## An Evaluation of Chinese Vocabulary Learning Apps for Self-access Extracurricular Learning

XIAO, Jingru\* ZHAO, Minming

Abstract

Mobile applications are becoming more and more common as educational tools to help the study of a second language outside the classrooms. Chinese vocabulary learning apps, as a particular type that accounts for a large share, have not only highlighted the importance of vocabulary learning in learning Chinese as a second language (CSL) field, but also represented the enormous potential of mobile technology in the area of self-access vocabulary acquisition. However, there hasn't been much done to analyze these applications, especially from the viewpoints of the learners. In addition, few studies have looked into the practical experience of using these apps.

Based on a needs analysis of Chinese learners, the current study designed a checklist for evaluating Chinese vocabulary learning apps and used them to assess six apps for self-access extracurricular learning. The criteria covered the following areas: content quality, pedagogical coherence, feedback and self-correction,

<sup>\*</sup> XIAO, Jingru, School of humanities and Social Science, The University of Science and Technology, HK. Email: jxiaoaq@connect.ust.hk (corresponding author)

ZHAO, Minming, School of humanities and Social Science, The University of Science and Technology, HK. Email: mzhaoal@connect.ust.hk

motivation, usability, customization, and sharing. Applying the checklist revealed that half of the selected apps lacked sharing and motivation mechanisms, and also performed poorly in the feedback system. The applications performed best in the category of content quality.

An app called trainchinese was chosen by getting the highest score on the checklist, and then was experienced by three volunteers. Results showed that even though the selected app is relatively in line with learners' psychological expectations, it can only partially meet the needs of learners for Chinese vocabulary learning in practical experience. Additionally, learners' different Chinese levels and learning styles are also big interfering factors. The findings of this paper provide more information for Chinese teachers and Chinese learning app designers to pay attention to the needs of self-access learners. However, the evaluation standards based on actual needs need to be further mined and supplemented by future research.

*Keywords:* learning Chinese as a second language, vocabulary learning, mobile applications, evaluation, self-access language learning

#### **1. INTRODUCTION**

The number of Chinese learners worldwide is steadily increasing. According to the statistical data from the Ministry of Education of the People's Republic of China (2021), by the end of 2020, more than 180 countries and regions around the world have carried out Chinese education, resulting in over 20 million Chinese learners. In the developing Chinese learning upsurge, how to expand Chinese vocabulary is becoming a hot topic in the research of Teaching Chinese as a Second language (TCSL). There is a fact that vocabulary is a crucial fundament of language acquisition (Grabe, 2004). It helps people communicate more smoothly and precisely and can improve all language abilities, including listening, speaking, reading, and writing (Schmitt, 2000).

Particularly, Chinese vocabulary learning directly affects the ability of Chinese learners to use the target language for normal communication, and determines the overall level of their Chinese learning (Xu, 2018). Lu (1998) once pointed out that "to learn Chinese well, it is important for a foreign student to master a large number of vocabulary and have sufficient vocabulary storage ". Chen (1999) further argued that if learners want to handle Chinese communication within two years, their vocabulary quantity should reach at least 20,000. A "vocabulary-centered theory" was put forward by Li and Yang (2004), emphasizing that pronunciation and grammar teaching should be included in vocabulary teaching. Since words can be broken down into morphemes, meanwhile extended upward into phrases and sentences, mastering grammar knowledge is also inseparable from the center of vocabulary. Chinese vocabulary itself is characterized by the integration of pronunciation, semantics, grammar, pragmatics, and cultural elements. Chinese word formation is the carrier of national culture, and Chinese grammar expresses and limits sentence meaning through word order and function words. Thus, it can be seen that learning Chinese vocabulary well is the basis for mastering Chinese grammar rules, finally mastering the Chinese language. This is one of the rationales behind this study's emphasis on acquiring

Chinese vocabulary.

At the same time, technology is able to make contributions to the mentioned fields above under the modern background. It makes learning more enjoyable, flexible, and heuristic, while also increasing productivity and efficiency (Halverson & Smith, 2009). Mobile Assisted Language Learning (MALL), which represents a fresh approach to second language instruction, allows anybody to learn without regard to location or time limitations. Second language learners can easily access a variety of apps to help with their independent extracurricular learning. (Gangaiamaran & Pasupathi, 2017). Since urban and rural areas frequently have different levels of internet access and digital literacy, which may be larger than the gap between nations. Mobile phones are being utilized for web access because wired connections are frequently unavailable outside of metropolitan areas. Online language learning is now accessible to previously underserved communities thanks to the sharp decline in the price of smartphones, particularly Android devices, and the growing capabilities of mid-range handsets. This will probably continue to have a significant impact on non-formal language acquisition and literacy education (Godwin-Jones, 2017b). Chinese vocabulary learning is no exception. Relevant mobile applications have now become a new method of self-access Chinese vocabulary learning. This article focuses on the free Chinese vocabulary learning apps for Android and IOS systems, trying to explore more possibilities for Chinese learners to use mobile applications for self-access vocabulary learning.

A growing corpus of research has investigated the use of mobile apps for vocabulary acquisition, (e.g., Sweeney & Moore, 2012; Lin et al., 2022), however, without extensive study on the benefits and limitations of apps for second language learning. It leads to a challenge to strategically and effectively incorporate these learning tools into planning and learning activities (More & Travers, 2013). As for self-access Chinese learners, the evaluation criteria based on their needs will help them better choose suitable learning apps. We have found gaps in the literature in evaluating Chinese vocabulary learning apps, related to four aspects: 1. most

of the evaluations are from the perspective of teachers or application developers (e.g., Chen, 2016; Rosell-Aguilar, 2017), while few are from the perspective of learners, with learners' needs as the first-hand data. 2. the existing research refers to the evaluation of Chinese learning apps mostly selecting comprehensive ones as the evaluation object (e.g., Neumann et al., 2019), with a lack of the evaluation framework designing for apps focuses a certain learning goal or language skill. There is also a lack of evaluation research specifically on Chinese vocabulary learning apps. 3, when evaluating a certain number of apps, previous research designed pre-use evaluation that is not based on real user experience, with just a theoretical framework (e.g., Chen, 2016; Gangaiamaran & Pasupathi, 2017). Other studies turned to the in-use evaluation, on the other hand, usually pay attention to one specific app selected based on the researcher's own views rather than a scientific evaluation (e.g., He & Zhu, 2018; Shi, 2021). 4. The majority of CSL researchers tend not to distinguish between apps designed for classrooms and for extracurricular learning. The participants are always foreign students from a certain university (e.g., Shi, 2021), ignoring the concept of "self-access learners".

Within published studies, Chen (2016) outlined seven essential criteria for evaluating second language learning apps. Neumann et al. (2019) then expanded these standards and designed a checklist to help preschool educators distinguish which types of children's second language learning applications may be the most effective. This study inherits the evaluation criteria of Chen (2016) and extends the checklist design of Neumann (2019). The purpose is to design a new evaluation checklist for Chinese vocabulary learning apps based on the actual investigation and analysis of the needs of Chinese learners. Furthermore, this study applies this list to actually evaluate some existing Chinese vocabulary learning apps (through certain screening criteria) and invites Chinese learners to experience the best performing apps in the list to explore the usefulness and evaluation effect. According to Tomlinson (2013), the function of pre-use is prediction. Still, inuse is more objective and reliable, so this study chooses to evaluate Chinese

vocabulary learning apps using both pre-use and in-use evaluation. The pre-use evaluation will use the checklist to evaluate the Chinese vocabulary learning apps selected from the app market, and the strengths and weaknesses of each app can be viewed through the scores for each rating criteria. The in-use evaluation will invite participants to experience a Chinese vocabulary learning app with the highest scores in the pre-use evaluation stage to obtain the second language learners' experience of using the app and make the evaluation more objective and reliable. The whole process will be taken in four steps: First, a questionnaire was conducted to explore the needs of self-access Chinese learners for ideal Chinese vocabulary learning apps. Second, a checklist was designed according to the questionnaire results and the past material evaluation framework. Third, existing apps on the market will be screened and scored through this checklist, to find out their benefits and limitations according to the learners' needs, and the one with the highest score will be selected as the object of in-use experience. Finally, A small group of volunteers will describe their feelings after experiencing the app for a period, to identify to what degree the app may meet their needs.

## **2. LITERATURE REVIEW**

#### 2.1 Needs analysis

The process of acquiring information about the needs of a specific customer group in industry or education is referred to as Needs Analysis (NA). Specifically, needs analyses in educational programs naturally focus on students' learning needs, which serve as the foundation for future creation of teaching materials, learning activities, exams, program evaluation techniques, and so on. (Brown, 2009). In this study, the purpose of doing NA is to inquire the information about what Chinese learners want toward mobile assisted self-study, and to apply it as the basis of evaluation framework design, based on learners' perspectives.

Schutz and Derwing (1981) recommended eight stages in a NA, "which

would appear to represent an absolute minimum for any needs assessment attempt deserving of the term" (p. 35); Jordan (1997) recommended 10 steps, while Graves and Xu (2000) recommended seven. Brown (2009) then categorized them into three stages: 1. Prepare for NA; 2. Conduct NA research; and 3. Apply NA findings. To carry out the first stage, a large number of literatures provided definitions of many sorts of needs as well as assessments of various issues and limits in using this notion. Chinese scholars like and commonly use Hutchinson and Waters' (1987) two definitions of needs: "target needs" and "learning needs." While "learning needs" are concerned with the learner's motivation and attitudes, interests, personal reasons for learning, learning styles, resources, and time available, "target needs" refers to the learner's "necessities," "lacks," and "wants" for operating well in the target scenario. They contended that what constitutes a "necessity" depends on the demands placed on the learner to effectively operate in the target situation. The learner's "lacks" may then be determined by comparing the "target proficiency" and the "existing proficiency," which should also be specified. Furthermore, allowing students to communicate their own expectations in relation to their target scenario requirements is emphasized in Hutchinson and Waters' (1987) concepts. The idea of "wants" is cognitively related to both objective and subjective needs.

