Comparing smartphone and tablet dictionary apps: The impact of dictionary interface

Toshiko Koyama

1. Background

Evolution of mobile technology has rapidly changed the learning environment in the last two decades. New gadgets such as smartphones or tablets have edged into L2 classrooms, and traditional audio-visual aids have been replaced by such mobile gadgets. The Institute for Information and Communications Policy (IIPC) reports the transition of information technology related to various equipment and infrastructure. The White Paper by IIPC summarizes the present status and challenges of utilization of digital technologies, and reports that digitalization in educational institutions accelerated by the COVID-19 pandemic in particular. When we direct our attention to personal lives, it reveals that Japanese use smartphones much more than personal computers as ICT devices in almost all generations. The paper says that 93.5% of Japanese between the age of 20 to 29, and 85.6% of Japanese between the age of 13 to 19 own smartphones respectively.

The use of dictionaries is no exception. As was mentioned in Koyama (2015), the appearance of a pocket E-dictionary was an epoch-making event around the turn of the 21st century in Japan. The convenience has been increasing sales of the dictionary since 2002, however, the advent of smartphones totally changed the situation. The shipments of pocket E-dictionaries are actually decreasing every year at present (JBMIA, 2022).

Koyama and Yamanishi (2019) investigated the transition in dictionary use from 2015 to 2018 with 291 of 1st year Japanese college students. According the results of the survey, the ownership rate of smartphones was almost 100%, and there was a noticeable trend toward earlier start using them each year. In addition, the number of the students who have used the web-based translation functions like Google Translate with their smartphones has increased year by year, reaching over 80% in 2018. On the other hand, it reported that pocket E-dictionaries still have deep-rooted popularity among middle school students in Japan, and approximately 50% students in the study responded that they mainly used them when performing reading and writing tasks for L2 learning.

2. The Previous Studies

Ma (2019) investigated how Hong Kong university L2 learners make use of dictionary apps. She reported that dictionary apps have become the most used learning tools among all only educational technologies. Then, what is a desirable dictionary interface for Japanese L2 learners? Some studies have been conducted to focus on the interface design between a pocket E-dictionary and a paper dictionary (e.g., Koyama & Takeuchi, 2003; 2004; 2005; 2007). Their research has confirmed that a pocket E-dictionary promoted learners' look-up frequency more than a paper dictionary did, and could reduce the time for FL reading. Yet, despite these advantages, they claimed that this higher look-up frequency by a pocket E-dictionary seems not to necessarily guarantee better reading comprehension nor retention of looked-up words. Koyama and Takeuchi concluded that the interface design of a paper dictionary might lead to higher word retention. This assertion was based on "the involvement load hypothesis" by Laufer and Hulstijn (2001), in which an elaborate process for acquiring new lexical information leads to higher retention.

On the basis of the findings, Koyama (2016) compared a pocket E-dictionary with smartphone dictionary apps in terms of look-up behavior, learning outcomes, and dictionary interface design. In the study, smartphone dictionary apps and pocket E-dictionaries the participants possessed were used, and both dictionaries included *Taishukan's Genius English-Japanese Dictionary* (4th edition) by which the participants were instructed to perform their assigned tasks in the experiment. This was to ensure that the conditions of the experiment should be the identical, and to especially focus on the dictionary interface design. She found that; 1) the time to complete the assigned task in using smartphone dictionary apps was the same as that of a pocket E-dictionary; 2) the number of lookups in using the pocket E-dictionary was the same as that of smartphone dictionary apps; 3) the retention of the looked-up words in using smartphone dictionary apps was almost the same as that of a pocket E-dictionary. Based on the findings, she concluded that the difference in interface design between pocket E-dictionaries and smartphone dictionary apps does not directly have an influence upon L2 learners' lookup behavior and the effect on their learning. However, the responses to the questionnaire in the study showed that the participants did not always prefer smartphone dictionary apps to a pocket E-dictionary, which resulted in an implication that the interface design of dictionaries might be somewhat an incentive to L2 learning. Even though current young L2 learners were skillful at smartphone use, they showed some preferences for a pocket E-

dictionary in terms of the size of screen display and a physical keyboard.

