



# European Journal of Educational Research

Volume 13, Issue 3, 1199 - 1213.

ISSN: 2165-8714

<http://www.eu-jer.com/>

## Design and Implementation of an Educational App as a Methodology to Improve Speaking Skills in EFL Students at B1 Level: A Case Study

Nelly Padilla\* 

Escuela Superior Politécnica de Chimborazo, ECUADOR

Yajaira Padilla 

Escuela Superior Politécnica de Chimborazo, ECUADOR

José Luis Andrade 

Escuela Superior Politécnica de Chimborazo, ECUADOR

Lorena Yumi 

Independent Researcher, ECUADOR

Received: June 29, 2023 • Revised: August 24, 2023 • Accepted: January 26, 2024

**Abstract:** The present study aimed to improve the speaking skills of university students at the B1 level who presented limitations in their oral competence. An educational methodology based on designing and implementing an application adapted to the Common European Framework of Reference was developed and applied to boost language performance. A case study was used to conduct the two stages of this research; the former had to do with a control group where intervention was carried out using non-probabilistic sampling with students of the Computing Faculty; a pretest was applied to test the knowledge acquired in their classroom sessions during the first quarter in 2020. The second process was tracking an experimental group, which was assessed after implementing the developed methodology using the app "4SkillsWeb". A posttest was used to evidence learners' progress during the COVID-19 lockdown, and the results showed improved oral competence in aspects such as grammar and vocabulary, discourse management, pronunciation, and interactive communication, with about 95% confidence in its validation. A qualitative-quantitative methodology was used to determine the influence of the English app. A t-students test was implemented to corroborate the data analysis taken by both groups through SOFTWARE JMP v 11.0.0G.

**Keywords:** *4SkillsWeb app, educational methodology, oral competence development.*

**To cite this article:** Padilla, N., Padilla, Y., Andrade, J. L., & Yumi, L. (2024). Design and implementation of an educational app as a methodology to improve speaking skills in EFL students at B1 level: A case study. *European Journal of Educational Research*, 13(3), 1199-1213. <https://doi.org/10.12973/eu-jer.13.3.1199>

### Introduction

Many mobile applications are downloaded every year, and educational apps are top-rated. For instance, Mauroux et al. (2014) said Duolingo became the most popular learning and education app, with about 98 million downloads worldwide. With the increasing demand of users, mobile technologies have become very popular in the educational field, especially in higher education. The advent of these technologies has provided educators with new platforms to engage students with class materials and promote more effective interactions Luna-Nevarez and McGovern (2018). The accessibility to technological devices among students means educators adapt to the new educational and learning trends that fit better with their interests. The immediateness and availability of information in educational apps make interaction and response more effective regardless of the surrounding conditions. In a study by Doardi (2021), the student's ability to access general university information and class content through mobile technology was examined, and the results determined that the communication experience had improved significantly by the end of the study.

Implementing new technologies in the educational field aims to expand the scope of skills students from this generation own as digital natives. Agati (2012) considers that Millennials born after 1982 have made technological devices the center of their daily activities.

With the advent of the COVID-19 pandemic, all the concepts about the use of new technologies took more relevance when students worldwide were forced to transition to online education, and with this, the emergence of a variety of tools to make the learning process as effective as possible regardless the adverse circumstances. García del Dujo et al. (2021) remark on the importance of educational apps and their role in the different measures and methods to achieve quality education during the lockdown since they consider the features of these apps allow participants to create a related structure among students, parents, and teachers.

\* Corresponding author:

Nelly Padilla, Escuela Superior Politécnica de Chimborazo, Ecuador. ✉ [npadilla@epoch.edu.ec](mailto:npadilla@epoch.edu.ec)

Educators have been pursuing to generate better results with the resources available daily. Regarding the language learning field, researchers have been conducting different studies on using apps and devices and their influence on improving oral skills. Välimaa and Hoffman (2008) conducted a study about the society discourse, making inferences about the effectiveness of using mobile apps for educational purposes, including EFL vocabulary learning, and the results showed that they could positively impact learning outcomes. Even though there are apps like Duolingo, which have been working on a large scale and are remarkably recognized, it is worth mentioning local works that have contributed to determining the importance of mobile apps. A study about the pedagogical use of WhatsApp to develop speaking skills determined the positive impact of using this app in the development of speaking skills in university students regarding communication, content, method, operability, attractiveness, and satisfaction.

As a result of the recurrent accountability education worldwide faces in terms of optimal outcomes, permanent updates in new technologies and adaptability are significant concerns among educators. The undeniable insertion of technology in all fields has led to the appearance of many independent learners, optimizing the advantages of the digital era when it comes to reducing time and working remotely. Educators must raise awareness about adapting their teaching methods for more autonomous pupils. There is an essential discussion about autonomy's implementation in language learning and education in general. Beyond the misconceptions about this topic, Benson (2011) remarks that the more autonomous the learners are, the more responsible and critical community members they become.

The challenge for educators, therefore, is to create the conditions and opportunities for learners to exercise control over their learning, so how practice is organized will directly affect the development of appropriate autonomous learning habits. In this sense, Computer Assisted Language Learning (CALL) has been a matter of study over the last few years for its constructivist approach that engages learners in their education process. According to Riveros and Mendoza (2008), Computer Assisted Language Learning provides stress-free environments for learners. It makes them more responsible by providing an independent learning condition using a computer or any other device. Internet access (Chilingaryan & Zvereva, 2017) and electronic devices become the most potent tools because English learners can have the opportunity to download apps (López-Leyva et al., 2020) to pursue independent practice. The present research study turns its attention to the use of an English-speaking application, 4SkillsWeb, where 115 learners from higher education levels were able to practice by themselves at anytime, anywhere, developing their autonomous learning (Beenen & Arbaugh, 2019) during the Covid-19 lockdown.

This research aims to answer the question: Does implementing the 4SkillsWeb app improve students' oral competence at the B1 level at ESPOCH? A review of existing literature to support the content of our research work in app development and its correlation with oral development in higher education was carried out. Then, the methodology used is presented, and the design of the app "4SkillsWeb" and its implementation as a methodological strategy to enhance students' performance are described. Next, the results with the corresponding analysis of the data collected from the control and experimental groups are presented to conclude with a discussion of these findings and the limitations of this study.

### Literature Review

Over the last few years, much research has been conducted about the shift education has made from teacher-centered classes to mobile learning environments. The objective of incorporating new methodologies in the field of education deals with the idea of making teaching and learning more concrete, tangible, attractive, interesting, practical, and meaningful; therefore, the urge to introduce appropriate resources to foster work is necessary in order to generate better retention in both in and out-of-school scenarios (Hussain et al., 2020), as well as boosting communicative competences necessary not only in educational environments but also in everyday situations.

#### *Importance of Orality in Language Learning*

"Communication occurs where there is speech" (Kieffer & Lesaux, 2008). Based on this statement, the author remarks on the importance of speaking as a tool for interaction among individuals to reach the smooth running of any system. Brown and Abeywickrama (2003) consider that students are judged mainly by how they communicate in real-life situations; therefore, developing meaningful and practical skills to fit those communication needs in terms of orality is essential.

