Indonesian Journal of Education (INJOE) https://felifa.net/index.php/INJOE
Vol. 6 No. 1, (2026), pages: 142~151 e-ISSN: 2810-059X

LEXICAL COMPLEXITY IN ENGLISH LISTENING MATERIALS OF THE CAKE APPLICATION

Zea Jelika Dewanda
Universitas Islam Negeri Sultan Maulana Hassanudin Banten
zeadewanda2004@gmail.com

Tatu Siti Rohbiah
Universitas Islam Negeri Sultan Maulana Hassanudin Banten
tatusiti.rohbiah@uinbanten.ac.id

Yazid Rukmayadi
Universitas Islam Negeri Sultan Maulana Hassanudin Banten
Yazid.rukmayadi@uinbanten.ac.id

Abstract

Lexical complexity plays an important role in determining the linguistic characteristics and difficulty of English listening materials. This study examines lexical complexity in English listening materials presented in the Cake Application, with particular attention to lexical diversity, lexical density, and lexical sophistication. A descriptive qualitative approach was employed, using listening transcripts, students' vocabulary findings, observations, and interviews as sources of data. The analysis focused on Type-Token Ratio (TTR), the proportion of content words, and the occurrence of sophisticated or less frequent vocabulary. The results indicated that the listening materials demonstrated relatively diverse vocabulary, with an average TTR of 64%. The lexical density analysis produced an average score of 51.6%, with scores ranging from 39% to 61%, suggesting a moderate to relatively high concentration of meaningful lexical items. In terms of lexical sophistication, several intermediate and less frequent words, including *exactly, journey, canoe, horrible, embarrassed,* and *ambiance*, were identified. These findings indicate that the Cake Application provides authentic listening input with varied lexical features and a manageable level of lexical challenge. Therefore, the density; lexical sophistication; Cake Application; listening materials potential as a supplementary digital resource for English listening instruction and vocabulary development.

Keywords: Lexical complexity; lexical diversity; lexical density; lexical sophistication; Cake Application; listening materials.

INTRODUCTION

In the digital age, English has become an essential skill for global communication. Social media has become an integral part of teenagers' daily lives, especially among students in Indonesia. Language plays a vital role in human life. Without language, students would struggle to navigate daily life.(Meylani Aulia, Journal of Education,Linguistics,Literature, 2024) Listening is the most important skill for understanding and effectively interpreting what a conversation partner says. Spontaneous conversations between native speakers at a normal pace.

Without listening skills, students will not learn effectively. According to Buck, “listening is a complex process in which the listener takes in incoming data and acoustic signals and interprets them based on a variety of linguistic and nonverbal knowledge”.Gary Buck, Menilai Mendengarkan, Seri Penilaian Bahasa Cambridge (Cambridge University Press, 2001).

Harmer has argued that the more you listen to and understand English, the better you will hear and grasp the correct tone and intonation, stress, and sounds, enabling you to understand individual words more quickly and combine them into coherent speech. Today, technology has become an integral part of our lives. Digital technology has been used in various fields, including communication, business, healthcare, and education. In the world of education, digital technology has played a major role, making it an important learning medium for students. Currently, Mobile-Assisted Language Learning (MALL) has emerged as a vital approach, referring to the use of mobile devices to facilitate flexible and easily accessible language learning. This has allowed students to access learning materials anytime and anywhere. In the context of English listening instruction using the Cake Application, MALL has provided various interactive features that support the development of language skills, particularly listening and speaking. This exposure has proven crucial for improving comprehension and helping learners become accustomed to natural pronunciation, intonation, and expressions.(Eko, J. S. S., & Novianty, 2023)

According to Jack C. Richards, the selection and evaluation of language teaching materials must take into account linguistic aspects and the level of difficulty of the texts used in the teaching process. In general, spoken language tends to differ from written language because it contains elements such as discourse markers, filler words, and frequent collocations. Therefore, it is important to analyze the lexical profiles of listening materials, including those in mobile learning apps, to ensure they are suitable for the target students.(Richards, 2001)

