

Vietnamese EFL Learners’ Productive Derivative Knowledge: Contextualized Derivative Recall Test Performance and Its Relationship with Derivative Misuse in Argumentative Essays

PHUONG NHI TRAN
PHAM THI HONG NHUNG
HUY VAN NGUYEN
BAO TRANG THI NGUYEN*

Faculty of English, University of Foreign Languages and International Studies, Hue University, Vietnam

*Corresponding author email: ntbtrang@hueuni.edu.vn

Article information	Abstract
Article history: Received: 16 Aug 2024 Accepted: 6 Oct 2025 Available online: 14 Oct 2025 Keywords: Productive derivative knowledge Contextualized derivative recall test Headwords counted (HC) Headwords not counted (HNC) Derivative errors Vietnamese EFL learners	<i>This study measured EFL learners’ productive derivative knowledge via a contextualized derivative recall test and explored its relationship with the misuse of derived forms in argumentative essays. Data were collected from 88 third-year English majors at a university in Vietnam, each writing an argumentative essay in their normal class hours and subsequently taking a contextualized derivative recall test of 30 target headwords. Test data were scored for the successful production of the target derivatives in two methods: headwords counted (HC) and headwords not counted (HNC). In addition, the argumentative essays were analyzed for derivative errors committed. Results indicated that these students were able to produce, on average, 52% and 44% of the target derivatives in the HC and HNC methods respectively. They were most successful with high-frequency target words, and accurate production rates were found to be subject to the scoring methods and derivatives’ parts of speech. Regarding erroneous use of derivatives in written essays, misuse of word parts of speech was more prevalent than others. In addition, there was a significantly positive correlation between the derivative recall test scores and the derivative errors in written essays, but only in the HC method. All these findings suggest that learners’ productive derivative knowledge may be influenced by numerous factors, and derivative recall test performance might not equate accurate use of derived forms in writing, thus providing important implications for assessing L2 productive derivative knowledge and instruction.</i>

INTRODUCTION

Derivative knowledge, knowledge about derived words or derivatives, is considered part of the multidimensional construct of vocabulary knowledge (Nation, 2013). It has been conceptualized either receptively or productively, in correspondence to the ability to recognize existing derived forms or produce derivatives of the headwords provided (Iwaizumi & Webb,

2021, 2022, 2023). According to Nation (2013), English-derived words are formed by adding affixes to a base word, which alters its meaning in the case of prefixed words (e.g., *possible*, *impossible*) and changes its part of speech and/or its meaning for suffixed words (e.g., *possibility*, *impossibility*). In this way, the varied grammatical or inflectional forms of the same word 'forget' (e.g., *forgets*, *forgetting*) are inflected, not derived words. Its derivations include *forgetfulness* (noun), *forgetful*, *forgettable*, *unforgettable* (adjective), and *forgetfully/unforgettably* (adverb). Knowledge about how to produce derived forms of a word "can help reduce the burden of learning new words and help learners improve their vocabulary knowledge, including form, meaning and use" (Sumalee & Sukying, 2024, p. 338). Drawing on knowledge of affixes, learners can recognize the meaning of given words by breaking them into different parts and creating related word forms, which facilitates vocabulary learning and development (Nation, 2013).

Derivative knowledge also plays a pivotal role in facilitating foreign or second language (L2) learners' communication (Iwaizumi & Webb, 2022; Nation, 2013) and it is instrumental in vocabulary expansion (e.g., Laufer, 2017; Nation, 2013; Sasao & Webb, 2017), pedagogy and materials design (Graves et al., 2019; Iwaizumi & Webb, 2022, 2023). In addition, knowledge of derivational morphology helps develop L2 reading proficiency (Kieffer & Lesaux, 2012) and writing ability (Leontjev et al., 2016). While research has shown that L2 learners often acquire substantial knowledge of receptive derivational affixes (Iwaizumi & Webb, 2023; Laufer et al., 2021; Matthews et al., 2023; Snoder & Laufer, 2022; Sukying, 2018), they tend to demonstrate a low level of productive derivative knowledge (i.e., the ability to comprehend and generate derivatives of base words) as measured by either contextualized (Iwaizumi & Webb, 2021; Schmitt & Zimmerman, 2002) or decontextualized recall tests (Iwaizumi & Webb, 2022, 2023). However, research on L2 learners' knowledge of derivatives is still limited (Iwaizumi & Webb, 2022, 2023; Leontjev et al., 2023) and further investigations are warranted with learners in different contexts. Some rare research (e.g., Leontjev et al., 2016) found that certain aspects of L2 learners' derivative knowledge correlated with their overall writing proficiency, highlighting the important role of derivative use in writing. However, the relationship between L2 learners' productive derivative knowledge and their use of derivatives in extended communication such as writing has largely remained unexplored.