There are important implications regarding orality and literacy when it has to do with language teaching and language acquisition, especially second-language acquisition (Hansen-Strain, L. 1989); while orality targets the interaction between communicator and audience aimed to share knowledge, literate tradition emphasizes the communicative function of language. Tannen (1982) summarizes this difference: Orality focuses on involvement, and literacy focuses on content. For Sarramona López (2009), orality has to do with speech, while literacy has to do with writing. From a realistic perspective, the traditional educational system in developing countries has been mainly focused on literacy, teaching preset contents, and persistently concentrating on reading and writing. At the same time, most ESL/EFL teachers continue using memorization and drills to teach speaking, which have been proven ineffective in developing proper orality when interacting in real-life situations.

Orality, referring to the acquisition of a second language, is a complex process where several factors intervene. It is stated that "The ability to speak involves not only knowledge of the characteristics of the language but also the ability to process

information and the use of language in context." (Dueñas Macias et al., 2015). In this sense, developing oral skills must involve a broad spectrum of resources that ensure adequate internalization. One of the most effective methods for this to happen has to do with interaction since speaking skills have sometimes been undervalued by simply considering it as a popular form of expression

However, speaking involves other aspects requiring attention, such as awareness and recognition of sounds, the relationship between letters and sounds, vocabulary, spelling, and comprehension in native and second languages (Kieffer & Lesaux, 2008). An appropriate interaction that meets these essential requirements will promote the proper development of the speaking skills of the participants of this study (Berns et al., 2016). For its achievement, the use of current, innovative, accessible, and essential resources that involve technological approaches is necessary (García del Dujo et al., 2021).

A shift to a more accurate teaching-learning process is necessary in our societies where speaking has been undervalued, even though employability depends more on communication than technology. According to Henriques et al. (2022), Ecuador is in the top two worst countries in Latin America regarding English language proficiency, with only Mexico scoring more poorly than Ecuador; in this criterion, it is imperative to understand the impact of English proficiency in our country; Yoon et al., (2021) remarked in his research that the analysis of the effect of English proficiency in overseas countries could create a moderating effect. In the study of the efficacy of English-speaking development (Henriques et al., 2022) in Ecuador, the English teaching model must be motivated by innovation and improvement of educational sources. Under these circumstances, comparing how speaking apps are developed worldwide, and their efficiency is significant. Terhorst asserted that access to apps in the learning process must be secure and have high standards of quality (Terhorst et al., 2021); in this scenery, learning apps can be well designed according to their content, and activities and avoid any possible failures (Kacetyl & Klímová, 2019).

### *The Rise of Educational Apps*

The way materials and information are presented in class has changed dramatically in recent decades. The increasing popularity of educational apps is evidenced worldwide by technological advances, having them become quite popular as a viable alternative to traditional classroom teaching, especially during the COVID-19 pandemic. Educational apps provide a personalized learning experience, which is crucial for effective learning, as they incorporate elements like multimedia, gamification, quizzes, and interactive exercises that enhance the learning experience. Regarding language learning, auxiliary media have always been used as valuable and central elements in the classroom (O'Donoghue et al., 2001), with greater or lesser success depending on the use given by the teacher, helping to enhance the learning of a new language (Mauroux et al., 2014), explicitly considering drawings, photos, images, and slides as visual aids, as well as recordings, cassettes, videos, and films, as an aid to develop the sense of hearing. Nevertheless, the advance of multimedia and the insertion of more sophisticated programming features have made it possible to reach levels of functionality, such as identifying graphical representations of speech patterns through voice recognition and detecting pronunciation details, which attempts to get more accuracy in oral performance.

Educational apps in language learning allow students to have more agency and ownership in their learning experience, which can lead to improved engagement and academic progress along with other benefits, such as effective time utilization and lower cost of operations. However, other considerations must be deliberated; for instance, the main focus should be on conceptualizing how language learning could be enhanced in new, innovative ways with the assistance of mobile devices (Godwin-Jones, 2011).

### *Technological Approaches to Language Acquisition*

It is undeniable that information and communication technologies have taken on an essential role in the teaching of foreign languages (Khorechko et al., 2015). Thus, teaching English as a foreign language is not indifferent to this new educational reality, where teachers and researchers have made all the necessary efforts to integrate computer advances in the different educational environments of the English language (Sarramona López, 2009). It is easy to see how this new technological reality is embodied in creating and developing mobile applications that help both directed and semi-directed learning and autonomous learning (Ahmadi, 2018).

One common approach that has been widely studied is the insertion of computer-assisted language learning (CALL) into class planning. It is not new, and its implementation has been mentioned previously in this paper; with this, Computer-assisted second language research (CASLR) has been brought up to the stage in order to study the effects of the learners'-computer programs interaction through tasks which are a standard part of instruction, assessment, or even research (Chapelle, 2007). SLA research uses such tasks to operationalize learning conditions and infer learners' knowledge and strategies.

Together with the abovementioned approach, the emergence of artificial intelligence (AI) has become significantly relevant in the language learning field. The appearance of a vast number of tools developed with AI has enabled teachers and students to access a significant number of applications and resources that are designed by utilizing machine learning

and natural language processing and were used to identify errors, provide feedback, and assess language abilities (Woo & Choi, 2021)

Indeed, teaching and learning processes are permanently being updated according to the new resources in the classroom (Doardi, 2021); at this point, teachers involve learners of English as a foreign language (EFL) to apply authentic material in the class (Shih et al., 2015). In this sense, the design of 4skills web app as a methodological approach for speaking skills improvement represents an innovative teaching tool to boost autonomous learning, breaking the repetitive direct mode of learning style. (Bouckaert, 2016).

## Methodology

### *Research Design*

The research design corresponded to a quantitative approach because it established a relationship between two variables (Kieffer & Lesaux, 2008), the data were quantified and analyzed through statistical methods to test the working hypotheses (Krapkov et al., 1997)

### *Working Methodology*

As a quasi-experimental study, the experimental and control groups were not randomly assigned, which was the missing criterion for it to be experimental. The research was based on comparing the relationship between the study variables and their analysis in different groups of students (Case group and Control group).

The level corresponded to relational research since it identifies the association between the two variables, the independent (4SkillsWeb application) and the dependent (Speaking), to explore how they interact and influence each other. These relationships were quantified and tested against the research hypotheses.

The independent variable (4SkillsWeb) was included in the experimental group but not included in the control group to allow comparison and analysis according to the statistical method. The equivalence of the groups was guaranteed, i.e., they were homogeneous groups with similar characteristics.

A pretest (before the intervention) and a posttest (at the end of the intervention) were applied. The test was designed to evaluate the application's pedagogical criteria, and the results obtained after using the 4skillweb application were compared.

The methodology, as well as the activities and resources included in this application, were created by the researchers through templates based on the PET exam (Preliminary English Test) (Moniuszko & Sciuk, 2013) and its corresponding format, later, it was programmed with an intuitive and user-friendly interface. In the application, the Model-View-View Model paradigm for mobile applications was implemented to use the Databinding class (Cabrero García & Richart Martínez, 1996), which allows direct data links with the graphical interface, thus not casting the components at the logical level. Subsequently, the data was uploaded with the respective layout of the activities, and an APK-type executable was generated. The programmers carried out Test activities and modifications to finally load the application in the Google Play Store (Dalvi et al., 2022), which is available for installation only for students with institutional email.

