Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Special Issue on CALL Vol. 17, No.12. July 2026 Pp. 345-364
DOI: https://dx.doi.org/10.24093/awej/call12.20

TEM-4 Vocabulary Learning Needs of Chinese PGUI Students: A Mixed-Methods Analysis for Mobile App Design

Haiping Duan
Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia
Email: p133196@siswa.ukm.edu.my

Melor Md Yunus
Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia
Corresponding Author: melor@ukm.edu.my

Hanita Hanim Ismail
Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia
Email: hanitaismail@ukm.edu.my

Received: 12/25/2026 Accepted: 04/28/2026 Published:07/15/2026

Abstract
Private general undergraduate institutions (PGUIs) account for 31.5% of undergraduate institutions in China, yet the vocabulary learning needs of their English majors remain underexplored, highlighting a critical research gap. This study adopts Hutchinson and Waters' (1987) needs analysis framework to examine PGUI students' Test for English Majors Band-4 (TEM-4) vocabulary learning needs. Employing a convergent parallel mixed-methods design, this study integrated questionnaire data from 240 students with semi-structured interviews from seven experienced teachers. The findings indicated that 77.9% of students recognized vocabulary as essential for TEM-4 performance, reflected in a high mean score for perceived necessity (M = 3.98, SD = 1.05). However, they reported moderate perceived lacks (M = 3.62) in current learning approaches. Students showed strong acceptance of mobile-assisted learning (M = 3.80) and high expectations for deep learning functions (M = 3.71). Qualitative analysis of 475 codes across 26 sub-themes revealed that teachers identified a critical 3:1 receptive-productive gap, with students recognizing approximately 6,000 words but actively using fewer than 2,000. All seven teachers confirmed that existing mobile applications, despite widespread use, fail to support deep vocabulary learning outcomes. The integration of quantitative and qualitative findings provided clear guidelines for developing TEM-4-specific mobile applications, highlighting systematic testing with immediate feedback, authentic TEM-4 examples, vocabulary discrimination exercises, and multimodal learning resources. The study provides empirical evidence for designing mobile applications that address the multidimensional vocabulary learning needs of PGUI students, a distinct and under-researched learner population.

Keywords: deep vocabulary learning, mixed-methods, mobile application design, mobile-assisted vocabulary learning, needs analysis, private general undergraduate institutions, receptive-productive gap, Test for English Majors Band-4

Cite as: Duan, H., Yunus, M. M., & Ismail, H.H.(2026).TEM-4 Vocabulary Learning Needs of Chinese PGUI Students: A Mixed-Methods Analysis for Mobile App Design. *Arab World English Journal (AWEJ) Special Issue on CALL17.* (12) 345-364. https://dx.doi.org/10.24093/awej/call12.20

Introduction

Over the past decades, English proficiency has grown increasingly important in China, driven largely by national policies that emphasise foreign language education. The Ministry of Education has required universities to effectively improve students' English proficiency and their ability to use English in academic and professional contexts (Cui, 2025). For English majors, Test for English Majors Band-4 (TEM-4) serves as the most authoritative national assessment, functioning both as a benchmark for teaching quality and as a credential that influences postgraduate admission and career prospects (Huang & Wang, 2024).

Central to success in such standardized tests is vocabulary knowledge, which plays a fundamental role in language acquisition across all skill areas. Nation (2006) stated that "98% text coverage…would be needed for most learners to gain adequate comprehension" (p. 73), while McCarthy (1984) argued that meaningful communication is unattainable without sufficient vocabulary regardless of grammatical competence. In the context of TEM-4, vocabulary is assessed both directly through the Language Knowledge section and indirectly through listening, reading, and writing components. This pervasive role of vocabulary underscores its critical importance for test success (Schmitt, 2010; Sun et al., 2023).

Vocabulary acquisition remains a persistent challenge for Chinese university students learning English as a foreign or second language (Cai, 2025). Studies have shown that the large-scale expansion of higher education enrolment since 1999 has intensified disparities in English vocabulary proficiency among students across different universities (Li, 2014). This gap becomes even more pronounced among students at Private General Undergraduate Institutions (PGUIs)—a rapidly expanding sector now comprising 31.5% of China's undergraduate institutions (Ministry of Education, 2024). PGUI students, who typically enter university with lower National College Entrance Examination scores, represent an academically vulnerable population whose English learning needs have been largely overlooked in existing research. Given their weaker foundational English proficiency, these students face substantial barriers in achieving the vocabulary threshold required for academic success and TEM-4 certification (Tao & Modehiran, 2023).

Recent advances in mobile technology have opened new avenues for addressing vocabulary learning challenges. Mobile-assisted Vocabulary Learning (MAVL) has gained widespread adoption due to its accessibility, personalization, and capacity for interactive engagement (Kukulska-Hulme & Shield, 2008). Popular applications such as Baicizhan and Shanbay have become ubiquitous among Chinese English learners. However, critical examination of existing applications reveals a significant limitation: most focus primarily on vocabulary retention and recognition through repetitive memorization, while neglecting deeper dimensions of vocabulary knowledge including semantic relationships, productive use, and contextual application (Li, 2024; Mohammadi et al., 2024). Wang (2019) proposed a comprehensive framework for deep vocabulary learning encompassing six dimensions—vocabulary breadth, vocabulary depth, vocabulary organisation, meta-vocabulary knowledge, receptive-productive skills, and semantic autonomy—yet few existing applications systematically address these dimensions.

Despite MAVL's potential for supporting comprehensive vocabulary development, limited research has investigated the specific needs of PGUI students for mobile-assisted deep vocabulary learning in the Chinese context. Understanding these needs through systematic analysis is essential for designing effective mobile applications that address the unique challenges faced by this underserved student population. Therefore, this study aims to conduct a needs analysis based

on Hutchinson and Waters' (1987) framework to inform the design of a TEM-4 vocabulary learning application. The research addresses the following question:

What are the needs of PGUI students for TEM-4 vocabulary learning through a mobile application in terms of necessities, lacks, wants, and learning needs?

Literature Review

The Role of Vocabulary in Second Language Acquisition

Vocabulary occupies a central position in language learning and serves as the fundamental building block for all forms of linguistic expression. As Wilkins (1972, p. 111) famously stated, "without grammar very little can be conveyed, without vocabulary nothing at all can be conveyed." Recent empirical research continues to validate this assertion. Akşit and Saygı (2024) found that among 536 EFL undergraduate students, vocabulary size was considerably below the levels needed for academic study, confirming that vocabulary remains a fundamental barrier to language competence. Moreover, the extent of one's vocabulary significantly influences learners' ability to communicate effectively in both spoken and written forms. Wang and Zhang (2025) confirmed through longitudinal analysis that both vocabulary breadth and depth make significant contributions to language comprehension among Chinese EFL learners, with vocabulary breadth serving as a particularly strong predictor.