The present study thus fills these gaps by measuring Vietnamese EFL students' knowledge of productive derivative knowledge via a contextualized recall test developed by Iwaizumi and Webb (2021) and explores its relationship with how these students used derived forms in essays. Understanding this empirical link will provide useful insights into L2 learners' productive derivative knowledge and inform derivative assessment and pedagogy. Correlations, if any, between students' ability to recall and produce derived forms via the recall derivative test and their active use of derivatives in writing may provide a useful reference for appropriate test extrapolations and instructional practices that assist students in derivative word development and enhance writing quality.

LITERATURE REVIEW

Measuring L2 productive derivative knowledge

A word family is a counting unit that consists of a base word (e.g., *prepare*), its inflected words (e.g., *prepares*, *prepared*, *preparing*), and derived words (e.g., *preparation*, *preparatory*) (Matthews et al., 2023; Sumalee & Sukying, 2024) and higher frequency word families tend to have more derived forms (Nation, 2013). Nation (2021) noted that “many learners have problems in dealing with word families in that they do not recognize that a complex word form is made of a known stem and affix (or affixes) and that the meaning of the complex form relates to the meaning of its parts” (p. 970). This further reiterates the need to measure the ability to produce derived forms accurately and the challenges learners encounter.

According to Schmitt and Zimmerman (2002), productive derivational knowledge refers to L2 learners’ ability to generate different parts of speech of a word. To date, productive derivative knowledge has been generally measured via decontextualized and/or contextualized recall tests. In a decontextualized derivative recall test, test takers are typically required to produce as many derivatives as possible for each of the given headwords (Iwaizumi & Webb, 2022, 2023). On the other hand, a contextualized derivative recall test gauges L2 learners’ capability to generate derived forms of words within sentence contexts (Iwaizumi & Webb, 2021). In particular, for each headword, hints (noun, adjective, verb, adverb) are provided in the context of gapped sentences. Participants are required to generate derived forms that fit the provided context.

Previous research has invariably demonstrated that producing English derivatives can be challenging for many groups of L2 learners (e.g., Iwaizumi & Webb, 2022, 2023; Leontjev et al., 2016). For example, according to Schmitt and Meara (1997), Japanese EFL students in their study were able to produce 15% of all the target derivatives in a decontextualized recall test of 20 headwords. In a recent study, Iwaizumi and Webb (2022) found that producing derivatives in a decontextualized recall test of 90 headwords was particularly daunting for L2 learners and even L1 speakers. The production rates were about 40% and 25% of the possible derivatives respectively for these two groups. In a study by Nguyen and Nguyen (2024), EFL students’ productive derivative knowledge was measured via a decontextualized derivative test and a low rate of accurate derivative production of about 25% of the time was found. This was the first study to examine the impact of the scoring method (e.g., headwords counted (HC) and headwords not counted (HNC)) on the accurate production of the target derived forms in a contextualized derivative test. The HC method includes the successful identification of the parts of speech (POS) of the target headwords as correct responses whereas the HNC excludes it. While the ability to assign the correct POS to a given prompt word indicates L2 learners’ derivational knowledge, it is not the same as the ability to generate its derived forms (Nguyen & Nguyen, 2024; Nguyen et al., 2025). For example, identifying the prompt word ‘*differ*’ as a verb is arguably less challenging than writing its derived forms (*different*, *differently*, *difference*). As such, the HC method might potentially overstate the derivative test results, underscoring the need to evaluate L2 learners’ productive derivative knowledge more appropriately. In this regard, the HNC method could align more with L2 learners’ ability to *produce* derived forms

of given headwords. The HC and HNC methods have yet to be applied in scoring the contextualized derivative recall test, though they have been used in scoring decontextualized derivative tests (Nguyen & Nguyen, 2024).

Research that measures L2 productive derivative knowledge via contextualized recall tests remains scarce. Using a contextualized derivative recall test with 16 headwords, Schmitt and Zimmerman (2002) found that undergraduate students were able to produce accurately about 55% of the target derivatives, which was lower than that of the postgraduate students (80%). That said, these production rates were higher than those found with the decontextualized derivative test in other research (e.g., Iwaizumi & Webb, 2022, 2023; Nguyen & Nguyen, 2024). By employing a contextualized derivative recall test with 30 headwords from different word frequency levels (1000-5000), Iwaizumi and Webb (2021) compared the productive derivational knowledge of 21 L1 speakers (university students), 18 ESL graduate students in Canada, and 61 EFL undergraduate students in Japan. The results showed that the EFL students were outperformed by the ESL and L1 speaker groups. On average, they produced correctly about 57%, 76% and 91% of the possible derived forms, respectively. These findings indicate that EFL students might find it more difficult to produce derived forms of the target headwords than other groups of learners who have exposure to input outside of the classroom.

In another context, Leontjev et al. (2016) targeted secondary students aged from 15–18 in Estonia and Finland. They measured the participants' productive derivative knowledge via three types of test: i) two decontextualized tests where learners provided derived forms of individual headwords, ii) a metalinguistic prompt-based task eliciting derivatives based on the metalinguistic information given, and iii) one contextualized test measuring production of the target derivatives in sentence contexts. The results showed that the learners were successful at 26%, 30%, and 64% of the time respectively, suggesting a higher success rate with contextualized than decontextualized derivative recall tests.

