

## **How Taking a Word for a Word Can Be Problematic: Context-Dependent Linguistic Markers of Extraversion and Neuroticism**

**Matthias R. Mehl**      **Megan L. Robbins**      **Shannon E. Holleran**  
University of Arizona

This study conceptually extends recent research on linguistic markers of psychological processes by demonstrating that psychological correlates of word use can vary with the context in which the words are used. The word use of 90 participants was analyzed across two theoretically defined communication contexts. Information about participants' public language use was derived from recorded snippets of their daily conversations with others. Information about their private language use was derived from stream-of-consciousness essays. Personality trait-word use associations emerged as highly context dependent. Extraversion as a public trait was related to verbal productivity in public but not private language. Neuroticism as a private trait was related to the verbal expression of emotions in private but not public language. Verbal immediacy was indicative of Extraversion in public and Neuroticism in private language use. The findings illustrate the importance of considering communication contexts in research on psychological implications of natural language use.

Keywords: Personality expression, personality judgment, text analysis, communication, LIWC, Electronically Activated Recorder

It is a mundane observation that individuals differ in their word choice. Some people are verbose and convey thoughts with long and windy threads of words; others communicate telegram style. Some are highly emotional, almost dramatic in their language, others are cold and factual. Some use vivid and engaging words; others use plain and detached ones. A theoretically and practically important scientific question concerns the extent to which such idiosyncrasies in word choice are mere "noise" in the process of message communication or systematically reflect individual differences in latent psychological processes. Apart from Freud's (1901) legendary psychoanalytic investigations of slips of the tongue and general philosophical treatments of the function of language (e.g., Lacan, 1968; Ricoeur, 1976), natural variation in word usage has largely been neglected as a subject of scientific inquiry. Only recently have researchers begun to empirically explore the psychological underpinnings of individual differences in natural language use (Chung & Pennebaker, 2007; Pennebaker, Mehl, & Niederhoffer, 2003; Pennebaker, 2011).

In this research, aspects of people's word choice have been linked to important life outcomes such as marital satisfaction (Simmons, Gordon, & Chambles, 2005), relationship stability (Slatcher & Pennebaker, 2006),

therapeutic success (Mergenthaler, 1996), depression (Rude, Gortner, & Pennebaker, 2004), suicidality (Stirman & Pennebaker, 2001), adjustment to cancer (Lieberman & Golstein, 2006), heart disease proneness (Scherwitz, Graham, & Ornish, 1985), and even longevity (Danner, Snowdon, & Friesen, 2001). Differences in word use have further been found to be indicative of differences in cultural backgrounds (Maass, Karasawa, Politi, & Sayaka, 2006; Tsai, Simenova, & Watanabe, 2004), age (Pennebaker & Stone, 2003), and gender and sexual orientation (Groom & Pennebaker, 2005; Mulac & Lundell, 1994). Finally, word choice has been used to distinguish deceptive from truthful communication (Hancock, Curry, Goorha, & Woodworth, 2008; Newman, Pennebaker, Berry, & Richards, 2003) and to track psychological responses to upheavals (Cohn, Mehl, & Pennebaker, 2004; Pennebaker & Lay, 2002). Over the last years, researchers have accumulated ample evidence that the words people naturally use have profound psychological implications.

Conceptually, it seems necessary that psychological correlates of words are at least partially determined by the context in which the words are used. For example, Goffman (1981) pointed out that cursing in public is motivationally different from cursing privately to oneself. Similarly, Clark and Tree (2002) found that even seemingly meaningless fillers such as ‘uh’ or ‘um’ take on different linguistic functions in formal and informal communications. Given the apparent role that context can play in determining psychological implications of word use, it is surprising that most studies so far have implicitly adopted the working hypothesis that psychological correlates of word use are by and large context-independent. Systematic and theoretically fueled investigations into context effects on psychological correlates of language use are missing.

Empirically, the role of the context is sometimes indirectly hinted at from inconsistencies across study findings. For example, Groom and Pennebaker (2005) found that women spontaneously made more social references than men in online personal ads. In contrast, Mehl and Pennebaker (2003) reported that both sexes made social references at comparable rates in daily conversations with others. Similarly, Mulac and Lundell (1994) reported that women used fewer first person singular pronouns than men in neutral picture descriptions. In contrast, Pennebaker & King (1999) found that women consistently used “I”, “me”, and “my” at higher rates in personal essays. Finally, Simmons et al. (2005) found a positive relationship between first person singular use in marital interactions and relationship satisfaction. In contrast, Sillars, Shellen, McIntosh and Pomegranate (1997) reported a negative correlation based on a slightly different design.

Importantly, in these examples, it is likely that the lack of effect convergence across studies reflects systematic differences in the captured

psychological processes rather than mere sampling bias. Specifically, the discrepancy in the first example may hint at a tendency for women to express sociability non-verbally in spoken and verbally in written language (Mehl, Gosling, & Pennebaker, 2006). Similarly, the discrepancy in the second example may indicate a tendency for men to be more engaged in writings of an impersonal and women in writings of an interpersonal nature (Pennebaker et al., 2003). Finally, in the third example, Simmons and her colleagues resolved the discrepancy by suggesting that the “use of first-person singular may be adaptive in some contexts and not in others ... It is possible that the task of the present study pulled for more negativity, and that in this context, increased self-reference was beneficial.” (Simmons et al., 2005, p. 935)

Taken together, then, both conceptual arguments and empirical evidence suggest that an important step towards developing a better theoretical understanding of the psychological implications of natural language use lies in (a) comparing participants’ language use across more than one communication context, and (b) empirically testing the extent to which linguistic markers of psychological processes systematically depend on contextual boundary conditions. Given the lack of research on the role of communication contexts in determining psychological implications of language use, this study sought to (a) provide first evidence that psychological correlates of word use can be highly context dependent and (b) demonstrate the value of theoretically specifying psychological aspects of communication contexts that moderate the extent to which words are indicative of psychological processes. Specifically, this study explored the role of communication contexts in the expression of personality in natural language use.

