Volume 18 Number 2 (2026)
July-December 2026
Page: 859-874

E-ISSN: 2656-9779
P-ISSN: 1907-6355
DOI: 10.37680/qalamuna.v18i2.8977

QALAMUNA
JURNAL PENDIDIKAN, SOSIAL, DAN AGAMA

A Lexical Profiling Analysis of State Islamic Higher Education Entrance Examinations: Implications for Vocabulary Breadth and Test Accessibility

Dhinuk Puspita Kirana [1], Aries Fitriani [2]

[1] Universitas Islam Negeri Kiai Ageng Muhammad Besari Ponorogo, Indonesia; dhinuk@uinponorogo.ac.id
[2] Universitas Islam Negeri Kiai Ageng Muhammad Besari Ponorogo, Indonesia; ariesfitriani@uinponorogo.ac.id

Received: 20/02/2026 Revised: 19/05/2026 Accepted: 16/07/2026

Abstract

Vocabulary knowledge plays a crucial role in reading comprehension and is widely recognised as one of the strongest predictors of successful reading performance among second language learners(Reynolds, 2021); however, the lexical demands of university entrance examinations may not always correspond to the vocabulary knowledge developed through secondary-school English curricula. This issue is particularly relevant in the State Islamic Higher Education Admission Test (UM-PTKIN), whose reading passages frequently incorporate Islamic and culturally specific content. Therefore, this study investigates the lexical profile of the English component of the UM-PTKIN examination and examines its implications for vocabulary breadth and test accessibility. Using a descriptive quantitative design, four English test scripts from the 2019 UM-PTKIN examination were analyzed through lexical profiling with RANGE, FREQUENCY, VocabProfile, and AntConc. The analysis focused on tokens, word types, and word families, with word families adopted as the primary unit because they provide a more realistic estimate of learners' vocabulary knowledge in educational assessment contexts. The findings revealed that the test relied predominantly on high-frequency vocabulary; however, a substantial proportion of vocabulary items fell outside established frequency lists. Many of these off-list items consisted of Islamic terminology, proper nouns, and culturally embedded expressions rather than advanced English vocabulary. These findings suggest that vocabulary demands in UM-PTKIN are influenced not only by lexical frequency but also by thematic and cultural familiarity. The study provides implications for test development, curriculum alignment, and vocabulary instruction in Islamic higher education contexts.

Keywords Lexical Profiling; Vocabulary Demands; UM-PTKIN; Islamic Higher Education; Test Acessiblity

Corresponding Author
Dhinuk Puspita Kirana
Universitas Islam Negeri Kiai Ageng Muhammad Besari Ponorogo, Indonesia; dhinuk@uinponorogo.ac.id

1. INTRODUCTION

Despite the well-established role of vocabulary knowledge in language assessment, empirical research examining the lexical characteristics of the UM-PTKIN English test remains extremely limited. Previous studies conducted in various educational contexts have demonstrated that lexical profiling can

provide valuable insights into text difficulty, vocabulary demands, assessment validity, and learner accessibility **(Castellano, 2024; Dang, 2022; Dang & Webb, 2016)**. Research on international proficiency examinations such as IELTS, TOEFL, and other standardized assessments has shown that lexical coverage is closely associated with reading comprehension and test performance, and that comprehension increases substantially as the percentage of known words in a text increases (Laufer, 2010; Schmitt, 2011)

However, little is known about the vocabulary demands of the English component of UM-PTKIN, despite its role as one of Indonesia's most important university entrance examinations for Islamic higher education. This lack of evidence represents more than a simple research gap. Without empirical information regarding the lexical demands of the examination, it is difficult to determine whether test performance reflects candidates' English proficiency or is disproportionately influenced by unfamiliar vocabulary. As a result, important issues related to test validity, accessibility, and fairness remain insufficiently understood. Recent language assessment research has emphasized that transparency regarding linguistic demands is a fundamental requirement for ensuring equitable assessment practices and valid score interpretations (Kremmel & Alderson, 2023; Aryoudust et al., 2024)

While lexical profiling studies have examined international tests such as TOEFL, IELTS, and EFL national examinations, no published study to date has systematically investigated the lexical characteristics of UM-PTKIN. More importantly, UM-PTKIN differs from most university entrance examinations because it incorporates Islamic-themed reading materials that may introduce culturally specific vocabulary absent from conventional frequency lists. As a result, findings from previous lexical profiling studies cannot be directly generalized to this context. This gap justifies the need for a dedicated investigation of UM-PTKIN vocabulary demands and their implications for assessment accessibility and fairness.

The need for such investigation is particularly important in the context of UM-PTKIN because the examination differs from general university entrance tests. Unlike assessments that predominantly employ general academic topics, UM-PTKIN reading passages frequently incorporate Islamic history, religious practices, Muslim public figures, and culturally specific educational themes. The inclusion of such content may introduce vocabulary items that are absent from conventional frequency lists and therefore contribute to lexical demands in ways that cannot be fully explained by frequency-based measures alone. Consequently, examining the lexical profile of UM-PTKIN is important not only for identifying vocabulary frequency levels but also for understanding how thematic and cultural content may influence reading accessibility and assessment fairness.

Although the present study analyzes the 2019 UM-PTKIN examination, the investigation remains academically relevant. Publicly available information and examination documentation indicate that the overall purpose, reading-oriented format, and assessment framework of UM-PTKIN have remained broadly stable across administrations. In language testing research, archived test materials are widely recognized as valuable resources for investigating lexical demands when the underlying assessment construct remains relatively consistent over time **(Bao & Peng, 2024)**. Therefore, the 2019 examination provides a meaningful benchmark for understanding the vocabulary demands of English language assessment within State Islamic Higher Education.

This study addresses these issues by conducting a systematic lexical profiling analysis of the English component of the UM-PTKIN examination. In addition to describing vocabulary frequency distributions, the study explores the extent to which culturally embedded and Islamic-related vocabulary contributes to the overall lexical profile of the test. By doing so, the study seeks to provide empirical evidence that may support test development, curriculum alignment, and vocabulary instruction for prospective university students.

