Article

The Impact of AI on Teaching EFL in Japan

Hiroyuki Obari

Abstract

A constructivist approach to flipped learning can motivate students by activating their brains to create new knowledge and reflect more consistently and deeply on their language learning experiences. The present study focused on assessing the use of the AI speakers Google Home Mini and Amazon Alexia as part of a flipped-learning environment to improve the English skills of two groups of native Japanese undergraduates. The participants were 47 native speakers of Japanese, all third-year Economics majors at a private university in Tokyo. Pretest and posttest TOEIC scores, as well as a post-training survey results, were used in evaluating the overall effectiveness of the program. Gains in TOEIC scores indicated the flipped-learning program that incorporated AI speakers improved the students' overall English skills, especially listening comprehension. The results suggest the integration of AI, along with social media and 21st-century skills, may be an effective way to improve the English language proficiency of adult ELT learners.

Keywords: AI speaker, constructivism, blended learning, flipped learning, 21st-century skills

1. Introduction

Alan Turing (1950) once said, "I propose to consider the question, 'Can machines think?'". Turing invented computers in the 1930s, and soon after became fascinated by the idea that computers might one day be intelligent. In 1950, he published a scientific paper on the question of whether a machine could 'think'. The paper introduced the 'Turing test' (Wooldridge, 2018, p. 6). Ninety years have passed since the invention of the computer. Nowadays AI has intervened

in every aspect of our lives in the 21st century. AI today is everywhere and it will soon become ever more prominent because AI software can reliably and efficiently make better decisions than people in a huge range of settings. In the future, AI will be invisibly embedded everywhere in which decisions are made (Wooldridge, 2018, p. 44).

Recently, artificial intelligence (AI) speakers can be experienced efficiently and smoothly using hand-held devices, which can enhance the construction of broader learning environments and viewpoints (Kepuska & Bohouta, 2018). AI/mobile technologies have succeeded in transforming learning methodologies. One such methodology adopted successfully in recent years is blended learning (BL) (Obari & Lambacher, 2014). BL combines traditional face-to-face classroom methods with computer-mediated activities, resulting in a more integrated approach to language learning. AI, mobile devices, and social media are the key components of the next generation of this novel wave of educational instruction. Digital content is also transforming and expanding as AI and mobile technologies continue to develop and improve. The growth of mobile social networking sites (mSNS) has enabled teachers to considerably increase the number of ideal learning opportunities through experiential learning activities with the help of these emerging technologies (Wong et al., 2015).

The main purpose of this study is to introduce two case studies carried out to ascertain the effectiveness of a BL/FL-training program that incorporated the AI speakers Google Home Mini and Amazon Alexia to improve the English proficiency of native Japanese, including 21st-century learning skills for developing international cultural awareness.

The following two research questions were targeted in the study:

- Could the participants improve their English proficiency and understanding of 21st-century skills after exposure to the AI/BL/FL activities?
- Was the application of AI/mobile learning and Content and Language Integrated Learning (CLIL) helpful in improving the participants' English skills?

Pedagogy of flipped learning and 21st century skills with AI

In 2008, Harvard researcher Tony Wagner identified the "7 Survival Skills"

necessary for the modern workplace (Wagner, 2008):

- Critical thinking and problem solving
- Collaboration
- Agility and adaptability
- Initiative and entrepreneurialism
- Effective oral and written communication
- Accessing and analyzing information
- Curiosity and imagination

In addition to the "7 Survival Skills" mentioned above, we added one more skill which we will refer to as "Coexistence with AI," resulting in 8C skills for 21st-century education. In this study, we attempt to integrate these 8C skills for developing ELT skills and intercultural awareness.

2. Method

2.1. Participants

The study was conducted over a period of two semesters (September 2018 to July 2019). A total of 47 undergraduates participated, all native speakers of Japanese. Participants were divided into two groups: Group 1 (n=24) during the first training period (Sept 2018 to Jan 2019), and Group 2 (n=23) during the second training period (April 2019 to July 2019).

2.2. Training procedure

Technologies utilized included Google Home Mini (Figure 1), Amazon Alexia (Figure 2), ATR CALL Brix, and the mSNS programs Facebook, Twitter, Line, and other on-line language learning programs. TOEIC was used as a measure to determine if the students' English skills improved and to help ascertain the overall effectiveness of the BL/FL program. TOEIC was administered to Group 1 as a pretest in September 2018 and posttest in January 2019, and to Group 2 as a pre-test in April 2019 and post-test in July 2019.

