



A Corpus-Based Analysis of *Take-Noun* Collocations in Essays by Korean University EFL Students

Jung-Yeo Chung (Ewha Womans University)
Eun-Joo Lee (Ewha Womans University)

Received: 25 February 2024
Revised: 30 March 2024
Accepted: April 4 2024

Chung, Jung-Yeo & Lee, Eun-Joo. (2024). A corpus-based analysis of *take-noun* collocations in essays by Korean university EFL students. *Modern English Education*, 25, 85-99.

Keywords

learner corpus, collocation, verb-noun collocation, delexical verbs, *take-noun* collocation
학습자 말뭉치, 연어, 동사-명사 연어, 경동사, take-명사 연어절

Jung-Yeo Chung (First author)
PhD Candidate
Department of English Education
Ewha Womans University
jungyeo77@gmail.com

Eun-Joo Lee (Corresponding author)
Professor
Department of English Education
Ewha Womans University
eunlee@ewha.ac.kr

Abstract

This study investigated the production of *take-noun* collocations in essay writings by Korean learners of English as a foreign language. A mixed-methods approach was employed, combining quantitative methods such as chi-square tests in R for collocation frequency analysis with qualitative methods. The qualitative analysis was constructed using Chang's (2018) framework to investigate how both interlingual and intralingual factors influence deviant collocations across various proficiency levels. The study findings showed a significant difference in the use of *take-noun* and delexical *take-noun* collocations among the three proficiency levels. The high-level group used these collocations the most, followed by the low and intermediate groups, with *care* being the most frequent collocate in all groups. Unlike earlier studies that highlighted L1 transfer, the present study identified a significant role of L2 factors in the misuse of *take*, such as replacing other delexical verbs (e.g., *take a meal* not *have a meal*), the substitution of specific verbs (e.g., *take the earphones* not *carry the earphones*), and limited knowledge of pairing specific nouns with particular verbs (e.g., *take a profit* not *make a profit*), especially among lower-level learners. These findings indicate the importance of L2 knowledge in learning verb-noun collocations for beginning and intermediate learners.

INTRODUCTION

Collocation, defined as a group of words that frequently occur together (Lewis, 2000), is essential in learning the English language. Research has shown that second language (L2) learners often struggle to learn and use collocations accurately, facing more difficulty with verb-noun collocations compared to other types (e.g., Altenburg & Granger, 2001; Laufer & Waldman, 2011; Nesselhauf, 2003, 2005; Puimège & Peters, 2019, 2020). Moreover, learners from various first language (L1) backgrounds have struggled in pairing high-frequency delexical verbs (e.g., *do*, *get*, *give*, *have*, *make*, and *take*) with suitable nouns (Juknevičienė, 2008; Kittigodin & Phoocharoenail, 2015; Sawaguchi & Mizumoto, 2022; Wang, 2016; Wang & Chen, 2023).

Korean learners of English as a Foreign Language (EFL) also frequently encounter challenges with the correct usage of high-frequency delexical verbs, as noted in various studies (Choi, 2019; Hong, 2021; Kim, 2002; Lee & Na, 2015; Ma &

Kim, 2013; M. Park, 2020). Many researchers have focused on delexical verb-noun collocations within the Korean EFL context, particularly their inclusion in educational materials. These studies have investigated the distribution and prevalence of such collocations in school textbooks (Choi & Chon, 2012; Kang, 2014; Kwon, 2012), English Language Teaching (ELT) textbooks (Jang, 2008), and College Scholastic Ability Test (CSAT) English textbooks (Y. Lee, 2020), comparing their frequency and types to those found in native speaker corpora. In terms of how Korean EFL learners produce delexical verb-noun collocations, various analyses have been conducted comparing learner corpora with native corpora (Choi, 2019; Hong, 2021; Kim, 2002; Lee & Na, 2015; M. Park, 2020). For instance, Kim (2002) and Lee and Na (2015) explored the usage of *make*, while Hong (2021) examined the grammatical structures of delexical verb-noun collocations used by both learners and native speakers. Choi (2019) identified constructional differences, attributing learners' collocational errors to both interlingual (L1 transfer) and intralingual (L2 influence) factors. M. Park (2020) incorporated a comparative analysis that included a judgement test for students to select appropriate delexical verb-noun collocations. Except for Choi's study (2019), most research has not divided or classified writers according to their proficiency levels using specific benchmarks like the Common European Framework of Reference for Languages (CEFR).

Despite extensive research on high-frequency delexical verb-noun collocations, there remains a gap in understanding how Korean EFL learners construct these collocations at different proficiency levels and how these levels interact with L1 and L2 influences to produce deviant constructions. This study aims to investigate the production of *take*-noun collocations among Korean English learners, examining the role of interlingual and intralingual factors in collocational errors across three levels of proficiency through in-depth quantitative and qualitative analysis.

The choice of *take*-noun collocations is motivated by the verb *take* being one of the most frequently used verbs by native English speakers (Nation, 2001) and its familiarity among L2 English learners compared to other delexical verbs across all proficiency levels (Kittigossin & Phoocharoensil, 2015). Unlike other common verbs such as *make*, *give*, and *have*, *take* has shown a significant correlation with learners' proficiency levels (M. Park, 2020), and its usage ratio by Korean EFL learners closely aligns with that of native speakers (Hong, 2021). Yet, despite its importance, *take* has received relatively less attention in research compared to *make*, as highlighted by previous studies (e.g., Altenburg & Granger, 2001; Kim, 2002; Lee & Na, 2015; Sawaguchi & Mizumoto, 2022).

LITERATURE REVIEW

Collocation

Collocations have been explored from diverse perspectives, particularly through corpus-based statistical and semantic-based phraseological approaches (Shi et al., 2023). The former approach highlights the frequency with which words co-occur (e.g., Vilkaitė & Schmitt, 2019), whereas the latter delves into the semantic connections between words within a collocation (e.g., Cowie, 1994; Nesselhauf, 2003, 2005). These viewpoints often converge in definitions of collocations, such as the one provided by Nation (2001), who described collocations as “items which frequently occur together and exhibit a certain level of semantic unpredictability” (p. 317). The present study adopts a view of collocations as groups of words that not only frequently co-occur but also share semantic or syntactic links (Nesselhauf, 2003, 2005).

Considering the extent of semantic or syntactic connections between words, Cowie (1994) presented a “collocational continuum model”, which classifies word combinations into four distinct categories: free combinations, restricted collocations, figurative idioms, and pure idioms. Free combinations permit words to be paired freely following structural rules (e.g., *like a candy*), while pure idioms are fixed expressions whose meanings are not directly deducible from the constituent words (e.g., *hit the road*). Restricted collocations, such as *take an action*, involve one component (e.g., *take*) being used in a figurative or delexicalized sense, while the other component (e.g., *action*) maintains its literal meaning. Figurative idioms, exemplified by *pay the price*, can be understood either figuratively as a whole or literally, based on the meanings of the individual words (Shi et al., 2023). This framework situates collocations on a spectrum ranging from freely formed expressions to fixed idiomatic phrases, with varying degrees of semantic or syntactic linkage (Laufer & Waldman, 2011).