Understanding what it means to know a word is essential for effective vocabulary instruction. Nation (2001) proposed three categories of vocabulary knowledge: word form (spoken form, written form, and word parts), word meaning (form-meaning connections, concepts, and associations), and word use (grammatical functions, collocations, and constraints). Beyond these categories, vocabulary knowledge operates along receptive and productive dimensions—the former involving recognition during listening or reading, the latter requiring retrieval and expression in speaking and writing (Laufer & Paribakht, 1998; Meara, 1990; Webb, 2008). This multidimensional conceptualisation suggests that effective vocabulary acquisition extends far beyond simple memorisation of definitions. Building on this view, Meara (2005) proposed that ideal L2 vocabulary competence comprises vocabulary size, vocabulary organisation, and automaticity, highlighting that development involves not merely accumulating words but developing rich associations and achieving fluent lexical access. This theoretical perspective has important implications for vocabulary instruction and assessment.

Despite its importance, vocabulary learning remains one of the most challenging aspects of second language acquisition (Nation & Meara, 2013). The learning process is often perceived as tedious and time-consuming, frequently leading learners to abandon their efforts (Dörnyei & Csizér, 2002). This challenge is particularly evident among Chinese university students. Hua et al. (2024), in their longitudinal study of English as a Medium of Instruction (EMI) students at a Chinese university, found that while students showed significant gains in vocabulary knowledge over a term, their overall levels remained below what successful academic engagement typically requires. Similarly, Yang and He's (2025) corpus-based analysis revealed that lexical richness—encompassing density, sophistication, and variation—remains a persistent weakness in Chinese students' academic writing. These challenges are likely more pronounced among students at less selective institutions, such as PGUIs, where students typically enter university with weaker English foundations. However, vocabulary research specifically targeting this population remains scarce, highlighting a notable gap in the literature.

Deep Learning of English Vocabulary

The lens of deep versus surface learning offers a useful framework for understanding approaches to vocabulary learning. The distinction was first articulated by Marton and Saljo (1976), who found that surface learning involves recognising and memorising information for immediate tasks, while deep learning requires learners to engage with "the underlying meaning of the text" (p.7). Applied to vocabulary acquisition, surface approaches involve rote memorisation of word-definition pairs, frequently resulting in rapid forgetting and failure to develop productive knowledge. In contrast, deep vocabulary learning involves "actively constructing knowledge through meaningful engagement with words, analysing morphological structures, and exploring semantic relationships" (Wang, 2019, p.56). Research in the Chinese context shows that traditional memorisation remains prevalent, with students relying heavily on dictionary use and repetition while underutilising semantic encoding and word association approaches (Sun et al., 2023).

Recognising the need for a comprehensive framework, Wang (2019) developed a model defining vocabulary deep learning as "the process in which learners invoke higher-order thinking to actively construct vocabulary knowledge and build and optimize the English mental lexicon network" (p. 48). The model identifies six essential dimensions: vocabulary breadth (quantity of words known), vocabulary depth (knowledge of multiple meanings, collocations, morphology, and usage), vocabulary organisation (systematic integration into interconnected networks), meta-vocabulary knowledge (awareness of what constitutes word knowledge and how to learn), receptive-productive skills, and semantic autonomy (understanding L2 words independently of L1 translation). Empirical research supports the relevance of these dimensions to Chinese learners' vocabulary challenges. Wang and Zhang (2024) investigated the contributions of morphological awareness and lexical inferencing ability to vocabulary knowledge among Chinese EFL learners using structural equation modelling and found significant gaps in learners' morphological processing and deep vocabulary knowledge. Students struggled particularly with morpheme recognition, word association, and semantic inferencing—reflecting deficiencies across multiple deep learning dimensions. These findings suggest that addressing the full spectrum of dimensions is necessary for meaningful vocabulary improvement.

Mobile-assisted Vocabulary Learning

Mobile-assisted Language Learning (MALL) has emerged as a promising approach, leveraging the ubiquity of mobile devices to support language acquisition through multimedia presentation, interactive exercises, personalised feedback, and anytime-anywhere accessibility (Kukulska-Hulme et al., 2017; Simonnet et al., 2025). Within this field, MAVL focuses specifically on lexical development. Theoretical support draws from dual coding theory, which posits that learning is enhanced through verbal and visual channels (Clark & Paivio, 1991), and spaced repetition research demonstrating that distributed practice leads to superior long-term retention compared to massed practice (Mahdi, 2018). The accessibility of mobile devices also supports autonomous learning and enables practice during otherwise unutilised time fragments (Barzani & Omar, 2024).

Recent empirical evidence supports MAVL effectiveness. A comprehensive Bayesian meta-analysis by Zhou and Zhou (2025), synthesising 65 quasi-experimental studies published between 2010 and 2024, found a large effect size ($g = 1.28$) for mobile applications in vocabulary learning when treatment duration exceeded 10 weeks. Teymouri's (2024) systematic review of digital flashcard-based MAVL studies confirmed that mobile-assisted approaches promote learner

autonomy, engagement, and long-term retention through spaced repetition and gamification. In the Chinese context, Li (2024) reviewed popular English vocabulary learning applications, identifying strengths in gamification and multimedia presentation.

Despite these encouraging findings, critical examination reveals important limitations in current MAVL applications. Mohammadi et al. (2024) found that while mobile-assisted learning improved both receptive and productive vocabulary knowledge among university students, most applications still emphasised recognition over production. Li (2024) concluded that existing Chinese vocabulary apps lack systematic integration of language learning theories supporting deeper lexical processing, despite incorporating various vocabulary learning strategies. Teymouri (2024) further stressed that although digital flashcards realise the interactive and personalised benefits of MAVL, current applications have not fully capitalised on these affordances for comprehensive vocabulary development. These critiques suggest that current applications largely fail to promote the deep learning dimensions that Wang (2019) identified as essential, even though MAVL demonstrates potential for supporting basic vocabulary knowledge.

Needs Analysis in Language Learning Material Development

The limitations in existing MAVL applications underscore the importance of systematic needs analysis in designing learning tools. Needs analysis is widely recognised as a systematic process of collecting and analysing data to identify and prioritise learners' needs, which serves as a fundamental basis for curriculum and material development (Mansor et al., 2025). In English language teaching, it plays a critical role in ensuring that instructional materials are aligned with the actual needs of target learners, rather than being based on intuitive or unverified assumptions (Mao & Zhou, 2024). Among various needs analysis frameworks, Hutchinson and Waters' (1987) model has been widely recognised as foundational in ESP research. They define target needs as comprising "necessities, lacks and wants" (p. 55), while learning needs refer to how learners can most effectively acquire the required knowledge. This framework continues to be widely applied in contemporary EFL research. For example, Gürler and Konca (2023) employed the same model to examine learners' necessities, lacks, and wants, identifying significant discrepancies between learners' existing proficiency and target needs.

