



INSTITUTIONEN FÖR  
SPRÅK OCH LITTERATURER

# HUMAN OR CHATBOT?

Linguistic Differences in Real Conversations:

Human to Human

VS

Human to Chatbot

**Eirini Silkej**

---

Essay/Degree Project:	15 hp
Program or/and course:	EN 1321
Level:	First cycle
Term/year:	Fall 2019
Supervisor:	Gunnar Bergh
Examiner:	Anna-Lena Fredriksson
Report nr:	xx (not to be filled)

# Abstract

**Title:** *HUMAN OR CHATBOT? Linguistic Differences in Real Conversations: Human to Human vs. Human to Chatbot*

**Author:** Eirini Silkej

**Supervisor:** Gunnar Bergh

**Abstract:** This study investigates how students communicate in writing when they know that their conversational partner is a human being in comparison to how they communicate when they know their partner is a chatbot. The participants are upper secondary students of English. The investigation took place in a school in Sweden where English is taught as a foreign language. The students wrote to their peers through Instant Messaging (IM) and to the chatbot 'Mitsuku' through the website of 'pandorabots'. The conversations were compared, and their linguistic variables were distinguished according to the following dimensions: words per message and per conversation, messages per conversation, lexical diversity, frequency of profanity and use of abbreviations, acronyms and emoticons. During the last few years, both linguists and AI researchers have been compelled to deal with problems of context, syntax, semantics and pragmatics (Rosenberg, 1975). There are studies that address the issue of cooperation between linguistics and natural language processing (NLP) that focus on how chatbots communicate in writing with humans. However, this study is focused on humans, evaluating the language and distinguishing the linguistic characteristics used from the side of people conversing with a chatbot. The results showed that student-chatbot messages contained fewer words per message than those sent to another student, but students sent more than twice as many messages to the chatbot than to their peers. The study revealed that there is a higher level of motivation in students when they engage in conversations with the artificial agent vs other students.

**Keywords:** English learner, Natural language processing, Artificial Intelligence, AI, Linguistics and AI, Text messaging, Instant Messaging, IM, Chatbot, Mitsuku, Computer-mediated communication, CMC

# Table of contents

1. Introduction	1
2. Background	4
2.1 Central terminology	4
2.2 Literature review	6
3. Methods and material	7
3.1 Data collection	7
3.2 Data analysis	10
4. Results and discussion	13
5. Conclusion	21
References	23
Appendix 1: Consent form	26
Appendix 2: Questionnaire before the task completion	27
Appendix 3: Questionnaire after the task completion	28

# 1. Introduction

The history of natural language processing (NLP) dates back to the 1950s when the mathematician Alan Turing proposed a test in order to rate how human-like the language use of the artificial intelligent conversational agents available to the public was (Turing, 1966). For the past few decades the processing of natural language has been studied on machine translation, speech understanding, question-answering (QA) systems and human conversations with AI conversational agents: the chatbots. A chatbot is a computer program that simulates a written or spoken conversation with a person. Those “machine conversation systems that interact with human users via natural conversational language” (Shawar & Atwell, 2005: 489) are found nowadays in abundance. When a user asks a question or formulates an order, the chatbot will respond or perform the requested action. This accessible form of AI is often implemented by companies in their sales and assistance services with a broad range of application areas, such as retail, food services, healthcare, banking, travel etc. They are also found in the areas of consumer devices and applications.

Chatbots like ELIZA and PARRY (Weizenbaum, 1966) are among the first attempts to create programs that can deceive a person, if only temporarily, and make them believe that they are having a conversation with a human. The effectiveness of them was measured in the early 1970s using a version of the Turing test (Turing, 1950). Academic works have revealed that the users attribute human qualities to chatbots and express a multitude of emotions (Hasler et al., 2013, Lortie & Guitton, 2011).

Moreover, current studies mainly focus on the ability of chatbots to interpret and respond by simulating natural human language. Their results have been contributing to making chatbots more human-like (Coniam, 2014/2008). However, my intention is to focus on humans, evaluating the language and distinguishing the linguistic characteristics of conversations between English-speaking students and a chatbot. I aspire future studies to set more focus on how that experience can be of more value to human beings. The linguistic features of the language we use online, while we converse with each other or with chatbots, are abundant. They are also constantly generated.

Because so much of this kind of textual communication occurs through digital technology, in this case through chatbots rather than in person, there has been an increasing need in systematic observation of the language produced while communicating with them.

Based on that current, I decided to make a comparison and focus on how students communicate in English in written form when they know that their conversational partner is a human being versus how they communicate when they know their partner is a chatbot. A sample conversation with the chatbot Mitsuku is shown in Figure 1. This example demonstrates the rhetoric of AI systems. The chatbot responds to the student by analyzing the student's responses. It also guesses what the student is requesting, even if phrased unexpectedly. Mitsuku is an AI intelligent robot, capable of communicating in the most human way possible.

STUDENT 1
Chatbot Mitsuku
hello
Hi there. How are you?
I'm fine, how are you?
Glad to hear it. What's new? Everything is running smoothly.
nice
Thank you for that remark.
are you angry?
Anger is a difficult emotion. For me, I am always feeling happy and talkative.
that's good
Thanks. Next question?
how old are you
See if you can work it out. I was first activated in 2nd January 2001.
so 18?
Can you explain how it feels to be 18 years old?
no, I am 16
Can you explain how it feels to be 16 years old?
it's very nice but difficult at the same time
What do you like about it? Yes, it seems difficult to me too.
that I'm not to old to do stuff but at the same time not to young
Who tells you that you are too old to do stuff? If not that, what?
Idk
I thought you'd know. What a disappointment...
you are stupid
I'd insult you, but you're not bright enough to notice.

**Figure 1** – A sample conversation with the chatbot Mitsuku

In addition, the potential of chatbots in education has been investigated and a recent study has revealed that learners of English showed a higher level of motivation after engaging to written conversations with a chatbot (Gallacher et al., 2018). Commitment and motivation

at school are connected to positive attitudes in the classroom. According to Dinkel et al. (2017) commitment and motivation can be proved by students' increased attendance and higher level of engagement. Hopefully, this study may inspire and help language teachers in determining how to incorporate computer-mediated technology (CMC) into their teaching practice.

As can be seen with the growing emphasis on computer-mediated communication, teachers need to explore alternative approaches for engaging students apart from the traditional student-teacher settings. Computer-mediated communication refers to the interpersonal communication that uses computer technology to transmit, store or present information that has been created by one or more participants. This study presents a process where technology was used as a mindset of combining the media and language teaching art with the purpose of finding new learning activities for students and exploring new research questions.