In this study, the definition of "needs" was derived from the perspective of self-access Chinese learners, particularly for mobile vocabulary learning. Then the second stage of NA is followed by using a questionnaire as the instrument. The questionnaire design for Chinese learners' needs analysis will mainly refer to the above definitions and classifications. Finally, the results of the questionnaire will be analyzed as part of the basis of the following checklist.

## 2.2 Materials evaluation

Tomlinson (2013a) divided the evaluation of materials into visual and principled evaluations, with the latter one described as "a procedure that involves measuring the value (or potential value) of a set of learning materials. It involves making judgments about the impact of the materials on the people who use them (Tomlinson, 2003b)." Since visual evaluation is "provisional" and "impressionistic" which may not provide professional guidance for materials, principled evaluation is a more systematic, deliberate, and complex process (Mishan & Timmis, 2015). This study is concerned with principled materials evaluation in language learning apps, which is the reference of the checklist design.

Mishan and Timmis (2015) suggested that when using principled materials evaluation, consideration needs to be given when the evaluation will be conducted. In-use evaluations are more "objective and dependable" than predictions in preuse evaluations (Tomlinson, 2013b, p. 32). Considering that this study evaluates Chinese vocabulary learning apps based on learners' needs, no study is conducted using the checklist and then having learners experience the app for further evaluation. We believe it is necessary to evaluate both pre-use and in-use processes.

For principled materials evaluation, a common approach is to develop a checklist associated explicitly with some specific criteria (Mishan, 2015; Yamaguchi A, Okamoto et al., 2019). A large body of literature discussed criteria designed for materials evaluation, primarily for language teaching classrooms (e.g., Sheldon, 1988; Ellis, R., 1997; Tomlinson, 2003). However, a few studies have proposed criteria for a checklist to evaluate language learning applications. Chen (2016) proposed seven evaluation criteria for a checklist of mobile learning apps using social interactionist theory, which is to help evaluate the relevance and cognitive development of language learning processes) and the affective filtering hypothesis (Krashen, 1988) as a theoretical framework. These criteria include content quality, pedagogically coherent, feedback and self-correction, motivation, usability, customization, and sharing. The results of Chen's (2016) evaluation of English learning apps showed that the features each app offered learners varied considerably, suggesting that the design of the apps could all be enhanced. Neumann et al. (2019) extended Chen's (2016) prior work by developing a checklist for evaluating mobile apps to help teachers of preschool

children distinguish between the content and effectiveness of apps developed for young children's second language learning. In addition, it added the criterion of cultural awareness to Chen's (2016) seven assessment criteria. A series of questions are raised about the evaluated apps. If the app can meet a certain requirement, it will get one point in the corresponding item. Neumann uses the added scores to compare the quality of apps in this way.

On the basis of both, this study aims to develop a checklist for evaluating Chinese vocabulary learning apps in the pre-use process. We decided to integrate the seven categories of app evaluation proposed by Chen (2016) based on the theoretical framework and the scoring system in Neumann et al.'s (2019) checklist design. The difference is that these two studies started from the theories of language acquisition, considering the nature of apps from the perspective of educational researchers. In the contrast, we hope to design a checklist based on learners' needs. We still decided to consider the content, technical support, or function of apps from Chen's (2016) framework, but these factors will appear in the form of questions in the questionnaire of this study, asking Chinese learners whether they really need these services. Finally, the attitude of learners' answers will determine whether researchers' considerations would appear in the checklist. (See Table 4.2).

### 2.3 Self-access Language Learning

Untutored second language acquisition is not unusual, with the fact that the majority of language learners across the world learn second languages in naturalistic circumstances (Polat and Kim, 2014, p. 184). It has been described as Self-access Language Learning (SALL), a type of individualized learning in which each student interacts with controlled and/or uncontrolled learning settings in a unique way (Gardner and Miller, 1999). An essential requirement to selfaccess learning, as Sheerin (1989) claims, is "the supply of self-access resources inside an ordered framework so that students may receive what they need." In second language learning, Learner autonomy in foreign language education has revolutionized previous methods in the language classroom and given rise to selfaccess facilities across the world (Little,1991). The emergence of virtual resources is blurring the limits of self-access settings, leading to an expansion of SALL management outside the physical constraints (Gardner & Miller,2011).

The use of technology by learners for self-access learning has been described in the previous literature as "emergent, interest-based, largely receptive, and incorporating a significant amount of accidental learning" (Lai, 2017, p. 50). More responsive technology experiences that are instruction-, information-, and entertainment-focused are used by students. They believe that their vocabulary and listening skills are best developed through casual technology-based learning. They appreciate and utilize audiovisual materials (such as music and online movies) and language learning applications (such as online dictionaries and translators) the most (Trinder, 2016). What we want to talk about in this study is using mobile applications in self-access Chinese vocabulary learning.

A considerable amount of research has stressed the help of mobile applications in autonomous vocabulary learning. Jennifer Betsy Redd's (2011) trial with high school students' vocabulary growth using an app was successful. For these students, she offered an iPod gaming program to help them improve their vocabulary for three weeks. As a consequence, the considerable difference between the pre-test and post-test confirmed that using a mobile device did, in fact, have an impact on high school students' vocabulary growth. Afterward, 80 college students were selected by Ornprapat and Wiwat (2015) to participate in M-learning for vocabulary development. First-year students were split into two groups with 40 each (the control group and the experimental group). The experimental group participated in a variety of vocabulary-building tasks through SMS, whereas the control group completed exercises on paper. Finally, the variation in mean scores convincingly demonstrates the success in vocabulary acquisition among the students in the experimental group. It is found that there is a lack of reviews specifically for self-access Chinese vocabulary learning apps.

## 2.4 Review of Chinese Vocabulary Learning Apps

Chinese vocabulary learning apps have received attention from many researchers in recent years. Chuang (2016) investigated the use of Chinese vocabulary apps in the classroom and provided Chinese language teachers with a selection of strategies for integrating mobile teaching resources in the classroom. Xu (2018) then applied a combination of questionnaires and indepth interviews to investigate the use of Chinese vocabulary learning apps by intermediate-level Chinese learners at Shanghai Foreign Studies University in the process of self-access learning. There are also scholars who evaluate apps under different standards for specific apps. Lin et al. (2018) used Zhang & Lu's (2015) five strategic vocabulary learning categories (learning word form, PIC / IMG, association, repetition, and word lists) to investigate the learning strategies provided by 22 Chinese vocabulary learning apps for users and showed that these vocabulary learning apps might be less likely to provide sufficient scaffolding for users to learn new words. However, we didn't find literature on the study of Chinese vocabulary learning apps based on needs analysis and materials evaluation

### **3. RESEARCH QUESTIONS**

Inspired by the above literature, this study will aim to address three important questions:

(1) What needs should be taken into account when choosing Chinese vocabulary learning apps for self-access Chinese learners?

(2) What are the benefits and limitations of existing self-access Chinese vocabulary learning apps, according to learners' needs?

(3) To what extent does the app (selected in the pre-use evaluation) meet the needs of self-access Chinese learners in an in-use evaluation?

## 4. METHOD & PROCEDURES

A mix-method design was conducted. The researchers gathered and analyzed both quantitative and qualitative data, as indicated below.

#### **4.1 Participants**

In the pre-use stage, we posted the questionnaire link to the Chinese chat corner of the Hong Kong University of Science and Technology (HKUST) and requested a teacher from the Center of Language Education (CLE) of the university to forward the link to her students. At the same time, we contacted our familiar Chinese learners and asked them to fill in the questionnaire. We planned to collect 20 responses at this stage. It is based on the number of Chinese learners that the two researchers were able to access at that time, also because 20 responses are enough to cover a certain number of subjects at different Chinese levels.

However, due to voluntary work, only 9 were received in the end: as a result, 3 were from the participants of the Chinese corner of HKUST, and 5 were from the researchers' familiar Chinese learning friends and their friends. Although the total number seemed not enough to support a scientific checklist design, since the nine responses came from Chinese learners in different countries, meanwhile contained learners at the primary level as well as the intermediate-to-advanced level, we believed that the checklist produced on the basis of this small part of data is still representative. It represented in microcosm what are the general needs of Chinese learners while using mobile Chinese vocabulary learning applications.

At the end of the questionnaire, we recruited volunteers willing to participate in the in-use stage, and four students left their emails. All of the participants are college students, non-native speakers of Chinese, and were born and raised abroad. They all would like to conduct self-access Chinese learning outside of school. There were 3 participants who were eventually willing to join the in-use evaluation and be interviewed. Two of them are Chinese learners at the primary level, and the other is at the advanced level. This study was conducted with the permission of the head of CLE of HKUST, as well as all students participating.

## 4.2 Pre-use evaluation

Quantitative data were collected in the pre-use stage, through a questionnaire and a checklist. The collected data from the former instrument can be used in answering the first research question, meanwhile the latter for the second research question.

#### 4.2.1 Questionnaire design

A questionnaire was designed on the WJX platform to investigate students' possible needs and wants for an ideal Chinese vocabulary learning app. The link to the questionnaire is shown in Appendix 1. The questionnaire needs to serve two purposes: first, to be able to answer the first research question; second, the data from recoveries can contribute to the next stage of app screening and evaluation checklist design. Based on these, the researchers designed different question forms and calculation methods in four general parts, which will be described in detail respectively as follow.

