

Chatbot-Human Interaction and Its Effects on EFL Students' L2 Speaking Performance and Anxiety

Fidel ÇAKMAK

Dr., Alanya Alaaddin Keykubat University, Antalya, TURKEY
ORCID: 0000-0002-3285-7661

Article information

Submission	23/06/2022	Revision received	28/09/2022
Acceptance	11/09/2022	Publication date	24/10/2022

Keywords: Chatbot, Replika, intelligent agents, L2 speaking anxiety, oral communication skills

Abstract: Recently, chatbot interactions have been used for oral communication practice in the field of foreign language education. Some existing studies have highlighted the use of chatbots in relation to specific L2 skills, yet unfortunately, the user experience component of chatbot interaction has not been empirically researched. This study investigates the effect of chatbot-human interaction with the chatbot Replika on L2 speaking performance and speaking anxiety. The participants in the current study were 90 EFL students from a state university in Turkey. They used the application Replika for twelve weeks practicing targeted subjects outside the class. Each student was assessed with the L2 speaking anxiety scale both before and after the intervention. Also, an open-ended questionnaire was administered to collect their perceptions of using chatbots for L2 speaking practice. The interactions with the chatbot were screen taped for a randomly chosen speaking task: Well-being. The findings reveal that there were negative perceptions and attitudes toward the chatbot interaction. Students reported facing difficulties in being understood precisely, which might have contributed to higher anxiety in L2 speaking. Notably, student performance with Replika was significantly better than their face-to-face peer interactions. Overall, the results show that although chatbot interaction is a novel way to provide speaking practice for students, the actual interaction with a chatbot might not be a reliable way to lessen their anxiety with L2 speaking.

Anahtar Sözcükler: Chatbot, Replika, ikinci dil konuşma becerisi, ikinci dil konuşma kaygısı, sözlü iletişim becerileri

Chatbot-İnsan Etkileşimi ve İngilizceyi Yabancı Dil Olarak Öğrenen Öğrencilerin İkinci Dil Konuşma Performansı ve Kaygısı Üzerine Etkisi

Özet: Son zamanlarda yabancı dil eğitimi alanında sözlü iletişim uygulamaları için chatbot etkileşimleri kullanılmaktadır. Bazı çalışmalar, chatbot kullanımını belirli ikinci dil becerileri ile ilişkili olarak vurgulamış, buna karşın chatbot etkileşiminin kullanıcı deneyimi boyutu deneysel olarak araştırılmıştır. Bu çalışma, chatbot Replika ile chatbot-insan etkileşiminin ikinci dil konuşma performansı ve konuşma kaygısı üzerindeki etkisini araştırmaktadır. Türkiye'deki bir devlet üniversitesinden İngilizceyi yabancı dil olarak öğrenen 90 öğrenci bu çalışmaya katılımcı oldu. Bu katılımcılar on iki hafta boyunca belirlenen konularda sınıfdışı konuşma pratiği yapmak için Replika uygulamasını kullandılar. Her öğrenci, uygulamadan önce ve sonra ikinci dil konuşma kaygısı ölçeği ile değerlendirildi. Ayrıca, öğrencilerin ikinci dil konuşma pratiği yaparken chatbot kullanma algılarını toplamak için açık uçlu bir anket uygulandı. Bulgular, chatbot etkileşimine yönelik olumsuz algı ve tutumların olduğunu ortaya koymaktadır. Öğrenciler, tam olarak anlaşılma zorluklarıyla karşılaştıklarını bildirdiler. Bu ikinci dille konuşmada daha yüksek kaygıya neden olmuş olabilir. Dikkat çekici bir sonuç olarak, Replika ile öğrenci performansı, yüz yüze akran etkileşimlerinden önemli ölçüde daha iyiydi. Genel olarak sonuçlar, chatbot etkileşiminin öğrenciler için konuşma pratiği sağlamanın yeni bir yolu olmasına rağmen, bir chatbot ile gerçek etkileşimin, ikinci dil konuşma kaygısını azaltmak için güvenilir bir yol olmayabileceğini göstermektedir.

To Cite This Article: Çakmak, F. (2022). Chatbot-human interaction and its effects on EFL students' L2 speaking performance and speaking anxiety. *Novitas-ROYAL (Research on Youth and Language)*, 16(2), 113–131.

1. Introduction

The prevalence of emerging technology integration in English language teaching has brought many affordances to overcome the limited classroom time available for practicing language knowledge and skills. Although Computer Assisted Language Learning (hereafter CALL) studies have dominated the recent volume of language learning studies and highlighted an improving effect on performances in language learning, the deployment of CALL was found to be limited. Experiments were not longitudinal and dealt with a limited number of language learning aspects (Zhao, 2003). The current focus on informal language learning, concurrent with developments in virtual and digital learning, and the use of mobile-assisted language learning applications, has expanded the scope of CALL to include innovative smart agents and chatbots as tools for language learning in personal spaces. As with most artificial intelligence-based applications, chatbots are recent industrial innovations that afford language learners the opportunity to interact with intelligent machine agents in an educational context. The definition of communication has been expanded to include not only human-to-human interaction but also human-to-computer interaction thanks to a volume of research on artificial intelligence and virtual reality. Additionally, recent trends in pedagogical approaches and curriculum design have highlighted the importance and effectiveness of personalized education adaptable to meet the needs of individual learners. In CALL learning, learners are provided with digital affordances to satisfy their learning needs. In today's world, personalized education has been augmented by novel applications such as virtual reality, AI-driven systems, and wearable computing (Chen, Zou, Xie, & Cheng, 2021). Chatbots have evolved into digital agents that can take part in interactions with human beings as interlocutors in a digital space. In the language learning context, chatbots can provide an opportunity for language learners to personalize their time and language practice through interactions with a virtual interlocutor. How these interactions are stimulated and maintained has attracted the interest of language learning researchers in recent years (Fryer, Coniam, Carpenter, & Lăpușneanu, 2020). They have explored how chatbots can be effectively used in language education with implications and solution suggestions for best practice (Fryer et al., 2020). Researchers have explored the benefits of chatbot-accompanied language practices for language acquisition, such as L2 writing (Kılıçkaya, 2020), L2 grammar teaching (Kim, Cha, & Kim, 2019), and L2 vocabulary learning (Bashori, van Hout, Strik, & Cucchiarini, 2021; Jiang et al., 2022). In these studies, chatbots have been found to be a helpful asset to facilitate learners' L2 development. However, the effect of chatbots on L2 language learning has not yet been researched empirically (Fryer, Nakao, & Thompson, 2019). The affective factors of chatbot interaction have also been studied in relation to chatbots' potential for making human beings (here students) feel more relaxed and engaged during language practice, which can, in turn, improve their motivation in language learning. Complementary research could be redesigned, taking the perceptions and emotions of the L2 users into account to see if there is any empirically positive effect of chatbot interactions in the practice of language skills.

