Language Learning & Technology
ISSN 1094-3501

February 2020, Volume 24, Issue 1
pp. 16–26

LANGUAGE TEACHING AND TECHNOLOGY FORUM

Evaluating intelligent personal assistants for L2 listening and speaking development

Gilbert Dizon, Himeji Dokkyo University

Abstract

While the use of intelligent personal assistants (IPAs) has exploded in recent years, little is known about their use to promote English as a foreign language (EFL) development. Thus, this study addresses this gap in the literature by examining the in-class use of the IPA, Alexa, among second language (L2) English students to support improvements in listening comprehension and speaking proficiency. The study utilized a quasi-experimental design with an experimental group (n = 13) which took part in a 10-week treatment of student-IPA interaction and a control group (n = 15) which did not. Results from the Mann-Whitney U test found that the experimental group was able to make more significant gains in L2 speaking proficiency. However, a significant difference was not found when comparing improvements in L2 listening comprehension. These findings suggest that IPAs may be a useful tool to promote L2 speaking skills and underscore the necessity for additional research on the emerging technology for language learning.

Keywords: *Intelligent Computer-Assisted Language Learning (ICALL), L2 Listening, L2 Speaking*

Language(s) Learned in This Study: *English*

APA Citation: Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology, 24*(1), 16–26. https://doi.org/10125/44705

Introduction

Grounded in an interactionist approach (Long, 1996), this study examined the use of intelligent personal assistant (IPAs), that is, cloud-based virtual assistants such as Siri and Alexa, to support improvements in second language L2 listening comprehension and speaking proficiency. In the interactionist perspective, IPAs seem to be a beneficial method of language learning as they promote interactions in the L2 and may direct learners to gaps in their linguistic knowledge, particularly when it comes to pronunciation (Dizon, 2017). While interactionist literature in L2 learning has primarily examined interactions between native language and L2 speakers or L2 peer interaction (e.g., Garcia Mayo & Pica, 2000; Gass & Varonis, 1990; Pica, Lincoln-Porter, Paninos, and Linnell, 1996; Philp & Iwashita, 2013), there has been a rise in the use of technology in interactionist research (Loewen & Sato, 2018). Specifically, much attention has been paid to the use of computer-mediated communication (CMC) and its effects on L2 writing. Although research pertaining to CMC and L2 speaking is still limited, studies have shown that CMC applications have the potential to promote negotiation of meaning and focus on form (Bueno-Alastuey, 2013; Yanguas, 2010), two important tenets of the interactionist approach.

The use of dialogue-based computer-assisted language learning (CALL) in interactionist research has also been popular in recent years due to technological advancements in natural language processing and automatic speech recognition (ASR). While some criticize ASR systems due to their inability to accurately understand L2 speech at the same rate as human listeners (Derwing, Munro, & Carbonaro, 2000), as McCrocklin (2016) notes, more recent studies have shown that the technology has been improving in this regard (Moustroufas & Digalakis, 2007). Compared to CMC in which a learner interacts through a computer, students interact with a computer in dialogue-based CALL. According to Bibauw, Francois, and Desmet (2019), research on dialogue-based CALL centers on the assumption that "meaningful practice of

a target language, as it occurs in conversation, leads to improvement in the learner's proficiency in that language, and that, even if a native speaker remains the ideal interlocuter, a computer can provide opportunities for such practice" (p. 3). As the researchers point out, the interactionist perspective features prominently in dialogue-based CALL systems, which include chatbots and ASR-based pronunciation training, as these types of systems afford opportunities for input and interaction, as well as pushed output (i.e., language that is produced when learners are encouraged to use the target language in a succinct and accurate manner; Ellis, 2003). In foreign language contexts, this is of particular importance because meaningful practice of the L2 is often limited to the confines of the language classroom (Fryer & Carpenter, 2006). Additionally, although the development of oral proficiency is time-consuming, ASR-based systems and CALL provide learners with the ability to practice their speaking skills for as long as they want in an anxiety-reduced environment, thereby leading to more speaking opportunities (Neri, Cucchiarini, & Strik, 2006; Wallace, 2015). Therefore, given the aforementioned benefits of CALL-based interaction on L2 acquisition, it is critical to investigate emerging technologies to understand the affordances they can provide language learners. With this goal, the present study investigated the use of Alexa, an IPA by Amazon, to promote improvements in EFL listening comprehension and speaking proficiency. In the context of this study, listening comprehension is defined as the ability to understand key details from spoken language, whereas speaking proficiency refers to the capacity to produce comprehensible speech with appropriate vocabulary, grammatical accuracy, and pronunciation (Payne & Whitney, 2002). The following research questions were addressed in the study:

