

Student Assessment of English Language Learning Apps

スティーヴン・E. クアシャ *Steven E. QUASHA*

Abstract

In this mobile-assisted language learning (MALL) research project, six business management majors from a women's university were introduced to five popular online apps for learning English. The participants were shown the apps during an informal 20-minute meeting. Then, participants explored using the apps over a three-week period simultaneously conducting assessments of the language learning apps using a five-point grading scale ranging from one to ten. The apps chosen for this study were Duolingo, Busuu, Memrise, FluentU, and Google Translate. Other than Google Translate, the remaining apps used in the study were new to the participants. Findings showed that learners especially enjoyed apps that featured gamification features, such as sounds and animations. These types of apps intrigued students and motivated them to increase their understanding of English. One caveat to this study was that students used free downloaded versions of the apps. As a result, many of the participants were confused by the constant upsell techniques offered nearly every time they accessed the apps.

Key Words : ☐ mobile-assisted language learning (MALL)

☐ language learning apps

☐ evaluation rubrics

☐ gamification and language learning

☐ motivation

Introduction

Never before in the history of language learning have students possessed the ability to utilize online resources for second language acquisition. Students today can access language dictionaries, audio pronunciation guides, and connect with online tutors, all in the palm of their hand. Yet, does all this add up toward producing more proficient language learners? Some key questions persist: For example, how can we determine the efficacy of these online learning tools? While some may appear to be indispensable tools for second language learners, a host of them could be categorized as online traffic. They represent content that users may glance at out of curiosity, but they cannot necessarily be quantified as a benefit to language

learning.

This paper's aim was to introduce students to online English language learning resources. In particular, the efficacy of smart phone apps was the main focus for this study. To understand students' personal experiences, interviews were conducted to help the author select which digital resources rated the highest for students. The interview group consisted of five students, with three first-year participants and two second-year students. All of the participants are business management majors at a four-year, private women's university located in central Japan. The interviews consisted of an initial 20-minute introduction of the research project with all five participants after week four of the semester. This was followed up one month later with two separate 30-minute open-ended discussions during the halfway point of the second academic semester. Further analysis of these interviews appears in the methods section of this paper.

Literature Review

Mobile-assisted language learning (MALL) is a derivation from mobile learning (m-learning) and computer-assisted language learning (CALL). MALL has gained a strong following among technologically savvy language educators because of its ability to support students' language learning through the use of mobile technologies such as mobile/smart phones. The foray into using mobile devices for education began in the 1990s when Brigham Young University-Hawaii used a cell phone to teach a distance education course with students in Tonga via telephone and computer (Green, Collier, & Evans, 2001). This was followed by the Stanford University learning lab using an integrated mobile phone to teach a Spanish course in 2001 (Chinnery, 2006).

In the Japanese university setting, Thornton and Houser (2005) developed some groundbreaking projects—including a course management system called Poodle—to provide SMS push messages and tasks for students to review and practice English outside of class. Their research findings yielded many positive results for students using the system. Vocabulary retention doubled, and students expressed a positive experience and preference for SMS instruction. Other educators have echoed that this progress might help usher language students from the formal classroom setting to a more informal mobile learning environment (Kukulska-Hulme, 2009). Taken a step further, Steel (2012) postulated that using apps for language learning is a highly desired activity for today's students, and educators would be wise to harness MALL into their curriculum.

Other advances using smart phones for language learning include Gromik's (2012) study which required students to utilize the video recording function on their device to produce short monologues in English. The author reported that students were effectively able to increase the

length of their video recording time over the course of the semester. The notion of push media versus pull—where students access the learning themselves—is succinctly summarized by Stockwell (2013). While pushing the learner into action certainly has its merits, the brunt of this research study will encompass the pulling that students encounter when they interact with language learning apps outside of the classroom. It is this arena of language learning using MALL that Godwin-Jones (2017) called the most important since it helps place learners into discovery mode and has potential for deeper learning.