In this way, the researchers carried out the intervention using the 4SkillsWeb application on the students' cell phones during their classes for an entire academic period.

It allowed to reinforce the learning of the English language in higher education students (Kacetl & Klímová, 2019) and develop skills in the oral production of the language through the use of this application developed with the help of ESPOCH teachers and students called 4SkillsWeb.

### *Research Field*

It was framed in the Socio-educational field since, through a systematic process, the authors' experience was combined with the reasoning and analysis of the results of the work to give scientific answers to the research question posed related to the level of influence of a web application as a tool for the development of oral competence in the teaching process of learning English.

### *Sample and Data Collection*

The data collection process was deliberately done by the teachers who took part in the study; the control and case groups were students of the Higher Polytechnic School of Chimborazo, majoring in different schools of the Faculty of Computing Science. From the 115 participants, 57 were set for the control group and 58 for the experimental group. Pretests were applied to both groups; these tests were designed based on the parameters established by the Common European Framework of Reference for Languages for level B1 of competence. The control group completed the course with no changes in the methodology applied during class sessions. On the other hand, the experimental group had direct instruction and follow-up assessment by implementing the 4SkillsWeb app, which was explicitly designed using the

methodology proposed for this study. At the end of the intervention stage, a posttest was applied to the control and experimental groups to verify the outcomes.

### *Analyzing of Data*

The data for this work were collected based on the criteria for the B1 level of performance of the CEFR. The parameters included in the tests correspond to the speaking part of the PET (Preliminary English Test) of the Cambridge Assessment, and it is displayed as follows:

- *Part 1: Personal information.*
- *Part 2: Describing a picture.*
- *Part 3: Discuss a topic.*
- *Part 4: Conversation about the topic.*

For each part of the test, several descriptors were assessed on a scale of 0-5, with 0 being the lowest and 5 being the highest score, which means that every part would fit a scale within 15 to 20 points. The descriptors aim to assess competence in the following aspects: Pronunciation, Grammar and Vocabulary, Interactive Communication, and Discourse Management. Hence, the descriptors for each section are grouped as follows:

### *Pronunciation*

1. Is the answer clear? Can the speaker be generally understood?
2. Is the speaker's intonation generally appropriate?
3. Does the speaker use sentence stress generally correctly? Is word stress correct?
4. Are individual sounds generally clear? Are they correctly produced?

### *Grammar and Vocabulary*

1. Does the speaker use simple grammatical forms with control?
2. Does the speaker attempt to use complex grammatical forms?
3. Does the speaker use a range of appropriate vocabulary to talk about familiar topics?

### *Interactive Communication*

1. Does the speaker start discussions? Does the speaker introduce new ideas?
2. Does the speaker react appropriately to what the interlocutor or other candidate says?
3. Does the speaker keep the interaction going? Does the speaker say more than the minimum? Does the speaker involve the other candidate?
4. Does the speaker try to move the interaction in an appropriate direction? ('develop the interaction and negotiate towards an outcome') Does the speaker need support?

### *Discourse Management*

1. Are the answers of an appropriate length for the task? Is there much hesitation?
2. Are the contributions relevant? Is there much repetition of ideas?
3. Does the speaker organize their contributions and use cohesive devices? Is there a range?

Cronbach's Alpha coefficient measured the reliability of an instrument's measurement scale; a result of .95 was considered very reliable or excellent.

The results were grouped into four main categories for better organization. Once the pretest and posttest results were gathered, they were compared and analyzed using a t-test through the SOFTWARE JMP v 11.0.0G for further discussion.

A  $p$ -value  $< .05$  was considered a statistically significant difference, and a  $p$ -value  $< .01$  was considered highly significant.

The Kolmogorov-Smirnov test was used for the normality test since the sample consisted of more than 30 subjects.

The effect size was measured with Cohen's  $d$ , and according to the classification, the result was considerable.

### Findings/Results

For the current study, four main criteria were the object of study, displayed in four sections or question types, which were arranged according to the parameters set in the CEFR for the level of performance B1; these criteria are Pronunciation, Grammar, and Vocabulary, Discourse Management, and Interactive Communication. The question types matched the speaking section of the Cambridge Preliminary English Test (PET), designed for assessing level B1 of performance. These criteria were statistically compared through the data gathered from pre and posttests.

#### Pretest Results

When analyzing the results of the pretest in the experimental and control groups, it was found that in all criteria, the mean is higher in the control group than in the experimental group. Nevertheless, when applying the student t-test, these differences were found to be statistically significant only in the Pronunciation and Interactive Communication criteria ( $p < .05$ ). At the same time, Grammar, Vocabulary, and Discourse management are not statistically significant ( $p > .05$ ). (See Table 1)

Table 1. Comparative analysis (means) Pretest Control vs. Experimental Group

| Section/question type        | Criteria                  | Group              | N     | Mean  | SD*    | p       |
|------------------------------|---------------------------|--------------------|-------|-------|--------|---------|
| Personal information         | Pronunciation             | Control group      | 57    | 13.81 | 3.68   | .001**  |
|                              |                           | Experimental group | 58    | 11.40 | 4.08   |         |
| Describe a picture           | Grammar & Vocabulary      | Control group      | 57    | 9.91  | 2.91   | .352    |
|                              |                           | Experimental group | 58    | 9.34  | 3.57   |         |
|                              | Pronunciation             | Control group      | 57    | 13.86 | 3.42   | .009**  |
|                              |                           | Experimental group | 58    | 11.81 | 4.75   |         |
| Discuss a topic              | Grammar & Vocabulary      | Control group      | 57    | 9.96  | 2.63   | .099    |
|                              |                           | Experimental group | 58    | 8.97  | 3.71   |         |
|                              | Pronunciation             | Control group      | 57    | 13.89 | 3.96   | .026**  |
|                              |                           | Experimental group | 58    | 12.03 | 4.85   |         |
| Conversation about the topic | Discourse Management      | Control group      | 57    | 8.86  | 2.61   | .806    |
|                              |                           | Experimental group | 58    | 8.72  | 3.25   |         |
|                              | Interactive communication | Control group      | 57    | 12.70 | 3.71   | .247    |
|                              |                           | Experimental group | 58    | 11.79 | 4.61   |         |
|                              | Pronunciation             | Control group      | 57    | 14.05 | 3.57   | <.001** |
|                              |                           | Experimental group | 58    | 11.00 | 4.26   |         |
| Interactive communication    | Control group             | 57                 | 12.61 | 3.64  | .012** |         |
| Experimental group           | 58                        | 10.64              | 4.61  |       |        |         |

\* Standard deviation; \*\* statistically significant

The results for each criterion were analyzed independently in each section of the test, and the findings are shown as follows:

#### Pronunciation

The expected outcome for this criterion at the B1 level is to assess an appropriate degree of intelligibility and some control in phonological features at both utterances and word levels. For assessment purposes, some specific performance descriptors were used throughout the execution of the four parts of the speaking test. For pronunciation criteria, four descriptors were graded out of 5 points each. As a result, 20 points would be obtained from the following descriptors: (a) Are the answers clear? Can the speaker be generally understood? (b) Is the speaker's intonation generally appropriate? (c) Does the speaker use sentence stress generally correctly? Is word stress correct? (d) Are individual sounds generally clear? Are they correctly produced?