The use of Authentic Materials in English language teaching has been a Subject of Research for several decades. Authentic materials (AM) refer to unmodified texts or other linguistic resources created for native speakers, such as newspapers, magazines, podcasts, or videos. Leading linguists such as Delly Hymes and Noam Chomsky in the 1960s and 1970s argued that gaining knowledge of language structure is highly beneficial for teachers and students who often face difficulties at various levels. And the most effective way to integrate authentic materials into practice has not yet been fully established.Ashley Renlund, "Lexical Complexity and English Teaching Materials: The Use of Authentic Texts vs. English Teachers' Textbooks," 2023. Gilmore points to the challenge of determining the difficulty level of idiomatic language and low-frequency vocabulary, which can hinder learning. However, these concerns have not deterred teachers from incorporating authentic materials into their teaching methodologies. (Gilmore,

2007)Therefore, it is difficult to determine how academic research can help teachers overcome the challenges associated with integrating authentic materials into their teaching.

One of the main factors influencing listening difficulty is lexical complexity, which includes lexical density, lexical diversity, and lexical sophistication. If the vocabulary is too difficult, students may experience cognitive overload, which can interfere with their comprehension. Batia Laufer and Paul Nation emphasize that lexical proficiency is a key indicator of overall language proficiency. If the listening material is too easy, students will not make significant progress. On the other hand, if the material exceeds learners' lexical comprehension, it can result in excessive cognitive load, hindering understanding. Therefore, finding the optimal balance of lexical difficulty is crucial to ensure that the input remains comprehensible while still providing a beneficial challenge for language acquisition.(Jack C, Richard, 2022)

Based on the phenomenon observed among students who still use the listening textbox, this study aims to explore in depth the dimensions of lexical complexity in the listening content of the Cake Application through linguistic corpus analysis. This analysis is conducted to identify the lexical density and complexity of the listening materials using the Cake Application. Therefore, this study is important for analyzing the linguistic profile of this material to ensure its effectiveness in supporting second language acquisition. By identifying these lexical patterns, this study can provide empirical evidence on how digital platforms can better bridge the gap between classroom learning and real-world language applications.

METHODS

Research Design

This study employed a descriptive qualitative research design to describe the lexical complexity of English listening materials in the Cake Application. The analysis focused on three dimensions of lexical complexity: lexical density, lexical diversity, and lexical sophistication. The descriptive qualitative approach was selected because the study aimed to provide an in-depth description of vocabulary characteristics in the selected listening materials rather than to generalize the findings to a wider population. The linguistic analysis of the transcripts served as the primary basis of the study, while classroom observations and interviews were used as supporting data to provide contextual information and strengthen the interpretation of the findings.

Research Site and Data Sources

The study was conducted at MAN 2 Pandeglang during the second semester of the 2025/2026 academic year. Data collection was carried out from March to April 2026. The primary data consisted of selected English listening materials from the Cake Application and their corresponding transcripts. Fifteen data samples were analyzed to identify patterns of lexical density, lexical diversity, and lexical

sophistication. Supporting data were obtained through classroom observations and interviews with students and an English teacher to gain additional information about students' experiences, vocabulary difficulties, and perceptions of the listening materials.

Research Instruments and Data Collection

The researcher served as the main instrument in collecting and analyzing the data. Three supporting instruments were employed: a document analysis sheet, an observation sheet, and interview guidelines.The document analysis sheet was used to examine the transcripts of the selected listening materials. It recorded the total number of words, content words, function words, lexical density, lexical diversity, and lexical sophistication. Lexical density was calculated by determining the proportion of content words to the total number of words in each transcript. Lexical diversity was measured using the Type-Token Ratio (TTR), which compares the number of different words (types) with the total number of words (tokens). Lexical sophistication was examined by identifying less frequent or relatively difficult vocabulary items appearing in the listening materials.

Classroom observations were conducted to obtain contextual information about students' engagement with the Cake Application, including their attention to the videos, use of subtitles, interaction with vocabulary features, and participation during listening activities. In addition, interviews were conducted with students and an English teacher to obtain supporting information regarding the perceived difficulty, vocabulary load, and suitability of the listening materials for the students' English proficiency level. These observations and interviews were not used as the primary basis for calculating lexical complexity but were used to support the interpretation of the document analysis.