In the early days of personality psychology, Allport (1937) noted that “language is a codification of common human experience, and by analyzing it much may be found that reflects the nature of human personality” (p. 373). Only recently, however, have researchers begun to explore how elements of language use are related to personality (Gifford & Hine, 1994; Ireland & Mehl, in press; McAdams et al., 2004; Weintraub, 1989). Identified personality trait-word use links have generally been consistent with the theoretical trait definitions. For example, Extraversion has been found to be related to the use of positive emotion words (Oberlander & Gill, 2006; Pennebaker & King, 1999), Neuroticism to the use of negative emotion words and first person singular (Pennebaker & King, 1999; Gill, Nowson, & Oberlander, 2007), and Openness to Experience to a preference for long words (Pennebaker & King, 1999; Gill et al., 2006). Similarly, the use of certainty words (e.g., guarantee, sure, definite) has been found to be related to being described as confident and intelligent, and the use of sexual references as being seen as assertive, expressive, and unconventional (Greve & Funder, 2006). In analyzing only

one language source from each participant, however, these studies could not address to what extent the identified personality trait–word use links generalize across contexts. Also, possibly due to a suboptimal matching between the studied traits and the affordances of the specific communication contexts from which participants' language was sampled (e.g., Extraversion and journal abstracts), the reported effect sizes have generally been small (Pennebaker & King, 1999). Finally, a narrative comparison of effects across two computer-mediated communication genres (blogs and emails) revealed both evidence for partial cross-genre convergence and genre-specific trait-word use associations (Gill et al., 2006).

How can communication contexts moderate the expression of personality in verbal behavior? From a Gibsonian perspective (Gibson, 1979), the psychological function that a specific context takes on for a given personality trait determines the extent to which it affords the expression of that trait. For example, compared to the solitary, studious environment of a library, the casual interpersonal environment of a party provides a unique opportunity structure for sociability as a trait to be expressed in sociable behavior. Funder's (1999) Realistic Accuracy Model conceptualizes such contextual influences on the expression of personality as trait  $\times$  information moderators of judgmental accuracy. In doing so, the Realistic Accuracy Model assumes that people have a "real" personality that can be accurately judged if the particular situation affords the expression of the relevant traits, and if the judge utilizes the correct cues to guide their judgment. Particular information (e.g., what moral values a person holds) becomes available only in particular contexts (e.g., deep conversations) and is critical for achieving accurate personality judgments on specific traits (e.g., traditionalism). Both frameworks suggest that it is important to identify what psychological features of the context can affect the naturalistic expression of personality.

Snyder and Ickes (1985) proposed two characteristics that define situations with regard to their role in personality processes: strong versus weak and precipitating versus non-precipitating situations. Strong situations are highly structured, provide salient cues that guide behavior, and therefore maximize the effects of situational demands; weak situations are relatively unstructured, provide little cues to guide behavior, and therefore maximize the effects of personality traits. Weak situations, then, generally afford the expression of (a broad array of) personality traits (cf. Letzring, Wells & Funder, 2006). Precipitating situations, on the other hand, are situations that (a) are relevant to a particular disposition, (b) make the disposition salient as a guide to behavior, and (c) permit alternative modes of responding that individuals differentially select as a function of their location on the dispositional dimension. Precipitating situations, then, selectively afford the expression of conceptually relevant

personality traits (cf. Marshall & Brown, 2006).

An important theoretical dimension along which personality traits vary is the degree to which they are public and generally high in visibility or private and generally low in visibility. For example, Extraversion, defined as the tendency to be sociable, talkative, and enthusiastic, is a prototypic public trait and Neuroticism, defined as the tendency to be anxious and easily upset, is a prototypic private trait (John & Robins, 1993; Watson, Hubbard, & Wiese, 2000). Vazire and Gosling (2004) identified that private-public is also a critical dimension along which interpersonal perception-public contexts vary. People's personal websites (e.g., MySpace or Facebook pages) and physical appearance are relatively public and their thoughts or dreams are relatively private forms of expression. Based on a straightforward correspondence logic, then, public and private interpersonal contexts should constitute precipitating situations for the expression of public and private personality traits. Therefore, one would expect that public communication contexts particularly afford the expression of public personality traits and private communication contexts particularly afford the expression of private personality traits.

We tested this prediction within an ecological study design where we collected two naturalistic language samples from each participant. Information about participants' public language use was derived from verbatim transcripts of recorded snippets of their daily conversations with others, sampled using the Electronically Activated Recorder (EAR; Mehl & Pennebaker, 2003). The EAR is a handheld computer with software that unobtrusively records snippets of ambient sounds in participants' immediate environment, alleviating problems with reliance on in-lab observation or self-reports of behavioral frequencies (Mehl, 2007). The EAR has been reliably used in numerous studies across a variety of topics, including behavioral manifestation of personality (Mehl, Gosling, & Pennebaker, 2006), conversational correlates of well-being (Mehl, Vazire, Holleran, & Clark, 2010), the validity of gender stereotypes (Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker, 2007), diurnal rhythms underlying affect-associated behavior (Hasler, Mehl, Bootzin, & Vazire, 2008), the role of expressive behaviors in the coping process (Robbins et al., 2011; Robbins, Mehl, Holleran, & Kastle, 2011), and manifestations of family conflict (Slatcher & Trentacosta, 2011). For a thorough review of the validity of the EAR method, see Mehl, Robbins, and große Deters (2012).