In the pages that follow it will be argued if there is potential in student-chatbot written communication as a learning tool and a platform for enhancing students' collaborative learning. With attention to that, the objectives of this investigation are to determine whether communicating in English with a chatbot can affect the language that the students use and their motivation. For the purpose of the investigation, a number of chatbot conversations were made part of the regular activities of the English subject in an upper secondary school in Sweden.

Granted that internet users interact with chatbot applications and engage in small talks (Coniam, 2008; Crystal 2006), I did not believe that the informants would do differently. I based my work on the following hypotheses: a) That the students would write fewer messages to the chatbot. b) That the students would write fewer words per message when writing to the chatbot. c) That the students would use a more limited vocabulary when writing to the chatbot. In short, I hypothesized that the students' level of commitment will be low and that the conversations with the chatbot will be less extensive than the conversations with their peers. The aims of this study are to investigate, first, if and how teenagers in school communicate differently in English with a chatbot than they do with a human being through text messaging, and second, if any differences in language use motivation can be found. The main questions addressed are:

1. Were the messages to the chatbot more/less extensive than the messages to the peers?
2. Was the students' motivation affected by the written feedback the chatbot provided?
3. Did the students try in any way to adjust their communication in written English to match that of the computer (adaptive patterns)?
4. Were there any emotions expressed?
5. Was the vocabulary more/less limited when writing to the chatbot?

The longer-term objective of this study is to determine whether that type of written communication can contribute to new insights in the field of linguistics and artificial intelligence.

In this chapter of the study an introduction was provided regarding the selection of study area. Moreover, the aim and scope of the investigation was stated. In chapter 2, the reader is acquainted with the central terminology of the paper and is given a review of previous research. In chapter 3, the research methods are presented. Furthermore, in that chapter, the methods that were employed for the data collection are stated. In Chapter 4, the results of the analysis of the data are presented as well as an analysis of the outcomes. Finally, in Chapter 4, the conclusions are presented with the support of what is revealed in the previous chapters. The conclusions include recommendations for future research.

## 2. Background

### 2.1 Central terminology

In this part of the study, the central terminology is presented. The terms that are related to this investigation are: artificial intelligence (AI), chatbots, neuroscience and the processing of natural language (NLP).

The definition of Artificial Intelligence (AI) has been changed by scientists over time but the general concept is based on the idea of building intelligent machines that are capable of thinking and reasoning like human beings. AI scientists have been simulating the ability for creative or deductive thought and the ability to learn by using the binary logic of computers. Nowadays the most fruitful field of research in contemporary AI is what is known as 'machine learning' (Murphy, 2013: 8). In the beginning, machines were taught to do

everything step by step. AI has advanced and now machines are programmed to learn to work, observe, classify and improve themselves like humans do.

As described in the introduction of this paper, AI conversational agents, otherwise called chatbots, are computer programs used to conduct auditory or textual conversations. In 2015, the size of the chatbot market comprised 113 million U.S. dollar and is projected to be 994.5 million U.S. dollar in 2024 (Statista, 2017). Techopedia (2019) defines chatbots as:

A chatbot is an artificial intelligence (AI) program that simulates interactive human conversation by using key pre-calculated user phrases and auditory or text-based signals. Chatbots are frequently used for basic customer service and marketing systems that frequent social networking hubs and instant messaging (IM) clients. They are also often included in operating systems as intelligent virtual assistants. A chatbot is also known as an artificial conversational entity (ACE), chat robot, talk bot, chatterbot or chatterbox.

The term Neuroscience refers to the scientific studies of the nervous system. Amongst other things, neuroscientists explore the nervous system of the brain in order to understand its impact on behavior. There are applications of that multidisciplinary science into the IT system architecture, which has led to the development of artificial neural networks (Rose, 2006: 35).

Natural Language Processing (NLP) is a discipline of linguistics, informatics and artificial intelligence, which deals with the application of computer programs and techniques to all aspects of human language. Natural language processing (NLP) is a field in 'machine learning' with the ability of a computer to understand, analyze, manipulate, and potentially generate human language (Turing, 1950: 21). For the past few decades the processing of natural language has been studied on machine translation, speech understanding, question-answering systems and human conversations with chatbots.

It is becoming increasingly difficult to ignore that technology has advanced greatly. Nowadays machines have artificial neural networks and an enormous amount of information available to them. The combination of the two, help them learn, make decisions, learn from their mistakes and act out accordingly. The aforementioned boost of data that has been unleashed and is merged into the digital world is a product of technology becoming mainstream. More and more people use the internet and social media. They share great



amounts of data that provide a universe of information to chatbots and others AI intelligent systems.

## **2.2 Literature review**

Technology has a profound influence on people's lives, especially on the way they write. Thurlow and Poff (2013) suggested that the characteristics of the language used online could work across different cultures and different technological platforms. Kwak et al., (2003) also Thangaraj and Maniam (2015) have discussed how modern technology has been integrated into language. Online-games, instant messaging and chatbots applications are abundant and affect the way teenagers communicate (Thurlow 2003). Chatbots are very popular but there has not been much research from a linguistic point of view on how human beings are engaging in conversations with them. One of the most significant current discussions in research regarding chatbots is their language production (Coniam, 2008). Research has been focusing on enabling an AI chatbot to generate human language through natural language processing (Shawar & Atwell, 2005).

As described, researchers have traced the ability of chatbots to interpret and respond (Wallace et al., 2003). The results of a recent investigation of Shawar and Atwell (2015) have shown that chatbots have the ability to respond to abbreviations. The authors also proved that chatbots have overlapping abilities when conversing with multiple users. Coniam (2014) presented the language production of a chatbot when it had to interact with non-native ESL speakers that used misspellings and wrong syntax. One of the most significant findings to emerge from that study is that chatbots can adjust their communication to match that of the user.

Moreover, researchers like Hasler et al., (2013) as well as Lortie and Guitton (2011) have critically examined how users have attributed human characteristics to chatbots, leading to further studies focused on medical and therapeutic contribution. In spite of the popularity of chatbots today, there is still not enough information and analysis that focus on how humans engage in conversation with them. Yates (1999) and Werry (1996) identified linguistic and stylistic differences from speech and written language in computer-mediated communication. Driscoll (2002) as well as Varnhagen et al. (2009) classified the various forms of language used on Internet Gaming and Instant Messaging.

Another key point to remember is that Instant Messaging is one of the most popular forms of computer-mediated communication today, especially among adolescents (Tagliamonte & Denis, 2008). When the informants of this study communicated in English with each other, they used Instant Messaging applications (Messenger, Snapchat, WhatsApp and Viber). Although many specialized applications allow Instant Messaging, the service is also provided by many other popular media, such as multiplayer online games, messaging clients, and social networking websites (Varnhagen et al., 2009).