Rise of the Apps

Since the introduction of the World-Wide Web in the early 1990s, computer aided language learning began its trajectory as a bona fide teaching tool. At first, learners had opportunities to use CD-ROMs and multimedia to bring about a sense of entertainment into the realm of foreign language learning. Many educators, the author included, envisioned this new method as an ideal way to motivate language learners to go beyond the classroom textbook, and potentially spend more hours studying and reviewing outside of class. However, creators of digital language acquisition programs were in direct competition with the burgeoning gaming industry. It is no secret that the latter boasts a much bigger budget, better skilled software programmers, and the ability to consistently deliver more entertaining content.

Although developers can now readily market language learning software programs directly to end users, introducing these programs through the school system represented a potentially higher rate of adoption. Therefore, having competent language teachers to help direct students to suitable materials was tantamount for this aim. Computer literate teachers are vital agents of change as Halverson and Smith (2009) elucidated in their study of an experimental high school in Israel. However, age is often a key determinant whether or not some educators are open toward using new technologies in school. In short, the older the teacher, the less likely they were toward trying new technology in the classroom.

With the introduction of the Apple i-Phone in 2007, using a smart phone as a learning tool took a giant leap forward. The advent of the i-Phone helped pave the way for developers to create applications—more commonly known as apps—which has helped catapult the smart phone as an indispensable daily source of information for people around the globe. With the rise of the smart phone, fewer people have personal computers at home, as apps have slowly begun replacing expensive software that once powered the personal computing revolution (Kaur, 2013). More importantly, this migration has resulted in a higher adoption rate for second language learners to use their smart phones as language learning tools (Kilmova, 2018).

Figure 1 shows the steady rise of app development for Apple, or iOS users, over the past

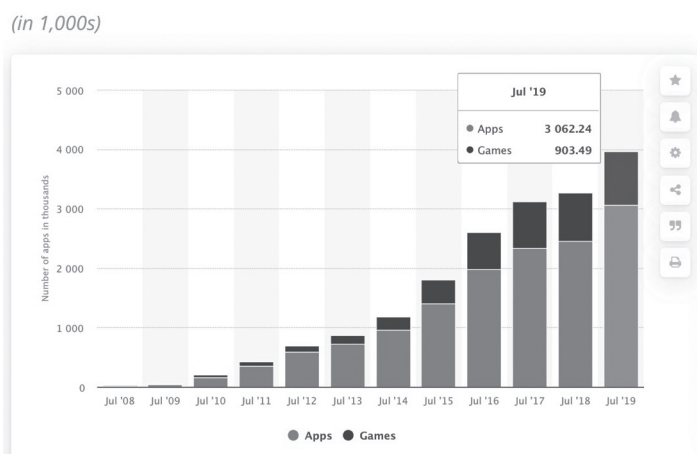


Figure 1. Number of Available Apps in the Apple App Store from 2008–2019

Note. Reprinted from Statista: Retrieved from <https://www.statista.com/statistics/268251/number-of-apps-in-the-itunes-app-store-since-2008/>

decade. While iPhone users downloaded 30 billion apps in 2018, Google Android smart phone users doubled that number with 76 billion downloads (Iqbal, 2019). With apps leading the way, the migration from the PC to mobile computing is clearly the driving force for the next digital revolution.

With so many applications available for downloading in the education category, it is imperative to use an assessment scale to determine the effectiveness for each app. Chen's (2016) assessment rubric for online language learning materials offered some salient points. In this study, the author created seven elements for evaluating apps. These consisted of content quality, pedagogical coherence, feedback and self-correction, motivation, usability, customization, and sharing. For this study, a modified version of Chen's assessment rubric featuring five elements instead of seven was deployed. The five rubrics used are listed in Table 1.

The main difference of this study vis-a-vis Chen's earlier work was that rather than have the author determine the rubric results for the various apps, a small group of students served as assessors for the language learning apps. Since students represent the end user—and are functioning in their second language—this approach seemed a more accurate measurement regarding the efficacy for using apps as language learning tools.