For section 1 (Personal information), the results obtained after comparing the performance of the control and experimental groups were statistically significant ( $p < .05$ ). The students of the control group reached a mean of 13.81 points out of 20. In comparison, the mean for the experimental group students was 11.40 points out of 20. The standard deviation for the control group was 3.68, and the experimental group's was 4.08. The results obtained by the control group in the pretest were 2.41 points above the experimental group.

For section 2 (Describe a picture), the results coming from the comparison of the control and experimental groups were statistically significant ( $p < .05$ ). The mean for the control group was 13.86 points, while for the experimental group, it was 11.81. The standard deviation for the former group was 3.42, and for the latter was 4.72. Similarly, the results obtained by the control group in the pretest were 2.05 points above the experimental group.

The results maintained the same trend for section 3 (Discuss a topic). The results were statistically significant ( $p < .05$ ). The mean for the control group was 13.89 points, while for the experimental group, it was 12.03. The standard deviation

for the control group was 3.96, and the experimental group's was 4.85. For this question, the results obtained by the control group in the pretest were 1.86 points above the experimental group.

Finally, for section 4 (Conversation about a topic), the results obtained for this criterion between the control and experimental group were statistically significant as well ( $p < .05$ ). The mean for the control group was 14.05 points, and for the experimental group, 11 points. The standard deviation was 3.57 for the control group and 4.26 for the experimental group. The results obtained by the control group in the pretest were 3.05 points above the experimental group.

### *Grammar and Vocabulary*

Grammar and vocabulary criteria assess the competence of participants to produce simple grammatical forms and their attempt at complex grammatical forms, as well as using a range of appropriate vocabulary to give and exchange information on familiar topics. The performance descriptors to assess these criteria are: (a) Does the speaker use simple grammatical forms with control? (b) Does the speaker attempt to use complex grammatical forms? (c) Does the speaker use a range of appropriate vocabulary to talk about familiar topics? Each descriptor was graded out of 5 points, which makes up a total of 15 points.

Grammar and vocabulary were assessed in 2 sections of the speaking test, and the results are shown as follows:

Section 1, corresponding to the "Personal information question," the results obtained after comparing the performance of both control and experimental groups were not statistically significant ( $p > .05$ ). The mean obtained by the control group was 9.91 with a standard deviation of 2.91, while the experimental group's mean was 9.34 with a standard deviation of 3.57. The results obtained by the control group were 0.57 points above the experimental group's results.

For section 2, corresponding to the question "Describe a picture," the results obtained after comparing the performance of both control and experimental groups were not statistically significant ( $p > .05$ ). The mean obtained by the control group was 9.96, with a standard deviation of 2.63. On the other hand, the mean of the experimental group was 8.97, with a standard deviation of 3.71. For this question, the control group was 0.99 points above the experimental group's results.

### *Interactive Communication*

Aspects such as the ability to initiate and respond to conversations were considered to assess this criterion. Additionally, participants were assessed on their ability to maintain and develop interactions and negotiate toward an outcome with little support. This section was graded out of 20 points, and the performance descriptors were worth 5 points each. They are as follows: (a) Does the speaker start discussions? Does the speaker introduce new ideas? (b) Does the speaker react appropriately to what the interlocutor or other candidate says? (c) Does the speaker keep the interaction going? Does the speaker say more than the minimum? Does the speaker involve the other candidate? (d) Does the speaker try to move the interaction appropriately? ('develop the interaction and negotiate towards an outcome') Does the speaker need support?

This criterion was assessed in two sections of the test, and the results are described below:

For section 3 (Discuss a topic), the criterion Interactive communication showed that the comparison of the results of the control and the experimental group was not statistically significant ( $p > .05$ ). The mean obtained by the control group was 12.70, while the experimental group's mean was 11.79. The standard deviation for the control and the experimental groups was 3.71 and 4.61, respectively. The results of the control group's performance are 0.91 points above the experimental group's results.

For section 4 (Conversation about a topic), the results turned out to be statistically significant ( $p < .05$ ). The mean for the control group was 12.61, and for the experimental group was 10.64. The standard deviation was 3.64 for the control group and 4.61 for the experimental group. The difference in results between the control and the experimental group was 1.97 points.

### *Discourse Management*

In this criterion, participants were assessed on their ability to produce extended stretches of language despite some hesitation and their contributions to achieving communication using appropriate, cohesive devices. This section was graded out of 15 points, and the performance descriptors used to assess this criterion were: (a) Are the answers appropriate for the task? Is there much hesitation? (b) Are the contributions relevant? Is there much repetition of ideas? (c) Does the speaker organize their contributions and use cohesive devices? Is there a range?

This criterion was assessed in section 3 of the test (Discuss a topic), where the comparison of the results of the control and experimental groups turned out not to be statistically significant ( $p > .05$ ). The mean for the control group was 8.86 points, and for the experimental group, it was 8.72 points. The standard deviation was 2.61 for the control group and 3.25 for the experimental group. The difference in results between the control and the experimental group is 0.14 points.

### Posttest Results

The results obtained from the posttest applied to the experimental and control groups show a significant variation from the ones gathered in the pretest. The mean for all criteria was higher in the experimental group. Additionally, there was a statistically significant difference in all the criteria assessed in the four test sections with a  $p$ -value  $<.05$ . The information can be seen in Table 2.

Table 2. Comparative analysis (means) Posttest Control vs. Experimental Group

| Section/question type      | Criteria                  | Group              | N    | Mean | SD*          | p            |
|----------------------------|---------------------------|--------------------|------|------|--------------|--------------|
| Personal information       | Pronunciation             | Control group      | 57   | 6.95 | 2.59         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 16.3 | 2.45         |              |
|                            | Grammar & Vocabulary      | Control group      | 57   | 4.88 | 1.77         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 11.8 | 1.98         |              |
| Describe a picture         | Pronunciation             | Control group      | 57   | 7.53 | 2.62         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 16.8 | 2.09         |              |
|                            | Grammar & Vocabulary      | Control group      | 57   | 5.00 | 1.78         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 11.8 | 1.74         |              |
| Discuss a topic            | Pronunciation             | Control group      | 57   | 7.18 | 2.52         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 16.4 | 2.06         |              |
|                            | Discourse Management      | Control group      | 57   | 5.14 | 1.65         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 11.4 | 1.97         |              |
| Interactive communication  | Control group             | 57                 | 6.88 | 2.59 | $<.001^{**}$ |              |
|                            | Experimental group        | 58                 | 15.5 | 2.64 |              |              |
| Conversation about a topic | Pronunciation             | Control group      | 57   | 6.91 | 2.41         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 16.5 | 2.05         |              |
|                            | Interactive communication | Control group      | 57   | 6.75 | 2.24         | $<.001^{**}$ |
|                            |                           | Experimental group | 58   | 15.4 | 2.49         |              |

\* Standard deviation; \*\* statistically significant

The analysis for these results was made for each criterion of the test. For the four sections included in the speaking part, the results are the following:

#### Pronunciation

For the section Personal information, the results were statistically significant ( $p < 0.05$ ) after comparing the performance of the control and experimental groups. The mean for the control group students was 6.95 points, while the mean for the experimental group students was 16.3 points out of 20. The standard deviation for the control group was 2.59, and the experimental group's was 2.45. The results obtained by the experimental group in the posttest were 9.35 points above the control group.