In this study, the verb *take* is combined with specific nouns, with *take* potentially assuming literal, figurative, or delexicalized connotations. Therefore, *take*-noun collocations in this analysis are positioned between free combinations and pure idioms, tending more towards restricted collocations.

Verb-noun Collocations

Research has shown that L2 learners are more inclined to produce verb-noun collocations than adjective-noun collocations, even though both types are frequently used by these learners (Altenburg & Granger, 2001; Chang, 2018; Laufer & Waldman,

2011; Nesselhauf, 2003, 2005; Puimège & Peters, 2019, 2020). Altenburg and Granger (2001) explored the use of *make* in corpora consisting of French and Swedish learners, as well as English native speakers, and found that challenges with this verb persisted even among advanced learners. Echoing this, Nesselhauf (2003, 2005) analyzed essays from the International Corpus of Learner English (ICLE) and observed that German advanced learners struggled to select the correct verb-noun collocations, especially when their native language's collocational patterns were misaligned with English.

Laufer and Waldman (2011) also highlighted the difficulty of acquiring verb-noun collocations for L1 Hebrew learners of English across three proficiency levels. Their findings indicated that L2 learners, regardless of their level, produced fewer collocations than native speakers, with many of these being atypical. Interestingly, they noted that the frequency of errors increased with proficiency, suggesting that L1 interference continues to affect even advanced level L2 learners without a clear link between proficiency level and the production of unusual collocations that could be influenced by L1. In a similar vein, Puimège and Peters (2019, 2020) discovered that EFL learners tend to learn fewer English verb-noun collocations than single words or adjective-noun collocations, as observed in a study on learning collocations from television. These studies consistently highlight the profound influence of L1 on the usage of non-standard collocations. Chang (2018) further explored how Korean L1 learners of English affects the use of unconventional collocations on six types of lexical pairings including verb-noun collocations in a comparative case study. The study identified several factors influencing these deviations, including direct L1 to L2 translation (e.g., “*do a role*” instead of “*play a role*”), discrepancies between what learners mean in their L1 and how they express these meanings using L2 collocations derived from L1 (e.g., “*lead a trend*” instead of “*set a trend*”), and the rephrasing of intended meanings by learners according to the conceptual structures of their L1 (e.g., “**use this time without good reason*” instead of “*waste time*”).

Regarding the syntactic structure of verb-noun collocations, these pairings have been analyzed irrespective of whether the noun functions as an object, complement, or adverbial within the clause, or of the presence of other elements (Nesselhauf, 2005). Verb-noun collocations encompass patterns where the noun may be pre- or post-modified (Nesselhauf, 2003); e.g., *take on a role* or *take care of*. In line with Nesselhauf's classification, this study will consider not only direct object nouns but also those preceded or followed by prepositions and more complex verb-noun structure such as *take defect/health into consideration* to comprehensively identify *take*-noun collocations.

Delexical Verb-noun Collocations

L2 learners from various linguistic backgrounds often encounter difficulties in correctly using high-frequency verbs known as “delexical verbs” or “light verbs” (e.g., *do*, *get*, *give*, *have*, *make*, and *take*) with appropriate nouns (Juknevičienė, 2008; Kittigossin & Phoocharoensil, 2015; Sawaguchi & Mizumoto, 2022; Wang, 2016; Wang & Chen, 2023). These verbs carry minimal meaning on their own and derive their semantic content primarily from the nouns they accompany (Sinclair, 1991). Nesselhauf (2005) described delexical verb-noun collocations as “stretched-verb constructions”, which lack independent verb meanings within the verb-noun structure. In these constructions, the noun, functioning as an object, conveys the core meaning, as seen in phrases such as *I took a shower*, *she made a mistake*, *let's have a look at it*, *I got some information about it*. Such instances are categorized as restricted collocations, where the noun is uniquely paired with a specific delexical verb, precluding its combination with other delexical verbs.

Research has consistently shown that L2 learners, irrespective of their L1 backgrounds, tend to produce atypical delexical verb-noun collocations due to interlingual or intralingual factors. Juknevičienė (2008) utilized a contrastive corpus analysis to reveal that Lithuanian English learners significantly underutilized typical academic register collocations and delexical verb-noun combinations, often resorting to their L1 constructions, which result in unconventional English collocations. Kittigossin and Phoocharoensil (2015) explored the learning strategies leading to irregular delexical-verb usage among Thai EFL learners of varying proficiency levels through gap-filling translation tests and interviews. They identified three strategies—native language transfer, synonymy, and overgeneralization—that contribute to the unusual use of English delexical verb-noun collocations, noting that *take* was the most recognized delexical verb among both high and low proficiency learners.

Sawaguchi and Mizumoto (2022) investigated the extent of L1 transfer effects on the misuse of *make*-noun collocations among Japanese EFL learners across different proficiency levels. Although their study did not find significant differences in the misuse of *make* collocates among the proficiency groups, the result suggested that collocational knowledge might expand with increasing proficiency. Wang (2016) compared the use of delexical verbs by Chinese and Swedish English learners to that of native speakers, examining the influence of L1 on target language acquisition through a comparison of L1, interlanguage (L2), and native speaker corpus data. The findings indicated that both learner groups overused delexical collocations but with less variety than native speakers. Wang and Chen (2023) analyzed verb-noun collocation errors using the Chinese Learner English Corpus, identifying four main error types: incorrect translations from Chinese, misuse of

semantically similar words, delexical verb errors, and part of speech errors. They observed that while translation and synonym-related errors decrease with improved proficiency, advanced learners do not necessarily enhance their delexical verb collocational or synonym usage. Across these studies, a recurrent theme is the production of unconventional delexical verb-noun collocations due to L1 factors such as direct translation from L1.

Difficulties of Korean EFL Learners with Delexical Verb-noun Collocations

Numerous studies have underscored the significant challenges that Korean EFL learners face when trying to grasp the nuances of high-frequency delexical verbs such as *do*, *get*, *give*, *have*, *make*, and *take* (Choi, 2019; Hong, 2021; Kim, 2002; Lee & Na, 2015; Ma & Kim, 2013; M. Park, 2020). While much previous research has predominantly focused on the distribution and usage of these collocations within educational materials like school and ELT textbooks, and CSAT English textbooks (Choi & Chon, 2012; Jang, 2008; Kang, 2014; Kwon, 2012; Y. Lee, 2020), comparisons with native speaker corpora such as the Louvain Corpus of Native English Essays (LOCNESS), the British National Corpus (BNC), and the Corpus of Contemporary American English (COCA) (Davies, 2008-present) reveal a notable discrepancy. Korean textbooks, despite featuring delexical verbs as frequently as native corpora, tend to offer less variety and fewer contextual usage examples. This discrepancy suggests a potential lack of exposure to the diverse contexts in which these verbs are used in natural English, potentially limiting learners' ability to fully internalize and use these verbs effectively in their own language production.