Methods

Five students majoring in business management at a private women's university served as participants for this study. During the middle of the second academic semester, all five of the students met together with the author where they were introduced to a host of online English language learning apps. After confirming that participants were able to download or access

Table 1. Evaluation Rubric for Language Learning Mobile Apps

Category	Least Suitable (1–3)	Average (4–7)	Most Suitable (8–10)
Motivation: Content is able to get the user to remain interested in the app.	Content does not foster self-directed learning by the user.	Limited amount of content, so users are mildly interested in the app.	Content helps users remain quite interested in the app. Promotes self-directed learning.
Usability: The app icons and menus are easy to navigate and understand.	Icons and menus are confusing. No on-screen help for users.	Icons and menus are clearly marked, but lack of on-screen help.	Icons and menus are clearly marked. On-screen help and tutorials are also available.
Quality of content: Content should help increase learners' English language ability.	Content does not help achieve learning goals or promote autonomous learning.	Content helps achieve learning goals, but does not promote autonomous learning.	Content helps achieve learning goals, autonomous learning, and relates previous learning to new content.
Feedback and self-correction: Users should receive feedback to promote self-regulated learning.	Feedback is highly limited for user responses.	Feedback is available and offers users a chance to try to correct their answers.	Feedback is quite specific and helps users improve their performance. Data is available for learners and teachers.
Sharing: Permits users to share their learning progress or concerns regarding the app.	Users progress is unavailable or lacks an ability to access.	User performance data is accessible within the app, however exporting that information is difficult.	Ability to save user progress within the app. It can be exported to classmates and/or the teacher.

each of the apps with their smart phones, students were then given a three-week time frame to use the apps and write down their results. These results were tabulated using the rubric that appears in Table 1.

In the introductory meeting, the author offered English-to-Japanese translations for each of the categories of the rubric handout to confirm that participants fully understood the grading rubric. All of the students wrote corresponding Japanese words to help them remember the information contained in each of the grading rubrics. At the end of the introductory meeting, participants were requested to give each app a score ranging from one to ten. This information was listed at the top of the rubric handout in the shaded boxes labeled; least suitable, average, and most suitable.

The free apps used in this study appear in Table 2. These five are listed by the name of the app, category, followed by their Google Play and Apple Store ratings. The reasoning behind choosing these five were as follows: Students would only be able to allocate time to explore a

Table 2. Name, Category, and Ranking of Apps

App Name	iOS Education Category Rank	Google Play Rating	Apple Store Rating
Duolingo www.duolingo.com	#2	4.7	4.7
Busuu: Learn English Speaking https://www.busuu.com/en/languages/english	#90	4.5	4.7
Memrise https://www.memrise.com/	#106	4.5	4.8
FluentU (free version) https://www.fluentu.com/english/	n/a	3.2	3.7
Google Translate https://translate.google.com/	#1 in Reference Category	4.5	4.5

maximum of five apps during the weeks of this research project. All of the students in the research project had part-time jobs outside of school, and some were also involved in club activities. As a result, much of participants' free time to review or study outside of class was rather limited.

Another issue to consider was that all of the apps had to be free since participants survived on a tight budget. The apps selected for this study were based on their popularity from internet search results. This list of apps which appear in Table 2 garnered the highest number of recommendations from a variety of language learning website reviews.

Results

Students were given their own assessment rubric to fill out with a notes section which appears in the appendix section of this paper. All participants were reminded that they were expected to review all of the apps on the list, and assign a numeric score ranging from one to ten for each app. The results of the participants scores appear in Table 3.

In the category of motivation, participants chose Duolingo as their top app. When queried for the reason during the interview session, students responded that the content seemed new and exciting. The app featured sounds that let them know if their answers were correct. They felt encouraged to continue using the app since it seemed more like a game. This notion of gamification helped keep learners on task and interested in the language app. Also, the pronunciation button was easy to see, and students liked having the ability to press the button to listen again. It was a form of built-in scaffolding. In addition, after they completed each stage, a congratulatory sound was made. This made participants want to explore the app more to play and improve.