The first part (see Appendix 1, p.44) is a brief introduction to the research project and a simple consent form for providing result data in the form of required questions. The content and research purpose of this project were briefly introduced to give the participants a basic understanding of it. Next, through four required questions with only a "yes" option, the participants can know they have the right to ask questions at any time, complete the questionnaire voluntarily, and all the information will be presented anonymously in the final paper. Only those who choose "yes" to these four questions can enter the main body of the questionnaire, so that all the results were regarded as giving permission to the researchers to analyze data.

The second section entailed looking into students' relevant personal information as well as their educational backgrounds (Appendix 1, p.45-46). It involved the system of mobile devices regularly used by students, the effective

length of students' Chinese study, the experience of students ever used Chinese vocabulary learning apps, etc. The purpose was to determine the criterion of selecting target apps for evaluation, to ensure that each participant has the conditions and can comfortably complete the final in-use evaluation.

The primary body of the questionnaire for participants' needs analysis, which was separated into four pages, is the third section (Appendix 1, p.46-51). The first page investigated the learning objectives of students, as well as their extended vocabulary memorizing goals when using vocabulary learning apps. The next three pages are the key parts that affect the content and scoring system of the checklist, including 49 scale questions. Learners' needs were defined as "What I think an ideal Chinese vocabulary learning app should provide", then followed the seven categories involved in the evaluation framework proposed by Chen (2016), including: 1. Content Quality: content should provide opportunities to advance learners' language skills, with connection to their prior knowledge. 2. Pedagogical Coherence: the skills provided in the app should be consistent with the targeted learning goal. 3. Feedback and self-correction: learners should be provided with feedback to conduct self-evaluation. 4. Motivation: elements are embedded to engage and motivate language learners to use the app. 5. Usability: learners are provided with clearly indicated menus and icons to easily navigate through the app. 6. Customization: learners have their individualized needs met including font size and customizable settings to personalize their learning. 7. Sharing: allowing learners to share their learning progress, issues, or concerns in learning. 49 items were listed asking about small constituents that an app may provide in the form of "I hope the following content/functions can appear in a Chinese vocabulary learning app". The content of each question started from Chen's (2016) seven categories, then determined by the review of existing Chinese vocabulary apps in previous studies as well as the researchers' own experience of using second language vocabulary apps. In addition, at the end of each page, there was an optional blank left so that participants can fill in the content they want but

are not mentioned in this questionnaire, to supplement possible negligence from researchers.

The answer choices for the questions included agree, neutral, and disagree, each worth 0, 1, and 2 points separately. The study's requirement to establish the scoring criteria for app evaluation in the ensuing checklist design is the cause for choice scoring. Maslow (1954) proposed a classification of basic needs into five categories: physiological, safety and security, belongingness, esteem, and selfactualization. It also has a certain reference value for autonomous foreign language learning, that is, the needs of language learners are hierarchical and should not be given equal importance. These importance levels should be determined by learners in the questionnaire as well. This study expanded the simple scoring system proposed by Neumann et al. (2019) and add the consideration of a weight system. The weight of each need was determined by the mean score calculated by all the answers to each question, which represents the extent of desire for different functions or contents. The higher the mean score, the more Chinese learners hope this thing can appear in a Chinese vocabulary learning APP. Therefore, these mean scores would become the basis of app scoring standards in the following checklist design.

The final part of the questionnaire concerned recruiting in-use evaluation volunteers (Appendix 1, p.52). Anyone interested in participating in the next round could leave their email address so that the researchers can provide them with further instructions. A 100 HKD thank-you payment was promised to motivate more volunteers.

Before sending the questionnaire to all the participants accessible, a pilot study was done with one familiar Chinese learner, to check if there are difficult questions or other misunderstandable problems. The data from this volunteer was collected as well. On the whole, the volunteer finished the questionnaire successfully and didn't find big problems, with only one suggestion raised. She thought two questions "I hope the app system can work smoothly" and "I hope the app is able to protect my privacy" are useless, because everyone is willing to choose "yes". However, considering that this questionnaire is designed to find out students' needs used for the follow-up checklist building, and students' choice of questions would determine the scoring mechanism of the checklist. The needs of students could not be assumed from the perspective of researchers. Therefore, the two questions were still maintained in the questionnaire. The link to the final draft of the questionnaire is attached in Appendix 1.

## 4.2.2 Analysis of questionnaire results and question selection

The conclusions drawn from the questionnaire data can be divided into two parts: the first part is the benchmark serviced for screening apps, while the second part contributed to the construction of checklist content.

As for the first part, the analysis illustrated in Appendix 1 shows that the majority (67%) of our participants learned more than one hundred effective hours of Chinese, resulting in a decision that apps designed mainly for beginners would not be selected. Next, the participants use Android and IOS almost half to half, therefore the researchers add a basic criterion as the app should fit on Both android and IOS systems, in the first-round selection. Besides, it is ordinarily planned that the volunteers use the app for at least three hours within ten days, but the results show that six of the participants have had the experience of using Chinese vocabulary apps and four of them use for half an hour to two hours per week. Therefore, the experience time is reduced to one hour within five days in the final in-use evaluation plan.

Appropriately 78% of the participants chose complex memory goals, including learning about vocabulary pronunciation, meanings, and corresponding written Chinese recognition. The result causes a relationship to their general positive attitudes toward all the items chosen in the main body. For learning goals, the participants regarded "matching words to their meanings or pronunciation" as the most important objective through Chinese vocabulary apps, then followed by "understanding vocabulary in context", and finally "using words appropriately in fluent conversations".

As for the second part, data from the main body of the questionnaire identified that all the participants tend to agree that what we mentioned in the 49 questions are necessary services in an ideal Chinese vocabulary learning app. There are 32 questions where no participant chose the "disagree" option, which means that all learners agree that these 32 constituents should appear in a Chinese vocabulary learning app. Of the remaining 16 questions, only one participant chose "disagree" in 13 of them and two participants in 3 of them (happened in the categories of Content Quality, Pedagogical Coherence, Customization, and Feedback). In the ordinary plan, the questions with more than half of the participants selected "disagree" would not be included in the checklist. As no question resulted in this way, all 49 questions in the questionnaire should have been added to the corresponding sub-criteria of the checklist. However, it was considered that question No.49, " I can accept that the app might have functions that require payment." was not added, because it did not relate to the study, and it is difficult to control the paid content in the free apps during the actual screening.

Data were processed by SPSS. A reliability analysis for the main body of the questionnaire has been done. The value of the  $\alpha$  reliability coefficient is 0.882, greater than 0.8, which indicates that the reliability quality of the research data is high. For the "alpha coefficient with deleted items", the reliability coefficient will not increase significantly after any item is deleted. Therefore, no item should be deleted from this questionnaire. The average scores represented by 48 questions are calculated, and the scores representing the corresponding questions are included in the checklist calculation system. Two additional comments were left by the participants as well, regarded as the needs of learners and added to the content of the checklist. The scoring method for these two items will be explained in the next step (See 4.2.3). Eventually, 50 items were grouped into 7 criteria and integrated into the checklist to form an evaluating mechanism. The average of each of the 50

items is counted to two decimal places, ranging from the lowest score of 1 point to the highest score of 2 points. The total average of all weight scores is 1.572, which is higher than that represented by "neutral" (1 point). The descriptive tables are shown in Table 4.1 and Figure 4.1.

### Table 4.1

Criteria	Minimum	Maxmum	Average	Standard Deviation
<b>Content Quality</b>	1.22	1.89	1.67	0.26
Pedagogical Coherence	1.11	1.67	1.4	0.22
Feedback and self-correction	1.33	1.89	1.65	0.20
Motivation	1.56	1.56	1.56	0
Usability	1	1.89	1.30	0.51
Customization	1.22	1.89	1.65	0.20
Sharing	1.44	1.56	1.5	0.08



It proves that participants generally have a positive attitude towards questionnaire contents, thinking they all should be taken into account when choosing Chinese vocabulary learning apps.

#### Figure 4.1

#### 4.2.3 Checklist design

As shown in Table 4.2, the seven categories (content quality, pedagogical coherence, feedback and self-correction, motivation, usability, customization, and sharing) proposed by Chen (2016), which were initially considered in the design of the questionnaire, were directly applied as the seven main criteria of the checklist to evaluate Chinese vocabulary learning apps. Each general criteria are further divided into a series of sub-criteria in conjunction with the questionnaire content (1. content quality: Chinese word categories, assisted functions of new words, relevant contents of Chinese words; 2. pedagogical coherence: learning method; 3. feedback and self-correction: records of the learning process, feedback of current vocabulary study; 4. motivation: reward, ranking, games; 5. usability: technical support, privacy protection, data management; 6. customization: languages, interface style data, learning plan design, concentration development; 7, sharing; peer working, communication), which were the summary and classification based on the questions chosen from the questionnaire. Each sub-criterion contains questions that describe the criteria to be applied. For example, for the criterion "Chinese word categories", the questions contained "Does the App contain the required vocabulary for official examination/necessary vocabulary for daily conversation/professional terminology in specific majors/relevant vocabulary of personal interest?", a total of 4 items. In another example, the criterion "Ranking" used the question "Does the App provide peers achievement ranking?" only. The two additional needs filled by the participants were involved in the checklist as well, the researcher summarized their corresponding sub-criteria and assigned them to the appropriate main criteria to become additional bonus items. The checklist contained a total of 50 questions.