1. Were there any significant differences in listening comprehension or speaking proficiency development between an experimental group that interacted with Alexa and a control group that did not?
2. What are EFL students views towards the use of Alexa as an in-class language learning activity?

Literature Review

Dialogue-based CALL

Dialogue-based CALL generally has a positive impact on affective factors in L2 learning, such as student interest and motivation (Anderson, Davidson, Morton, & Jack, 2008; Jia & Chen, 2009; Kaplan, Sabol, Wisher, & Seidel, 1998; Sha, 2009). In a survey-based study on chatbots, Sha (2009) found that L2 English students had a high level of interest when interacting with the target CALL system, an AI-based chatterbot called Verbot. Jia and Chen (2009) also examined students' interest in and attitudes towards an English computer dialogue system through surveys. Based on their results, the researchers concluded that the program enhanced the participants' interest in language learning and also increased their confidence. However, it is important to note that age and experience played a role in the results of the researchers' study —university students had more positive views of the CALL system than middle and high school students. In another study, Kaplan et al. (1998) examined the use of an intelligent tutoring system with L2 learners of Arabic. All of the learners perceived the system to be a good method for L2 learning, with less proficient learners having more favorable views than those with advanced proficiency. Anderson et al. (2008) also investigated the use of a dialogue-based CALL system, specifically, a self-access CALL program which incorporated a virtual world with L2 learners from a variety of language backgrounds. The results of their research revealed that the majority of the learners perceived it to be engaging, useful, and enjoyable. The system also had positive effects on motivation, particularly among L2 students who were intrinsically, rather than externally, motivated.

As stated by Bibauw et al. (2019), effectiveness studies related to language learning outcomes and dialogue-based CALL often have resulted in positive findings in support of such systems. Kim (2016) examined the influence of voice chat on Korean learners of L2 English and, according to the results of the study, the system had a positive, significant effect on oral proficiency. Form-focused instruction in dialogue-based CALL also seems to have a positive influence on learners. Wilske and Wolska (2011) examined the use of computer-based dialogue on L2 learners and found that it had a positive impact on L2 German target

structures. It was also found that direct feedback had a greater positive effect than implicit feedback, demonstrating that the former is ideal in dialogue-based CALL instruction. In a study on computer assisted pronunciation training (CAPT), it was determined by Chiu, Liou, and Yeh (2007) that L2 English CAPT training supported significant improvements in the teaching of speech acts. In another study on CAPT, Liakin, Cardoso, and Liakina (2015) found that mobile ASR-based pronunciation training had a more significant impact than non-ASR training in the development of L2 French segmental features. Kaplan et al. (1998) found that the intelligent tutoring program had a significant positive effect on translation skills, even after just one hour of training. Despite the positive results found in these studies, Bibauw et al. (2019) cautions that more rigorous empirical research is needed to better assess the efficacy of dialogue-based CALL to enhance language learning, as most studies suffer from small sample sizes, as well as potential conflicts of interest as researchers often evaluate systems that they themselves developed.

IPAs in L2 Research

Research on the use of IPAs for L2 learning is scarce and limited to small-scale studies. Moussalli and Cardoso (2016) conducted a feasibility study on the use of Alexa as a tool to extend language learning outside of the classroom with four female L2 English students. The learners in the study interacted with the virtual assistant for about 30 minutes and asked it pre-determined and leaner-generated questions. The researchers also found that they enjoyed interacting with the virtual assistant and viewed it as a helpful tool for language learning. Specifically, student comments in interviews suggested that the learners were able to receive implicit feedback from the IPA. One drawback that was found in their study was that some students commented on the difficulty of being understood by the virtual assistant.