To achieve these objectives, the study addresses the following research questions:

a. What are the lexical characteristics of the UM-PTKIN English test in terms of tokens, word types, and word families?

b. How are vocabulary items distributed across different frequency levels and off-list categories?

c. To what extent do off-list vocabulary items reflect Islamic-related and culturally specific content within the examination texts?

d. How does the lexical profile vary across the four UM-PTKIN test scripts, and what implications do these lexical characteristics have for vocabulary breadth and test accessibility among prospective university students?

2. METHODS

Research Design

This study employed a descriptive quantitative design using corpus-based lexical profiling techniques to examine the vocabulary characteristics of the English component of the State Islamic Higher Education Admission Test (UM-PTKIN). Lexical profiling was selected because it provides a systematic means of identifying vocabulary frequency distributions, lexical coverage, and vocabulary demands within assessment texts. This approach has been widely applied in vocabulary and language assessment research to evaluate the accessibility and difficulty of educational materials and standardized examinations

Corpus and Data Source

The corpus consisted of four English test scripts administered in the 2019 UM-PTKIN examination (codes 1911, 1913, 1915, and 1917). These scripts were selected because they represent officially standardized test forms used within the national admission system for State Islamic Higher Education institutions in Indonesia. The corpus contained reading passages, questions, and relevant textual content encountered by test takers during the examination.

Although the corpus originated from the 2019 administration, its use remains methodologically justified. Publicly available documentation indicates that the general purpose, reading-oriented format, and assessment framework of UM-PTKIN have remained broadly stable across subsequent administrations. Consequently, the 2019 examination provides a meaningful baseline for investigating vocabulary demands in this high-stakes assessment context.

Data Preparation and Corpus Cleaning

Prior to analysis, all test scripts were converted into plain-text format and manually proofread to ensure textual accuracy. To avoid distortions in lexical frequency counts, several non-lexical elements were excluded from the corpus, including item numbers, answer-option labels (A, B, C, and D), numerals, punctuation symbols, measurement units, website protocols (e.g., http, https, www), and other non-linguistic strings.

Proper nouns were retained in the corpus because they form part of the authentic lexical input encountered by examinees during test administration. Excluding such items could underestimate the actual vocabulary demands experienced by test takers. However, because proper nouns frequently appear as off-list items in lexical profiling, their occurrence was subsequently examined qualitatively to distinguish culturally specific vocabulary from genuinely low-frequency English vocabulary. This procedure was intended to avoid overestimating lexical difficulty while preserving the authenticity of the corpus.

Lexical Profiling Instruments

Three complementary analytical tools were employed.

First, RANGE was used to calculate lexical coverage and determine the distribution of tokens, word types, and word families across frequency bands. RANGE served as the primary instrument for quantitative lexical profiling.

Second, VocabProfile(Cobb, 2020) was employed to verify and visualize vocabulary distribution across frequency levels. Although both RANGE and VocabProfile perform lexical profiling functions, they were used for different purposes. RANGE **(Nation & Heatley, 2002)**generated quantitative statistics, whereas VocabProfile facilitated validation of vocabulary classification and lexical coverage patterns.

Third, AntConc(Anthony, 2022) was used as a concordancing tool to examine word occurrences in context and to identify thematic patterns among off-list vocabulary items. AntConc(Anthony, 2017) was particularly useful for determining whether off-list words consisted primarily of proper nouns, Islamic terminology, geographical references, or genuinely low-frequency English vocabulary.

The combined use of these tools enhanced analytical reliability through methodological triangulation, allowing quantitative profiling results to be verified and interpreted through contextual corpus analysis.

Lexical Unit of Analysis

The present study adopted word families as the primary unit of analysis. Following (Nation, 2022a, 2022b), a word family consists of a base word together with its inflected and derived forms that are assumed to be recognisable by language users possessing a reasonable level of morphological knowledge. For example, teach, teacher, teaching, and taught are treated as members of the same word family.

Word families were selected rather than lemmas because university entrance examination candidates are generally expected to possess sufficient morphological awareness to recognize common inflectional and derivational relationships. Consequently, word-family analysis provides a more realistic estimate of vocabulary knowledge and lexical coverage in educational assessment contexts. The use of word families is also consistent with established procedures in vocabulary research and lexical profiling studies that investigate the relationship between vocabulary knowledge and text comprehension(Nation, 2011; Webb, 2008). This approach is also widely employed in vocabulary-size and lexical-threshold research, thereby facilitating comparison with previous studies.

Reference Word Lists

Lexical profiling was conducted using the General Service List (GSL) (West, 1953) and the Academic Word List (AWL)(Coxhead, 2000), which remain among the most widely used reference lists in vocabulary research and educational assessment. Although newer resources such as the New General Service List (NGSL)(EAP Foundation, 2022) and BNC/COCA-based lists(I. S. P. Nation, 2017) have been developed, the GSL and AWL continue to provide a useful benchmark for comparison because a substantial body of lexical profiling research has employed these lists.

To complement the GSL-AWL analysis, AntConc (Anthony, 2022) was also used to examine vocabulary distribution across broader frequency levels derived from BNC/COCA-based frequency lists. This dual approach enabled the study to identify both the proportion of high-frequency and academic vocabulary and the extent to which the corpus contained lower-frequency and specialized lexical items.

Data Analysis

The analysis proceeded in four stages. First, corpus statistics were calculated to identify the number of tokens, word types, and word families, as these measures provide a fundamental description of corpus composition and lexical complexity (Anandi, 2024; Tahir, 2025). Second, lexical profiling was

conducted to determine vocabulary distribution across frequency bands and off-list categories, a procedure commonly employed to investigate vocabulary demands and text accessibility (Synchak, 2025). Third, concordance analysis was performed to identify the characteristics and contextual use of off-list vocabulary items, following corpus-based approaches that utilize keyword-in-context (KWIC) analysis to examine lexical patterns (Anthony, 2022; Weisser, 2012). Finally, the results were compared across the four test scripts to examine lexical consistency and variation among parallel examination forms, an approach frequently adopted in corpus-assisted studies to identify similarities and differences across texts (Anandi, 2024; Tahir, 2025).

