Korean Journal of English Language and Linguistics, **Vol 26, March 2026, pp. 418-440**
DOI: 10.15738/kjell.26..202603.418

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Effects of Lexical Coverage on L2 Incidental Vocabulary Learning During Listening

Hyeeun Park · Josephine Mijin Lee (Ewha Womans University)

Received: February 12, 2026
Revised: March 2, 2026
Accepted: March 2, 2026

Park, Hyeeun (First author)
Graduate Student, Department of English Education
Ewha Womans University
Email: ninjahyeni@ewha.ac.kr

Lee, Josephine Mijin (Corresponding author)
Associate Professor, Department of English Education
Ewha Womans University
A407 Education Building, 52 Ewhayeodae-gil, Seodaemun-gu
Seoul, Korea
Tel: +82-2-3277-2666
Email: leejosephine@ewha.ac.kr

ABSTRACT

Park, Hyeeun and Josephine Mijin Lee. 2026. Effects of lexical coverage on L2 incidental vocabulary learning during listening. *Korean Journal of English Language and Linguistics 26*, 418-440.

This study examined the effect of lexical coverage (90%, 95%, 98%) on incidental vocabulary learning (IVL) in listening. Seventy-seven Korean EFL middle school students listened to passages in which coverage was manipulated by using pseudowords. Vocabulary knowledge (form, grammar, meaning) was measured immediately and after three weeks. Results showed a non-linear effect in that the 98% condition significantly outperformed the 90% condition, while the 90% and 95% groups showed no significant difference. Across dimensions of vocabulary knowledge, form and grammar were acquired more readily than meaning. A significant Coverage × Time × Dimension interaction revealed that for form and grammar knowledge, higher coverage led to better initial learning, but all groups experienced attrition over time. For meaning, only the 98% condition facilitated initial vocabulary learning, but this knowledge showed steep attrition. These findings suggest that near-complete coverage is needed for IVL, especially for acquiring meaning knowledge, in listening.

KEYWORDS

lexical coverage, second language, incidental vocabulary learning, listening

1. Introduction

Incidental vocabulary learning (IVL) has received increasing attention in second language (L2) acquisition research because of its crucial role in supporting L2 vocabulary growth (Kurokawa et al. 2025, Teng and Uchihara 2024, Webb and Nation 2017). IVL is central to L2 development given that vocabulary knowledge is both multidimensional (Nation 2001) and incremental (van Zeeland and Schmitt 2013), so learners need more than intentional learning to acquire different aspects of word knowledge. However, in EFL contexts, learners' exposure to meaningful input is often limited, creating an impoverished environment for IVL to occur (Teng 2023). Identifying the conditions that maximize and facilitate IVL in EFL settings thus remains an empirical challenge.

Previous research has reported word-related factors (Crossley et al. 2016, Elgort and Warren 2014, Peters and Webb 2018, Puimège and Peters 2019, Vidal 2003), contextual factors (Pellicer-Sánchez and Schmitt 2010, Penno et al. 2002, Peters et al. 2016, van Zeeland and Schmitt 2013, Vidal 2011), and learner-related factors (Teng 2025) to facilitate IVL. Among these factors, lexical coverage—defined as the proportion of known words in a text (Webb 2021)—provides a text-sensitive indicator of learning conditions that may not be fully reflected in learner-level measures alone. Despite its theoretical relevance, relatively little research has examined the relationship between lexical coverage and IVL. This gap is particularly noteworthy in the domain of listening. Vocabulary gains from listening are generally smaller than those from reading (Brown et al. 2008), suggesting that listening may impose additional challenges for IVL. Although listening is a primary medium for L2 exposure, it has received only limited attention in IVL research (Webb 2016).

Given that lexical coverage reflects the degree to which the lexical demands of spoken input are accessible to learners, it offers a useful framework for examining how input conditions shape opportunities for incidental learning during listening. The present study thus aims to address this gap by investigating the impact of lexical coverage on IVL in the listening condition. The present study investigates IVL under listening comprehension-oriented instructions in which unfamiliar forms were made perceptually salient but not explicitly taught. In doing so, it aims to provide pedagogical insights that inform vocabulary and listening instruction, particularly with regard to setting realistic learning goals and selecting listening materials that promote IVL. The research questions are as follows:

1. Does lexical coverage (90%, 95%, 98%) affect learners' performance on three measures of incidental vocabulary knowledge (form, grammar, and meaning) on the immediate vocabulary test?
2. Does lexical coverage (90%, 95%, 98%) affect the retention of incidentally learned vocabulary knowledge three weeks after exposure?
3. Does lexical coverage interact with time (immediate vs. delayed) in incidental vocabulary learning?

2. Literature Review

2.1 L2 Incidental Vocabulary Learning

IVL has been defined in two complementary ways. Methodologically, it refers to vocabulary learning that occurs when learners are not forewarned of a subsequent vocabulary retention test. Pedagogically, it is defined as vocabulary learning that emerges as a by-product of meaning-focused activity rather than from intentional study (Hulstijn 2001). L2 research has demonstrated that IVL can occur across a range of input modes, including reading

(e.g., Waring and Takaki 2003), listening (e.g., Jin and Webb 2020), and viewing audiovisual input (e.g., Peters and Webb 2018).

The outcomes of IVL cannot be attributed to a single factor; rather, they are shaped by the interaction of multiple variables. Accordingly, most IVL research has focused on identifying factors that facilitate or constrain learning, such as word exposure frequency, L2 proficiency (particularly vocabulary knowledge), input mode, text type, and lexical coverage. One extensively investigated variable is exposure frequency, yet no consensus exists regarding the number of encounters required for IVL to occur. Estimates vary widely across studies, largely due to differences in how vocabulary knowledge is conceptualized and measured. Despite this variability, findings consistently indicate that increasing the number of encounters promotes vocabulary acquisition (Vidal 2003). Moreover, previous research suggests that more encounters are generally required for IVL through listening than through reading (Brown et al. 2008). For instance, Penno et al. (2002) found that learners began to demonstrate some knowledge of target words after a single exposure, whereas Vidal (2011) reported that the greatest vocabulary gains in listening occurred at five to six encounters. Similarly, van Zeeland and Schmitt (2013) showed that limited aspects of word knowledge, such as spoken form recognition and part-of-speech information, developed after relatively few encounters (e.g., around seven), while more complete and durable word knowledge required over fifteen encounters.

L2 proficiency, particularly vocabulary knowledge, has also been shown to exert a robust influence on IVL (Uchihara et al. 2019, Webb et al. 2023). Learners with larger vocabulary sizes tend to achieve better comprehension of input (Peters and Webb 2018), which in turn is assumed to facilitate inference of unknown words from context and subsequent learning (Ellis 1999). However, lexical coverage—the proportion of words in a text that are already known to the learner (Pellicer-Sánchez et al. 2024)—may provide a more precise indicator of the knowledge required to understand and potentially acquire vocabulary from a specific text than general vocabulary size alone. Unlike measures of general vocabulary knowledge that estimate overall vocabulary size or knowledge across frequency bands, lexical coverage directly reflects learners' familiarity with the actual words in the input. This distinction is important because a small number of topic-related words can account for a substantial proportion of lexical coverage in a given text (Webb 2021). Furthermore, research has shown that while general vocabulary knowledge predicts reading comprehension, it does not reliably predict listening comprehension; in contrast, text-specific vocabulary knowledge significantly predicts both reading and listening comprehension (Mehrpour and Rahimi 2010). Because lexical coverage captures knowledge of the specific lexical items present in the input, it offers a potentially more conclusive account of the vocabulary knowledge relevant for IVL.

Despite extensive research on IVL, little is known about how learners' lexical knowledge interacts with input characteristics to support vocabulary acquisition. The continued reliance on global proficiency and vocabulary size measures obscures the fact that comprehension and learning are driven by access to text-specific lexical items rather than abstract knowledge estimates. Lexical coverage, therefore, can operate as a theoretically motivated and empirically grounded construct for examining IVL, particularly across different input modes. Building on this view, the present study operationalizes lexical coverage to examine its role in shaping IVL outcomes.

2.2 Lexical Coverage and Incidental Vocabulary Learning in L2 Listening

Building on findings that L2 proficiency and vocabulary knowledge shape IVL, recent research has increasingly pointed to lexical coverage as a more precise construct for capturing the lexical conditions under which IVL occurs. Lexical coverage refers to the proportion of running words in a text that are known to a learner (Nation 2001, Nation and Waring 1997). It has been widely recognized as a key indicator of input adequacy, as it captures the extent to which learners can access meaning in spoken or written input using their existing lexical knowledge.