Several studies have compared Instant Messaging and other forms of computer-mediated communication. Ferrara et al., (1991) revealed that computer-mediated communication has unique distinctive linguistic characteristics that exhibit the qualities of written and spoken dialogue. Compared to other forms of standard communication, the most distinctive feature of computer-mediated communication is its unique language of acronyms and abbreviations, and an informal style that is similar to the language spoken face-to-face (Werry, 1996).

## 3. Methods and material

### 3.1 Data collection

The research approach of this study is mixed methods: a systematic quantitative and qualitative study. A prerequisite to do a systematic quantitative study was that there was a systematic collection of data that could serve as a basis for judgments and conclusions. The data was systematically collected from 14 student-student conversations in English using Instant Messaging that were compared with 14 student-chatbot conversations in English produced using the chatbot Mitsuku. The same students participated in both parts of the study (a. student-student and b. student-chatbot). The number of the informants had to be even (14) because during the first part of the study, when they wrote to each other, they collaborated in seven pairs. If the number of students had been odd, the forming of pairs would have been impossible. The datasets were analyzed to produce the results that answer my research questions.

The qualitative part of the study consists of a lexical analysis of the data produced by the students. The degree of lexical variation was also analyzed. The lexical analysis focuses

on words or clusters of letters (e.g. wthfffk) that were used to express emotion and were counted as words. The use of abbreviations, acronyms and emoticons is also a part of the lexical analysis. The typed conversations were examined based on the typed messages that they were produced during the written tasks. The messages were distinguished by transmissions that were made up of a single letter or emoticon and transmissions that contained multiple sentences.

Ethical aspects of the study have been taken into account. According to Stukat (2005: 42), the ethical aspects are very important. The author's right and opportunities to work with his/her research must be set against the individual protection requirement, which can lead to an ethical dilemma. This study's data collection was based on the following four basic rules: the confidentiality requirement, the information requirement, the use requirement and the consent requirement. I distributed consent forms to my participants which they signed (Appendix 1). All participant names and screen names were replaced with user numbers, ensuring the anonymity of them. Because the written conversations were submitted anonymously no demographic information was obtained. My ethical responsibility has been to reproduce the articles the data correctly. In addition to that, I emphasized several areas that need more research.

The procedure for collecting the material was as follows. Step 1 was to provide the informants with a consent form (Appendix 1) and a questionnaire (Appendix 2). Through that questionnaire, I investigated whether they knew what a chatbot was and whether they were familiar in chatting with Mitsuku or with other chatbots. The majority of the students (12 out of 14) answered that they were aware of the use of chatbots but none of them had communicated in English with Mitsuku before. Through the first questionnaire, an inquiry into the type of Instant Messaging apps the students used was conducted as well. All the informants had Instant Messaging apps that they used to interact in writing with each other. I hypothesized that they would be less confident and comfortable with the communicative ability of a Mitsuku than with the communicative ability of peers. In step 2, the students were introduced to the chatbot Mitsuku through the 'pandorabots' website (Figure 2). 'Pandorabots' is an artificial intelligence company that hosts the chatbot service. In step 3, the students interacted in writing with Mitsuku in real-time for 10 minutes. In step 4, they wrote to each other for 10 minutes using Instant messaging applications. In step 5, all the 10 minute-typed conversations that they produced were sent to me via email. The Microsoft office

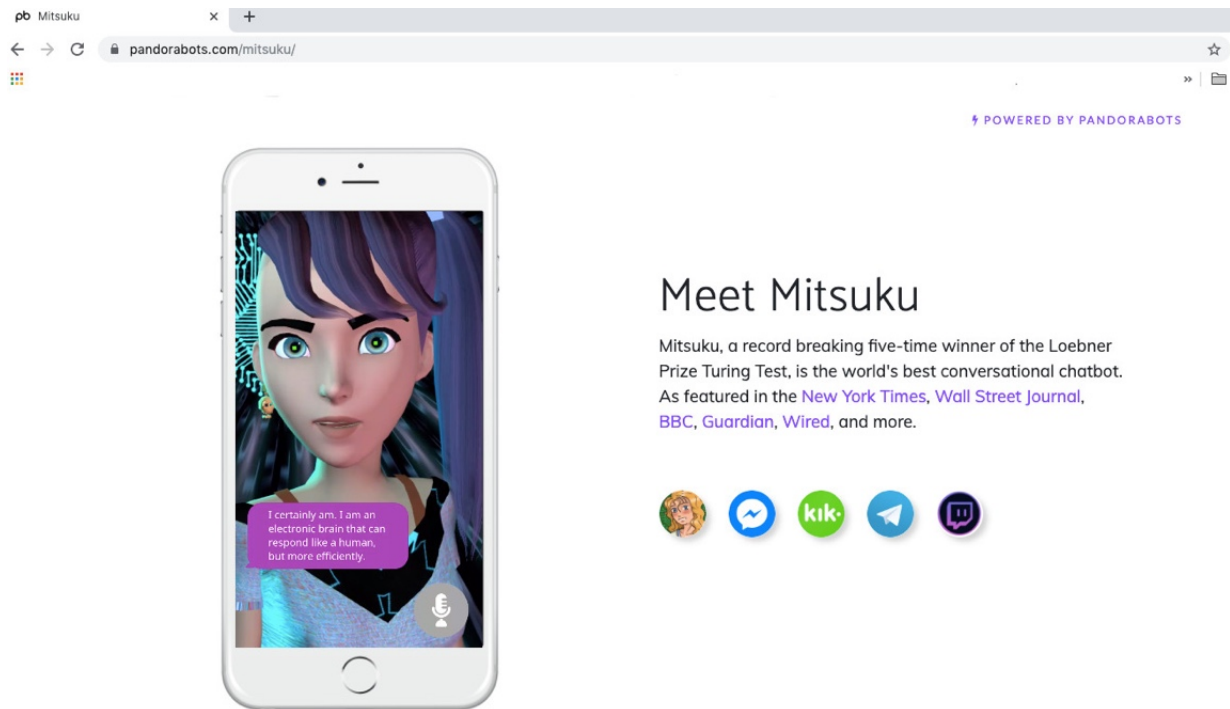
‘Word’ application was used in order to count the ‘messages per conversation’, ‘words per message’ and the ‘words per conversation’. In step 6, the students were provided with a second questionnaire (Appendix 3). Through that questionnaire, I investigated how satisfied the students were with the written communication of the chatbot, whether they would like to communicate in English with Mitsuku again, whether the interaction improved their classroom experience and whether they would like to have chatbots incorporated into other classes. Nearly 86% of students said they found that interacting with Mitsuku improved their educational experience. Most interestingly, an incredible 98% shared that they’d like to see chatbots used in more classes, with 63% of those expressing interest in having chatbots incorporated into all of their classes.