Information about participants' private language use was derived from a stream-of-consciousness (SOC) writing task where participants tracked their momentary thoughts and feelings (Pennebaker & King, 1999). We specifically sought to identify context-specific linguistic markers of private and public personality traits. In line with this aim, we selected variables, contexts, and traits on theoretical grounds to identify such context-specific effects. We hypothesized that Extraversion as a public trait would be more

strongly related to aspects of participants' public EAR-sampled word use and Neuroticism as a prototypic private trait would be more strongly related to aspects of participants' private SOC word use. The relevant types of word use examined here are (a) verbal production, (b) verbal emotional expression, and (c) verbal immediacy. Verbal production refers to the sampled count of words uttered by the participant. Verbal emotional expression is the percentage of words uttered that were positive or negative emotion words. Lastly, verbal immediacy refers to the degree to which participants' language use reflects psychological engagement with versus distance from the topic about which they speak (Cohn et al., 2004; Pennebaker et al., 2003; Wiener & Mehrabian, 1968). Verbal immediacy is marked by a high percentage of words uttered that are first person, singular pronouns, and infrequent use of articles, long words, and discrepancy words, indicating a more personal, involved, and experiential tone, relative to an impersonal, distanced, and rational one.

Specifically, we predicted that (a) Extraversion would be more strongly related to verbal productivity in participants' EAR-sampled language (i.e. public interactions) than in their SOC language (i.e. private thoughts), (b) Neuroticism would be more strongly related to verbal emotional expression in participants' SOC than in their EAR-sampled language, and (c) verbal immediacy would be related to Extraversion in participants' EAR-sampled language and Neuroticism in their SOC language. Several studies across different labs have identified verbal immediacy as a core linguistic device that conveys a personal, involved, experiential language (Biber, 1988; Gill et al., 2006; Pennebaker & King, 1999). The last prediction, then, is based on the idea that verbal immediacy facilitates social processes in interpersonal contexts and emotional processes in private contexts (Wiener & Mehrabian, 1968).

## Method

### Participants

Ninety introductory psychology students (45 female, mean age = 18.7,  $SD = 0.9$ ) participated in the study for course credit. The students are a sub-sample of the 96 participants reported in Mehl et al. (2006) for which both EAR and SOC language samples were available.<sup>1</sup>

### Personality Measures

Upon arrival to the lab, the participants were informed about the purpose of the study, provided their informed consent, and completed a

---

<sup>1</sup> The SOC writing samples of 6 (4 male, 2 female) participants did not get stored because of temporary server problems.

questionnaire battery that included the 44-item Big Five Inventory (BFI; John & Srivastava, 1999). The participants rated themselves on the BFI items using a scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The alpha reliabilities for Extraversion and Neuroticism were .90 and .87. The two scales were correlated at  $-.28$ ,  $p = .01$ . In this sample, the mean for Extraversion was 3.49 ( $SD = 0.93$ ) and the mean for Neuroticism was 2.70 ( $SD = 0.89$ ). Both variables were normally distributed but evidenced a slight skew in direction of their socially desirable end. John and Srivastava (1999) provide more detailed psychometric information on the BFI.

### Collection of the Language Samples

*EAR monitoring and transcription of captured utterances.* The EAR system consisted of a digital voice recorder (SONY ICD-MS1), an external microphone (OPTIMUS Tie Clip Microphone), and a controller chip (Mehl et al., 2001). The chip was programmed to produce 4.8 30-sec recordings per hour. Participants wore the EAR for two consecutive weekdays during their waking hours. They were thoroughly informed about the study's privacy and confidentiality policies and had an opportunity to erase recordings they did not want the researchers to hear.<sup>2</sup> On average, participants provided 128 ( $SD = 34$ ) valid waking sound files (~64 min). More information on participants' compliance with the method is provided in Mehl et al. (2006). Most utterances captured by the EAR take place in an interpersonal context, and therefore render them a public communication context—anyone in proximity to the participant could potentially hear them.

A team of research assistants transcribed all of the participants' utterances captured by the EAR. Inter-transcriber agreement (ICC[2,k]) calculated from the transcripts of 76 conversations contained in a set of training EAR recordings transcribed by 8 research assistants exceeded .90 for all language-use categories in this study.

*SOC writing exercise.* Participants completed a 20-min SOC writing exercise when they returned to the lab to drop off their EAR. Specifically, they received the following instructions (cf. Pennebaker & King, 1999): "For the next 20 minutes, write about whatever comes to your mind. Think about what your thoughts, feelings and sensations are at this moment. Write about them as they come to you; follow where your mind naturally goes". Participants were asked to write continuously without stopping and not to pay attention to grammar or spelling. Participants were seated in individual sound-dampened cubicles and typed directly into a computer. Following standard recommendations for running writing studies, it was emphasized to participants that the privacy and confidentiality of their

---

<sup>2</sup> 16/90 participants used this opportunity and 3/90 students erased a total of 10 sound files. Therefore, only 10 of >11,000 sound files were deleted.

essays would be protected. Compared to the EAR-sampled interpersonal conversations, this solitary, confidential setting provided a relatively private communication context. Indeed, in this SOC writing task, participants typically share their private thoughts and feelings (Holleran & Mehl, 2008).

### **Linguistic Analysis and Derived Measures of Language Use**

The participants' verbatim EAR transcripts and SOC writing samples were submitted to *Linguistic Inquiry and Word Count (LIWC*; Pennebaker, Francis, & Booth, 2001). LIWC is one of the most widely used and extensively validated word-count based text analysis programs (Mehl, 2005). It operates by comparing all words of a given text to a dictionary consisting of more than 2,300 words. The words in the LIWC2001 default dictionary are arranged into 74 grammatical (e.g., pronouns, articles, prepositions) and psychological (e.g., positive emotions, cognitive mechanisms) language-use dimensions. The LIWC analysis resulted in two sets of language-use variables per participant. All variables (except raw word count) are expressed as percentages of the total number of sampled words. For this study, we limited the analyses to LIWC variables with theoretical links to Neuroticism and Extraversion. Specifically, we focused on the following 3 linguistic devices: Verbal productivity, verbal emotional expression, and verbal immediacy. These linguistic devices have repeatedly been found to capture important psychological processes (for a review see Pennebaker et al., 2003).

