What Makes a Digital Language Learning Platform Work? A Usability and Quality Review

Irfan Rifai[1*], *Alex* Jhon[1], *Kristianus* Oktriono[2], and *Elisa Carolina* Marion[3]

[1]English Department, Faculty of Humanities, Bina Nusantara University, Jakarta, Indonesia, 11480
[2]Language Center, Tourism Department, Faculty of Humanities, Bina Nusantara University, Jakarta, Indonesia, 11480
[3]Japanese Literature Department, Faculty of Humanities, Bina Nusantara University, Jakarta, Indonesia, 11480

Abstract. Due to the escalating usage of the internet and digital applications in recent years, foreign language learning experiences has become easier, simpler, and more fun for users. Consequently, to maintain those factors, usability and the quality of the platform are two main concerns for platform developers. To provide a personalized, structured, and gamified platform that offers opportunities for social interaction among users, researchers developed a digital language learning platform in middle of 2020. In 2021, the platform was tested on 3200 students who provided feedback on its usability through System Usability Scale. In addition, a small group of students were involved in an FGD session where they shared their experiences in using the platform. The usability test of the platform scored at 58, which implied that the platform still has some room for improvement. Through the FGD session, users could find information and functions related to the application easily and they expressed satisfaction toward the platform's performance and valued the feedback opportunities. This research contributes to the field of a digital language learning platform development, especially when it is used by the students of higher education who wish to improve their foreign language mastery conveniently.

1 Introduction

The interest on foreign language learning has been perpetually high, with English being the most popular language studied at a global scale, followed by French, Spanish, German, Japanese, Italian, and Korean [1]. The reasons and motives people learn a foreign or additional language have been mostly about pleasure, doing business, and educational purposes. Whatever motivates modern people would be, Meticulous Research [2] predicted that that online foreign language learning market in 2027 will reach up to 21.2 billion in value, showing that the interest in learning a foreign language in the future remains high and promising for language learning service providers.

The current leading online language learning platform is Duolingo, which has generated a revenue of more than 250 million in 2021 [3]. Duolingo is not the only platform that offers language learning service. In recent years, similar platforms have offered similar services, such as Memrise, FluentU, Babbel, Busuu, Livemocha, byki, Rosetta Stone, LingQ, and two Indonesian local platforms: Bahaso and Cakap. Each of these platforms is competing to offer unique language learning experience for such a large market to win the hearts of millions of potential users. For example, Duolingo offers an easy-to-use application and Grammar Translation Method concept, which is outdated but

works surprisingly well [4]. In addition to the solid foundation on the chosen theory and method, Duolingo offers its users features such as instant feedback and XP point system to increase users' engagement, which may have been the other contributing factors for its success.

In comparison, Babble, the second most popular online language learning platform with 10 million subscribers (about the population of Jakarta), offers three methods that are grounded in real-life conversations, authenticity of interactions, and a variety of supporting learning tools such as podcasts, stories, and games. Unlike Duolingo, Babble strongly bases its cognitive-based language learning theory to emphasize scaffolded proficiency in vocabulary and phrases required in real conversation.

Studies found that online / web-based language learning or mobile based language learning has been an efficient method in reducing anxiety and increasing motivation [5], interactive and helpful [6], or even improving inclusiveness [7]. With an attempt to develop an online language learning platform, a team of language experts and IT professionals at a private university in Jakarta worked together to assemble an online language learning platform. The platform, available in mobile and web, was built on the team's interpretation of Krashen's Second Language Acquisition theory that revolves around the five hypotheses, i.e., the acquisition-learning hypothesis, monitoring hypothesis,

---

* Corresponding author: irifai@binus.edu

input hypothesis, the affective filter hypothesis, and the natural order hypothesis. [8] and aspects of sociocultural theory in learning [9]. Thus, it was designed with the ultimate understanding that students need to have personalized, structured, and gamified learning experiences, while simultaneously having the opportunities for social interaction with other users.

The development of the platform was finished in 2021. To test the usability i.e., the capacity of a system to provide a condition for its users to perform the tasks safely, effectively, and efficiently while enjoying the experience [10] we conducted two rounds of surveys to measure the usability of the platform and the quality of the platform judging from the users' experiences. To improve our understanding of the usability data survey we gathered quantitatively, we conducted an FGD session to a group of 29 users to whom we explored the usability of the platforms in qualitatively.