During training, participants of both groups were divided into eight subgroups, with half using Google Home Mini and the other half using Amazon Alexia. The AI speakers were literally integrated into the daily lives of the participants over the four-month training period. A timer was regularly set while interacting with the AI speakers to practice English listening, speaking, and vocabulary skills



Figure 1: Google Home mini



Figure 2: Alexa

using a variety of software programs.

Group 1 used Google Home Mini (Figure 1) on a daily basis to improve their English listening and speaking skills using the following software applications: 'Best Teacher', "Travel English", "Let's play around with English", and BBC/CNN news. Group 2 used Alexia (Figure 2) and Google Home Mini on a daily basis to improve their listening comprehension and vocabulary skills using the following programs: Kikutan, English Quiz by Arc, Liberty English, and Kindle. Participants of both groups also used VR Goggles to interact within a variety of authentic ELT learning environments.

While studying with the AI speakers, participants recorded short movie clips of their learning experiences which were uploaded to Facebook. The participants also kept written diaries with their observations about the contents and duration of their studies, periodically recording their thoughts using a smartphone.

At the end of the training period, participants of all eight subgroups in Group A delivered presentations of their impressions of the BL/FL-lesson training using the AI speakers, with a majority indicating the training had a positive effect on their learning experience. They carried out their own research on how to effectively use AI speaker in learning English. In February 2019, all eight subgroups of Group 1 wrote research papers about their project to prove the effectiveness of using AI speakers and made presentations about the results of their own research.

Including the above AI training procedure, the following additional tasks were incorporated in training:

(1) Practiced English using the AI speakers Google Home Mini and Amazon Alexia.
(2) Watched online TED Talks, including Rick Warren's "Purpose-Filled Life" (Warren, 2012) and other popular subjects, using their mobile devices; wrote 300-word summaries, created PowerPoint presentations, and discussed their summaries with a group of English native speakers two times during the course of 4 months.
(3) Studied worldviews after viewing online lectures by several Oxford University scholars, which focused on ontological and epistemological issues; and then delivered PPT presentations and created digital stories with their iPads. They presented their summaries with a group of English native speakers several times and further discussed the comparative worldviews between East and West.
(4) Use ATR CALL Brix to practice TOEIC and measure their improvements of English proficiency.
(5) Studied the theory behind 21st-century skills with iPads and delivered PowerPoint presentations summarizing the contents.
(6) Interacted with English native speakers who assessed their presentations and discussed worldviews and cultural issues – the native speakers included eight university students from the United States.
(7) Periodically participated in supplementary interactions and discussions.
(8) In both group 1 and 2, flipped lessons were put into practice focusing on the presentations and discussions in the class after they received lots of comprehensible input about the assigned tasks every week.
(9) Every student assessed his or her presentation based upon 6 items; a. structure and visual aids, prosody and pronunciation, c. eye contact and gestures, d. delivery and volume, e. coherence and logic, f. enthusiasm and content. This year we started to use the computer evaluation called "PeerEval" in which each student uses the smartphone with access code and id number to assess presentations (Figure 3). By using this evaluation system, students were able to better understand their strengths and weaknesses among the 6 items. After the presentation, they could easily examine the average score of each category as assessed by the evaluators in timely fashion (see Figure 4). This evaluation system helped the students to improve their presentation skills.

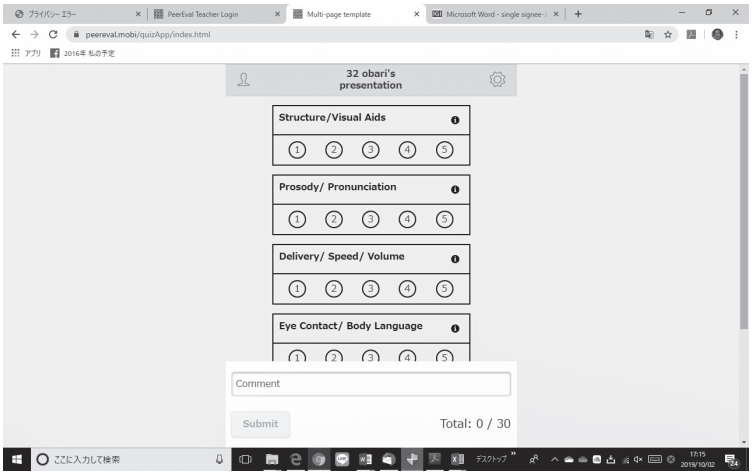
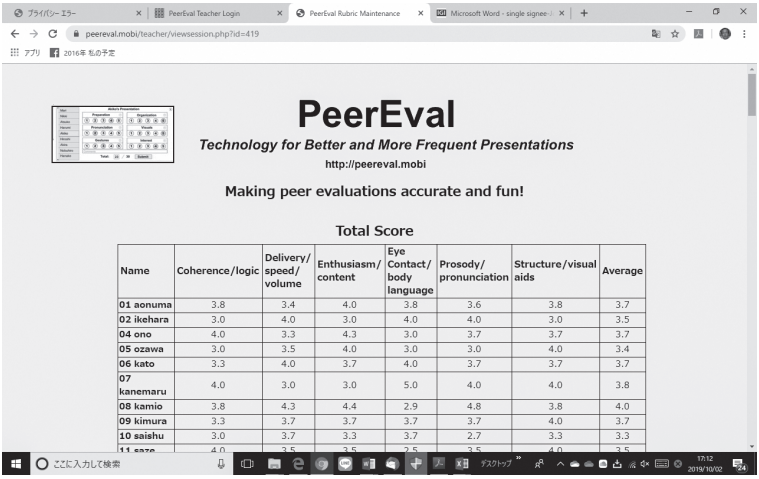