Investigations into Korean EFL learners' production of verb-noun collocations have been conducted mostly in comparative analysis using learner corpora and native corpora. Notably, two studies focused solely on the verb *make* (Kim, 2002; Lee & Na, 2015). Kim (2002) analyzed college students' writings, comparing their usage with that of French and Swedish learners and native English speakers as studied by Altenburg and Granger (2001). The results showed that Korean learners, at high-beginner to intermediate levels, tended to overuse *make* in its productive and causative senses, often overlooking its delexical and idiomatic uses. Lee and Na's study, which compared Korean students' TOEFL essay writings with American students' writings from the LOCNESS corpus, confirmed these tendencies. The Korean students in this latter study were considered at intermediate-high to advanced proficiency levels by the authors. Ma and Kim (2013) assessed Korean high school students' understanding of delexical verb collocations and their prevalence in *English II* textbooks, uncovering a significant gap between the students' self-assessed knowledge and their actual proficiency, as evidenced by a 38% average score on a relevant test.

Expanding the scope, Choi (2019) examined the construction of four delexical verb-noun collocations (i.e., *make*, *take*, *give*, *get*) among Korean learners and native speakers using the Yonsei English Learner Corpus (YELC) and LOCNESS, investigating how interlingual and intralingual factors influence collocational errors. Despite noting an overuse of high-frequency verbs across three proficiency levels, Choi (2019) did not find significant differences in the usage of delexical collocations across proficiency levels. Moreover, the qualitative analysis for identifying the sources of deviant collocation was primarily confined to lower proficiency data, leaving the impact of L1 and L2 factors on higher proficiency learners ambiguous. M. Park (2020) also investigated the use of these four delexical verbs in Choi's (2019) study by examining essays from Korean EFL college students (KOCOE) and the LOCNESS, including a "phrase preference test" to discern Korean college students' inclination towards delexical-verb structures or single-verb structures. The findings indicated a notable difference in the usage of the four delexical verbs between the two corpora. Interestingly, while students generally favored single-verb over delexical verb constructions regardless of proficiency, the use of delexical *take* collocations was significantly associated with proficiency levels; higher-level students used them more than the lower-level group. Hong (2021) analyzed the use of *give*, *make*, and *take* among Korean English learners compared to native English speakers, drawing on data from the ICLE and LOCNESS corpora. The findings revealed that, although Korean learners underutilized *give* and *make*, their usage of *take* was comparable to that of native speakers. Additionally, there was a marked avoidance of passive constructions (be + p. p.), especially with *make*, which is likely influenced by L1 characteristics, given that passive constructions are rarely used in Korean.

Chang (2018), along with Choi (2019) and Hong (2021), explored the challenges Korean learners face with English delexical verb-noun collocations. Chang (2018) pointed out three main interlingual factors causing difficulties. Firstly, there is the inappropriate substitution of delexical verbs, such as using *do* instead of *make* in *do a promise*. This error highlights the nuanced differences between verbs that might seem interchangeable. Secondly, the misuse of delexical verbs where specific verbs are required, exemplified by *take success* instead of *achieve success*, indicates a struggle with selecting verbs that accurately convey the intended meaning. Lastly, the need for precise verbs with certain nouns, as exemplified by preferring *acquire knowledge* over *get knowledge*, underscores the importance of verb choice in facilitating effective communication. These findings shed light on the subtle yet significant challenges encountered by Korean learners in

mastering delexical verb-noun collocations in English.

The Present Study

Despite extensive research on high-frequency delexical verb-noun collocations, few studies have explored how Korean EFL learners construct these collocations across different proficiency levels, or how proficiency levels affect the influence of L1 and L2 in producing deviant constructions (e.g., Chang, 2018; Hong, 2021; Kim, 2002; Lee & Na, 2015). This study aims to investigate the production of *take*-noun collocations among Korean English learners at low, intermediate, and high proficiency levels, examining the role of interlingual and intralingual factors in collocational errors across three levels of proficiency through detailed quantitative and qualitative analysis.

This study focuses on *take*-noun collocations because *take* is one of the highest frequency verbs by native English speakers (Nation, 2001) and the most recognizable lexical verb among L2 English learners across all proficiency levels (Kittigossin & Phoocharoensil, 2015), in addition, *take* uniquely correlates with learner proficiency, setting it apart from verbs such as *make*, *give*, and *have* (M. Park, 2020) and its usage by Korean EFL learners shows no notable difference from that of native speakers (Hong, 2021). However, research on *take* is less extensive compared to *make*, despite its significance, as evidenced by prior research (Altenburg & Granger, 2001; Kim, 2002; Lee & Na, 2015; Sawaguchi & Mizumoto, 2022).

The current study categorizes proficiency levels more distinctly based on TOEIC scores within the same CEFR bands: A2, B1, and B2-C1. This nuanced approach aims to highlight the specific characteristics of each proficiency level. Given the broad range within CEFR bands (e.g., A2: TOEIC 225–545, B1: 550–780, B2: 785–940, C1: 945–990), collecting data near the score boundaries (e.g., 540 and 550, 770 and 785) might affect the results, such as making it challenging to identify distinct features between levels (e.g., Choi, 2019; Sawaguchi & Masumoto, 2022). Furthermore, this study will illuminate the performance of novice beginners in using verb-noun and delexical verb-noun constructions, comparing it with the performance of learners at other proficiency levels. This focus on the lower end of the proficiency spectrum (TOEIC scores 225–300) addresses a gap in the literature on verb-noun collocations. To fill these gaps, the present study addresses the following research questions:

1. What are the most frequent *take*-noun collocations in English essay writings by EFL Korean college students at low, intermediate, and high proficiency levels?
2. To what extent do L2 proficiency levels influence the use of delexical *take*-noun collocations in English essay writings of EFL Korean college students?
3. To what extent can deviant delexical *take* + noun collocations be explained by inter-language (L1 influence) or intra-language (L2 influence) factors, and is there a variance in the error ratio across the three proficiency levels?

METHOD

Corpus

The Gacheon Learner Corpus (GLC), serving as the primary data source for this study, explores the usage of *take*-noun collocations among Korean EFL learners. Compiled between 2012 and 2014 at Gacheon University in Korea, the GLC amasses a total of 2.5 million words, marking it as the largest corpus of its kind in Korea to date. Participants were undergraduate students who spoke Korean as their L1 and were enrolled in the same mandatory English course, covering a variety of majors. Their essays, which covered various topics, were responses to class assignments written throughout the course. Participants were tasked with selecting one out of twenty diverse topics, spanning genres such as argumentative, narrative, and expository; e.g., “What topics should people avoid during small talk? Why?”, “Have you ever engaged in an activity that was particularly exhilarating or terrifying? Describe your experience.”, “When choosing a hotel, what factors are most critical to you and why?”. They then articulated their thoughts in essays ranging from 100 to 150 words via Google Forms. In 2015, the corpus was finalized, encompassing 25,073 texts authored by 2,500 participants. The participants primarily fell within the low to intermediate proficiency levels, according to Yoo and Shin (2022), with only 369 texts contributed by learners at or above the B2 level of the CEFR.

To conduct a comprehensive analysis of the subtle linguistic features typical of various proficiency levels, this study organizes learners into three groups: low, intermediate, and high proficiency. These groups are delineated based on specific TOEIC score ranges, which correspond to CEFR bands. The defined ranges and their associated text counts are as follows: A2 (230–300) with 442 texts, B1 (550) with 771 texts, and B2-C1 (785–945) with 369 texts. This classification scheme is

notable for including the lower end of the proficiency spectrum (TOEIC scores 230–300), an area often overlooked in previous research. Due to the limited number of texts from high-level students (only 369), to ensure balance, the intermediate level was represented solely by texts from a single score band, specifically TOEIC 550. Originally, there were 774 texts at the TOEIC 550 level, but this number was adjusted to 771 after removing two duplicates and one text written in Korean. Table 1 outlines the three sub-corpora, each corresponding to a distinct proficiency level.