Our research was aimed to answer the questions on how the users rate the usability and the quality of the newly built language learning platform and how users experience their online language learning through the newly built platform. The users' responses on the usability test and their experiences in using the platform were limited to the use of one single platform that the team of developers has just developed.

2 Literature Review

Ever since the Covid – 19 pandemics struck in early 2020, there has been a surge of numbers of students who have turned to online platforms for their language learning. A lot of these platforms integrated concepts such as gamification and social interactions among learners to engage participants and elevate their language learning experience [11-12]. They have also gotten more choices to improve their language proficiency at their own time and space. A concept that a Second Language Acquisition (SLA) scholar, Krashen [8] has long identified.

Krashen hypothesized five theories on second language acquisition: the acquisition-learning hypothesis, the Monitor hypothesis, the Input hypothesis, the Affective Filter hypothesis, and the Natural Order hypothesis. The first hypothesis of Krashen's theory of second language acquisition is the learning – acquisition thesis which stated that there are two independent systems of foreign language learning: "acquired system" and "learned system". In the acquired system, learners learn the language subconsciously. Just like children when they learn a language, the learners need to be involved in a natural set of communication in order to acquire the language. The learned system, on the other hand, is a formal instruction where learners are involved informal instruction to study about the language. The learned system, on the other hand, mostly occurs at schools.

With people are now exposed to English and other foreign languages in the forms of movies, digital news, podcasts, music videos, books of different forms, all of those within reach and in the palms of their hands. Brian Immanuel or Rich Brian, an Indonesian-based rapper currently enjoying his career in United States, is a concrete example of how constant exposures to English and in- context usage of the language through these various media can "teach" someone another language effectively [13]. There are, of course, pros and cons as to what constitutes a better opportunity to language learning: acquisition, like the case of Brian, or by learning systematically, through guidance of teachers.

The next hypothesis is the monitor hypothesis. In this hypothesis, learners are subject to learn consciously to monitor the correctness of the language that they learn. Some studies associate the monitoring hypothesis with the personality of learners [14]. They stated that most introverts, due to lack of confidence, overuse the monitoring function in their language learning. Krashen's third hypothesis is the input hypothesis. Krashen believes that to support students' learning, teachers or in this case, a platform, must provide learning material that is one step beyond their current knowledge. Thus, the input, or the knowledge being introduced, should be comprehensible for students. Krashen calls it "comprehensible input". For input to be comprehensible, they must be interesting and fun for learners.

The fourth hypothesis is the affective filter hypothesis. Krashen hypotheses that having affective factors, or the lack of them, like motivation, self-confidence, and good self-image can be influential to the success or failure for one's efforts in learning a foreign language. Lack of confidence and low self-image can enhance affective filter that may eventually form "mental blocks" in learning.

Gamification is not a new concept to language learning. Rooted in Computer Assisted Language Learning (CALL), gamification has gotten into language classrooms effortlessly. Gamification Assisted Language Learning or GALL [15], is a term conceptualized by the author to emphasize the elements that he believes would benefit language learners, i.e., users are happy and engaged, existing goals to achieve, rules to limit the game, background information or recorded achievement, and the psychological elements that would increase the motivation of users. The last hypothesis is the natural order hypothesis. In this hypothesis, Krashen theorizes that the acquisition of grammar in one's language follows a natural order, which is predictable. It all depends on the age, their first language background, and exposure.

The sociocultural theory as well as the concept of gamification place learners as the agents of their own learning. We would use both concepts as main references as we build the content and the platform for digital language learning. The idea of play is instrumental in the perspective of sociocultural theory of learning [9]. From the perspective of sociocultural theory, human activities are mediated by cultural tools

be it signs (language) or external tools. Human activities are understood to be a purposeful interaction of the subject with the world (the object), and through the enactment of activities, human transforms the object while the object transforms human as well x [16]. Gamification, thus, fits perfectly as both tools for social interactions and tool to increase personal motivation.

Reviewing the five hypotheses, we laid the conceptual foundation for the development of a new online language learning platform (Table 1).

Table 1. The conceptual foundation of the platform

Krashen's Hypothesis	Platform Conceptual Foundations	Platform Features
Acquired vs Learned	Personalized, Structured	Rich texts and learning resources: videos, MOOCS, etc. Leveled competencies
Monitor	Gamified, Structured	Feedback, Check points, and Course Mastery
Comprehensible Input	Structured	Scaffolded, leveled Content
Affective Filter	Personalized, Social, Gamified	XP points, Discussion Forum, & Video Conference
Natural Order	Structured, Gamified	Scaffolded, leveled Content

In addition to the four conceptual foundations, we have also developed other features such as gender [7], cross-cultural understanding [17], and accessibility [18] features to ensure that the newly developed platform meets the needs of all users.