Previous studies have successfully applied needs analysis to develop digital English learning materials. Abd Rahman et al. (2024) identified that educators recognized significant challenges in students' synthesis writing, particularly related to language proficiency and the complexity of the writing process, and emphasized the need for mobile application modules that support instruction through accessible and technology-enhanced learning tools. Most relevant to the present study, Fisal and Said (2023) examined English as a Second Language (ESL) pre-university learners' needs for web-based academic vocabulary resources using Hutchinson and Waters' framework, revealing positive attitudes toward technology-enhanced learning and dissatisfaction with current classroom instruction. However, no study has systematically investigated the needs of PGUI students in China for mobile-assisted vocabulary deep learning in TEM-4 preparation. Given this population's unique characteristics—lower baseline proficiency, limited exposure to effective instruction, and specific test requirements—targeted research is needed. The present study addresses this gap by applying Hutchinson and Waters' (1987) framework to investigate PGUI students' necessities, lacks, wants, and learning needs for TEM-4 vocabulary learning through a mobile application designed to support deep vocabulary learning.

Method

Research Design

This study employed a convergent parallel mixed-methods design to investigate the TEM-4 vocabulary learning needs of English majors at PGUIs in China. Following Hutchinson and Waters' (1987) needs analysis framework, the research examined four dimensions: necessities, lacks, wants, and learning needs. The study collected quantitative data from 240 students through a structured questionnaire and qualitative data from seven teachers through semi-structured interviews. The parallel design enabled independent data collection from two complementary stakeholder groups whose perspectives were subsequently integrated to identify patterns of convergence and complementarity.

Participants

The quantitative strand involved 240 English major students from different PGUIs in China. Demographic characteristics of these participants are presented in Table 1 below.

Table 1. *Demographic profile of the survey participants (N = 240)*

Items	Category	Frequency	Percentage (%)
Gender	Female	191	79.6
	Male	49	20.4
Year of Study	First year	80	33.3
	Second year	117	48.8
	Third year	43	17.9
Years Learning English	6-7 years	46	19.2
	8-9 years	71	29.6
	10-11 years	103	42.9
	12 years or more	20	8.3
Total		240	100

The sample consisted predominantly of female students (79.6%), reflecting the typical gender distribution of English major programmes in China. Second-year students constituted the largest group (48.8%), followed by first-year (33.3%) and third-year students (17.9%). Most participants reported 10–11 years of English learning experience (42.9%).

In the qualitative strand, seven experienced English teachers (T1–T7) from five PGUIs participated in semi-structured interviews. The teachers were purposively selected based on their experience in vocabulary instruction and TEM-4 preparation, with teaching experience ranging from 9 to 14 years across diverse geographical regions (Table 2).

Table 2. *Profile of teacher interviewees (N = 7)*

Participants	Gender	Years of teaching experience	Location in China
T1	Female	12	Southwest
T2	Male	10	Southwest
T3	Female	9	Northeast
T4	Female	13	Central
T5	Male	14	Central
T6	Female	11	Southeast
T7	Male	10	Northwest

Instruments

Two instruments were developed according to Hutchinson and Waters' (1987) framework. The student questionnaire comprised 37 items organised into four parts: Part A

(Necessities, five items) measured perceived importance of TEM-4 vocabulary for various language skills; Part B (Lacks, five items) assessed perceived deficiencies in current vocabulary learning situations; Part C (Wants) was divided into C1 (APP Attitudes, three items) measuring attitudes towards mobile-assisted vocabulary learning, and C2 (Deep Learning Expectations, six items adapted from Wang's 2019 framework); Part D (Learning Needs, 15 items) measured preferences for input content, exercise types, and media types. All items used 5-point Likert scales. The teacher interview protocol consisted of 10 open-ended questions exploring observations of student vocabulary learning behaviours, perceived challenges and deficiencies, attitudes towards mobile learning technologies, and recommendations for effective interventions.

Validity and Reliability

Both instruments underwent rigorous validation. The questionnaire was reviewed by two experts in English language teaching who evaluated item clarity, relevance, and alignment with the theoretical framework. Following expert review, a two-phase pilot test was conducted: the first with 15 students identified items requiring clarification, and the second with 34 students assessed reliability. Cronbach's alpha values in the pilot demonstrated acceptable to excellent internal consistency: Part A (α = .902), Part B (α = .686), Part C1 (α = .933), and Part C2 (α = .959). Table 3 shows the Cronbach's alpha values for the main study, which all exceeded the 0.70 threshold recommended by Taber (2018). Two experts also reviewed the interview protocol, which was then piloted with two teachers to evaluate question effectiveness and duration.

Table 3. *Reliability statistics of the main study questionnaire (N=240)*

Constructs	Items	Cronbach's α
Part A: Necessities	5	.867
Part B: Lacks	5	.882
Part C1: APP Attitudes	3	.791
Part C2: Deep Learning Expectations	6	.867
Part D: Learning Needs	15	N/A Formative

Note. Part D is treated as a formative construct; internal consistency is not applicable (Diamantopoulos & Winklhofer, 2001).

Data Collection and Analysis

Quantitative data were collected over two weeks via Wenjuanxing. Participants were informed of the study's purpose, assured of anonymity, and asked to provide informed consent. Concurrently, semi-structured interviews were conducted online via Tencent Meeting, lasting 30–45 minutes each. Interviews were conducted in Mandarin Chinese, audio-recorded with consent, and subsequently transcribed and translated into English.

Quantitative data were analysed using SPSS Version 26, with descriptive statistics including frequencies, percentages, means, and standard deviations calculated for all items. Mean scores were interpreted based on equal interval classification, where the scale range (1–5) was divided into three levels: low (1.00–2.33), moderate (2.34–3.67), and high (3.68–5.00). Qualitative data were analysed through thematic analysis following Braun and Clarke's (2006) six-phase approach, with codes organised according to the framework dimensions. Member checking with two participating teachers confirmed the accuracy and representativeness of interpretations. Following independent analysis of both strands, quantitative and qualitative findings were integrated through a joint display organised by framework dimensions to identify patterns of convergence and complementarity.

Results and Findings

Quantitative Findings: Student Questionnaire

As shown in Table 4, students reported high levels of perceived necessities (M = 3.98, SD = 1.05), APP attitudes (M = 3.80, SD = 1.08), and deep learning expectations (M = 3.71, SD = 1.09). In contrast, perceived lacks (M = 3.62, SD = 1.19) and overall learning needs (M = 3.65, SD = 1.26) were rated at a moderate level.

Table 4. *Summary of student questionnaire findings (N = 240)*

Construct	Items	M	SD	Level
Part A: Necessities	5	3.98	1.05	High
Part B: Lacks	5	3.62	1.19	Moderate
Part C1: APP Attitudes	3	3.80	1.08	High
Part C2: Deep Learning Expectations	6	3.71	1.09	High
Part D: Learning Needs	15	3.65	1.26	Moderate

Note. Level interpretation: Low (1.00-2.33), Moderate (2.34-3.67), High (3.68-5.00).