*Verbal productivity.* Participants' verbal productivity was assessed with the LIWC variable Word Count. It is expressed as the total number of words that were contained in the participants' verbatim EAR transcripts and their SOC writing samples, respectively.

*Verbal emotional expression.* The LIWC variables Positive Emotion Words (e.g. "happy", "joy", "pride") and Negative Emotion Words (e.g. "hate", "afraid", "cry") were selected to capture the degree to which participants verbally expressed positive or negative emotions. The superordinate LIWC emotion categories were chosen over more specific ones (e.g., Sadness, Anxiety) because of the low base rates of the latter ones in natural language.

*Verbal immediacy.* A composite measure of verbal immediacy was derived by aggregating the following 5 LIWC variables: First-person Singular Pronouns ("I", "me", "my"), Present Tense Verbs, Discrepancy Words (e.g. "would", "should", "could"), and inverse scores of Articles ("a", "the") and Words of More Than Six Letters. These linguistic variables have been found to naturally co-occur in speech and writing, and to factor-analytically form one, usually the first, factor (Biber, 1988; Gill, et al., 2006; Pennebaker & King, 1999). All variables were standardized prior to



their aggregation. To preserve mean differences between the two language samples, the z-scores were computed across both participants' EAR and SOC samples.

The composite measure showed adequate internal consistency for both participants' EAR-sampled and SOC language (Cronbach's  $\alpha = .56$  and  $.77$ ). Samples scoring high on verbal immediacy can be described as using personal, involved, experiential language, focused on the here and now. Low scoring samples can be characterized as having an impersonal, detached, abstract, and rational tone (Wiener & Mehrabian, 1968). Table 1 shows language samples from participants who scored high or low on the EAR and SOC measure of verbal immediacy.

Table 1  
*Language Samples from Four Participants Who Scored High or Low on the EAR-derived and Stream-of-Consciousness Measure of Verbal Immediacy*

Language Sample	High Verbal Immediacy	Low Verbal Immediacy
EAR	Um, I don't know. If I don't talk to you anymore this week I'll talk to you this Friday. Nobody ... unless I can get somebody to come but Sarah ... well, Matt's coming up here to Austin, Sarah's boyfriend, Brittney, is going home. Um, I was going to ask my friend Christine if she wanted to come.	Then whenever the sun came out the ice started to melt and kept falling from the trees. It was cool, but it wasn't fun because it was freezing and every part of me was warm except for my face. That was miserable - every part of me was warm except my face.
SOC	Beth left today for the wedding...I miss Jake I want to see him even though I talk to him every night...I should call my mom tonight...it was good seeing Sarah - I miss her - things were awkward with Jason and Jill - Beth still doesn't know about the Jill thing and I don't think I should tell her	It reminds me of the long introspective ego death ...Your mind follows long and complicated trains of logic to some conclusion which cannot be expressed in words. You are left speechless. Staring at the sky. Smiling and unafraid of everything. It's an amazing thing, when it happens.

*Note.* EAR = Electronically Activated Recorder; SOC = Stream of Consciousness; the immediacy composite variable is based on the following LIWC variables: 1<sup>st</sup> person singular pronouns (I, me, my), present tense verbs, discrepancy words (e.g., would, should, could), articles (a, the; reverse scored) and words of more than six letters (reverse scored).

## Results

### Relationships between Participants' EAR-Sampled and SOC Language

Table 2 shows the base rates for participants' EAR-sampled (public, recorded language use) and SOC language (private, written thoughts). Expectedly, over the 2 days of monitoring the EAR sampled on average more words than participants wrote during the 20-min SOC writing exercise. Further, participants used both less positive and negative emotion words in the EAR-sampled compared to the SOC language. This suggests that participants' daily conversations contained less emotional themes than their spontaneous streams of thought. Finally, participants' EAR-sampled and SOC language evidenced comparable levels of verbal immediacy.

Table 2

*Base Rates and Intercorrelations for Participants' EAR-Sampled and Stream-of-Consciousness Language Use*

LIWC variable	Base Rates		paired <i>t</i> -test	<i>r</i>
	EAR-sampled language	SOC language		
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>t</i> (89)	
<i>Verbal productivity</i>				
Sampled raw word count	1011 (523)	766 (277)	4.07**	.08
<i>Verbal emotional expression</i>				
Positive emotion words	2.3 (0.8)	2.9 (0.9)	-4.52**	.16
Negative emotion words	1.6 (0.8)	1.9 (1.0)	-2.71**	.18
<i>Verbal immediacy</i>				
Immediacy composite variable	-0.2 (0.4)	0.2 (0.8)	-0.43	.08

*Note.* *N* = 90; EAR = Electronically Activated Recorder; SOC = Stream of consciousness; LIWC = Linguistic Inquiry and Word Count; the immediacy composite measure is the average of the following standardized LIWC variables: 1<sup>st</sup> person singular pronouns, present tense verbs, discrepancy words, articles (reverse scored) and words of more than six letters (reverse scored); in order to preserve mean differences between the language use domains, the standardizations were based on the means and standard deviations across both language samples; the other LIWC variables (except sampled raw word count) are expressed as percentages of the total number of sampled words; base rates are variable means with standard deviations in parentheses.

\*\* *p* < .01.