For the section Describe a picture, the results from comparing the control and experimental groups were statistically significant ( $p < 0.05$ ). The mean for the control group was 16.8 points, while for the experimental group, it was 7.53. The standard deviation for the control group was 2.62, and for the experimental group was 2.09. The results obtained by the experimental group in the posttest were 9.27 points above the control group.

For the section Discuss a topic, the results were also statistically significant ( $p < 0.05$ ) when comparing both groups. The mean for the control group was 7.18 points, while for the experimental group was 16.4 points. The standard deviation for the control group was 2.52, and for the experimental group was 2.06. For this section, the results obtained by the experimental group in the posttest were 9.22 points above the control group.

Finally, for the section Conversation about a topic, the results obtained from comparing the control and experimental groups were similarly statistically significant ( $p < 0.05$ ). The control group had a mean of 6.91 points, and the experimental group had 16.5 points. The standard deviation was 2.41 for the control group and 2.05 for the experimental group. The results obtained by the experimental group in the posttest were 9.59 points above the control group.

#### Grammar and Vocabulary

For section 1, the results obtained after comparing the performance of the control and experimental groups were statistically significant ( $p > 0.05$ ). The mean obtained by the control group was 4.88, while the experimental group's mean was 11.8. The standard deviation was 1.77 for the control group and 1.98 for the experimental group. The results obtained by the experimental group were 6.92 points above the control group's.

For section 2, the comparison of the control and experimental groups was statistically significant ( $p > 0.05$ ). The mean obtained by the control group was 5.00 with a standard deviation of 1.78, while the mean for the experimental group was



11.8 with a standard deviation of 1.74. For this question, the results obtained by the experimental group were 6.8 points above the control group's.

#### *Interactive Communication*

For section 3, the performance comparison of the results of the control and the experimental group was statistically significant ( $p > .05$ ). The mean obtained by the control group was 6.88 points, while the experimental group's mean was 15.5 points. The standard deviation for the control group was 2.59, and for the experimental groups was 2.64. The results of the experimental group's performance were 8.62 points above the results of the control group.

For section 4, the results obtained of the performance of the control and experimental groups were statistically significant ( $p < .05$ ). The mean for the control group was 6.75 points, and for the experimental group was 15.4 points. The standard deviation was 2.24 for the control group and 2.49 for the experimental group. The difference in results between the control and the experimental group was 8.65 points.

#### *Discourse Management*

Section 3: The comparison results for the performance of the control and experimental groups were statistically significant ( $p > .05$ ). The mean for the control group was 5.14 points, and for the experimental group, it was 11.4 points. The standard deviation was 1.65 for the control group and 1.97 for the experimental group. The experimental group's results were 6.26 points above the control group's.

Due to the significant difference between the control group's Pre and Posttest results, a One-factor variance analysis was carried out to contrast the variations in the four criteria assessed during the two research periods.

*Table 3. One-Factor Variance Analysis: Pronunciation.*

| Groups               | Count          | Sum                | Mean         | Variance |            |         |
|----------------------|----------------|--------------------|--------------|----------|------------|---------|
| Pre                  | 57             | 792.5              | 13.90        | 10.79    |            |         |
| Post                 | 57             | 407                | 7.14         | 5.50     |            |         |
| VARIANCE ANALYSIS    |                |                    |              |          |            |         |
| Source of variations | Sum of squares | Degrees of freedom | Mean squares | F        | p          | F value |
| Between groups       | 1303.60        | 1                  | 1303.60      | 159.98   | $p < .001$ | 3.93    |
| Within groups        | 912.60         | 112                | 8.15         |          |            |         |
| Total                | 2216.20        | 113                |              |          |            |         |

*Table 4. One-Factor Variance Analysis: Grammar and Vocabulary.*

| Groups               | Count          | Sum                | Mean         | Variance |            |         |
|----------------------|----------------|--------------------|--------------|----------|------------|---------|
| Pre                  | 57             | 566.5              | 9.94         | 7.15     |            |         |
| Post                 | 57             | 281.5              | 4.94         | 2.88     |            |         |
| VARIANCE ANALYSIS    |                |                    |              |          |            |         |
| Source of variations | Sum of squares | Degrees of freedom | Mean squares | F        | p          | F value |
| Between groups       | 712.5          | 1                  | 712.5        | 142.10   | $p < .001$ | 3.93    |
| Within groups        | 561.57         | 112                | 5.01         |          |            |         |
| Total                | 1274.07        | 113                |              |          |            |         |

*Table 5. One-Factor Variance Analysis: Interactive Communication.*

| Groups               | Count          | Sum                | Mean         | Variance |            |         |
|----------------------|----------------|--------------------|--------------|----------|------------|---------|
| Pre                  | 57             | 721.5              | 12.68        | 11.92    |            |         |
| Post                 | 57             | 388.5              | 6.82         | 5.49     |            |         |
| VARIANCE ANALYSIS    |                |                    |              |          |            |         |
| Source of variations | Sum of squares | Degrees of freedom | Mean squares | F        | p          | F value |
| Between groups       | 972.71         | 1                  | 972.71       | 111.75   | $p < .001$ | 3.93    |
| Within groups        | 974.89         | 112                | 8.70         |          |            |         |
| Total                | 1947.61        | 113                |              |          |            |         |

*Table 6. One-Factor Variance Analysis: Discourse Management.*

| Groups | Count | Sum | Mean | Variance |
|--------|-------|-----|------|----------|
| Pre    | 57    | 505 | 8.86 | 6.80     |
| Post   | 57    | 392 | 6.88 | 6.72     |

Table 6. Continued

| VARIANCE ANALYSIS    |                |                    |              |       |            |         |
|----------------------|----------------|--------------------|--------------|-------|------------|---------|
| Source of variations | Sum of squares | Degrees of freedom | Mean squares | F     | p          | F value |
| Between groups       | 112.01         | 1                  | 112.01       | 16.57 | $p < .001$ | 3.93    |
| Within groups        | 757.02         | 112                | 6.76         |       |            |         |
| Total                | 869.03         | 113                |              |       |            |         |

### Discussion

The focus of this study is on the results obtained in the two phases of the research: First, the pretest was carried out in classroom environments to all the participants using a traditional approach to English teaching, and second, the posttest was carried out during confinement after implementing the methodology proposed through the use of the 4SkillsWeb app to the experimental group. The following discussion will be held for the control and experimental groups regarding the criteria of pronunciation, grammar and vocabulary, interactive communication, and discourse management.

#### Pronunciation Criterion

The assessment of this criterion is relevant to the general outcomes of the study, as it is found in all sections of the test. As shown in Table 1, the results obtained in the pretest do not differ significantly between the control and experimental groups. In this first section, the control group participants scored slightly higher than the experimental group students in every test section. The average performance for the control group for this criterion was 13.7 points, equivalent to 68.5%, which is considerably acceptable in terms of achievement. On the other hand, the average performance for the experimental group reached 11.6 points, equivalent to 58% of achievement. The pretest results showed that the control group had a better performance in terms of pronunciation (2.1 points) than the experimental group. Both groups obtained the highest scores in the questions that involved interaction. However, the overall performance of both study groups needed to be improved to reach the B1 level of proficiency.