Necessities

Students demonstrated high agreement regarding the necessity of TEM-4 vocabulary (M = 3.98, SD = 1.05). According to Table 5, TEM-4 test performance received the highest agreement (M = 4.13), followed by reading comprehension (M = 4.09) and listening comprehension (M = 3.99). Class debates and presentations (M = 3.89) and writing quality (M = 3.77) received somewhat lower but still substantial support. These findings highlight that PGUI English majors clearly recognise the significance of vocabulary across multiple language skills, with particularly strong recognition of its role in examination success and receptive skills development.

Table 5. *Descriptive statistics for part A: necessities (N = 240)*

Item	M	SD
Q4. TEM-4 vocabulary is useful for me to read English texts.	4.09	0.98
Q5. TEM-4 vocabulary is useful for me to improve the quality of my English writing.	3.77	1.09
Q6. TEM-4 vocabulary is useful for me to listen to English audio.	3.99	1.06
Q7. TEM-4 vocabulary is useful for me to participate in formal class debates and presentations.	3.89	1.07
Q8. TEM-4 vocabulary is useful for me to achieve good results in the TEM-4 exam.	4.13	1.01
Overall	3.98	1.05

Lacks

Students reported a moderate level of lacks in their current vocabulary learning situations (M = 3.62, SD = 1.19). Table 6 shows that the most prominent concerns were ineffective learning strategies (M = 3.73) and insufficient classroom vocabulary input (M = 3.71). Not utilising supplementary resources (M = 3.62), difficulty achieving deep learning (M = 3.55), and insufficiency of textbook-only learning (M = 3.49) were also acknowledged. This suggests that students face challenges stemming from both inadequate instructional support and limitations in their self-directed vocabulary learning approaches.

Table 6. *Descriptive statistics for Part B: Lacks (N = 240)*

Item	M	SD
Q9. Classroom teaching does not provide sufficient TEM-4 vocabulary input for me.	3.71	1.17
Q10. Learning TEM-4 vocabulary only from textbooks is not enough for me.	3.49	1.21
Q11. I do not use other supplementary resources to learn TEM-4 vocabulary other than textbooks.	3.62	1.20
Q12. I use other supplementary to learn TEM-4 vocabulary, but I find it difficult to achieve deep	3.55	1.17

vocabulary learning.		
Q13. The current ways in which I learn vocabulary are not effective enough for TEM-4 exam preparation.	3.73	1.17
Overall	3.62	1.19

Wants

Table 7 suggests students held highly positive attitudes toward mobile-assisted vocabulary learning (M = 3.80, SD = 1.08). The majority expressed enjoyment in APP-based vocabulary learning (M = 3.98). A considerable proportion found APP learning motivating (M = 3.81) and enjoyed doing vocabulary exercises on applications (M = 3.62). The results indicate strong receptiveness to mobile-assisted learning among PGUI English majors.

Table 7. *Descriptive statistics for Part C1: APP attitudes (N = 240)*

Item	M	SD
Q14. I like learning TEM-4 vocabulary on apps.	3.98	1.03
Q15. I like doing TEM-4 vocabulary exercises on apps.	3.62	1.10
Q16. I find learning TEM-4 vocabulary on apps very motivating.	3.81	1.07
Overall	3.80	1.08

Table 8 presents the deep learning expectations based on Wang's (2019) framework of vocabulary deep learning dimensions. Students expressed the strongest preferences for receptive-productive skill integration (M = 3.90), indicating a desire to apply learned vocabulary in both comprehension and production tasks. Vocabulary breadth (M = 3.83), vocabulary organisation through network building (M = 3.76), and meta-vocabulary knowledge (M = 3.66) also received high ratings. Expectations for vocabulary depth (M = 3.52) and semantic autonomy (M = 3.59) were relatively lower but still approached the high range. These findings indicate that students want mobile applications that facilitate comprehensive vocabulary development rather than mere word memorisation.

Table 8. *Descriptive statistics for Part C2: deep learning expectations (N = 240)*

Item	M	SD
Q17. I hope TEM-4 vocabulary apps can help me master word pronunciation, spelling, and basic meanings.	3.83	1.02
Q18. I hope TEM-4 vocabulary apps can help me master various aspects of vocabulary knowledge such as collocations, usage, roots/affixes, and derivatives.	3.52	1.07
Q19. I hope TEM-4 vocabulary apps can help me clarify what aspects of knowledge I need to master for each word.	3.66	1.10
Q20. I hope TEM-4 vocabulary apps can help me connect learned vocabulary knowledge to form a systematic vocabulary network.	3.76	1.10
Q21. I hope TEM-4 vocabulary apps can help me use learned words in reading, listening comprehension, as well as speaking and writing.	3.90	1.04
Q22. I hope TEM-4 vocabulary apps can help me understand the differences in meaning and usage between English words and their Chinese equivalents.	3.59	1.15
Overall	3.71	1.09

Learning Needs

Students' learning needs concerning preferred content for the proposed mobile application received moderate-to-high importance ratings (M = 3.65, SD = 1.26) across three sub-dimensions: input content preferences, exercise type preferences, and media type preferences. Table 9 presents the detailed findings for each sub-dimension.

Table 9. *Descriptive statistics for Part D: learning needs (N = 240)*

Sub-dimension	Item	M	SD
Input Content	Q23. The definition of a word	4.04	1.00
	Q24. Pronunciation	3.84	1.14
	Q25. Sample sentences	3.78	1.24
	Q26. Vocabulary organisation	3.82	1.24
	Q27. Vocabulary discrimination	3.87	1.18
	Sub-total	3.87	1.16
Exercise Types	Q28. Matching	2.53	1.06
	Q29. Multiple-choice	3.88	1.21
	Q30. True/false	3.85	1.24
	Q31. Sentence translation	3.73	1.25
	Q32. TEM-4 vocabulary discrimination quizzes	3.68	1.29
	Q33. Vocabulary memorization games	2.61	1.20
	Sub-total	3.38	1.34
Media Types	Q34. Text	3.85	1.22
	Q35. Image	3.81	1.20
	Q36. Audio	3.74	1.20
	Q37. Video	3.74	1.12
	Sub-total	3.79	1.19

For input content, the definition of a word (M = 4.04) and vocabulary discrimination (M = 3.87) received the highest ratings, followed by vocabulary organisation (M = 3.82), sample sentences (M = 3.78), and pronunciation (M = 3.84). For exercise types, students showed a clear preference for structured, examination-oriented formats—multiple-choice (M = 3.88), true/false (M = 3.85), sentence translation (M = 3.73), and TEM-4 vocabulary discrimination quizzes (M = 3.68)—while matching (M = 2.53) and memory games (M = 2.61) received notably lower ratings. This pattern suggests students prioritise assessment-relevant practice over gamified activities. For media types, all four received similarly high ratings, indicating preference for multimodal learning experiences.

Qualitative Findings: Teacher Interviews

Semi-structured interviews with seven experienced English teachers (T1-T7) from PGUIs yielded 475 codes across 26 sub-themes, organised into the four main themes of Hutchinson and Waters' (1987) framework. All seven teachers contributed codes to each main theme, demonstrating robust inter-rater consensus and enhancing the reliability of the findings.