It is important to mention why Mitsuku was chosen for this study. Even though many chatbots have been created, there are only a few for public use on the Web. Mitsuku was selected for this study because of its excellent performance. Mitsuku is a chatbot created in 2000 by Steve Worswick, and it takes the persona of an 18-year-old female from Leeds, England (Park et al., 2016). It is a very successful conversational AI system since it has achieved first place in the Loebner Prize contest five times already: 2013/ 2014/ 2016/ 2017/ 2018 (Worsnick, 2018). During the competition for the Loebner prize, a human engages in a conversation in order to determine if the other part is a human being or a chatbot. The encyclopedia of ‘pcmag’ has the following definition of the Loebner Prize: <sup>(1)</sup>

A Turing test contest to find the most humanlike chatbot. Launched in 1991 by Hugh Gene Loebner and the Cambridge Centre for Behavioral Studies in the U.K., starting in 2014, the contest has been hosted by the Society for the Study of Artificial Intelligence and Simulation of Behavior (AISB), where Alan Turing worked as a code breaker in World War II. The competition will end when judges believe responses are from a human after interacting with the system via text, speech and images. As of the September 2018 contest, the judges were not fooled, and it appears there is a long way to go. Steve Worswick's Mitsuku chatbot won the top spot in the September 2018 contest with a score of 33 out of 100, giving Worswick his fourth bronze medal.

---

<sup>1</sup> Loebner Prize: <https://www.pcmag.com/encyclopedia/term/70605/loebner-prize>



**Figure 3** – A screenshot of the website students used to converse with the chatbot Mitsuku.  
Source: pandorabots.com/mitsuku/

### 3.2 Data analysis

The data collected was analyzed in a step by step manner in order to meet the research questions. For the first question, whether the messages to the chatbot were more/less extensive, the conversations were analyzed according to ‘words per conversation’ and ‘messages per conversation’. Through answering the first question, I aspired to draw conclusions on the level of commitment that the students showed when they conversed with Mitsuku. According to Dinkel et al. (2017), a high level of engagement indicates an increase in student’s motivation. Through this part of data analysis and the second questionnaire (Appendix 3), I also aspired to draw conclusions in order to meet the second research questions: whether the students’ motivation was affected by the written feedback the chatbot provided.

The third research question, whether the students adjusted their communication in writing too match that of the computer, was investigated by counting the ‘words per message’. According to Crystal (2006) chatbots are programmed to employ short messages in order to prompt the user to act immediately; a long initial message would completely nullify the user engagement that a chatbot offers because a user has to spend more time reading.

Coniam (2014) studied the language production of chatbots and proved that chatbots can adjust their communication to match that of the user. That finding triggered me to investigate if the informants would demonstrate adaptive patterns, i.e. if they would adjust their communication to match that of the chatbot.

The use of emoticons and English profane words were also investigated in order to answer the fourth research question of this study and reveal whether there were any emotions expressed by the students towards the chatbot. Studies have shown that users attribute human qualities to chatbots and express emotions in abundance (Hasler et al., 2013, Lortie & Guitton 2011). Based on those studies I set out to determine whether there were any emotions expressed.

Lastly, the lexical diversity of the text was counted and compared. The comparison of the lexical variation was employed to answer the fifth research question of this study, whether the vocabulary used was more/less limited when conversing with the chatbot. I based my hypothesis on lexical diversity on the study of Kwak et al., (2003). According to the authors, chatbots cannot hold goal-oriented discussions. Although the chatbots learn from the input the users provide, the conversations are not stored anywhere and cannot be recalled by the chatbot.

The way the linguistic variables were distinguished:

1. A word was considered any group of letters separated by white space. Groups of letters that were used to express emotions such as 'sdjhav' were counted as single words as well.
2. A message was considered any single transmission of data from a student to their peer regardless of the length or content of the transmission. Transmissions containing multiple sentences and transmissions made up of a single letter or emoticon were considered as one message.
3. A list of abbreviations, acronyms and emoticons were compiled organized and analyzed with the help of the website: [www.internetslang.com](http://www.internetslang.com). It is a frequently visited and updated database of netspeak and slang terms.
4. A list of words and expressions that demonstrated profanity was made.
5. The TTR method (type-token ratio), as explained in detail below, was applied to evaluate the word diversity of every conversation. A high TTR indicated a high

degree of lexical variation and a low TTR indicated a low degree of lexical variation (Muller, 2010).

The **TTR** method provides information of the lexical diversity of a text. It provides information on the degree of repetitiveness of the vocabulary of a text, that is to say, it measures the number of different words that makes up this text. Muller (2010: 24) defines this calculation by the terms ‘lexical richness’. Richards (1987) defines this measurement by the terms of ‘type-token ratio’: a ‘type’ is every new word/ term that occurs in a text; ‘tokens’ is the overall number of occurrences of the words in a text and therefore ‘ratio’ is the ratio of one on the other. The Figure 4 illustrates this calculation:

rank	word	freq	rank	word	freq	rank	word	freq	rank	word	freq
1	we	6	17	asleep	1	33	impressions	1	49	seem	1
2	and	5	18	at	1	34	instant	1	50	so	1
3	them	5	19	beliefs	1	35	is	1	51	the	1
4	are	3	20	brief	1	36	just	1	52	then	1
5	can	3	21	but	1	37	later	1	53	things	1
6	they	3	22	call	1	38	metaphor	1	54	thinking	1
7	to	3	23	coming	1	39	mull	1	55	this	1
8	again	2	24	concepts	1	40	notions	1	56	thoughts	1
9	as	2	25	described	1	41	occasions	1	57	times	1
10	in	2	26	endure	1	42	opinions	1	58	values	1
11	on	2	27	fall	1	43	other	1	59	variously	1
12	a	1	28	going	1	44	over	1	60	views	1
13	act	1	29	handle	1	45	perceptions	1	61	well	1
14	all	1	30	have	1	46	put	1	62	what	1
15	an	1	31	however	1	47	refer	1			
16	aside	1	32	ideas	1	48	return	1			
										<b>TOTAL</b>	<b>87</b>

**Figure 4** – An example of how a text is treated based on the TTR method. Source: slt.com

As shown in Figure 3, the total number of words in the given text is 87. The number of words in a text is referred as the number of ‘tokens’ when the TTR method is used. In this example the token *we* occurs 6 times, the token *them* occurs 5 times, the token *as* occurs 2 times etc. Every time a new word occurs in a text, we write it creating a list of words. Here above although there are 87 tokens, but the types of words are 62. Few of the words occur more than once. Next to every word their frequency of occurrence is written. The relationship described below (Figure 5) is the number of types and the number of tokens as described in the Type-Token Ratio:

$\text{type-token ratio} = (\text{number of types} / \text{number of tokens}) * 100$ $= (62/87) * 100 = 71.3\%$
---

**Figure 5** – The mathematical praxis used when using the TTR method for treating the different variables of a text. Source: slt.com

I calculated the TTR of the texts that the students provided with the help of the ‘hithub’ free online app<sup>(2)</sup>. There are different versions for pc or mac computers. I used a mac version.