1.1. Chatbot-Human Interaction and L2 Speaking Competence

Eliza was the first prototype pioneered by Weizenbaum (1966) that could interact with human beings with limited language tools such as few question formatted structures. This interaction was necessarily inflexible and led to very limited discussions. Over time other chatbots such as Parry, Jabberjack, and Alice were improved to make use of contextual patterns to create responses and express a limited number of emotions. At the turn of the millennium, chatbot technology evolved to provide practical information such as weather,

news, key dates, etc. Then another phase of interaction emerged with personal voice assistants such as Siri, Alexa, IBM Watson Assistant, Google Now, and Woebot. Most of the interactions between chatbots and humans were in the form of questions or directions from humans to chatbots who would retrieve concrete information. The human interlocutor is generally the one who initiates the interactions by making inquiries. Generally, the personal assistant responds meaningfully with a digital voice responding to the inquiries. The evolution of this interaction has proceeded smoothly from a limited number of pre-programmed responses to meaningful, specific, adaptive, and personalized responses. There is still room for improvement to overcome notable barriers, such as failing to understand heavy accents, user spelling errors, or misunderstanding due to the failure to understand the context of conversation (Adamopoulou & Moussiades, 2020). As the technology has continued to integrate into daily life through PDAs, (e.g., Siri, Alexa), the application of chatbots has been welcomed in the field of education. Over the last decade, there has been a genuine interest in interacting with virtual personas through games and the use of targeted applications for voice chats. However, voice chatbots still struggle with misunderstandings and by consequence off-topic responses. In natural form, depending on the proximity and willingness to communicate between interlocutors, the search for clarification comes through nonverbal as well as verbal communication. Although the nature of human oral interaction involves processing the utterances, it is generally considered that interlocutors communicate most effectively when verbal linguistic input and non-linguistic signaling input are processed simultaneously and contextually. The immediacy of this process is also mimicked to an extent in human-chatbot interactions depending on the smartness of the technology and design parameters. However, it is often observed that chatbots may proceed in conversation through repetitions or fail to respond, yielding sudden and inappropriate breaks in conversation (Adamopoulou & Moussiades, 2020).

L2 speaking competences entail knowledge of language and discourse, as well as communication strategies through which core speaking skills are applied naturally (Burns, 2019). Citing the holistic framework of L2 speaking competence components by Goh and Burns (2012), Burns (2019) revisited the components to clarify the concept of the L2 speaking competence. Knowledge of language and oral discourse entails mastery of the language's sounds, grammar and vocabulary and recognizing how speech is organized socially and pragmatically in the specific culture. Core speaking skills refer to developing the aural skills to process spoken language immediately and the oral skills to produce utterance with fluency. Fluency further requires that one negotiate and organize the oral text while appropriately managing the flow of speech. Effective communication strategies involve developing compensatory cognitive strategies to work around language limitations (e.g., rephrasing, gestures), metacognitive strategies (e.g., thinking and pre-planning of the utterances), and interaction strategies (e.g., asking for confirmation, clarity, repetitions). In teaching L2 speaking competence, these components are taught either through direct instructions with an emphasis on skills and strategies for conversational practice in a controlled way, or indirectly through instructions to practice using the language functionally to communicate meaning (Thomas, 2019).

Chatbots have recently emerged as a tool for L2 language learning (Huang, Hew, & Fryer, 2022). This has provoked renewed interest in the instruction of L2 speaking competence and of how the cycle of teaching L2 speaking could be better implemented when students have the autonomy to engage with available technology. Obviously, teaching the L2 speaking cycle via chatbot is not the same experience as face-to-face with human partners in class. However, such self-regulated practice in the form of conversations with chatbots could afford learners

an additional opportunity to practice speaking skills learned in the classroom context. The utility of this practice will vary with the L2 competence of the learner and the sophistication of the specific chatbot design. It is also useful to note the difference in the learning environment afforded by chatbots in conventional L2/FL settings. The use of chatbots in class might not show the potential of the technology since the classroom time application is generally limited in time and unstructured.

Classroom practice of oral communicative tasks through chatbots takes place in a formal and necessarily constrained environment, including time pressures which may force students to terminate tasks prematurely. The success of accomplishing the task in a set time depends on individual cognitive and affective factors. Chatbot interactions outside the conventional setting by contrast, are less time constrained as the human both initiates and terminates the conversation according to their own time constraints. This affords humans the ability to self-regulate the pace of practice according to their needs at the moment. When it comes to the efficiency of chatbot interaction in terms of L2 language gains in L2 speaking competence, Dizon (2020) demonstrated with an empirical study that chatbot interaction significantly improved L2 speaking proficiency of EFL learners who were in interaction with an intelligent personal assistant, Alexa when compared to a control other with no such interaction. However, the same positive outcome was not observed in both groups' listening performances. This improvement could stem from the ready opportunities that artificially intelligent agents such as chatbots provide for L2 learners to access oral practice whenever and wherever it is wanted. Along with self-paced accessibility, a chatbot providing interactive and personalized feedback also allows the L2 users to receive feedback in real-time. These factors make chatbots a useful tool for L2 oral practice and development. Kim et al. (2019) reviewed a number of chatbots for their effectiveness in terms of foreign language learning. They demonstrated that chatbots have a positive impact on students' oral communication skills by providing L2 language inputs and opportunities to interact and construct meaning.

In a similar vein, Yang, Kim, Lee, and Shin (2022) explored the efficiency of the chatbot, Ellie, for its design and performance as an English virtual interlocutor. They also explored students' perceptions of using Ellie in EFL class and found that the chatbot application encouraged and facilitated L2 learners to make conversation in class successfully. Fryer et al. (2019) uncovered some parameters to consider when applying a virtual partner for conversational practice. They explored the human-chatbot interactions of 122 EFL students in a Japanese higher education context for a semester to investigate their speaking performances, students' task interest, and learning experiences. Interestingly, the findings showed that prior language competence correlated to more interest in chatbot rather than human conversations and that the perceptions of learning with the chatbots were in direct relation to students' task interest. The authors acknowledged the novelty effect as influential as expected for short-term task interest.