In a recent follow-up study, Moussalli and Cardoso (2019) again examined the use of Alexa with a greater number of L2 students (N = 11) and found that they used a variety of strategies, namely, repetition, rephrasing, or abandonment, to resolve breakdowns in communication. However, the authors found that the IPA did not have difficulty understanding accented speech, which contradicts the findings in their previous study. There may be several reasons behind the discrepancy in intelligibility between the researchers' two studies. First, the small sample size in both studies could have been a contributing factor. Differences in language proficiency could have also played a role. Additionally, the researchers employed different methodologies in the two studies; the earlier study relied on survey and interview responses to assess comprehensibility, while their more recent research utilized the transcripts, a more objective measure, along with a survey and interviews. Furthermore, as noted by Bajorek (2019), voice recognition technologies such as Alexa suffer from gender- and race-related biases, which in turn, may negatively affect intelligibility for female users. Although unclear, this might have influenced Moussalli and Cardoso's (2016) study because it only had female participants.

Lastly, a study by Dizon (2017) examined using Alexa for L2 learning in a case study with four Japanese EFL learners. The study found mixed results regarding the capacity of the virtual assistant to understand L2 speech. However, according to data obtained in interviews, the students had generally favorable perceptions of its use for EFL learning. Specifically, student comments indicated that the IPA supported learner effectiveness through indirect pronunciation feedback and also provided better access to conversational opportunities in the target language.

While the aforementioned research highlights the potential of IPAs for L2 learning, to date, no studies have investigated their use to enhance L2 listening and speaking skills. In addition, student perceptions towards virtual assistants for in-class language learning are not yet known. Thus, the present study works toward filling these gaps in CALL literature.

Methodology

Participants

A total of 37 first- and second-year EFL students at a Japanese university participated in the study during

the fall semester of 2018. Based on their scores on the EIKEN test, a standardized English-proficiency exam, the students' ability levels ranged from A1 to B2 on the CEFR proficiency scale. The participants were made up of both male and female students ranging in age from 18 to 20 years old. All of the students were enrolled in the same elective English course which focused on speaking development, namely, conversation, discussion, and presentation skills. The course met three times a week in 90-minute classes over a 15-week semester. Four intact classes took part in the study, two of which were taught by the researcher and were part of the experimental group, while the remaining two classes were taught by different instructors and were part of the control group. Those in the experimental group were surveyed at the start of the semester regarding prior experience with smart speakers and all of them responded that they had never used one before. Each group, experimental and control, was made up of both first- and second-year students. Due to attendance issues, only 28 students (experimental = 13, control = 15) were able to take both the pre- and post-tests and as a result, only the data from these students was examined in the study.

Target IPA: Alexa

Alexa, an IPA developed by Amazon, was first introduced in 2014 when the company launched its smart speaker, the Echo. Interacting with Alexa is similar to interacting with other IPAs: Users need to say the wake word, in this case, *Alexa*, and then give it a command or make a request. Since its introduction, popularity for smart speakers has soared, with a quarter of U.S. households owning at least one smart speaker, according to a Nielsen (2018) report. Another interesting finding from the report was how people use IPAs. Unsurprisingly, the most common uses were listening to music and searching for real-time and factual information. More surprising, however, was that 68% of users talked with their smart speakers for fun, illustrating the conversational potential of voice recognition technologies.

Alexa was designated as the target IPA for two reasons: Affordability and versatility. Alexa-enabled devices come in at a range of prices, with the cheapest smart speaker by Amazon priced at $25. As Bateson and Daniels (2012) correctly point out, cost should not be overlooked when it comes to CALL, as students may have limited financial resources to spend on language learning. The second reason Alexa was chosen was due to the number and quality of third-party skills that can be used (over 70,000 as of January 2019), particularly in terms of ones that have language learning potential. For instance, skills such as Magoosh Vocabulary Builder and Word of the Day Quiz Game can be used by L2 English learners to support vocabulary development. Interactive storytelling skills like Earplay, in which users are guided through an audio story and have to orally respond to prompts, also give learners opportunities for English input and output practice. Furthermore, Alexa offers socialbots that students can have conversations with. These socialbots, which are developed by university teams from around the world, are challenged with having an engaging and coherent conversation with a human for 20 minutes. These skills and socialbots were included in the list of suggested commands the students could use when interacting with Alexa. Lastly, it is important to note that all the skills and features offered by Alexa are free to enable, thereby adding more value to it as a cost-effective way to study a language.