Table 3. Student Evaluations for Language Learning Mobile Apps

Category	Duolingo	Busuu	Memrise	FluentU	Google Translate
Motivation	5.0	3.2	4.5	4.8	3.2
Usability	5.5	4.3	4.3	6.5	8.3
Quality of content	5.0	2.7	4.5	5.5	5.3
Feedback and self-correction	5.0	2.3	6.0	5.0	2.3
Sharing	1.2	1.0	2.0	3.7	1.3
Total Score	21.7	13.5	21.3	25.5	24.4

FluentU received similar positive reviews regarding its layout and design which led to a higher motivation score. However, participants—when asked to choose between the two—preferred Duolingo by a slight margin. One thing that students did not enjoy was the constant upsell within the Duolingo app. They were constantly asked to try the upgrade for seven days. Doing so would unlock quizzes and other new information within the app. One student found this to be demotivating, and felt that it caused her anxiety when using the app.

For usability, Google Translate was the clear-cut winner. Part of this could stem from the fact that students were already accustomed to using the app on their smart phones. Since there was no learning curve compared to the other apps introduced in this study, users had very little trouble navigating how to translate words or phrases with Google Translate. According to participants, FluentU beat out Duolingo for the second spot due to its layout. However, since FluentU only offered a two-week trial, a few of the participants were unhappy about having to agree purchasing and then cancelling the app. Many mumbled that this maneuver was troublesome. Some were worried about having to pay for many months of fees. Additionally, a few of the participants harbored concerns that the free material might not be as in-depth as the paid version of FluentU. Therefore, participants remained divided over the actual usability of their version of the FluentU app.

The quality of content category once again showed that FluentU and Duolingo were the best of the apps. This was followed closely by Google Translate which was an interesting choice. Participants assumed that the translations offered by machines—or Google in this case—were wholly accurate. Many of us who are second language speakers and learners know that machine translation can be a very inexact science. While it may help us with simple translations, other times machine translation churns out attempts at word-for-word interpretation that does not come close to the meaning in its original language. The interesting element of this category's result is that participants assumed Google Translate was error free.

When queried for a more in-depth response, participants said they believed that Google Translate offered them a better chance to understand unknown words or expressions in English. Without the convenience of the app, students said they would feel lost.

In the area of feedback and self-correction, participants selected Memrise as the best app. Student responses during the interview revealed that Memrise asked learners to type their answers more often than the other apps. Participants felt that this allowed for more interaction and less of the feeling of simple listening practice. Also, users received instant feedback whether or not their answers were correct within the Memrise app. Both Duolingo and FluentU received identical scores in the feedback and self-correction category. According to participants, nothing really separated Duolingo and FluentU in this category. They both had similar approaches in the way they offered feedback. Upon further inquiry, students answered that the feedback and self-correction was good, but not outstanding.

The last category was sharing. Unfortunately, most of the apps used in the study did not fare well in this category. FluentU led the pack, but again this comes with a slight disclaimer since the free version likely does its best to entice users to upgrade to the full version which costs money. Although many of the apps automatically recycle student errors, and thereby force learners to eventually make the correct choice, participants seemed unsatisfied by the current state of sharing within the apps. Learners expressed the need for language learning apps to have a separate screen so they can check on their mistakes. Interestingly, none of the participants expressed a desire to export the information or share it with a teacher. Rather, users said they would like to keep the information for themselves within the app to check on their errors to gauge their progress.

For total score, the FluentU app slightly outpaced Google Translate. As aforementioned, familiarity and ease of use—with no learning curve nor symptoms of anxiety—likely led to the high score by participants in this study for Google Translate. FluentU was the second choice as students enjoyed its usability and engaging content. Overall, students felt that the Busuu app was inferior in many of the categories compared to the other apps. Another reason echoed by users regarding Busuu was that, similar to FluentU, they were constantly being needed to upgrade to the paid version in order to access the quizzes and better content. Many of the participants expressed a desire to be left alone while exploring and playing with the app. They mentioned that after they finish with the free version, they would then decide to purchase additional content from the app developer. They felt the apps took away some of the joy by constantly reminding them to sign up for the paid version.