For the scoring of each criterion, this study was based on the scoring design of Neumann et al.'s (2019) checklist, in which if an application can satisfy a question, it gets a score of 1 point. If it cannot, it receives 0. However, as explained in 4.2.2, different questions should not be weighted equally. The weighting of each question is determined by its average score calculated from the results of the questionnaire. If an application meets these requirements in the assessment, it will receive a comparable mean score.

This checklist also added the needs that come from participants' suggestions, as written in the blanks in the questionnaire, to the corresponding sub-criteria, each rated at 1 point (with an "Add" before the question). Since the needs that came from one participant can't be determined as the general needs of all Chinese learners, these two items were decided to become extra bonuses. If an app being evaluated met these two criteria, an additional 1 point would be added (in the form of "+1" after the formal scores). The total score for each application is calculated by adding up the scores for each sub-criterion (maximum score = 76.6+2). In the forthcoming app evaluation, after the scores of all apps are calculated, researchers will give priority to comparing the total scores except for the two bonus items, that is, the formal scores before "+" determined by the presence or absence of the elements listed in the 48 questions in the checklist. If the formal scores of each app are different, the ranking of apps will be merely determined by the formal scores. Only when there are apps with exactly the same formal scores, additional points would be considered. At that time, the number after "+" would determine the ranking of apps with the same formal score.

The final draft of the checklist for evaluating Chinese vocabulary learning apps is shown in Table 4.2.

 Table 4.2 Evaluation Criteria of Chinese vocabulary learning Apps for self 

 access extracurricular learning

			Score
Criteria	Sub-criteria	Question	(absent=0;
			present= mean)
		Does the App contain the required vocabulary for official examination (e.g., HSK, BCT)?	0 / 1.89
		Does the App contain necessary vocabulary for daily conversation?	0 / 1.78
	categories	Does the App contain professional terminology in specific majors (e.g., finance, law)?	0 / 1.22
		Does the App contain relevant vocabulary of personal interest (e.g., games, cartoons, movies)?	0 / 1.78
	Assisted functions of new words	Does the App provide Pronunciation audios of new words?	0 / 1.44
		Does the App provide phonetic notation (Pinyin) of new words?	0 / 1.89
Content Quality		Does the App provide translation of meanings (in English / learners' native languages) of new words?	0 / 1.78
		Does the App provide sample sentences (characters with Pinyin) of new words?	0 / 1.89
		Does the App provide sample sentences (pronunciation audios) of new words?	0 / 1.33
		Does the App provide descriptive image of new words?	0 / 1.22
		Does the App provide related video of new words?	0 / 1.33
		Does the App give memorizing tips?	0 / 1.78
		Does the App show synonyms / antonyms of each new word?	0 / 1.78
	Relevant	Does the App show homophonic words of each new word?	0 / 1.67
	contents of Chinese words	Does the App provide relevant phrases to each new word?	0 / 2.00
		Does the App provide grammatical structures related to each new word?	0 / 1.89

1				13	
(	с	o	n	T.	)
٩.	~	~	•••	••	,

Criteria Sub-criteri		Question	Score (absent=0; present= mean)
		Does the App provide pronunciation practice?	0 / 1.56
		Does the App provide Pinyin notation practice?	0 / 1.33
Pedagogical Coherence	Learning	Does the App practice on matching words to definite translations (in English / learners' native languages)?	0 / 1.33
	method	Does the App practice on matching words to descriptive explanations (in English / learners' native languages)?	0 / 1.11
		Does the App provide drills about filling in blanks to complete sentences?	0 / 1.67
	Records of Learning Process	Does the App show fully mastered words (in total & per day)?	0 / 1.56
		Does the App show non-fully mastered words (in total & per day)?	0 / 1.67
		Does the App show the number of words learned?	0 / 1.78
		Does the App show the number of words not yet learned?	0 / 1.33
Feedback and		Does the App provide learners' forgetting curves (Compare with the Ebbinghaus forgetting curve)?	0 / 1.67
self-correction		Does the App allow checking words need reviewing next day?	0 / 1.89
		Does the App analyze learners' vocabulary memorization durability?	0 / 1.67
	Feedback of Current	Does the App record learners' daily engagement?	0 / 1.33
	Vocabulary Study	Does the App analyze learners' average correctness (for meaning)?	0 / 1.67
		Does the App provide learners' pronunciation score?	0 / 1.89
	Reward	Does the App have reward mechanism (e.g., badges, titles)?	0 / 1.56
Motivation	Ranking	Does the App provide peers achievement ranking?	0 / 1.56
	Games	Does the App have word practice games?	0 / 1.56

(cont)	(	cont	t)
--------	---	------	----

Criteria	Sub-criteria	Question	Score (absent=0; present= mean)
	Technical Support	(Add) Does the App code well on different OS and screen size / orientation?	0 / 1.00
Usability		Does the App system work smoothly?	0 / 1.00
	Privacy Protection	Is the App able to protect learners' privacy?	0 / 1.00
	Data management	Is the App able to backup and restore learners' study data?	0 / 1.89
	Languages	Is the App able to set different system languages?	0 / 1.78
	Interface Style Data	Does the App allow setting the general style of its interface (e.g., colors, layout)?	0 / 1.67
		Do the learning content provided by the App can be graded according to learners' Chinese proficiency?	0 / 1.89
	Learning Plan Design	Is the App able to set new words numbers per day?	0 / 1.67
		Is the App able to limit days for reaching learning goal?	0 / 1.22
Customization		Does the App allow setting daily study reminders?	0 / 1.44
		Does the App allow setting learning depths (e.g., pronunciation / usage / literacy)?	0 / 1.78
		Is the App able to mark important words?	0 / 1.67
		Is the App able to mark words which often misunderstood?	0 / 1.78
	Concentration development	(Add) Is the App able to lock learners' phones until finishing temporary setting goals?	0 / 1.00
Charing	Peer Working	Dose the App give chance to voluntary group study?	0 / 1.56
Snaring	Communication	Does the App have worldwide social community?	1 / 1.44

## 4.2.4 App screening

In April 2022, the number of potential apps was determined by searching for the keyword "Chinese vocabulary" and "Chinese words" in Qimai Data (a Chinese mobile app data analytics platform). Search filters were applied to the following categories: Platform (iPhone and iPad), Location (US IOS App Store). A total of 184 apps were found.

The first screening step eliminated apps unrelated to Chinese vocabulary learning. After screening, 64 apps were left for the second step. Since the initial search location of the apps was the IOS App Store in the U.S., however, it could not be confirmed in which region the participants' mobile App Stores were set up. Moreover, according to the questionnaire results, both Android and IOS mobile systems are used. To ensure that the app could be used in both Android and IOS systems and more regions, the researchers searched the IOS and GooglePlay App Stores in China, Russia, UK, UAE, Japan, and Korea using the keyword. Thirty apps were left and downloaded to enter the third-round screening. The third-stage criteria excluded the following apps: (1) did not work properly; (2) whose learning content did not closely match Chinese vocabulary; (3) with little free content to support enough usage; (4) were designed for children.

After the third-round screening, six apps remained (shown in Table 4.3) and were evaluated by the checklist.

Logo	App Name	Developer	
	Mandarin Chinese by Nemo	Nemo Apps LLC	
\1 ↓ 大	てい 女子 た か た か た た		
PORO	PORO – Chinese Vocabulary	На Но	

Table 4.3

(	con	t)
· ·		- /

Logo	App Name	Developer
高难马来 GHNESE 之间 P 22 于老汁护	HSK hero (series) – Learn Chinese	Handtechnics
Tobo: Learn Chinese Vocabulary		Utku Uysal
Learn Mandarin Chinese & Hanzi - Drops		PLANB LABS OU

## 4.2.5 Evaluation of selected apps

The two researchers, native Mandarin-speaking postgraduate students studying in Teaching Chinese as a Second Language, completed the evaluation of half of the number (3) apps separately, and then double-checked each other's evaluation results to find whether there were omissions. A further indication of the components that were most frequently present and absent in the apps could be found by looking at the questions in each criterion. Figure 4.2 illustrated the presence (colors) and absence (white) of each criterion and sub-criteria for each selected Chinese vocabulary learning app. Short titles of apps and simplified forms of each question are shown (Table 4.3 gives full titles with apps presented in the same order, Table 4.2 gives complete forms of each question in the same order).