How L2 learners benefit from using chatbots for their L2 speaking skill development has been explored in relation to their existing proficiency levels. Kim, Cha, and Kim (2021) found that the chatbots affect the L2 speaking skills of students at two different L2 proficiency levels: low level and intermediate. The pretest and posttest analysis proved that both groups of students yielded statistically significant improvement in English speaking. The two groups did not differ in the scores on pronunciation, yet, there was a marked difference in intonation and observed stress in read-aloud task performances between the two groups. Chen et al. (2021) discussed the potential of Google Assistant from the ELF learners' perspective. The students in the study believed that Google Assistant could be beneficial for improving their

L2 listening and speaking skills. Although perceived improvements were reported, interactional strategies were also analyzed. During the interactions, high proficient students were better at dealing with communication breakdowns. They tended to rephrase the utterances that caused communication breakdowns, whereas low proficiency students repeated the same utterances to repair the communication breakdowns. Mutual comprehensibility was higher between the high proficient students and Google Assistant while low proficiency students were prone to pronunciation errors that caused misunderstandings in communication.

1.2. Chatbot-Human Interaction and L2 Speaking Anxiety

Recent studies on chatbot intervention in language learning have highlighted the specific technology used in chatbots to enable an interactive conversation in L2. The idea of applying artificially intelligent (AI) designs in L2 learning tools sounds appropriate as today's L2 users are presumed to have familiarity with virtual environments. However, this virtual technology does not yet furnish learners with all the necessary conditions for learning. Beyond the technical design parameters, it is expected that, to a certain extent, an AI interlocutor should respond to the needs of L2 learners by recognizing the cognitive, social, and affective factors innate to the process of language learning. Any learning environment is prone to be affected by human psychology, which plays an immense role in language learning (Ortega, 2014). As a psychological construct, speaking anxiety is a factor that greatly impacts L2 learning. It can preempt the process of learning and affect the initiation of interaction, the response quality, and participant engagement in L2 interactions. L2 speaking anxiety manifests itself under the construct of foreign language anxiety. Foreign language anxiety is "a distinct complex construct of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process." (Horwitz, Horwitz, & Cope, 1986, p. 128). L2 speaking anxiety occurs when L2 language learners have the above-mentioned state of mind and feel stressed and anxious while in oral communication with interlocutors. Several studies have demonstrated significant effects of anxiety on L2 language learning and speaking skills (Çağatay, 2015; He, 2013; Öztürk & Gürbüz, 2014; Tercan & Dikilitaş, 2015) and offer solutions from research outcomes. The use of drama techniques (Ataş, 2015; Galante, 2018), scaffolding feedback (Zarei & Rezandoust, 2020), and positive psychology (Oxford, 2016) are some of the recommendations for overcoming foreign language anxiety. The current pedagogical practice recommends applying positive psychology in language classes and acknowledging L2 learners' psychological and social barriers as important elements in mitigating the stress and anxiety of L2 language learning and practice. As presented by Toyama & Yamazaki's (2021) review of classroom interventions and foreign language anxiety, there has been a shift from traditional psychology to positive psychology, which has resulted in bringing focus to individual and interactional interventions in language education. These include self-management and mood-boosting activities, student-student interactions, and interactional teacher-student interventions. Individual intervention encourages L2 learners to navigate their anxiety level through positive self-talk, self-regulated strategies, exploiting gaming, and music opportunities. The study also explicitly highlights the interactional benefits of keeping a dialogue journal, peer review/feedback, participating in video chats, portfolio assessment and such. Technology-driven interventions to overcome speaking anxiety in EFL classes include the use of podcasts (Hamzaoğlu & Koçoğlu, 2016), ASR-based websites (Bashori et al., 2021), Google Assistant (Bao, 2019), and specific mobile applications (Xiangming, Liu, & Zhang, 2020). These studies advocate the use of emerging technology to reduce stress and improve language learning outcomes.

In line with this advocacy, Tai and Chen (2020) explored a two-week intervention using an intelligent personal assistant - Google Assistant, for adolescent EFL learners' oral interaction practice. The study showed that even if it was a short intervention, the IPAs intervention increased student confidence in communication while reducing their speaking anxiety. The authors posit that this positive influence of affective factors on L2 learning might have boosted their L2 language gains. On a similar note, El Shazly (2021) explored the role of chatbots in the speaking practice of 48 EFL learners for eight weeks in both formal and informal settings. He found that students' foreign language anxiety levels were not reduced when interacting with the chatbots. The study discussed the possibility that applying chatbots for speaking practice in the formal academic context might have caused them to feel embarrassed when speaking with the chatbots publicly in the classroom settings. Additionally, the short duration of intervention might have compromised familiarity with the chatbot. This finding was contradictory to the findings of other studies favoring technology-driven systems or tools to combat the L2 speaking anxiety. As seen above, the studies on L2 speaking anxiety have focused on contexts that require face-to-face conversation in physical presence in combination with various other elements such as native or non-native interlocutors, L2 proficiency, and so forth. Given that these studies about reducing L2 speaking anxiety through technological systems or devices have yielded contradictory results, further research is needed to determine if the level of foreign language speaking anxiety makes any difference in speaking performance with human interlocutors and technology-driven interlocutors.

1.3. Chatbot-Human Interaction and L2 Learners' Perceptions

As chatbots have emerged in the business context, the designs of chatbots have become more socially responsive. They have been tailored to establish social relationships and bonds with humans (Ciechanowski, Przegalinska, Magnuski, & Gloor, 2019). This is likely due to customer demands and the desire to increase the effectiveness of the chatbots for messaging. The existing studies on chatbot intervention in L2 language education have not only dealt with the designs and pedagogical implications, but also the language learners' perceptions about the effectiveness of the chatbots. In most of the existing chatbot studies on L2 language learning, the users (L2 learners) reported positive reactions to having utilized the chatbots (Thai & Chen, 2020; Yang et al., 2022). This positivism is thought to have stemmed from the novelty effect (Fryer, Ainley, Thompson, Gibson, & Sherlock, 2017). When chatbots are used in the class or a part of a course shortly, and for the first time, students generally report having enjoyed using the chatbots. Another reason that could contribute to this positivism is that the chatbot could give feedback to human partners in conversations. This can lead the human interlocutor to feel that they are active agents humanizing the chatbot (Adamopoulo & Moussiades, 2019, p. 1). Costa (2018) pointed out that chatbots are no longer considered a technical tool for acquiring specific information, but rather have evolved to act as a virtual buddy that a human partner could feel close to through anthropomorphization. Citing Weber (2005), Costa (2018) supports the idea that anthropomorphization facilitates problem resolution at the social-emotional level, which represents a shift from a coldly rational, human-controlled interaction to a more social and emotional interaction. Human-machine interactions, such as those with chatbots, can evoke in users "the feelings of intimacy, closeness, and empathy" (p. 62). While in an empathetic state, users likely feel less judged when interacting with the chatbot, and are more likely to exercise a willingness to disclose (Lucas, Gratch, King, & Morency, 2014). Alternatively, it may also stem from the feeling that they have improved their L2 language skills (Yang et al., 2022)