Lexical coverage is critical for IVL insofar as it determines whether learners can allocate sufficient attentional resources to unfamiliar words for learning to occur. According to Sweller's (1988) Cognitive Load Theory, working memory capacity is limited, which constrains the cognitive resources available for vocabulary learning. Higher lexical coverage, however, increases the likelihood of comprehension success (Webb 2021), and when learners are provided with the lexical coverage necessary for basic comprehension, they are more likely to devote attentional resources to processing unknown lexical items (Godfroid et al. 2018, Webb and Rodgers 2009). In contrast, when lexical coverage is low, learners need to distribute their attention across many unfamiliar words, reducing the depth of processing afforded to any single item. Furthermore, lexical coverage also plays a significant role in L2 lexical inferencing (Kim 2023), a mechanism that is central to L2 vocabulary acquisition (Nation 2001). Thus, these considerations highlight the theoretical significance of investigating the relationship between lexical coverage and IVL.

Despite its theoretical relevance, however, direct empirical evidence linking lexical coverage to IVL remains limited, as most research has examined lexical coverage primarily in relation to comprehension rather than vocabulary learning. For instance, Bonk (2000) found that learners could achieve adequate comprehension at lexical coverage levels below 95%, while Stæhr (2009) found that high-level listening comprehension was possible at 98% lexical coverage. Supporting these findings, van Zeeland and Schmitt (2013) reported that adequate listening comprehension was achievable at 95% and 98% lexical coverage. Together, these studies indicate that 95% coverage is sufficient for good comprehension and 98% is necessary for high-level comprehension. Although these studies have contributed to establishing widely cited coverage thresholds, they were not designed to examine vocabulary learning outcomes. A handful of studies have reported incidental vocabulary gains alongside coverage estimates (e.g., Godfroid et al. 2018, Webb and Rodgers 2009), but they have primarily examined vocabulary learning as a by-product of exposure under relatively naturalistic conditions, without systematically manipulating or comparing different levels of lexical coverage as an independent variable. As a result, the lexical coverage levels that optimally support IVL remain insufficiently specified.

Research that has explicitly examined lexical coverage in relation to IVL is further limited by its focus on reading, leaving listening comparatively underexplored. This gap is especially salient given that learners may encounter spoken input more frequently than written input outside the classroom (Teng 2024). Moreover, vocabulary gains from listening tend to be smaller than those from reading (Brown et al. 2008), likely because spoken input is transient (Brown 2001) and characterized by continuity, variability, and rapid temporal unfolding (Weber and Scharenborg 2012). These properties of speech processing increase the likelihood that unfamiliar words go unnoticed in listening contexts (van Zeeland 2014), thereby constraining opportunities for IVL. For instance, Rodgers (2013) found no significant effect of lexical coverage on vocabulary learning during television viewing. This null effect has been attributed to the presence of visual information, which may compensate for low lexical coverage by providing additional contextual support and reducing learners' reliance on lexical knowledge alone. Such findings suggest that the role of lexical coverage may vary across input modes. Addressing this gap, the present study examines the role of lexical coverage in IVL during listening. By manipulating lexical coverage through the use of pseudowords, the study isolates coverage effects while controlling content, thereby allowing for a precise investigation of how different coverage levels shape learning conditions in listening input. In doing so, this study contributes to IVL research by clarifying the lexical conditions under which listening input can effectively support IVL.

3. Methods

3.1 Participants

A total of 77 Korean third-grade middle school students (L1 = Korean) participated in the study. The participants were drawn from three intact classes that were randomly assigned to one of three lexical coverage conditions: 98%, 95%, or 90%. The 98% condition included 25 students (10 female, 15 male), the 95% condition included 26 students (12 female, 14 male), and the 90% condition included 26 students (10 female, 16 male). All participants had studied English as a foreign language for a minimum of seven years in accordance with the national curriculum.

To ensure that the intended lexical coverage levels were accurately implemented, it was necessary to confirm that participants had mastered all lexical items in the listening passages except for the pseudowords. For this purpose, a Vocabulary Levels Test (VLT; Webb et al. 2017) was administered. The VLT is a well-established and rigorously validated measure of vocabulary knowledge (Read 2000) and has been widely used to estimate learners' lexical coverage and vocabulary levels due to its strong face validity (Rodgers 2013). The Korean version of the 1,000-word-level VLT was administered in a paper-and-pencil format. Following Webb et al. (2017), a score of 29 out of 30 or higher was taken to indicate mastery of the 1,000-word level. Six participants who did not meet this criterion were excluded from the study.

3.2 Instruments

3.2.1 Target words

Target words were replaced by pseudowords to meet the lexical coverage levels. Monosyllabic pseudowords were created with the *ARC Nonword Database* (Rastle et al. 2002), while polysyllabic pseudowords were created using Wordoid (n.d.), a tool that generates pronounceable pseudowords. All pseudowords conformed to the phonologic and orthographic constraints of English, and had the same number of syllables, characters, inflections, and suitable suffixes as the original words to sound like the word class of the target words. The target words consisted of nouns, verbs, and adjectives. Adverbs were excluded because their characteristic endings could give away the grammatical functions of the words.

Each target pseudoword appeared four times within the listening passage. Because the passage was played twice during the treatment session, participants were exposed to each target pseudoword a total of eight times. This decision was informed by previous research showing that the largest gains in listening-based vocabulary learning occur around five to six encounters (Vidal 2011) and that knowledge of spoken forms and part-of-speech information emerges after approximately seven encounters (van Zeeland and Schmitt 2013). Importantly, the listening passages differed in the number of pseudowords introduced to achieve the intended lexical coverage levels, thus number of target items tested differed by lexical coverage condition. Specifically, the 98% condition included 3 target pseudowords, the 95% condition included 8 target pseudowords, and the 90% condition included 15 target pseudowords. Only the pseudowords introduced in each participant's assigned coverage condition were included as target items in the vocabulary tests. Because the number of target items differed across coverage conditions, all vocabulary test scores were converted to percentage accuracy scores to allow comparison across groups.

Prior knowledge of target vocabulary can be controlled using several methods, including selecting low-frequency words, consulting instructors familiar with learners' proficiency levels, administering vocabulary

pretests, or using pseudowords (e.g., Horst 2005, Schmitt 2010, van Zeeland and Schmitt 2013, Waring and Takaki 2003). Among these options, the present study adopts the pseudoword approach to manipulate lexical coverage while controlling all remaining vocabulary, thereby enabling lexical coverage to be examined in isolation. Although the use of pseudowords has been criticized for limited ecological validity (Giordano 2023, Webb 2007) and its potential to inflate learning outcomes (Mohamed 2018), it remains a methodologically robust approach for experimental research because it minimizes the likelihood of incidental exposure to target items outside the treatment session (Teng 2016). Moreover, the pseudoword approach ensures that participants have no prior knowledge of the target items, allowing any observed vocabulary gains to be attributed solely to the treatment and rendering separate vocabulary pretests unnecessary (Webb et al. 2023). Details of the pseudowords used in this study are provided in Table 1.

Table 1. Target Words and their Substituted Pseudowords

	Target words	Pseudowords
1	morning	*scroove*
2	clean	*pring*
3	happy	*crong*
4	think	**trisk**
5	class	**slamp**
6	room	**varm**
7	talk	**wonk**
8	friend	**sprock**
9	school	glanch
10	walk	yarm
11	look	pake
12	teacher	stranch
13	homework	spleague
14	tired	plomp
15	lunch	grosh

Note. The unmarked pseudowords and the pseudowords in bold and italics were all in the 90% version. The pseudowords in bold, and the pseudowords in italics were the pseudowords in the 95% version. The pseudowords in italics were the pseudowords in the 98% version.

3.2.2 Listening passage

The listening passage was a fictional narrative depicting an ordinary school day of a student. An initial version was generated by ChatGPT (OpenAI 2025), based on a prompt requesting a narrative closely related to the everyday experiences of Korean middle school students. Narrative texts were selected because they provide more favorable conditions for IVL than expository texts (Gardner 2004).

The initial version was subsequently edited to control the frequency and contextual richness of the target words. Specifically, each target word was embedded four times, and the surrounding contexts were adjusted to allow their meanings to be inferred from context. To minimize unintended comprehension difficulty prior to lexical manipulation, the base narrative was simplified so that all lexical items fell within the most frequent 1,000 words of English, as verified using the Vocabulary Profiler (Cobb 2002). This ensured that, before pseudoword substitution, the passage imposed minimal lexical demands relative to participants' demonstrated vocabulary mastery. Consequently, any changes in processing demands across conditions can be attributed specifically to the systematic lexical coverage manipulation rather than uncontrolled lexical difficulty in the base text. The final simplified version of the passage consisted of 600 words.