Research has begun to demonstrate the significant contribution of derivatives' POS in L2 derivative production. For instance, nouns and verbs were found to be more accurately produced than others in a decontextualized derivative recall test of 90 headwords performed by EFL undergraduates in Japan (Iwaizumi & Webb, 2022). With the same decontextualized derivative test, Nguyen et al.'s (2025) study uncovered that Vietnamese EFL university students were more successful with verb derivatives than nouns, adjectives, and adverbs when counting headwords as correct responses. However, the pattern shifted with verbs being the least successful in the HNC method. These findings were explained with reference to the fact that more than 50% of the target verb derivatives were prompt words, whereas the proportion of prompt words that are nouns, adjectives, and adverbs was much lower (about 14.8 %, 10.5%, 0.9% respectively). This contributed to the higher success rate of verb derivative production in the HC method and a much lower proportion of it when identifying the correct POS of prompt words was not counted. Overall, the findings suggest the impact of scoring methods that need to be considered in scoring derivative tests.

Turning to the contextualized derivative format, verb derivatives were more successfully produced than other word classes in a contextualized recall test of 16 headwords performed

by ESL undergraduates and graduates (Schmitt & Zimmerman, 2002). This again indicates that derivatives' POS is a potential contributing factor that needs further exploration. In a contextualized derivative test, recognizing the POS of the given headwords needed to complete the intended gaps is arguably less challenging than writing their derived forms that fit those gaps. It is therefore highly necessary to utilize both HC and HNC scoring methods to better measure L2 derivative knowledge.

Misuse of derivatives in L2 writing

Knowledge about word forms is one key word knowledge component (Nation, 2013) and as such, investigating derivative errors, one common type of lexical errors in written language production (Llach, 2011), will shed light on learners' stage of vocabulary acquisition. As Laufer (1991) puts it, "lexical errors are a reliable instrument to investigate the organization of the mental lexicon in L2 and to find out more about vocabulary development" (p. 270). Findings about misuse of derived forms in EFL writing will reveal gaps about what learners wish to communicate and the language forms they have to express the intended message, thus producing 'pushed output' that drives language development (Swain, 2005). The lexical errors learners commit also significantly predict the quality of their written text (e.g., Llach, 2011, 2017), which has important pedagogical value in informing instructors' feedback and remedial approaches.

With such importance of understanding lexical errors, considerable research attention has been channeled to examining general errors in written essays by EFL learners in different educational contexts. For example, Chan (2010) scrutinized the writing samples of Hong Kong Cantonese ESL learners and found that lexical errors were quite common, and errors in the deployment of affixes were the most frequent of derivative erroneous use. Suffixes were among other lexical errors committed by Thai EFL learners in their argumentative essays (Hemchua & Schmitt, 2006), whereas incorrect use of prefixes and parts of speech was frequent in EFL writing by Mongolian English majors at a Chinese university (Li, 2015).

There are only a handful of studies focusing exclusively on derivative errors in written texts by learners in varied instructional contexts. For example, Ibsheeyah et al. (2023) recently investigated the errors related to English derivational morphemes used by English-major university students in Libya in their written compositions. They found more errors with suffixes than prefixes and with verbs and nouns than with other word classes. In contrast, in Ali et al.'s (2021) study, university students in Pakistan often omitted suffixes to form noun and adjective derivatives in a derivative-eliciting task. Using affixes is also problematic for Grade 11 students in Tangerang in their narrative compositions, and errors that involve noun derivatives were common (Kusumawardhani, 2018). Generally, the findings of these studies indicate that using derivatives in writing is still difficult for many learners, though variation exists regarding the misuse of different word parts. This could be due to the different tasks employed to elicit the written derivatives and the different groups of learners whose committed errors might be influenced by their different L1 backgrounds.

Misuse of derived forms could be explained from different perspectives. Firstly, the ways in which learners acquire derivatives could be a contributing factor. Derivative acquisition

typically occurs in the item-based and system-based manners (Boers, 2021; Schmitt & Zimmerman, 2002). As its name suggests, item-based learning refers to learning a given derived form as a whole word without decomposing it into different elements. In contrast, in the system-based manner, learners rely on their derivational affix knowledge to analyze a word, segmenting it into the different components (e.g., disappointment: dis-, appoint-, -ment). Next, interlingual and intralingual perspectives commonly used to explain lexical errors (Boers, 2021) could apply to account for derivative errors, one type of lexical misuse. In the former, the influence of learners' first language (L1) is quite often cited as learners transfer the process of word formation from their L1 to the target language. In this regard, a lack of congruency in the target language features between the learners' L1 and the target L2 might contribute to errors. The latter refers to L2 influence or 'intra-lexical interference' (Boers, 2021) where misformation of words is viewed to result from incomplete knowledge of the target language features. Different accounts of intralingual factors in the misuse of derivatives have been centered around learners' limited knowledge pertaining to derivational patterns and rules (e.g., Al-khresheh, 2016) or insufficient exposure to authentic language input (Heydari & Bagheri, 2012) that might lead to errors such as overuse of affixes, misaffixation and misuse of POS (Chan, 2010; Hemchua & Schmitt, 2006; Li, 2015). Although these factors are only possible post-hoc explanations, they pinpoint where errors lie to offer pedagogical recommendations (Ellis & Barkhuizen, 2005).