The use of chatbots for improving L2 speaking competence can lead to feelings of motivation and personal agency in learning (Moussalli & Cardoso, 2020). In alignment with the studies highlighting positive outcomes of chatbot use from the affective aspect, Moussalli and Cardoso (2016) found that L2 learners enjoyed the interactions and felt comfortable with the interactions. Similarly, another study by Underwood (2017) supported this finding, stating that EFL learners favored artificial intelligence interactions, finding the interaction motivating and enjoyable. Thai and Chen (2020) also explored EFL students' perceptions of Google Assistant for learning English. They articulated that the EFL students had a high level of motivation, engagement, and comfort during their interactions with Google Assistant. Interestingly, unlike Yang et al.'s (2022) study findings, these positive feelings did not reflect actual improvement in English speaking ability. In line with the studies that displayed positive perceptions of the L2 learners while interacting with the chatbots, Chen et al. (2021) explored how EFL students at different proficiency levels perceive the educational potential of Google Assistant. The result showed that students enjoyed interacting with Google Assistant and found it useful for practicing English. Yang et al. (2022) also surveyed the potential of the chatbot, Ellie, in EFL settings and found that most students reported a positive impression of the chatbot. They reported feeling understood by the chatbot and found the interactions fun and interesting. In a prominent study from the Turkish context, Kılıçkaya (2020) found that students using Replika for their writing practice through interactive oral communication also showed enthusiasm for interacting with their virtual partner for L2 practice. Participants reported finding the application interesting and encouraging. He suggested that in addition to writing practice the application should be utilized for other L2 productive skills, notably L2 speaking practice. He also explored the problematic areas of using Replika for written interactions. Among a few deficits, he found that erroneous responses and lack of emotional expressions could be disadvantageous to learners.

Not all the existing studies yielded positive results. Some chatbot studies have reported negative perceptions from the L2 learners' perspective. One factor contributing to negative perceptions could be the language proficiency of learners. Qinghua and Satar (2020) found that in contrast to their low proficiency colleagues, high-level language learners expressed dissatisfaction with chatbots. Another factor could be due to the uncanny valley effect wherein human interlocutors feel a certain degree of discomfort in interacting with a chatbot closely mimicking a human (Ciechanowski et al., 2019). Gallacher, Thompson, and Howarth (2018) explored Japanese EFL learners' perceptions of the chatbot Cleverbot and postulated that learners did not perceive the chatbot to be an amiable replacement for human interlocutors. Unlike human partners, chatbots are not as expressive as humans in showing emotions and unable to negotiate nuanced language, which could affect interaction.

Literature on chatbot use in L2 learning with an emphasis on L2 communication skills has shown that the majority of chatbot studies employ self-reported surveys highlighting their perceived usefulness to claim effectiveness. Therefore, it is necessary to more rigorously explore how chatbots could be integrated into EFL learning processes with an experimental research design to quantify their impact on affective factors such as L2 speaking anxiety and learners' perceptions. This study aims to examine the human-chatbot interaction in terms of humans' speaking skill performance and L2 speaking anxiety when communicating with chatbots. The study also explores student perceptions and emotions over their conversation with the chatbots. The current study seeks to answer these primary questions:

1. Are there any differences in the L2 speaking performance of students when they interact with human beings as peers versus the chatbot Replika?
2. Are there any differences in the anxiety level of L2 speakers before and after interactions with the chatbot Replika?
3. What are the students' perceptions of the chatbot interactions?

2. Method

2.1. Research Design

The current study adopted a mixed methods approach utilizing both quantitative and qualitative data collection instruments before and after the intervention. The pretest and posttest compared participants' scores in the L2 speaking task and their L2 speaking anxiety. Before inferential analyses were conducted, the descriptive statistics of the test scores were obtained, and their distributions were examined for normalcy. The outliers were identified and excluded. Students' background information was analyzed by using descriptive statistics. Students' performances with human peers and chatbots, and students' L2 speaking anxiety were analyzed before and after the intervention by using paired t-tests in order to compare students' scores on two different variables, fluency of conversation and observed speaking anxiety.

2.2. Participants

The study was conducted in the Spring semester of the 2020-2021 academic year in the ELT program at a state university in Turkey. It was conducted throughout the course Oral Communication Skill II, which was provided to freshmen in the spring semester as a follow-up to the course titled Oral Communication Skills I. The students had two hours of course per week. There were 90 ELT freshmen (57 female and 33 male students) who volunteered to participate in the study. The language proficiency test was administered before the intervention (Allan, 2004), and their language proficiency level was found to be B1 at the entry point. They had an age range varying from 18 to 23.

2.3. Data Collection Instruments and Procedure

The study included data from four instruments:

2.3.1. Background questionnaire (BQ)

At the beginning of the intervention, the participants were asked to fill in a short background questionnaire to collect data on the learners' profiles. The questionnaire consisted of seven questions regarding social use of English, and their perceived L2 language learning skills.