In addition to lexical control, the overall difficulty of the listening passage was evaluated to ensure its suitability for the participants' proficiency level. Two experienced English teachers of third-grade middle school students

independently rated the passage in terms of topic difficulty, grammatical complexity, and the amount of background knowledge required, using a 10-point scale (1 = very low, 10 = very high). Both teachers assigned a rating of 1 across all three dimensions, indicating that the passage was well within the learners' proficiency range and required minimal background knowledge.

Based on the simplified version, three versions of the listening passage were created to represent different levels of lexical coverage (98%, 95%, and 90%) by systematically replacing words with pseudowords. Following Reynolds' (2015) distinction among degrees of word form variation, only identical word forms were used, as previous research has shown that IVL is greater when no inflectional or derivational variation is present. Accordingly, each target word appeared in an identical form four times throughout the passage. To achieve the intended coverage levels, 12 words were replaced by pseudowords in the 98% condition, 30 words in the 95% condition, and 60 words in the 90% condition. Each target pseudoword was embedded in semantically supportive narrative contexts containing local syntactic cues, situational redundancy, and surrounding lexical information intended to facilitate meaning inference.

All three versions of the listening passage were recorded using the same AI text-to-speech voice simulating a native American English female speaker, generated via ElevenLabs (n.d.). To minimize variation across recordings, identical pitch settings were used, and the speech rate was fixed at 155 words per minute. This rate was selected with reference to the EBS National Middle School English Listening Ability Assessment, in which the speech rate for third-grade listening passages typically ranges from approximately 129 to 139 words per minute. Moreover, prior research indicates that speech rates between 126 and 172 words per minute do not adversely affect listening comprehension (Diehl et al. 1959). Accordingly, the passages were recorded at 155 words per minute, resulting in a total duration of 3 minutes and 50 seconds. Consistent with the EBS assessment, all recordings were produced in an American English accent. An overview of the three listening passages is provided in Table 2.

Table 2. Overview of the Three Listening Passages

Lexical coverage (%)	Real word count	Pseudoword count	Recording time (min.s)
98	588 words	12 words	3.54 min
95	570 words	30 words	3.57 min
90	540 words	60 words	3.57 min

3.2.3 Vocabulary test

This study adopted the dimensions approach to vocabulary assessment (Read 2000), which measures multiple aspects of word knowledge in order to capture even small increments in vocabulary learning. Vocabulary development is widely understood as incremental and componential, involving the gradual accumulation of phonological, orthographic, grammatical, and semantic knowledge. This approach is particularly appropriate for listening-based IVL, where vocabulary gains are typically smaller than those observed in reading contexts (Brown et al. 2008, Vidal 2011). Accordingly, vocabulary knowledge was assessed across three dimensions: (a) spoken word form recognition, (b) part-of-speech (grammar) recognition, and (c) meaning recall. In listening contexts, learners must not only recognize the spoken form of a word but also process its syntactic role within the unfolding sentence. Sensitivity to distributional and grammatical cues may therefore emerge prior to stable form–meaning mapping. Previous research on listening-based IVL (e.g., van Zeeland and Schmitt 2013) has similarly treated grammatical category knowledge as an early-developing dimension of lexical acquisition. Accordingly, part-of-speech recognition in the present study is interpreted not as a terminal learning outcome, but as an intermediate developmental indicator reflecting learners' integration of unfamiliar forms into emerging syntactic representations. Spoken word form recognition and part-of-speech recognition were measured using multiple-

choice recognition tests, while meaning knowledge was assessed through a recall test. During the form recognition and part-of-speech recognition tests, participants were presented with both spoken and written forms of the target pseudowords. In the meaning recall test, only the written forms of the target pseudowords were provided.

To ensure that the tests measured vocabulary knowledge rather than comprehension of English definitions, all word definitions were presented in Korean. The test also included five comprehension questions. This procedure ensured that participants approached the listening task with a focus on comprehension rather than vocabulary learning. In order to evaluate whether the lexical coverage manipulation affected global comprehension, comprehension question accuracy was calculated separately for each coverage condition. A one-way ANOVA on comprehension accuracy revealed no significant differences among the 90%, 95%, 98% lexical coverage groups (F (2, 74) = 0.25, p = .780, η^2p = .007). Importantly, comprehension accuracy remained comparable across conditions (90%: 88.46%, 95%: 87.69%, 98%: 92.00%), suggesting that although processing demands may have increased under lower coverage conditions, participants were still able to construct a coherent representation of the narrative. Therefore, differences in vocabulary learning cannot be attributed solely to global comprehension breakdown. Although answers to these questions were not included in the main IVL analyses, they provide contextual evidence regarding the processing demands imposed by different coverage levels. The tests were administered in a paper-and-pencil format and replicated the test formats used in van Zeeland and Schmitt (2013). The same vocabulary tests were administered at both the immediate and delayed tests. A sample of the vocabulary test is provided in Appendix, and the three dimensions of vocabulary knowledge are explained below.

3.2.3.1 Form recognition

The word form recognition test required participants to identify the pseudoword that sounded familiar from four spoken options, one of which corresponded to a target word. An additional option, *"I don't remember any of these,"* was provided and explicitly recommended to reduce blind guessing when none of the options sounded familiar. Because recognizing the spoken form of a word is a prerequisite for vocabulary learning through listening, this receptive test format was considered appropriate. To ensure that the test measured spoken word form recognition only, response options were presented orally and labeled as A, B, C, or D, while participants were provided with an answer sheet listing the corresponding letters. Each item was read twice by the same AI speaker used to record the listening passages.

3.2.3.2 Part-of-speech (grammar) recognition

The part-of-speech recognition test assessed participants' knowledge of the grammatical category of the target words, which constitutes an integral component of vocabulary knowledge. Participants heard each target word and were asked to indicate its part of speech on paper or to select an *"I don't know"* option if unsure. To support grammatical identification without revealing meaning, three example sentences illustrating different possible parts of speech for the target word were provided in written form. Each target word was presented twice by the same AI speaker used to record the listening passages.

3.2.3.3 Meaning recall

In the meaning recall test, participants were presented with all 15 target items and asked to write down anything they knew about their meanings, including a Korean translation, a synonym, or a brief explanation. Following the assumption in van Zeeland and Schmitt (2013), it was assumed that participants had already formed a link between

the spoken and written forms of the target words, as both forms had been presented in the part-of-speech recognition test. To support recall without providing semantic cues, the target words were embedded in the same neutral sentences used in the part-of-speech recognition test, as recalling the meanings of target words in complete isolation has been shown to be overly demanding (Donkaewbua 2007).

3.2.4 Procedure

All participants completed the Korean version of the VLT at the 1,000-word level. Those who scored 29 out of 30 or higher were assigned to one of three lexical coverage conditions (98%, 95%, or 90%), with 25, 26, and 26 participants in each group, respectively. To maintain the incidental nature of vocabulary learning, participants were not informed in advance of the vocabulary tests, consistent with the methodological definition of IVL (Hulstijn 2001). Instead, the task was framed as a listening comprehension activity, with instructions emphasizing understanding the overall meaning of the passages. Participants were informed that the passages contained pseudowords and were encouraged to comprehend the stories well enough to answer the comprehension questions. Although participants were not informed of any subsequent vocabulary test, they were told that the passages contained pseudowords and were encouraged to understand them in order to answer comprehension questions. This instruction likely increased the perceptual salience of unfamiliar forms relative to fully naturalistic listening conditions. Therefore, the present study operationalizes incidental learning in the methodological sense (i.e., absence of forewarning of a vocabulary test; Hulstijn 2001), but under comprehension-oriented instructions with salient unknown forms.

The experiment was conducted during a regular 45-minute English lesson using classroom stereo equipment. The three vocabulary tests were administered sequentially in the following order: form recognition, part-of-speech recognition, and meaning recall. Because the part-of-speech test presented the written form of each pseudoword, participants received an additional orthographic exposure prior to completing the meaning recall task. This sequencing may have strengthened spoken–written form connections during the immediate test. However, as the procedure was identical across lexical coverage conditions, any additional exposure effects were constant across groups. Immediately after listening, participants completed five comprehension questions, followed by the three vocabulary tests. The same vocabulary tests were administered again three weeks later to assess retention, as a three-week delayed posttest is indicative of stable and durable learning (Schmitt 2010). After both the immediate and delayed tests, participants were fully debriefed regarding the purpose of the study and informed that the target items were pseudowords rather than real English words.

3.2.5 Data analysis

All test responses were scored dichotomously as either correct or incorrect, with one point awarded for each correct response. Raw scores were converted into percentage accuracy scores to allow for meaningful comparison across conditions.