3. FINDINGS AND DISCUSSIONS

Lexical Characteristics of the UM-PTKIN English Test

To determine the lexical characteristics of the 2019 UM-PTKIN English examination, all four versions of the test scripts were subjected to corpus-based analysis. Each test version consisted of four short reading passages accompanied by multiple-choice questions and contained between approximately 757 and 874 running words. On average, each test exposed candidates to around 827 running words, indicating a relatively consistent lexical load across test versions. This level of exposure aligns with findings from recent studies suggesting that standardized entrance examinations typically present learners with between 800 and 1,000 tokens per test to balance reliability and processing demands (Pellicer-Sánchez, 2024).

Table 1. Results of the RANGE Analysis of the UM-PTKIN English Tests

Word List Category	Tokens n (%)	Types n (%)	Word Families
First 1,000 Words	2,245 (67.8)	428 (47.9)	331
Second 1,000 Words	183 (5.5)	92 (10.3)	76
Academic/Mid-Frequency Words (List 3)	146 (4.4)	69 (7.7)	60
Off-List Words	735 (22.2)	304 (34.0)	–
Total	**3,309 (100.0)**	**893 (100.0)**	**467**

The lexical profiling analysis revealed that the UM-PTKIN English test corpus was predominantly composed of high-frequency vocabulary, with the first 2,000 word families accounting for the largest proportion of lexical coverage. This finding is broadly consistent with previous studies of standardized English examinations, which have reported that high-frequency vocabulary forms the lexical foundation of assessment texts and plays a central role in supporting reading comprehension **(Dang, 2022; Dang & Webb, 2016)**. The predominance of high-frequency vocabulary suggests that the examination was designed to remain accessible to learners who possess a solid foundation in core English vocabulary.

However, the finding that approximately one-fifth of the corpus consisted of off-list vocabulary deserves particular attention. At first glance, such a proportion appears unusually high for a standardized entrance examination. Previous lexical-threshold research has suggested that readers typically require approximately 95% lexical coverage for minimally adequate comprehension and around 98% coverage for comfortable and independent reading (Laufer, 2010; Nation, 2022b, 2022a; Schmitt, 2011; Schmitt, 2010). Consequently, the lexical profile identified in the present study indicates that mastery of the first 2,000 word families alone is unlikely to guarantee successful comprehension of UM-PTKIN reading passages.

Nevertheless, the interpretation of lexical difficulty requires careful consideration. The proportion

of off-list vocabulary identified in the UM-PTKIN corpus (22.2%) appears considerably higher than that reported in several studies of international English proficiency examinations, where off-list vocabulary typically represents a smaller proportion of total lexical coverage (McLean, 2021; Nguyen, 2022). However, unlike TOEFL and IELTS reading texts, which primarily employ academic and general-interest topics, UM-PTKIN passages frequently contain Islamic terminology, religious institutions, and culturally specific references. Consequently, the elevated percentage of off-list vocabulary reflects thematic specialization rather than excessive linguistic complexity. This finding suggests that lexical difficulty in UM-PTKIN is partially shaped by domain-specific discourse, highlighting the need to consider cultural and contextual factors alongside frequency-based measures when evaluating test accessibility.

From a pedagogical perspective, the findings suggest that prospective UM-PTKIN candidates require more than knowledge of general high-frequency vocabulary. Successful performance may also depend on the ability to infer meaning from context and to process culturally specific vocabulary encountered in Islamic educational discourse. This observation supports recent arguments that vocabulary demands in high-stakes assessments should be interpreted not solely through frequency-based measures but also through broader considerations of contextual and disciplinary literacy.

Table 2. Topics of Reading Passages Across the Four UM-PTKIN Test Versions

Reading Passage	Code 1911	Code 1913	Code 1915	Code 1917
Text 1	Idul Fitri	Sekaten	Welcoming a Child in Islam	Hinduism
Text 2	Muhammad Ali	Yusuf Islam	Abdurrahman Wahid	Malala Yousafzai
Text 3	Masjid Qiblatain	Masjid Qiblatain	Masjid Qiblatain	Masjid Qiblatain
Text 4	Rain	Mountains	The Invention of Miswak	Al-Azhar University

An examination of thematic content revealed that the four test versions shared one identical passage on Masjid Qiblatain, positioned consistently as the third reading passage in each test. Such repetition reflects test standardization practices commonly adopted to ensure equivalence across parallel test forms (Green, 2021). Of the thirteen distinct passage topics identified across the test scripts, only three were general in nature—namely Hinduism, Rain, and Mountains. The remaining ten passages were closely associated with Islamic contexts and could be grouped into three thematic categories: Islamic figures (e.g., Muhammad Ali, Yusuf Islam, Abdurrahman Wahid, and Malala Yousafzai), Islamic practices and cultural events (e.g., Idul Fitri, Sekaten, and welcoming a child in Islam), and Islamic history and institutions (e.g., the invention of miswak, Masjid Qiblatain, and Al-Azhar University).

This thematic pattern helps explain the substantial proportion of off-list vocabulary identified in the lexical profiling analysis. Many off-list items consisted of Islamic-related terms, proper nouns, and culturally specific references that are not typically included in general high-frequency word lists such as the GSL or AWL. Therefore, the high proportion of off-list words should not be interpreted solely as evidence of advanced lexical difficulty, but rather as a reflection of the religious and cultural orientation of the test content. From a test fairness perspective, the inclusion of culturally embedded vocabulary may provide an advantage to candidates who are already familiar with Islamic concepts and terminology, while potentially increasing processing demands for those with limited background knowledge of such topics. This finding supports previous research suggesting that topic familiarity and cultural knowledge can influence reading comprehension performance independently of linguistic proficiency **(Koda, 2005; S. McLean & Kramer, 2015; Sung, 2021)**. Consequently, vocabulary demands in UM-PTKIN should be interpreted not only in terms of word frequency levels but also in relation to the cultural and disciplinary contexts represented in the examination texts.