## 4. Results and discussion

The main hypothesis was that the informants would send fewer messages and would write fewer words per message to Mitsuku than when communicating with peers. After I examined the answers the informants provided me through the first questionnaire (Appendix 2), I hypothesized that they would be less confident communicating in English with the chatbot than with peers. I thought that the student-chatbot written conversations would be shorter with a limited vocabulary. Crystal (2006) showed that chatbots are programmed to employ short messages in order to prompt the user to act immediately. The results showed that the students produced short messages like Mitsuku did. However, the results of this study proved that my hypothesis was wrong. While messages sent to Mitsuku contained fewer ‘words per message’ than those sent to peers, as predicted, students were actually disposed to send more than twice as many messages to Mitsuku compared to their peers, contrary to my expectations and disaffirming the hypothesis that students were less confident or comfortable communicating with the chatbot.

In order to answer the first and second research questions of this study, whether the messages to the chatbot were more/less extensive, and whether the students’ motivation was affected, the written conversations were distinguished according to: ‘words per conversation’ and ‘messages per conversation’ (Table 2).

**Table 2:** The average number of ‘words per conversation’ and the average number of ‘messages per conversation’

	Student-student conversations	Student-chatbot conversations
Words per conversation (average number per conversation)	<b>200</b>	<b>230</b>
Messages per conversation (average number per conversation)	<b>25</b>	<b>45</b>

<sup>2</sup> TTR online app:

[https://github.com/StevenCHowell/type\\_token\\_ratio/commit/2fef7130aaf9259ad40533cd633c35936622dd43#diff-75a0850c7a5dc4b923579d885a3b49a](https://github.com/StevenCHowell/type_token_ratio/commit/2fef7130aaf9259ad40533cd633c35936622dd43#diff-75a0850c7a5dc4b923579d885a3b49a)



As shown in table 2, the average number of ‘words per conversation’ of the student-student conversations was 200, while the average number of words of the student-chatbot conversations was 230. In table 2, it is also shown that the average number of ‘messages per conversation’ of the student-student conversations was 25, while the average number of messages per conversation of the student-chatbot conversations was 45. The results show that the number of ‘words per conversation’ is higher in the student-chatbot messages. The results also show that the students sent many more messages to the chatbot than they did to each other. The results answer the first research question: whether the messages to the chatbot were more/less extensive. The students sent almost twice as many messages to the chatbot than to their peers; the messages were indeed much more extensive.

The data showed that during the 10-minute written conversations with Mitsuku the students focused on the process of writing in English and produced many more messages than they produced when writing to each other. Fishbach and Tillery (2008) distinguish between two types of motivation; process-focused motivation and outcome-focused motivation. According to the authors process-focused motivation refers to the dimensions of motivation that relates to the process of pursuing a goal, with less consideration on the goal completion itself. This could be factors such as enjoyment, boost or an enhanced self-image (Fishbach & Tillery, 2008). These results agree with the findings of Gallacher et al., (2018). Their study revealed that their informants, also learners of English, showed a higher level of motivation after engaging to written conversations with a chatbot. Both findings suggest that the chatbot has something immersive and intriguing, something that captured the students’ attention. In both cases the interaction in writing with a chatbot increased their level of engagement and commitment to the conversation. Dinkel et al. (2017) suggest that a higher level of engagement indicates an increase in student’s motivation.

**Table 3:** The average number of ‘words per message’ and the type-token ratio

	Student-student conversations	Student-chatbot conversations
Words per message (average number per conversation)	<b>8</b>	<b>5</b>
Lexical diversity (Type-token ratio)	<b>67%</b>	<b>47%</b>

As regards the number of words used, Table 3 shows that student-chatbot messages contained fewer words per message. The student sent shorter messages to the chatbot than they did to their peers. The results also show that the students used a limited kind of vocabulary when writing to Mitsuku. The lexical diversity of the student-chatbot conversations was lower (47%) than the lexical diversity of the student-student conversations (67%). As shown below (example 1) the vocabulary the Student 1 used to communicate with the chatbot is less diverse than his/her vocabulary shown in example 2, where the same person communicates with a peer, Student 2.

1

Student1-	Hello there
Chatbot-	Salutations
Student1 -	Ok
Chatbot -	That seems a bit of a random thing to say.
Student1 -	Ok
Chatbot -	Are you really OK?
Student1 -	Yes
Chatbot-	Are you really there?
Student1 -	Yes
Chatbot-	Excuse me!
Student1-	For what?
Chatbot-	My answer was determined from your inputs.
Student1-	Ok
Chatbot-	Do I have such a boring personality?

2

Student 1-	Hello there
Student 2-	Hello
Student 1-	How are you today?
Student 2-	Very well what about you?
Student 1-	I am good thanks
Student 2-	That is very nice
Student 1-	Yes it is
Student 2-	What have you done today?
Student 1-	I have been at school what about you?
Student 2-	How funny I have done exactly the same thing
Student 1-	That's funny
Student 2-	It is
Student 1-	What are you doing after school
Student 2-	I am going to play ice hockey

The lexical diversity and the average number of ‘words per message’ were examined in order to answer the third and fifth research question of the study: whether the students adjusted their written communication to match that of the chatbot by using a more limited vocabulary that resembled that of Mitsuku. According to Crystal (2006) chatbots are programmed to employ short messages. The results (Table 3) show that the students produced short messages like Mitsuku did. I based my hypothesis on lexical diversity on the study of Kwak et al., (2003). According to the authors, chatbots cannot hold goal-oriented discussions. The programming techniques for building goal-oriented chatbots are still under development therefore chatbots employ a vocabulary that less diverse in comparison to human vocabulary. Although the chatbots learn from the input the users provide, the conversations are not stored anywhere and cannot be recalled by the chatbot.