Statistical analyses were conducted to address the three research questions. To address research question 1, which examined whether lexical coverage (90%, 95%, 98%) affected learners' performance on the three dimensions of incidental vocabulary knowledge (form recognition, part-of-speech recognition, and meaning recall) on the immediate test, a mixed-design repeated-measures ANOVA was conducted. Lexical coverage was treated as a between-subjects factor, and vocabulary knowledge dimension as a within-subjects factor. To address question 2, which investigated whether lexical coverage influenced the retention of incidentally learned vocabulary three weeks after exposure, the same mixed-design repeated-measures ANOVA framework was applied to the delayed

test data, again with lexical coverage as the between-subjects factor and vocabulary knowledge dimension as the within-subjects factor. Question 3, which examined whether lexical coverage interacted with time (immediate vs. delayed) in affecting IVL across vocabulary knowledge dimensions, was examined with a three-way mixed-design repeated-measures ANOVA. In this analysis, lexical coverage was included as a between-subjects factor, while time and vocabulary knowledge dimension were included as within-subjects factors. Where significant main effects or interactions were observed, Bonferroni-adjusted post hoc comparisons and simple-effects analyses were conducted to further interpret the results. Effect sizes were calculated using Cohen's *d* for pairwise comparisons and partial eta squared (η^2p) for ANOVA effects. Following Cohen's (1988) conventions, effect sizes of 0.20 were interpreted as small, 0.50 as moderate, and 0.80 as large.

4. Results

This section reports the statistical analyses addressing the three research questions. Descriptive statistics summarize vocabulary performance across lexical coverage conditions and vocabulary knowledge dimensions. Mixed-design repeated-measures ANOVAs examine the effects of lexical coverage and vocabulary knowledge dimension on immediate and delayed test performance, followed by a three-way mixed repeated-measures ANOVA testing the interaction among lexical coverage, time, and vocabulary knowledge dimension. Where appropriate, post hoc pairwise comparisons and simple-effects analyses are conducted.

4.1 Descriptive Statistics

Tables 3 and 4 present the descriptive statistics for the immediate and delayed vocabulary tests, organized by lexical coverage condition and vocabulary knowledge dimensions. According to Table 3, a clear coverage-related trend is evident, with mean scores increasing as lexical coverage increases. The 98% coverage group consistently outperformed the 90% and 95% groups across all three dimensions, particularly in the meaning dimension (M = 56.00), whereas performance in the 90% and 95% conditions was lower. Across dimensions, scores for form recognition and grammatical category recognition were similar, while meaning recall was markedly weaker, indicating that meaning knowledge was more difficult to acquire incidentally than form or grammatical knowledge immediately after exposure.

Table 3. Descriptive Statistics of the Immediate Vocabulary Test

Coverage	Dimension	M	*SD*	N
90%	Form	55.13	24.48	26
	Grammar	54.10	22.65	26
	Meaning	24.62	25.70	26
	Total	44.62	20.78	26
95%	Form	57.69	27.17	26
	Grammar	63.94	28.58	26
	Meaning	24.04	31.80	26
	Total	48.56	25.74	26
98%	Form	74.67	29.31	25
	Grammar	70.67	33.78	25
	Meaning	56.00	38.15	25
	Total	67.11	28.95	25

Table 4. Descriptive Statistics of the Delayed Vocabulary Test

Coverage	Dimension	M	*SD*	N
90%	Form	30.77	23.25	26
	Grammar	24.10	23.10	26
	Meaning	4.87	9.53	26
	Total	19.91	15.18	26
95%	Form	31.25	25.56	26
	Grammar	31.73	21.57	26
	Meaning	9.13	17.52	26
	Total	24.04	16.51	26
98%	Form	46.67	23.57	25
	Grammar	40.00	31.92	25
	Meaning	17.33	27.42	25
	Total	34.67	20.37	25

Table 4 displays the delayed vocabulary test performance administered three weeks after the initial exposure. Overall performance declined across all conditions compared to the immediate test. Nevertheless, a clear ordering by lexical coverage remained evident. The 98% coverage group showed the highest retention across all three dimensions (Form: $M = 46.67$, $SD = 23.57$; Grammar: $M = 40.00$, $SD = 31.92$; Meaning: $M = 17.33$, $SD = 27.42$), whereas the 90% and 95% coverage groups exhibited similarly low levels of retention. Across vocabulary knowledge dimensions, form recognition and grammatical category recognition were retained to a greater extent than meaning knowledge. Retention of meaning knowledge was particularly limited, indicating that most participants failed to recall the meaning of the target words after three weeks. Although the 98% coverage group showed higher meaning scores than the other groups, the overall level of retention remained low, reflecting attrition in this dimension.

Overall, the descriptive statistics reveal a gradual increase in scores as lexical coverage increases, with the 98% lexical coverage group showing the highest mean scores across measures. The differences between the 90% and 95% coverage groups were relatively small. Across vocabulary knowledge dimensions, performance on form recognition and grammatical category recognition was comparable, whereas performance on meaning recall was markedly lower. Across all groups, scores on the delayed test decreased by approximately 45–50% compared to the immediate test, particularly in the meaning dimension.

4.2 Effect of Lexical Coverage on Immediate L2 Incidental Vocabulary Learning

Table 5 presents the results examining the effects of lexical coverage and vocabulary knowledge dimension on learners' performance on the immediate vocabulary test. The analysis revealed a significant main effect of lexical coverage, $F(2, 222) = 12.723$, $p < .001$, $\eta^2p = .103$, indicating that lexical coverage had a statistically significant effect on IVL in the immediate test. The effect size was medium according to Cohen (1988), suggesting that the effect was not only statistically significant but also meaningful. A significant main effect of vocabulary knowledge dimension was also observed, $F(2, 222) = 23.275$, $p < .001$, $\eta^2p = .173$, indicating that IVL outcomes differed depending on the vocabulary knowledge dimension. The effect size was large. However, the interaction between lexical coverage and vocabulary knowledge dimension was not statistically significant, $F(4, 222) = 1.235$, $p = .297$, $\eta^2p = .022$, indicating that the effect of lexical coverage was relatively consistent across the three vocabulary knowledge dimensions.

Table 5. Results of the Mixed-Design Repeated-Measures ANOVA on the Immediate Vocabulary Test

Source	*SS*	*df*	*MS*	*F*	*p*	η^2p
Coverage	21,942.14	2	10,971.07	12.723	< .001**	.103
Dimension	40,138.94	2	20,069.47	23.275	< .001**	.173
Coverage × Dimension	4,258.57	4	1,064.64	1.235	.297	.022
Error	191,425.52	222	862.28			
Total	257,765.17	230				

Note. ** $p < .001$

To further examine the significant main effect of lexical coverage observed in the results for the immediate test, Bonferroni-adjusted post hoc analyses were conducted to identify pairwise differences among the three coverage conditions. Bonferroni adjustment was used for all post hoc comparisons due to its conservative nature and applicability to both equal and unequal sample sizes. Table 6 presents the results. The results indicated a clear advantage for the 98% lexical coverage condition. Specifically, the 98% coverage group outperformed the 90% coverage group by an average of 22.50 points ($p = .007$, Cohen's $d = -0.896$), representing a very large effect size. A borderline effect was observed between the 95% and 98% coverage groups. Although this difference did not reach statistical significance after Bonferroni adjustment ($p = .058$), the effect size was moderate ($d = -0.678$), suggesting potential practical significance. In contrast, no meaningful difference was found between the 90% and 95% coverage groups, either statistically or practically ($p = 1.000$, $d = -0.169$).

Table 6. Post Hoc Pairwise Comparisons for the Main Effect of Lexical Coverage

Comparison	Mean Diff	*t*	*p*(Bonf)	Cohen's *d*	Interpretation
90% vs 95%	-3.94	-0.608	1.000	-0.169	n.s.
90% vs 98%	-22.50	-3.197	.007*	-0.896	$p < .05$
95% vs 98%	-18.55	-2.421	.058	-0.678	n.s.†

Note. * $p < .05$

To further examine the significant main effect of vocabulary knowledge dimension, Table 7 presents the results. No significant difference was found between form recognition and grammar recognition ($p = 1.000$, Cohen's $d = -0.016$). In contrast, performance on the meaning dimension was significantly lower than on both the form dimension (mean difference = 27.73, $p < .001$, $d = 0.873$) and the grammar dimension (mean difference = 28.19, $p < .001$, $d = 0.875$), with both comparisons yielding very large effect sizes.