The aforementioned studies have examined different aspects of lexical and derivative errors in written language production and provided important implications for writing instruction and vocabulary development. However, no research to date has explored the relationship between the derivative errors L2 learners commit in EFL writing and their productive derivative knowledge as measured through a contextualized derivative recall test. It is necessary to not only understand what learners know but also how they use derivatives so as to measure their lexical knowledge more comprehensively (Nation, 2013). This line of research is pressing to inform derivative test development and pedagogy. Any found relationship between recall test performance and derivative misuse in meaning-focused activities (i.e., argumentative essay writing) may advise practitioners on how to appropriately measure productive derivative knowledge and develop L2 derivatives for meaningful communication.

This study thus explored the extent to which Vietnamese EFL learners produced the derivatives of the target prompt words in a contextualized derivative recall test. It also seeks to find the correlation between derivative test scores and erroneous use of derivatives in argumentative essays. In other words, the following three research questions (RQs) are addressed:

RQ1: To what extent do EFL students produce derivatives of the target headwords in a contextualized derivative recall test?

RQ2: What derivative errors do EFL students make in their argumentative essays?

RQ3: What is the relationship between the derivative errors EFL students committed in argumentative writing and their productive derivative knowledge measured via the contextualized derivative recall test?

METHODOLOGY

Participants

A convenience sampling method was adopted to select participants in the present study. A total of 88 students in their third year of a Bachelor's English-major program at a Vietnamese university, willingly participated in this research. They were enrolled in two writing classes taught by the fourth author, and expected to achieve a higher CEFR-B2 level as the course learning outcome. They were aged around 20 on average, and a majority of them (75) were female and all had completed a prior writing course that targeted a low CEFR-B2 level. That said, as judged by the class teacher, their proficiency levels varied from a low B2 to a low C1 level. Competent argumentative essay writing was one of the targets of the course.

Data collection

The contextualized derivative recall test

To assess the participants' productive knowledge of derivatives, the contextualized derivative recall test developed by Iwaizumi and Webb (2021) was employed. Each target derivative was accompanied by 3-5 sentences that elicit the derived forms according to the given parts of speech. As described by Iwaizumi and Webb (2021), the sentences were sourced from dictionaries and contained high-frequency words to ensure that language is not an issue for test completion. Both prefixed and suffixed words were targeted. The complete test version with the answer keys can be downloaded from <https://onlinelibrary.wiley.com/doi/abs/10.1002/tesj.640>. An example related to the prompt word "**arrange**" is extracted from the test for illustration as follows. The students were asked to complete each gap with the correct derived form of the word provided.

arrange

- Noun (1) I like the [*arrangement*] of the tables in this room.
- Noun (2) She is an [*arranger*].
- Verb (1) I [*arranged*] the meeting.
- Verb (2) I [*rearranged*] the meeting because I could not make it last time.

As depicted in Iwaizumi and Webb (2021), the target headwords in this test were sourced from the most frequent 1000–5000 word lists in the British National Corpus (BNC)/ the Corpus of Contemporary American English (COCA) (Nation, 2012). In total, the test contained 30 headwords, 10 from each level: 1000, 2000, 3000-5000 with more headwords as prompt words from the first two levels ($k = 20$) than the 3000-5000 level ($k = 10$) (see Appendix B). Iwaizumi and Webb (2021, p. 8) reported three main reasons for which more prompt words are targeted in the 1000 and 2000 word lists. Firstly, they contain more frequent words than the 3000–5000 list. Secondly, more frequent words have a more extensive coverage in English spoken and written texts and test takers of different proficiency levels could be familiar with them due to their higher frequency. Thirdly, the inclusion of fewer words in the 3000-5000 group was because lower-frequency words are of a similar challenge for learners.

A printed version of the test was administered to the students in their normal writing classes. Following Iwaizumi and Webb (2021), before commencing the test, clear instructions were provided to the participants, with further explanation of the grammatical terms such as nouns, adjectives, adverbs, and verbs as needed. Participants were instructed to read and complete each gapped sentence with the correct derived form of the word provided. The participants had as much time as they wanted for test completion so that time pressure would not be a factor affecting recall of derived forms. On average, these students spent approximately 45 minutes on this test.