On the other hand, the results obtained for this criterion in the posttest brought exciting findings to the study. At this stage, the app 4SkillsWeb was designed and implemented as an educational methodology to foster speaking competence in the experimental group. The results at this stage differ significantly from the ones obtained in the pretest. Hence, for the Pronunciation criterion, the control group obtained an average score of 7.14 points out of 20, which represents 35.7% of performance achievement. In comparison, the experimental group reached an average score of 16.5 points out of 20, representing 82.5% of performance achievement.

Related studies, such as those by Calvo Benzie (2017), showed similar results. The implementation of technologies for the development of pronunciation was of great benefit to improve this skill. In her study, the use of ICTs combined with an appropriate methodology meant a potential way to develop motivation and engagement to work not only on controlled activities but also to promote autonomy in their work; the use of technology also influenced the choices of students when it comes to pronunciation; in general terms, the students in the research linked the strategies used through the use of technology to their language learning outcomes.

Although studies have evidence of the positive influence of using TICs, some factors can partially affect the achievement of learning goals since students showed a loss of motivation and engagement in using electronic devices and technologies along the study. This idea matches the statement of Bennett et al. (2008), who argue that students prefer technology for personal purposes to educational issues. Besides intrinsic factors, the present study also showed the influence of extrinsic considerations to be taken into account. For instance, the results from the pretest revealed that the level of interaction decreased during remote classes due to COVID-19, affecting their performance in terms of one-to-one communication. Breaux et al. (2023) confirm this statement in a study on isolation on adolescent social functioning.

As for this study, the concern was the effectiveness of implementing the proposed methodology. It is worth mentioning that the general overview for this criterion in the posttest showed that the experimental group did better than the control group. With a difference of 9.36 points and 82.5% global achievement, it reaches level B1 of performance according to the assessment scale of the CEFR for the PET test.

#### Grammar and Vocabulary Criterion

The difference between the two groups in the Personal Information section of the pretest was 0.57 points. In performance achievement, the two study groups had a minor difference of 0.7 points. However, neither the control nor the experimental group reached the level B1 of performance.

In the posttest, however, there is an essential shift in the results for the study groups. The experimental group reached an average performance of 11.8 points out of 15 for the Personal Information and Describe a Picture section. In comparison, the control group reached 4.8 and 5 points for the Personal Information and Describing Picture sections correspondingly. Consequently, the experimental group reached 78.7% of performance achievement in this criterion,

while the control group only obtained 32% of performance achievement. Moreover, during the posttest, the experimental group improved by 2.6 points compared to the pretest. Based on this information, the experimental group reached level B1 of performance after the educational intervention.

These results agree with the statement that educational apps improve grammar and vocabulary. Guaqueta and Castro-Garces (2018) evidenced similar results in a study where several mobile apps were used to enhance student performance. In their study, aspects like the behavioral change towards technology, the level of confidence and assurance, the immediate feedback, and the motivation involved, among others, were evidenced when using apps to practice grammar and vocabulary during their ESL process.

#### *Interactive Communication Criterion*

This criterion was assessed in two test sections (Discuss a topic and Conversation about a topic). Likewise, the results for this criterion were very similar for both groups. In section three, there was a slight difference of 0.91 points where the control group performed better than the experimental group. In section four, the difference between both groups was 1.97 points, with the control group doing better again. The average performance for the control group was 12.6 points out of 20, representing 63% of achievement. The experimental group averaged 11.2 points, equivalent to 56% of achievement. Neither the control nor the experimental group obtained the necessary punctuation to reach the B1 level of performance.

The posttest revealed a dramatic change in results. The general average of performance for the control group for this criterion went from 12.6 points in the pretest to 6.8 points in the posttest. On the contrary, the experimental group improved the average performance, going from 11.2 points in the pretest to 15.4 points in the posttest. The performance achievement for the control group was 34%, while for the experimental group, it was 77%. According to these results, after the posttest, the experimental group reached the B1 level of performance, while the control group did not.

#### *Discourse Management*

This criterion was assessed in the section Discuss a topic. The difference in means for this section was 0.14 points, which shows that both study groups were at almost the same level of performance. These figures represented 59% and 58.1% of achievement for the control and experimental groups, respectively. However, neither of the groups obtained the required punctuation to reach level B1 of performance.

The posttest results evidenced a big difference in performance between the two groups. The control achieved 34.2% of performance, while the experimental group obtained 76%. The results showed that after the intervention, the experimental group reached performance level B1.

In general terms, there was a drastic change in the results during the two moments of the research. Although the level of performance was quite similar for the pretest for the two groups, the results needed to be better to meet the goal of this research. Nevertheless, the posttest results evidenced that the implementation of technology through apps can positively affect the improvement of oral skills. Studies like the one conducted by Deng and Trainin (2015) support that mobile apps can be used to achieve learning outcomes, especially in the field of language, as they increase students' motivation to produce extended written and spoken texts, especially in circumstances with limited interaction, like the worldwide pandemic of COVID-19 (Gael & Elmiana, 2021). In this sense, implementing the mobile app 4SkillsWeb during the intervention stage allowed students to practice the required skills to achieve the linguistic competence required for all the criteria during the study (Berns et al., 2016).

The present study suggests that integrating educational apps can create a comprehensive methodology, as shown in the contributions of Kacatl and Klímová (2019). This study is crucial as it emphasizes that this approach complements a better learning process. At the same time, the app designed for this project offers practical exercises and personalized feedback, creating a holistic learning experience.

Regardless of the positive findings related to implementing the educational methodology, which was the object of this study, it would be worth analyzing the results obtained for the control group. According to the ANOVA test carried out for one-factor analysis, the performance of the control group decreased dramatically, and the reasons why this phenomenon occurred may be related to several factors linked to the pandemic. Firstly, the level of interaction with the control group was limited to remote sessions, where the amount of participation dropped dramatically due to the features of this modality. Students were not fully engaged with classes, and the adaptation process did not occur overnight, especially with those students who did not have full access to technological facilities. Secondly, during the study, the number of hours provided for English sessions was reduced to half, which affected the amount of exposure to the subject and, hence, the amount of practice time. Thirdly, for the control group, significant changes in the strategies were not adapted to class sessions. Also, the methodologies and resources were kept the same as when they were working on face-to-face classes. There was a clear disadvantage, considering the adverse conditions the group had to work in. Finally, several studies reveal the pandemic's psychological impact on the educational field. Loss of motivation, some mental health issues such as Major Depressive Disorder (MDD), and Generalized Anxiety Disorder (GAD) were

some of the factors that adversely affected students during online classes (Irawan et al., 2020)

### **Conclusions**

The dramatic change education suffered during the second quarter of the 2020s led to important considerations regarding how educators deal with technological issues. On the one hand, the imminent switch from face-to-face to virtual education made technology essential for achieving the expected goals at each stage. Thus, according to the results obtained in this study, it is clear that the lack of an appropriate methodology that keeps students focused on the target outcomes and, overall, keeps them motivated can have profound effects when their general level of performance decreases from one modality to another. Based on the results obtained in the four criteria assessed throughout this study, it was evidenced that the implementation of a new methodology addressed to develop independent study and permanent practice in the participants of the experimental group had a positive impact firstly on the way they embraced their educational process by the time they improved their linguistic competences. The chances of practice increased since the participants kept working on developing their speaking skills through the methodology designed and applied for specific purposes in the mobile app. Regardless of the limited physical contact and peer interaction while working in online classes, the students could carry out activities designed to pursue a B1 level of competence, according to CEFR. Pronunciation activities were implemented throughout the app to boost phonological accuracy and intelligibility, positively resulting in the estimated outcomes.