Data Analysis

The data were analyzed systematically through four main stages: data preparation, data reduction, data display, and interpretation. First, the selected listening materials were collected and transcribed, and the transcripts were checked for accuracy. Second, relevant linguistic data were identified and categorized according to the three dimensions of lexical complexity. For lexical density, the proportion of content words to total words was calculated using the following formula:

Lexical Density = (Number of Content Words / Total Number of Words) × 100

Lexical diversity was measured using the Type-Token Ratio:

TTR = (Number of Types / Number of Tokens) × 100

For lexical sophistication, vocabulary items that were relatively infrequent or unfamiliar to the learners were identified and described based on their occurrence and contextual use in the listening materials. The results of these analyses were then organized into tables to facilitate comparison across the fifteen data samples.

Finally, the findings from the lexical analysis were interpreted descriptively

and supported by observations and interview responses. Data triangulation was employed by comparing the primary document analysis with the supporting observation and interview data. This process was intended to strengthen the credibility of the interpretation and provide a more comprehensive understanding of lexical complexity in the Cake Application listening materials.

RESULTS AND DISCUSSIONS

Education Internal Quality Assurance System

RESULTS

Lexical Diversity

This section presents the data used to measure the lexical complexity of the English listening materials in the Cake Application.To evaluate the lexical richness and the variety of vocabulary produced by the students, the researcher employed the Type-Token Ratio (TTR) method. This method measures the ratio of unique words (Types) to the total number of words (Tokens) generated by the 15 respondents after engaging with the Cake Application materials. The data gathered from the students is presented in Table 4.1 below.

Table 4.1 Score Lexical Diversity

No	Responden	Total Words (Tokens)	Unique Words (Types)	TTR Score
1.	Student1	52	34	65%
2.	Student 2	47	30	64%
3.	Student 3	51	38	75%
4.	Student 4	68	42	62%
5.	Student 5	62	48	77%
6.	Student 6	75	40	53%
7.	Student 7	64	45	71%
8.	Student 8	55	39	71%
9.	Student 9	42	22	56%
10.	Student 10	55	31	56%
11.	Student 11	70	49	70%
12.	Student 12	36	18	50%
13.	Student 13	65	43	66%
14.	Student 14	90	58	64%
15.	Student 15	52	29	56%
Average		59	38	64%

As shown in Table 4.1, the average TTR score was 64%, with an average of 59 tokens and 38 types. Individual TTR scores ranged from 50% to 77%, indicating variation in the lexical diversity of the students' responses. Student 5 obtained the highest TTR score (77%), whereas Student 12 recorded the lowest score (50%). These differences suggest that the students demonstrated varying degrees of vocabulary variety when engaging with the listening materials.

Overall, The findings indicate variation in lexical diversity among the respondents Materials involving specific themes, such as *Wildlife Conservation* and *Smartphone Addiction*, provided opportunities for students to encounter varied lexical items, while more conversational materials tended to contain more frequently repeated vocabulary. Thus, the variation in TTR scores reflects differences in the lexical diversity of the materials and students' vocabulary use.

Lexical Density

Lexical density in this study is determined by calculating the proportion of content words which carry the main meaning of the text, such as nouns, main verbs, adjectives, and adverbs against the total number of words (tokens). The analysis of the production of the 15 respondents is presented in Table 4.2 below.

Table 4.2 Lexical Density of the Selected Listening Materials

No	Name	Total Word	Content Word	Lexical Density
1.	Student 1	52	27	52%
2.	Student 2	47	23	49%
3.	Student 3	51	29	57%
4.	Student 4	68	33	49%
5.	Student 5	62	37	60%
6.	Student 6	75	35	47%
7.	Student 7	64	33	52%
8.	Student 8	55	32	58%
9.	Student 9	42	18	43%
10.	Student 10	55	24	44%
11.	Student 11	70	38	54%
12.	Student 12	36	14	39%
13.	Student 13	65	36	55%
14.	Student 14	90	55	61%
15	Student 15	52	22	42%

Lexical density was analyzed based on the proportion of content words to the total number of words in 15 selected listening materials from the Cake Application. As shown in Table 4.2, the lexical density ranged from 39% to 61%, indicating variation from relatively conversational to more information-rich materials.