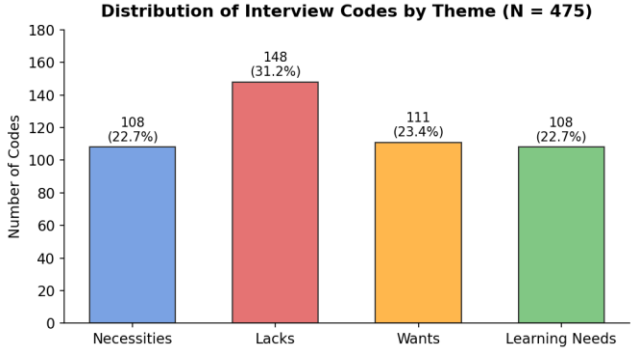


Figure 1. Distribution of interview codes by theme (N = 475)

As shown in Figure 1, Lacks represented the most prominent theme (31.2%, 148 codes), followed by Wants (23.4%, 111 codes), while Necessities and Learning Needs each accounted for 22.7% (108 codes). This distribution pattern suggests that the gap between students' current vocabulary abilities and TEM-4 requirements constitutes the most critical concern among experienced teachers.

Necessities

All seven teachers confirmed that students recognise the fundamental importance of vocabulary learning for TEM-4 success (21 codes, 7/7 teachers). As T4 articulated, "Students generally believe that vocabulary learning occupies a very, very important position in learning English, and it is also an important factor in whether they can pass TEM-4." However, exam pressure emerged as the dominant motivational force (69 codes, 7/7)—the single most frequently coded sub-theme across the entire analysis. Teachers also identified declining motivation over time (5 codes, 4/7), presenting a significant pedagogical challenge. T6 explained: "In the early stages, everyone is particularly enthusiastic... However, as time goes on, because of the long preparation cycle for vocabulary, their interest will decline." This erosion of initial enthusiasm is particularly concerning given that sustained engagement is essential for achieving the vocabulary breadth and depth required for TEM-4 success.

Lacks

The Lacks theme emerged as the most substantial category in the analysis (148 codes, 31.2%), indicating that vocabulary learning deficiencies constitute the most critical challenge in TEM-4 preparation from teachers' perspectives. Teachers identified deficiencies across all dimensions of vocabulary knowledge. Figure 2 illustrates the sub-theme distribution.

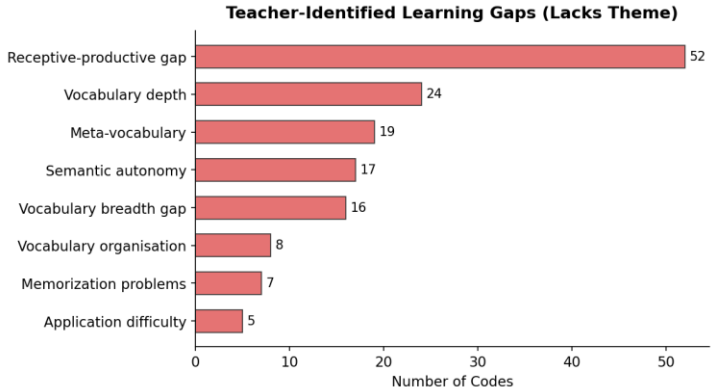


Figure 2. Teacher-identified learning gaps (Lacks Theme)

The most severe deficiency concerns the receptive-productive gap (52 codes, 7/7 teachers). T3 provided specific quantification: "Students can recognise around 6,000 words, but when it comes to productive use, it might be less than 2,000." This 3:1 ratio represents a fundamental obstacle to TEM-4 success, as the examination includes substantial productive components. Teachers also identified vocabulary depth deficiency (24 codes, 7/7), with T2 noting students "only remember one or two common meanings"; meta-vocabulary knowledge gaps (19 codes, 7/7), where students cannot clarify aspects of knowledge they need to master for each word; semantic autonomy weakness (17 codes, 6/7), manifested in over-reliance on translation; and breadth gaps of 25-30% below the 8,000-word requirement (16 codes, 7/7).

Wants

All seven teachers reported that students demonstrate a strong preference for mobile vocabulary learning applications (25 codes) and make extensive use (49 codes) of popular platforms such as Baicizhan, Momo, and Shanbay. Students value convenience and the ability to utilise fragmented time, as T4 stated: "They are very willing to use apps because apps are very convenient, allowing them to learn vocabulary anytime, anywhere." However, all seven teachers identified a critical limitation: existing applications cannot achieve deep vocabulary learning outcomes (16 codes, 7/7 teachers). T3 directly articulated: "Current apps cannot achieve deep learning." Teachers cited two primary reasons for this inadequacy. First, T5 explained: "These apps are not specifically designed for TEM-4; they're just general vocabulary apps with TEM-4 tags added." Second, T6 noted the focus on superficial learning: "Current apps just help with memorisation; they're too boring and don't provide application scenarios."

Learning Needs

Teachers provided concrete, evidence-based guidance for designing effective TEM-4 vocabulary learning interventions (108 codes). Figure 3 shows the priority ranking of recommended features.

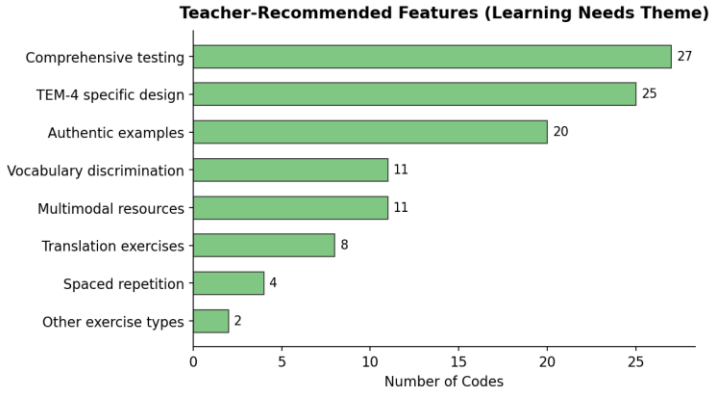


Figure 3. Teacher-recommended features (Learning Needs Theme)

Three features received universal endorsement (7/7 teachers), indicating essential rather than optional requirements: comprehensive testing with immediate feedback (27 codes), TEM-4 specific design (25 codes), and authentic example sentences from real examinations (20 codes). T2 stated that "testing is necessary—it lets students verify whether they have mastered the vocabulary." T1 recommended "example sentences from real TEM-4 tests to help students understand how vocabulary is actually used in the exam." Teachers identified vocabulary discrimination practice (11 codes, 7/7) and translation exercises (8 codes, 5/7) as high-priority exercise types. T1 explained: "Vocabulary discrimination practice is critical—students lose many points on this question type." Multimodal resources also received universal mention (11 codes, 7/7), with teachers advocating for integrated audio, video, and visual materials to support different learning styles.