A systemic study on the usability of a language learning platform such as Duolingo found that the challenging tasks, reward incentives, systematic levels, and the ranking of users according to their achievements have been the gamification components that supported its popularity. The foci of the studies exploring MALL have been emphasized on the creation of tools rather than the process and outcomes of language learning from using these tools [19]. There has been a dearth of studies that explored users' voices on their experiences in using the tools or their language learning achievement after the experience. By integrating usability tests with FGD to hear students' opinions and experiences, this study aims to fill that gap.

3 Methodology

This is a study of two round surveys to measure the usability of the platform and the quality of the platform from the users' experiences. To improve our understanding of the usability data survey, we conducted an FGD session to a group of 29 to whom we explored the usability of the platforms in qualitative manner.

3.1 System Usability Scale (SUS)

Application development must pay attention to several important aspects, such as its design. Technically, creating the interface design of an application requires actual testing to peruse the elements required, such as user experience or usability. Usability testing in this research uses the System Usability Scale (SUS). SUS was developed by John Brooke in 1986 Lewis [20] explained the System Usability Scale (SUS) as standardized questionnaire for the assessment of perceived usability. Bangor et al (2008) [21] stated that the analysis and experiment indicated that the SUS was an exceptionally robust and multi-purpose instrument for professionals. In detail, SUS consists of 10 questions and 5 answer options. The answer choices ranged from strongly disagree to strongly agree.

To get feedback on the usability on the newly built platform, the following questions in table 2 were used.

Table 2. SUS Rating System

No	Questions	1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
1	I think that I would like to use this system.					
2	I found the system unnecessarily complex.					
3	I thought the system was easy to use.					
4	I think that I would need the support of a technical person to be able to use the system					
5	I found the various functions in the system were well integrated.					
6	I thought there was too much inconsistency in this system.					
7	I would imagine that most people would learn to use this system very quickly					
8	I found the system very cumbersome to use.					
9	I felt very confident using the system.					
10	I needed to learn a lot of things before I could get going with this system					

3.2 Focus Group Discussion (FGD)

We invited students taking part in the pilot project to join for a discussion session on their experience in using the platform. There were 29 users, 15 female and 14 males of 17 – 18 years old who joined the discussion session. All of them took the English course.

There were five questions that were asked to the 10 students. Here are the five questions:
1. What do you like or do not like about the platform?
2. What kind of obstacles did you experience when accessing the platform?
3. What features do you know are offered by the platform?
4. Did you have any obstacles with the features offered by the platform?
5. Do you have any suggestions for the platform?

The FGD session was led by a faculty who have been involved in developing the content and the platform. The questions were only the question to trigger students' responses. Students' responses were noted, coded, and highlighted for further development of the platform.

4 Results and Discussion

There were 32 users who filled the SUS, 21 identified themselves as female and 12 identified themselves as male. All were university students who took the English course at the platform.

4.1 Usability

Table 3. Usability Scores

Item	(1) Strongly Disagree	(2) Disagree	(3) Neutral	(4) Agree	(5) Strongly Agree
Q1: Accessibility			3.5625		
Q2: Complexity		2.6875			
Q3: Simplicity			3.46875		
Q4: Technical Support		2.6875			
Q5: Integrated function			3.4375		
Q6: inconsistency			3.5		
Q7: easy to learn			3.75		
Q8: complicatedness		2.59375			
Q9: confidence			3.40625		
Q10: pre-learning process		2.875			

Based on System Usability Scale (SUS), the 32 respondents explained the following experiences qualitatively where:

1. The app is easy to access is in the range of neutral to agree
2. The app is too complex is in the range of disagree to neutral
3. The app is easy to use is in the range of neutral to agree
4. The needs technical assistance from other people to be able to use this App is in the range of disagree to neutral
5. The various functions are well integrated in this App is in the range of neutral to agree
6. They feel a lot of inconsistency in this App is in the range of neutral to agree
7. They imagine others will find it easy to learn this App quickly is in the range of neutral to agree
8. They find this App inconvenient/complicated to use is in the range of disagree to neutral
9. they feel confident when using this App is in the range neutral to agree
10. they need to learn many things before they can use this App is in the range of disagree to neutral

As it is shown by table 3, the System Usability Scale (SUS) contains 10 statement items where odd numbers represent positive tone and even numbers represent negative tone. The respondents gave an assessment of the usability of the product on a scale of 1 to 5. A scale of 1 means strongly disagree and a scale of 5 means strongly agree. Calculation of positive items is the score

of each statement obtained from the user's score was reduced by 1. For each question with an even number, the final score was obtained from the score of 5 minus the score of the user's statement. Following that, the SUS score was multiplied by 2.5 to get the total score. A product is considered to have good usability if the overall SUS score equal to or above 68. In this research, the usability score is 58.20. It signifies that the application has room for improvement although the score states that it is on average quality.