TABLE 1

Description of The Three Selected Sub-Corpora from GLC

Proficiency level	Number of essays	Type	Token
Low (A2, TOEIC 230-300)	442	8,114	46,907
Intermediate (B1, TOEIC 550)	771	12,904	81,787
High (B2-C1, TOEIC 785-945)	369	7,211	43,617
Total	1,582	28,229	172,311

Data Analysis Procedure

In order to analyze *take*-noun collocations within the GLC, AntConc (4.2.4) is employed, leveraging its N-gram and KWIC (Key Word In Context) functions. The verb *take* was lemmatized to include its variations such as *takes*, *taking*, *took* and *taken*. Similarly, both plural and singular forms of nouns are lemmatized (e.g., *change* and *changes* are lemmatized as *change*). Consistent with Nesselhauf's (2005) classification, *take*-noun collocations are identified, including those where nouns are accompanied by prepositions (e.g., *take care of*, *take into consideration*). As discussed in Nesselhauf's (2005) study, "even more idiomatic expressions such as *take place* were analyzed as verb + noun combinations" (p.47). Therefore, expressions like *take place* are categorized within *take*-noun combinations. However, free combinations involving proper nouns, as illustrated by sentences such as *Prince Hans tracked her back to take Elsa to the frozen world*, are excluded from this analysis.

In categorizing these collocations, nouns of similar meanings are grouped under a single lemma; for example, *a class*, *a design class*, and *English class* are all counted classified under *class*, and both *care* and *personal care* are aggregated under *care*. The phrase *take care* in both *take care of* and *take care* contexts is grouped under the *take-care* collocation, though they are semantically distinct and analyzed separately to pinpoint deviant uses. Pronouns referring to earlier mentioned nouns are also considered as noun collocates of *take*; for instance, in *I would tell the cashier that the change is too much. If I take it, the cashier would make great losses*, *it* refers back to *change* and is counted as a collocate of *take*. The analysis extends to identifying the subjects and objects associated with *taken*, as in *cosmetics should never be taken lightly*.

This study focuses on the use and potential misuse of *take*-noun collocations, incorrect grammatical usage unrelated to these combinations, such as spelling mistake, misuse of the base form in place of the participle, voice errors, are overlooked. For instance, the sentence *People usually want to wear pretty clothes when take a picture at their tourist attraction* is included for analysis despite its grammatical inaccuracies (an error in voice). Across three proficiency levels, a total of 195 instances of *take*-noun collocations were identified, among which six sentences contained ungrammatical usages. These comprised one spelling mistake. Notably, this instance also involved incorrect use of *take* instead of *taking* (*people run away take motorcycle, even airport staff). There were two additional instances where *take* was used incorrectly instead of *taking*, and three errors in voice. The raw frequencies of *take*-noun and delexical *take*-noun collocations across the three sub-corpora are normalized to an adjusted frequency per 1,000,000 words. This normalization enables comparison and follows the method used in Wang's (2016) study, facilitating the recognition of the normalized frequencies (NFs) with larger, more intuitive whole numbers, which might otherwise appear as smaller, less intuitive decimal numbers. The NFs of these collocations at three proficiency levels are then analyzed using chi-square tests in R version 4.3.2 to identify significant differences among the levels. An initial chi-square test is conducted to identify statistical differences among the three groups, followed by pairwise chi-square tests to pinpoint significant differences between each pair of groups, resulting in three sets of pairwise comparisons. The Bonferroni correction method is applied to each set to reduce the risk of Type I errors.

To identify delexical *take*-noun collocations, this study utilizes Wang's (2016) comprehensive list of noun collocates for high-frequency verbs from the British National Corpus (BNC), and cross-references with the Online Longman Dictionary of Contemporary English, Cambridge Online Dictionary, and Oxford Online Collocation Dictionary. Wang's list, which offers a broader range of delexical verb-noun collocations than any single dictionary, reduces the chance of including irrelevant collocates. It features 400 noun collocates for each verb, organized according to their frequency in both present and past tense forms, and ranked by log-likelihood (LL) and mutual information (MI) scores. Notably, many nouns recur

across both verb tenses. Consequently, akin to Choi (2019), the current study focuses on the distinct types of noun collocates, utilizing approximately 200 collocates for analysis

Collocations that are not listed in Wang’s (2016) compilation or the aforementioned dictionaries, and that have fewer than 100 occurrences in COCA—criteria used by Sawaguchi and Mazumoto (2022) to identify deviant *make*-noun collocations—are classified as deviant. Additionally, this study will identify collocational errors by examining the top 10 noun collocates of *take* from COCA, using a window size of 4 words on both the left and right sides of *take*. This approach aims to investigate their verb collocates to ascertain potential misuses in lieu of *take*. Table 2 presents the top 10 noun collocates of *take* in COCA.

TABLE 2

The Top 10 Noun Collocates of “Take” in COCA (Retrieved on January 20th, 2024)

Noun Collocates	Frequency	Mutual Information (MI)
Care	67926	4.33
Place	60846	3.07
Look	36545	4.00
Step	29402	3.77
Advantage	22634	4.53
Action	21500	2.93
Picture	16231	2.78
Break	13537	3.55
Account	12310	3.09
Breath	12040	4.01

Note. MI presents the level of co-occurrence strength between the node and its collocates.

The influence of interlingual and intralingual factors on producing deviant *take*-noun collocations will be examined based on the classifications of L1 and L2 factors that Chang (2018) provided as the sources of deviant lexical collocations that Korean adult learners of English: L2 factors—inappropriate substitution of other delexical verbs, the misuse of delexical verbs where specific verbs are required, and lack of knowledge of certain nouns needed for specific verb or other delexical verb; L1 factors—direct translation from L1, differences between the intended meanings in L1 and the semantic representations of L2 combinations translated from L1, and learners’ paraphrasing of intended meanings based on L1 semantic frameworks.

FINDINGS AND DISCUSSION

The Most Frequent *Take*-noun Collocations at Three Proficiency Levels

The following table summarizes the frequency of *take*-noun collocations across three proficiency groups. The Normalized Frequency in Table 3 indicates that the *take*-noun collocations are most frequently used by the high proficiency group, followed by the low and intermediate groups, respectively.

TABLE 3

The Frequency of “Take-noun” Collocations by Three Proficiency Groups

	Low	Intermediate	High
Type	39	43	45
Token	57	68	70
Normalized Frequency	1215*	831*	1605*

Note. An asterisk (*) denotes statistical significance at $p < .001$. The alpha level for the Chi-square pairwise tests is set at $p < .0167$, applying the Bonferroni correction method. This adjustment is made to mitigate the risk of Type I errors, considering that three sets of pairwise comparisons are conducted.

