“My robot is an idiot!” – Students’ perceptions of AI in the L2 classroom

Andrew Gallacher[1], Andrew Thompson[2], and Mark Howarth[3]

Abstract. Japanese university students (N=253) conversed with human and Artificially intelligent (AI) chatbot partners then recorded their perceptions of these interactions via open-ended written feedback. This data was qualitatively analyzed to gain a better understanding of the merits and demerits of using chatbots for English study from the students’ perspective. Results suggest that, in its current state, students perceive the chatbot used in this study as a novelty rather than a legitimate language-learning tool and that it lacks the richness of interaction they could achieve with their peers. Ultimately, it is argued that educators should be more critical of incorporating AI technology in the second/foreign language (L2) classroom before it is ready for use.

Keywords: CALL, artificial intelligence, AI, chatbot, student perception.

1. Introduction

AI ‘chatbots’ provide an opportunity for independent conversational practice (Atwell, 1999) without the need for continual one-on-one support (Parker, 2007), making them an area of great interest within Computer-Assisted Language Learning (CALL) based research (Goda et al., 2014). Studies into the efficacy of chatbots as language learning tools have demonstrated a need for improvement before warranting greater adoption (Coniam, 2014; Fryer & Carpenter, 2006; Goda et al., 2014). However, the exponential rate of development of AI justifies routine and frequent ‘checking in’ with current iterations of these products, as ongoing

1. Kyushu Sangyo University, Fukuoka, Japan; gallacher@ip.kyusan-u.ac.jp
2. Kyushu Sangyo University, Fukuoka, Japan; thompson@ip.kyusan-u.ac.jp
3. Sojo University, Kumamoto, Japan; markwhowarth@gmail.com

How to cite this article: Gallacher, A., Thompson, A., & Howarth, M. (2018). “My robot is an idiot!” – Students’ perceptions of AI in the L2 classroom. In P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny (Eds), *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 70-76). Research-publishing.net. https://doi.org/10.14705/rpnet.2018.26.815

advancements make their feasibility of use for language practice simply a matter of time (Atwell, 1999; Fryer & Carpenter, 2006; Goda et al., 2014; Gulz, 2004; Morton & Jack, 2007; Parker, 2007).

With this in mind, the current study was designed to investigate the impact of using an artificially intelligent chatbot, known as *Cleverbot*, in a Japanese L2 classroom. As a part of this research, students' perceptions of AI versus human conversational partners were examined. The intervention utilized a task-based format with students performing spoken interviews with both AI and human partners in alternating sessions. Qualitative results obtained after each session were coded, and are presently discussed in relation to ongoing trends in CALL.

2. Methods

2.1. Participants

First and second year Japanese university students (N=253), studying English as a compulsory subject, took part in this study. Placement test results indicate a range of fluencies roughly equivalent to A1 to B2 on the Common European Framework of Reference for languages (CEFR).

