of California, Institute of Personality and Social Research.

Lee, C. H., Shim, J., & Yoon, A. (2005). The review about the development of Korean linguistic inquiry and word count. *Korean Journal of Cognitive Science, 16,* 93–121.

McCrae, R. R., & Costa, P. T. Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52,* 509–516.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.

Nevid, J. S., & Rathus, S. A. (2005). *Psychology and the challenges of life: Adjustment in the new millennium* (9th ed.). Hoboken, NJ: John Wiley.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count (LIWC): A computerized text analysis program.* Mahwah, NJ: Erlbaum.

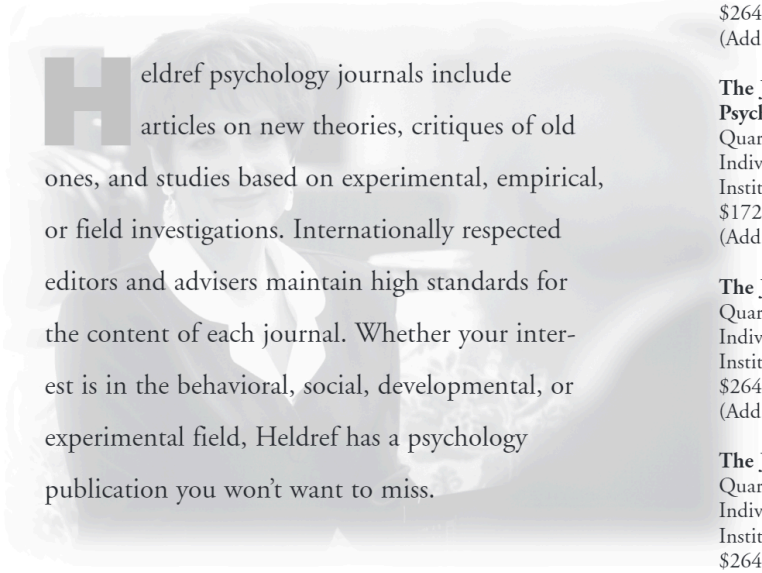
Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77,* 1296–1312.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54,* 547–577.

Yoon, A. S., & Kwon, H. C. (1997, July). *Rule-based morphological disambiguation using statistics in an agglutinative language.* Paper presented at the 16th International Congress of Linguists, Paris.

Manuscript submitted August 26, 2006
Revision accepted for publication January 15, 2007

Psychology



Heldref psychology journals include articles on new theories, critiques of old ones, and studies based on experimental, empirical, or field investigations. Internationally respected editors and advisers maintain high standards for the content of each journal. Whether your interest is in the behavioral, social, developmental, or experimental field, Heldref has a psychology publication you won't want to miss.

To place an order,
please copy this form and mail or fax; visit our Web site at **www.heldref.org** or call customer service at **(866) 802-7059.**

Heldref Publications
Subscription Office Address:
PO Box 830350, Birmingham, AL 35283-0350
Toll Free Number: (866) 802-7059 • Fax Number: (205) 995-1588
Fax Number for outside of the US and Canada: (205) 995-1567
Email address: heldref@ebsco.com

PUBLICATION #1		COST	
PUBLICATION #2		COST	
	Please include appropriate international shipping costs.	TOTAL COST	
ACCOUNT #	(MasterCard/VISA)	SECURITY CODE	EXP. DATE
SIGNATURE			
NAME/INSTITUTION			
ADDRESS			
CITY/STATE/ZIP		COUNTRY	

Genetic, Social, and General Psychology Monographs
Quarterly
Individual: $180 online only, $189 print and online
Institutional: $219 online only, $219 print only, $264 print and online
(Add $14 for postage outside the U.S.)

The Journal of Group Psychotherapy Psychodrama, and Sociometry
Quarterly
Individual: $71 online only, $74 print and online
Institutional: $144 online only, $144 print only, $172 print and online
(Add $13 for postage outside the U.S.)

The Journal of General Psychology
Quarterly
Individual: $180 online only, $189 print and online
Institutional: $219 online only, $219 print only, $264 print and online
(Add $14 for postage outside the U.S.)

The Journal of Genetic Psychology
Quarterly
Individual: $180 online only, $189 print and online
Institutional: $219 online only, $219 print only, $264 print and online
(Add $14 for postage outside the U.S.)

The Journal of Psychology:
Interdisciplinary and Applied
Individual: $207 online only, $217 print and online
Institutional: $251 online only, $251 print only, $301 print and online

The Journal of Social Psychology
Bimonthly
Individual: $207 online only, $217 print and online
Institutional: $251 online only, $251 print only, $301 print and online
(Add $16 for postage outside the U.S.)

Psychoanalysis & Psychotherapy
Quarterly
Regular Annual Subscription Rates:
Individual: $71 online only
Institutional: $137 online only

ReVision:
A Journal of Consciousness and Transformation
Quarterly
Individual: $47 online only, $49 print and online
Institutional: $103 online only, $103 print only, $124 print and online
(Add $14 for postage outside the U.S.)

heldref PUBLICATIONS
www.heldref.org

Heldref Publications is a division of the nonprofit Helen Dwight Reid Educational Foundation

SUB08