2.3.2. State of L2 Learning Survey (SL2LS)

This current study employed a scale (SL2LS in this context) used by Liu & Jackson (2008). The SL2LS had a 5-point response scale ranging from strongly disagree (SD) to strongly agree (SA). The participants used the scale to complete a questionnaire consisting of 57 questions divided among four subsections: the 20-item Unwillingness to Communicate Scale (UCS) developed by Burgoon (1976), the six-item Language Class Risk-Taking (LCR), the four-item Language Class Sociability (LCS) scales designed by Ely (1986), and Foreign Language Classroom Anxiety Scale (FLCAS) developed by Horwitz et al. (1986). Gürsoy and Korkmaz (2018) revised a version of the FLCAS to reflect contextual elements of L2

speaking anxiety specific to EFL learning in the Turkish context. It was this 27-item revised version that was used in this study. As the aim of the study is to explore the effect of L2 speaking anxiety on EFL learners' speaking performances, the revised version of FLCAS (L2 Speaking Anxiety Scale) was used before and after the chatbot intervention. The reliability coefficient of the pilot study in the original version of the L2 Speaking Anxiety Scale was .89 and in the current study .87.

The values for Cronbach's Alpha for the overall SL2LS of 57 items were $\alpha = .735$ before the intervention and $\alpha = .786$ after the intervention, which indicates an acceptable level of internal consistency for the scale.

2.3.3. L2 Speaking Performance Assessment (L2SPA)

Participants were assessed for their L2 oral communication performance by three raters. The raters used a grid comprising five qualitative aspects of spoken language use as defined by the Common European Framework of Reference for Language (CEFR 3.3) (Council of Europe, 2018): Range, accuracy, fluency, interaction, and coherence. Each of these aspects is 20 points maximum with a total possible of 100 points. The mean score of the three raters scores was the final score of the students' L2 speaking performance. There were a few disagreements; however, they were resolved through negotiation among the raters. The grid was used to assess both the L2 speaking performance of participants with human peers (pretest) and with the chatbot Replika (posttest) on the same L2 speaking task.

2.3.4. Perception Questionnaire (PQ)

The participants were asked to complete a revised questionnaire modeled on the study by Nordberg et al. (2019). The questionnaire aimed to quantify the participants' experiences with the chat session. It consisted of nine questions, five of which used a scale for rating some features of the chatbot, ranking from 1 (negative) to 7 (positive). The other four questions required open-ended, free-text responses.

The data for the current research was collected during the Spring Semester. The participants were registered for a departmental course Oral Communication Skills II. The course includes a variety of communication-oriented activities to improve conversational strategies through discussions, debates, persuasive presentations, and interviews. It is a compulsory course for the Spring Semester in the ELT programs where freshmen focus on L2 speaking practice in their first year. Preparation for the chatbot intervention for the course was arranged after all the students had taken Oral Communication Skill 1 in the previous semester. That course gave them a common background of communicative oral practice working in groups and pairs. Then, at the very beginning of the Spring semester, before the intervention, the students were made aware of the application Replika which could be used for additional opportunities for their speaking practice outside the class. Ninety students voluntarily subscribed to the program; presumably, they were interested in using it to augment their practice. The intervention started in the 2nd week and ended in the 14th week of the Spring Semester. The instructor assigned the students to carry out a conversational task in pairs and talked about a randomly selected topic (expression of being) chosen from the topics existing in Replika. For this task, a prompt (an outline) was provided to the students to improvise their conversation. It was prepared similarly to an outline in Replika that was relevant to the topic selected. In the 3rd week, students were given approximately 7 to 10 minutes to briefly discuss well-being in pairs. The same conversation topic was given to the students as an assignment to be completed with the chatbot. In the 11th week, they were asked to begin

saving their communications with Replika. The students were at liberty to use the chatbot Replika at their convenience to complete assigned tasks from the 11th week up to the 13th week. After they had completed the tasks, they shared the link with the instructor for assessment. The researcher collected all the videos of students' chats with Replika and assessed them with the other two assessors. The interval between two speaking tasks on the same topic was offset sufficiently to avoid any practice effects. Finally, in the 14th week, the students were given the perception questionnaire and asked to share their review of the interactions with the chatbot. The complete implementation plan is given below in Table 1.

Table 1.

Implementation Plan

Weeks	Research Action	Data Collection
1	Intro to the course informing Ss about the intervention	BQ, SL2LS, L2SPA
2	Intro to Replika app	–
3	Pair work assignment: Well-being	Pretest
4-11	In-class activities	–
11-13	Replika assignment: Well-being	Posttest
14	Sharing chatbot experiences with peers	PQ

2.4. Data Analysis

The analysis of four instruments used in the study was computed quantitatively with SPSS 23. Only two questions (#10 and #11) in the Perception Questionnaire were qualitatively assessed following the thematic analysis (Braun & Clark, 2006). In order to ensure the accuracy in the responses of the participants, expert reviews on the themes were cross-checked (Creswell & Clark, 2018) as follows: A sample of 80% of the data was coded by an L2 language instructor independently. Inter-rater reliability was then calculated in percentage and was found to be high (92%). Additionally, two language instructors working in a similar EFL context examined the themes that the researcher and the L2 language instructor came up with from the participants' responses. Minor disagreements were resolved in conferences.

3. Findings

Students' background information was analyzed to see what personal elements the participants brought in L2 language learning before the intervention. Table 2 shows that almost half of the students do not have English-speaking friends in their social circles. They were mostly positive about their reading and speaking L2 language skills; however, they perceived their listening and writing skills as relatively poor and lower than the other two skills.

Table 2.

Descriptive Statistics of Participants' L2 Language Background

Profile Info	n	%
I have English-speaking friends		
Yes	47	52.8
No	42	47.2
Reading Ability		
Poor	2	2.2
Okay	14	15.7
Good	63	70.8
Very Good	10	11.2
Listening Ability		
Poor	49	55.1
Okay	40	44.9
Good	–	–
Very Good	–	–
Writing Ability		
Poor	31	34.8
Okay	21	23.6
Good	30	33.7
Very Good	7	7.9
Speaking Ability		
Poor	6	6.7
Okay	33	37.1
Good	41	46.1
Very Good	9	10.1
Overall English Proficiency		
Poor	3	3.4
Okay	26	29.2
Good	48	53.9
Very Good	12	13.5
Total	89	100

To address the first research question: Are there any differences in L2 speaking performance of students when they interact with human beings as peers versus the chatbot Replika? The students' L2 speaking performances scores were analyzed. As illustrated in Table 3, a paired t-test was applied to analyze the differences between students' L2 speaking performance with peers in pair work and their performances with a chatbot.

Table 3.

Results of the Paired-t Test (L2 Speaking Performance)

Group	n	Mean	SD	t	df	p
SP-Chatbot	89	74.30	8.61	5.947	88	0.000
SP-Peer	89	69.70	11.09			

Table 3 shows that there was a significant difference between the posttest mean score (74.30) and the pretest mean score (69.70) of the speaking performance test scores ($t(88) = 5.947$, $p < .05$).