Discussion

Although Duolingo claims to be the most popular language learning platform with more

than 150 million users, participants in this study ranked it third. For the most part, Duolingo is a translation software product that has been gamified. Perhaps participants in this study recognized these weaknesses which then justified the lower score.

One area that caused demotivation and some anxiety was the numerous attempts by many of the app learning programs to upsell participants from the free version to a monthly or yearly fee. Therefore, it is little wonder why Google Translate received such a high usability score from participants. The app never advertises, nor does it try to upsell users. In hindsight, combining Google Translate with the rest of the language learning apps was problematic from an analysis standpoint. Google Translate is a reference app, not an education one. As such, there was little new discovery among users because all of them knew the product. Additionally, Google Translate does not offer feedback or sharing in its current digital incarnation and could not compete in that category.

Better preparation by the administrator of this research study during the initial meeting with students—with screen shots explaining the upsell techniques—may have helped alleviate some of the anxiety participants experienced. Nevertheless, in the final meeting between students, many participants expressed a desire to purchase a paid version of the course for two of the apps featured. However, rather than access the course during the school semester, participants felt that using the apps while on the longer school breaks of August-September and/or February-March would be more compatible with their schedules.

Finally, a larger sample size of this study might generate different results than the findings of this paper. However, given the systemic restraints of limited free time learning outside of class for many Japanese university students, organizing a sample size of 50 or more participants could prove to be quite challenging. Despite any shortcomings of this study, this MALL pulling approach for mobile phones was quite formidable. According to students, it created more interest in English language learning among all of the group participants, and many were curious how students from other countries learned English online. Perhaps this curiosity will lead to more exploration and discovery using their smart phones as an important learning tool.

References

- Chen, X. (2016). Evaluating language-learning mobile apps for second-language learners, *Journal of Educational Technology Development and Exchange*, 9(2), Article 3. Retrieved from <https://aquila.usm.edu/cgi/viewcontent.cgi?article=1002&context=jetde>
- Chinnery, G. (2006). Going to the MALL: mobile assisted language learning. *Language Learning and Technology*, 10(1), 9–16.
- Godwin-Jones, R. (2017). Smartphones and language learning. *Language Learning & Technology*, 21(2), 3–17. Retrieved from <https://scholarspace.manoa.hawaii.edu/>

- bitstream/10125/44607/1/21_02_emerging.pdf
- Green, B., Collier, K., & Evans, N. (2001). Teaching tomorrow's class today: English by telephone and computer from Hawaii to Tonga. In L.E. Henrichsen (Ed.), *Distance-learning program* (pp. 71–82). Alexandria, VA: TESOL.
- Gromik, N. (2012). Cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, 58(1), 223–230.
- Halverson, R., & Smith, A. (2009). How new technologies have (and have not) changed teaching and learning in schools. *Journal of Computing in Teacher Education*, 26 (2), 49–55.
- Iqbal, M. (2019). App download and usage statistics. Retrieved from <https://www.businessofapps.com/data/app-statistics/>
- Kaur, S. (2013). The revolution of tablet computers and apps: A look at emerging trends. *IEEE Consumer Electronics Magazine*, 2(1), 36–41.
- Kilmova, B. (2018). Mobile phones and/or smartphones and their apps for teaching English as a foreign language. *Education and Information Technologies*, 23(3), 1091–1099.
- Kukulska-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL*, 21(2), 157–165.
- Statista (2019). Retrieved from <https://www.statista.com/statistics/268251/number-of-apps-in-the-itunes-app-store-since-2008/>
- Steel, C. (2012). Fitting learning into life: Language students' perspectives on benefits of using mobile apps. In M. Brown, M. Hartnett & T. Stewart (Eds.), *Future challenges, sustainable futures*. Proceedings ascilite Wellington 2012. (pp. 865–880). Retrieved from https://www.researchgate.net/profile/Caroline_Steel/publication/288308534_Fitting_learning_into_life_Language_students%27_perspectives_on_benefits_of_using_mobile_apps/links/5a4c1904aca2729b7c8950da/Fitting-learning-into-life-Language-students-perspectives-on-benefits-of-using-mobile-apps.pdf
- Stockwell, G. (2013). Tracking learner usage of mobile phones for language learning outside of the classroom. In P. Hubbard, M. Schulz & B. Smith (Eds), *Learner-computer interaction in language education: A festschrift in honor of Robert Fischer* (pp. 118–136). San Marcos, TX: CALICO.
- Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, 21, 217–228.