Analysis at the word-family level identified a total of 467 word families across the corpus. Of these, 331 word families belonged to the first 1,000 frequency band, followed by 76 word families from the second 1,000 and 60 word families from the third frequency level. Overall, approximately 73.3% of the vocabulary in the UM-PTKIN English tests was drawn from the first 2,000 most frequent word families. Although this finding confirms the substantial contribution of high-frequency vocabulary to the test texts, the coverage remains considerably below the lexical thresholds commonly associated with independent reading comprehension. Previous studies suggest that readers typically require at least 95% lexical coverage for minimally adequate comprehension and approximately 98% coverage for comfortable, unassisted reading (I. S. P. Nation, 2006, 2022a, 2022b). Consequently, mastery of only the first 2,000 word families is unlikely to guarantee successful comprehension of the UM-PTKIN reading passages. Rather than rendering the texts entirely unreadable, this gap indicates that test takers must rely on additional vocabulary knowledge, contextual inference, and background knowledge to compensate for unfamiliar lexical items. The finding is particularly relevant in the context of UM-PTKIN, where culturally embedded and Islamic-related vocabulary contributes to the lexical demands of the texts and may further influence reading accessibility among prospective university students.

Taken together, the vocabulary used in the 2019 UM-PTKIN English tests was dominated by high-frequency words, particularly those belonging to the first 2,000-word levels. However, the tests also contained a substantial number of off-list items, many of which were proper nouns and Islamic-related terms. This indicates that the lexical profile of the examination extended beyond general high-frequency vocabulary and reflected the distinctive thematic orientation of State Islamic Higher Education entrance examinations.

Vocabulary Distribution Across Frequency Levels and Off-List Categories

An analysis of the frequency list revealed that the definite article *the* (268 occurrences) was the most frequently occurring lexical item across the four UM-PTKIN English test versions. Other highly frequent words included prepositions (*in*, *of*, *to*, *for*, *with*, *by*, *on*, and *towards*), conjunctions (*and* and *but*), pronouns (*he*, *it*, *she*, and *they*), and forms of the verb *be* (*is*, *was*, and *be*). Such a distribution is consistent with findings from large English corpora and standardized assessment texts, where function words dominate frequency lists because of their essential role in establishing grammatical relationships, cohesion, and meaning construction (Biber et al., 1999; I. S. P. Nation, 2022a, 2022b; Reppen, 2008). The predominance of these items suggests that the overall lexical structure of the UM-PTKIN corpus largely reflects general patterns of English language use.

Table 3. Top Islamic-Related Vocabulary in the UM-PTKIN Corpus

Word	Frequency	Category
Muhammad	24	Islamic Figure
Prophet	22	Religious Figure
Masjid	16	Religious Institution
Muslims	15	Religious Community
Islamic	13	Religious Descriptor
Islam	13	Religion
Quran	10	Sacred Text
Jerusalem	10	Religious Location

Word	Frequency	Category
Revelation	10	Religious Concept
Makkah	8	Sacred Place

The prominence of Islamic-related vocabulary represents one of the most distinctive findings of the present study. Unlike international proficiency examinations such as TOEFL or IELTS, which predominantly employ general academic topics, the UM-PTKIN examination incorporates reading passages closely associated with Islamic history, religious practices, Muslim public figures, and educational institutions. As a result, vocabulary items such as Quran, Masjid, Qibla, Kaabah, Makkah, and Prophet appeared frequently throughout the corpus.

This finding provides an important explanation for the substantial proportion of off-list vocabulary identified in the lexical profiling analysis. Many of the words categorized as off-list were absent from conventional frequency lists not because they were linguistically difficult but because they represented culturally specific concepts. Consequently, lexical difficulty in UM-PTKIN cannot be fully explained through frequency-based approaches alone.

The findings may also be interpreted through schema theory, which suggests that readers draw upon prior knowledge structures to facilitate comprehension and meaning construction (Koda, 2005). For candidates who possess familiarity with Islamic concepts and discourse, culturally embedded vocabulary may support comprehension despite falling outside general frequency lists. Conversely, candidates with limited background knowledge may encounter additional processing demands even when the vocabulary itself is not linguistically complex. This observation raises important considerations regarding test accessibility and fairness, as reading performance may be influenced not only by English proficiency but also by topic familiarity and cultural knowledge.

The results therefore suggest that lexical profiling in Islamic higher education contexts should extend beyond simple frequency counts and consider the interaction between vocabulary knowledge, disciplinary discourse, and cultural familiarity. Such an approach provides a more comprehensive understanding of vocabulary demands in context-specific language assessments.

More noteworthy, however, is the prominence of culturally and religiously specific content words among the most frequent lexical items. The words *Muhammad* (24 occurrences) and *prophet* (22 occurrences) appeared repeatedly across the four test versions, while other frequently occurring terms included *Masjid* (16), *Muslims* (15), *Islam* (13), *Islamic* (13), *Jerusalem* (11), and *revelation* (10). To examine this pattern further, an additional frequency analysis identified forty Islamic-related lexical items in the corpus, including religious concepts (*Quran*, *Surah*, *Allah*, *Kaabah*, and *Qibla*), locations (*Makkah*, *Madinah*, *Jerusalem*, and *Cairo*), religious practices (*prayer*, *Dhuhr*, *Asar*, and *Fitri*), and Islamic educational institutions (*Al-Azhar* and *Al-Zaytunah*).

These findings help explain the relatively high proportion of off-list vocabulary reported in the lexical profiling analysis. Many of the words categorized as off-list were not advanced English vocabulary items but rather proper nouns, Islamic terminology, and culture-specific references that are absent from general English frequency lists. Consequently, the lexical demands of the UM-PTKIN English tests cannot be interpreted solely through frequency-based measures. Instead, they should also be understood in relation to the thematic and cultural content of the texts.