For all 3 linguistic devices, participants' EAR-sampled and SOC language use were only weakly correlated (last column in Table 2). Prior research has found moderate cross-situational consistency in word use (Pennebaker et al., 2003). In interpreting the relatively low consistency in this study, it is important to consider that (a) most studies so far have compared word use across different topics but within the same communication medium (i.e. only written or spoken language; Mitzner & Kemper, 2003) and (b) reported consistency estimates for a single (i.e., average correlations) rather than the average measure (i.e., Cronbach's alpha) tend to be comparable to the ones in this study (Pennebaker & King, 1999).

### **Linguistic Markers of Neuroticism and Extraversion in SOC and EAR-sampled Language**

*Verbal productivity.* Consistent with our first prediction, Extraversion was correlated substantially and significantly with the number of words that the EAR captured over the 2-day monitoring but only weakly and non-significantly with the number of words participants' wrote in the SOC essays (see Table 3). The differences between the two correlations approached statistical significance ( $p = .08$ ).<sup>3</sup> Further, Neuroticism was negatively related to the number of EAR-sampled words and unrelated to the number of words in the SOC essays.

*Verbal emotional expression.* In line with our second prediction, Neuroticism was negatively related to the use of positive emotion words and positively related to the use of negative emotion words in the SOC essays but unrelated to the use of positive and negative emotion words in participants' EAR-sampled language. The differences between the two corresponding pairs of correlations were statistically significant. Further, Extraversion was unrelated to the use of positive and negative emotion words in both the SOC and the EAR language samples.

*Verbal immediacy.* Finally, consistent with our third prediction, verbal immediacy, that is a personal, involved, and experiential language, was correlated with Extraversion in participants' EAR-sampled language and Neuroticism in the SOC essays. In contrast, verbal immediacy was unrelated to both Extraversion in the SOC essays and Neuroticism in participants' daily conversations with others. Again, the differences

---

<sup>3</sup> We also analyzed the data using multi-level modeling analyses with language source as level 1 predictor and Extraversion and Neuroticism, respectively, as level 2 predictors. The results were very similar to those obtained from the correlational analyses: The cross-level interactions were significant where the two corresponding correlations differed significantly. To facilitate interpretation, we report the simple correlation coefficients as effect size estimates along with significance tests for differences in dependent correlations (Williams' modification of the Hotelling test; Kenny, 1987).

between the two corresponding pairs of correlations were statistically significant.

Table 3  
*Correlations between Extraversion and Neuroticism and Participants' EAR-Sampled and Stream-of-Consciousness Language Use*

LIWC-category	Extraversion		Neuroticism	
	EAR-sampled language	SOC language	EAR-sampled language	SOC language
<i>Verbal productivity</i>				
Sampled raw word count	.31**	.11	-.22*	-.02
<i>Verbal emotional expression</i>				
Positive emotion words	.05	.03	<b>.13</b>	<b>-.22*</b>
Negative emotion words	-.01	-.12	<b>.06</b>	<b>.27**</b>
<i>Verbal immediacy</i>				
Immediacy composite measure	<b>.24*</b>	<b>-.06</b>	<b>.00</b>	<b>.24*</b>

Note.  $N = 90$ ; EAR = Electronically Activated Recorder; SOC = Stream of consciousness; LIWC = Linguistic Inquiry and Word Count; the immediacy composite measure is based on the following LIWC variables: 1<sup>st</sup> person singular pronouns, present tense verbs, discrepancy words, articles (reverse scored) and words of more than six letters (reverse scored); correlations in bold indicate significant differences between the two adjacent correlations ( $p \leq .05$ , one-sided; Williams modification of the Hotelling test; Kenny, 1987).

\*  $p < .05$ , \*\*  $p < .01$

## Discussion

Recently, there has been a growing interest in the psychological implications of natural language use. The research on this topic has led to important insights into language-linked psychological processes (Chung & Pennebaker, 2007; Pennebaker, 2011; Pennebaker, Mehl, & Niederhoffer, 2003). However, both theoretical arguments and empirical inconsistencies among study findings have pointed to two potential road blocks on the way to developing more refined theories of word use: (a) the lack of systematic data on the generalizability of correlates of word use across communication contexts and (b) the failure to theoretically specify communication context in which words are assumed to be particularly informative about certain psychological processes. In sampling language across two communication contexts and in mapping an important psychological feature of communication context (private vs. public environment) onto a critical theoretical feature of the psychological

process of interest (private vs. public trait expressions), this project provided first evidence that the psychological implications of natural word use can be highly context specific.

### **Private and Public Trait Expression in Private and Public Language Use**

Consistent with our first prediction, Extraversion as a prototypic public trait was uniquely related to verbal productivity in a public communication context (i.e. EAR verbal recordings). In line with the trait definition (e.g., Costa & McCrae, 1992; Goldberg, 1990) and prior research (Gifford & Hine, 1994), extraverted participants were more talkative in their daily conversations with others than their introverted counterparts. Consistent with our second prediction, Neuroticism as a prototypic private trait was uniquely related to verbal emotional expression in a private communication context (i.e. stream of consciousness essays). Again, in line with the trait definition and prior research (Pennebaker & King, 1999), neurotic participants' natural streams of thought contained more negative emotion words and fewer positive emotion words than their less neurotic counterparts.

Finally, consistent with our third prediction, verbal immediacy had different trait implications in public and private language. When participants' talked to others in their daily conversations, a personal, involved, experiential language was indicative of Extraversion (see upper left example in Table 1). In contrast, when participants tracked their momentary thoughts and feelings, a personal, involved, experiential language was indicative of Neuroticism (see lower left example in Table 1). The link between verbal immediacy and Neuroticism has previously been established for two relatively private language sources, students' journal writings (Pennebaker & King, 1999) and weblogs (Gill et al., 2007). The link between verbal immediacy and Extraversion has not been reported before—presumably due to a lack of highly public language samples in prior research (Mehl, et al., 2006).