Table 7. Post Hoc Pairwise Comparisons for the Main Effect of Vocabulary Knowledge Dimension

Comparison	Mean Diff	*t*	*p*(Bonf)	Cohen's *d*	Interpretation
Form vs Grammar	-0.47	-0.101	1.000	-0.016	n.s.
Form vs Meaning	27.73	5.415	< .001**	0.873	Sig. (L)
Grammar vs Meaning	28.19	5.428	< .001**	0.875	Sig. (L)

Note. ** $p < .001$

4.3 Effect of Lexical Coverage on Retention of L2 Incidental Vocabulary Learning

Table 8 presents the results examining the effects of lexical coverage and vocabulary knowledge dimension on learners' performance on the delayed vocabulary test administered three weeks after the immediate test. The analysis revealed a significant main effect of lexical coverage, $F(2, 222) = 8.118$, $p < .001$, $\eta^2p = .068$, indicating

that lexical coverage continued to have a statistically significant effect on the retention of incidentally learned vocabulary after three weeks. However, the effect size was smaller than that observed on the immediate test (η^2p = .103), suggesting an attenuation of the coverage effect over time. A significant main effect of vocabulary knowledge dimension was also observed, F (2, 222) = 26.978, p < .001, η^2p = .196. Notably, the effect size was larger than that found on the immediate test (η^2p = .173), indicating increased divergence in retention across vocabulary knowledge dimensions over time. The interaction between lexical coverage and vocabulary knowledge dimension was not statistically significant, F (4, 222) = 0.268, p = .899, η^2p = .005, indicating that the effect of lexical coverage on delayed vocabulary retention was relatively consistent across the three vocabulary knowledge dimensions.

Table 8. Results of the Mixed-Design Repeated-Measures ANOVA for the Delayed Vocabulary Test

Source	*SS*	*df*	*MS*	*F*	*p*	η^2p
Coverage	8,819.34	2	4,409.67	8.118	< .001**	.068
Dimension	29,309.23	2	14,654.61	26.978	< .001**	.196
Coverage × Dimension	581.29	4	145.32	0.268	.899	.005
Error	120,591.59	222	543.21			
Total	159,301.45	230				

Note. ** p < .001

Table 9. Post Hoc Pairwise Comparisons for the Main Effect of Lexical Coverage

Comparison	Mean Diff	*t*	*p*(Bonf)	Cohen's *d*	Interpretation
90% vs 95%	-4.12	-0.937	1.000	-0.260	n.s.
90% vs 98%	-14.75	-2.940	.015*	-0.824	p < .05
95% vs 98%	-10.63	-2.051	.137	-0.574	n.s.

Note. * p < .05

To further examine the significant main effect of lexical coverage observed in the results of the delayed test, Bonferroni-adjusted post hoc analyses were conducted to identify pairwise differences among the three lexical coverage conditions. Table 9 presents the results of the Bonferroni-adjusted post hoc pairwise comparisons for the main effect of lexical coverage on the delayed vocabulary test. The results indicated a persistent advantage for the 98% lexical coverage condition. Specifically, the 98% coverage group demonstrated significantly higher vocabulary retention than the 90% coverage group (mean difference = 14.75 points, p = .015, Cohen's d = −0.824), representing a large effect size. Although the difference between the 95% and 98% coverage conditions did not reach statistical significance after Bonferroni correction (p = .137), the effect size was moderate (d = −0.574).

Table 10. Post Hoc Pairwise Comparisons for the Main Effect of Vocabulary Knowledge Dimension

Comparison	Mean Diff	*t*	*p*(Bonf)	Cohen's *d*	Interpretation
Form vs Grammar	4.25	1.029	.915	0.166	n.s.
Form vs Meaning	25.74	7.079	< .001**	1.141	Sig. (L)
Grammar vs Meaning	21.48	5.716	< .001**	0.921	Sig. (L)

Note. ** p < .001

To further examine the significant main effect of vocabulary knowledge dimension revealed for the delayed test, Bonferroni-adjusted post hoc analyses were conducted to determine which dimensions differed in terms of vocabulary retention. Table 10 presents the results of the Bonferroni-adjusted post hoc pairwise comparisons for the main effect of vocabulary knowledge dimension on the delayed vocabulary test. As in the immediate test, no

significant difference was observed between retention of form and grammar knowledge (p = .915, Cohen's d = 0.166). In contrast, retention in the meaning dimension was markedly lower. Performance on the meaning dimension was lower than that on both the form dimension (mean difference = 25.74 points, d = 1.141) and the grammar dimension (mean difference = 21.48 points, d = 0.921), with both comparisons yielding very large effect sizes.

4.4 Interaction between Time and Lexical Coverage in L2 Incidental Vocabulary Learning

Table 11 summarizes the main and interaction effects of lexical coverage, time, and vocabulary knowledge dimension on vocabulary performance across the immediate and delayed tests. The analysis revealed a main effect of time, F (1, 74) = 202.421, p < .001, η^2p = .732. This large effect size confirms a substantial decline in performance over time, with an average decrease of approximately 27.22 points from the immediate test to the delayed test. Beyond this overall decline, a significant Coverage × Time interaction, F (2, 148) = 8.320, p < .001, η^2p = .036, showed that retention over time varied as a function of lexical coverage. In addition, lexical coverage interacted significantly with vocabulary knowledge dimension, F (4, 148) = 6.096, p < .001, η^2p = .052, suggesting that coverage influenced the different knowledge dimensions in different ways. The interaction between time and vocabulary knowledge dimension was also significant, F (2, 148) = 4.892, p = .008, η^2p = .022. Lastly, the significant three-way interaction among Coverage, Time, and Dimension, F (4, 148) = 3.191, p < .05, η^2p = .079, points to a complex pattern in which the effects of lexical coverage on specific vocabulary dimensions varied across testing occasions.

Table 11. Results of the Three-Way Mixed Repeated-Measures ANOVA for Lexical Coverage, Time, and Vocabulary Knowledge Dimension

Source	*SS*	*df*	*MS*	*F*	*p*	η^2p
Between-subjects						
Coverage	29,206.35	2	14,603.17	6.041	<.004*	.140
Within-subjects						
Time	85,162.63	1	85,162.63	202.421	< .001**	.732
Dimension	68,533.76	2	34,266.88	70.144	< .001**	.487
2-way Interactions:						
Coverage × Time	1,555.13	2	777.56	8.320	< .001**	.036
Coverage × Dimension	2,278.95	4	569.74	6.096	< .001**	.052
Time × Dimension	914.41	2	457.20	4.892	< .008*	.022
3-way Interaction:						
Cov × Time × Dimension	2,560.92	4	640.23	3.191	< .015*	.079
Error	41,495.32	148	93.46			

Note. * p < .05, ** p < .001, *F, p, and* η^2p values for main effects and the three-way interaction are from R mixed-design ANOVA with separate error terms (between-subjects error *df* = 74*;* within-subjects error *df* = 74 or 148). *SS and MS* reflect effect sums of squares.

To unpack this three-way interaction, follow-up simple-effects analyses were conducted separately for each vocabulary knowledge dimension (form, grammar, and meaning), examining how the effect of lexical coverage differed across distinct time points within each dimension.

Table 12. Simple Effects of Lexical Coverage × Time Interaction for the Form Dimension

Coverage	Time	M	*SD*	Mean Diff	*p*	Cohen's *d*
90%	Immediate	55.13	24.48	24.36	< .001**	1.02
	Delayed	30.77	23.25			
95%	Immediate	57.69	27.17	26.44	< .001**	1.00
	Delayed	31.25	25.56			
98%	Immediate	74.67	29.31	28.00	< .001**	1.05
	Delayed	46.67	23.57			

Note. ** $p < .001$

Table 13. Simple Effects of Lexical Coverage × Time Interaction for the Grammar Dimension

Coverage	Time	M	*SD*	Mean Diff	*p*	Cohen's *d*
90%	Immediate	54.10	22.65	30.00	< .001**	1.31
	Delayed	24.10	23.10			
95%	Immediate	63.94	28.58	32.21	< .001**	1.27
	Delayed	31.73	21.57			
98%	Immediate	70.67	33.78	30.67	< .001**	0.93
	Delayed	40.00	31.92			

Note. ** $p < .001$

Table 14. Simple Effects of Lexical Coverage × Time Interaction for the Meaning Dimension

Coverage	Time	M	*SD*	Mean Diff	*p*	Cohen's *d*
90%	Immediate	24.62	25.70	19.75	< .001**	1.02
	Delayed	4.87	9.53			
95%	Immediate	24.04	31.80	14.91	< .008*	0.58
	Delayed	9.13	17.52			
98%	Immediate	56.00	38.15	38.67	< .001**	1.16
	Delayed	17.33	27.42			

Note. * $p < .05$, ** $p < .001$

Tables 12, 13, and 14 present the effects of the Coverage × Time interaction within each vocabulary knowledge dimension. For the form dimension (Table 12), a significant time-related decline was observed within each lexical coverage condition, with the 98% coverage group ($M = 74.67$) outperforming the 90% ($M = 55.13$) and 95% ($M = 57.69$) groups. However, at the delayed test, while the 98% group maintained higher scores ($M = 46.67$), differences in mean performance between coverage conditions appeared smaller at the delayed test. The decline from immediate to delayed was relatively uniform across coverage conditions ($d = 1.02$, 1.00, and 1.05 for 90%, 95%, and 98% respectively), indicating that higher lexical coverage facilitated initial form learning but did not confer a differential advantage for long-term retention.