Table 4. Examples of Off-List Vocabulary Identified in the Corpus

Category	Examples
Proper Nouns	Malala, Yousafzai, Cairo, Agra

Category	Examples
Islamic Terms	Quran, Kaabah, Qibla, Masjid
Religious Locations	Makkah, Madinah, Jerusalem
Educational Institutions	Al-Azhar, Al-Zaytunah
Website References	Khanacademy, Biographyonline

From the perspective of schema theory, readers draw upon prior knowledge structures to facilitate comprehension and meaning construction (Koda, 2005). Familiarity with Islamic concepts, historical figures, and religious practices may therefore support comprehension even when some lexical items fall outside high-frequency word lists. This suggests that vocabulary demands in UM-PTKIN are influenced by both lexical knowledge and topic familiarity. While the inclusion of Islamic-related content is consistent with the institutional context of State Islamic Higher Education, it also indicates that reading performance may partly depend on candidates' background knowledge of Islamic discourse. Therefore, lexical accessibility in UM-PTKIN should be evaluated not only in terms of vocabulary frequency but also in relation to culturally embedded content that shapes text comprehension.

Off-List Vocabulary and Islamic-Related Content

The analysis of vocabulary distribution across the four UM-PTKIN test scripts (1911, 1913, 1915, and 1917) revealed a remarkably consistent lexical pattern. Across all test versions, the largest proportion of vocabulary belonged to the first 1,000-word frequency level, accounting for approximately 65.9% to 69.5% of all tokens. This finding indicates that the reading texts primarily relied on high-frequency vocabulary, which is generally considered essential for basic language comprehension and forms the foundation of English reading proficiency (Nation, 2011; Webb, 2023). The dominance of the first 1,000-word level suggests that the test was designed to assess reading comprehension through largely familiar vocabulary rather than through extensive use of rare lexical items.

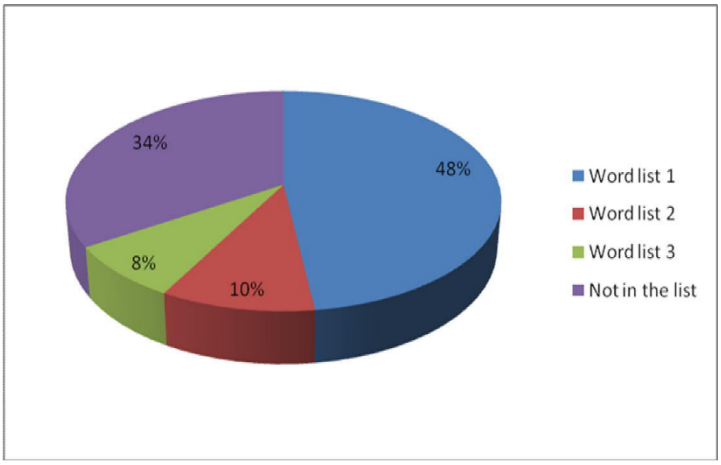

Figure 1. Types of word used in the tests

The second and third 1,000-word frequency levels represented a much smaller proportion of the texts. The second frequency band accounted for between 5.1% and 6.5% of the tokens, while the third frequency band contributed between 3.7% and 4.9%. Although relatively limited in number, these vocabulary items are likely to increase lexical complexity and distinguish test takers with broader vocabulary knowledge from those who possess only basic lexical competence. Previous studies have demonstrated that mastery of the most frequent 2,000–3,000 word families is an important prerequisite for successful comprehension of general English texts **(Dang, 2022; Dang & Webb, 2016; Nation, 2006)**.

Therefore, the presence of vocabulary from the second and third frequency bands suggests that the UM-PTKIN English test targets learners beyond the elementary level and requires a moderate degree of lexical sophistication.

Table 5. Distribution of each test script

Test Script	Tokens	WL1 (%)	WL2 (%)	WL3 (%)	Not in List (%)	Types	Families
1911	844	66.8	5.2	4.5	23.5	344	197
1913	874	68.9	5.1	3.7	22.3	352	202
1915	834	69.5	6.5	4.7	19.3	359	234
1917	757	65.9	5.3	4.9	23.9	311	188

A particularly notable finding is the consistently high proportion of words classified as "not in the list." Across the four test scripts, this category accounted for approximately 19.3%–23.9% of tokens and 25.3%–33.0% of word types. The relatively large proportion of out-of-list vocabulary can be attributed to the inclusion of proper nouns, geographical names, religious terminology, acronyms, and culture-specific expressions. Examples such as *Al-Baqarah*, *Qibla*, *Ulama*, *Makkah*, and *Madinah* reflect the distinctive Islamic context of the examination. While these lexical items may not necessarily increase linguistic difficulty for students familiar with Islamic discourse, they introduce domain-specific vocabulary that may affect text accessibility for learners with different educational or cultural backgrounds.

The comparison among the four test scripts also demonstrates a high degree of lexical consistency. The proportions of vocabulary across frequency levels vary only slightly between test forms, suggesting that the different versions were constructed according to similar lexical standards. Such consistency is important in high-stakes testing because it supports comparability across test forms and reduces the possibility that one version is substantially more difficult than another due to vocabulary differences. From a language assessment perspective, this finding provides preliminary evidence that the lexical demands of the UM-PTKIN English test were relatively stable across test administrations.

Overall, the results indicate that the UM-PTKIN English test combines high-frequency vocabulary with a substantial proportion of context-specific lexical items related to Islamic, historical, and cultural topics. This lexical profile suggests that successful performance depends not only on mastery of general English vocabulary but also on familiarity with specialized vocabulary frequently encountered in Islamic educational contexts. Consequently, vocabulary instruction for prospective UM-PTKIN candidates should address both high-frequency English vocabulary and relevant academic or domain-specific lexical items.

The relatively stable distribution of vocabulary across frequency bands suggests that the four test forms were developed according to comparable lexical specifications. Such consistency is particularly important in high-stakes testing because substantial variation in lexical difficulty across parallel forms may compromise fairness and score comparability (Kremmel, 2023).