3. Purposes of the Present Study

The primary objective of the study, therefore, was to explore the potential to enhance language learning using mobile devices. As one means to this end, the present study examined how the differences in interface design between a smartphone and a tablet-based dictionary apps affected their look-up behavior, learning outcomes, and their perceptions of the dictionary apps.

4. Methodology

4.1. Participants

The participants were 36 undergraduate students (21 males and 15 females, aged 21 to 22 years) whose majors were Education or Social Sciences. They were heavy smartphone users, and were familiar with tablets. All of them were enrolled in teacher training course, and planning to obtain an English Teaching License after graduation. Based on the result of a 45-item cloze test conducted in advance, their English proficiency levels were considered to be ranging from intermediate to lower intermediate levels (M = 20.72, SD = 3.19). Although almost all the participants have used free dictionary apps such as Google Translate or Weblio for their daily L2 learning up to that time, they had to purchase the paid dictionary apps for the experiment beforehand. Therefore, they were paid for their participation in the end.

4.2. Dictionary Apps and Tasks assigned

The same dictionary apps, *Sanseido's Wisdom English-Japanese Dictionary* (3rd edition), was used for both smartphones and tablets. The *Wisdom,* just like the *Genius,* is one of the most popular learner dictionaries in Japan, and are sold in three types such as printed, equipped in a pocket E-dictionary, and apps.

To assess their learning outcomes while looking up dictionaries, two kinds of quizzes were used including ten word definition questions and five reading comprehension questions based on each text respectively. Given the participants' proficiency, these questions were chosen from the 2nd grade test of *Eiken* of 2003 and 2004, a well-known English proficiency test in Japan, which were regarded including several words and phrases that were judged to be unfamiliar to the participants. The similar topics for each

text were carefully selected, and each readability was approximately the same level as shown in Table 1.

Table 1
Readability of the texts

Task	Flesch Reading Ease	Flesch-Kincaid Grade Level	Word Count	Paragraph Count
text A	63.4	9.3	370	4
text B	65.0	8.7	360	4

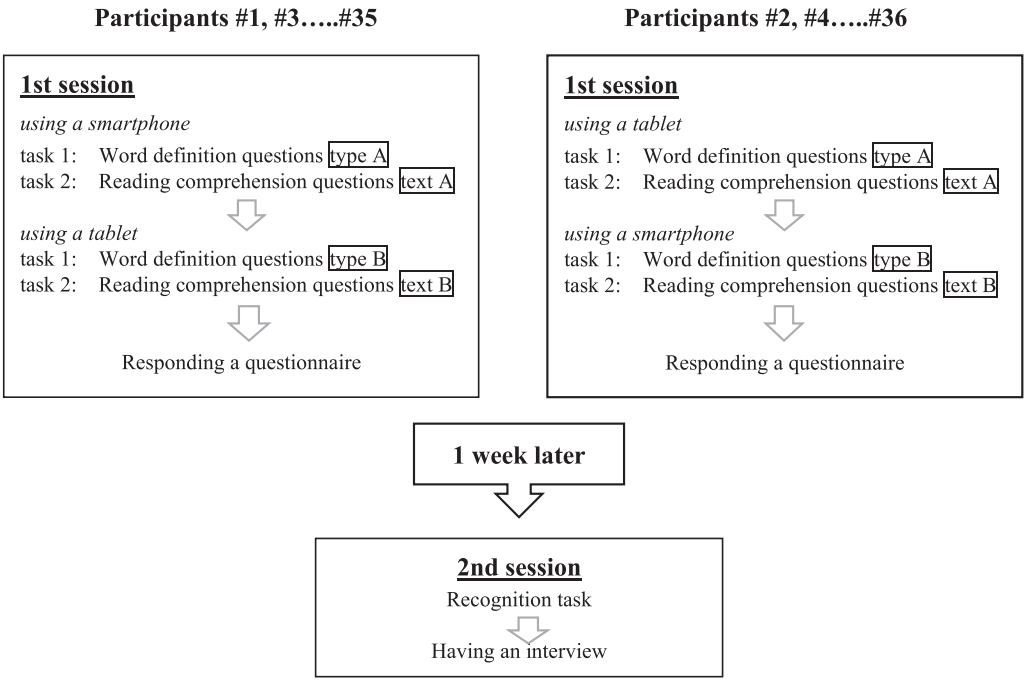
4.3. Procedure

To focus on the differences in both interface design and the effects on learning, the same procedure in Koyama (2016) was adapted. Before the experiment, all the participants were given the explanation and consent forms, and the author obtained their written approval for the experiment. They were given sufficient time to get used to the dictionary apps.