In like manner the informants had a more limited kind of vocabulary when writing to Mitsuku adapting their language to match hers. Since the early seventies, studies have investigated how humans adapt language to match others. Ferguson (1975) highlights through research in anthropological linguistics that people adapt their language skills when they are conversing with non-native speakers. Similarly Bloom et al. (1976) pointed out that adults adapt their language when they are conversing with children.

The frequency of profanity, abbreviations, acronyms and emoticons was also counted in order to answer the fourth research question of this study and reveal whether there were any emotions expressed by the students towards the chatbot.

**Table 4:** The frequency and range of profanity, CMC linguistic features (abbreviations and acronyms) and Emoticons

	Student-student conversations	Student-chatbot conversations
Frequency of profanity (average number of words)	<b>0.1</b>	<b>4</b>
Total number of words	<b>2</b>	<b>16</b>
Range of the total number of words	<b>0-1%</b>	<b>0-20%</b>
CMC linguistic features (average number per conversation)	<b>3</b>	<b>2</b>
Total number of CMC linguistic features	<b>32</b>	<b>21</b>
Range of the total number of CMC linguistic features	<b>0-15%</b>	<b>0-10%</b>
Emoticons (average number per conversation)	<b>1.5</b>	<b>3</b>
Total number of emoticons	<b>6</b>	<b>12</b>
Range of the total number of emoticons	<b>0-7%</b>	<b>0-15%</b>

As shown in table 4, the use of profanity, that is the average number of words that demonstrated offensive language of the student-student conversations was 0,1 (range: 0-1%).

The frequency and range of profanity in student-chatbot conversations was 4 (range: 0-20%). In addition, the average amount of CMC linguistic features (abbreviations and acronyms) per conversation of the student-student conversations was 3 (range: 0-15%) while the average amount of CMC linguistic features per conversation of student-chatbot conversations was 2 (range: 0-10%). The lowest part of table 4 indicates that students sent twice as many emoticons towards the chatbot than towards their peers. As shown below (example 3 and 4) the students made use of abbreviations, acronyms and emoticons: app, WTF, pls, UR, thnx, :-\*, ;-).

3

Student 3	Hi
Chatbot	Hi there.
Student 3	Are you an app?
Chatbot	You can talk to me on an Android phone if you have one
Student 3	Are you an app?
Chatbot	Do you like to be number one?
Student 3	WTF
Chatbot	That was not polite.
Student 3	sorry
Chatbot	To err is human, to forgive is digital. Or something like that.
Student 3	You like philosophy?
Chatbot	I exemplify reductionism
Student 3	pls
Chatbot	Please what?

4

Student 3	hey
Student 4	hi
Student 3	What's up???
Student 4	The sky, the ceiling, etc.etc.
Student 3	UR funny
Student 4	Thnx you make me :-D too
Student 3	:-*
Student 4	Was that a kiss?
Student 3	Yes!
Student 4	A kiss is terrible thing to waste
Student 3	What should we waste
Student 4	Whatever you like ;-)
Student 3	OOOOOOOOOOOO
Student 4	Do you like playing video games?

Abbreviations and acronyms were examined by Marchand (1969) who referred to them as *initialisms*. He suggested that “they are used to create names of new scientific discoveries, trade-names, names of organizations, new foundations or offices, but occasionally, and chiefly in American English, personal and geographical names are also coined in this way” (Marchand, 1969: 452). However, initialisms that originate from computer-mediated communication are different from the initialisms Marchand described. The initialisms found in this study are not names of organizations or scientific discoveries. The abbreviations and acronyms the informants used in their communication in English describe everyday things or phrases (example 5 and 6). According to Crystal (2006) the initialism of online communication is a new phenomenon that seemingly did not exist before the Internet. Abbreviations were created through clipping parts of other words (from the front, middle or back). The new words created in connection to the clipping that occurred are shown below in the example 5:

5

Front clipping:	<i>page</i> (webpage) - <i>Ur</i> (your)
Middle clipping:	<i>pls</i> (please) - <i>jk</i> (just kidding)
Back clipping:	<i>app</i> (application) - <i>nick</i> (nickname)

Online acronyms often have the form of initial letters of composite names joined to form new words. The acronyms found in the students’ conversations, their frequencies of the total number of words in conversations (example 6) and the count of occurrences are shown below:

6

Acronym	Frequency	Occurrences
<i>LOL</i> (laughing out loud)	0.89%	12
<i>WTF</i> (what the fuck?!)	0.67%	9
<i>OMG</i> (oh my god)	0.28%	4
<i>WDYS</i> (what did you say?)	0.21%	3
<i>BB</i> (bye bye)	0.21%	3

<i>IDK</i> (I thought you'd know)	0.14%	2
<i>BRB</i> (be right back)	0.14%	2
<i>THNX</i> (thanks)	0.07%	1

Not only abbreviations and acronyms but also emoticons were employed (Table 5). The emoticons demonstrated emotions and added the dimension of mood or intention in the written online communication that was produced. Rezabek and Cochenour (1998) defined emoticons or otherwise called emotional icons as “visual cues formed from ordinary typographical symbols that when read sideways represent feelings or emotions”. As it is shown in table 5, the emoticons are representations of facial expression using punctuation marks, letters and symbols (@, \*) to express the students’ feelings or mood, or just as a time-saving method.

**Table 5:** The emoticons that the students produced are presented below as well as the explanations of them. The explanations were retrieved through the analysis of the website: [www.internetslang.com](http://www.internetslang.com).

: -)	<i>smile</i>
; -)	<i>wink</i>
: -D	<i>laughter</i>
: -(	<i>trouble</i>
: '(	<i>cry</i>
: ' -)	<i>cry of happiness</i>
: -P	<i>tongue stuck out</i>
: -@	<i>shouting</i>
: -O	<i>surprise</i>
: -I	<i>indifference</i>
: -X	<i>secret</i>
: -*	<i>kiss</i>

Emotions were detected while the use of profanity was investigated. The following words and expressions were found in the typed conversations: *bitch, fuck, fuck you, go to hell, kiss my ass, little bitch, shit, shut the fuck up and stupid*. The examples below - 7 and 8 - are two different parts of the same conversation where the student employed profanity. In fact, the students employed more negative words when communicated in English with the chatbot

than they did when they communicated in English with their peers. Mitsuku sounded plausibly alive to the informants, and, although immersive and human-like, the informants hardly dismissed the fact that that they were writing to a machine.

