



Cues in computer-mediated communication: A corpus analysis

Monica A. Riordan*, Roger J. Kreuz

The University of Memphis, 202 Psychology Building, Memphis, TN 38152, USA

ARTICLE INFO

Article history:

Available online 27 July 2010

Keywords:

Nonverbal
Email
Chat
Blogs
Cues
Online communication

ABSTRACT

An analysis of five contemporary corpora examines the use of several different cues in four channels of computer-mediated communication. With an in-depth corpus analysis, we show that a wealth of cues is available in online communication, and that these cues are often matched with words that have particular functions and/or semantic meanings. Using the Linguistic Inquiry and Word Count text analysis software (Pennebaker et al., 2007), we found the two largest categories represented by cue-laden words involved affect and cognitive mechanisms, suggesting that cues are largely used to indicate emotion or to disambiguate a message. We argue that learning the meaning of these cues is central to learning how people communicate nonverbally while online.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

In face-to-face communication, nonverbal cues such as eye contact, gaze, vocal intonation, and gestures can be reliable indicators of a speaker's personality, abilities, sexual orientation, and gender (for a review, see Ambady, Bernieri, & Richeson, 2000). Such nonverbal behavior may provide information, regulate interaction, and express intimacy (e.g., Ekman & Friesen, 1969; Harrison, 1973). Many of these cues are absent in text-based computer-mediated communication (CMC), making communication potentially ambiguous. Kiesler, Siegel, and McGuire (1984) assert that many nonverbal behaviors present in face-to-face communication are used to regulate, modify, and control the message being communicated; the absence of these behavioral cues in CMC may result in miscommunication. Indeed, Burgoon, Buller, and Woodall (1996) reported that nonverbal cues account for more of a receiver's perception of a sender's affect than verbal content does.

Earlier research on this topic suggested that CMC cannot transmit emotion (e.g., Calhoun, 1991; Dubrovsky, Kiesler, & Sethna, 1991; Siegel, Dubrovsky, Kiesler, & McGuire, 1986); however, more recent research and emerging theories suggest that CMC is able to do so—sometimes unintentionally (Thompson & Foulger, 1996; Walther & D'Addario, 2001). This recent research suggests that users of CMC adapt to the lack of cue systems available as opposed to face-to-face communication. However, little research exists to explain *how* communicators make this adaptation. In this study, we attempt to address this gap in the research by examining five different CMC corpora for nonverbal cues. We also relate these cues to specific word functions and semantic meanings, showing that cues are not used indiscriminately in CMC.

2. Literature review

Harris and Paradice (2007) assert that in CMC—as in face-to-face communication—two types of cues, nonverbal and verbal, are available to encode and decode emotions. Verbal cues consist of the same language that is available in face-to-face communication; the words and sentences that are spoken or written. Some studies have examined verbal CMC, exploring word counts within semantic categories (e.g., Hancock, Landrigan, & Silver, 2007), or the use of assertions and qualifiers (e.g., Guiller & Durnell, 2006). However, researchers have failed to explore the second important part of communication: the nonverbal cues.

Carey (1980) identified five categories of nonverbal cues in CMC: vocal spelling, lexical surrogates, spatial arrays, manipulation of grammatical markers, and minus features. Vocal spelling such as “weeeeeelllllll” and lexical surrogates such as “mhhh” use nonstandard spelling that imitates vocal intonation or tone. Examples of spatial arrays include emoticons such as :-) and are generally a sequence of keyboard characters that represent nonverbal behaviors, such as facial expressions. Manipulated grammatical markers such as additional punctuation and capital letters may indicate pauses (...), express attitude (!!!), or signal tone of voice (SHOUT). Minus features refer to an absence of certain language standards that are present in normal writing such as a lack of capitalization at the beginning of a sentence. Harris and Paradice (2007) suggest that characteristics of CMC such as those outlined above provide information about the type and degree of emotion the message sender intends to convey.

Crystal (2001) describes a host of ways in which symbols are used as cues in online conversation without any actual data on the use of the cues themselves. Researchers who have collected such data have largely focused on non-naturalistic CMC generated in a lab (e.g., Hancock et al., 2007), conversations that were

* Corresponding author.

E-mail address: mschpers@memphis.edu (M.A. Riordan).

generated with the knowledge they would be read by researchers (e.g., Fox, Bukatko, Hallahan, & Crawford, 2007), or CMC with a higher level of formality than those in which cues might be expected (e.g., Kalman & Gergle, 2009, who examined business emails in the Enron corpus). Indeed, Derks, Bos, and von Grumbkow (2008) found that emoticons are used more in social contexts than in task-oriented contexts such as is found in business emails.

The research to date exploring the use of nonverbal cues in CMC has focused on only two specific kinds: emoticons (e.g., Walther & D'Addario, 2001) and chronemic cues (time-related cues; e.g., Walther & Parks, 2002). Yet other cues have been identified or speculated upon; Crystal (2001) notes the existence of several types of paralinguistic and prosodic cues available online such as asterisks, capitalized words, and repeating letters. Fox et al. (2007) examined instant message conversations for italicized words, repeating letters, and exclamation points. Kalman and Gergle (2009) examined the use and role of vocal spelling and repeating punctuation marks. However, the variation in sample sizes and types of CMC studied in previous research does not offer conclusive evidence about what kinds of cues are used and how frequently they are employed.