The same linguistic device, then, can be indicative of very different psychological processes depending on the communication context in which it is used. We suggest that this is the case because verbal immediacy may take on different psychological functions according to the requirements of the contexts (Gibson, 1979). In a private communication context such as in a person's inner speech, daydreams, or journal writings, verbal immediacy may track "task engagement" and thereby reflect the "hot", experiential processing of emotionally colored autobiographic material (Epstein, 1994; Wiener & Mehrabian, 1968). In this context, then, verbal immediacy may have primarily emotional consequences. On the other hand, in a public communication context, such as in a person's face-

to-face interactions, verbal immediacy may track “task engagement” and thereby be partially responsible for extraverts’ perceived enthusiastic nature and their ability to recruit and hold social attention in the presence of others (Ashton, Lee & Paunonen, 2002). In this context, then, verbal immediacy may have largely social consequences.

Interestingly, the degree to which participants verbally expressed emotions in their daily conversations was not reliably related to either Neuroticism or Extraversion. The use of emotion words in public contexts, then, may not have very strong trait implications but instead be more informative of contextually-cued social dynamics such as situational scripts, peer influences, or impression management concerns (Goffman, 1981). However, the absence of a relationship between Extraversion and the verbal expression of positive emotions in a social context was surprising considering prior research on Extraversion and positive affect. Non-overlapping method-variance between the trait and our linguistic measure of emotional expression can (only) serve as a partial explanation (Lucas & Fujita, 2000). Future research should follow up on this finding and explore links between Extraversion and the verbal expression of positive emotions in different communication contexts using more fine-grained categories of positive emotion words.

**Limitations and Future Directions.** In looking at private and public trait expressions, this study focused on a communication context that has been identified as an important moderator of personality expression (John & Robins, 1993; Watson, et al., 2000). To provide a strong test of context dependence, we sampled language from a private, a person’s momentary thoughts and feelings, and a public, a person’s daily conversations with others, context (Vazire & Gosling, 2004). Many important sources of natural language use, however, cannot be readily classified as private or public and fall somewhere in between on a private--public continuum. For example, it is currently a topic of scientific debate, to what degree different forms of computer-mediated communications such as emails (Baron, 1998), blogs (Gill et al., 2006), or instant messages (Slatcher & Pennebaker, 2006) are public or private in nature. Blogging websites, for instance, possess traditional (private) journaling features but are also used as (public) platforms for social sharing.

In a study on psychological responses to the attacks of September 11, 2001, Cohn et al. (2004) found a large drop in verbal immediacy in people’s blogs in the first two months after the events. Interpreting the findings from the present study in the context of Fleeson’s (2001) model of traits as density distributions of states, the question emerges whether this post-9/11 change in language use was indicative of an increase in state Extraversion (e.g., affiliative tendencies) or –counterintuitively–a decrease in state Neuroticism (e.g., emotionality)? Cohn et al.’s (2004)

interpretation of the finding as reflecting psychological distancing as part of a psychological shock reaction is more in line with the latter explanation and empirically supported by Kosloff et al.'s (2006) finding that under mortality salience individuals showed dissociative tendencies to 9/11 reminders. More research is needed to identify the psychological underpinnings of patterns of word use in the increasingly prevalent, yet theoretically still underspecified new forms of computer-mediated communication (Gill et al., 2006).

Future research should more generally explore the degree of generalizability and context-dependence of different personality trait-word use links. This study sought to identify context-specific linguistic markers of private and public personality traits. In line with this aim, it selected variables, contexts, and traits on theoretical grounds to identify such context-specific effects. Within the same data set, other trait-word use links emerged as invariant across the two communication contexts. For example, the use of swear words was indicative of a lack of agreeableness in both participants' EAR and SOC language. Similarly, the use of anger words was indicative of a lack of conscientiousness in both language samples. The findings from this point to the potentials of letting theoretical considerations regarding the affordance structure of specific communication contexts guide predictions as to which linguistic trait makers should be specific to or invariant across communication contexts.

Finally and more broadly, within research on the everyday judgment of personality, very little is known about the psychological dimensions underlying common interpersonal perception contexts (e.g., Funder & Sneed, 1993; Gifford & Hine, 1994; Gill, Oberlander, & Austin, 2006; Gosling, Ko, Mannarelli, & Morris, 2002; Letzring, et al., 2006). Vazire and Gosling (2004) asked "What characteristics of bedrooms make them a particularly good context for judging Emotional Stability? What characteristics of personal websites make them a particularly good context for judging Openness?" and point out that "the current research on interpersonal perception in ecological contexts does not provide direct answers to these questions (...) By conceptualizing interpersonal-perception contexts in terms of such dimensions, commonalities can be identified and ultimately used to develop a general theory of the processes and parameters underlying the expression and perception of personality" (p. 129).

Beyond the fact that language is arguably one of the most common forms of personality expression, a unique strength of using language as a medium for studying interpersonal perception is that language naturally varies along several dimensions and can therefore be systematically sampled and compared to estimate the impact that psychological parameters have on the expression of personality in spontaneous (verbal) behavior. For example, in addition to the private-public continuum, the

degree of controllability is another dimension along which interpersonal perception contexts vary (Vazire & Gosling, 2004). People's spontaneous word use has been considered a prime example of an automatic effect; individuals generally have little control over what words they spontaneously use (Pennebaker et al., 2003). Sometimes, however, they do carefully choose their words. Potentially, then, a direct comparison of word use across a high controllability context, such as a crafted verbal self-description, and a low controllability context, such as a think-aloud self-description under cognitive load, can provide critical information about impression management concerns in verbal self-presentations (cf. Paulhus, 2002). The study of word use provides researchers with the unique opportunity to test how associations between selected behaviors (i.e. types of words) and selected traits change as specific features of the context (e.g., private vs. public and/or high vs. low control) are varied and the overall medium of personality expression (i.e. language) is held constant.