On the other hand, it is necessary to raise awareness of the possible effects of the transition to virtual education in the educational field. Even though technology has become an essential tool to boost skills and abilities, the lack of interaction among participants affected other areas, such as their social skills. In this sense, it was evident that students outside the experimentation process significantly decreased their level of oral competence. Then, it is proved that through adequately purposed tools and a guided methodology, it is possible to achieve positive outcomes regardless of the constraints external circumstances can lead to. The ultimate purpose of education should be the permanent innovation that allows individuals to face adverse situations and turn them into opportunities to function better in society.

### **Recommendations**

The findings of the present work evidenced the importance of adaptation to new environments. In a world where trends catch the attention of more and more people every time, it is essential to consider the remarkable advance and scope technology has reached in all fields, and education must be open to these changes. Today, many teachers are turning their classes into more technological spaces to achieve better results as students feel more engaged when they work with means they are familiar with. In some societies, it can still generate some resistance; as Tannen (1982) mentioned in a study on teachers' fear of technology, even experienced educators can feel reluctant to use it and reject the use of available technologies in their classrooms. Nevertheless, educators must overcome fears and implement technology into class sessions and activities. As demonstrated during the present work, using resources designed and adapted to students' needs brings positive outcomes when used correctly.

Regarding this issue, it is worth mentioning the appropriate use of methodologies; an accurate diagnosis of students' needs is essential to design and stick to an approach to meet the aimed outcomes. In this case, the development of an educational app was necessary based on the reality that everyone had to adapt due to the unforeseen circumstances of the COVID-19 pandemic. The activities included in the app were addressed to develop students' independence and confidence as they advanced at their own pace. It was essential to reinforce their motivation when achieving their educational goals and improving their oral competence. Thus, permanent training in technological issues and methodologies is compulsory for teachers willing to keep updated with the current educational and learning trends. Finally, more studies should be conducted to determine which strategies, approaches, and methodologies are more effective when it comes to language learning, especially with students of current generations who are primarily immersed in technological issues and whose interests and learning styles can be somewhat different to the ones teachers expect them to have.

Lastly, the posttest results evidenced a dramatic decrease in the performance of the control group students. The factors influencing this phenomenon are essential for future research.

### **Limitations**

Among the limitations this work dealt with, the most relevant was providing person-to-person feedback. Restrictions due to the sanitary crisis worldwide forced participants to switch from classroom sessions to virtual environments, which, at some point, turned out to be the appropriate scenario to test the proposed methodology since it is directly related to the assessment of mobile learning as a way to improve oral performance. Nonetheless, the sudden change in educational conditions profoundly impacted the perception of communication during online class sessions. In this sense, even when students who participated in the study improved their oral performance, the virtual conditions limited the participants to have a more effective interaction among partners and teachers. Additionally, academic dishonesty in virtual scenarios became more popular, which was a significant concern among teachers who participated in the study.

Regarding this fact, participants do better in natural environments where the evaluation process can be more realistic and efficient.

### Authorship Contribution Statement

N. Padilla: Conceptualization, data analysis, and interpretation. Y. Padilla Y: Drafting, data acquisition. Andrade: Conceptualization, drafting, and design. Yumi: Critical revision, technical support supervision

### References

- Agati, H. A. (2012). The millennial generation: Howe and Strauss disputed—the College of William and Mary.
- Ahmadi, M. R. (2018). The use of technology in English Language Learning: a literature review. *International Journal of Research in English Education*, 3(2), 115-125. <https://doi.org/10.29252/ijree.3.2.115>
- Beenen, G., & Arbaugh, B. (2019). Flipping class: Why student expectations and person-situation fit matter. *International Journal of Management Education*, 17(3), Article 100311. <https://doi.org/10.1016/j.ijme.2019.100311>
- Bennett, S., Maton, K., & Kervin, L. (2008). The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, 39(5), 775-786. <https://doi.org/10.1111/j.1467-8535.2007.00793.x>
- Benson, P. (2011). *Teaching and researching: autonomy in language learning* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315833767>
- Berns, A., Isla-Montes, J.-L., Palomo-Duarte, M., & Doderó, J.-M. (2016). Motivation, students' needs, and learning outcomes: a hybrid game-based app for enhanced language learning. *SpringerPlus*, 5, Article 1305. <https://doi.org/10.1186/s40064-016-2971-1>
- Bouckaert, M. (2016). Designing a materials development course for EFL student teachers: principles and pitfalls. *Innovation in Language Learning and Teaching*, 10(2), 90-105. <https://doi.org/10.1080/17501229.2015.1090994>
- Breaux, R., Cash, A. R., Lewis, J., Garcia, K. M., Dvorsky, M. R., & Becker, S. P. (2023). Impacts of COVID-19 quarantine and isolation on adolescent social functioning. *Current Opinion in Psychology*, p. 52, Article 101613. <https://doi.org/10.1016/j.copsyc.2023.101613>
- Brown, H. D., & Abeywickrama, P. (2003). *Language assessment: Principles and classroom practices*. Pearson Education.
- Cabrero García, J., & Richart Martínez, M. (1996). El debate investigación cualitativa frente a investigación cuantitativa [The quantitative vs. qualitative research debate, clinical nursing]. *Enfermería Clínica*, 6(5), 212-217. <https://bit.ly/421dal5>
- Calvo Benzie, Y. J. (2017). Contributions of new technologies to the teaching of English pronunciation. *Language Value*, 9(1), 1-35. <https://doi.org/10.6035/LanguageV.2017.9.2>
- Chapelle, C. A. (2007). Technology and second language acquisition. *Annual Review of Applied Linguistics*, 27, 98-114. <https://doi.org/10.1017/S0267190508070050>
- Chilingaryan, K., & Zvereva, E. (2017). Methodology of flipped classroom as a learning technology in foreign language teaching. *Procedia - Social and Behavioral Sciences*, 237, 1500-1504. <https://doi.org/10.1016/j.sbspro.2017.02.236>
- Dalvi, A., Siddavatam, I., Thakkar, V., Vedpathak, A., & Patel, A. (2022). Recognizing child-unsafe apps through user reviews on the Google Play Store. In M. Bianchini, V. Piuri, S. Das, & R. N. Shaw (Eds.), *Advanced computing and intelligent technologies* (pp. 111-120). Springer. [https://doi.org/10.1007/978-981-16-2164-2\\_9](https://doi.org/10.1007/978-981-16-2164-2_9)
- Deng, Q., & Trainin, G. (2015). Learning vocabulary with apps: From theory to practice. *The Nebraska Educator: A Student-Led Journal*, 29, 49-69. <https://digitalcommons.unl.edu/nebeducator/29>
- Doardi, D. (2021). Digital environment and generation Z. *Teoría de La Educación*, 33(2), 27-47. <https://doi.org/10.14201/teri.25224>
- Dueñas Macías, F. A., Cardozo Becerra, D. X., & Peña Ibarra, C. M. (2015). Developing speaking skills through speaking-oriented workshops. *Enletawa Journal*, 8(2), 35-45. <https://rb.gy/9n9by2>
- Gael, K. E., & Elmiana, D. S. (2021). Mobile-Assisted Language Learning (MALL) in English language acquisition: a critical literature review. *Journal of English Language Teaching Innovations and Materials*, 3(2), 76-86. <https://doi.org/10.26418/jeltim.v3i2.49813>
- García del Dujo, Á., Vlieghe, J., Muñoz-Rodríguez, J. M., & Martín-Lucas, J. (2021). Pensar la (teoría de la) educación, desde la tecnología de nuestro tiempo [Thinking of (the theory of) education from the technology of our time]. *Teoría De La Educación. Revista Interuniversitaria*, 33(2), 5-26. <https://doi.org/10.14201/teri.25432>