For the grammar dimension (Table 13), a similar pattern emerged. Learners in the 98% coverage condition demonstrated the highest performance at both immediate ($M = 70.67$) and delayed ($M = 40.00$) tests. The effect sizes for the time-related decline were large across all coverage conditions ($d = 1.31$, 1.27, and 0.93 for 90%, 95%, and 98% respectively). It should be noted that the 98% coverage group showed a smaller decline in effect size ($d = 0.93$) compared to the other groups, suggesting a slightly more durable retention of grammar knowledge under higher coverage conditions, albeit a modest difference.

The most notable difference emerged in the meaning dimension (Table 14). At the immediate test, the 98% coverage group ($M = 56.00$) showed higher mean scores than the 90% ($M = 24.62$) and 95% ($M = 24.04$) groups,

which showed nearly identical performance. This pattern aligns with the significant three-way interaction in Table 11. Figure 1 presents boxplots with individual data points illustrating the distribution of meaning recall scores across lexical coverage conditions at the immediate and delayed tests. For the lower coverage conditions, the proportion of zero scores was 1 out of 26 (3.8%) in the 90% condition and 0 out of 26 (0.0%) in the 95% condition at the immediate test, and 5 out of 26 (19.2%) and 3 out of 26 (11.5%) at the delayed test, respectively.

However, the 98% group also showed the largest decline over time (mean difference = 38.67, d = 1.16), compared to the 90% (mean difference = 19.75, d = 1.02) and 95% groups (mean difference = 14.91, d = 0.58). At the delayed test, while the 98% group still maintained higher meaning scores (M = 17.33) than the other groups, the relative advantage had decreased. The 95% coverage group showed the smallest decline in effect size (d = 0.58), a pattern that likely reflects very low baseline performance in the lower coverage conditions rather than superior retention.

These dimension-specific analyses reveal that the significant Coverage × Time × Dimension interaction reflects different patterns of lexical coverage effects across vocabulary knowledge types. For form and grammar knowledge, higher lexical coverage provided an advantage that diminished over time. In contrast, the meaning dimension exhibited a different pattern wherein the 98% coverage condition produced higher initial learning (M = 56.00) compared to the 90% (M = 24.6) and 95% (M = 24.3) conditions. This differential pattern across dimensions—uniform attrition for Form and Grammar versus coverage-dependent attrition for Meaning—is the drive of the significant three-way interaction (see Figure 2).

While the 98% group subsequently showed the largest decline over time (mean difference = 38.67), this should not be interpreted as the 98% condition being vulnerable to forgetting. Rather, the apparent stability of the 90% and 95% groups is attributable to very low initial learning of meaning, which resulted in limited knowledge available for subsequent attrition. This interpretation is supported by the distributional evidence shown in Figure 1, which indicates an increased incidence of zero-score responses at the delayed test in the 90% and 95% conditions. Consequently, the 98% condition was the only level sufficient to facilitate the initial construction of form-meaning links. However, the attrition observed in this group indicates that while high coverage is a prerequisite for establishing these initial mappings, the resulting connections remain fragile and subject to decay.

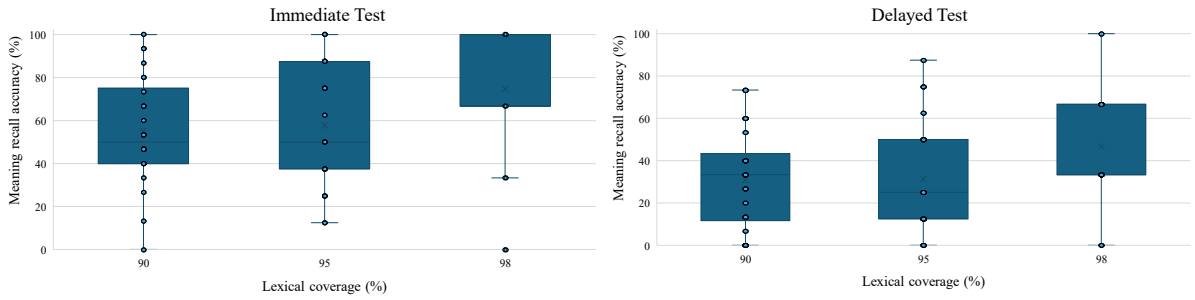

Figure 1. Distribution of Meaning Recall Scores on the Immediate and Delayed Tests by Lexical Coverage Condition

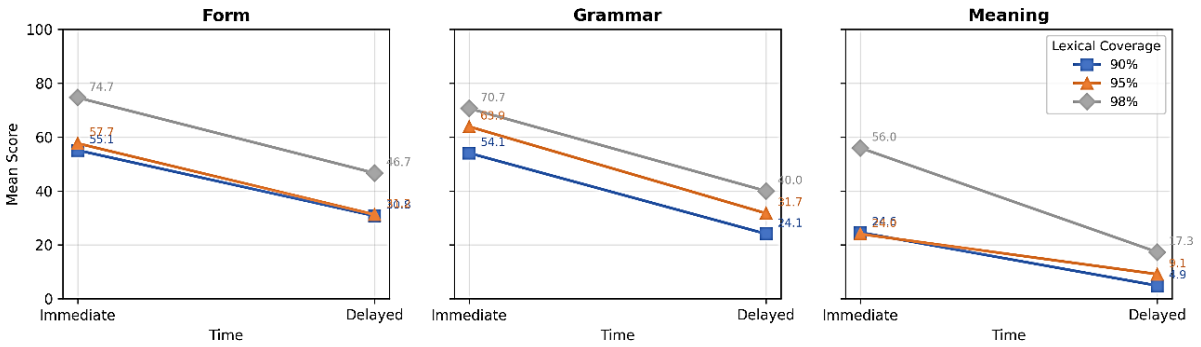

Figure 2. Interaction Between Time (Immediate vs. Delayed) and Lexical Coverage (90%, 95%, and 98%) across Vocabulary Knowledge Dimensions (Form, Grammar, and Meaning)

5. Discussion

5.1 Effect of Lexical Coverage on Immediate L2 Incidental Vocabulary Learning

The immediate test showed a significant effect of lexical coverage on IVL. Participants in the 98% coverage condition significantly outperformed those in the 90% condition, whereas no significant difference was found between the 90% and 95% conditions or the 95% and 98% conditions. This finding shows that the 98% lexical coverage condition had the highest gains in IVL, implicating that the relationship between coverage and learning is non-linear rather than involving a gradual increase across levels. While the contrast between 95% and 98% was statistically non-significant, the overall pattern reveals that a near-complete lexical coverage leads to more advantages when it comes to IVL. According to the Cognitive Load Theory proposed by Sweller (1988), working memory is limited. Thus, a split-attention effect occurs when allocating attention between various sources of input as comprehension takes place (Schnotz and Kürschner 2007). This theoretical explanation can be applied to the pattern found in the present study. In a text with high lexical coverage, it is possible to allot attention to the few unknown words, whereas for low coverage text, less cognitive resources are available for attending to new words. As sufficient attention to unfamiliar items in input is necessary for acquisition (Hulstijn 2001), high lexical coverage can provide a facilitative environment for IVL.

In addition, Ye (2024) claimed that one precondition for building L2 lexical knowledge is textual comprehension. Comprehension is not an automatic result of L2 vocabulary acquisition but a prerequisite for integrating unfamiliar vocabulary items. Although the exact threshold required for comprehension remains inconclusive, L2 listening studies have associated lexical coverage ranging from as low as 90% (Giordano 2023) to 98% (Stæhr 2009) with adequate listening comprehension. From this perspective, the higher vocabulary scores of the 98% lexical coverage group can be attributed to the possibility of better comprehension. That is, IVL becomes possible when ample lexical support exists, as listeners are less likely to experience breakdowns in processing and have more cognitive resources available for attending to new vocabulary. IVL, then, does not automatically arise from exposure alone; IVL happens when lexical coverage sustains comprehension of input.

Another finding involves the effect of IVL across different vocabulary knowledge dimensions. The vocabulary scores for meaning were lower than those for form and grammar knowledge, which suggests that more difficulty is involved in incidentally acquiring meaning than knowledge of form or grammar. It has been shown that the relative order of acquisition moves from the largest gains in form recognition and the smallest in meaning recall

(Brown et al. 2008, Hatami 2017). This is in line with van Zeeland and Schmitt's (2013) study where form and grammar recognition was thought of as developing before "high level of knowledge" such as meaning recall. Similarly, Chen and Truscott's (2010) study also reported that more exposures were needed for acquiring vocabulary meaning than other knowledge types like orthography and grammar. Overall, the acquisition profile from previous studies displayed similar patterns with the present study, with knowledge of meaning being particularly difficult to acquire.