Treatment

Those in the experimental group took part in a 10-week in-class treatment consisting of 12 minutes of Alexa interaction each week via the Echo Dot, a smart speaker by Amazon. Students interacted with the IPA individually with their own Echo Dot or in pairs. Prior to the beginning of the treatment, the researcher gave a short training session, showing the students how to use the speaker and demonstrated possible commands they could give Alexa. A list of commands and skills was provided to give them an idea of what types of questions and requests could be given to the IPA. The list contained approximately 70 commands and skills, including the aforementioned vocabulary and interactive storytelling skills (see Table 1 below for a sample interaction), and went through several iterations before it was finalized. With that said, students were strongly encouraged to come up with their own commands when interacting with Alexa. If Alexa failed to understand a command or request, the students were instructed to reflect on what part of their speech was misunderstood and adjust their output (e.g., speak more clearly, or more slowly or modify

pronunciation of a specific word in order to increase intelligibility). In other words, the learners were responsible for testing hypotheses regarding how to say specific commands or requests and for monitoring their own production for gaps in linguistic knowledge, key aspects of Swain's (1993) Output Hypothesis as well as interactionist theory (Chapelle, 2005).

Table 1. *Sample interaction from interactive storytelling skill*

Speaker	Transcript
Narrator:	A young woman you've never seen in your life sits down at your table. She looks anxious, "Pretend you know me." Do you ask questions or play along?
Student:	Ask questions.
Narrator:	You ask her "What's going on?" "I think I'm being followed. I've done something I probably shouldn't have. Five minutes, that's all I ask then I'm out of your life forever." A server comes over and puts two drinks on your table. It's not the same waiter who's been refilling your water glass. And you never ordered any drinks. You can send them back, take a drink, or leave them alone.
Student:	Take a drink.

During the training period, it became clear that noise was going to be a problem in the classroom. Students were having difficultly hearing Alexa's responses due to the number of students interacting with the IPA at the same time. In order to alleviate this issue, the researcher had the students interact with Alexa in three small private rooms in the back of the classroom during the treatment. This ensured that the students would not be distracted by hearing other smart speakers or their peers. In the first week of the treatment, the learners interacted with Alexa in pairs, whereas in week two, the students interacted with the virtual assistant individually. This pattern, alternating between individual use and student dyads, followed for the duration of the treatment, and was implemented to understand which condition the students were most comfortable with when interacting with the IPA.

Research Instruments

Listening Assessment

Two listening comprehension tests consisting of 12 items each were created according to a similar procedure outlined by Chang and Read (2006), as their study provided a clear blueprint on how to develop a well-designed EFL listening assessment. Moreover, it gave insight into what factors support listening comprehension, specifically, repeated input (3 exposures), topic preview, and question preview. Due to time constraints, only question preview was implemented to support listening comprehension in the current study. Following the procedure of Chang and Read (2006), two pilot studies were conducted at the same university in the previous semester prior to the main study. In the first pilot, two EFL students helped evaluate potential audio passages for the listening tests. The students rated five passages from an intermediate English listening textbook (Chase, 2013) and rated the talks on topic familiarity, speech rate, vocabulary, and comprehension. Afterward, a second pilot was conducted with two other EFL students to assess the procedure of the listening assessments and clarity of test items. Based on feedback provided by the students, several items were modified in order to enhance clarity, and L1 instructions were added to increase understanding of the testing procedure. In the present study, the first-year students took version A of the assessment prior to the treatment and version B after the treatment, while the second-year students took version B for the pre-test and version A for the post-test. This was done so that student memory of each listening assessment would not affect performance on the tests.