Three Chi-square pairwise tests show that there are significant differences between groups in using *take*-noun collocations: Low and Intermediate ($\chi^2=72.07, p < .001$), Low and High ($\chi^2=53.936, p < .001$), and Intermediate and High ($\chi^2=245.93, p < .001$). In terms of the variety of collocational types, the low proficiency group exhibits the smallest range, while the high proficiency group demonstrates more diverse usage. However, caution is warranted in interpreting these results. Notably, the intermediate group has a uniform TOEIC score of 550, in contrast to the other groups, which comprise a wider range of scores. This uniformity within the intermediate group might account for the differences in collocational frequency observed.

Table 4 presents the most frequent *take*-noun collocations across three proficiency groups, focusing on nouns that are paired with *take* more than twice. ‘Frequency’ in the table refers to the number of occurrences of each noun collocate within each proficiency level, and ‘percentage’ is calculated based on the raw frequency of tokens of *take*-noun collocations within each of the three groups. The total types of noun collocates of *take* are detailed in Appendix A. *Care* emerges as the most frequently used noun collocate across all groups, consistent with findings from COCA. *Picture* also ranks high in frequency across the three groups. Among the top 10 noun collocates of *take* in COCA, seven nouns — namely *care, place, look, advantage, action, picture, break* — are present in the three sub-corpora of this study while *step, account, and breath* are absent. The prevalent use of *care* with *take*, especially among the top 10 collocates in COCA, may be influenced by a specific Korean educational context. Korean school textbooks, while including delexical verbs as often as native English materials, tend to offer less variety and fewer contextual usage examples. This pattern has been noted in prior research (Choi & Chon, 2012; Jang, 2008; Kang, 2014; Kwon, 2012; Y. Lee, 2020).

TABLE 4

The Most Frequent Noun Collocates of “Take” at the Three Proficiency Groups

Proficiency level	Noun	Frequency (%)
Low	Care	9 (15.97)
	Picture	4 (7.08)
	Rest	3 (5.26)
	Exercise	3 (5.26)
	Haircut	2 (3.51)
	Action	2 (3.51)
	Airplane	2 (3.51)
	Car	2 (3.51)
	Time	2 (3.51)
Intermediate	Care	12 (17.65)
	Bus	5 (8.8)
	Picture	4 (5.88)
	Photo	3 (4.41)
	Look	2 (2.94)
	Walk	2 (2.94)
	Computer	2 (2.94)
Drink	2 (2.94)	
High	Care	9 (12.86)
	Change	8 (11.43)
	Time	3 (4.29)
	Class	3 (4.29)
	Picture	3 (4.29)
	Test	2 (2.86)
	Boat	2 (2.86)
	Information	2 (2.86)
Bus	2 (2.86)	

While the prevalence of other noun collocates is generally similar, *bus* in the intermediate group and *change* in the high proficiency group stand out for their higher frequency. This could be due to the specific essay prompts more commonly used in these groups. In other words, this higher frequency of *bus* and *change* may be attributed to the specific essay prompts

used more frequently in these groups. Although this study focuses on *take*-noun and delexical *take*-noun collocations, it does not differentiate essay prompts by proficiency level. Nonetheless, it suggests that different essay genres in GLC might influence the choice of specific noun collocates (J. Lee, 2021).

The Use of Delexical *Take*-noun Collocation at Three Proficiency Levels

The frequency of delexical *take*-noun collocations across three proficiency levels is illustrated at Table 5. Normalized Frequency data show that the high group uses delexical *take*-noun collocations notably more frequently than the other two groups (low-831, intermediate-685, high-1284).

TABLE 5
Frequency of Delexical “Take-noun” Collocation at the Three Proficiency Groups

	Low	Intermediate	High
Type	23	32	32
Token	39	56	56
Normalized Frequency	831*	685*	1284*
Ratio to total collocation	68.42%	82.35%	80.00%

Note. An asterisk (*) denotes statistical significance at $p < .001$. The alpha level for the Chi-square pairwise tests is set at $p < .0167$, applying the Bonferroni correction method. This adjustment is made to mitigate the risk of Type I errors, considering that three sets of pairwise comparisons are conducted.

Three Chi-square pairwise tests show that there are significant differences between groups in using delexical *take*-noun collocations: Low and Intermediate ($\chi^2=14.061, p < .001$), Low and High ($\chi^2=97.026, p < .001$), and Intermediate and High ($\chi^2=182.22, p < .001$), while intermediate and high group use the same number of the types of delexical *take*-noun construction. The ratio to total *take*-noun collocation reveals a significant portion of *take* usage as delexical across all groups. The intermediate group exhibits the highest usage of delexical *take* at 82.35%, followed by the high group at 80.00%, and the low group at 68.42%.

The significant differences in using delexical *take*-noun between groups suggests that the increase of proficiency level promotes using delexical *take*-noun collocations, which is in line with the findings of M. Park (2020) and Sawaguchi and Mizumoto (2022). Contrary to previous research indicating that L2 learners often overuse high-frequency verbs but underutilize them as delexical verbs (Kim, 2002; Kittigodin & Phoocharoensil, 2015; Hong, 2021), the study’s findings show a considerable use of delexical *take* across proficiency levels. This may suggest that learners, regardless of proficiency, are relatively familiar with the delexical application of *take*, echoing the insights from Kittigodin and Phoocharoensil (2015). This observation could be attributed to the extensive range of noun collocates of *take* in Wang’s (2016) list, which was employed to identify delexical *take*-noun collocations in this research. Moreover, the prevalent use of specific nouns like *as care, picture, change, bus* may have increased the proportion of delexical verbs. A complete list of delexical *take*-noun collocations across the three proficiency levels can be found in Appendix B.

The Influence of L1 and L2 factors on Deviant Delexical *Take*-noun Collocation

The frequency of deviant delexical *take*-noun collocations and their ratio to the total number of *take*-noun collocations is presented at Table 6.

TABLE 6
Frequency of Deviant Delexical “Take-noun” Collocations

Proficiency level	Frequency	The ratio to the total tokens
Low	13	22.80%
Intermediate	12	17.65%
High	12	17.14%

The data show that the low proficiency group produces the most deviant construction the most, followed by the intermediate group, with the high proficiency group producing the least although the proportions for the intermediate and high groups are similar. This result suggests that writers in the low proficiency group face greater challenges in constructing correct *take*-noun collocations. This finding is not surprising, considering they are at the lower end of the proficiency spectrum.

