



# Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations



Jennifer Hill <sup>a,\*</sup>, W. Randolph Ford <sup>b,1</sup>, Ingrid G. Farreras <sup>c,2</sup>

<sup>a</sup> School of Engineering and Applied Science, The George Washington University, 800 22nd St. NW, Washington, D.C. 20052, USA

<sup>b</sup> Data Analytics Program, Graduate School, 3501 University Boulevard East, University of Maryland University College, Adelphi, MD 20783, USA

<sup>c</sup> Psychology Dept., Hood College, 401 Rosemont Ave., Frederick, MD 21701, USA

## ARTICLE INFO

### Article history:

Available online 20 March 2015

### Keywords:

CMC  
Instant messaging  
IM  
Chatbot  
Cleverbot

## ABSTRACT

This study analyzed how communication changes when people communicate with an intelligent agent as opposed to with another human. We compared 100 instant messaging conversations to 100 exchanges with the popular chatbot Cleverbot along seven dimensions: words per message, words per conversation, messages per conversation, word uniqueness, and use of profanity, shorthand, and emoticons. A MANOVA indicated that people communicated with the chatbot for longer durations (but with shorter messages) than they did with another human. Additionally, human–chatbot communication lacked much of the richness of vocabulary found in conversations among people, and exhibited greater profanity. These results suggest that while human language skills transfer easily to human–chatbot communication, there are notable differences in the content and quality of such conversations.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Artificial intelligence's (A.I.) efforts in the last half century to model human language use by computers have not been wildly successful. While the idea of using human language to communicate with computers holds merit, A.I. scientists have, for decades, underestimated the complexity of human language, in both comprehension and generation. The obstacle for computers is not just understanding the meanings of words, but understanding the endless variability of expression in how those words are collocated in language use to communicate meaning.

Nonetheless, decades later, we can find an abundance of natural language interaction with intelligent agents on the internet, from airline reservation systems to merchandise catalogs, suggesting that humans have little or no difficulty transferring their language skills to such applications. Because so much of this communication occurs through digital technology rather than in person, computer-mediated communication (or “CMC”) has become a prominent area of research in which to explore this simulation of natural human language.

One of the most popular forms of CMC today, particularly among adolescents and teenagers, is instant messaging (IM) (Tagliamonte & Denis, 2008). While many specialized applications enable instant messaging, the service is also provided through many other popular media, such as multiplayer online games, email clients, and social networking websites (Varnhagen et al., 2009).

Several studies have compared IM and other forms of CMC to other forms of language. Ferrara, Brunner, and Whittemore (1991) determined that CMC possesses uniquely distinguishing linguistic features that display qualities of both written and spoken dialogue. Compared to other standard forms of communication, CMC's most distinctive trait is its unique, shortened-form language of acronyms and abbreviations, and an informal discursive style that is similar to face-to-face spoken language (Werry, 1996). CMC differs from spoken communication, however, in its lack of cues from features such as body language, communicative pauses, and vocal tones (Hentschel, 1998). Despite this absence of cues, however, CMC has been found to be able to communicate emotion as well as or better than face-to-face communication (Derks, Fischer, & Bos, 2008).

Although CMC has been compared to other forms of communication, few studies have compared different forms of CMC to one another. Perhaps the most noteworthy of these studies is Baron's (Baron, 2007) comparison of the linguistic characteristics of IM and text (or SMS) messages – another form of CMC – among

\* Corresponding author. Tel.: +1 301 693 4579.

E-mail addresses: [jenhill@gwu.edu](mailto:jenhill@gwu.edu) (J. Hill), [rand.ford@umuc.edu](mailto:rand.ford@umuc.edu) (W. Randolph Ford), [farreras@hood.edu](mailto:farreras@hood.edu) (I.G. Farreras).

<sup>1</sup> Tel.: +1 240 684 5606.

<sup>2</sup> Tel.: +1 301 696 3762.

American college students, which found that the average text message contained more words, characters, sentences, abbreviations, and contractions than the average instant message.

To our knowledge, however, no research has investigated the linguistic characteristics of a different form of CMC: chatbot communication. Chatbots, or chatterbots, are another widespread domain of CMC. Chatbots are “machine conversation system[s] [that] interact with human users via natural conversational language” (Shawar & Atwell, 2005, p. 489). Users interact with these applications primarily to engage in small talk. Functionally, their approach to natural language processing is an extension of the same technique used in Weizenbaum’s ELIZA (Weizenbaum, 1966). A variety of new chatbot architectures and technologies (e.g., Ultra Hal, ALICE, Jabberwacky, Cleverbot) have arisen recently, each attempting to simulate natural human language more accurately and thoroughly (Carpenter, n.d. a; Shawar & Atwell, 2007; Wallace, Tomabechei, & Aimless, 2003; Zabaware., n.d.).