Speaking Assessment

A speaking proficiency test developed by Payne and Whitney (2002) was adopted in this study for two reasons. Firstly, although the oral proficiency interview (OPI), which is based on ACTFL Proficiency

Guidelines, is the most commonly used test to measure oral proficiency, it is not able to detect improvements in speaking that may develop in a single semester (Payne & Whitney, 2002). Secondly, the speaking test created by Payne and Whitney centers around a more simplified definition of oral proficiency than the ACTFL OPI, specifically, the ability to produce comprehensible speech through the use of appropriate vocabulary, grammatical accuracy, and pronunciation. This is the kind of speech that would likely be understandable by an IPA, as these technologies were explicitly created for use by native speakers. During the assessment, students selected one of three prompts at random. Each prompt was written in Japanese so that the task was fully understandable for the students. They were then instructed to speak about the given topic for at least 1.5 minutes. The topics related to tasks that L2 students at this level should have mastered (talk about a trip; talk about their weekly routine; talk about their weekend). If a student ran out of things to say about a topic before reaching 90 seconds, the learner repeated the task with a different prompt until he or she reached the desired length of response.

Survey

A 10-item Likert-scale survey ranging from *strongly disagree* (1) to *strongly agree* (5) was developed to assess the students' perceptions of Alexa for in-class L2 learning (see Appendix). The survey was first created in English by the researcher and then translated into the L1 by a Japanese colleague so that the items would be clear and understandable. The first eight items were adapted from the technology acceptance model (TAM) literature (Davis, 1989; Huang, Huang, Huang, & Lin, 2012). The TAM is one of the most impactful and widely utilized research frameworks used to study a person's acceptance of a given technology (Lee, Kozar, & Larsen, 2003), thus it was adopted as the primary construct within which to examine the learners' views of the IPA. Four of the main constructs from the TAM (Huang, et al., 2012) were measured in the study: Perceived ease of use (PEOU), perceived usefulness (PU), attitudes towards use (ATU), and behavioral intention (BI). As defined by Davis (1989) PEOU is a user's belief that using a particular technology will be free of effort, while PU is a user's belief that a technology will enhance that person's performance. ATU refers to a user's perception of a tool's overall favorability (Fishbein & Ajzen, 1975). Lastly, BI is an individual's assessment of how likely s/he is to perform a specific behavior (Fishbein & Ajzen, 1975). A strong BI is a reflection that a user will accept and use a technology in the classroom (Yi, Jackson, Park, & Probst, 2006). The last two items on the survey were concerned with the participants' use of Alexa (individual vs. in pairs) to better understand how the L2 learners preferred to interact with the virtual assistant.

Data Collection and Analysis

In order to evaluate the effectiveness of IPAs on L2 listening comprehension and speaking development, pre- and post-tests were administered to the students. Assessment of both groups was administered in the same classroom and time periods. The listening assessments were completed as a group while the speaking tests were audio-recorded and administered individually. After testing was completed, the experimental group also took a post-treatment survey in order to understand their views of Alexa as a formal language learning activity.

Following data collection, two native English speaking university instructors who did not work at the same institution where the participants studied were asked to rate the students' speaking test responses. The raters evaluated the participants' responses according to an oral production scale (maximum = 50 points) created by Payne and Whitney (2002). The scale had five criteria: Comprehensibility, fluency, vocabulary use, syntax and grammar, and pronunciation. On the first day of scoring, the raters compared their evaluations of 10 students' pre- and post-test responses. They were then instructed to rate the remaining test responses individually. Interrater reliability was 0.80. The scores of each rater were averaged to produce pre- and post-test scores. Because of the small sample size and non-normal distribution of some of the data, the Mann-Whitney U test was used to determine if there were significant differences between any potential gains that were made in L2 listening and speaking skills between the control group and the experimental group. Results of the pre-test indicated that there was not a significant difference in L2 English listening ($U = 83.5$, $p = .535$) or speaking skills ($U = 81.5$, $p = .477$) between the two groups. In other words, the groups

were equivalent in ability regarding the two variables studied at the outset of the study.