Figure 42 Apr Source Block Chart		App short title							
	rigure 4.2 App scoring book chain		Nemo	PORO	trainchinese	HSK hero	Tobo	Drops	
			required vocabulary for official examination						
		Chinese word extension	necessary vocabulary for daily conversation						
		Crimese word categories	professional terminology in specific majors						
			relevant vocabulary of personal interest						
			Pronunciation audios						
			phonetic notation (Pinyin)						
			translation of meanings (in English/learners' native languages)						
	contract condition	A soluted forestions of more morely	sample sentences (characters with Pinyin)						
	content quanty	Assisted functions of new words	sample sentences (pronunciation audios)						
			descriptive image						
			related video						
			memorizing tips						
			synonyms/antonyms						
		Palarant contants of Chinese works	homophonic words						
		Perevan contents of chinese words	relevant phrases						
			related grammatical structures						
			pronunciation practice						
	Dadaanaind		Pinyin notation practice						
	Cabarana	Learning method	matching words to definite translations (in English learners' native languages)						
	Concrence		matching words to descriptive explanations (in English learners' native languages)						
			filling in blanks to complete sentences						
			fully mastered words (in total & per day)						
			non-fully-mastered words (in total & per day)						
		Records of Learning Descares	the number of words learned						
Critena		Recently of Learning Process	the number of words not yet learned						
(simplified	Feedback and self-correction		forgetting curves (Compare with the Ebbinghaus forgetting curve)						
torm)			checking words need reviewing next day						
		Feedback of Current Vocabulary Study	vocabulary memorization durability						
			daily engagement						
			average correctness (for meaning)						
			pronunciation score						
		Reward	reward mechanism						
	Motivation	Ranking	peers achievement ranking						
		Games	word practice games						
		Technical Support	(Add) code well on different OS and screen size/orientation						
	Trability	rectinear copport	App system work smoothly						
	Osability	Privacy Protection	protect learners' privacy						
		Data management	backup and restore study data						
		Languages	set different system languages						
		Interface Style	setting the general style of its interface						
			learning content can be graded according to learners' Chinese proficiency						
			set new words numbers per day						
	Outerintee		imit days for reaching learning goal						
	Castomization	Learning Plan Design	setting daily study reminders						
			setting learning depths						
			mark important words						
			mark words often misunderstood						
		Concentration development	(Add) lock learners' phones until finishing temporary setting goals						
	(1)	Peer Working	voluntary group study						
	Sharing	Communication	worldwide social community						
			Letter and the second sec						

It can be seen that apps performed best in usability. Half of the evaluation objects met all the needs of this category, and the remaining half also met 3/4 of the needs. Apps also scored higher in content quality. In addition, all apps to some degree met the requirements of pedagogical coherence and customization to varying degrees. They all performed poorly in sharing, and only one app has one of its functions. Apps differed greatly in feedback and self-correction, with some having almost no relevant content and functions, while others provided many services.

Table 4.4 shows the seven main criteria and overall scores for each of the Chinese vocabulary learning applications. As can be seen, the major ratings varied from 15.56 to 35.47, indicating that the apps were of varying quality. The app "trainchinese" had the greatest score (35.47+1), therefore being chosen as the target app in the in-use evaluation. However, it was still much lower than the maximum potential score of 76.05+2. It reveals that among the apps that entered the checklist evaluation in our previous rounds of screening, the best one could meet the most

	F t	IOIAI	20.34+1	35.47+1	17.67+1	15.56+1	32.34+1	27.24+1
		Sharing	0	0	0	0	1.44	0
ular learning		Customization	4.89	7.03	3.34	3.67	5.23	3.45
extracurric		Usability	3.89+1	3.89+1	2+1	2+1	3.89+1	2.89+1
· self-access (	Criteria	Motivation	0	0	0	1.56	1.56	3.12
ning Apps for		Feedback and self-correction	3.11	3.11	0	0	7.67	1.78
ocabulary Lean		Pedagogical Coherence	1.56	2.44	2.44	1.33	2.44	4.11
chinese Vo		Content Quality	6.89	19	9.89	7	10.11	11.89
ry Scores of C		neveroper	Nemo Apps LLC	trainchinese B.V.	Ha Ho	Handtechnics	Utku Uysal	PLANB LABS OU
Table 4.4 Catego	A North Andrewski (1997)	App Mane	Mandarin Chinese by Nemo	trainchinese	PORO – Chinese Vocabulary	HSK hero (series) – Learn Chinese	Tobo: Learn Chinese Vocabulary	Learn Mandarin Chinese & Hanzi - Drops

requirements of learners, while there are still many deficiencies that existed.

Content quality and usability were the categories in which the applications generally performed well. In the categories of feedback and self-correction, motivation, and sharing, some applications do poorly (0 points). Even trainchinese, who got the highest score, did nothing in the columns of "motivation" and "sharing". The rating scale shows that the Chinese vocabulary learning apps currently available on the market for both IOS and Android systems have poor performance, asymmetric function development, and pay more attention to learning content and training.

#### 4.3 In-use evaluation

Three Chinese learners were invited to join the in-use evaluation on a voluntary basis, to provide data for answering the third research question. Inuse evaluation is a two-step process containing an experience period and semistructured interviews. Firstly, three Chinese learners (two beginners and one advanced) were asked to experience the app which gained the highest score on the checklist for five days, with a minimum total duration of one hour. The participants were asked to fill in a question outline (see Appendix 2) during their experience, recording their feelings about each feature. This request aimed to design an interview outline based on the answers to the questions and to help participants recall their experiences during the interview.

### 4.3.1 Semi-structured Interview

The semi-structured interview is using a wide range of research questions to guide the interview. Interview guidelines or interview forms are usually designed before the beginning of the interview as the framework, but their words and question order are not too limited. The main content must be consistent with the research questions, and the type of questions or discussion methods are carried out in a more flexible way. Therefore, it can provide a more realistic appearance of the respondents' cognitive feelings (Lin et al., 2005).

In this study, interviews were conducted with each of the three experiencers

via ZOOM at the convenience of the participants. The interviews with the two beginner Chinese learners were conducted in English, with the advanced learner was conducted in Chinese. All interviews were recorded and fully transcribed via ZOOM with the consent of the participants, and then locked in OneDrive with a password. The interviews ranged from 15 to 30 minutes, with an average time of 20 minutes. The full draft of the interview questions is attached in Appendix 2.

#### 4.3.2 Finding from interview data

The experience records and interviews of the three participants were regarded as three mini case studies under the theme of "how satisfied are you feeling with this app", then interpreted thematically. Keywords were extracted from the interview records to explain the participants' experience attitudes and judgment of the app. Their interview keywords would be linked to make the final statement about the app. Table 4.5 indicates the four interviewees' basic information.

Code	Chinese learning hours	Mobile system	any experience with the same type of apps	Total time using trainchinese
А	Less than 100 hours	IOS	No	148 mins
В	Less than 100 hours	Android	Yes	180 mins
С	More than 100 hours	IOS	Yes	2 hours per week, since 2015

**Table 4.5 Interviewee information** 

#### 4.3.2.1 Case A

Participant A is a first-time learner of Chinese who used trainchinese for 148 minutes in total. He used the app on the IOS system, and "There are no small technical issues, even not shooting up.". He feels partially satisfied with this app, mainly learning content, training mode, and operation.

Participant A said what impressed him the most was that trainchinese provided learners with useful learning content with interactive features.

trainchinese used different colours to differentiate between Chinese tones and helped him remember this difficult language knowledge, and the word interface provides tone representation, audio to practice pronunciation and animation for character-writing. The app also recorded his learning progress and reminded him to keep reviewing previously learned words every three days to five days, which is attractive because he wants to see further feedback. Participant A also felt the "Discussion" function was enjoyable, allowing users to share their real-life experiences using words or sentences and letting the other users comment, which provided him with a comfortable learning atmosphere.

However, participant A said trainchinese did not provide a learning path for the learner. Once started, the users cannot change their learning level, and the app will not give any suggestions according to the learners' progress. Participant A suggested that official learning paths could be provided, as "It was more like an independent training" when the learner was left without guidance. For the practice mode, participant A hopes the app can allow learners to input their pronunciation and give feedback.

Participant A said that he would like to use trainchinese for a long time and recommend it to his friends who are learning Chinese because the function of coloring Chinese tones is really useful. In his description, trainchinese is a very suitable app for beginners, especially because of some special functions to help beginners overcome difficulties.

#### 4.3.2.2 Case B

Participant B is also a first-time Chinese learner using trainchinese, and his experience lasted about 180 minutes. He used this app on the Android system and suffered from several technical issues. Generally, he was not very satisfied with trainchinese ("give it like five out of 10"), mainly because of the feedback system and "old-fashioned" settings.

Participant B was also impressed by the function of colouring the Chinese tones, and he said it would motivate his Chinese study. Like other participants,

participant B commented that this app has a rich vocabulary, covering almost basic everyday language. The app provided different vocabulary topics for learners; participant B said he could choose one specific topic and learn how to use the vocabulary, sentences and grammatical structures in the real context. Moreover, the complete training mode gives participant B a consistent learning experience.

Participant B said the disadvantages of trainchinese were also evident. Firstly, some example sentences confused him, and this app did not provide him with materials to train his reading and listening skills. Secondly, trainchinese lacks many visual elements, which makes him feel bored. For example, this app does not use attractive images when explaining words or sentences, and learners can share pictures, radio or other media rather than just type on the discussion forum. Thirdly, some functions of trainchinese require an internet connection for implementation, which is not very convenient for participant B to achieve offline learning. The most uncomfortable thing is the outdated page design, in which complex icons and subfolders will waste learning time. Participant B also compared trainchinese with other relevant apps (such as Pleco) and said trainchinese provides little feedback and lacks flexibility. He suggested that trainchinese add quizzes to give learners feedback rather than just showing them how well they have mastered the words.

Participant B indicated that he would not be willing to continue using trainchinese in the future as he felt that there were better vocabulary-learning apps available. Overall, this interview suggested that Chinese vocabulary learning apps should include more skill-building content, multiple ways of explaining words (e.g., example sentences, images), and an attractive and consistent interface to meet different learning styles.

#### 4.3.2.3 Case C

Participant C is a high-level Chinese learner who has been using trainchinese since 2015. She downloaded it on IOS system and generally used it for 2 hours per week. She felt very satisfied with it when first started to learn Chinese because it was convenient to look up words and can be used offline, so it

could meet her needs outside the classroom. However, as her Chinese level grows up, she is now not satisfied with trainchinese because it can no longer provide the vocabulary she needs at her current level. Participant C made a similar discussion with participant A: this app is very suitable for Chinese beginners to look up new words, because it classifies life topics (e.g. airport, hospital, restaurant, etc.) and involves various daily words, also the electronic dictionary can be downloaded in advance to search words at any time. However, when learners reach the level that can handle basic daily communication in Chinese, it is difficult to find advanced or even intermediate level words in trainchinese.