The highest lexical density was found in Student 14's data (61%), consisting of 55 content words out of 90 total words, while the lowest was found in Student 12's data (39%), with 14 content words out of 36 words. Overall, the findings indicate that the Cake Application provides varied lexical input that can expose tenth-grade EFL learners to both everyday conversational and more information-rich English.

Lexical Sophistication

Lexical sophistication was examined by identifying less frequent, context-specific, and relatively advanced vocabulary encountered by the 15 respondents in the Cake Application listening materials. The analysis identified a range of sophisticated lexical items across different thematic materials, including *scared, relieved, riddance, canoe, horrible, journey, ambiance, frustrated, display, handle,* and *embarrassed*. The distribution of these lexical items across the respondents is presented in Table 4.3.

Table 4.3. Sophisticated Vocabulary Identified by Respondents

No	Responden	Word
1.	R1, R3, R7	Scared
2.	R2, R5, R9	Relieved
3.	R4	Riddance
4.	R1,R2,R5,R8,R10,R12	Canoe
5.	R1,R3,R4,R6,R8,R11	Horrible
6.	R2,R5,R7,R9,R10,R13	Journey
7.	R6, R14	Ambiance
8.	R3, R8, R15	Frustrated
9.	R1, R4, R6, R9, R12	Display
10.	R2, R5, R8, R11	Handle
11.	R3, R7, R10, R14	Embarrassed
12.	R1,R2,R4,R5,R6,R9,R13	Exactly
13.	R2, R4, R7, R11, R15	Fault

The findings show that the respondents encountered varied lexical items beyond basic everyday vocabulary. Several words appeared repeatedly across the students' data, indicating that particular vocabulary items were more frequently encountered in the listening materials. As summarized in Table 4.4, Exactly had the highest frequency, occurring seven times, followed by *Journey, Canoe,* and *Horrible*, each occurring six times. *Fault* and *Display* appeared five times, while *Embarrassed* and *Handle* occurred four times. Other identified items included *Scared* and *Relieved*, each with a frequency of three.

Table 4.4. Frequency of Frequently Identified Sophisticated Words

No	Word	Frequency
1.	Exactly	7
2.	Journey	6
3.	Canoe	6
4.	Horrible	6
5.	Fault	5
6.	Display	5

7.	Embarrassed	4
8.	Handle	4
9.	Scared	3
10	Relieved	3

Interview Findings

The teacher reported that some vocabulary items in the Cake Application were challenging for students, but these unfamiliar words helped increase their awareness of vocabulary gaps. The teacher also considered the application suitable for students' listening development because its visual and interactive features provided additional support for learners with limited listening proficiency.

The students generally responded positively to the Cake Application. They found it interesting, interactive, and accessible, while videos, subtitles, pronunciation practice, and repetition features helped them understand the listening materials. Students also reported encountering new vocabulary, expressions, and phrases through authentic, everyday contexts. Overall, the interview findings suggest that the Cake Application can support students' vocabulary development and listening engagement.

DISCUSSIONS

The findings indicate that the English listening materials in the Cake Application demonstrate moderate to high lexical complexity across three dimensions: lexical density, lexical diversity, and lexical sophistication. The average lexical density was approximately 51.6%, while the average Type-Token Ratio (TTR) was 64%. The materials also contained several relatively unfamiliar or low-frequency vocabulary items, such as *exactly, journey, canoe, horrible,* and *embarrassed*. These findings suggest that the Cake Application provides varied and information-rich lexical input while maintaining a level of complexity that can still be accessed by tenth-grade EFL learners.