Integration of Quantitative and Qualitative Findings

Following the parallel mixed-methods design, this study integrated the quantitative and qualitative findings through a joint display to identify patterns of convergence, complementarity, and divergence. This integration serves to validate findings across methods and

provide a richer understanding through multiple perspectives.

Table 10. *Joint display: integration of student questionnaire and teacher interview findings*

Dimension	Student Questionnaire (N = 240)	Teacher Interviews (N = 7)	Integration
Necessities	High recognition (M = 3.98); test performance highest (M = 4.13)	Universal confirmation (21 codes, 7/7); exam pressure dominant (69 codes); declining motivation	Convergence on recognition; Complementarity: teachers reveal sustainability challenges
Lacks	Moderate level (M = 3.62); ineffective strategies (M = 3.73); difficulty achieving deep learning (M = 3.55)	Most prominent (148 codes, 31.2%); 3:1 receptive-productive gap; 25-30% breadth gap	Convergence on gaps; Complementarity: teachers provide specific quantification
Wants: APP	High attitudes (M = 3.80); enjoy APP learning (72.1%)	High acceptance (25 codes, 7/7); apps cannot achieve deep learning (16 codes, 7/7)	Convergence on acceptance; Critical gap: desire vs. tool capability
Wants: Deep Learning	High expectations (M = 3.71); receptive-productive skills (M = 3.90)	Current apps lack TEM-4 specificity (5/7); focus on superficial memorisation	Complementarity: students express needs; teachers explain tool failures
Learning Needs	Input: word meaning (M = 4.04), discrimination (M = 3.87); Exercise: MC (M = 3.88); LOW: games (M = 2.61)	Testing essential (27 codes); TEM-4 design (25 codes); discrimination critical (11 codes)	Strong convergence on structured exercises and the importance of discrimination

The integration of data in Table 10 reveals three key patterns. First, both strands show substantial convergence regarding recognition of vocabulary importance, acceptance of mobile learning, and preference for structured exercises. Students and teachers independently arrived at consistent conclusions about TEM-4 vocabulary learning needs. Second, the qualitative data complement the quantitative findings through specific quantification—teachers identified the 3:1 receptive-productive ratio and 25-30% breadth gap that student self-reports did not capture, while also explaining why current applications fail to meet student expectations. Third, a critical gap emerges from integrating the Wants dimension: students demonstrate high APP acceptance (M = 3.80) and high expectations for deep learning functions (M = 3.71), yet teachers unanimously report existing applications cannot achieve these outcomes, identifying a significant opportunity for TEM-4-specific intervention development.

The mixed-methods needs analysis reveals a comprehensive picture of PGUI English majors' TEM-4 vocabulary learning needs. Students demonstrate high awareness of vocabulary necessity (M = 3.98), perceive moderate learning gaps (M = 3.62), and express strong acceptance of mobile-assisted learning (M = 3.80) with high expectations for deep learning functions (M = 3.71). Teachers corroborate these findings and provide crucial additional insights. They highlight the multidimensional nature of vocabulary deficiencies—particularly the 3:1 receptive-productive gap—and confirm that existing mobile applications fundamentally fail to achieve deep learning outcomes despite widespread use. The integration of both data strands establishes clear specifications for developing a TEM-4-specific mobile application incorporating comprehensive testing with immediate feedback, authentic TEM-4 examples, vocabulary discrimination practice, translation exercises, and multimodal resources.

Discussion

The most striking finding concerns the profound gap between receptive and productive vocabulary knowledge among PGUI students. Teacher interviews quantified this

disparity at approximately 3:1—students can recognise around 6,000 words yet deploy fewer than 2,000 productively. This finding resonates with Li and Hafner's (2022) investigation, which examined mobile-assisted vocabulary learning among Chinese EFL learners, reporting that although both receptive and productive vocabulary knowledge improved, the enhancement was more substantial in receptive knowledge. However, the present study reveals that this imbalance is considerably more pronounced in the PGUI context than previously documented. While Webb (2008) and subsequent researchers (Mohammadi et al., 2024; Yilmaz & Kavanoz, 2025) have established that receptive vocabulary typically exceeds productive vocabulary, the magnitude of the gap identified here—with students recognising three times more words than they can actively use—suggests that PGUI students face qualitatively different challenges from their counterparts at more selective institutions. This finding challenges the implicit assumption in much vocabulary research that the receptive-productive relationship follows consistent patterns across institutional contexts and calls for targeted interventions specifically designed for this underserved population.

The multidimensional nature of vocabulary deficiencies identified in this study provides empirical support for Wang's (2019) deep vocabulary learning framework. Students demonstrated gaps not merely in vocabulary breadth—falling 25-30% below the 8,000-word TEM-4 requirement—but also in vocabulary depth, organisation, and meta-vocabulary knowledge. These findings corroborate Wang and Zhang's (2024) research showing that Chinese EFL learners struggle particularly with morphological awareness and semantic encoding. Yet the present study extends this work by demonstrating how these deficiencies interconnect: students' reliance on surface-level memorisation, as reported by both quantitative ($M = 3.55$ for difficulty achieving deep learning) and qualitative data, perpetuates isolated word knowledge rather than the interconnected lexical networks that Nation (2001) and Meara (2005) identify as essential for authentic language use. This interconnection suggests that vocabulary interventions must address multiple dimensions simultaneously rather than treating breadth and depth as sequential developmental goals.

A particularly noteworthy finding is the critical disconnect between students' high acceptance of mobile-assisted vocabulary learning ($M = 3.80$) and teachers' collective assessment that existing applications cannot achieve deep learning outcomes. This finding extends Li's (2024) critique that Chinese vocabulary applications lack systematic integration of language learning theories by providing stakeholder evidence from both learner and instructor perspectives. The meta-analysis by Zhou and Zhou (2025), reporting large effect sizes (Hedges' $g = 1.28$) for mobile vocabulary applications, suggests that MAVL can be highly effective—yet the present findings indicate that this potential remains largely unrealised in the Chinese context. Teachers specifically criticised existing applications for their generic design and focus on superficial memorisation, echoing Mohammadi et al.'s (2024) observation that most applications emphasise recognition over production. This convergence of evidence suggests that the problem lies not with mobile technology itself but with application design, which prioritises engagement features and gamification over theoretically grounded pedagogical approaches.

The learning needs data offer important insights that both confirm and complicate existing assumptions about vocabulary pedagogy. Students strongly preferred structured, examination-oriented exercises such as multiple choice, true/false, and sentence translation, while gamified elements received notably lower ratings. This preference challenges the prevailing enthusiasm for gamification in the MAVL literature (Teymouri, 2024). This preference pattern may reflect the examination-driven learning culture in Chinese higher education (Wu et al., 2025).

However, it also suggests that PGUI students possess pragmatic awareness of what supports their actual learning goals. Simonnet et al.'s (2025) systematic review of technology-assisted vocabulary learning emphasised the importance of aligning tool features with learner needs—a principle echoed in the present findings. Teachers' strong endorsement of vocabulary discrimination practice, a question type associated with significant score loss, suggests that examination-aligned design may be more impactful than generic vocabulary building.