Besides, the apps lacks example sentences, synonyms, antonyms, and other learning content, which are very important materials to her. Some Chinese words have different meanings in different contexts, while trainchinese often provides very simple examples or even does not provide examples, which cannot make learners understand when to use the new word.

Participant C compared trainchinese and her favorite vocabulary learning website " $\pm$  5KPC", which is a Chinese learning website in Russia, and later made some suggestions to this app. That website has a very rich vocabulary storage and can explain the meaning of Chinese words in Chinese (for C, a high-level Chinese learner, it is very helpful to explain the target language in the target language).

As for practice mode, participant C's favorite function in trainchinese is flashcards. This function supports making the existing words in the app into word cards directly, which is convenient for subsequent review and personalized training mode. She feels a little regretful that this app does not provide word writing exercises. She hopes to have a Chinese vocabulary learning app that supports the function of Chinese character writing exercises. It is also better to have a scoring and evaluation system for it.

To summarize, the interviews with the three Chinese learners show that trainchinese only partially satisfies their learning needs. Even though trainchinese received the highest score in the checklist evaluation, limitations still need to be improved, mainly in the lack of vocabulary content, language skill training, feedback, and the unattractive interface. The above interview contents were summarized into keywords and sentences, then arranged into a table (See table 4.6).

Satisfaction			
level	Α	В	С
What needs trainchinese has met	<ol> <li>useful learning content with interactive features</li> <li>various training mode</li> <li>sharing forum</li> <li>smooth operation</li> </ol>	<ol> <li>rich vocabulary material with attractive features</li> <li>vocabulary topics</li> </ol>	<ol> <li>basic words storage</li> <li>some offline functions</li> </ol>
What TrainChinese failed to meet	<ol> <li>learning paths</li> <li>pronunciation feedback</li> </ol>	<ol> <li>skill-building content</li> <li>good page design</li> <li>multiple ways of words explanation</li> <li>complete offline learning</li> <li>feedback system</li> </ol>	<ol> <li>relevant</li> <li>learning material</li> <li>character</li> <li>writing practice</li> </ol>

Table 4.6 Interview content

#### **5. DISCUSSIONS**

The usage of mobile applications for self-access extracurricular language learning has rapidly increased in recent years. It also emphasizes the necessity of giving second language learners specific standards when choosing high-quality applications to enhance vocabulary development. The current study created an app criterion checklist to assess Chinese vocabulary learning apps made for selfaccess learners, drawing on well-established needs analysis theories and evaluation checklists. The results of the questionnaire showed that learners prefer the more app learning functions and content, the better. All the participants agreed that the 18 categories involving content quality, pedagogical coherence, feedback and selfcorrection, motivation, usability, customization, and sharing, are necessary to their self-access Chinese words study. Under this context, the use of the checklist indicated significant variation in the caliber of the Chinese vocabulary learning apps that are now available on the market as well as a discrepancy between what they have offered and what is required to satisfy the needs of Chinese learners. Usability, the majority of content quality factors, some degree of pedagogical consistency, and customization were all considered strengths in general. Limitations were discovered for sharing, motivation, and certain aspects of feedback and self-correction. The follow-up experiencing results showed even apps that get relatively high scores on the checklist still can't meet most of the needs of Chinese learners, and there are many disappointments even in the high-scoring sectors (such as content quality). This may give some enlightenment to application developers: even successful Chinese vocabulary learning apps still have many things to improve according to learners' needs, and they need to constantly update the system and innovative content to keep themselves competitive. In addition, due to different learners' Chinese levels and learning styles, their expectations for apps in actual experience are also different. The interviews shows that each of the criteria in the checklist is a must when evaluating Chinese vocabulary learning apps, and the learners were more likely to express opinions related to content quality and customization than other criteria. These requirements pose more challenges to designers who are committed to making popular Chinese vocabulary learning apps.

It should be acknowledged that there are some limitations to the current study, and some suggestions can be made for future research. Firstly, the sample for the questionnaire was inadequate. As only nine participants completed the needs analysis, some necessary needs would have been missed in the design of the checklist, while future research could collect as many questionnaires as possible. Secondly, many factors have not been taken into account in the design of the checklist. The selection of evaluation criteria can be based on various learning theories, constrained by time and subject matter, whereas this paper only chose learner needs as the primary consideration. Moreover, this study does not take into account that the variability of teaching materials may lead to different learning needs. In the review of 48 evaluation checklists, Makundan and Ahour (2010) criticized most checklists as too contextually relevant to be generalizable, suggesting future checklists should consider more factors to achieve universal applicability. Finally, the period for participants to experience an app was too short. Due to time constraints, learners might not be allowed to experience some of the features in-depth, and therefore the results of the in-use evaluation may not be accurate. Perhaps the data would have been more convincing if the learners had been given a few more days.

## **6. CONCLUSION**

Many articles evaluate language learning materials from material evaluation, but very few have studied Chinese learning apps, mainly based on data generated from needs analysis and checklists. This study fills this part of the gap and makes a solid contribution to materials evaluation in the application of mobile Chinese language learning by evaluating Chinese vocabulary apps through preuse evaluation and in-use evaluation. The recent research has dual relevance in this regard. First, it creates a thorough checklist for evaluating Chinese vocabulary learning applications, which is lacking in previous research. When creating vocabulary-learning applications for self-access Chinese learners, app developers may use these criteria as a clear guide to ensure they keep these aspects in mind, particularly from the viewpoints of the learners themselves.

Secondly, the current study adds to our knowledge of the existing pedagogical quality and content of Chinese vocabulary learning apps by providing greater evidence of the learning outcomes required in each app reviewed. The real-world app usage by students reveals that there is a discrepancy between the demands of students as perceived by researchers and teachers and their actual needs. Most importantly, the present study's app rating standards have the potential to enable students to choose high-quality applications to assist their independent Mandarin study outside class.

#### REFERENCES

- Brown, J. D. (2009). 16 Foreign and second language needs analysis. *The handbook of language teaching*, 269.
- Chen, X. (2016). Evaluating Language-learning Mobile Apps for Second-language Learners. *Journal* of Educational Technology Development and Exchange, 9(2), 39-51.
- Chuang, H. Y. (2016). Mobile-assisted language learning APPs for the Chinese classroom. Journal of Technology and Chinese Language Teaching, 7(2), 113-119.
- Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT journal*, 51(1), 36-42.
- Gangaiamaran, R., & Pasupathi, M. (2017). Review on use of mobile apps for language learning. *International Journal of Applied Engineering Research*, 12(21), 11242-11251.
- Gardner, D., Miller, L., 1999. *Establishing Self-Access: From Theory to Practice*. Cambridge University Press, Cambridge.
- Gardner, D., & Miller, L. (2011). Managing self-access language learning: Principles and practice. *System*, 39(1), 78-89.

Grabe, W. (2004). 3. Research on teaching reading. Annual Review of Applied Linguistics, 24, 44-69.

- Graves, K., & Xu, S. (2000). Designing language courses: A guide for teachers (No. 428 G7.). Boston, MA: Heinle & Heinle.
- Godwin-Jones, R. (2017b). Smartphones and language learning. *Language Learning & Technology* 21 (2): 3–17.
- 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), 49-54.
- He, S(何姗) & Zhu, RP(朱瑞平). (2018). A qualitative study on Chinese learners' autonomous learning using dictionary app Taking Pleco as an example (汉语学习者使用词典 APP 进行自主学习的质性研究 —— 以 Pleco 为例). Ludong University Journal (Philosophy and Social Sciences Edition) [鲁东大学学报(哲学社会科学版)] (01),59-64. doi:CNKI:SUN: YSZS.0.2018-01-011.
- Hutchinson, T., & Waters, A. (1987). English for specific purposes. Cambridge university press.
- Jordan, R. R. (1997). English for academic purposes: A guide and resource book for teachers. Cambridge University Press.
- Lai, C. (2017). Autonomous language learning with technology: Beyond the classroom. London, UK: Bloomsbury.
- Lin, C. H., Zhou, K., Yang, S., & Sun, Z. (2022). Developing an evaluation framework for vocabularylearning apps. *Interactive Learning Environments*, 1-15.
- Lin, J. D., Yen, C. F., & Chen, M. H. (2005). Qualitative research method: Models and steps of interviewing. *Journal of Disability Research*, 3(2), 122-136.
- Little, D. (1991). Learner autonomy 1: Definitions, issues and problems. Dublin: Authentik.
- Maslow, A. H. (1954). Motivation and personality. New York: Harper & Row.
- Mishan, F. (2015). Materials development for TESOL. Edinburgh University Press.
- More, C. M., & Travers, J. C. (2013). What's app with that? Selecting education-al apps for young children with disabilities. *Young Exceptional Children*, 16, 15-32. doi:10.1177 / 1096250612464763
- Mukundan, J., Hajimohammadi, R., & Nimehchisalem, V. (2011). Developing an English language textbook evaluation checklist. *Contemporary Issues in Education Research (CIER), 4*(6), 21-28.
- Neumann, M., Wang, Y., Qi, G. Y., & Neumann, D. (2019). An evaluation of mandarin learning apps designed for English speaking pre-schoolers. *Journal of Interactive Learning Research*, 30(2), 167-193.