Appendix. Evaluation Rubric for Language Learning Mobile Apps

Category	Least Suitable (1–3)	Average (4–7)	Most Suitable (8–10)
<p>Motivation: Content is able to get the user to remain interested in the app.</p> <p>My score: _____</p> <p>Notes: _____</p> <p>_____</p> <p>_____</p>	<p>Content does not foster self-directed learning by the user.</p>	<p>Limited amount of content, so users are mildly interested in the app.</p>	<p>Content helps users remain quite interested in the app.</p> <p>Promotes self-directed learning.</p>
<p>Usability: The app icons and menus are easy to navigate and understand.</p> <p>My score: _____</p> <p>Notes: _____</p> <p>_____</p> <p>_____</p>	<p>Icons and menus are confusing. No on-screen help for users.</p>	<p>Icons and menus are clearly marked, but lack of on-screen help.</p>	<p>Icons and menus are clearly marked. On-screen help and tutorials are also available.</p>
<p>Quality of content: Content should help increase learners' English language ability.</p> <p>My score: _____</p> <p>Notes: _____</p> <p>_____</p> <p>_____</p>	<p>Content does not help achieve learning goals or promote autonomous learning.</p>	<p>Content helps achieve learning goals, but does not promote autonomous learning.</p>	<p>Content helps achieve learning goals, autonomous learning, and relates previous learning to new content.</p>
<p>Feedback and self-correction: Users should receive feedback to promote self-regulated learning.</p> <p>My score: _____</p> <p>Notes: _____</p> <p>_____</p> <p>_____</p>	<p>Feedback is highly limited for user responses.</p>	<p>Feedback is available and offers users a chance to try to correct their answers.</p>	<p>Feedback is quite specific and helps users improve their performance. Data is available for learners and teachers.</p>
<p>Sharing: Permits users to share their learning progress or concerns regarding the app.</p> <p>My score: _____</p> <p>Notes: _____</p> <p>_____</p> <p>_____</p>	<p>Users progress is unavailable or lacks an ability to access.</p>	<p>User performance data is accessible within the app, however exporting that information is difficult.</p>	<p>Ability to save user progress within the app. It can be exported to classmates and/or the teacher.</p>

---【著者略歴】---

Steven E. QUASHA (スティーヴン・E. クアシャ)

1962年 Brookline, MA. U.S.A. 生まれ

所 属・現 職 梶山女学園大学現代マネジメント学部現代マネジメント学科・講師

最終学歴・学位 修士 (Master of Arts) : Asian Students (1994) & Applied Linguistic-TESOL (2007)

所 属 学 会 Japan Association for Language Teaching (JALT) (全国言語教育学会) ;
American Council on the Teaching of Foreign Languages (ACTFL) ;
International Language Testing Association (ILTA)

主 要 著 書 “A Single Step to English Communication—3rd Edition” (単著) (2018.3)
Gifu, Japan: Akebono Press.

“Are We Listening? Student Reflections on the use of Music in the English
Language Classroom” (単著) (2018.3) 梶山女学園大学現代マネジメント学部紀要「社会とマネジメント」第15巻, pp. 15–27.

“Destination Weddings: The Allure of Maui” (単著) (2017.3) 梶山女学園大学現代マネジメント学部紀要「社会とマネジメント」第14巻,
pp. 1–10.