Collectively, these findings point toward a clear design specification for TEM-4 vocabulary applications. Rather than replicating the superficial approaches of existing tools, effective applications must be purpose-built around Wang's (2019) deep learning dimensions while addressing the specific receptive-productive gap that characterises PGUI learners. The integration of comprehensive testing with immediate feedback, authentic TEM-4 examples, and multimodal resources—features universally supported by teachers—provides a concrete foundation for intervention development. The emphasis on vocabulary discrimination practice and contextual application, combined with spaced repetition mechanisms supported by dual coding theory (Clark & Paivio, 1991), offers a pathway for bridging the gap between recognition and production that current applications fail to address. However, perhaps the most significant contribution of this study lies in its identification of PGUI students as a distinct population with specific needs that differ meaningfully from the general Chinese EFL learner population that dominates existing research. This recognition should inform not only application design but also future research directions in vocabulary acquisition and mobile-assisted language learning.

Conclusion

This study employed Hutchinson and Waters' (1987) needs analysis framework to investigate the TEM-4 vocabulary learning needs of English majors at PGUIs in China. Through a convergent parallel mixed-methods design integrating questionnaire data from 240 students with semi-structured interviews from seven experienced teachers across five institutions, the research identified substantial vocabulary deficiencies characterised by a striking 3:1 receptive-productive gap, high learner acceptance of mobile-assisted vocabulary learning coupled with dissatisfaction toward existing applications, and clear preferences for examination-aligned learning content over gamified approaches.

The findings carry several important implications for stakeholders in vocabulary education. For mobile application developers, the critical disconnect between learner expectations and current tool capabilities suggests that effective TEM-4 vocabulary applications must be purpose-built around Wang's (2019) deep learning framework rather than adopting generic vocabulary-building approaches. This entails systematic attention to vocabulary depth through morphological analysis and collocation practice, vocabulary organisation through semantic network construction, and productive skill development through contextualised output tasks—dimensions that current applications largely neglect in favour of superficial memorisation features. For language educators, the evidence of ineffective classroom vocabulary instruction and students' reliance on surface-level memorisation strategies points to the need for pedagogical interventions that explicitly teach vocabulary learning strategies and promote deeper lexical processing. Teachers should consider integrating well-designed mobile tools as supplements to classroom instruction, particularly for addressing the fragmented study time characteristic of contemporary university life. For curriculum designers and policymakers, the findings underscore the importance of recognising PGUI students as a distinct learner population whose vocabulary development

needs differ meaningfully from those at more selective institutions, warranting targeted resource allocation and instructional support.

This study contributes to the mobile-assisted language learning literature by challenging the prevailing assumption of homogeneous learner needs across institutional contexts. It offers concrete design specifications grounded in both student and teacher perspectives rather than researcher assumptions alone. However, several limitations should be acknowledged. The sample was drawn exclusively from PGUIs, which may constrain generalisability to other institutional settings such as key universities or vocational colleges. The reliance on self-reported data may not fully capture actual vocabulary proficiency levels, and the cross-sectional design precludes causal inferences about the relationship between identified needs and learning outcomes.

These limitations, rather than diminishing the study's contribution, illuminate promising directions for future inquiry. Subsequent research could incorporate objective vocabulary assessments such as the Vocabulary Levels Test to triangulate self-reported data. Researchers could also extend the investigation to diverse institutional contexts for comparative analysis and employ longitudinal designs to track how learner needs evolve across different stages of TEM-4 preparation. Most importantly, the design specifications emerging from this needs analysis provide a robust foundation for developing and empirically evaluating TEM-4-specific vocabulary applications—applied research that could directly benefit the rapidly growing yet underserved PGUI student population. As mobile technology continues to reshape language learning landscapes in China and beyond, ensuring that technological innovations address authentic learner needs rather than assumed ones becomes increasingly critical. This study represents a meaningful step toward that goal, offering both theoretical insights into vocabulary learning challenges and practical guidance for bridging the gap between learner needs and mobile-assisted solutions in Chinese higher education.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: During the preparation of this manuscript, the authors used Grammarly to improve language clarity and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

About the Authors:

Haiping Duan is a PhD candidate in Teaching English as a Second Language (TESL) at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM), with nine years of teaching and research experience in higher education. Her research focuses on technology-enhanced language learning, particularly in mobile-assisted vocabulary acquisition in EFL contexts. ORCID: https://orcid.org/0009-0004-3979-5928

Melor Md Yunus is a Professor at the Faculty of Education, Universiti Kebangsaan Malaysia. She holds a M.A. in TESL from the Arizona State University, USA. She then earned her Ph.D. in Education (TESL) from the University of Bristol, UK. Her areas of concentration are TESL, language pedagogy and the use of technology in TESL. ORCID: https://orcid.org/0000-0001-7504-7143

Hanita Hanim Ismail is a senior lecturer at the Faculty of Education of UKM, specialising in TESL and English literature. Her research focuses on enhancing teaching methods, curriculum development, and materials for English instruction. ORCID: https://orcid.org/0000-0003-3121-8822

References

Abd Rahman, E., Md Yunus, M., Hashim, H., & Khadirah Ab. Rahman, N. (2024). A Digital Approach to Teach Synthesis Writing for a Defence University: A Needs Analysis. *Arab World English Journal, 15*(1), 327–351. https://doi.org/10.24093/awej/vol15no1.21

Akşit, Z., & Saygı, Ş. (2024). Which components of word knowledge do EFL learners learn? *Language Learning in Higher Education, 14*(1), 253–268. https://doi.org/10.1515/cercles-2024-2001

Barzani, S. H. H., & Omar, F. R. (2024). WhatsApp-mediated vocabulary learning: A study of Kurdish EFL learners. *Arab World English Journal, 15*(4), 364–376. https://doi.org/10.24093/awej/vol15no4.23

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cai, Z. (2025). Enhancing EFL vocabulary and psychological well-being in Chinese undergraduates through adaptive digital games with mind mapping and Runge–Kutta modeling. *Frontiers in Psychology, 16*, 1644162. https://doi.org/10.3389/fpsyg.2025.1644162

Cui, X. (2025). Investigating students' academic language-related challenges and their interplay with English proficiency and self-efficacy on EMI success in Transnational Education (TNE) programs in China. *English for Specific Purposes*, *78***,** 53–69. https://doi.org/10.1016/j.esp.2024.12.002

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–210. https://doi.org/10.1007/BF01320076

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Dörnyei, Z., & Csizér, K. (2002). Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied Linguistics, 23*(4), 421–462. https://doi.org/10.1093/applin/23.4.421

Fisal, F. A. M., & Said, N. E. M. (2023). ESL pre-university learners' needs analysis for web-based English academic vocabulary learning resource. *International Journal of Learning, Teaching and Educational Research, 22*(4), 195–216. https://doi.org/10.26803/ijlter.22.4.12