No interaction was observed between lexical coverage and vocabulary knowledge dimension. In other words, the advantage of the 98% condition was similar across form, grammar, and meaning, rather than being restricted to or amplified in a particular dimension.

5.2 Effect of Lexical Coverage on Retention of L2 Incidental Vocabulary Learning

The effect of lexical coverage was significant in the delayed test. The 98% group performed significantly better than the 90% coverage group even after three weeks, while the difference between 95 and 98% coverage conditions was not statistically significant. Meanwhile, retention of meaning knowledge was lower across all lexical coverage conditions. That is, meaning knowledge was particularly susceptible to forgetting over time. This finding contrasts with those of van Zeeland and Schmitt (2013), which revealed that knowledge of meaning was most retained, whereas knowledge of grammar was more prone to attrition. One possible reason for this contrasting result can be because in their study, the initial meaning recall scores were much lower (8.5%) than those for form (45.8%) and grammar (33.7%). Because the level of meaning knowledge was limited from the start, there was little room for further decline to occur; thus, retention of meaning appeared relatively stable compared to other types of knowledge.

As in the immediate test, no interaction between lexical coverage and vocabulary knowledge was found in the delayed test, indicating that the relative difficulty of retaining meaning knowledge, compared to form and grammar knowledge, was consistent across all lexical coverage conditions. Higher lexical coverage led to better retention of L2 vocabulary knowledge but did not lessen attrition of meaning knowledge.

5.3 Interaction between Time and Lexical Coverage in L2 Incidental Vocabulary Learning

A significant interaction between lexical coverage and time was found, indicating that the effect of lexical coverage differed between the immediate and delayed tests. The magnitude of the effect of lexical coverage was reduced in the delayed tests. Specifically, while the 98% coverage group outperformed the other groups at both immediate and delayed tests, the difference between coverage conditions was smaller in the delayed test. This finding reveals that the IVL advantage associated with higher lexical coverage weakened over time. A similar pattern was also found in Vidal (2011), where learners acquired more vocabulary from reading than listening, but this difference disappeared in the delayed test. This collective result implies that input conditions may exert a stronger influence on early stages of IVL than on long-term retention.

Also, a significant interaction was found in Coverage × Time × Vocabulary Knowledge Dimension, showing that the effects of lexical coverage over time differed across vocabulary dimensions. For the form dimension, the effect of lexical coverage attenuated at the delayed test; while higher coverage facilitates initial form recognition, its advantage for retention is limited. For the grammar dimension, higher lexical coverage facilitated both initial learning and retention, while retention in the 98% coverage group was relatively more stable than the lower coverage groups. For the meaning dimension, the effect was more pronounced. The 98% coverage group had higher immediate tests results compared to the 90% and 95% conditions that both showed low performance.

However, the 98% group showed the largest decline over time. This interaction effect suggests that while high lexical coverage (98%) affects early stages of IVL, the acquired knowledge can be fragile. The relative "stability" of the lower coverage groups, in contrast, is attributed to their limited initial learning, leaving little knowledge available for subsequent attrition. While van Zeeland and Schmitt (2013) reported form and grammar knowledge to be more sensitive to attrition than meaning knowledge, the results of the present study showed a different pattern. In this study, meaning knowledge showed the most decline over time across all lexical coverage conditions, indicating that meaning knowledge was more vulnerable to attrition than form and grammar knowledge in IVL. This ties in with Waring and Takaki's (2003) reading study, where knowledge of meaning was lost faster than the other types of word knowledge.

6. Conclusion

The present study examined the effect of lexical coverage (90%, 95%, and 98%) on incidental vocabulary learning (IVL) from spoken narrative input. A total of 77 Korean EFL middle school students were assigned to one of three lexical coverage conditions, manipulated using pseudowords within a listening passage. The study employed a dimensions approach to assessment (Read 2000), measuring knowledge of spoken form, grammatical part of speech, and meaning recall both immediately after exposure and after a three-week delay to determine if varying levels of known words in a text affect initial learning and retention.

The first research question concerned immediate learning, and the results indicated that lexical coverage facilitated IVL, but only at the highest level. Learners in the 98% coverage condition significantly outperformed those in the 90% condition, while the 90% and 95% groups did not significantly differ. This pattern suggests a non-linear effect in which IVL becomes more possible as lexical coverage approaches near-complete levels, rather than increasing gradually across coverage conditions. This finding aligns with Cognitive Load Theory (Sweller 1988), suggesting that high coverage is necessary to prevent the split-attention effect and allow sufficient attentional resources for processing unknown items. Consistent with previous research on the incremental nature of vocabulary acquisition (e.g., Hatami 2017, van Zeeland and Schmitt 2013), knowledge of form and grammar was acquired more readily than meaning. While Pellicer-Sánchez and Schmitt (2010) previously found form recognition to be hard to acquire, the present study aligns more closely with Brown et al. (2008) and Chen and Truscott (2010), confirming that meaning recall is the most difficult dimension to acquire incidentally from listening.

Regarding the second research question on retention, the advantage of the 98% lexical coverage condition persisted three weeks after exposure, although the effect magnitude weakened over time. Similar to the results on immediate testing, retention of meaning knowledge was lower than that of form and grammar across all conditions. This contrasts with findings by van Zeeland and Schmitt (2013), who reported that meaning knowledge was relatively stable while form knowledge decayed over time. However, as noted in the discussion, this discrepancy is likely due to a methodological problem; in van Zeeland and Schmitt (2013), initial meaning scores were very low to begin with, leaving little room for decline.

Lastly, the third research question analyzed the interaction between coverage and time. A significant three-way interaction revealed that the levels of retention differed by dimension. For form and grammar, higher coverage provided an initial advantage that diminished over time. However, for meaning knowledge, only the 98% condition was facilitated IVL. This high vulnerability of incidentally learned meaning aligns with Waring and Takaki's (2003) reading study, which found meaning knowledge to be lost faster than other knowledge types, supporting

the componential view that meaning requires deeper processing than surface features (Nation 2001).

These findings hold significant pedagogical implications for EFL listening instruction. While previous listening research suggests that 95% coverage may suffice for adequate comprehension (e.g., Bonk 2000, van Zeeland and Schmitt 2013), this study indicates that near-complete coverage (98%) appears to provide favorable conditions for IVL in listening, particularly for establishing meaning. Teachers should be aware that materials designed for practicing listening comprehension may not all be sufficient for vocabulary development. To facilitate IVL, instructors should select listening materials where unknown words are sparse. Furthermore, given the fragility of meaning knowledge observed in this study, incidental exposure alone can be insufficient. Instead, meaning-focused listening may need to be supplemented with intentional post-listening activities for form-meaning connections to occur.

This study is not without limitations. The use of pseudowords, while providing experimental control, may not fully reflect the ecological complexity of natural language processing, as in the potential clues that cognates provide. Additionally, the study's design involved a single listening session, while future research could examine the effects of multiple exposures over longer periods. It should also be noted that written forms of the target pseudowords were presented during the part-of-speech recognition and meaning recall tasks, which could have provided orthographic exposure that differs from a strictly listening-only condition. There is the possibility that presentation of written forms may have partially created a listening-reading hybrid context during the testing stages. However, because this procedure was identical across lexical coverage conditions and occurred only after the listening phase, it is unlikely to have significantly affected the comparison of coverage effects. One last limitation concerns contextual inferability. Although the target pseudowords were introduced with contextual clues to facilitate inference, inferability was not independently validated through a separate rating task. Future research may incorporate independent inferability measures to create contextual conditions more precisely. Despite these limitations, this study contributes to IVL research by investigating the effect of lexical coverage in the domain of L2 listening, revealing that the fleeting nature of spoken input requires near-complete coverage.

Acknowledgement

During the preparation of this manuscript, the authors used ChatGPT (GPT-5.3, OpenAI) for language editing and stylistic refinement. All conceptualization, analysis, interpretation, and conclusions are solely those of the authors.

References

Bonk, W. J. 2000. Second language lexical knowledge and listening comprehension. *International Journal of Listening* 14(1), 14-31.

Brown, H. D. 2001. *Teaching by Principles: An Interactive Approach to Language Pedagogy* (2nd edition). White Plains, NY: Pearson Education.

Brown, R., R. Waring and S. Donkaewbua. 2008. Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language* 20(2), 136-163.

Chen, C. and J. Truscott. 2010. The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics* 31(5), 693-713.

Cobb, T. 2002. *Web Vocabprofile.* Accessed 1 December 2025 Available online at http://www.lextutor.ca/vp/

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, NJ: Erlbaum.

Crossley, S., K. Kyle and T. Salsbury. 2016. A usage-based investigation of L2 lexical acquisition: The role of

input and output. *The Modern Language Journal* 100(3), 702-715.