In the first session of the study, they were assigned a word definition and a reading comprehension tasks using either English-Japanese dictionary apps on the smartphone or the tablet respectively. The time they needed for the tasks, the number of their lookups, and their quiz scores were compared. After the both tasks were completed, they answered the questionnaire about each dictionary interface. In the second session, which was held on a week after the first session, a recognition test was conducted to investigate how much their looked-up words were retained. Also, the participants were interviewed for their impressions on each dictionary interface. As can be seen in Figure 1, the combination of the use of the gadget (a smartphone or a tablet) and two kinds tasks were properly balanced.

The entire session lasted about 120 minutes. The experiment was carried on individually during the 2017 academic year so that the participants worked at their own pace. Given the number of the participants in the experiment, the non-parametric Wilcoxon signed-ranks test was adopted to analyze the collected data.

Figure 1
Procedure of the Experiment



5. Results and Discussion

5.1. Lookup behavior

Table 2 compares the time they completed the assigned tasks and the number of looked-up words for the tasks in the first session in two conditions. Although the mean value of the time needed seems to show some differences, no significant difference was found in the time to perform the tasks in either condition at the .05 level in the Wilcoxon signed-ranks test ($p = .07963$ with a small effect size, $r = .207$). However, the difference in the number of lookups between two conditions was statistically supported as shown in the table ($p = .04119$ with a small effect size, $r = .241$). We, therefore, can claim that the participants could look up more words in a shorter period of time with tablet dictionary apps when performing the word definition and the reading comprehension tasks.

Table 2
The Comparison of Variables of Look-up Behavior

	Smartphone			Tablet		
	M	*SD*	*median*	*M*	*SD*	*median*
Time to perform tasks (min.)	43.33	12.04	44.0	40.44	10.72	39.5
The number of lookups	33.72	14.78	31.5	37.56*	18.05	33.0

*$p < .05$

5.2. Learning outcomes

Table 3 shows the variables for ten word definition questions (Task 1), five reading comprehension questions (Task 2), and total scores in the first session. One point was given to each correct answer, with full marks being ten, five, and 15 respectively. The difference in Task 1 between smartphone use and tablet use was not found ($p = .381$ with a small effect size, $r = .103$). Nevertheless, the reading comprehension scores in Task 2 revealed a significant difference between two conditions ($p = .0172$ with a small effect size, $r = .281$). In short, the results suggested that the participants performed the reading comprehension task better using smartphone apps. This might be attributed to the type of each task. Since Task 1 was composed of ten word definition questions, the participants might have answered by simply looking up the unknown words in each dictionary app. In Task 2, however, they had to read and comprehend each text while doing the look up, and thus, this process might have affected the differences of the outcomes. Accordingly, more detailed data will be needed to make a decisive assertion.

Table 3
The Comparison of Variables of the Quiz Scores

	Smartphone			Tablet		
	M	*SD*	*median*	*M*	*SD*	*median*
Task 1 (Word Definition Questions)	8.53	1.52	9	8.33	1.71	9
Task 2 (Reading Comprehension Questions)	2.67*	1.20	3	2.14	1.15	2
Total (Task 1 + Task 2)	11.19*	2.24	11	10.47	2.13	11

*$p < .05$

In grading the recognition task, one point was given if the participants marked the words they had actually looked up in the first session. Comparing the rate in Table 4, the mean value was especially larger when using smartphone apps. This difference was statistically supported by the result of Wilcoxon signed-ranks test ($p = .03603$ with a small effect size, $r = .247$). This indicates that the looked-up words using the smartphone apps resulted in better retention than those using a tablet, even though the participants looked up more