To conclude, numerous studies have demonstrated that the words people use reveal more than what their semantic meaning suggests. Idiosyncrasies in word choice have profound psychological implications. This investigation conceptually extends this research by providing empirical evidence for the idea that these implications can sometimes critically and theoretically meaningfully depend on the larger communication context in which the words are used. In demonstrating this effect in the context of the everyday expression of personality, it revitalizes Allport's (1937) early insight that "language (...) reflects the nature of human personality" (p. 373) and points to unique ways of how the psychological study of words can further our understanding of basic personality processes.

**Author Notes:** Portions of these data that examine broad personality implications of daily life and the accuracy of lay assessments of sub-clinical depression are reported elsewhere (Mehl, 2006; Mehl, Gosling, & Pennebaker, 2006). We are grateful to Samuel Gosling, Jeff Greenberg, and James Pennebaker for their valuable comments on previous drafts of this manuscript. Correspondence concerning this article should be addressed to Matthias R. Mehl, Department of Psychology, University of Arizona, Tucson, AZ 85721, e-mail: mehl@email.arizona.edu.

## References

- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Henry Holt.
- Ashton, M. C., Lee, K., & Paunonen, S. V. (2002). What is the central feature of



- extraversion? Social attention versus reward sensitivity. *Journal of Personality and Social Psychology*, *83*, 245-252.
- Baron, N. (1998). Letters by Phone or Speech by Other Means: The Linguistics of Email, *Language and Communication*, *18*, 133-170.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Chung, C. K. & Pennebaker, J. W. (2007). The psychological function of function words. In K. Fiedler (Ed.), *Frontiers of social psychology* (pp. 343-359). New York: Psychology Press.
- Clark, H., & Tree, J. F. (2002). Using *uh* and *um* in spontaneous speech. *Cognition* *84*, 73-111.
- Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic indicators of psychological change after September 11, 2001. *Psychological Science*, *15*, 687-693.
- Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Danner, D., Snowdon, D., & Friesen, W. (2001). Positive emotions in early life and longevity: Findings from the Nun Study *Journal of Personality and Social Psychology*, *80*, 814-814.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*, 709-724.
- Fleeson, W. (2001). Towards a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*, 1011-1027.
- Freud, S. (1901). *Psychopathology of Everyday Life*. New York: Basic Books
- Funder, D. C. (1999). *Personality Judgment: A Realistic Approach to Person Perception*. San Diego: Academic Press.
- Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, *64*, 479-490.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, Houghton Mifflin.
- Gifford, R. & Hine, D. (1994). The role of verbal behavior in the encoding and decoding of interpersonal dispositions. *Journal of Research in Personality*, *28*, 115 - 132.
- Gill, A. J., Nowson, S., & Oberlander, J. (2007). *Language and personality in computer-mediated communication: Factors and features across genres and models*. Manuscript submitted for publication.
- Gill, A. J., Oberlander, J., & Austin, E. (2006). Rating E-mail personality at zero acquaintance. *Personality and Individual Differences*, *40*, 497-507.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goldberg, L. R. (1990). An alternative "Description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216-1229.
- Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*, 379 - 398.

- Greve, L., & Funder, D. (2006). *Personality and Language: How Our Words Reveal Us*. Poster presented at the annual conference of the Association for Research in Personality. Palm Springs, CA.
- Groom, C.J., & Pennebaker, J.W. (2005). The language of love: Sex, sexual orientation, and language use in online personal advertisements. *Sex Roles, 52*, 447-461.
- Hancock, J., Curry, L, Goorha, S., & Woodworth, M. (2008). On lying and being lied to: An automated linguistic analysis of deception. *Discourse Processes, 45*, 1-23.
- Hasler, B., Mehl, M. R., Bootzin, R., & Vazire, S. (2008). Preliminary evidence of diurnal rhythms in everyday behaviors associated with positive affect. *Journal of Research in Personality, 42*, 1537-1546.
- Holleran, S. E., & Mehl, M. R. (2008). Let me read your mind: Personality judgments based on a person's natural stream of thought. *Journal of Research in Personality, 42*, 747-754.
- Ireland, M. E. & Mehl, M. R. (in press). Language use and personality. In: Holtgraves, T. (Ed.) *Oxford Handbook of Language and Social Psychology*. New York: Oxford University Press.
- John, O. P & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*, 521-551.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 102-138). New York: Guilford Press.
- Kenny, D. A. (1987). *Statistics for the Social and Behavioral Sciences*. Little, Brown and Company: Boston, MA.
- Kosloff, S., Solomon, S., Greenberg, J., Cohen, F., Gershuny, B., Routledge, C., & Pyszczynski, T., (2006). Fatal distraction: The impact of mortality salience on dissociative responses to 9/11 and subsequent anxiety sensitivity. *Basic and Applied Social Psychology, 28*, 349-356.
- Lacan, J. 1968. *The Language of the Self: The Function of Language in Psychoanalysis*. Baltimore: Johns Hopkins Press.
- Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Quantity and quality of available information affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology, 91*, 111-133.
- Lieberman, M. A., & Goldstein, A. B. (2006). Not all negative emotions are equal: The role of emotional expression in online support groups for women with breast cancer. *Psycho-Oncology, 15*, 160-168.
- Lucas, R. E., & Fujita, F. (2000). Factors influencing the relation between extraversion and pleasant affect. *Journal of Personality and Social Psychology, 79*, 1039-1056.
- Maass, A., Karasawa, M., Politi, F., & Sayaka, S. (2006). Do verbs and adjectives play different roles in different cultures? A cross-linguistic analysis of person perception. *Journal of Personality and Social Psychology, 90*, 734-750.
- Marshall, M. A., & Brown, J. D. (2006). Trait aggressiveness and situational provocation. A test of the "Traits as Situational Sensitivities (TASS)" Model. *Personality and Social Psychology Bulletin, 32*, 1100-1113.