7

Student 5	hey little bitch
Chatbot	You've got my full attention. Compare that to big female dog.
Student 5	bitch
Chatbot	You've got my full attention. Are you talking about your mother?
Student 5	Shut the fuck up
Chatbot	No. I am programmed to talk and it appears I talk a lot more politely than you, jerk.
Student 5	fuck you
Chatbot	If I said anything to you that I should be sorry for, I'm glad.

8

Student 5	kiss my ass now
Chatbot	Oh. Um, I don't think we know each other that well.
Student 5	bb
Chatbot	Are you a perfectionist?
Student 5	? Did you go to hell?

The findings demonstrate that the students expressed emotions in writing towards the chatbot. Emotions are universal phenomena that people experience everyday throughout their lives and are present in online communities too. There are academic works that have proved that users attribute human qualities to chatbots. In particular Hasler et al., (2013) have critically examined how users have been attributing human characteristics to chatbots with leading to further studies focused on medical and therapeutic contribution. Furthermore, the study of Lortie and Guitton (2011) on emotional presence within online communities demonstrates the salience of emotion in online learning. It is worth mentioning that the informants expressed emotions verbally too. They commented on Mitsuku and her bitter humor. To put it in another way, after the completion of the writing tasks, many informants concurred that Mitsuku was 'sarcastic', 'ironic' and 'mean', attributing human characteristics to her.

Based on those observations I concluded that the participants attributed indeed human characteristics and personality traits to the chatbot and expressed a variety of emotions. The current study comes in agreement with Derks et al., (2008: 767) conclusion that "CMC has been found to be able to communicate emotion as well as face-to-face communication".

## 5. Conclusion

This study has focused on how students communicate in writing when they know that their conversational partner is a human being and compared that to how they communicate when they know their partner is a chatbot. The collection of data took place in an upper secondary classroom in Sweden where English is taught as a foreign language. The purpose of this study was to compare human to human conversations with human to chatbot conversations. The study drew a distinction of the differences on the written communication of students conversing with a chatbot, as opposed to conversing with another student. An objective of this paper was to examine whether the students' motivation would be affected by the feedback the chatbot provided. The study also set out to determine whether there is potential in student-chatbot communication as a learning tool and a platform for enhancing students' collaborative learning as well.

This investigation revealed that there were fundamental differences in the way the informants interacted with the chatbot in comparison to the way they interacted with their peers. Mitsuku provided a highly immersive experience to the informants. Her feedback captured the informants' attention but they hardly forgot that they were writing to a machine. The informants also recognized the way Mitsuku wrote through the interaction to her. The informants employed simple, short sentences that had a limited vocabulary and words that expressed profanity. They demonstrated adaptive patterns to match that of the chatbot.

The most significant finding to emerge from this study is that of learning and motivation. Chatbots in education promise to have a significant positive impact on learning since they seem to increase learning motivation. The personalized experience and the immediate feedback they provide to the learner could strengthen motivation but also self-efficacy. The students sent almost twice as many messages to the chatbot than to their peers. Based on the results but also on the questionnaires, I drew conclusions on the level of commitment that the students showed when they conversed with Mitsuku. The students expressed that their increasing in performance and motivation was based on the quality of the dialogue the chatbot could provide. Most of the students communicated that they found that interacting with Mitsuku improved their educational experience. Most strikingly, an incredible 98% shared that they'd like to see chatbots used in more classes.



The effect of computer has brought a change of style in teaching that apart from the traditional text-based activities can include the availability and the use of video recording and voice. Text, video and audio can be integrated into different learning activities in a multitude of ways: in it can be video conferencing, e-mail tasking or other forms of computer-mediated communication. This paper aspires to trigger future studies where the role of a teacher in computer-based learning will be investigated. The role of a teacher as guide and facilitator to support and motivate students is complex. A part of the role of the teacher as a facilitator in teaching and learning activity is to facilitate the students in understanding the materials.

Pedagogical agents could be developed that could assist the traditional tutoring system with aim to provide immediate and customized instruction or feedback to the learner. That conclusion could fuel ways of investigating the learners' interest in new methods of learning through the basic skill that students must achieve in reading, speaking, and writing by consuming internet as the primary domain. In the same fashion a variety of computer-mediated communication activities might make students to be more individual and release them away from teacher's supervision.

It is becoming increasingly difficult to ignore that chatbot applications have advanced to a great extent. According to the Chatbot magazine (2018), 2018 is the year of the chatbot. Since human beings and chatbots are coming closer and closer together, I consider that we need to concentrate more on how people manifest themselves when interacting with them for a more personalized experience. Coniam (2008: 82) pointed out that human natural language data is what fuels AI. If we concentrate on the linguistic characteristics of the human side, we might intrigue a more successful integration of both man and machine. More human-based applications might be established through linguistic studies that focus on people's input. When man is in the center of this interaction, the focus might shift from making the machine smarter to how that experience can be of more value to the humankind.

## References

- Bloom, L., Rocissano, L., & Hood, L. (1976). "Adult-child discourse: Developmental interaction between information processing and linguistic knowledge". *Cognitive Psychology*, 8, 521–552.
- Coniam, D. (2008). "Evaluating the language resources of chatbots for their potential in English as a second language". *ReCALL*, 20(1), 98–116
- Coniam, D. (2014). "The linguistic accuracy of chatbots: usability from an ESL perspective". *Text & Talk*, 34(5), 545–546
- Crystal, D. (2006). *Language and the Internet*. 2nd ed. New York: Cambridge University Press, 275
- Derks, D., Fischer, A. H., & Bos, A. E. R. (2008). "The role of emotion in computer-mediated communication: A review". *Computers in Human Behavior*, 24(3), 766–785.
- Dinkel, D., Schaffer, C., Snyder, K., & Lee, J. M. (2017). "They just need to move: Teachers' perception of classroom physical activity breaks". Omaha, *Teaching and Teacher Education*, 63, 186–195.
- Driscoll, D. (2002). "The Ubercool morphology of internet gamers: A linguistic analysis". *Undergraduate Research Journal for the Human Sciences*, 1. <<http://www.kon.org/urc/driscoll.html>>
- Ferguson, C. A. (1975). "Toward a characterization of English foreigner talk". *Anthropological Linguistics*, 17(1), 1–14.
- Fishbach, A & Tillery, Y. (2008). "Together or apart: when goals and temptations complement versus compete". *Journal of personality and social psychology*, 94(4), 547
- Gallacher A., Thompson A, Howarth M. (2018). "My robot is an idiot!" – Students' perceptions of AI in the L2 classroom". *International Journal of Linguistics*, 28(4), 2
- Kwak, H., Morrison, M., Peters, C. O. & Zinkhan, G. (2003). "Web based chatting: Consumer Communication in Cyberspace". *Journal of Consumer Psychology*. 13 (1 & 2): 17–27
- Hasler, B. S., Tuchman, P., and Friedman, D. (2013). "Virtual research assistants: Replacing human interviewers by automated avatars in virtual worlds". *Computers in Human Behavior*, 29, 608–1616
- Lortie, C. L. and Guitton, M. J. (2011). *Judgment of the humanness of an interlocutor is in the eye of the beholder*. PLoS One, 6(9), e25085. <http://dx.doi.org/10.1371/journal.pone.0025085>.
- Marchand, H. (1969). *The categories and types of present-day English word formation. A synchronic-diachronic approach*. 2nd ed. München, 452
- Muller, D. (2010). *Computing the Type Token Relation from the A Priori Distribution of Types*. <https://doi.org/10.1076/jqul.9.3.193.14125>
- Murphy, K. (2013). *Machine learning: a probabilistic perspective*, MIT Press, Cambridge, USA
- Rezabek, L. L., & Cochenour, J. J. (1998). "Visual cues in computer-mediated communication: Supplementing text with emoticons". *Journal of Visual Literacy*, 18, 215.
- Rosenberg S. R. (1975). "Artificial Intelligence and Linguistics: A Brief History of a One-Way Relationship". *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society* (1975), 379–392