Argumentative essays

The present study focused on argumentative essays for two main reasons. First, they were one major component of the target writing courses where data were collected for the present study. Furthermore, essay writing provides students ample opportunities to use their lexical resources, and thus derived words, to convey intended meanings and express their positions.

Students were instructed to write an argumentative essay of at least 250 words on a given paper sheet in their normally scheduled writing classes as a practice test and in as much time as they needed, without external resources, such as dictionaries or assistance from the teacher or peers. According to Laufer and Nation (1995), a minimum word count of 250 words would supply reliable results in view of lexical use. Since the two writing classes were scheduled on the same date, one following the other and taught by the same teacher, two different topics were given to prevent inter-class exchanges: one class wrote about the topic of salary and the other TV game shows (see Appendix A for the task prompts). In total, there were 88 scripts written by those who also subsequently completed the contextualized derivative recall test. They spent 65 minutes on average writing the essays. In order to avoid the practice effects of students being aware of the use of derivatives, students wrote an argumentative essay before completing the contextualized derivative recall test.

Data analysis

Scoring the contextualized derivative recall test

The answer keys, composed by Iwaizumi and Webb (2021) were used as a guide for scoring the contextualized recall test. Each correct answer was awarded 1 point and an incorrect one or no response received 0 point. In some research (e.g., Iwaizumi & Webb, 2021; McCutchen & Stull, 2015; Schmitt & Zimmerman, 2002), the participants received one point for responses that contain spelling mistakes in derivatives, as long as those misspelled forms were phonetically accurate. However, as we aimed to find the correlation between learners' derivative test performance and derivative errors in written essays, we did not count this kind of misspelling and there were only four of them in the data. Following Iwaizumi and Webb (2021), adjectives that are formed by inflectional morphemes (e.g., *varied*, *arranged*) were counted as derived words. These applied in both HC and HNC methods, but the scoring differed in these two scoring methods as follows.

In the HC method, the correct use of the prompt word as it is to complete each gap was included, “because identifying parts of speech of the prompt words and correctly filling out the blanks required participants to demonstrate their knowledge about the items” (Iwaizumi & Webb, 2021, p. 10). However, correctly recognizing the POS of a given item and producing its derivatives are not the same (Nguyen & Nguyen, 2024). For example, using the given headword ‘art’ as a noun to fill in the intended gap is arguably less demanding than writing its derived form (‘artist,’ ‘artistic,’ ‘artistically’). Accordingly, in the HNC method, if the required gaps are the same as prompt words, they were not included. The maximum number of derived forms in the contextualized recall test is 119 (following Iwaizumi and Webb’s (2021) answer keys) and 89 (excluding correct recognition of given POS) in the HC and HNC methods respectively. The Cronbach alpha values of $\alpha = .916$ (HC) and $\alpha = .887$ (HNC) for the derivative test data in the present study demonstrated a high level of internal consistency for the contextualized derivative recall test.

Identifying derivative errors in the written essays

Each essay was manually scrutinized for the errors related to derivative use that emerged. The focus was exclusively on derived forms and a descriptive taxonomy of derivative errors was employed as it is practical to inform pedagogy (Ellis & Barkhuizen, 2005). Since derived words are formed by means of affixation that might alter the meaning, and/or the part of speech of the newly affixed items (Nation, 2013; Sumalee & Sukying, 2024), these errors in the essay data were linguistically categorized into: misuse of affixes, misuse of POS, and mixed errors. Following Nation’s (2013) definition of derivatives, if adding inflectional morphemes to a base word changes that word’s part of speech (e.g., *arranged* as an adjective), it was counted as a derived form. Definitions and examples of each type of errors made as they were in the original essays are presented as follows, and the correct options are provided in square brackets.

Misuse of affixes involves applying incorrect prefixes or suffixes to words, resulting in the formation of non-standard words. In our study, this kind of error did not count misspellings that were phonetically accurate representations (e.g., *communicater*, *communicator*) (c.f. Iwaizumi & Webb, 2021) to be consistent with the derivative recall test scoring. Erroneous use of prefixes in the current data mainly involved creating prefixed words with opposite meanings that are not target-like:

*It is **irreasonable** to let children become famous when they are too young. (E11)*
[unreasonable]

*They will have **imsatisfaction** with their work. (E3)[**dissatisfaction**]*

Misuse of suffixes entailed using words that contain incorrect suffixes in the written texts:

*Those doing the job usually have a chance to communicate with customers as well as flight **attendance**... (E52) [**attendant**]*

*Kids may have a chance to work as a **cooker**. (E7) [**cook**]*

*Ideal jobs can bring humans **excition**. (E19) [**excitement**]*

Suffix misuse could include adding suffixes where they are not necessary:

*Salary helps them work **hardly** and effectively. (E30) [hard]*

*Young people want to find out a company having professional environment and especially **highly** salary. (E87) [high]*

*Doing a job with **joyness** will make people relaxed and enjoy their work. (E60) [joy]*

*They have chance to work with **mentorships** who expert in this field. (E72) [mentors]*

Misuse of parts of speech involved the incorrect use of the POS of the intended words. Examples of this kind of erroneous use are presented in Table 1.