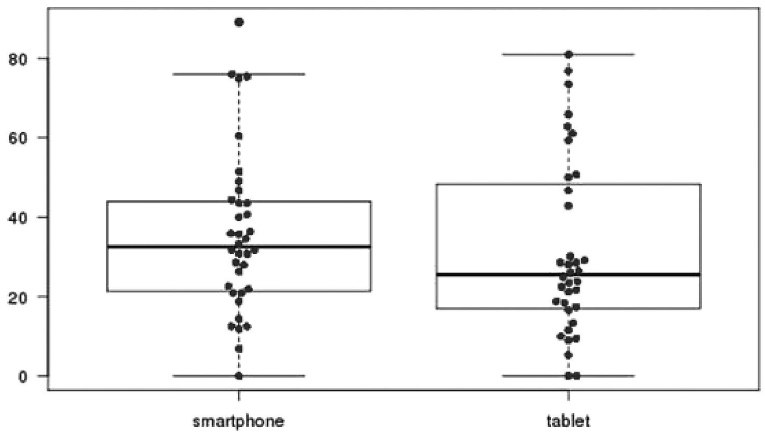
words in a shorter period of time with tablet dictionary apps. Nevertheless, in addition to the relatively small effect size, the data in the box-and-whisker plot (see Figure 2) shows a wider distribution range when the participants used the tablet dictionary apps. Consequently, further examination will be considered necessary.

Table 4
Results of Rate of Recognition

	Smartphone			Tablet		
	M	*SD*	*median*	*M*	*SD*	*median*
Rate of recognition (%)	35.61*	20.42	32.55	31.54	22.22	25.55

*$p < .05$

Figure 2
Box plots with individual data points



5.3. Learners' perception and preference

After completing the first session of the experiment, a 16-item questionnaire was administered to assess their perception of the two types of dictionary apps. In the questionnaire, the participants rated their impressions on a scale of one to five. Table 5 represents some excerpts from the questionnaire items which showed considerable differences between the two dictionary apps.

Table 5
Some Excerpts from Responses to the Questionnaire

Item #	Questionnaire Items	Smartphone app	Tablet app
1	This dictionary provided me with much information at first sight.	3.2	4.6
6	This dictionary is easy to see lexical information when looking up.	3.2	4.7
12	I enjoyed using this dictionary.	3.3	4.4

The questions were originally given in Japanese.

As shown in Table 5, overall, the tablet apps received a higher evaluation from the participants in terms of the dictionary interface. It is worth noting that there is no remarkable difference in the ratings for questionnaire item #13, *"I would like to use this dictionary app again when I perform the similar task as this* (smartphone app; 3.8, tablet app; 4.1)*,"* although they gave a favourable evaluation of the tablet app in #12, *"I enjoyed using this dictionary."* It appears that the participants did not much care about which dictionary apps would be available when performing this kind of L2 task. This interpretation is supported by their response to the item #15, *"I would like to continue using this app after graduating from the college,"* in which their rating was the same.

Table 6
Some Excerpts from Participants' Feedback

Participant #	Feedback
10	I used the keyboard of a tablet, and used flick input when using the smartphone. Either will work for me.
11	I prefer a paper version because a liquid crystal display makes my eyes tired.
14	I was a bit confused because lexical information was shown all at once when using tablet app.
19	I am used to use a physical keyboard like PC. It takes a great deal of trouble for me to input by flick operation when using a smartphone.
20	I feel that the smaller screen of the smartphone makes lexical information condense so that I could focus on the information I wanted to see.
32	I have been using a paper dictionary since I was a junior high school student. I always made notes on the dictionary pages, so I would recommend a paper dictionary for beginners.

The feedback was originally given in Japanese.

After the recognition test in the second session, the participants commented on each dictionary app while reviewing the tasks they had performed a week before.

Some comments in terms of dictionary interface design are shown in Table 6. The responses on the questionnaire seems to show that the participants regarded the tablet dictionary app as appropriate for L2 learning compared to the smartphone app mainly due

to its screen size. Interestingly enough, some of their feedback indicated that they did not necessarily consider the tablet dictionary interface as desirable for EFL learners. This contradictory feedback should be further investigated with a more extensive survey.

6. Concluding Remarks

Since the study was one of the first attempts to compare the lookup behavior and learning outcomes between two types of dictionary apps focusing on the dictionary interface, there are some limitations. One is the number of participants was comparatively small, and the other is the effect sizes of the collected data were small.