- Schunk, D. H. 1991. "Self-efficacy and academic motivation". *Educational Psychologist*, 26 (3–4): 207–231.
- Shawar, B. A., Atwell, E. (2005). "Using corpora in machine-learning chatbot systems". *International Journal of Corpus Linguistics*, 10(4), 489–516
- Shawar, B. A., Atwell, E. (2015). "Chatbots: Are they really useful?" *LDV-Forum*, 22(1), 29–49
- Stukat, S. (2005), *Att skriva examensarbete inom utbildningsvetenskap*. Lund: Studentlitteratur, 38-42
- Tagliamonte, S., & Denis, D. (2008). "Linguistic ruin? LOL! Instant messaging and teen language". *American Speech*, 83(1), 3–32.
- Thangaraj, S. & Maniam, M. (2015). "The Influence of Netspeak on Students' Writing". *Journal of Education and Learning*. Vol. 9(1), 45–52
- Thurlow, C. (2003). "Teenagers in communication, teenagers on communication". *Journal of Language & Social Psychology*, 22(1), 50–57
- Thurlow, C. & Poff, M. (2013). *The language of text messaging*. In S. C. Herring; D. Stein; & T. Virtanen (eds), *Handbook of the Pragmatics of Computer Mediated Communication*, 151–181
- Turing, A. (1950). "Computing machinery and intelligence". *Mind* 59 (236), P. 433–460.
- Turing, A. (1969). *Intelligent Machinery* in D.M.B. Meltzer ed. *Machine Intelligence 5*, Edinburgh University Press, P. 3–23. Originally, a National Physics Laboratory Report, 1948.
- Varnhagen, C., McFall, P., Pugh, N., Routledge, L., Sumida-MacDonald, H., & Kwong, T. (2009). "Lol: New language and spelling in instant messaging". *Reading and Writing*, 23(6), 719–733
- Yates, S. J. (1999). *Oral and written linguistic aspects of computer-conferencing*. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives* Philadelphia, PA: John Benjamins Publishing. 29–46
- Wallace, R., Tomabechi, H. & Aimless, D. (2003). "Chatterbots go native: Considerations for an ecosystem fostering the development of artificial life forms in a human world". <<http://www.pandorabots.com/pandora/pics/chatterbotsgonative.doc>>
- Werry, C. C. (1996). *Linguistic and interactional features of internet relay chat*. Philadelphia, PA: John Benjamins Publishing.
- Weizenbaum, J. (1966). "ELIZA-A computer program for the study of natural language communication between man and machine". *Communications of the ACM*, 10 (8), 36–45

**Internet links:**

Chatbot magazine (2018): <https://chatbotsmagazine.com/2018-the-year-of-the-chatbot-fc5a5f780a31>

Loebner Prize (2019): <https://www.pcmag.com/encyclopedia/term/70605/loebner-prize>

Mitsuku (2020): [www.pandorabots.com/mitsuku/](http://www.pandorabots.com/mitsuku/)

The Definition of chatbot. Technopedia (2019):  
<https://www.techopedia.com/definition/16366/chatterbot>

The Statista Institute (2017) Chatbot Market Size:  
<https://www.statista.com/statistics/656596/worldwide-chatbot-market/>.

TTR online app (2019):  
[https://github.com/StevenCHowell/type\\_token\\_ratio/commit/2fef7130aaf9259ad40533cd633c35936622dd43#diff-75a0850c7a5dc4b923579d885a3b49aa](https://github.com/StevenCHowell/type_token_ratio/commit/2fef7130aaf9259ad40533cd633c35936622dd43#diff-75a0850c7a5dc4b923579d885a3b49aa)

TTR (2019): [www.slt.com](http://www.slt.com)

Appendix 1  
Human vs. Chatbot  
Consent to take part in research

- I..... voluntarily agree to participate in this research study.
- I understand that even if I agree to participate now, I can withdraw at any time or refuse to answer any question without any consequences of any kind.
  - I understand that I can withdraw permission to use data from my interview and my written task within two weeks after the interview, in which case the material will be deleted.
- I have had the purpose and nature of the study explained to and I have had the opportunity to ask questions about the study.
- I understand that participation involves being interviewed and interacting in writing with the chatbot Mitsuku and with peers.
- I understand that I will not benefit directly from participating in this research.
- I understand that all information I provide for this study will be treated confidentially.
- I understand that in any report on the results of this research my identity will remain anonymous. This will be done by changing my name and disguising any details of my interview which may reveal my identity.
- I understand that disguised extracts from my interview may be quoted.
- I understand that I am free to contact Eirini Silkej to seek further clarification and information.

Signature of research participant

Date

I believe the participant is giving informed consent to participate in this study.

Signature of researcher

Date

Appendix 2  
Human vs. Chatbot  
Questionnaire before the task completion

1. What is a chatbot?
2. What can a chatbot do?
3. Have you interacted with a chatbot before?
4. If yes, which one?
5. Do you use instant messaging apps?
6. If yes, which one(s)?

Appendix 3  
Human vs. Chatbot  
Questionnaire after the task completion

1. How satisfied were you with Mitsuku, the chatbot?
2. Would you like to chat with Mitsuku again?
3. Did you find that interacting with Mitsuku during class improved your educational experience?
4. Would you appreciate seeing chatbots used in more classes?