To address the second research question: Are there any differences in anxiety level of L2 speakers before and after interactions with the chatbot Replika? Students' anxiety levels before and after the chatbot interaction were analyzed. A paired t-test was applied to analyze the differences between pretest and posttest speaking anxiety test scores.

Table 4.

Results of the Paired-t Test (L2 Speaking Anxiety)

Group	n	Mean	SD	t	df	p
SA Post	89	3.27	.205	10.636	88	0.000
SA Pre	89	2.95	.147			

Table 4 shows that there was a significant difference between the posttest mean score (3.27) and the pretest mean score (2.95) of the speaking anxiety test scores ($t(88) = 10.636$, $p < .05$).

To answer research question 3: What are the students' perceptions of the chatbot interactions? Five questions were analyzed quantitatively while four questions were analyzed qualitatively. The items and mean scores are given below. As seen from the first part of Table 5, students were positive about their ability to express themselves accurately; however, they reported negative impressions of Replika and the likelihood of using Replika in the future. Overall, students reported above-average interactions and experiences with Replika. The second part of Table 5 shows students' feeling mostly anxious and awkward when interacting with Replika.

Table 5.

Students' Perceptions of the Chatbot Interactions

Part 1: Quantitative section	Item	Mean
	• What was your impression of Replika?	2.48
	• How would you rate your interactions with Replika for the conversation?	3.89
	• How useful did you experience that Replika was for the conversation?	3.78
	• To which extent did you experience getting to say what you wanted to?	6.05
	• How likely will you use Replika for your L2 speaking practice?	2.19
Part 2: Qualitative section	Positive (f)	Negative (f)
	feeling amused (14)	fearful (13)
	being funny (16)	annoyed (17)
	joyful (21)	feeling absurd (22)
	excited (24)	feeling awkward (38)
		being anxious (43)

4. Discussion and Conclusion

The current study examined the human-chatbot interaction in terms of the humans' speaking skill performance and experience of L2 speaking anxiety when communicating with the chatbot. The study also briefly explored student perceptions and emotions over their conversations with the chatbot. The results suggest that students' speaking performance was significantly better with the chatbot. This was in line with the other studies which reported similar findings (Kim et al., 2019; Yang et al., 2022). The reason for this outperformance could have been due to the fact that conversational chatbot designs have a limited but corrective range of responses when compared with human interlocutors and thus can be easier to manage with regards to the flow of conversation. EFL students' communication skills could have been well-practiced as the chatbot provided predictable responses giving the student an extended opportunity to negotiate meaning (Kim et al., 2019). The enhanced performance could also be due to task interest and task familiarity that students already experienced success in L2 speaking practice with the chatbot. As seen from the demographic info, the majority of students perceived they had good L2 speaking skills, and half of them had English-speaking friends. This could have created a positive atmosphere conducive to good performance. Also, the audio/visual affordances of chatbot could have been meaningful in interactive oral language practice. Students might also have considered the chatbot partner to be a native speaker, a virtual buddy providing exposure to more authentically accented English. That might have enhanced students' task interest leading to more focused performances during the interactions with the chatbot. Although the enhanced task interest could have been a novelty effect, the first-time intervention might have had a resilient positive impact on learning outcomes. Further observation would be required to confirm that. Also, during the conversation, the Politeness Effect might have played a role in focusing attention and produced better learning outcomes for students when they were conversationally communicating with the chatbot rather than communicating through written text (Johnson & Wang, 2010).

Students' performances were significantly better with the chatbot than with their human counterparts. This was to be expected since novice students with low or medium language proficiency might have advanced more on their L2 oral communication skills through repeated practice with a virtual partner rather than a less predictable one-time interaction with a human partner in a controlled speaking task in class. This expected benefit from the chatbot by novice learners with low language proficiency was also observed by Qinghua & Satar (2022) and Wang et al. (2008). These findings suggest that novice background EFL students can further L2 communication skills and performances through interactions with chatbots.

The second finding of this study observed that students' L2 speaking anxiety was actually higher after the intervention. This was an interesting occurrence since participants performed well with the chatbot. The expectation for lower anxiety is based on the results of earlier studies (e.g., Alemi, Meghdari, & Ghazisaedy, 2015; Bao, 2019; Fryer & Carpenter, 2006) where participants were observed to have a reduced anxiety purportedly as the application created positive affective factors. However, the result of the current study is in line with the study by El Shazly (2021), which discusses the unexplored possibility of increased anxiety playing a facilitative role in making the most of students' cognitive and linguistic potential. It is also possible to say that the novelty effect of L2 oral practice with a chatbot, and frustration with inevitable communication breakdowns as well as a flatter than expected effect of the chatbot might have increased the anxiety level during the chatbot interaction.

As to the third finding, students reported a negative perception of the chatbot as an English conversation partner. Although they perceived they had done well to express what they meant and that the chatbot helped them improve their speaking performance, their overall impression was negative. This was a surprise in the study and conflicts with the study by Yang et al. (2022). It does support the claims about the limitations of chatbot interactions regarding responsiveness and coherence (Bibauw, François, & Desmet, 2022). Off-topic or incoherent parts of the chatbot interaction could have negatively colored the perceptions of EFL students in the study context. While the current findings conflict with the study by Kılıçkaya (2020) and Bao (2019), it underscores the importance of feeling understood properly in a meaningful interaction. While chatbots afford opportunities for convenient L2 practice, the quality of those interactions is important. A positive affective state for L2 learners, such as that produced by a healthy conversation (with a minimum of breakdowns and misunderstandings), feeling understood, and being able to understand the interlocutor is important to managing L2 speaking anxiety. Overcoming this anxiety could be possible with more fully developed and advanced systems of dialogue-based conversational chatbots.