With those limitations in mind, let me summarize the findings in the study. First, the students looked up more words in a shorter period of time with tablet dictionary apps when performing the task assigned in the experiment. Second, they performed the reading comprehension task better when using smartphone dictionary apps. In addition, the looked-up words using the smartphone apps resulted in better retention than those with a tablet one. Third, the tablet apps received a higher evaluation from the students mainly due to its screen size. It is noteworthy, however, that despite the high evaluation, the students did not necessarily consider the tablet dictionary interface as desirable for EFL learning.

One important thing that should be clearly stated here is the data collection in the study was made during the 2017 academic year. Owing to the COVID-19 pandemic since 2020, the EFL learning environment in Japan has changed considerably, and the gadgets that L2 learners utilize might be different. In fact, Koyama and Yabukoshi (2022; 2019) reported in their two-year research that free smartphone dictionary apps (i.e., Weblio and Google Translate) have remained popular and pocket E-dictionaries have become less popular. Surprisingly enough, the proportion of no dictionary users has slightly increased in two years. Consequently, more quantitative and qualitative studies should be conducted to explore the impact of the dictionary interface design on L2 learning.

*This article is a revised version of the paper presented at the 5th WorldCALL Conference held in Concepcion, Chile from 13 to 16 November in 2018.

Acknowledgements

This research is partially supported by the Grant-in-Aid for Scientific Research (C) 19K00777 in the fiscal years from 2019 to 2021. The author would like to express her gratitude to Professor, Atsushi Mizumoto of Kansai University and Professor Emeritus, Thomas Robb of Kyoto Sangyo University for their insightful comments on earlier version of this article.

References

IICP. White Paper in 2022. Tokyo: Institute for Information and Communications Policy, 2022. https://www.soumu.go.jp/johotsusintokei/whitepaper/

Koyama., T. (2016). How do mobile technologies affect learning environments? *LEXICON 46.* 27-44.

Koyama., T. (2015). The impact of E-dictionary strategy training on EFL class. *Lexicography ASIALEX 2.* 35-44 DOI: 10.1007/s 40607-015-0018-3.

Koyama, T. & Takeuchi, O. (2003). Printed dictionaries vs. electronic dictionaries: A pilot study on how Japanese EFL learners differ in using dictionaries. *Language Education & Technology, 40,* 61-79.

Koyama, T. & Takeuchi, O. (2004). Comparing electronic and printed dictionaries: How the difference affected EFL learning. *JACET Bulletin, 38,* 33-46.

Koyama, T. & Takeuchi, O. (2005). Does an assigned task result in better retention of words?: Two empirical studies on hand-held electronic dictionaries. *Language Education & Technology 42,* 119-132.

Koyama, T. & Takeuchi, O. (2007). Does look-up frequency lead to better comprehension of EFL learners?: Two empirical studies on electronic dictionaries. *Calico Journal. 25,* 110-125.

Koyama., T. & Yabukoshi, T. (2019). University L2 learners' dictionary use in multiple-choice vocabulary quiz: a pilot study. *The Bulletin of Education, Osaka Ohtani University 45.* 17-27.

Koyama., T. & Yabukoshi, T. (2022). The effect of dictionary/app usages in M/C vocabulary task. *LEXICON 52.* 1-16.

Koyama, T. & Yamanishi, H. (2019) A transition in the use of dictionaries by Japanese college students; the effects of the revision of authorized English text books. *The Abstract book of the 45th JASELE national conference in Hirosaki, Aomori.* 472-473.

Laufer, B. & Hulstijn, J. H. (2001). Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics, 22*(1), 1-26.

Ma, Q. (2019). University L2 learners' voices and experience in making use of dictionary apps in mobile assisted language learning (MALL). *International Journal of Computer-Assisted Language Learning and Teaching, 9*(4), 18-36.

Statistics in 2022 by Japan Business Machine and Information System Industries Association. https://www.jbmia.or.jp/whatsnew/topics_detail.php?id = 1555

Appendix

Text A

1. 次の（1)～(10）までの（　　）に入れるのに最も適切なものを一つ選んで、その数字を○で囲みなさい。辞書で確認した語は、○で囲むこと。

1)　A: Did you hear that Susan is in the hospital? She fell off her bike and broke her arm.
B: How (　　)! We should go visit her.
1. awful　　2. empty　　3. needless　　4. jealous

2)　A: Here's the 20 dollars I (　　) you. Thanks again for lending it to me.
B: No problem, Bill.
1. owe　　2. blame　　3. doubt　　4. thrill