Gürler, İ., & Konca, M. Y. (2023). Lacks, Necessities and Wants of Prospective English Teachers in ELT Departments: A Needs Analysis and Sustainability. *Discourse and Communication for Sustainable Education, 14*(1), 24–35. https://doi.org/10.2478/dcse-2023-0003

Hua, Y., Guo, Q., & Luo, Y. (2024). Vocabulary learning in EMI courses: A longitudinal study of Chinese university students over a term. *System, 12*6, 103479.

https://doi.org/10.1016/j.system.2024.103479
Huang, X., & Wang, L. (2024). Investigating the contribution of student engagement and foreign language enjoyment to English majors' TEM4-related positive emotions: A positive reinforcement perspective. *Learning and Motivation, 88,* 102043. https://doi.org/10.1016/j.lmot.2024.102043
Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.
Kukulska-Hulme, A., Lee, H., & Norris, L. (2017). Mobile learning revolution: Implications for language pedagogy. *The handbook of technology and second language teaching and learning*, 217-233. https://doi.org/10.1002/9781118914069.ch15
Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL, 20*(3), 271–289. https://doi.org/10.1017/S0958344008000335
Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*(3), 365–391. https://doi.org/10.1111/0023-8333.00046
Li, C. (2014). Development, problems and solutions: A critical review of current situation of college English language education in mainland China. *Arab World English Journal, 5*(3), 133–148. Retrieved from https://awej.org/images/AllIssues/Volume5/Volume5number3September/22.pdf
Li, C. (2024). A review of theories, pedagogies and vocabulary learning tasks of English vocabulary learning apps for Chinese EFL learners. *Journal of China Computer-Assisted Language Learning, 4*(2), 346–375. https://doi.org/10.1515/jccall-2023-0026
Li, Y., & Hafner, C. A. (2022). Mobile-assisted vocabulary learning: Investigating receptive and productive vocabulary knowledge of Chinese EFL learners. *ReCALL, 34*(1), 66–80. https://doi.org/10.1017/S0958344021000161
Mansor, N. S., Abu Hassan, N. A. S., Amat, R., & Tambunan, A. R. S. (2025). Bridging the gap: A systematic review of needs analysis in English language courses in higher education. *International Journal of Research and Innovation in Social Science, IX*(VIII), 1786–1800. DOI: https://doi.org/10.47772/IJRISS.2025.908000150
Mahdi, H. S. (2018). Effectiveness of mobile devices on vocabulary learning: A meta-analysis. *Journal of Educational Computing Research, 56*(1), 134–154. https://doi.org/10.1177/0735633117698826
Mao, F., & Zhou, J. (2024). A needs analysis of ESP courses in colleges of art and design: Consensus and divergence. *PLOS ONE, 19*(6), e0305210. https://doi.org/10.1371/journal.pone.0305210
Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology, 46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x
McCarthy, M. (1984). A new look at vocabulary in EFL. *Applied Linguistics, 5*(1), 12–22. https://doi.org/10.1093/applin/5.1.12
Meara, P. (1990). A note on passive vocabulary. *Second Language Research, 6*(2), 150–154. https://doi.org/10.1177/026765839000600204
Meara, P. (2005). Lexical frequency profiles: A Monte Carlo analysis. *Applied Linguistics,*

26(1), 32–47. https://doi.org/10.1093/applin/amh037

Ministry of Education. (2024). *Statistics on China's higher education institutions*. Ministry of Education of the People's Republic of China.

Mohammadi, M., Valizadeh, M., Zohdi Jalal, P., & Xodabande, I. (2024). University students' academic vocabulary development through mobile-assisted learning: Exploring the impacts on receptive and productive knowledge. *Heliyon, 10*(4), e28103. https://doi.org/10.1016/j.heliyon.2024.e28103

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Nation, I. (2006). How Large a Vocabulary is Needed For Reading and Listening? *The Canadian Modern Language Review, 63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nation, P., & Meara, P. (2013). 3 Vocabulary. In *An introduction to applied linguistics* (pp. 44-62). Routledge.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan. https://doi.org/10.1057/9780230293977

Sun, D., Chen, Z., & Zhu, S. (2023). What affects second language vocabulary learning? Evidence from multivariate analysis. *Frontiers in Education, 8*, Article 1210640. https://doi.org/10.3389/feduc.2023.1210640

Simonnet, E., Loiseau, M., & Lavoué, É. (2025). A Systematic Literature Review of Technology-Assisted Vocabulary Learning. *Journal of Computer Assisted Learning*, *41*(1), e13096. https://doi.org/10.1111/jcal.13096

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tao, X., & Modehiran, P. (2023). Experiences and preferences of Chinese EFL students in vocabulary learning and mobile applications in vocabulary learning in a private international university in China. *The New English Teacher, 17*(2), 1–25. https://www.proquest.com/openview/c6e26547d85093a1151a46721a5392f6/1?pq-origsite=gscholar&cbl=4531125

Teymouri, R. (2024). Recent developments in mobile-assisted vocabulary learning: A mini review of published studies focusing on digital flashcards. *Frontiers in Education, 9,* 1496578. https://doi.org/10.3389/feduc.2024.1496578

Wang, Z. (2019). *A study on deep learning of English vocabulary based on mobile-assisted language learning*. Guangming Daily Press.

Wang, T., & Zhang, H. (2024). Contributions of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge among Chinese EFL learners: A structural equation modeling analysis. *Applied Linguistics Review, 15*(3), 1047–1085. DOI: https://doi.org/10.1515/applirev-2022-0022

Wang, T., & Zhang, H. (2025). Reciprocal effects of vocabulary breadth, vocabulary depth, and reading comprehension: A cross-lagged panel analysis in Chinese-speaking EFL learners. *Humanities and Social Sciences Communications, 12*(1), 1-10. https://doi.org/10.1057/s41599-025-04694-2

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition, 30*(1), 79–95. https://doi.org/10.1017/S0272263108080042

Wilkins, D. A. (1972). *Linguistics in language teaching.* Edward Arnold.

Wu, J. H., Chen, P. F., & Zhang, N. (2025). Applying transformative learning theory to Chinese college English listening and speaking education programs. *Language Teaching Research*, 13621688251352272. https://doi.org/10.1177/136216882513522

Yang, Y., & He, X. (2025). Lexical richness in Chinese university students' EFL writing: A corpus-based comparison. *Humanities and Social Sciences Communications, 12*, 1043. https://doi.org/10.1057/s41599-025-05560-x

Yilmaz, S., & Kavanoz, S. (2025). Measuring the effect of receptive and productive vocabulary size on foreign language skills. *PORTA LINGUARUM*, (44). https://doi.org/10.30827/portalin.vi44.31870

Zhou, Y., & Zhou, M. (2025). A meta-analysis on mobile-assisted vocabulary learning: Do mobile applications help?. *ReCALL*, 1-19. https://doi.org/10.1017/S0958344025100335