Table 1
Types of derivative errors and examples from the original essays

Type of POS misuse	Examples
Using noun instead of verb	<i>The increasing popularity of entertainment on TV or Internet devices is the condition to production game shows for young children. (E12) [produce]</i>
Using noun instead of adjective	<i>Children will be awareness of learning soft skills. (E16) [aware]</i>
Using noun instead of adverb	<i>They have to work hardness to earn a living. (E61) [hard]</i>
Using adjective instead of noun	<i>Working with a huge passionate can bring people a sense of satisfaction. (E77) [passion]</i>
Using adjective instead of verb	<i>Many parents mandatory their children to take part in TV shows... (E30) [mandate]</i>
Using adjective instead of adverb	<i>They can free develop their skill. (E11) [freely]</i>
Using verb instead of noun	<i>Young people now tend to put an emphasize on well-paid jobs... (E54) [emphasis]</i>
Using verb instead of adjective	<i>Kids are allowed to get involve in TV programs. (E28) [involved]</i>
Using adverb instead of noun	<i>Motivation will help people overcome any hard in their jobs. (E64) [hardship]</i>
Using adverb instead of adjective	<i>They have opportunities to receive comments from professionally and pedagogical experts. (E38) [professional]</i>

Mixed errors refer to errors that involved more than one of the types above:

*Some citizens will be **inawareness** of the drawbacks of TV shows and continue to allow their children to join in. (E16) [unaware]*

This example contains a combination of incorrect prefixes (**in-** instead of **un-**) and POS (noun, **awareness** instead of adjective, **aware**). Similarly, the next example involves the misuse of prefixes '**im-**' for '**un-**' and POS (noun instead of adverb):

*When they just work for salary but not for passion, they may work **imconfidence** because that job is not their field. [E50] [unconfidently]*

The accuracy of derived forms in the written essays was checked by reference to several dictionaries, namely the Cambridge Dictionary (<https://dictionary.cambridge.org/dictionary/english/>), the Collins Dictionary (<https://www.collinsdictionary.com/dictionary/english/>), the Longman Dictionary of Contemporary English (<https://www.ldoceonline.com/dictionary/english/>) and the Oxford Learner's Dictionary (<https://www.oxfordlearnersdictionaries.com/definition/english/>). If any derivative item does not appear in any of these dictionaries, it was further checked in other English corpora such as BNC and COCA for its occurrences. If no such derived form exists, it was considered erroneous.

Inter-reliability and statistical analyses

Fifteen percent of the test and essay data were randomly selected for inter-coding. The first author, and an EFL student with an IELTS score of 8.0 and prior experience in doing language education research, performed independent test scoring and error identification. The agreement percentages for the former were 90%, 93%, 88%, and 100% for noun, adjective, verb, and adverb derivatives respectively. The two coders also reached a large consensus on the latter with a high agreement of from 87% to 94% for the categories of derivative misuse in the essays. By reference to a minimum benchmark of 80% (Yin, 2015), all these figures indicate high inter-reliability. Instances of uncertainty and differences were resolved through discussions between the two coders. Subsequently, the remaining data were coded by the first author.

The test scores were prepared in the form of an Excel spreadsheet, inputted for each target derivative associated with each target headword. The data were then analyzed using the Statistical Package for the Social Sciences (SPSS) (Version 20.0) for descriptive and inferential statistics. As recommended by Field (2018), the normality of the data was checked via the Shapiro-Wilk test and the results indicated non-normality (all p values < .05). In order to compare the mean percentages between derivatives of different word classes in each scoring method (HC and HNC), Friedman non-parametric tests were performed. Post-hoc Wilcoxon signed-rank tests were further conducted and Bonferroni correction significance level of $p < .008$ (.05/6) for multiple comparisons were applied, with the effect sizes (r) additionally reported. The relationship between the derivative test scores and the derivative errors was examined via non-parametric Spearman's rho correlation tests because the normality assumption was violated. As guided by Field (2018), the correlation coefficients themselves were effect sizes: small ($r_s = .10$), medium ($r_s = .30$) and large ($r_s = .50$).

FINDINGS

RQ1. To what extent do Vietnamese EFL students produce derivatives of the target headwords in a contextualized derivative recall test?

The first research question aimed to measure the students' ability to produce derivatives in a contextualized derivative recall test of 30 headwords. The results (Table 2) indicated that they were successful, on average, at 52% and 44% of the time in the HC and HNC methods

respectively. In both methods, the accuracy rate at the 1000 word frequency level was the highest, followed by the 3000-5000 and 2000 levels. The Friedman test results indicated statistically significant differences in the correct production of derivatives at the different levels in both HC ($\chi^2(2) = 59.730$, $p < .001$) and HNC ($\chi^2(2) = 48.31$, $p < .001$) methods. In particular, as shown by the post-hoc Wilcoxon signed ranks test results (Table 3), greater success was found with the 1000 level (Mdn = .56) than the 2000 and 3000-5000 levels (Mdn = .46 and Mdn = .52 respectively) in the HC method ($p < .001$). In the HNC method, while more derived forms were produced at the 1000 than 2000 levels ($p < .001$), with a large effective size ($r = .63$), there were no statistical differences in production rates between the 1000 and 3000-5000 levels. Notably, learners produced more derivatives at the 3000-5000 than 2000 levels in both methods ($p < .001$). Clearly, word frequency and scoring methods could play important roles in L2 learners' successful production of target derivatives.