4.2 Quality – Users' Perspectives

During the FGD, we asked several questions that targeted users' responses after the experience of using the platform. Even though we have thoroughly surveyed the usability of the platform, during the FGD session, we asked their impressions about the usability of the platform.

In terms of usability, users were asked some questions on what they like about the platform. Many of them believed that the platform was "practical" and "easy to understand". They also stated that the "using the applications was like playing games while learning". In addition, they stated that "the questions are interesting", "they have unique learning system", "easy to find", has "limited processing time", "simple", "interesting", and "has fun design". Some students emphasized that it was "easy to hone language skills".

On the other side, users also expressed what they did not like about the platform. In this part, the users stated that they encountered problems in using voice recognition, one of the features at the platform. There was also widescreen limitation when users accessing the platform through smartphones. They also complained about the size of the font, and lavk of variation in the activities. This occurs, mainly due to the adjustment of features used in different versions (mobile or web-based). For example, the width of the screen, physically is smaller on the mobile version. Users were more comfortable in using the web-based version. Above all, the users also experienced limited explanation or manual in using the platform.

This research also adopted quality criteria for the assessment and criterion of learning objects [20]. The pedagogic and evaluation of linguistic is used as the perimeter of educational application quality. There are 10 rubrics for apps were applied i.e., cognitive value and pedagogic coherence, content quality, capacity to generate learning, interactivity and adaptability, motivation, format & layout, usability, accessibility, visibility, and interoperability. This study indicated that the highest score was in pedagogic coherence and the lowest score was in the accessibility.
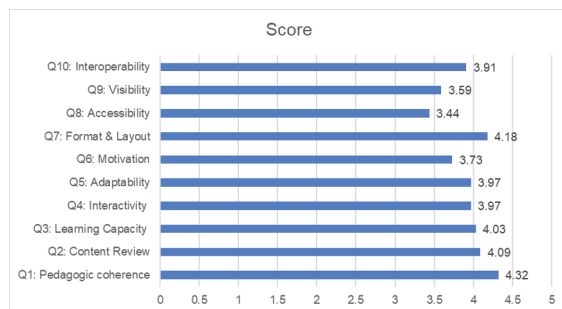
Fig. 1. The quality of the platform

In terms of user's perspective, there are 10 statements representing quality (figure 1). The first one is related to interoperability in the platform with the score above the average (3.91). Then, the platform's visibility scores above average (3.59). Following that is the platform's accessibility scores that was below average (3.44). the platform format and layout have a good interactive level with score above average (4.18). The platform's content can increase language learning motivation received above average perspective (3.73). The platform's adaptability and interactivity also depict above average indicators respectively (3.97). Learning Capacity and content review also indicates above average perspective with the score 4.03 and 4.09 consecutively. In terms of pedagogic coherence, the users reveals that the level of perspective is quite high with the level of 4.32.

5 Conclusion

An ideal online language learning platform development is one that not only supports learners in meeting their language learning objectives, but also one that engages because it is fun, structured, social, and personalized. In terms of usability, the platform being developed was considered to have average quality. The platform's usability score was 58.20. A significant, yet it has plenty of rooms for improvement. We also asked the usability of the platform by asking users some questions on that point. Users stated that they liked some practicality and the easy-to-use features of the platform, yet complained about some other features, such the lack of variation in the activities.

In terms of quality of the platform, this study indicated that the highest score was in pedagogic coherence and the lowest score was in the accessibility. The platform needs to improve its appearance and performance to create better engagement from users, i.e., by fixing its visibility, interoperability, and interactivity. By improving these areas, the platform should be able to improve the motivation of the users in using the platform, which was also low. Further studies can also be done to test the performance of the platform from other angles, such as content development issues from the non-technical aspect.