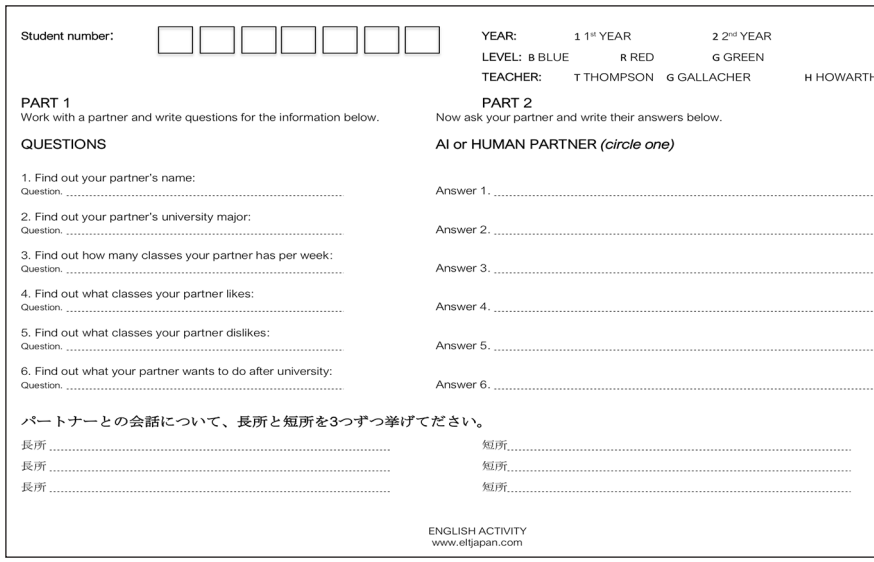
2.2. Technology

Cleverbot (www.cleverbot.com) was created by Rollo Carpenter in 1988 (originally called *Jabberwacky*, Wikipedia contributors, 2018). It was designed to simulate human chat using AI algorithms that match input phrases and themes to those found within a continuously growing database of previously held conversations. It has had more than 279 million conversations to date, and passed the Turing Test in 2011 (Wikipedia contributors, 2018).

2.3. Instrument

Participants created questions asking for specific information using the prompts on their worksheets (see Figure 1). They then verbally asked these questions to either a human partner or *Cleverbot*. Partner responses were recorded on the worksheets in written form, as were students' perceived merits and demerits of having completed the activity with either an AI or human partner. The latter open-ended feedback was later translated then coded for analysis.

Figure 1. Conversation task worksheet



Student number: ☐☐☐☐☐☐☐

YEAR: 1 1st YEAR 2 2nd YEAR
LEVEL: B BLUE R RED G GREEN
TEACHER: T THOMPSON G GALLACHER H HOWARTH

PART 1
Work with a partner and write questions for the information below.

QUESTIONS

1. Find out your partner's name:
Question. .....

2. Find out your partner's university major:
Question. .....

3. Find out how many classes your partner has per week:
Question. .....

4. Find out what classes your partner likes:
Question. .....

5. Find out what classes your partner dislikes:
Question. .....

6. Find out what your partner wants to do after university:
Question. .....

PART 2
Now ask your partner and write their answers below.

AI or HUMAN PARTNER *(circle one)*

Answer 1. .....

Answer 2. .....

Answer 3. .....

Answer 4. .....

Answer 5. .....

Answer 6. .....

パートナーとの会話について、長所と短所を3つずつ挙げてください。

長所 ..... 短所 .....
長所 ..... 短所 .....
長所 ..... 短所 .....

ENGLISH ACTIVITY
www.eltjapan.com

2.4. Procedure

Four intervention sessions were conducted (two sessions each with AI and human partners) with participants alternating between partner types for each session. Previous to the initial intervention, participants were shown how to access and orally interact with *Cleverbot* (using the voice recognition function on their smartphones). After every session, participants were given five minutes to complete their conversation worksheets. Each session lasted approximately 15 minutes.

2.5. Analysis

Completed worksheets were translated from Japanese into English where necessary. Students' reported merits and demerits of both AI and human partners were then coded in preparation for qualitative analysis. Codes for both partner types follow:

Human merits

- *Confirmable*: able to confirm understanding.
- *Contiguous*: can follow up conversation and ask for more detail.
- *Quick and reliable*: can complete quickly and rely on partner for response.

- *Natural*: facial expressions / emotions / voice easily understood.
- *Enjoyable/interesting*: interesting / fun / can make friends.
- *Discoverable*: can discover shared interests and learn about partner.
- *Forgiving*: more forgiving of mistakes (compared to voice capture on phone).
- *Appropriate*: stays on topic (appropriate responses to question asked).

Human demerits

- *Dependent on others*: cannot practice by yourself.
- *Limited responses*: receive one-word answers.
- *Misunderstanding*: hard to understand partner's pronunciation (ability).
- *Inability to check accuracy*: cannot check pronunciation / spelling / correctness (with phone's speech to text function).
- *Use of Japanese*: partner responds in Japanese.
- *Time consuming*: takes more time (because of follow up).
- *Lack of language ability*: lack of ability to use English (non-native partners).
- *Emotionally awkward*: feel shy or awkward / could lead to arguments.