Among the four test scripts, only minor differences were observed in the proportions of high-frequency and off-list vocabulary. This pattern indicates that lexical demands remained relatively stable across administrations and supports the interpretation that differences in candidate performance are unlikely to be attributable to substantial variation in vocabulary difficulty between test forms.

From an assessment perspective, lexical consistency constitutes an important aspect of test quality because it helps ensure that examinees are evaluated according to comparable linguistic demands. The findings therefore provide preliminary evidence that the lexical design of the UM-PTKIN English test

reflects an effort to maintain equivalence across parallel examination forms.

Lexical Variation Across Test Scripts and Its Implications for Vocabulary Breadth and Test Accessibility

Lexical Variation Across Test Scripts

The comparative analysis reveals a remarkably consistent pattern across all four test scripts. In every version, the majority of vocabulary belonged to the first 1,000-word frequency band, accounting for approximately 65.9%–69.5% of all tokens. This finding indicates that the lexical demands of the UM-PTKIN English test were primarily based on high-frequency vocabulary, which is generally considered essential for reading comprehension and accessible to most EFL learners

Tabel 6. Comparative Vocabulary Distribution across UM-PTKIN Test Scripts

Test Script	Tokens	Types	Word Families	WL1 (%)	WL2 (%)	WL3 (%)	Off-list (%)
1911	844	344	197	66.8	5.2	4.5	23.5
1913	874	352	202	68.9	5.1	3.7	22.3
1915	834	359	234	69.5	6.5	4.7	19.3
1917	757	311	188	65.9	5.3	4.9	23.9

Among the four test scripts, code 1915 contained the highest proportion of first-level vocabulary (69.5%) and the lowest proportion of off-list words (19.3%), suggesting that it may have been slightly more accessible lexically than the other versions. In contrast, script 1917 contained the smallest number of tokens (757) and the highest proportion of off-list vocabulary (23.9%), indicating a somewhat greater presence of proper nouns, culture-specific terms, and vocabulary outside the established frequency lists.

Despite these minor variations, the overall lexical composition remained highly comparable across test forms. The percentages of vocabulary from the second and third 1,000-word frequency bands showed only limited fluctuation, ranging from 5.1%–6.5% and 3.7%–4.9%, respectively. Similarly, the number of word families varied only moderately, from 188 to 234. Such consistency suggests that the different test versions were constructed according to similar lexical standards, thereby supporting comparability and fairness across administrations.

Another noteworthy finding concerns the relatively high proportion of off-list vocabulary, which ranged from 19.3% to 23.9% of all tokens. This proportion is substantially higher than the percentages observed for the second and third frequency bands. A closer examination revealed that many of these items consisted of proper nouns, geographical names, Islamic terminology, acronyms, and culturally specific expressions. Therefore, although the tests relied predominantly on high-frequency English vocabulary, successful performance may also have depended on familiarity with domain-specific lexical items associated with Islamic and educational contexts.

Overall, the findings demonstrate that the lexical profiles of the four UM-PTKIN test scripts were highly similar. The dominance of high-frequency vocabulary combined with a moderate proportion of off-list words indicates a deliberate balance between accessibility and contextual specificity, which is consistent with the objectives of a high-stakes university entrance examination.

Although minor differences were identified among the four test scripts, the overall lexical profiles remained remarkably similar. Script 1915 exhibited the highest proportion of first-level vocabulary and the lowest percentage of off-list items, suggesting slightly lower lexical demands compared with the other versions. In contrast, script 1917 contained the highest proportion of off-list vocabulary, which may have increased processing demands for some test takers. Nevertheless, these differences were relatively small and did not substantially alter the overall lexical composition of the examination.

The consistency observed across the four test scripts is an important finding from a language assessment perspective. In high-stakes testing, substantial variation in lexical demands across parallel forms may threaten score comparability and compromise perceptions of fairness among examinees. The relatively stable distribution of high-frequency vocabulary, academic vocabulary, and off-list items identified in the present study suggests that the different UM-PTKIN test versions were developed according to comparable lexical specifications. This finding supports the assumption that the various test forms were designed to measure similar language abilities despite differences in reading topics and textual content.

A closer examination of the reading passages further revealed that variation across test forms was influenced primarily by thematic content rather than by overall vocabulary difficulty. Test scripts containing passages related to Islamic history, religious institutions, or prominent Muslim figures tended to include a larger number of proper nouns and culturally specific expressions, which contributed to higher off-list percentages. However, these lexical differences reflected contextual variation rather than substantial differences in linguistic complexity. Consequently, the lexical profiles suggest that the examination maintained a balance between contextual diversity and lexical comparability across forms.

The findings therefore provide preliminary evidence that lexical variation among the four UM-PTKIN test scripts was limited and systematically controlled. Such consistency is desirable in large-scale entrance examinations because it helps ensure that differences in candidate performance are more likely to reflect language proficiency than unintended differences in vocabulary demands between test versions. This observation strengthens the argument that the UM-PTKIN English test demonstrates an acceptable degree of lexical equivalence across parallel forms.

Implications of the Lexical Profile for Vocabulary Breadth and Test Accessibility

The lexical profile identified in the present study provides important insights into the vocabulary knowledge required for successful performance in the UM-PTKIN English examination. Although the test corpus was dominated by high-frequency vocabulary, the presence of a substantial proportion of off-list items indicates that mastery of core vocabulary alone may not be sufficient to ensure optimal comprehension. This finding suggests that prospective candidates require a broader vocabulary repertoire extending beyond the most frequent word families commonly taught in secondary-school English programs.

From the perspective of vocabulary breadth, the results imply that students who possess knowledge limited primarily to high-frequency vocabulary may encounter difficulties when processing culturally specific expressions, proper nouns, and academic or disciplinary terms embedded within the examination texts. Previous research has consistently demonstrated that broader vocabulary knowledge contributes significantly to reading comprehension, inferencing ability, and overall test performance (Dagnaw, 2023; Stæhr, 2008; Uchihara, 2020; Wang, 2025; Webb, et.al, 2023). In particular, vocabulary size has been found to be a strong predictor of learners' reading and listening abilities, suggesting that vocabulary breadth constitutes a fundamental component of language proficiency (Stæhr, 2008).Consequently, vocabulary breadth should be understood not merely as knowledge of frequent words but also as the capacity to recognize and interpret vocabulary encountered across diverse thematic and educational contexts.