Results and Discussion

As Table 2 depicts, both the control and experimental groups made slight gains in L2 listening development. However, the difference between the groups' improvements was not significant (U = 97, p =.989). This result suggests that the use of IPAs such as Alexa may not have a significant impact on L2 listening comprehension. A possible explanation for this is that the students may not have been able to adequately understand Alexa's responses. The IPA was developed for use by native speakers, thus, its rate of speech and level of vocabulary may make it difficult for L2 listeners to sufficiently comprehend the virtual assistant. According to interactionist theory, one of the primary benefits of interaction is obtaining modified input (Chapelle, 2005). Yet, at this point in their development, IPAs are unable to provide any modifications to its output that would promote enhanced L2 comprehension. Another potential reason for the lack of significance is that the learners might have focused more on speaking practice during interactions and did not pay much attention to Alexa's responses. Therefore, it may be beneficial to have students write down or orally repeat an IPA's responses in order to confirm comprehension. Moreover, the use of interactive storytelling skills and socialbots should be encouraged as these features emphasize both listening and speaking practice.

Table 2. *Descriptive Statistics of Listening and Speaking Test Results*

		Pre-test		**Post-test**	
Variable	**Group**	***M***	***SD***	***M***	***SD***
Listening	Control	5.27	2.02	6.20	1.66
	Experimental	5.85	2.51	6.69	1.93
Speaking	Control	19.97	5.81	19.20	5.38
	Experimental	18.77	5.56	20.46	6.08

Results from the speaking tests revealed that while the experimental group was able to make a small gain, the control group performed worse on the post-test compared to the pre-test. The Mann-Whitney U test also revealed that the difference between the two groups' gains, or lack thereof, was significant (U = 45.5, p = .017). Even without considering the significance between the control and experimental groups' performance, the fact that the latter made gains in L2 speaking highlights the potential of IPAs to support foreign language development, as well as reinforces previous findings regarding the positive effects dialogue-based CALL can have on language learners (Bibauw et al., 2019). This is particularly important in foreign language contexts such as Japan where learners likely have few opportunities for meaningful speaking opportunities outside of the classroom.

Table 3. *Survey Results*

Construct	***M***	***SD***
Perceived ease of use	4.00	0.89
Perceived usefulness	3.96	0.82
Attitudes towards use	3.96	0.69
Behavioral intention	3.88	0.86
Individual use	3.38	1.04
Pair use	3.23	0.73

The students' responses to the survey, which assessed their views of Alexa, are summarized in Table 3. Descriptive results indicate that the learners' perceptions were moderately favorable, with all four TAM constructs having mean values close or equal to 4 (*agree*). This implies that the students not only enjoyed using the IPA for L2 learning, but also perceived it to be a useful tool to study English, which supports past research on the use of virtual assistants in foreign language contexts (Dizon, 2017; Moussalli & Cardoso, 2016, 2019). In terms of the students' use of Alexa, there was not a strong preference in support for either individual or pair use. While unclear, this finding may indicate that L2 learners feel equally as comfortable using Alexa individually as they do with a partner. Nevertheless, the role that individual or pair use had on the participants is not known; thus, a future study incorporating IPAs for in-class use could examine how each condition affects the learning process or student interaction with virtual assistants.

Conclusion

L2 interaction plays a prominent role in the development of listening and speaking skills. Yet, due to several reasons, including anxiety, lack of time, or mere reluctance (Wallace, 2015), students may not make the required effort to interact with another speaker. Thus, it is vital to examine ASR-based technologies such as IPAs since they seemingly offer a practical solution to these issues (Neri, Cucchiarini, & Strik, 2006). Hence, the goal of this preliminary investigation was to evaluate the influence that Alexa could have on L2 listening comprehension and oral proficiency, as well as to investigate students' opinions of their use for formal language learning. Although it was found that the use of Alexa may not have contributed to the development of L2 listening, results from the study suggest that IPAs may promote improvements in L2 speaking. However, the results from the current study should be met with some caution, as the small sample size limits the generalizations that can be made from its findings. Thus, future studies should incorporate large groups taken from a randomized student population. Moreover, teacher effect might have impacted the results, so this is an additional limitation of the study. Nevertheless, the current study illustrates that IPAs may have some value to foreign language learners. As a result, more attention needs to be placed on investigating how they can be used to promote language development.