AI merits

- *Independent*: can practice alone.
- *Quick response*: AI responds quickly and reliably to questions.
- *Ease of understanding*: replies in simple, easy English.
- *Smoothness of interaction*: one question, one answer no deviating.
- *Enjoyable/interesting*: enjoyable / fun / interesting.
- *Exposure to varied responses*: exposure to new vocabulary and phrases.
- *Speech capture useful*: speech-to-text was interesting and useful.
- *Accountability (error detection)*: valuable to English learning (checking pronunciation and spelling via the speech to text functions).

AI demerits

- *Non-contiguous*: cannot ask follow up questions.
- *Inappropriate response*: answers received often off topic.
- *Unnatural*: no facial, verbal, emotional cues.
- *Lack of audio response*: text only response / no listening practice.
- *Unintelligible response*: unable to understand the given response (ability?).
- *Audio capture unforgiving*: speech-to-text unforgiving for L2 pronunciation.

- *Uninteresting*: felt very industrial / like talking to a machine.
- *Unable to confirm understanding*: could not check meaning of responses.

3. Results and discussion

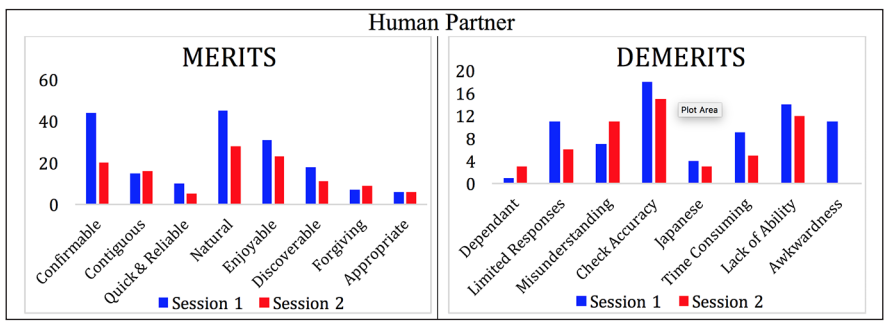
Table 1 depicts the overall number of responses for both human and AI partners. There were more reports of human partner merits than demerits overall, with the discrepancy between them being more than double. Conversely, AI partner demerits outnumbered merits, suggesting that students saw humans as better conversational partners than *Cleverbot*.

Table 1. Total number of responses

TOTAL RESPONSES		
	MERITS	**DEMERITS**
HUMAN Partner		
Session 1	176	75
Session 2	118	55
AI Partner		
Session 1	110	141
Session 2	92	107

Figure 2 and Figure 3 show the distribution of student responses for both human and AI partners between sessions. This illustrates the aspects of conversational interaction that participants found to be of benefit and/or detriment to their language practice.

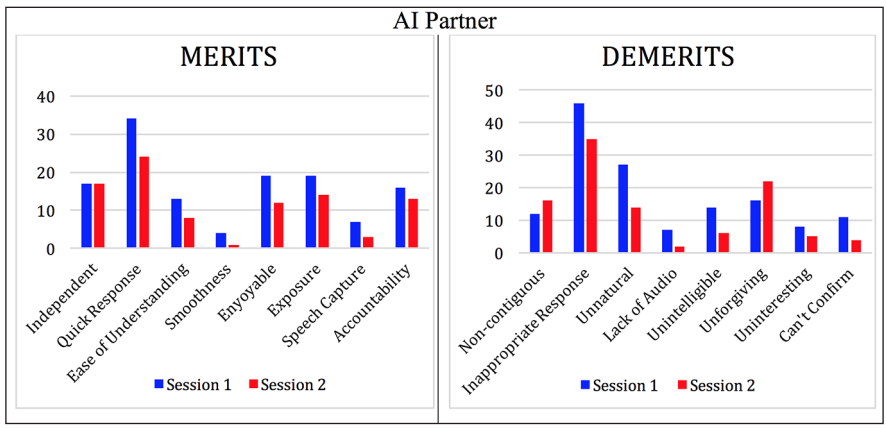
Figure 2. Merits and demerits of human partner



Naturalness of communication, the ability to confirm understanding/meaning, and the enjoyment of interaction were the most frequently reported merits of working

with a human partner. The inverse of the first two were also reflected in the most commonly reported demerits of conversing with a human partner, namely: the inability to check accuracy (without the presence of an input/output field) and the misunderstanding of meaning. The third most reported demerit of human partners was a lack of English ability. Awkwardness disappeared as a category of response in the later session. Perhaps due to the students having become more comfortable with their classmates over time.