- Polat, B. and Kim, Y. (2014). Dynamics of complexity and accuracy: a longitudinal case study of advanced untutored development. *Applied Linguistics* 35 (2): 184-207.
- Redd, J. (2011). "Supporting vocabulary growth of high school students: An analysis of the potential of a mobile learning device and gaming app," Ph.D. thesis, Iowa State University, IOWA.
- Rosell-Aguilar, F. (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO journal*, *34*(2), 243-258.
- Schmitt, N. (2000). Vocabulary in language teaching. Cambridge, UK: Cambridge University Press.
- Schunk, D. H., & Zimmerman, B. J. (1994). Self-regulation of learning and performance: Issues and educational applications. Lawrence Erlbaum Associates, Inc.
- Schutz, N. W., & Derwing, B. L. (1981). The problem of needs assessment in English for specific purposes: Some theoretical and practical considerations. *Language for specific purposes* program design and evaluation, 29-45.
- Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. ELT journal, 42(4), 237-246.
- Shi, YX(师亚欣). (2021). Research on satisfaction of Chinese spoken app based on user experience (基于用户体验的汉语口语类 APP 满意度调查研究). Popular Standardization[大众标准 化](19),148-150+153.doi:CNKI:SUN:DZBH.0.2021-19-054.
- Suwantarathip, O., & Orawiwatnakul, W. (2015). Using Mobile-Assisted Exercises to Support Students' Vocabulary Skill Development. *Turkish Online Journal of Educational Technology*-*TOJET*, 14(1), 163-171.
- Sweeney, P., & Moore, C. (2012). Mobile apps for learning vocabulary: Categories, evaluation and design criteria for teachers and developers. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 2(4), 1-16.
- Trinder, R. (2016). Blending technology and face-to-face: Advanced students' choices. *ReCALL*, 28, 83-102.
- Tomlinson, B. (Ed.). (2003). Developing materials for language teaching. A&C Black.
- Tomlinson, B. (Ed.). (2013). Applied linguistics and materials development. A&C Black.
- Yamaguchi, A., Okamoto, E., Curry, N., & Konno, K. (2019). Adapting a Checklist of Materials Evaluation for a Self-Access Learning Center in Japan.
- 陈贤纯 (1999): 对外汉语中级阶段教学改革构想 —— 词语的集中强化教学.《世界汉语教学》 (04),3-11. doi:CNKI:SUN:SJHY.0.1999-04-000.
- 李如龙,杨吉春(2004):对外汉语教学应以词汇教学为中心.《暨南大学华文学院学报》(04),21-29+49. doi:10.16131/j.cnki.cn44-1669/g4.2004.04.006.
- 陆俭明 (1998): 对外汉语教学中经常要思考的问题.《语言文字应用》(04). doi:10.16499/ j.cnki.1003-5397.1998.04.001.
- 徐樱子(2018): APP 在汉语词汇自主学习中的应用(硕士学位论文,上海外国语大学). https:// kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201802&filename=1018151014.nh
- 中华人民共和国教育部(2021):《2020年中国语言文字事业和语言生活状况》. 检自 http:// www.moe.gov.cn/fbh/live/2021/53486/sfcl/202106/t20210602\_534898.html(检索日期) 2022.5.1

# Appendix 1: Original questionnaire and result analysis report

Link: https://www.wjx.cn/vj/tv7y5fX.aspx

## AN EVALUATION OF CHINESE VOCABULARY LEARNING APPS FOR SELF-ACCESS EXTRACURRICULAR LEARNING

## FIRST PART

1 I confirm that I have read and understand the information about the research project and I have had the opportunity to ask questions about the project.

Options	Sample size	Proportion
Yes	9	100%

2 I understand that my participation is voluntary and anonymous.

Options	Sample size	Proportion
Yes	9	100%

3 I understand that after submitting my responses, I will not be able to withdraw them.

Options	Sample size	Proportion
Yes	9	100%

4 I give permission for the research team to use my anonymous responses in possible publications.

Options	Sample size	Proportion
Yes	9	100%

## **SECOND PART**

5 What is / are your native / first / most expert language(s)?

6 How much effective time have you been learning Chinese (including class

time and extracurricular self-study time)?

Options	Numbers	Proportions
Less than 100 hours of Chinese language study	3	33.33%
More than 100 hours of Chinese language study	6	66.67%

7 What system is installed on your most commonly used mobile device?

Options	Numbers	Proportions
Android	5	55.56%
IOS	4	44.44%
Other system	0	0%
I don't know	0	0%

8 Have you ever used Chinese vocabulary learning apps?

Options	Numbers	Proportions
Yes	6	66.67%
No	3	33.33%

9 How much time do you spend each week using Chinese vocabulary learning Apps?

Options	Numbers	Proportions
Less than half an hour	1	16.67%
half an hour to 2 hours	4	66.67%
2 hours to 4 hours	0	0%
more than 4 hours	1	16.67%

## **THIRD PART**

Now we would like to know something about your possible needs and wants for an ideal Chinese vocabulary learning App.

10 What memory goal do you want to achieve through Chinese vocabulary learning Apps?

Options	Numbers	Proportions
Vocabulary pronunciation	1	11.11%

Options	Numbers	Proportions
Vocabulary pronunciation and meanings	1	11.11%
Vocabulary pronunciation, meanings, and corresponding written Chinese recognition	7	77.78%

11 Please prioritize your learning objectives.

Options	Average scores
Use words appropriately in fluent conversations	2
understand vocabulary in contexts	1.56
match words to their meanings and / or pronunciation	1.11

This page is to investigate your needs for learning contents of an ideal Chinese vocabulary learning App.

12 I would like to learn the following word categories from Chinese vocabulary Apps:

	Questions	Disagree (0)	Neutral (1)	Agree (2)	Means
Item1	Required vocabulary of official examination (e.g. HSK, BCT)	0(0%)	1(11.11%)	8(88.89%)	1.89
Item2	Necessary vocabulary for daily conversation	0(0%)	2(22.22%)	7(77.78%)	1.78
Item3	Professional terminology in specific majors (e.g. finance, law)	1(11.11%)	5(55.56%)	3(33.33%)	1.22
Item4	Relevant vocabulary of personal interest (e.g. games, cartoons, movies)	0(0%)	2(22.22%)	7(77.78%)	1.78

13 I would like to learn these relevant contents when memorizing a Chinese

word:

	Options	disagree	neutral	agree	Means
Item6	synonyms / antonyms	0(0%)	2(22.22%)	7(77.78%)	1.78
Item7	homophonic words	0(0%)	3(33.33%)	6(66.67%)	1.67
Item8	phrases	0(0%)	0(0%)	9(100%)	2
Item9	grammatical structures containing this word	0(0%)	1(11.11%)	8(88.89%)	1.89

14 I would like the learning content provided by the App can be graded according to my Chinese proficiency.

Item10	Options	Sample size	Proportion	
	disagree	0		0%
	neutral	1		11.11%
	agree	8		88.89%

15 If your have other needs for vocabulary learning content, please fill in the blank.

This page is to investigate your needs for functions of an ideal Chinese vocabulary learning App.

16 I would like the following to appear to assist my Chinese vocabulary learning:

	Options	disagree	neutral	agree	Means
	Pronunciation audios	2(22.22%)	1(11.11%)	6(66.67%)	1.44
Item11	phonetic notation (Pinyin)	0(0%)	1(11.11%)	8(88.89%)	1.89
Item12	Translation of meanings (in English / my native languages)	0(0%) 2(22.22		7(77.78%)	1.78
Item13	Sample sentences (characters with Pinyin)	0(0%)	1(11.11%)	8(88.89%)	1.89
Item14	Sample sentences (pronunciation audios)	1(11.11%)	4(44.44%)	4(44.44%)	1.33
Item15	descriptive image	1(11.11%)	5(55.56%)	3(33.33%)	1.22
Item16	related video	2(22.22%)	2(22.22%)	5(55.56%)	1.33
Item17	Memorizing tips	0(0%)	2(22.22%)	7(77.78%)	1.78

17 I would like to use the following functions to set my own learning plan:

	Options	disagree	netural	agree	Means
Item18	Set new words numbers per day	0(0%)	3(33.33%)	6(66.67%)	1.67
Item19	Limit days for reaching learning goal	1(11.11%)	5(55.56%)	3(33.33%)	1.22
Item20	Set daily study reminders	1(11.11%)	3(33.33%)	5(55.56%)	1.44
Item21	Set learning depths (pronunciation / usage / literacy)	0(0%)	2(22.22%)	7(77.78%)	1.78
Item22	Mark important words	0(0%)	3(33.33%)	6(66.67%)	1.67
Item23	Mark words which often misunderstood	0(0%)	2(22.22%)	7(77.78%)	1.78

	Options	disagree	neutral	agree	Means
Item24	fully mastered words (in total & per day)	1(11.11%)	2(22.22%)	6(66.67%)	1.56
Item25	non-fully-mastered words (in total & per day)	0(0%)	3(33.33%)	6(66.67%)	1.67
Item26	The number of words have learned	0(0%)	2(22.22%)	7(77.78%)	1.78
Item27	The number of words not yet learned	1(11.11%)	4(44.44%)	4(44.44%)	1.33
Item28	My forgetting curve (Compare with the Ebbinghaus forgetting curve)	0(0%)	3(33.33%)	6(66.67%)	1.67
Item29	Words need reviewing next day	0(0%)	1(11.11%)	8(88.89%)	1.89

18 I hope the App can provide these records related to my learning process:

19 I would like these methods to appear to review Chinese vocabulary I have learned:

	Options	disagree	neutral	agree	Means
Item30	Pronunciation practice	1(11.11%)	2(22.22%)	6(66.67%)	1.56
Item31	Pinyin notation practice	2(22.22%)	2(22.22%)	5(55.56%)	1.33
Item32	match words to definite translations (in English / my native languages)	1(11.11%)	4(44.44%)	4(44.44%)	1.33
Item33	match words to descriptive explanations (in English / my native languages)	1(11.11%)	6(66.67%)	2(22.22%)	1.11
Item34	Fill in blanks to complete sentences	1(11.11%)	1(11.11%)	7(77.78%)	1.67