Despite the popularity of chatbots today, we are not aware of any research analyzing how humans converse with them, particularly from a linguistic perspective. Several extant studies on chatbots have focused on developing or improving their ability to interpret and respond meaningfully to human language: one study examined a chatbot’s ability to respond correctly when faced with common CMC features like abbreviations and overlapping utterances from multiple speakers (Shawar & Atwell, 2005), while another examined a chatbot’s robustness when faced with unconventional linguistic features from non-native ESL speakers (such as misspellings and incorrect word order) (Coniam, 2008). Another area of research has focused on evaluating users’ attribution of human qualities or personality traits to the chatbots they converse with, and how that may lead to greater disclosure in medical, research, or therapeutic settings (Hasler, Tuchman, & Friedman, 2013; Holtgraves, Ross, Weywadt, & Han, 2007; Lortie & Guitton, 2011). Lortie and Guitton (2011) specifically investigated how judges go about distinguishing between humans and computers when interacting in a formal Turing Test. They tracked several descriptive and cognitive parameters along with indicators of interest using the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Chung, Ireland, Gonzales, and Booth (2007)) program. Their results suggested that judges’ determinations of humanness were associated with communication that contains more words per message, a higher percentage of articles, and a higher percentage of words that were longer than six letters. Such communication, however, is biased by the goal-directed behavior of judges trying to figure out if they are talking to a computer.

## 2. Current study

The purpose of this study was to investigate how users’ explicit and implicit transfer and expectations of human language is manifested in human–computer interaction. Specifically, we sought to answer an unexplored question in the fields of both computer-mediated communication and chatbot development: do humans communicate differently when they know their conversational partner is a computer as opposed to another human being?

To accomplish this, we compared 100 random human IM conversations against 100 random human conversations with Cleverbot. First launched on the internet by Rollo Carpenter in 1997, Cleverbot is one of the most popular chatbots in existence today. It is designed to simulate natural conversation by learning from human interaction, relying on feedback to communicate (Carpenter, n.d. b). Using a synchronous one-to-one online interface reminiscent of an instant messenger, a human user types a statement and then waits for Cleverbot to send a response.

**User:** Hi.

**Cleverbot:** Do you like cheese?

**User:** No.

**Cleverbot:** Why not?

**User:** Not a big fan.

**Cleverbot:** You like animals?

**User:** Yes. I like cats.

**Cleverbot:** Do you like me the best though?

**User:** I have to get to know you more to see.

**Cleverbot:** Do you like herman cain?

**User:** No.

**Cleverbot:** Why?

**User:** Cause's I dont know nothing about him.

**Cleverbot:** My name is Cleverbot.

**Fig. 1.** A typical conversation with Cleverbot.

Cleverbot is unique in that it “learns” what humans have said to it in previous conversations and uses this knowledge to determine how to respond in new conversations. Figure 1 demonstrates a typical conversation with Cleverbot.

Cleverbot passed the 2011 Turing Test at the Techniche Techno-Management Festival held by the Indian Institute of Technology Guwahati. Of the 1334 volunteers who participated in four-minute typed conversations with either Cleverbot or real humans, 59% rated Cleverbot as human while 63% rated the real humans as human (Aron, 2011), suggesting that Cleverbot is one of the most advanced and human-like conversational agents currently available to the public.

In our comparison of 100 random human IM conversations against 100 random human conversations with Cleverbot, we examined the amount of written content, the uniqueness of the words used, the frequency of profane language, and the use of standard CMC linguistic features such as shorthand phrases (e.g., “lol”) and emoticons (e.g., “:-”). We expected the linguistic profiles of the human–human IM conversations to be consistent with those of previous studies. We also hypothesized that people would send fewer messages, write fewer words per message, and exhibit a more limited vocabulary when communicating with a chatbot compared to another person. We expected this based on three characteristics found in human–chatbot conversations that are not found in human–human conversations. First, people have less experience communicating with chatbots than with humans, and so would be less confident and comfortable with the communicative ability of a chatbot. Second, while chatbots are designed to sustain conversations, they are limited in their ability to have an

Download English Version:

<https://daneshyari.com/en/article/350323>

Download Persian Version:

<https://daneshyari.com/article/350323>

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