3) There is a growing (　　) in many countries for governments to protect the rights of nonsmokers.
1. vice　2. tendency　3. compromise　4. bargain

4) I couldn't (　　) on reading because the children kept making so much noise.
1. fasten　2. operate　3. depend　4. concentrate

5) Yesterday, a storm hit a coastal town in Chiba, (　　) damaging many houses there.
1. properly　2. frequently　3. quietly　4. severely

6) Debbie enjoys her new job very much. Her (　　) are friendly and the working conditions are good.
1. colleagues　2. citizens　3. patriots　4. substitutes

7) Brad took a beautiful photograph showing the (　　) of trees in a lake.
1. reflection　2. inspiration　3. extinction　4. distinction

8) The doctor told me that I can't play any sports until my injury (　　) completely.
1. heals　2. improves　3. transforms　4. converts

9) A: Do you know how far Paris is from Berlin?
B: No but we can look it up in my (　　).
1. circuit　2. compass　3. atlas　4. ruler

10) A: What time should I visit you this evening?
B: About seven, if that (　　) you.
1. follows　2. suits　3. matches　4. fits

2. 次の英文を読んで（1）から（5）までのそれぞれの英語に続けるのに最も適切なものを選び、数字を○で囲みなさい。辞書で確認した語は、○で囲むこと。

Overcrowding on highways in the United States has long made driving to and from work a slow and frustrating experience for many Americans. About 30 years ago, government agencies in the United States introduced high-occupancy vehicle (HOV) lands on some highways to try and reduce traffic levels. These lanes were reserved for cars with three or more people. Because fewer cars used HOV lanes, traffic in those lanes moved faster than in the other lanes/ The government's intention was to reduce the amount of gasoline being used by having friends and co-workers travel together to work.

Soon after HOV lanes were set up on some highways between Washington, D.C., and its suburbs, however, something unexpected began to happen. Single drivers started picking up strangers at bus stops in order to make use of the faster lanes. This practice, which became known as slugging, gradually grew in popularity. These are now a number of different routes with specific pick-up and drop-off points where "slugs" can be seen waiting patiently in line for a ride. Slugging is not actively promoted by the government, but some individuals have set up websites that give details of slug routs and invite more people to get involved.

Over time, a number of rules have developed that are designed to make slugging safer and more pleasant. On arrival at the pick-up point, the driver first calls out his or her destination rather than asking where the slug wants to go. Slugs can refuse a ride if they are suspicious of the driver or if taking the ride would mean leaving a woman waiting alone. In the car, there should be no taking unless the driver begins the conversation, and no money should be exchanged. At the end of the ride, the driver and the slugs all say "thank you."

Everybody benefits from slugging. Drivers who pick up slugs are able to get to work faster. And by

not taking their cares into town, slugs save money on fuel and parking. Slugging also reduces the number of cars on the road. So, as many commuters in Washington D.C., now know, taking a slug to work is a great way to help improve the environment.

(1) In the United States about 30 years ago,

A) driving to work not a slow and frustrating experience.
B) lanes were created for cars carrying at least three people.
C) government agencies wanted more cars to use the highways.
D) government agencies built new highways to reduce traffic levels.

(2) What happened soon after HOV lanes were introduced?

A) Drivers began giving rides to people waiting at bus stops.
B) Websites were developed to give information about bus routes.
C) The government set up a number of pick-up points for slugging.
D) More people began commuting to work by bus instead of by car.

(3) According to the rules of slugging,

A) drivers may charge for the ride if the slug does not say “thank you.”
B) slugs should first tell the driver where they want to go.
C) slugs should not refuse a ride if the driver is a woman.
D) conversations should only by started by the driver.

(4) The practice of slugging

A) means that drivers now spend less on parking.
B) saves people money and helps the environments
C) allows drivers to travel longer distances to work.
D) has enabled the government to build more bus stops.

(5) Which of the following statements is true?

A) HOV lanes were not actively promoted by the government.
B) HOV lanes were introduced to make highways in the U.S. safer.
C) Drivers who pick up slugs are able to reach their destinations more quickly.
D) Slugging was started by individuals who no longer wanted to work in Washington, D.C.

From the 2nd Grade test of EIKEN, 2004