Figure 3: PeerEval



Name	Coherence/logic	Delivery/ speed/ volume	Enthusiasm/ content	Eye Contact/ body language	Prosody/ pronunciation	Structure/visual aids	Average
01 aonuma	3.8	3.4	4.0	3.8	3.6	3.8	3.7
02 ikehara	3.0	4.0	3.0	4.0	4.0	3.0	3.5
04 ono	4.0	3.3	4.3	3.0	3.7	3.7	3.7
05 ozawa	3.0	3.5	4.0	3.0	3.0	4.0	3.4
06 kato	3.3	4.0	3.7	4.0	3.7	3.7	3.7
07 kanemaru	4.0	3.0	3.0	5.0	4.0	4.0	3.8
08 kamio	3.8	4.3	4.4	2.9	4.8	3.8	4.0
09 kimura	3.3	3.7	3.7	3.7	3.7	4.0	3.7
10 saishu	3.0	3.7	3.3	3.7	2.7	3.3	3.3
11 saze	4.0	3.5	3.5	2.5	3.5	4.0	3.5

Figure 4: Assessment of PeerEval

3. Results

3.1 Pre-test/post-test TOEIC

The pretest and posttest results of TOEIC of Group 1 (n=24) during the first training period (September 2018 to January 2019) increased from a mean score

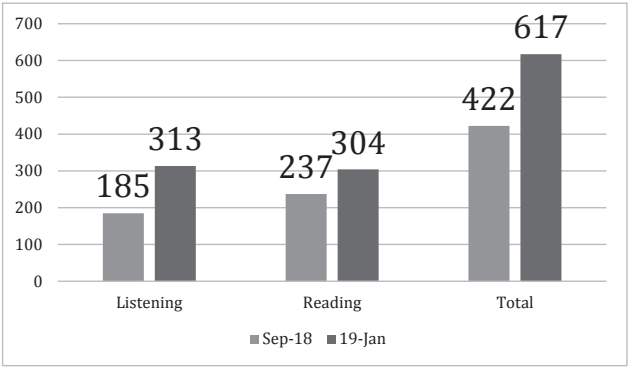


Figure 5: Group 1 pretest vs. posttest TOEIC scores.

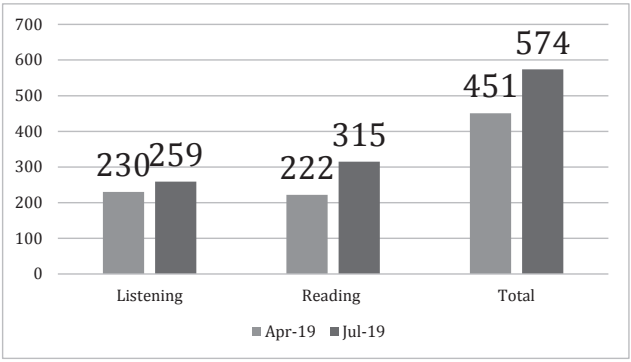


Figure 6: Group 2 pre-test vs. post-test TOEIC scores.

of 422 (*SD:115*) to 617 (*SD:114*), or equivalent to B1 CEFR.

As shown in Figure 5, Group 1 improved both their listening and reading scores. Their listening scores, in particular, increased from a mean of 185 (*SD=55*) to 313 (*SD=76*). Similarly, the overall TOEIC scores of Group 2 participants (n=23) increased from a mean of 451 (*SD:125*) to 574 (*SD:93*), or equivalent to B1 CEFR (Figure 6). The pre-/post-test TOEIC results were analyzed using a series of t-tests, indicating these differences were statistically significant ($p < .01$).