TABLE 7
Frequency of L1 and L2 Factors on Deviant Delexical “Take-noun” Collocations

Proficiency level	L1(%)	L2(%)
Low	1 (7.69%)	12 (92.31%)
Intermediate	1 (8.33%)	11 (91.67%)
High	1 (8.33%)	11 (91.67%)

Table 7 illustrates the proportion of L1 and L2 factors influencing the production of deviant delexical *take*-noun collocations. The analysis reveals a significant impact of L2 factors on deviant use of delexical *take*-noun collocations while the influence of L1 transfer appears to be minimal, with only one identified case across three proficiency levels. It is worth noting that certain atypical collocations might be attributed to both L1 and L2 factors. For instance, collocations such as *take the bus*, **take the motorcycle*, and **take a Xiyro Swing* could be influenced by a lack of L2 vocabulary knowledge in using *take* as well as the direct L1 translation of *take* into “ta-da” in Korean. This translation is combined with words for bus, motorcycle, or amusement park rides in Korean (e.g., “beo-seu-leul ta-da,” “o-to-ba-i-leul ta-da,” “nol-i-gi-gu-leul ta-da”). In this study, these instances are categorized under L2 factors because they exemplify the overgeneralization of *take* in the context of “traveling using a vehicle” (e.g., *take the bus*, *take the train*). All misuse of *take*-noun collocation across three proficiency levels are listed in Appendix C. Among L2 factors, *take* is frequently misused instead of other delexical verbs as illustrated in Examples (1)–(5):

- (1) *I love hobbit character in the movie who take five meals a day. (< have a five meals)
- (2) *They shouldn’t think that the driver who has taken a drink (had a drink)
- (3) *I think the best way to improve his or her appearance is take a plastic surgery.
(< get a plastic surgery)
- (4) *And if they are take a massage, they are more active (< get a massage)
- (5) *I don’t take Manicures, Pedicures or Facials. (< get Manicures, Pedicures)

Examples (1)–(5) demonstrate that writers choose other delexical verbs such as *have* and *get* instead of *take*, indicating confusion in selecting the appropriate delexical verb. This tendency to confuse delexical verbs has been documented in previous studies (Chang, 2018; Choi, 2019).

Take is also incorrectly used in place of specific (heavy) verbs with distinct meanings, rather than as a delexical verb, as illustrated in Examples (6)–(11):

- (6) *I want to take a relax with awesome places. (< relax)
- (7) *I regularly take a shampoo and conditioner. (< use a shampoo)
- (8) *I always take my earphones everywhere (< carry my earphones)
- (9) *It was so nice to go down the hill taking a snowboard. (< snowboarding)
- (10) *Then I take a light airplane and rose into the sky. (< get on/board a light airplane)
- (11) * When I take the bus, the car behind the bus very surprise me. (get < on the bus)

Example (6) could be explained as an example of overgeneralization of delexical *take* as in *take a rest*. Examples (7)–(9) demonstrate that *take* is misused to replace other particular verbs that convey precise meanings within the given context. Example (10) and (11) are the instances that Korean EFL learners are often confused in using *take* as the meaning of “to travel somewhere by using a particular form of transport or a particular vehicle, route.” *Take airplane* or *the bus* means that use ‘airplane’ or ‘bus’ as a vehicle for traveling. However, in these examples, the speaker is actually on the airplane or bus, making the use of *take* inappropriate. These examples show that Korean EFL learners often struggle with selecting the correct verbs, considering the nuanced meanings of *take* in different contexts

Deviant *take*-noun collocations can also arise from limited knowledge of pairing specific nouns with particular verbs, including other delexical verbs, as demonstrated in Examples (12)–(14):

- (12) *people run away take motorcycle, even airfort staff. (< ride/drive a motorcycle)
 (13) *So when i go to the Lotte World, then i never take a Xiyro Swing made me terrible feelings.
 (< go on/ ride a Xiryro Swing)
 (14) * You should take a 'real profit'(< make/ earn a profit)

In Examples (12) and (13), the writers appear to lack knowledge of how to appropriately pair the noun *motorcycle* with specific verbs such as *ride* or *drive*, and how to describe engaging with amusement park rides using phrases such as *go on* or *ride*, rather than incorrectly using *take*. Example (14) suggests that the writer may not be aware that *profit* is typically combined with verbs like *make* or *earn* to convey the concept of “generating earnings,” rather than with *take*.

L1 factors, such as direct translations from L1, play a significant role in the misuse of delexical *take*-noun collocations, as evidenced by Examples (15)–(17). This may stem from differences between the intended meanings in L1 and the semantic representations of L2 combinations translated from L1 (Chang, 2018).

- (15) *I and my friends take room and play there all night.
 (16) *We take a room for tomorrow's sun rise.
 (17) *it reminds me that the fear of scary in that time, too, took it for a few months due to stress gastritis.

In English, *take* often means “to get hold of something with your hands.” Its Korean equivalent, “jab-da,” is used in Examples (15) and (16), where the writers might translate *take a room* literally into Korean as “bang-eul jab-da.” In Korean, “jab-da” can mean “to make a reservation,” a usage not applicable to *take* in English. In Example (17), *take* is used to mean “to have or come to have a particular feeling or opinion.” The Korean equivalents, “ga-ji-da” or “deul-da,” can be paired with feelings like fear, translating into “gong-po-gam-eul ga-ji-da” or “gong-po-gam-i deul-da.” However, in English, *take* is not typically collocated with *fear* but rather with nouns like *offense*, *pity*, or an *interest*, which denote specific feelings.

An analysis of the top 10 noun collocates of *take* from the COCA shows that *look* and *action* are frequently used with *have*, where *take* would be more suitable. This misuse is illustrated in Example (18):

- (18) *Did not have any action. (< take any action)

The findings suggest that even high-level Korean EFL learners find it challenging to use delexical *take*-noun collocations appropriately. This is aligned with the findings of Laufer and Waldman (2011), and Wang and Chen (2023), in which advanced L2 learners did not enhance their usage in combining high-frequency delexical verbs with nouns correctly. Moreover, this analysis shows that there is a minor influence of L1 factors and a significant impact of L2 factors on deviant *take*-noun collocations across all proficiency levels. These findings contrast with those of previous studies that highlighted the crucial role of L1 interference in the deviant use of high frequency delexical verbs (Altenburg & Granger, 2001; Jukneviene, 2008; Sawaguchi & Mizumoto, 2022; Wang, 2016; Wang & Chen, 2023). We should be cautious when interpreting the results, as the proficiency levels in this study are lower than those in previous studies. Previous studies have included participants at advanced or high-intermediate levels, whereas this study includes data from the lower end of A2 band (230–300 on the TOEIC). This segment has been less explored in learner corpus research. The intermediate level is also categorized at the lower end within the B1 band (550 on the TOEIC), and a significant portion of the high-level scores (785–820) falls into the lower range of the B2 band. This suggests that L2 knowledge, particularly at lower proficiency levels, is a crucial factor influencing the learning of collocations, as these learners possess less L2 knowledge than their advanced counterparts.

CONCLUSION

The current study investigates the usage of *take*-noun collocations and delexical *take*-noun collocations among Korean EFL college students, examining the effects of L1 and L2 factors on collocational errors across three proficiency levels. The findings show a significant difference in the use of these collocations among the three proficiency levels. Specifically, the high-level group uses these collocations the most, followed by the low and intermediate groups, respectively. *Care* emerges the most common collocate across all levels, with *picture* also appearing frequently. The intensive pairing of *care* with *take*, particularly among COCA's top 10 collocations, might be influenced by the distinct educational context of Korean education. Korean school textbooks, while including delexical verbs as often as native English materials, tend to offer less variety and fewer contextual usage examples. This trend has been supported by earlier studies (Choi & Chon, 2012; Jang, 2008; Kang,

2014; Kwon, 2012; Lee, Y., 2020). Learners across all proficiency levels significantly utilize delexical *take*-noun collocations, which challenges previous findings that L2 learners often overuse high-frequency verbs but underutilize them in delexical contexts (Hong, 2021; Kim, 2002; Kittigodin & Phoocharoensil, 2015). Notably, learners at lower proficiency levels display a higher ratio of deviant *take*-noun constructions compared to their total usage of *take*-noun constructions, while intermediate and advanced learners show similar usage patterns.