Table 2

Descriptive statistics for the correct responses (mean percentages) in the contextualized recall test (N = 88)

	1000		2000		3000-5000		Total	
	M	SD	M	SD	M	SD	M	SD
HC	.57	.16	.46	.16	.52	.21	.52	.17
HNC	.47	.20	.37	.17	.46	.22	.44	.18

Table 3

Wilcoxon signed ranks test results (N = 88)

		2000 vs .1000	3000-5000 vs. 1000	3000-5000- vs. 2000
HC	Z	-6.793 ^a	-4.046 ^a	-4.617 ^b
	Asymp. Sig. (2-tailed)	.000	.000	.000
	r	.72	.43	.50
HNC	Z	-5.945 ^a	-.839 ^a	-5.753 ^b
	Asymp. Sig. (2-tailed)	.000	.402	.000
	r	.63	-	.26

Note. a. Based on positive ranks.

b. Based on negative ranks.

Table 4 further presents the descriptive statistics (mean percentages) of the contextualized derivative recall test performance across different word parts. In the HC method, the learners' accurate production rates of noun, adjective, verb, and adverb derivatives were 55%, 43%, 68%, and 49% respectively. In the HNC, these rates decreased to 49% for nouns, 38% for adjectives, 31% for verbs while the rate for adverb derivatives remained unchanged.

In order to compare production rates of derivatives among word parts, non-parametric Friedman tests were performed for the data non-normality. The results revealed statistically significant differences between derivatives of different parts of speech, $\chi^2(3) = 147.406$, $p < .001$ (HC) and $\chi^2(3) = 92.068$, $p < .001$ (HNC). In the HC method, the post-hoc Wilcoxon signed-rank test results indicated that verbs (Mdn = .74) were more accurately produced than nouns (Mdn = .54), $Z = -6.361$, $p < .001$, $r = .68$, adjectives (Mdn = .40) ($Z = -7.980$, $p < .001$, $r = .85$) and adverbs (Mdn = .53), $Z = -7.136$, $p < .001$, $r = .76$. Meanwhile, nouns were more accurately produced than adjectives ($Z = -7.674$, $p < .001$, $r = .82$), and adverbs ($Z = -4.006$, $p < .001$, $r = .43$). In addition, the learners were more successful in forming adverb than adjective derivatives

($Z = -4.501$, $p < .001$, $r = .48$). These results point to the impact of POS on the accurate production of target derivatives with verbs and nouns being more successful than other word parts when headwords were counted.

Turning to the HNC method, a different pattern was found with verbs being the least successful of all. Their accurate production ($Mdn = .33$) was significantly lower than nouns ($Mdn = .47$) ($Z = -5.919$, $p < .001$, $r = .63$), adjectives ($Mdn = .33$) ($Z = -3.028$, $p = .002$, $r = .32$) and adverbs ($Mdn = .53$) ($Z = -5.599$, $p < .001$, $r = .60$). Noun derivatives were more successfully generated than adjectives, $Z = -6.930$, $p < .001$, $r = .74$ while adverbs were more accurately produced than adjectives, $Z = -6.509$, $p < .001$, $r = .69$. In addition, there was no significant difference between nouns and adverbs, $Z = -.213$, $p = .832$. These findings collectively suggest that L2 learners' ability to produce target derivatives in a contextualized recall test might be highly subject to the target items' POS and the scoring methods (HC/HNC).

Table 4
Descriptive statistics (mean percentages) of the contextualized derivative recall test
across different word parts

Scoring method	Part of speech	Min	Max	M	SD
HC	Noun	.15	.95	.55	.17
	Adjective	.05	.81	.43	.18
	Verb	.18	.94	.68	.18
	Adverb	.00	1.00	.49	.23
HNC	Noun	.10	.97	.49	.19
	Adjective	.00	.81	.38	.18
	Verb	.00	1.00	.31	.29
	Adverb	.00	1.00	.50	.22

RQ2. What derivative errors do EFL students make in their argumentative essays?

Table 5 presents the different types of derivative errors found in the students' essays. In total, misuse of POS constituted more than half of the derivative errors followed by misuse of affixes and mixed errors. Erroneous use of suffixes was more frequent than that of prefixes, at 19.81% and 3.1% respectively. Of the POS-related errors (Table 6), misuse of adjectives and nouns was substantial (about 38% each) while that of verbs (about 19%) and adverbs (about 5%) was much less common.