3.2 Post-training questionnaire

Post-training surveys were administered to both groups at the end of their

respective AI/BL/FL training for the purpose of ascertaining their overall impressions of the program. Responses to a few of the survey questions are summarised as follows:

(Q1) The online lectures were beneficial in improving my English proficiency: 88% agreed (n=47).

(Q2) mSNS (Facebook, Line, Twitter) was useful in learning English: 92% agreed (n=47).

(Q3) This program was useful in learning 21st-century skills: 90% agreed (n=47).

(Q4) CCC members helped students to change their worldviews through face-to-face discussions: 88% agreed (n=47).

(Q5) Collaboration works helped students to improve their learning and English proficiency: 94% agreed. (n=47)

(Q6) The AI speaker was useful in improving my English skills: 84% agreed (n=47).

(Q7) The AI speaker was useful in improving listening skills: 87% agreed (n=23).

(Q8) The AI speaker was useful in improving speaking skills: 57% agreed (n=23).

(Q9) The AI speaker was useful in improving reading skills: 13% agreed (n=23).

(Q10) The AI speaker was useful in improving writing skills: 4.3% agreed (n=23).

(Q11) The presentation practice with PowerPoint slides helped students to improve English proficiency: 100% agreed (n=47).

3.3 Post-training comments

(S1) I could set up the AI speaker to a particular time to make studying a part of my daily routine by listening to the news. Alexa has many useful applications for studying English.

(S2) The AI speaker helped me to improve my English listening comprehension and speaking skills. It was much easier to study, and I was more motivated to study English with this AI speaker. I was increasingly exposed to listening to English, which was the significant advantage of using an AI speaker.

(S3) The AI speaker has several advantages – for example, we have more opportunities to listen to authentic English and practice more English speaking with native speakers. We feel as if we are having a conversation with a native speaker of English.

(S4) The AI speaker has two advantages – First, it is easy to use with only one's

voice to study English. Second, it helps to improve English speaking ability. I could be more actively engaged in learning English with the AI speaker. I felt I could improve my English skills over a shorter period.

(S5) I did not have to sit at the desk to study English. I could easily engage in learning English with Alexa. Alexa has many applications for learning English.

(S6) Hurdles to learn English decreased to make learning a part of my daily routine with great pleasure. I could practice English loudly without hesitation.

(S7) The AI speaker could create a virtual environment where I felt I was speaking with a native speaker of English. However, it did not help to improve my reading and writing skills in English. The study is where I usually sit at the desk. However, after I used the AI speaker, my image of studying changed to make me feel the joy of learning.

(S8) We could acquire a variety of news information without using a smartphone by using the AI speaker. Throughout using the AI speaker during the first term, I felt the AI speaker would be very advantageous for early childhood English education due to its many applications. An AI speaker could attract children by motivating them to learn with great pleasure, not just sitting at a desk.

3.4 Integration of keywords as reported by students

All the students gave some comments as a type of feedback about their use of AI, which were incorporated in a Wordcloud. The results are shown below in Figure 7, which included the following words: English, Possible, Learning, AI, and so on. It can be assumed from these words that the students felt they could effectively learn English using the AI speakers (Figure 7).



Figure 7: http://wordcloudjp.com/

4. Discussion

Google Home Mini, Amazon Alexia, ATR CALL Brix, the mSNS programs, Facebook, Twitter, Line, and other language learning programs were used during my seminar course. TOEIC was used as a measure to determine if the students' English skills improved and to help ascertain the overall effectiveness of the BL/FL program. In both Group 1 and Group 2, the mean TOEIC score significantly improved only for one semester. The total average mean score of Group 1 improved 195 points from 422 (SD, 115) to 617 (SD, 114). The mean listening scores (Group 1) improved 128 points from 185 (SD, 55) to 313 (SD, 76), whereas the mean reading scores improved only 67 points. The fact that the students in Group A made more progress in listening than reading comprehension could be due to the effect of using the AI speaker to some extent.

On the other hand, the total mean scores of Group 2 improved 123 points from 451 (SD, 125) to 574 (SD, 93). Both the mean listening and reading scores improved. However, reading scores of Group 2 improved more than their listening scores even though they used AI speakers. According to the questionnaire, 84% agreed that the AI speaker was useful in improving English skills, particularly listening skills, with 87% agreement. On the other hand, only 13% (Group 2, n=23) agreed that the AI speaker was useful in improving reading skills, although the average reading scores improved more than listening scores. Therefore, in the Group 2 study, it may be difficult to claim that the AI speaker had helped the students to improve listening skills more than reading skills.