References

Anderson, J. N., Davidson, N., Morton, H., & Jack, M. A. (2008). Language learning with interactive virtual agent scenarios and speech recognition: Lessons learned. *Computer Animation and Virtual Worlds*, *19*(5), 605–619. https://doi.org/10.1002/cav.265

Bateson, G., & Daniels, P. (2012). Diversity in technologies. In G. Stockwell (Ed.), *Computer-assisted language learning: Diversity in research and practice* (pp. 127–146). New York, NY: Cambridge University Press. https://doi.org/10.1017/cbo9781139060981

Bajorek, J. P. (2019). Voice recognition still has significant race and gender biases. Retrieved from https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases

Bibauw, S., Francois, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, *31*(8), 1–51. https://doi.org/10.1080/09588221.2018.1535508

Bueno-Alastuey, M. C. (2013). Interactional feedback in synchronous voice-based computer mediated communication: Effect of dyad. *System 41*(5), 543–559. http://dx.doi.org/10.1016/j.system.2013.05.005

Chang, A. C-S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly*, *40*(2), 375–397. https://doi.org/10.2307/40264527

Chapelle, C. (2005). Interactionist SLA theory in CALL research. In J. L. Egbert & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 53–64). Mahwah, NJ: Lawrence Erlbaum.

Chase, B. T. (2013). *Pathways 1: Listening, Speaking, & Critical Thinking*. Boston, MA: National Geographic Learning.

Chiu, T.-L., Liou, H.-C., & Yeh, Y. (2007). A study of web-based oral activities enhanced by automatic speech recognition for EFL college learning. *Computer Assisted Language Learning*, *20*(3), 209–233. https://doi.org/10.1080/09588220701489374

Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Derwing, T. M., Munro, M. J., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly*, *34*(3), 592–603. https://doi.org/10.2307/3587748

Dizon, G. (2017). Using intelligent personal assistants for second language learning: A case study of Alexa. *TESOL Journal*, *8*(4), 811–830. https://doi.org/10.1002/tesj.353

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, *10*(3), 8–14. Retrieved from https://www.lltjournal.org/item/2555

Garcia Mayo, M. P., & Pica, T. (2000). L2 learner interaction in a foreign language setting: Are learning needs addressed? *International Review of Applied Linguistics in Language Teaching*, *38*(1), 35–58. https://doi.org/10.1515/iral.2000.38.1.35

Gass, S. M., & Varonis, E. (1990). Miscommunication in nonnative speaker discourse. In N. Coupland, H. Giles & J. Wiemann (Eds.), *"Miscommunication" and Problematic talk* (pp. 121–145). London: Sage Publications.

Huang, Y. M., Huang, Y. M., Huang, S. H., & Lin, Y. T. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education*, *58*(1), 273–282. https://doi.org/10.1016/j.compedu.2011.08.008

Jia, J., & Chen, W. (2009). The further development of CSIEC project driven by application and evaluation in English education. *British Journal of Educational Technology*, *40*(5), 901–918. https://doi.org/10.1111/j.1467-8535.2008.00881.x

Kaplan, J. D., Sabol, M. A., Wisher, R. A., & Seidel, R. J. (1998). The Military Language Tutor (MILT) program: An advanced authoring system. *Computer Assisted Language Learning*, *11*(3),265–287. http://doi.org/10.1076/call.11.3.265.5679

Kim, N.-Y. (2016). Effects of voice chat on EFL learners' speaking ability according to proficiency levels. *Multimedia-Assisted Language Learning*, *19*(4), 63–88. Retrieved from http://journal.kamall.or.kr/wp-content/uploads/2017/01/Kim_19_4_03.pdf

Lee, Y., Kozar, K. A., & Larsen, K. R.T. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems*, *12*, 752–780. https://doi.org/10.17705/1CAIS.01250

Liakin, D., Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal*, *32*(1), 1–25. https://doi.org/10.1558/cj.v32i1.25962

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, *51*(3), 285–329. https://doi.org/10.1017/S0261444818000125

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413–468). New York: Academic Press.