The current study presents an empirical study of chatbot-human interaction compared with human interaction, and the effects of chatbot interactions on learners' L2 speaking anxiety. It also touches on learners' perceptions of using chatbots for conversational practice. The study found that there was a significant difference between the posttest mean score and the pretest mean score of the L2 speaking performance test scores and L2 Speaking Anxiety scores. The participants expressed mostly negative impressions about using Replika as a virtual communication partner. This came as a surprise since participants performed significantly better with chatbot than with human interlocutors in the speaking tasks. It was interesting that the chatbot seemed to provoke L2 speaking anxiety. This suggests that, contrary to intent, the application did not provide a non-threatening environment for L2 users to reduce their L2 speaking anxiety. The students' negative impressions about using the chatbot are, presumably, due to linguistic factors (e.g., failures of communication, miscommunication, and so forth) that were apparent in the chatbot interaction. These failures can be reduced with improved chatbot design. A major implication of these findings which could enlighten language teachers is that the smart opportunities for practice afforded by technological advancement are not always favored by students. When they plan to bring such technological opportunity in their classes, it would be wise to foresee fading task interest and to use such interventions in a considered and limited fashion. Another implication is that chatbots specifically could be better suited to extracurricular practice for novice students with limited exposure opportunities to improve L2 speaking practice, rather than as an in-class intervention. Overcoming L2 speaking anxiety requires successful interactions over time. It should be noted that no single intervention is a panacea for L2 learning and chatbots, while potentially useful may be best viewed as one of many tools available for L2 instruction. They can provide access to authentic accents and be programmed to provide targeted exposure to targeted elements of language like vocabulary and syntax structures. This, however, is just a form of practice. Successful L2 mastery with chatbots still requires meaningful interactions over time in both receptive and expressive aspects of the target language. Chatbots show potential as a tool to that end, but one which should be used with mindful intent.

One of the limitations of the present study is that it was not possible to track students' use of Replika in their informal settings. The current study was set up to observe how they interacted with the chatbot at their convenience for the randomly assigned task only. Tracking students' whole interaction with the chatbot was never intended out of respect for

their self-regulated mode of L2 oral practice. Since the whole interaction was not trackable, how often they were using Replika before they rated their perceptions and shared them with other students and the researcher was unknown. Future studies could be set up to explore student perceptions over a longer period and to see if perceptions vary according to task. Alternatively, another study could be set up to factor in more personal background elements such as personality, language proficiency and aptitude to relate it to the differences in L2 speaking performances both live and with chatbots.

Statements on Open Data, Ethics and Conflict of Interest

The data used in this research contain assessment scores. The author is, therefore, unable to make this data publicly available. The participants' personal information was kept private and protected during the research process. The participation was voluntary, and the consent forms were obtained. Some of the data were presented in the 8th Annual AZCALL Conference, held online on October 11–16, 2021. There is no conflict of interest existing in work reported above.

Acknowledgement

The author thanks the referees for the careful and insightful review of the manuscript. She also expresses her special thanks and profound gratitude to Madlen Kamhi for inspiring her to conduct the current research and supporting her in her research pursuits.

References

- Adamopoulou, E., Moussiades, L. (2020). An overview of chatbot technology. In I. Maglogiannis, L. Iliadis & E. Pimenidis (Eds.), *IFIP Advances in information and communication technology*, 584 (pp. 373–383). Springer. https://doi.org/10.1007/978-3-030-49186-4_31
- Alemi, M., Meghdari, A., & Ghazisaedy, M. (2015). The impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. *International Journal of Social Robotics*, 7(4), 523–535. <https://doi.org/10.1007/s12369-015-0286-y>
- Allan, D. (2004). *Oxford placement test*. Oxford: Oxford University Press.
- Ataş, M. (2015). The reduction of speaking anxiety in EFL learners through drama techniques. *Procedia- Social and Behavioral Sciences*, 176, 961–969. <https://doi.org/10.1016/j.sbspro.2015.01.565>
- Bao, M. (2019). Can home use of speech-enabled artificial intelligence mitigate foreign language anxiety: Investigation of a concept. *Arab World English Journal*, 5, 28–40. <https://dx.doi.org/10.24093/awej/call5.3>
- Bashori, M., van Hout, R., Strik, H., & Cucchiari, C. (2021). Effects of ASR-based websites on EFL learners' vocabulary, speaking anxiety, and language enjoyment. *System*, 99. <https://doi.org/10.1016/j.system.2021.102496>
- Bibauw, S., François, T., & Desmet, P. (2022). Dialogue systems for language learning: Chatbots and beyond. In N. Ziegler & M. González-Lloret (Eds.), *The Routledge handbook of second language acquisition and technology* (pp. 121–134). Abingdon, Oxon: Routledge. <https://doi.org/10.4324/9781351117586-12>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

- Burgoon, J. K. (1976). The unwillingness-to communicate scale: Development and validation. *Communication Monographs*, 43, 60–69. <https://doi.org/10.1080/03637757609375916>
- Burns, A. (2019). Concepts for teaching speaking in the English language classroom. *Learn Journal: Language Education and Acquisition Research Network*, 2(1), 1–11. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1225673.pdf>
- Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92, 539–548. <https://doi.org/10.1016/j.future.2018.01.055>
- Costa, P. (2018). Conversing with personal digital assistants: On gender and artificial intelligence. *Journal of Science and Technology of the Arts*, 10(3), 59–72. <https://doi.org/10.7559/citarj.v10i3.563>
- Council of Europe. (2018). Qualitative aspects of spoken language use - Table 3 (CEFR 3.3): Common reference levels. Retrieved from <https://www.coe.int/en/web/common-european-framework-reference-languages/table-3-cefr-3.3-common-reference-levels-qualitative-aspects-of-spoken-language-use>
- Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning: Topic modeling and knowledge mapping. *Educational Technology & Society*, 24(1), 205–222. <https://www.jstor.org/stable/26977868>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE.
- Çağatay, S. (2015). Examining EFL students' foreign language speaking anxiety: The case at a Turkish state university. *Procedia - Social and Behavioral Sciences*, 199, 648–656. <https://doi.org/10.1016/j.sbspro.2015.07.594>
- Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*, 24(1), 16–26. Retrieved from <https://www.lltjournal.org/item/10125-44705/>
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), 1–15. <https://doi.org/10.1111/exsy.12667>
- Ely, C. M. (1986). An analysis of discomfort, risk-taking, sociability, and motivation in the L2 classroom. *Language Learning*, 36, 1–25. <https://doi.org/10.1111/j.1467-1770.1986.tb00366.x>
- Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, 10(3), 8–14. Retrieved from <https://www.learntechlib.org/p/74498/>
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. *Computers in Human Behavior*, 75, 461–468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, 279–289. <https://doi.org/10.1016/j.chb.2018.12.023>
- Fryer, L. K., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, 24(2), 8–22. Retrieved from <https://www.lltjournal.org/item/10125-44719/>
- Galante, A. (2018). Drama for L2 speaking and language anxiety: Evidence from Brazilian EFL Learners. *RELC Journal*, 49(3), 273–289. <https://doi.org/10.1177/0033688217746205>

- Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. New York: Cambridge University Press.
- 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). <https://doi.org/10.14705/rpnet.2018.26.815>
- Gürsoy, E., & Korkmaz, H. (2018). Speaking anxiety of freshmen and senior prospective ELT teachers. *Eğitimde Kuram ve Uygulama*, 14(1), 48–60. <https://doi.org/10.17244/eku.346886>
- Hamzaoğlu, H., & Koçoğlu, Z. (2016). The application of podcasting as an instructional tool to improve Turkish EFL learners' speaking anxiety. *Educational Media International*, 53, 313–326. <https://doi.org/10.1080/09523987.2016.1254889>
- He, D. (2013). What makes learners anxious while speaking English?: A comparative study of the perceptions held by university students and teachers in China. *Educational Studies*, 39(3), 338–350. <https://doi.org/10.1080/03055698.2013.764819>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <http://dx.doi.org/10.2307/327317>
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Jiang, M. Y., Jong, M. S., Wu, N., Shen, B., Chai, C. S., Lau, W. W., & Huang, B. (2022). Integrating automatic speech recognition technology into vocabulary learning in a flipped English class for Chinese college students. *Frontiers in Psychology*, 13, 902429. <https://doi.org/10.3389/fpsyg.2022.902429>
- Johnson, W. L., & Wang, N. (2010). The role of politeness in interactive educational software for language tutoring. In C. C. Hayes & C. A. Miller (Eds.), *Human-computer etiquette: Cultural expectations and the design implications they place on computers and technology* (pp. 91–113). New York: Auerbach Publications. <https://doi.org/10.1201/b10420>
- Kılıçkaya, F. (2020). Using a chatbot, Replika, to practice writing through conversations in L2 English: A case study. In M. Kruk & M. Peterson (Eds.), *New technological applications for foreign and second language learning and teaching* (pp. 221–238). New York: IGI Global.
- Kim, N. Y., Cha, Y., & Kim, H.-S. (2019). Future English learning: Chatbots and artificial intelligence. *Multimedia Assisted Language Learning*, 22(3), 32–53. <https://doi.org/10.15702/mall.2019.22.3.32>
- Kim, H.-S., Cha, Y. & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korean Journal of English Language and Linguistics*, 21, 712–734. <https://doi.org/10.15738/kjell.21..202108.712>
- Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal*, 92(1), 72–86. <https://doi.org/10.1111/j.1540-4781.2008.00687.x>
- Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94–100. <https://doi.org/10.1016/j.chb.2014.04.043>
- 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. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, 33(8), 865–890. <https://doi.org/10.1080/09588221.2019.1595664>
- Nordberg, O. E., Wake, J. D., Nordby, E. S., Flobak, E., Nordgreen, T., Mukhiya, S. K., & Guribye, F. (2019). Designing chatbots for guiding online peer support conversations for adults with ADHD. In *Chatbot research and design– Third International Workshop, CONVERSATIONS 2019* (pp. 113–126). Cham, Switzerland: Springer.
- Ortega, L. (2014). *Understanding second language acquisition*. New York, NY: Routledge.
- Oxford, R. L. (2016). Powerfully positive: Searching for a model of language learner well-being. In D. Gabryś-Barker & D. Galajda (Eds.), *Positive psychology perspectives on foreign language learning and teaching* (pp. 21–37). <https://doi.org/10.1007/978-3-319-32954-3>
- Öztürk, G., & Gürbüz, N. (2014). Speaking anxiety among Turkish EFL learners: The case at a state university. *Journal of Language and Linguistic Studies*, 10(1), 1–17. Retrieved from <https://dergipark.org.tr/en/pub/jlls/issue/9938/122913>
- Qinghua, Y., & Satar, M. (2020). English as a foreign language learner interaction with chatbots: Negotiation for meaning. *International Online Journal of Education and Teaching (IOJET)*, 7(2), 390–410. Retrieved from <http://iojet.org/index.php/IOJET/article/view/707>
- Tai, T. Y., & Chen, H. H. J. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 1–18. <https://doi.org/10.1080/10494820.2020.1841801>
- Tercan, G., & Dikilitaş, K. (2015). EFL students' speaking anxiety: a case from tertiary level students. *ELT Research Journal*, 4(1), 16–27. Retrieved from <https://dergipark.org.tr/en/download/article-file/63653>
- Thomas, N. (2019). Teaching L2 speaking: Recommending a holistic approach. *rEFLlections: A Journal of Language Education*, 26(1), 134–145. <https://files.eric.ed.gov/fulltext/EJ1271172.pdf>
- Toyama, M., & Yamazaki, Y. (2021). Classroom interventions and foreign language anxiety: A systematic review with narrative approach. *Frontiers in Psychology*, 12, 614184. <https://doi.org/10.3389/fpsyg.2021.614184>
- Underwood, J. (2017). Exploring AI language assistants with primary EFL students. In K. Borthwick, L. Bradley, & S. Thouëсны (Eds.), *Short papers from EUROCALL 2017* (pp. 317–321). Southampton, United Kingdom: Research-Publishing.net.
- Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human Computer Studies*, 66, 96–112. <https://doi.org/10.1016/j.ijhcs.2007.09.003>
- Weber, J. (2005). Helpless machines and true loving care givers: A feminist critique of recent trends in human-robot interaction. *Journal of Information, Communication and Ethics in Society*, 3, 309–218. <https://doi.org/10.1108/14779960580000274>
- Weizenbaum, J. (1966). ELIZA-A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Xiangming, L., Liu, M., & Zhang, C. (2020). Technological impact on language anxiety dynamic. *Computer & Education*, 150, 1–12. <https://doi.org/10.1016/j.compedu.2020.103839>
- Yang, H., Kim, H., Lee, J., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL FirstView*, 1–17. <https://doi.org/10.1017/S0958344022000039>

- Zarei, A. A., & Rezaoust, H. (2020). The effects of scaffolded and unscaffolded feedback on speaking anxiety and self-efficacy. *Journal of Modern Research in English Language Studies*, 7(4), 111–132. <https://doi.org/10.30479/JMRELS.2020.13464.1655>
- Zhao, Y. (2003). Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO Journal*, 21(1), 7–27. Retrieved from <http://www.jstor.org/stable/24149478>