References

1. C. Blanco, Duolingo Language Report. Available https://blog.duolingo.com/2021-duolingo-language-report/, (2021)

2. Meticulous Research Online language learning market to reach 21.2 billion by 2027. Available https://www.globenewswire.com/news-release/2021/03/18/2195339/0/en/Online-Language-Learning-Market-to-Reach-21-2-billion-by-2027-Exclusive-Report-by-Meticulous-Research-Covering-Emerging-Growth-Factors-Latest-Trends-and-Forecasts-and-Pre-and-Post-.html on May 8, 2021.

3. L. Ceci, Duolingo: annual revenue 2019-2021 Retrieved from https://www.statista.com/statistics/1247833/annual-duolingo-revenue/, 2022.

4. C. H. Karjo, and W. Andreani, Learning foreign languages with Duolingo and Memrise. In Proceedings of the 2018 International Conference on Distance Education and Learning, pp. 109-112, (2018)

5. N. Cong-Lem, Web-based language learning (WBLL) for enhancing L2 speaking performance: A review. Advances in Language and Literary Studies, 9(4), 143-152, (2018)

6. S. Saud, M. Basri, and A. A. Ramly, and AA. Patak, Using online resources technology for foreign language learning: strategies, impact, and challenges. International Journal on Advanced Science Engineering and Information, 10(4), 1504-1511, (2020)

7. R. R. Simanjuntak, U. P. Dewi, and I. Rifai, Gender-related model for mobile-based learning. In IOP Conference Series: Earth and Environmental Science **126**, No. 1, p. 012060, IOP Publishing.

8. S. Krashen, Principles and practice in second language acquisition, (1982)

9. L. S. Vygotsky, The collected works of LS Vygotsky: The fundamentals of defectology **2**, Springer Science & Business Media, (1987)

10. L. J. Yeon; K. J. Young; Y. S. Ju; K. Y. Soo; K. H. Yeon; K. J. Hyun; K. Sohye; P. J. Ha; H. J. Soo; K. Siye; K. Hyerim."Development and Usability of a Life-Logging Behavior Monitoring Application for Obese Patients". Journal of Obesity & Metabolic Syndrome. 28 (3): 194–202, (2019)

11. H. Dehghanzadeh, H. Fardanesh, J. Hatami, E. Talaee, and O. Noroozi, Using gamification to support learning English as a second language: a systematic review. In Computer Assisted Language Learning. Routledge. https://doi.org/10.1080/09588221.2019.1648298 , (2019)

12. S. Kibonggsong, and B. Johnnburton, Advances in Game-Based Learning Gamification in Learning and Education Enjoy Learning Like Gaming. http://www.springer.com/series/13094, n.d.

13. N. Jenkins, Meet 17-Year-Old Indonesian Rapper Rich Chigga | Time. Available,

https://time.com/4532948/brian-imanuel-rich-chigga-jakarta-indonesia-rap-hiphop/ , (2016)

14. J. M. Dewaele, Personality: Personality traits as independent and dependent variables. In Psychology for language learning (pp. 42-57). Palgrave Macmillan, London, (2012)

15. Y. Udjaja, Gamification Assisted Language Learning for Japanese Language Using Expert Point Cloud Recognizer. International Journal of Computer Games Technology, (2018) https://doi.org/10.1155/2018/9085179

16. Chee Siang Ang, P. Zaphiris, and S. Wilson, "Computer Games and Sociocultural Play: An Activity Theoretical Perspective," *Games and Culture*, vol. 5, no. 4, pp. 354–380, Oct. 2010, doi: https://doi.org/10.1177/1555412009360411.

17. E. C. Marion, C. Sobarna, J. A. Rumeser, and R. L. Ardiati, Forms of Address to Superior and Organizational Culture in Japan Creative Industry. Opción: Revista de Ciencias Humanas y Sociales, (93), 79-95, (2020)

18. A. Iglesias, L. Moreno, P. Martínez, and R. Calvo, Evaluating the accessibility of three open-source learning content management systems: A comparative study. Computer Applications in Engineering Education, 22(2), 320-328, (2014)

19. M. Shortt, S. Tilak, I. Kuznetcova, B. Martens, and B. Akinkuolie, Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020, Computer Assisted Language Learning, (2021)

20. J. R. Lewis, The system usability scale: past, present, and future. International Journal of Human–Computer Interaction, 34(7), 577-590, (2018)

21. A. Bangor, P. T. Kortum, and J. T. Miller, "An Empirical Evaluation of the System Usability Scale," *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, Jul. 2008, doi: https://doi.org/10.1080/10447310802205776.