The findings also have important implications for test accessibility. Recent meta-analytic evidence further confirms that vocabulary knowledge is one of the strongest predictors of second language reading comprehension across diverse educational contexts (Teng, 2023). Accessibility in language assessment refers to the extent to which test content can be reasonably understood by members of the intended test-taking population without introducing unnecessary linguistic barriers. While the predominance of high-frequency vocabulary suggests that the UM-PTKIN English test remains broadly accessible to senior secondary school graduates, the concentration of Islamic-related and culturally

embedded vocabulary introduces an additional dimension of difficulty that may not be captured by conventional lexical frequency analyses.

This observation is particularly relevant when considered through the lens of schema theory. Readers draw upon prior knowledge and background experiences to support comprehension and meaning construction (Koda, 2005). Candidates who are familiar with Islamic concepts, institutions, and historical references may benefit from existing schemata that facilitate text processing, even when encountering vocabulary classified as off-list. In contrast, candidates with more limited exposure to such content may experience greater cognitive demands despite possessing comparable levels of English proficiency. Therefore, test accessibility in the UM-PTKIN context appears to be influenced by an interaction between vocabulary knowledge and topic familiarity.

From an assessment perspective, these findings do not necessarily indicate a threat to test validity. Rather, they highlight the distinctive nature of language assessment within Islamic higher education contexts, where disciplinary and cultural content forms part of the construct being assessed. Nevertheless, careful consideration should be given to the proportion and distribution of culturally specific vocabulary to ensure that lexical demands remain aligned with the intended proficiency level of prospective students. Excessive reliance on context-specific vocabulary may unintentionally increase construct-irrelevant variance and affect perceptions of fairness among examinees from diverse educational backgrounds.

Overall, the lexical profile suggests that the UM-PTKIN English test is generally accessible to students with a solid foundation in high-frequency English vocabulary. However, successful performance is likely to depend on a combination of vocabulary breadth, contextual inferencing skills, and familiarity with Islamic and educational discourse. This finding is in line with **(Webb & Nation, 2017)** who argue that successful comprehension depends not only on the number of words known but also on learners' ability to infer meaning and integrate lexical knowledge with contextual information. These findings underscore the importance of aligning vocabulary instruction, curriculum development, and test design with the actual lexical demands encountered in high-stakes university entrance examinations.

The present study contributes theoretically to lexical profiling and language assessment research by demonstrating that lexical demands in context-specific examinations cannot be fully explained through frequency-based vocabulary measures alone. In the case of UM-PTKIN, a substantial proportion of off-list vocabulary originated from culturally embedded and religiously specific discourse rather than advanced low-frequency English words. This finding extends previous lexical profiling research by highlighting the interaction between lexical frequency, cultural familiarity, and assessment accessibility in specialized educational contexts (Castellano, 2024; Pellicer-Sánchez, 2024; Sung, 2021).

4. CONCLUSION

This study examined the lexical profile of the English component of the 2019 UM-PTKIN examination through an analysis of tokens, word types, word families, and vocabulary frequency distributions across four parallel test scripts. The findings indicate that the lexical composition of the test is predominantly based on high-frequency English vocabulary, while also incorporating a substantial number of words that fall outside established frequency lists. A closer examination revealed that many of these off-list items consisted of proper nouns, geographical references, Islamic terminology, and culturally embedded expressions associated with the thematic orientation of the examination rather than advanced English vocabulary alone.

The results further demonstrate a high degree of lexical consistency across the four test versions, suggesting that the different forms were constructed according to comparable lexical standards. From a language assessment perspective, this consistency supports the fairness and comparability of the

examination across administrations. At the same time, the findings indicate that successful performance in the UM-PTKIN English test depends not only on mastery of high-frequency English vocabulary but also on the ability to infer meaning from context and to comprehend vocabulary embedded within Islamic, historical, and educational discourse.

Overall, the lexical profile suggests that the UM-PTKIN English test is generally appropriate for senior secondary school graduates who possess an adequate foundation in high-frequency English vocabulary. However, candidates with limited exposure to academic vocabulary, Islamic-related terminology, and culturally specific content may experience additional comprehension challenges. Therefore, lexical accessibility in UM-PTKIN should be understood not solely in terms of vocabulary frequency levels but also in relation to the background knowledge and thematic familiarity required by the texts.

This study contributes to the growing body of lexical profiling and language assessment research by providing empirical evidence from a high-stakes university entrance examination that has received limited scholarly attention. Unlike internationally oriented proficiency tests such as TOEFL or more general university entrance examinations such as UTBK, the UM-PTKIN English test combines high-frequency English vocabulary with a substantial proportion of Islamic and culturally specific lexical items. This distinctive lexical profile represents a unique characteristic of language assessment within the context of Islamic higher education in Indonesia and constitutes the principal contribution of the present study.

Several limitations should be acknowledged. First, the study focused exclusively on lexical characteristics and did not investigate item difficulty, distractor effectiveness, readability, or broader aspects of test validity. Second, the analysis was based on the 2019 UM-PTKIN examination. Although the overall purpose, format, and reading-oriented nature of the test have remained broadly comparable across administrations, lexical demands may evolve over time. Future studies are therefore encouraged to compare the lexical profiles of the 2019 examination with more recent computer-based UM-PTKIN tests in order to examine possible changes in vocabulary demands and lexical complexity. Further research should also combine lexical profiling with Vocabulary Size Test data from prospective candidates so that the relationship between vocabulary demands and learners' actual vocabulary knowledge can be examined more directly and empirically.

REFERENCES

Anandi, D. G. A. (2024). Lexical Richness in Essays: A Corpus-Based Study Among Indonesian Junior High School Students. *Journal of English Education*, *9*(2), 120–134.