Figure 3. Merits and demerits of AI partner



As for AI partners, the most cited merits were: the speed of response, independence of use, and the exposure to new forms of language. Top demerits included: the inappropriateness of responses, technical difficulties (speech to text function on smartphones), the unnaturalness of interaction, and a lack of being able to carry on the conversation. All aspects that are reminiscent of what separates man from machine, and that pose the most difficult hurdles to overcome in developing AI technology of this sort (Coniam, 2014; Fryer & Carpenter, 2006; Goda et al., 2014).

4. Conclusions

In accordance with Atwell (1999) and Parker (2007), students saw the independence of conversing with *Cleverbot* as being beneficial for their English study. Other merits reported for AI partners had more to do with the speech-to-text function of the smartphone rather than the AI itself. Which suggests that the pre-existing functions on students’ smartphones might provide the same benefits as some

iterations of AI, without the added confusion of having to learn a new software platform.

In any case, it is clear that students do not see *Cleverbot* as a viable replacement for communication with another human being. Its lack of emotion, visible cues, and inability to confirm understanding were reported to be some of the major drawbacks to its form of interaction. Therefore, educators should be skeptical of incorporating current AI technology in the L2 classroom as frustration from interaction might outweigh any benefits to its inclusion within an English Curriculum. Future research would benefit from development of a quantifiable survey utilizing the categories presented herein that could provide a more consistent analysis across AI chatbot platforms.

References

Atwell, E. (1999). *The language machine: the impact of speech and language technologies on English language teaching*. British Council.

Coniam, D. (2014). The linguistic accuracy of chatbots: usability from an ESL perspective. *Text & Talk, 34*(5), 545-567. https://doi.org/10.1515/text-2014-0018

Fryer, L., & Carpenter, R. (2006). Emerging technologies. Language in action: from webquests to virtual realities. *Language Learning & Technology, 10*(3), 8-14.

Goda, Y., Yamada, M., Matsukawa, H., Hata, K., & Yasunami, S. (2014). Conversation with a chatbot before an online EFL group discussion and the effects on critical thinking. *Information and Systems in Education, 13*(1), 1-7. https://doi.org/10.12937/ejsise.13.1

Gulz, A. (2004). Benefits of virtual characters in computer based learning environments: claims and evidence. *International Journal of Artificial Intelligence in Education, 14*(3), 313-334.

Morton, H., & Jack, A. M. (2007). Scenario-based spoken interaction with virtual agents. *Computer Assisted Language Learning, 18*(3), 171-191. https://doi.org/10.1080/09588220500173344

Parker, L. (2007). Technology in support of young English learners in and out of school. In L. Parker (Ed.), *Technology-mediated learning environments for young English learners* (pp. 213-250). Routledge.

Wikipedia contributors. (2018, October 5). *Cleverbot*. In Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Cleverbot&oldid=862672526

Published by Research-publishing.net, a not-for-profit association
Contact: info@research-publishing.net

Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018
Edited by Peppi Taalas, Juha Jalkanen, Linda Bradley, and Sylvie Thouësny

Publication date: 2018/12/08

Typeset by Research-publishing.net
Cover theme by © 2018 Antti Myöhänen (antti.myohanen@gmail.com)
Cover layout by © 2018 Raphaël Savina (raphael@savina.net)
Drawings by © 2018 Linda Saukko-Rauta (linda@redanredan.fi)

ISBN13: 978-2-490057-22-1 (Ebook, PDF, colour)
ISBN13: 978-2-490057-23-8 (Ebook, EPUB, colour)
ISBN13: 978-2-490057-21-4 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, UK: British Library.
Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: Décembre 2018.