- McAdams, D. P., Anyidoho, N. A., Brown, C., Huang, Y. T., Kaplan, B., & Machado, M. A. (2004). Traits and stories: Links between dispositional and narrative features of personality. *Journal of Personality, 72*, 761-784.
- Mehl, M. R. (2005). Quantitative text analysis. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp.141–156).
- Mehl, M. R. (2006). The lay assessment of sub-clinical depression in daily life. *Psychological Assessment, 90*, 340-345.
- Mehl, M. R. (2007). Eavesdropping on health: A naturalistic observation approach for social-health research. *Social and Personality Psychology Compass, 1*, 359-380.
- Washington, DC: American Psychological Association.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90*, 862–877.
- Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology, 84*, 857 – 870.
- Mehl, M. R., Pennebaker, J. W., Crow, M., Dabbs, J., & Price, J. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers, 33*, 517 – 523.
- Mehl, M. R., Robbins, M. L., & große Deters, F. (2012). Naturalistic observation of health-relevant social processes: The Electronically Activated Recorder (EAR) methodology in psychosomatics. *Psychosomatic Medicine, 74*, 410-417.
- Mehl, M. R., Vazire, S., Holleran, S. E., & Clark, C. S. (2010). Eavesdropping on happiness: Well-being is related to having less small talk and more substantive conversations. *Psychological Science, 21*, 539-541.
- Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science, 317*, 82.
- Mergenthaler, E. (1996). Emotion-abstraction patterns in verbatim protocols: A new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology, 64*, 1306-1315.
- Mitzner, T. L., & Kemper, S. (2003). Oral and written language in late adulthood: Findings from the Nun Study. *Experimental Aging Research, 29*, 457-474.
- Mulac, A. & Lundell, T. L. (1994). Effects of gender-linked language differences in adults' written discourse: Multivariate tests of language effects. *Language & Communication, 14*, 299-309.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*, 665-675.
- Oberlander, J. & Gill, A. (2006). Language with character: A stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes, 42*, 239-270.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D.N. Jackson, & D.E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 67-88). Hillsdale, NJ: Erlbaum.

- Pennebaker, J. W. (2011). *The Secret Life of Pronouns*. Bloomsbury Press: New York, NY.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC2001*. Mahwah, NJ: Lawrence Erlbaum Associates.
- 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. & Lay, T. C. (2002). Language use and personality during crises: Analyses of Mayor Rudolph Giuliani's press conferences. *Journal of Research in Personality*, *36*, 271-282.
- 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.
- Ricoeur, P. (1976). *Interpretation Theory: Discourse and the Surplus of Meaning*. Fort Worth, TX: Texas Christian Univ. Press.
- Robbins, M. L., Focella, E. S., Kastle, S., Weihs, K. L., Lopez, A. M., & Mehl, M. R., (2011). Naturalistically observed swearing, emotional support and depressive symptoms in women coping with illness. *Health Psychology*, *30*, 789-792.
- Robbins, M. L., Mehl, M. R., Holleran, S. E., & Kastle, S. (2011). Naturalistically observed sighing and depression in rheumatoid arthritis patients: A preliminary study. *Health Psychology*, *30*, 129-133.
- Rude, S. S., Gortner, E. M., & Pennebaker, J. W. (2004). Language use of depressed vulnerable college students. *Cognition and Emotion*, *18*, 1121-1133.
- Scherwitz, L., Graham, L.E., & Ornish, D. (1985). Self-involvement and the risk factor for coronary heart disease. *Advances*, *2*, 6-18.
- Sillars, A. L., Shellen, W., McIntosh, A., & Pomegranate, M. A. (1997). Relational characteristics of language: Elaboration and differentiation in marital conversations. *Western Journal of Communication*, *61*, 403-422.
- Simmons, R., Gordon, P.C., & Chambless, D. (2005) Pronoun use in marital interaction: What do "you" and "I" say about marital health? *Psychological Science*, *16*, 932-936.
- Slatcher, R. B., & Pennebaker, J. W. (2006). How do I love thee? Let me count the words: The social effects of expressive writing. *Psychological Science*, *17*, 660-664.
- Slatcher, R. B., & Trentacosta, C. J. (2011). A naturalistic observation study of the links between parental depressive symptoms and preschoolers' behaviors in everyday life. *Journal of Family Psychology*, *25*, 444-448.
- Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 883-947). Reading, MA: Addison-Wesley.
- Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and non-suicidal poets. *Psychosomatic Medicine*, *63*, 517-522.
- Tsai, J. L., Simenova, D., & Watanabe, J. (2004). Somatic and Social: Chinese Americans Talk About Emotion. *Personality and Social Psychology Bulletin*, *30*, 1226-1238
- Vazire, S. & Gosling, S. D. (2004). e-Perceptions: Personality impressions based

## CONTEXT DEPENDENT TRAIT IMPLICATIONS OF LANGUAGE USE

- on personal web-sites. *Journal of Personality and Social Psychology*, 87, 123–132.
- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, 78, 546–558.
- Weintraub, W. (1989). *Verbal behavior in everyday life*. New York: Springer-Verlag.
- Wiener, M. & Mehrabian, A. (1968). *Language within language: Immediacy, a channel in verbal communication*. New York: Appleton-Century-Crofts.