- Godwin-Jones, R. (2011). Mobile apps for language learning. *Language Learning and Technology*, 15(2), 2-11. <http://doi.org/10125/44244>
- Guaqueta, C. A., & Castro-Garces, A. Y. (2018). The use of language learning apps as a didactic tool for EFL vocabulary building. *English Language Teaching*, 11(2), 61-71. <https://doi.org/10.5539/elt.v11n2p61>
- Hansen-Strain, L. (1989). Orality/literacy and group differences in second-language acquisition. *Language learning*, 39(4), 469-496. <https://doi.org/10.1111/j.1467-1770.1989.tb00900.x>
- Henriques, C. O., Chavez, J. M., Gouveia, M. C., & Marcenaro-Gutierrez, O. D. (2022). Efficiency of secondary schools in Ecuador: A value-based DEA approach. *Socio-Economic Planning Sciences*, 82, Article 101226. <https://doi.org/10.1016/j.seps.2022.101226>
- Hussain, A., Mkpojiogu, E. O., & Babalola, E. T. (2020). Using mobile educational apps to foster work and play in learning: A systematic review. *International Journal of Interactive Mobile Technologies*, 14(18), 178-194. <https://doi.org/10.3991/ijim.v14i18.16619>
- Irawan, A. W., Dwisona, D., & Lestari, M. (2020). Psychological impacts of students on online learning during the COVID-19 pandemic. *KONSELI: Jurnal Bimbingan dan Konseling*, 7(1), 53-60. <https://doi.org/10.24042/kons.v7i1.6389>
- Kacatl, J., & Klímová, B. (2019). Use of smartphone applications in English language learning challenge for foreign language education. *Education Sciences*, 9(3), Article 179. <https://doi.org/10.3390/educsci9030179>
- Khorechko, U. V., Scherbinin, A. I., Lebedeva, I. O., Gui, L. B., & Sentsov, A. E. (2015). Basic methodical grounds for teaching international students Chinese field-specific terms in Technical Institutes of Higher Education (in the Context of Polymer Production). *Procedia - Social and Behavioral Sciences*, 215, 43-52. <https://doi.org/10.1016/j.sbspro.2015.11.572>
- Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading and writing*, pp. 21, 783-804. <https://doi.org/10.1007/s11145-007-9092-8>
- Krapkov, A. A., Trapeznikov, L. P., Vovkushevskii, A. V., Kassirova, N. A., Petrov, V. A., Rukavishnikova, T. N., & Yudelevich, A. M. (1997). Use of existing and development of new computer technologies to analyze the statics and dynamics of concrete water-development works. *Hydrotechnical Construction*, 31, 301-306. <https://doi.org/10.1007/BF02768781>
- López-Leyva, J. A., Talamantes-Álvarez, A., Sanabria-Vincent, E., Aguilera-Silva, L., Gastelum-Rodríguez, G., & Meza-Arballo, O. (2020). Soil and environmental monitoring on a vineyard in the Guadalupe Valley as a tool for processes of precision viticulture based on ZigBee technology to improve the e-agriculture. In F. Narváez, D. Vallejo, P. Morillo, & J. Proaño (Eds.), *Smart technologies, systems and applications* (pp. 29-39). Springer. [https://doi.org/10.1007/978-3-030-46785-2\\_3](https://doi.org/10.1007/978-3-030-46785-2_3)
- Luna-Nevarez, C., & McGovern, E. (2018). On the use of mobile apps in education: The impact of digital magazines on student learning. *Journal of Educational Technology Systems*, 47(1), 17-31. <https://doi.org/10.1177/0047239518778514>
- Mauroux, L., Könings, K. D., Zufferey, J. D., & Gurtner, J.-L. (2014). Mobile and online learning journal: Effects on apprentices' reflection in vocational education and training. *Vocations and Learning*, 7, 215-239. <https://doi.org/10.1007/s12186-014-9113-0>
- Moniuszko, A., & Sciuk, A. (2013). Practice test # 3: Difficulty level-hard. In A. Moniuszko & A. Sciuk (Eds.), *PET and PET/CT study guide: A Review for passing the PET specialty exam* (pp. 153-225). Springer. [https://doi.org/10.1007/978-1-4614-2287-7\\_4](https://doi.org/10.1007/978-1-4614-2287-7_4)
- O'Donoghue, J., Singh, G., Caswell, S., & Molyneux, S. (2001). Pedagogy vs. Technocentrism in virtual universities. *Journal of Computing in Higher Education*, pp. 13, 25-46. <https://doi.org/10.1007/BF02940943>
- Riveros, V. S., & Mendoza, M. I. (2008). Theoretical considerations of the use of the Internet in education. *Omnia*, 14(1), 27-46. <https://www.redalyc.org/articulo.oa?id=73714102>
- Sarramona López, J. (2009). Los retos de las nuevas tecnologías para la educación a distancia [The challenges of new technologies for distance education]. *Teoría De La Educación. Revista Interuniversitaria*, 12, 13-27. <https://doi.org/10.14201/2864>
- Shih, R.-C., Lee, C., & Cheng, T.-F. (2015). Effects of English spelling learning experience through a mobile line app for college students. *Procedia - Social and Behavioral Sciences*, 174, 2634-2638. <https://doi.org/10.1016/j.sbspro.2015.01.945>

- Tannen, D. (1982). Exploring orality and literacy. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 1–16). Praeger.
- Terhorst, Y., Messner, E.-M., Schultchen, D., Paganini, S., Portenhauser, A., Eder, A.-S., Bauer, M., Papenhoff, M., Baumeister, H., & Sander, L. B. (2021). Systematic evaluation of content and quality of English and German pain apps in European app stores. *Internet Interventions*, 24, Article 100376. <https://doi.org/10.1016/j.invent.2021.100376>
- Välimaa, J., & Hoffman, D. (2008). Knowledge society discourse and higher education. *Higher Education*, 56, 265-285. <https://doi.org/10.1007/s10734-008-9123-7>
- Woo, J. H., & Choi, H. (2021). 기반의 언어 학습 도구들에 대한 체계적인 검토 [Systematic review for AI-based language learning tools]. *Journal of Digital Contents Society/디지털콘텐츠학회논문지*, 22(11), 1783-1792. <https://doi.org/10.9728/dcs.2021.22.11.1783>
- Yoon, H., Choi, K., Kim, J., & Jang, Y. (2021). Neighborhood walkability, personal active travel, and health in Asian Americans: Does English proficiency matter? *Journal of Transport and Health*, 21, Article 101082. <https://doi.org/10.1016/j.jth.2021.101082>