Contrary to previous research that emphasized L1 influence in the misuse of delexical verbs (e.g., Altenburg & Granger, 2001; Choi, 2019; Nesselhauf, 2003, 2005; Wang, 2016), this study reveals a more pronounced impact of L2 factors. This includes the substitution of specific verbs and a limited understanding of which nouns pair well with particular verbs, particularly among lower-level learners. The deviation in collocations may also stem from differences between L1 intended meanings and the semantic representations of L2 combinations translated from L1. This discrepancy may be attributed to the inclusion of novice lower-level learners not typically present in previous research and a much wider range of intermediate learners, potentially highlighting the importance of L2 knowledge in learning collocations and addressing collocational errors among beginners and intermediate learners.

The findings suggest important pedagogical implications, highlighting the necessity for targeted instruction on verb-noun collocations for EFL/ESL learners. The results indicate that novice beginners may be more influenced by L2 knowledge than by L1 factors when learning *take*-noun collocations. In this context, Data-Driven Learning (DDL) method, such as analyzing concordance lines, could be highly effective in improving learners' grasp of context-specific collocation usage. However, given the challenges posed by corpus-based learning, such as time consumption and data overload, more accessible alternatives like teacher-selected paper concordances (Johns, 1994) could serve as a practical option for introducing high-frequency delexical verbs to L2 learners.

Specific activities can be tailored based on learners' proficiency levels to improve their mastery and use of *take*-noun collocations. Novice-beginner learners, who may struggle with acquiring L2 knowledge, could benefit from teacher-selected paper concordances featuring the word *take*. This would aid in grasping the correct usage of *take*-noun collocations within various contexts. Intermediate-level students might also benefit from teacher-selected paper concordances, which could include a broader range of examples than those for beginners, potentially highlighting instances of incorrect collocation usage. Teachers can guide these students to identify such deviant collocations and to understand the contextually correct usage of *take*-noun collocations. Moreover, explaining the reasons behind the incorrect usage by Korean learners could help them recognize and rectify their common mistakes with *take*-noun collocations. In this phase, DDL facilitates both teacher-led instruction and student-centered learning (E. Park, 2019). Lastly, advanced learners could be encouraged to explore *take*-noun collocations using concordance tools like COCA, moving towards more autonomous, student-centered learning. This enables them to uncover a broader spectrum of noun collocates for *take* that they have not previously utilized, thereby expanding their lexical knowledge in diverse contexts. Given the context of whole-class or group-based English learning in Korean public schools, the first two strategies—providing teacher-selected paper concordances and guiding students through them to identify deviant collocations—are more practical for implementing DDL. On the other hand, more personalized or small group-based classes, or English writing courses at the college level with smaller student numbers, could employ concordance tools more effectively.

Several limitations should be considered when interpreting this study's findings. Firstly, the size of the corpora is small, as the sub-corpora for the three proficiency levels were selected from a relatively limited portion of the total GLC, a scope narrower than that in previous corpus studies. Analyzing a broader range of GLC data could yield more robust results. Additionally, the lack of identified essay prompts across proficiency levels might have skewed the frequency of specific noun collocates, such as *bus* in the intermediate group and *change* in the high proficiency group, compared to other collocates. Although this study focuses on Korean learners' use and misuse of *take*-noun collocations at three distinct proficiency levels, incorporating and comparing data from native speakers' corpora would offer a more comprehensive understanding of the patterns exhibited by Korean EFL learners in their use of *take*-noun collocations. This study delves into how Korean learners of English produce *take*-noun collocations and scrutinize the relationship between proficiency levels and both L1 and L2 factors in the use of deviant collocations across different proficiency levels—a topic that has received limited attention in previous research. A future study, incorporating a broader analysis of learner data categorized by essay genres and a comparing it with native speaker corpora, could provide deeper insights into the patterns of high-frequency verb-noun collocations, such as, *have*, *get*, *give*, *do*, and *take*, and the use of deviant collocations by L2 learners at various proficiency levels.

References

- Altenburg, B., & Granger, S. (2001). The grammatical and lexical patterning of *make* in native and non-native student writing. *Applied Linguistics*, 22(2), 173-194.
- Chang, YiBoon. (2018). Features of lexical collocations in L2 writing: A case of Korean adult learners of English. *English Teaching*, 73(2), 3-36.
- Choi, Hye-Young, & Chon, Yuah. (2012). A corpus-based analysis of collocations in tenth grade high school English textbooks. *Multimedia-Assisted Language Learning*, 15(2), 41-73.
- Choi, Wonkyong. (2019). A corpus-based study on *delexical verb + noun* collocations made by Korean learners of English. *Journal of Asia TEFL*, 16(1), 279-293.
- Cowie, A. P. (1994). *Phraseology*. In R. E. Asher (Ed.), *The encyclopedia of language and linguistics* (pp. 3168-3171). Pergamon.
- Davies, M. (2008–present). *The Corpus of Contemporary American English (COCA): 520 million words, 1990-present*. <http://corpus.byu.edu/coca/>.
- Hong, Sinchul. (2021). A corpus-based comparative study of light verb constructions between native learners and Korean EFL learners from the grammatical structure perspective. *Studies in Foreign Language Education*, 35(4), 125-143.
- Jang, Eun Young. (2008). *An analysis of collocation in the elementary ELT reading textbooks: With a focus on delexical verbs*. [Unpublished master's thesis]. Yonsei University.
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *ELT Journal*, 4, 27-45.
- Jukneviene, R. (2008). Collocations with high-frequency verbs in learner English: Lithuanian learners vs native speakers. *KALBOTYRA*, 59(3), 119-127.
- Kang, Min-Kyung. (2014). A corpus-based analysis of Korean high school English textbooks based on the semantic distribution of high-frequency verb collocation. *English Language Teaching*, 26(1), 113-130.
- Kim, Myung-Hee. (2002). Uses of *make* in Korean EFL learner writing: A corpus-based study. *English Teaching*, 57(4), 297-314.
- Kittigossin, R., & Phoocharoensil, S. (2015). Investigation into learning strategies and delexical verb use by Thai EFL learners. *3L: The Southeast Asian Journal of English Language Studies*, 21(2), 63-72.
- Kwon, Hae-Young. (2012). *An analysis of 'get, have, take' collocations in high school 10 textbooks corpus-based study*. [Unpublished master's thesis]. Kyung Hee University.
- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647-672.
- Lee, JongBong. (2021). Using corpus analysis to extend experimental research: Genre effects in L2 writing. *System*, 102563.
- Lee, Sujeong, & Na, Yoon-Hee. (2015). The use of the verb *make* in the corpora of American NS students and Korean EFL learners. *English 21*, 28(1), 401-426.
- Lee, Youmin. (2020). *A corpus-based analysis of the light verb construction in 2019 EBS English textbooks: Focusing on make and take*. [Unpublished master's thesis]. Sogang University.
- Lewis, M. (2000). *Teaching collocation: Further development in the lexical approach*. Language Teaching Publications.
- Ma, Jee Hyun, & Kim, Youngsu. (2013). Korean high school English learners' knowledge of collocations: Focusing on delexical verbs: *Make, get, and take*. *Language Research*, 49(1), 45-71.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. John Benjamins.
- Park, Eunjeong. (2019). The effectiveness of corpus-aided instruction to improve second language college students' academic writing. *Modern English Education*, 20(3), 15-25.
- Park, Myongsu. (2020). Comparative study of delexical verbs in an EFL learner corpus. *Studies in Linguistics*, 56, 195-216.
- Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audiovisual input: An exploratory study into incidental learning of single words and formulaic sequences. *The Language Learning Journal*, 47(4), 424-438.
- Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 42(3), 525-549.
- Sawaguchi, R., & Mizumoto, A. (2022). Exploring the use of *make + noun* collocations by Japanese EFL learners through a bilingual essay corpus. *Corpora*, 17(Supplement), 61-77.
- Shi, J., Peng, G., & Li, D. (2023). Figurativeness matters in the second language processing of collocations: Evidence from a self-paced reading experiment. *Language Learning*, 73(1), 47-83.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Vilkaitė, L., & Schmitt, N. (2019). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics*, 40(2), 329-354.
- Wang, Y. (2016). *The idiom principle and L1 influence: A contrastive learner-corpus study of delexical verb + noun collocations*. John Benjamins.
- Wang, Y., & Chen, Y. (2023). A comparative study on verb-noun collocation errors of Chinese EFL learners with different language proficiency. *Journal of Literature and Art*, 13(8), 585-592.
- Yoo, Isaiah WonHo, & Shin, Yu Kyung. (2022). English lexical bundles in a learner corpus of argumentative essays written by Korean university students. *Corpora*, 17, 23-42.