The lexical density findings show that the materials generally contain a balanced proportion of content and function words. This is consistent with Halliday's view that lexical density reflects the amount of lexical information contained in a text. The moderate level found in this study is particularly relevant to listening instruction because the materials combine authentic spoken language with sufficient contextual support. Thus, students are exposed to natural English without being consistently overwhelmed by highly dense linguistic input. (Bell, B., & Farmer, n.d.) The TTR findings further indicate considerable lexical variation in the materials. An average TTR of 64% suggests that students encountered a relatively broad range of vocabulary rather than highly repetitive lexical patterns. This supports the importance of lexical diversity in expanding learners' vocabulary repertoire. However, differences among the fifteen data samples also indicate that

students' engagement with vocabulary varied across topics and materials.

Regarding lexical sophistication, the presence of less frequent and unfamiliar words demonstrates that the Cake Application exposes learners to vocabulary beyond basic everyday expressions. Words such as *journey, canoe, frustrated,* and *embarrassed* provide opportunities for learners to develop vocabulary knowledge through authentic contexts. The combination of audio, subtitles, visual information, and contextualized dialogue may help learners infer meaning and process unfamiliar lexical items.

The observation and interview findings support these linguistic results. Students generally showed positive engagement with the application and reported encountering new vocabulary through the listening activities. The teacher also indicated that some vocabulary was challenging, particularly for students with lower listening proficiency, but considered the application useful because its audiovisual features provide contextual support. These findings suggest that the Cake Application can function as a supplementary Mobile-Assisted Language Learning (MALL) resource when appropriate scaffolding is provided.

Overall, the findings suggest that the Cake Application offers authentic listening input with manageable but meaningful lexical challenges. Its moderate lexical density, relatively high lexical diversity, and exposure to sophisticated vocabulary can support vocabulary development and listening comprehension. However, teachers should provide vocabulary preparation, subtitles when necessary, and repeated listening activities to ensure that authentic materials remain accessible to tenth-grade EFL learners.

CONCLUSION

This study analyzed the lexical complexity of English listening materials in the Cake Application by examining three dimensions: lexical density, lexical diversity, and lexical sophistication. Appropriate for tenth-grade EFL learners with an average lexical density of 51.6%, an average Type–Token Ratio (TTR) of 64%, and the presence of several intermediate and low-frequency words. These findings indicate that the materials provide authentic and varied language input while remaining appropriate for tenth-grade EFL learners.

The results suggest that the Cake Application can serve as a valuable supplementary learning resource for English listening instruction. Its authentic listening materials expose students to diverse vocabulary and meaningful language use in real-life contexts, supporting both listening comprehension and vocabulary development. Although students encountered some unfamiliar lexical items, these words offered opportunities to expand their vocabulary knowledge through contextualized learning. Therefore, integrating the Cake Application into English language teaching may help enrich classroom instruction by providing authentic listening experiences and encouraging continuous vocabulary growth among EFL

learners.

REFERENCES

Bell, B., & Farmer, D. (n.d.). *Oral/Interpersonal Communication*. https://wtcs.pressbooks.pub/ communications/%0APreface

Buck, G. (2001). *Assessing listening, Cambridge Language Assesment Series*. Cambridge University Press.

Eko, J. S. S., & Novianty, L. (2023). the effect of cake's online application on listening skill. *Journal of Engineering Research*, *10*(1), 35–45.

Gilmore, A. (2007). *Authentic materials and authenticity in foreign language learning SURVEYS AND STUDIES*. Cambridge University Press Cambridge. https://doi.org/https://doi.org/10.1017/S0261444807004144

Jack C, Richard, J. P. (2022). *Teaching and Learning in English Medium Instruction: An Introduction*. Taylor & Francis. https://books.google.co.id/books?id=jKpYEAAAQBAJ

Meylani Aulia, Journal of Education , Linguistics , Literature, and A. (2024). The role of Indonesian as a communication tool in learning. *Journal of Education , Linguistics , Literature , and Art*, *2*(2). https://doi.org/10.62568/ella.v2i2.128

Renlund, A. (2023). *Lexical Complexity and English Teaching Materials : English Teachers' use of Authentic vs. Textbook Texts*.

Richards, J. C. (2001). *Curriculum Development in Language Teaching*. Cambridge University Press. https://books.google.co.id/books?id=tVhx_B3a3qQC.