CMC allows for use of underlined and capitalized text along with a range of other cues; use of special symbols such as brackets, braces, and asterisks may also play a role in determining the message sender's intentions. Indeed, Kreuz (1996) asserts that certain cues, such as capitalization, underlining, and emoticons, play a role in denoting irony in written communication.

Despite the availability of nonverbal cues in CMC, research from the 1990s shows that their actual use remains infrequent (e.g., Rezabek & Cochenour, 1998; Witmer & Katzman, 1997). However, more recent research refutes these findings, showing cue use to be relatively frequent. For example, Riordan and Kreuz (submitted for publication) established that participants believe they use cues quite often: their results show that on average, 30% or more of participants indicated having employed a variety of cues in their use of email and chat programs. It may be that as technology has become more pervasive during the past decade, cue use has increased. The increasing use of these nonverbal cues in CMC permits an exploration of the amount and kind of cues being used, further demonstrating how rich CMC can be to its users.

While several researchers have paid lip service to the existence of nonverbal cues in online conversation, many offering anecdotal evidence, few have attempted to establish a rate of use of these cues or to examine the role of these cues in the conversation. While Rezabek and Cochenour (1998) and Witmer and Katzman (1997) found low rates of cue use, more than a decade has passed since these baselines were established. The increasing pervasiveness of CMC as a standard in business, scholarship, and social lives, including the increasing rate of use of text messaging via mobile phones, suggests that the rate of cue use may have risen and/or that the type of cues used in CMC may have become more diverse.

In the current study, we examine four research questions:

RQ1: What are the base rates of certain cues in CMC?

RQ2: Are cues used differentially among different CMC channels?

RQ3: Are cues used in conjunction with words that have specific functions?

RQ4: Are cues used in conjunction with specific semantic categories?

These questions are addressed using five different corpora, within which the great majority of messages are social in character and naturalistic in nature (i.e., not created with the expectation of being examined by researchers).

3. Methods

Five corpora were examined for vocal spelling, emoticons, and the manipulation of several forms of punctuation: repeating exclamation points and question marks, angled brackets, curly braces, asterisks, underscores, tildes, and capitalized words.

3.1. The corpora

Five corpora were downloaded from the Internet in January 2010. These corpora represent several different CMC channels (blogs, email, chat rooms) and topics as well as methods of collection (a study with college students, a World Wide Web crawl, and listservs). All five are exchanges among people in social contexts. Only one corpus contains CMC data that were collected from participants with the knowledge that its contents would be read by researchers.

3.1.1. AIR-L corpus

This listserv consists of emails among members of the Association of Internet Researchers. The archive begins in May 2001. The corpus used here, which contains 5770 emails, begins in April 2008 and extends to January 2010 and can be downloaded at <http://www.listserv.aoir.org/pipermail/AIR-L-aoir.org/>. This corpus consists of 2,001,256 words and 14,253,014 characters (duplicates not removed).

3.1.2. British Columbia conversation corpus

The BC3 consists of all email threads from the World Wide Web Corpus. The W³C is composed of over 200,000 files gathered from a "crawl" of the World Wide Web Consortium's sites in 2005 and 2006. The files include mailing lists, public webpages, text from .pdf, .ppt, and .doc files, and more. The BC3 is a subset of this larger corpus, consisting only of email. More specifically, 40 email threads exist in the BC3 corpus. This corpus consists of 43,374 words and 382,751 characters (duplicates not removed).

3.1.3. Chalkhills corpus

The Chalkhills Digest is a listserv originally dedicated to discussing the 1980s band XTC, though its discussion has broadened to movies and music of many types. The archive, which can be accessed at <http://www.chalkhills.org/digests/index-01.html>, begins in April 1989. The corpus used here consists of 391 emails, spanning January 2008 to January 2010, with 83,037 words and 643,226 characters (duplicates not removed).

3.1.4. Loyola College corpus

The Loyola College corpus consists of over 900 texts generated between September 2006 and December 2007. Participants were given one of six predetermined topics for discussion: gender discrimination, the legalization of marijuana, gay marriage, pedophilia in the Catholic Church, privacy rights in schools, and the Iraq war. The participant was to address the topic in one of six channels of communication: blog, online chat, discussion, email, essay, and interview. The corpus webpage and information text file explicitly note that the corpus was not modified in any way (although discussions were transcribed, of course).

While the corpus includes six types of communication channels, only three were used in this analysis: blog, online chat, and email. These three channels are computer-mediated and are not subject to any possible subjectivity that may occur in transcription. The analyzed portions of the corpus consist of blogs including 54,594 words and 310,377 characters (among 24 participants with 1 blog each); online chat consisting of 177,556 words and 973,933 characters (among six chat groups with six participants each); and

Download English Version:

<https://daneshyari.com/en/article/351532>

Download Persian Version:

<https://daneshyari.com/article/351532>

[Daneshyari.com](https://daneshyari.com)