Appendix A

The total types of noun collocates of *take* at three proficiency levels

Proficiency Level (total types)	Noun collocates of <i>take</i>
Low (39)	care, picture, rest, exercise, action, break, bus, camera, airplane, look, profit, room service, class, attendance, attention, haircut, wallet, manicure, pedicure, motorcycle, information, thinking, photos, shelter, present, clothes, lead, health, service, (cosmetic)surgery, buck, time, offer, place, part, money, fear, car, test
Intermediate (43)	care, picture, photo, bus, wheel, drink, look, time, meal, relax, rest, shampoo, shower, tour, walk, action, interest, (cosmetic) surgery, defect, money, penny, change, glass, consideration, room, (public)transportation, computer, subway, health, opportunity, cosmetics, students, space, class exercise, notice, fun, taxi, trip, minutes, part, possession, years
High (45)	care, picture, boat, bus, part, action, pill, (cosmetic) surgery, rest, swing, advantage, opportunity, change, transportation, time, meal, hour, medicine, money, earphones, eyes, cars, thumbs, offense, authority, drivers, test, message, information, photographs, movie, (digital) camera, exam, lecture, shower, class, license, antidepressant, tranquilizer, journey, room, walk, minutes, notebook, fire extinguisher,

Appendix B

The total types of noun collocates of delexical *take* at three proficiency levels

Proficiency Level (total types)	Noun collocates of delexical <i>take</i>
Low (23)	care (9), picture (4), rest (3), action (2), time (2), car (2) break, bus, look, profit, class, attendance, attention, information, photos, shelter, service, lead, offer, place, part, money, test
Intermediate (32)	care (12), bus (6), picture (4), photo (3), look (2), walk (2), computer (2), wheel, meal, rest, shower, tour, action, interest, defect, money, penny, change, glass, consideration, (public) transportation, subway, opportunity, space, class, notice, taxi, trip, minutes, part, possession, years
High (32)	care (9), change (8), picture (3), time (3), class (3), boat (2), bus (2), test (2), information (2), part, action, pill, rest, advantage, opportunity, transportation, hour, medicine, money, cars, offense, photographs, movie, exam, lecture, shower, antidepressant, tranquilizer, journey, walk, minutes, notebook

Note. The numbers in parentheses indicate the frequency of each noun's collocation with the delexical use of *take*. Nouns without parentheses have collocated with the delexical *take* only once.

Appendix C

The deviant delexical *take*-noun collocations at three proficiency levels

Proficiency Level (total types)	Deviant <i>take</i> -noun collocations
Low (13)	<p>*So slim people <u>take exercise</u> for staying.</p> <p>*But a lot of people to <u>take their health</u>.</p> <p>*We need to <u>take care ourself</u> from the pollution.</p> <p>*you should <u>take a 'real profit'</u></p> <p>*I don't <u>take Manicures, Pedicures</u> or Facials</p> <p>*But I regularly <u>take Haircut</u> because I hate messy hair.</p> <p>*I <u>take hair cut</u> monthly.</p> <p>*Then i <u>take a light airplane</u> and rose into the sky.</p> <p>*people run away <u>take motorcycle</u>, even airfort staff</p> <p>*it reminds me that the fear of scary in that time, too, took it for a few months due to stress gastritis.</p> <p>*Black color art <u>take old thinking</u> to people.</p> <p>*Women <u>take this service</u>.</p> <p>*she was <u>taken cosmetic surgery</u>.</p>
Intermediate (12)	<p>*When I <u>take the bus</u>, the car behind the bus very surprise me.</p> <p>*In the day many old people <u>take the bus</u> so drivers wait for them to sit.</p> <p>*people were starting to <u>take their health</u></p> <p>*If you want to happen to the rapid effect, <u>Take cosmetic surgery</u></p> <p>*I want to <u>take a relax</u> with awesome places.</p> <p>*They should not <u>take one glass of alcohol</u> for granted</p> <p>*but I regularly <u>take a shampoo and conditioner</u> because my hair is very damaged by several dying and perm.</p> <p>*I and my friends <u>take room</u> and play there all night.</p> <p>*Tired customer conveniently <u>take a meal</u> in the room through 24-hour room service.</p> <p>*people have been working on their weight and <u>taking exercise</u></p> <p>*They shouldn't think that the driver who has <u>taken a drink</u> is the person who has not <u>taken a drink</u> yet</p>
High (12)	<p>*I think the best way to improve his or her appearance is <u>take a plastic surgery</u></p> <p>*I love hobbit character in the movie who <u>take five meals</u> a day.</p> <p>*I always <u>take my earphones</u> everywhere</p> <p>*So when i go to the Lotte World, then i never <u>take a Xiyro Swing</u></p> <p>*So I always <u>take care</u>, when I drinking and do a computer game or web surfing.</p> <p>*If parents <u>take care their child</u> very carefully</p> <p>*It was so nice to go down the hill <u>taking a snowboard</u>.</p> <p>*And if they are <u>taken a massage</u>, they are more active.</p> <p>*I have heard that <u>it was taken</u> in New Zealand. Background of nature in the movie was so spectacular.</p> <p>*if the hotel is far away from my destination or have to <u>take complicate transportation transfers</u> from the hotel to the destination</p> <p>*And we <u>took a room</u> for tomorrow's sun rise.</p> <p>*He planned to kill her to <u>take the authority</u> to govern the world after she make the curse ended.</p>