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System*, *57*, 25–42. https://doi.org/10.1016/j.system.2015.12.013

Moussalli, S., & Cardoso, W. (2016). Are commercial 'personal robots' ready for language learning? Focus on second language speech. In S. Papadima-Sophocleous, L. Bradley, & S. Thouësny (Eds.), *CALL communities and culture – short papers from EUROCALL 2016* (pp. 325–329). https://doi.org/10.14705/rpnet.2016.eurocall2016.583

Moussalli, S., & Cardoso, W. (2019). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, *32*, 1–26. https://doi.org/10.1080/09588221.2019.1595664

Moustroufas, N., & Digalakis, V. (2007). Automatic pronunciation evaluation of foreign speakers using unknown text. *Computer Speech and Language*, *21*(1), 219–230. https://doi.org/10.1016/j.csl.2006.04.001

Neri, A., Cucchiarini, C., & Strik, H. (2006). ASR-based corrective feedback on pronunciation: Does it really work? In *Proceedings of the 9th International Conference on Spoken Language Processing INTERSPEECH 2006* (pp. 1982–1985), Pittsburgh, USA.

Nielsen. (2018). (Smart) speaking my language: Despite their vast capabilities, smart speakers are all about the music. Retrieved from https://www.nielsen.com/us/en/insights/article/2018/smart-speaking-my-language-despite-their-vast-capabilities-smart-speakers-all-about-the-music/

Payne., J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, *20*(1), 7–32. https://doi.org/10.1558/cj.v20i1.7-32

Philp, J. & Iwashita, N. (2013). Talking, tuning in and noticing: Exploring the benefits of output in task-based peer interaction. *Language Awareness*, *22*(4), 353–370. https://doi.org/10.1080/09658416.2012.758128

Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly 30*(1), 59–84. https://doi.org/10.2307/3587607

Sha, G. (2009). AI-based chatterbots and spoken English teaching: A critical analysis. *Computer Assisted Language Learning*, *22*(3), 269–281. https://doi.org/10.1080/09588220902920284

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, *50*(1), 158–164. https://doi.org/10.3138/cmlr.50.1.158

Wallace, L. (2015). Reflexive photography, attitudes, behavior, and CALL: ITAs improving spoken English intelligibility. *CALICO Journal*, *32*(3), 449–479. https://doi.org/10.1558/cj.v32i3.26384

Wilske, S., & Wolska, M. (2011). Meaning versus form in computer-assisted task-based language learning: A case study on the German dative. *Journal for Language Technology and Computational Linguistics*, *26*(1), 23–37.

Yanguas, I. (2010). Oral computer-mediated interaction between L2 learners: It's about time. *Language Learning & Technology*, *14*(3), 72–93. Retrieved from https://www.lltjournal.org/item/2701

Yi, M. Y., Jackson, J. D., Park, J. S., & Probst, J. C. (2006). Understanding information technology acceptance by individual professionals? Toward an integrative view. *Information & Management*, *43*(3), 350–363. https://doi.org/10.1016/j.im.2005.08.006

Appendix. Alexa Survey

Construct	Item
Perceived ease of use	1A: I think that Alexa was clear and understandable.
	1B: Learning how to use Alexa was easy.
Perceived usefulness	2A: I think that Alexa can improve my English ability.
	2B: I think that Alexa can improve my motivation to use English.
Attitudes towards use	3A: I like using Alexa to learn English.
	3B: I have a positive attitude toward using Alexa.
Behavioral intention	4A: If I had access to an Echo, I would use it to learn English.
	4B: If I owned an Echo, I would be happy to use it to study English.
Individual usage	5: I feel comfortable using Alexa on my own.
	6: I feel comfortable using Alexa with a partner.

About the Author

Gilbert Dizon is a lecturer at Himeji Dokkyo University, Japan and a doctoral student in the Department of Literacy, Culture, and Language Education at Indiana University. His research interests are focused on the use of technology to promote formal and informal second language learning.

E-mail: gilbert.dizon.jr@gmail.com