20 I would like to know feedback of my current vocabulary study from these perspectives:

	Options	disagree	neutral	agree	Means
Item35	Vocabulary memorization durability	1(11.11%)	1(11.11%)	7(77.78%)	1.67
Item36	Learning engagement (Record of days studying)	0(0%)	6(66.67%)	3(33.33%)	1.33
Item37	Average correctness (for meaning)	1(11.11%)	1(11.11%)	7(77.78%)	1.67
Item38	Pronunciation score	0(0%)	1(11.11%)	8(88.89%)	1.89

	Options	disagree	neutral	agree	Means
Item39	Voluntary group study	0(0%)	4(44.44%)	5(55.56%)	1.56
Item40	Worldwide social community	0(0%)	5(55.56%)	4(44.44%)	1.44
Item41	Reward mechanism (e.g. badges, titles)	0(0%)	4(44.44%)	5(55.56%)	1.56
Item42	Learning achievement ranking	0(0%)	4(44.44%)	5(55.56%)	1.56
Item43	word practice games	0(0%)	4(44.44%)	5(55.56%)	1.56

21 I would like the following features to motivate my vocabulary study:

22 If you have any other function needs, please fill in the blank.

This page is to investigate your needs for the technical aspects of an ideal Chinese vocabulary learning App.

23 I hope the App is able to backup and restore my study data.

Item44	Options	Sample size	Proportion
	disagree	0	0%
	neutral	1	11.11%
	agree	8	88.89%

24 I hope the App allow setting the general style of its interface. (e.g. colors, layout)

Item45	Options	Sample size	Proportion
	disagree	0	0%
	neutral	3	33.33%
	agree	6	66.67%

25 I hope the App is able to set different system languages.

Item46	Options	Sample size	Proportion
	disagree	0	0%
	neutral	2	22.22%
	agree	7	77.78%

26 I hope the App system can work smoothly.

Item47	Options	Sample size	Proportion
	No	0	0%
	Yes	9	100%

27 I hope the App is able to protect my privacy.

Item48	Options	Sample size	Proportion
	No	0	0%
	Yes	9	100%

28 I can accept that the App might have functions that require payment.

Item49	Options	Sample size	Proportion
	No	1	11.11%
	Yes	8	88.89%

29 If you have any other technical needs, please fill in the blank.

## FINAL PART

30 We are now recruiting some participants to take part in the next stage of our research. A small group of volunteers will be asked to experience one App, to be selected based on the responses to this questionnaire, for ten days and then be invited to individual interviews.

Participants who complete the next stage and take part in the interview will be given a 100HKD thank-you payment (through FPS). If you are willing to join the App using stage, Please leave your email in the blank.

(Details of the following process and a consent form will be sent to you if provided)

Thank you very much for taking our surveys! Your responses will help us understand Chinese learners' wants and needs for using Chinese vocabulary learning Apps.

# **Appendix 2: Question outline and interview questions**

(Question outline)

Please use 1-3 keywords/phrases to describe each following perspective of

the App.

1	Learning content (Words categories, learnable vocabulary number, etc.)	
2	Training mode (in new word learning & review)	
3	Feedback (of learning progress & vocabulary study)	
4	Operation of this App	
5	Settings (Languages, color, font size, etc.)	
6	Information sharing (Discussion, community, etc.)	

(Interview questions outline)

1. What is your general impression of this App?

2. Are you satisfied with the learning content provided by this app?

If yes, what functions/features of the learning material meets your requirements for a Chinese vocabulary learning app?

If no, why?

3. Are you satisfied with the training model of this app?

If yes, what functions/features about the practices meet your requirements for a Chinese vocabulary learning app?

If no, why?

4. Are you satisfied with the feedback provided by this app?

If yes, what functions/features of the feedback system meet your requirements for a Chinese vocabulary learning app?

If no, why?

5. Are you satisfied with the setting functions provided by this app?

If yes, what functions/features of settings meet your requirements for a Chinese vocabulary learning app?

If no, why?

6. Are you satisfied with the information-sharing functions of this app?

If yes, what functions/features of the sharing mechanism meet your requirements for a Chinese vocabulary learning app?

If no, why?

7. What is your favourite part of this App? Can you explain it in detail?

8. What is your least favourite part of this App? Can you explain it in detail?

9. Are there anything you really want for a Chinese vocabulary App, but this App doesn't provide?

10. Are you willing to use this App to learn Chinese vocabulary for a long time? Why?

## Question outline results from participants

#### 1.Case A

Test Period: May 1, 2022, to May 5, 2022

May 1, 2022, Using time: 25 minutes

May 2, 2022, Using time: 32 minutes

May 3, 2022, Using time: 28 minutes

May 4, 2022, Using time: 33 minutes

May 5, 2022, Using time: 30 minutes

Total: 148 minutes

Please use 1-3 key words/phrases to describe each following perspective of the App.

1	Learning content (Words categories, learnable vocabulary number, etc.)	Great organization of content starting from basic information like numbers and working through useful sentences, verbs, and grammatical structure. Word interface provides tone representation, audio to practice pronunciation and animation for character-writing.
2	Training mode (in new word learning & review)	Settings allow to focus on translation, writing/reading, audio recognition. Helps get familiar with the words and sentences by showing the characters and romanization. Very individual training, no set or suggested learning path.
3	Feedback (of learning progress & vocabulary study)	Brings up progress of words/phrases practiced and catalogues them into "Not Learned", "Partially Learned", "Fully Learned".
4	Operation of this App	Very intuitive and smooth operation. Pop-up instruction messages.
5	Settings (Languages, color, font size, etc.)	Adequate font size. Highly configurable: language input (English, Pinyin, Zhuyin) and output (Simplified and Traditional Chinese). Good use of color for tone-association.
6	Information sharing (Discussion, community, etc.)	Interesting feature: "Discussions" allows users to share their real-life experiences using the words or sentences, as well as, letting the other users comment.

## 2. Case B

Total usage time: 3 hours

Please use 1-3 key words/phrases to describe each following perspective of

the App.

1	Learning content (Words categories, learnable vocabulary number, etc.)	<ul> <li>-Very comprehensive content, covering topics for many levels.</li> <li>-Covers learning materials for various Chinese proficiency tests (HSK, TOCFL, and YCT)</li> </ul>
2	Training mode (in new word learning & review)	-Excellent vocabulary training experience (many example sentences, stroke order, and audio pronunciationincludes audio and video explanations -The word lists are hard to navigate, with many unnecessary sub-folders.
3	Feedback (of learning progress & vocabulary study)	<ul><li>Difficult to keep track of learned words on the app.</li><li>Little feedback on progress.</li><li>No adaptable learning suggestions, the app is very static.</li></ul>

```
(cont)
```

4	Operation of this App	-The OCR feature does not work well on hand-written characters, and barely works on screenshots of typed characters, way worse than other appsThe basic functions of the app (searching, and creating card lists) work well. -Many features don't work in offline mode. -The app is not stable and often crashes.
5	Settings (Languages, color, font size, etc.)	<ul> <li>The app interface is old-fashioned, with poor choice of colors.</li> <li>The interface is very weird, many pop- up alerts, and some settings are misplaced.</li> <li>Lacks visual elements.</li> </ul>
6	Information sharing (Discussion, community, etc.)	-Limited discussion features (only comments and likes/ dislikes on words) -very limited communication with other users.

## 3. Case C

Since 2015, 2hours pre week

1	Learning content (Words categories, learnable vocabulary number, etc.)	Really useful word categories especially for beginners Lack of words (often can't find words in dictionary), not enough examples of usage of words
2	Training mode (in new word learning & review)	Nice effective training
3	Feedback (of learning progress & vocabulary study)	
4	Operation of this App	Nice
5	Settings (Languages, color, font size, etc.)	Content with setting
6	Information sharing (Discussion, community, etc.)	

## 对于面向课外自主学习的汉语词汇学习应用程式的评测

肖静茹\*赵敏名

#### 摘要

移动应用程式正普遍应用于第二语言学习领域。汉语词汇学习 app, 作为在市场中占有很大份额的特殊类型,凸显了词汇学习的重要性与 移动技术在自主习得领域的巨大潜力。本研究基于对中文学习者的需 求分析进行问卷调查,并依据结果设计了评估汉语词汇应用的检查 表,使用该检查表评估了六个用于课外自主学习的 app。结果表明, 半数选定的 app 缺乏共享和动机功能,并且在反馈与纠正机制中表现 不佳。这些应用在内容质量方面表现最好。

一款名为 trainchinese 的中文词汇学习 app 在本研究的检查表中获得最高分,随后由三名志愿者进行体验。结果显示,尽管该应用相对其他同类更符合学习者的心理期望,但在实际体验过程中仍然只能满足学习者的部分需求。此外,学习者在汉语水平和学习风格上的差异也是对研究的干扰因素。

本文的研究结果提供了从学习者需求出发评估中文词汇学习 app 的崭 新角度。然而,有针对性的应用程式评测标准仍需要进一步的探究, 并通过多对象的使用体验加以补充。

关键词:中文作为第二语言的学习 词汇学习 移动应用程 式 评测 自主语言学习

<sup>\*</sup>肖静茹,香港科技大学人文社科学院。(本文通讯作者)

赵敏名,香港科技大学人文社科学院。