Fifty-seven percent of students agreed the AI speaker was useful in improving their speaking skills, but at present there is no empirical evidence for this claim, although Group 2 students took the OPIc computer speaking test in April and will take this test again in January 2020.

Students mainly used the AI speaker to study English listening and speaking. That is why only 13% agreed the AI speaker was useful in improving reading skills and only 4.3% in improving writing skills. As the AI speaker was mainly used for listening and for interaction with the AI native speakers of English, the students implied the AI speaker was not useful in studying writing and reading skills.

In addition, the following additional 7 tasks (2) to (8) in Section 2 (Method) were incorporated in training. In both Groups 1 and 2, flipped lessons were put

into practice focusing on the presentations and discussions in the class after students gained sufficient comprehensible input about the assigned tasks each week. As a result, their listening and speaking skills, and writing skills to some extent, were reinforced.

The students of both groups studied worldviews and discussion with CCC (Campus Crusade for Christ), which helped them become more internationally-minded and changed their worldviews, as 88% agreed the interactions with CCC members influenced their worldviews through face-to-face interactions.

Learning 21st century skills was integrated into the seminar and 90% of students agreed that this program was useful in learning 21st century skills. In Group 1 and Group 2, a variety of digital technologies were used for students to be more engaged in learning a variety of topics. The distinction between the classroom activities and outside-class activities was clearly drawn so that each student could get as much as comprehensible input and be more engaged in interaction and discussions with critical and creative thinking.

This research included some limitations in attempting to ascertain the effectiveness of using AI in teaching language and communication. It was a challenge to carry out this research by exclusively focusing on the role of AI and how effectively it improved the students' English skills, e.g., what area of learning and how it was affected, etc., as a result of integrating various types of technologies with many learning strategies. What is most important in learning English is face-to face interaction and how much comprehensible input we can gain from various learning sources whether it be analogue or digital environments.

5. Conclusions

An assessment of the pretest and posttest TOEIC scores revealed a combination of the AI/BL/FL lessons had a positive effect on their English language learning. Students' listening and oral communication skills improved, which may have been due to the integration of the language learning activities which focused on a social constructivist approach while utilizing the AI smart speakers Alexia and Google Home Mini. Additionally, the post-course questionnaire revealed both groups were satisfied with the online course materials and motivated by the AI/BL/FL environment incorporating 21st-century skills. Students cited improved recognition of ambiguities within cross-cultural contexts and improved global

communication skills. The students' oral summaries also improved through interaction with the AI smart speakers and native English speakers. Taken as a whole, these results would seem to indicate the integration of AI smart speakers into the BL/FL training program played a role in improving the students' overall language proficiency and expanding their worldviews. However, we should mention that this study was merely a modest start and further research utilizing AI speakers exclusively in training will be required to more accurately measure native Japanese improvement in English proficiency.

6. Acknowledgements

This work was supported by a grant from Economic Research Institute in the College of Economic and was integrated from my presentations and proceedings of several international conferences. Finally, I would like to express my hearty thanks to Professor Steve G. Lambacher of Social Informatics who checked this paper and revised some of my writings to bring this paper to completion.

References

A. M. Turing (1950). Computing Machinery and Intelligence. *Mind 49*: 433–460.

Kepuska, V., & Bohouta, G. (2018, January). Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference* (CCWC) (pp. 99–103). IEEE. https://doi.org/10.1109/ccwc.2018.8301638

Obari, H., & Lambacher, S. (2014). Impact of a blended environment with m-Learning on EFL skills. In S. Jager, L. Bradley, E. J. Meima & S. Thouësny (Eds), *CALL design: principles and practice - proceedings of the 2014 EUROCALL Conference, Groningen, The Netherlands* (pp. 267–272). Research-publishing.net. https://doi.org/10.14705/rpnet.2014.000229

Thomas Rob, PeerEval, https://peereval.mobi/(2019.12.20)

Wagner, T. (2014). *The global achievement gap: why even our best schools don't teach the new survival skills our children need-and what we can do about it*. Hachette UK.

Wagner, T. (2016). 7 *Survival skills*. http://www.21stcenturyschools.com/tony-wagner.html. Retrieved 2018–11–03

Warren, R. (2012). *The Purpose Driven Life: what on earth am I here for?* Zondervan.

Wooldridge, Michael. (2018). *Artificial Intelligence*. A Ladybird Expert Book., Penguin Random House, UK.

Wong, C. H., Tan, G. W. H., Loke, S. P., & Ooi, K. B. (2015). Adoption of mobile social networking sites for learning? *Online Information Review*, 39(6), 762–778. https://doi.org/10.1108/oir-05-2015-0152