Diehl, C. F., R. C. White and K. W. Burk. 1959. Rate and communication. *Communications Monographs* 26(3), 229-232.

Donkaewbua, S. 2007. *The Effects of Previous Partial Word Knowledge on Vocabulary Learning through Listening.* Unpublished manuscript, Victoria University of Wellington, New Zealand.

ElevenLabs. n.d. *ElevenLabs Text-to-Speech*. Accessed 1 December 2025 Available online at https://elevenlabs.io/app/speech-synthesis/text-to-speech

Elgort, I. and P. Warren. 2014. L2 Vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning* 64(2), 365-414.

Ellis, R. 1999. *Learning a Second Language through Interaction*. Amsterdam: John Benjamins.

Gardner, D. 2004. Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics* 25(1), 1-37.

Giordano, M. J. 2023. Lexical coverage in dialogue listening. *Language Teaching Research* 27(6), 1504-1531.

Godfroid, A., J. Ahn, I. Choi, L. Ballard, Y. Cui, S. Johnston, S. Lee, A. Sarkar and H. Yoon. 2018. Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition* 21(3), 563-584.

Hatami, S. 2017. The impact of learner-related variables on second language incidental vocabulary acquisition through listening. *Vocabulary Learning and Instruction* 6(1), 1-20.

Horst, M. 2005. Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review* 61(3), 355-382.

Hulstijn, J. H. 2001. Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson, ed., *Cognition and Second Language Instruction*, 258-286. Cambridge: Cambridge University Press.

Jin, Z. and S. Webb. 2020. Incidental vocabulary learning through listening to teacher talk. *The Modern Language Journal* 104(3), 550-566.

Kim, J. 2023. *The Effects of EFL College Students' L1 Print Exposure and L2 Lexical Coverage in English Texts on L2 Lexical Guessing*. Doctoral dissertation, Seoul National University Graduate School, Korea.

Kurokawa, S., A. M. Hein and T. Uchihara. 2025. Incidental vocabulary acquisition through captioned viewing: A meta-analysis. *Language Learning* 75(4), 939-987.

Mehrpour, S. and M. Rahimi. 2010. The impact of general and specific vocabulary knowledge on reading and listening comprehension: A case of Iranian EFL learners. *System* 38(2), 292-300.

Mohamed, A. A. 2018. Exposure frequency in L2 reading: An eye-movement perspective of incidental vocabulary learning. *Studies in Second Language Acquisition* 40(2), 269-293.

Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, I. S. P. and R. Waring. 1997. Vocabulary size, text coverage and word lists. In N. Schmitt and M. McCarthy, eds., *Vocabulary: Description, Acquisition and Pedagogy*, 6-19. Cambridge: Cambridge University Press.

OpenAI. 2025. *ChatGPT* [Large Language Model]. Available online at http://chat.openai.com

Pellicer-Sánchez, A. and N. Schmitt. 2010. Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language* 22(1), 31-55.

Pellicer-Sánchez, A., S. Webb and A. Wang. 2024. How does lexical coverage affect the processing of L2 texts? *Applied Linguistics* 45(6), 953-972.

Penno, J. F., I. A. G. Wilkinson and D. W. Moore. 2002. Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the matthew effect? *Journal of Educational Psychology* 94(1), 23-33.

Peters, E., E. Heynen and E. Puimège. 2016. Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System* 63, 134-148.

Peters, E. and S. Webb. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40(3), 551-577.

Puimège, E. and E. Peters. 2019. Learners' English vocabulary knowledge prior to formal instruction: The role of learner-related and word-related variables. *Language Learning* 69(4), 943-977.

Rastle, K., J. Harrington and M. Coltheart. 2002. 358,534 Nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* 55A(4), 1339-1362.

Read, J. A. 2000. *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Reynolds, B. L. 2015. The effects of word form variation and frequency on second language incidental vocabulary acquisition through reading. *Applied Linguistics Review* 6(4), 467-497.

Rodgers, M. P. 2013. *English Language Learning through Viewing Television: An Investigation of Comprehension, Incidental Vocabulary Acquisition, Lexical Coverage, Attitudes, and Captions*. Doctoral dissertation, Victoria University of Wellington, New Zealand.

Schmitt, N. 2010. *Researching Vocabulary: A Vocabulary Research Manual*. Basingstoke, UK: Palgrave Macmillan.

Schnotz, W. and C. Kürschner. 2007. A reconsideration of cognitive load theory. *Educational Psychology Review* 19(4), 469-508.

Stæhr, L. S. 2009. Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition* 31(4), 577-607.

Sweller, J. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12(2), 257-285.

Teng, F. 2016. The effects of word exposure frequency on incidental learning of depth of vocabulary knowledge. *GEMA Online® Journal of Language Studies* 16(3), 53-70.

Teng, M. F. 2023. Understanding incidental vocabulary learning in practice: An interview with Stuart Webb. *Asian Journal of English Language Teaching* 32(1), 155-166.

Teng, M. F. 2024. Modality of input and factors affecting incidental vocabulary learning: Reading, listening, and viewing with captions. *Applied Linguistics Review* 16(4), 1607-1635.

Teng, M. F. 2025. Incidental vocabulary learning from listening, reading, and viewing captioned videos: Frequency and prior vocabulary knowledge. *Applied Linguistics Review* 16(1), 477-507.

Teng, M. F. and T. Uchihara. 2024. Factors affecting incidental vocabulary learning. In M. F. Teng and B. L. Reynolds, eds., *Researching Incidental Vocabulary Learning in a Second Language*, 131-148. New York: Routledge.

Uchihara, T., S. Webb and A. Yanagisawa. 2019. The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning* 69(3), 559-599.

van Zeeland, H. 2014. Lexical inferencing in first and second language listening. *The Modern Language Journal* 98(4), 1006-1021.

van Zeeland, H. and N. Schmitt. 2013. Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System* 41(3), 609-624.

Vidal, K. 2003. Academic listening: A source of vocabulary acquisition? *Applied Linguistics* 24(1), 56-89.

Vidal, K. 2011. A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning* 61(1), 219-258.

Waring, R. and M. Takaki. 2003. At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language* 15(2), 130-163.

Webb, S. 2007. The effects of repetition on vocabulary knowledge. *Applied Linguistics* 28(1), 46-65.

Webb, S. 2016. Learning vocabulary through meaning-focused input: Replication of Elley (1989) and Liu & Nation (1985). *Language Teaching Research* 20(1), 99-116.
Webb, S. 2021. Research investigating lexical coverage and lexical profiling: What we know, what we don't know, and what needs to be examined. *Reading in a Foreign Language* 33(2), 278-293.
Webb, S. and P. Nation. 2017. *How Vocabulary is Learned*. Oxford: Oxford University Press.
Webb, S. and M. P. Rodgers. 2009. The lexical coverage of movies. *Applied Linguistics* 30(3), 407-427.
Webb, S., Y. Sasao and O. Balance. 2017. The updated vocabulary levels test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics* 168(1), 33-69.
Webb, S., T. Uchihara and A. Yanagisawa. 2023. How effective is second language incidental vocabulary learning? A Meta-analysis. *Language Teaching* 56(2), 161-180.
Weber, A. and O. Scharenborg. 2012. Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science* 3(3), 387-401.
Wordoid. n.d. *Wordoid.* Accessed 1 December 2025 Available online at https://wordoid.com
Ye, Z. 2024. The effects of lexical densities and task types on Chinese students' incidental vocabulary acquisition. *Journal of Language Teaching* 4(3), 1-10.

Examples in: English
Applicable Languages: English
Applicable Level: Secondary, Tertiary

Appendix. Example Questions of Vocabulary Test

1. Form recognition test:
The participant sees on paper:
1. ☐ A ☐ B ☐ C ☐ D ☐ I don't remember any of these.
At the same time the participant hears on the recording:
'Number one [1 sec.] A, glanch [1 sec.] B, preave [1 sec.] C, smarge [1 sec.], D, shroke [2 sec.]'

2. Part-of-speech recognition test:
The participant sees on paper:

1.	**glanch**	noun	☐ The **glanch** is good.	☐ I don't know.
		verb	☐ I will **glanch** quickly.	
		adjective	☐ She looks very **glanch**.	

At the same time the participant hears on the recording:
'Number one [1 sec.] glanch [6 sec.]'

3. Meaning recall test:

1. The **scroove** was quiet.	_____	☐ I don't know.
2. I will **pring** later.	_____	☐ I don't know.
3. I feel **crong** today.	_____	☐ I don't know.
4. He began to **trisk** carefully.	_____	☐ I don't know.
5. He enjoys the **slamp**.	_____	☐ I don't know.
6. The **varm** was large.	_____	☐ I don't know.