Anthony, L. (2017). Corpus Linguistics and Vocabulary : A Commentary on Four Studies. *Vocabulary Learning and Instruction*, *6*(2), 79–87. https://doi.org/10.7820/vli.v06.2.Anthony

Anthony, L. (2022). *AntConc [Computer software]* (4.20). Waseda University.

Bao, Zhen., & Peng, Cheng. (2024). The effects of EFL wordlist and proficiency on vocabulary knowledge. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1289106

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman Grammar of Spoken and Written English. In *Cambridge University Press* (Number viii). Longman. https://doi.org/10.1017/S0022226702211627

Castellano-Risco, I. (2024). Lexical coverage and reading comprehension in English language assessment. *Language Assessment Quarterly*, *21*(1), 45–63.

Cobb, T. (2020). *VocabProfile: Lexical profiling tool.* . https://www.lextutor.ca/vp/

Coxhead, A. (2000). A new Academic Word List. *TESOL Quarterly*, *34*(2), 213–238.

Dagnaw, A. T. (2023). Revisiting the role of breadth and depth of vocabulary knowledge in reading comprehension. *Cogent Education*, *10*(1).

Dang, T. N. Y. , C. A. , & W. S. (2022). Lexical coverage and learners' text comprehension: Revisiting the 95% and 98% thresholds. *Language Learning*, *72*(1), 38–69.

Dang, T. N. Y., & Webb, S. (2016). Evaluating lists of high-frequency words. *ITL - International Journal of Applied Linguistics*, *167*(2), 132–158. https://doi.org/10.1075/itl.167.2.02dan

EAP Foundation. (2022). *General Service List (GSL)*.

Green, A. (2021). *Exploring language assessment and testing: Language in action*. Routledge.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Kremmel, B. , & A. J. C. (2023). *Re-examining language assessment: Validity, fairness and test use*. Cambridge University Press.

Laufer, B. , & R.-K. G. C. (2010). Lexical threshold revisited. *Reading in a Foreign Language*, *22*(1), 15–30.

McLean, S. , & K. B. (2021). Lexical profiles of standardized English tests: Implications for vocabulary learning and assessment. *Language Assessment Quarterly*, *18*(4), 406–425.

McLean, S., & Kramer, B. (2015). *The Creation of a New Vocabulary Levels Test*. Retrieved from https://www.lvlt.info

Nation, I. S. P. (2006). Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? Canadian Modern Language Review, 63(1), 59–82. *Canadian Modern Language Review*, *63*(1), 59–82.

Nation, I. S. P. (2017). *The BNC/COCA Level 6 Word Family Lists*.

Nation, I. S. P. (2022a). *Learning Vocabulary in Another Language*. Cambridge University Press.

Nation, I. S. P. (2022b). What do 20 years of vocabulary research tell us? . *Language Teaching: Cambridge University Press*, *55*(2), 173–190. DOI: 10.1017/S0261444821000098

Nation, I. S. P. , & A. Heatley. (2002). *RANGE : A Program for the Analysis of Vocabulary in Texts [Computer software]*. Victoria University of Wellington.

Nation, I. S. P. , & W. R. (1997). *Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), Vocabulary: Description, Acquisition and Pedagogy* (Schmitt & M. McCarthy, Ed.; pp. 6–19). Cambridge University Press.

Nation, P. , & W. S. (2011). *Researching and Analyzing Vocabulary*. Heinle.

Nguyen, T. M. H. , & B. F. (2022). Vocabulary demands of EFL reading tests and textbooks: A corpus-based comparison. *Journal of English for Academic Purposes*, *56*.

Pellicer-Sánchez, A. (2024). How does lexical coverage affect the processing of L2 texts? . *Applied Linguistics*, *45*(6), 953–975–953–975.

Reppen, R. (2008). O'Keeffe, A., McCarthy, M. & Carter, R. (2007). From Corpus to classroom: Language use and language teaching . *International Journal of Corpus Linguistics*. https://doi.org/10.1075/ijcl.13.4.09rep

Reynolds, B. L. (2021). Vocabulary knowledge and reading comprehension: A review of recent evidence. *Reading in a Foreign Language*, *33*(2), 282–304.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research manual*. Palgrave Macmillan.

Schmitt, N. , & S. D. (2020). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N. , J. X. , & G. W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26–43.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139–152.

Sung, K. (2021). Topic familiarity and lexical difficulty in EFL reading assessment. *Reading in a Foreign Language*, *33*(2), 180–201.

Synchak, O. , S. V. , B. M. , & S. M. (2025). Corpus-based vocabulary profiling for Ukrainian: From lexical analysis to the PULS digital learning platform. *Proceedings of ELex 2025*, 472–489.

Tahir, M. R. P. , S. S. , & K. L. (2025). Students' high-frequency vocabulary profile: A corpus study at Manado State Institute of Islamic Studies. *Journal of Islamic Integration*, *17*(1), 1–15.

Teng, F. (2023). Vocabulary knowledge and second language reading comprehension: A meta-analysis. *System*, 117.

Uchihara, T. , & C. J. (2020). Investigating the role of vocabulary size in second language reading ability. *Language Teaching Research*, *24*(4), 540–556.

Wang, T. (2025). Reciprocal effects of vocabulary breadth, vocabulary depth, and reading comprehension among EFL adolescents. *Humanities and Social Sciences Communications*, *12*(46).

Webb & Nation. (2017). *How Vocabulary Is Learned*. Oxford University Press.

Webb, S. , N. P. , & C. A. (2023). Vocabulary and language assessment: Recent developments and future directions. *Language Teaching*, *56*(4), 521–540.

Webb, S. N. I. S. P. (2008). Evaluating the vocabulary load of written text. *TESOLANZ Journal*, *16*, 1–19.

Weisser, M. (2012). *Concordancers: An overview. In The Encyclopedia of Applied Linguistics.* Wiley.

West, M. (1953). *A General Service List of English Words*. Longman, Green and Co.