Table 5
Frequency of derivative errors in students' argumentative essays

Types of errors	Frequency	Total
Misuse of POS	N for other POS	327 (63.50%)
	Adj for other POS	
	V for other POS	
	Adv for other POS	
Misuse of affixes	Misuse of prefixes	118 (22.91%)
	Misuse of suffixes	
Mixed errors	70 (13.59%)	70 (13.59%)
Total	515 (100%)	515 (100%)

Note. POS: parts of speech; N: noun, V: verb; Adj: adjective; Adv: adverb

Table 6
Distribution of misuse of POS errors in students' argumentative essays

Misuse of POS		Frequency	Percentage	Total
Misuse of nouns	N for V	36	11.01	124 (37.92%)
	N for Adj	84	25.69	
	N for Adv	4	1.22	
Misuse of adjectives	Adj for N	98	29.97	126 (38.53%)
	Adj for V	9	2.75	
	Adj for Adv	19	5.81	
Misuse of verbs	V for N	43	13.15	61 (18.66%)
	V for Adj	18	5.51	
Misuse of adverbs	Adv for N	4	1.22	16 (4.89%)
	Adv for Adj	12	3.67	
Total		327	100	327 (100%)

Note. N: noun, V: verb; Adj: adjective; Adv: adverb

RQ3. What is the relationship between the derivative errors students make and their productive derivative knowledge?

RQ3 asked about the relationship between the derivative errors made by the students and their productive derivative knowledge. The descriptive statistics (Table 7) show that, on average, these Vietnamese EFL students accurately produced 61.09 derived forms out of 119 of the 30 target headwords and 38.34 out of 89 target items in the HC and HNC methods respectively. They committed an average of 6.69 derivative errors. Note that there was great individual variation (large gaps between minimum and maximum values and large standard deviations).

Table 7
Descriptive statistics for the contextualized derivative recall test scores and derivational errors in written essays

		Min	Max	M	SD
Contextualized derivative recall test scores	HC	14	106	61.09	19.70
	HNC	4	77	38.34	16.25
Errors in essays		0	24	5.85	4.51

The Spearman's rho correlation tests were conducted for the non-normally distributed data and the results for the HC method (Appendix C) revealed a positive relationship between the contextualized derivative recall test scores and the derivational errors that students made in their essays, $r_s = .225$, $p = .035$, though the effect size was small. In other words, the more derived forms correctly produced in the recall test, the more derivational errors committed in EFL argumentative essays. In particular, verb derivatives in the contextualized test had the highest correlation with the derivative errors ($r_s = .348$, $p = .001$), followed by nouns ($r_s = .218$, $p = .042$) and adverbs ($r_s = .217$, $p = .042$), though all with small to medium effect sizes. The correlation between adjectives and the derivative errors in argumentative essays was not statistically significant ($r_s = .178$, $p = .098$).

Regarding the HNC method, the results (Appendix D) followed a different pattern, indicating no significant correlation between the overall derivative test performance and the derivative

errors in EFL essays ($r_s = .183$, $p = .089$), with only adverb derivatives having a weak positive relationship with derivative misuse ($r_s = .217$, $p = .042$). The findings altogether indicated the significant role of the scoring methods in assessing L2 learners' derivative test performance and its relationship with erroneous use of derivatives in essays.

DISCUSSION

Students' productive derivative knowledge measured by the contextualized derivative recall test

Overall, the students achieved an average accuracy rate of 52% and 44% of the time in the contextualized derivative recall test using the HC and HNC methods, respectively. The HC finding is comparable with previous research (Iwaizumi & Webb, 2021; Leontjev et al., 2016; Schmitt & Zimmerman, 2002) showing learners were successful within the 50% to 65% range in a contextualized derivative test. Moreover, these rates were higher than those in a decontextualized derivative test (e.g., Iwaizumi & Webb, 2022; Nguyen & Nguyen, 2024) which found a success rate of about 25% of the time. This suggests that students may perform better when contextual cues are given because in the contextualized derivative test, each gapped sentence provides the context for students to guess the meaning of the needed derived form, thus facilitating its recall. In contrast, the decontextualized derivative test is more challenging as it requires students to produce all the possible derivatives of given individual words without any context to understand their meaning. However, further research is necessary to verify the impact of different test formats on derivative production with the same group of learners (Iwaizumi & Webb, 2024).

The findings further demonstrated that the highest percentage of accurate derivative usage was observed at the 1000 frequency level. This is congruent with previous research (e.g., Ishii & Schmitt, 2009; Iwaizumi & Webb, 2021), which found greater success with headwords of higher frequency (1000 level) because of their familiarity and common usage in daily communication. However, it is surprising that the performance at the 3000-5000 frequency level demonstrated a higher proportion of accurate production than the 2000 level. This could be perhaps because the average frequency of target derivatives might be highest at the 3000 level, followed by that at the 2000 level in both academic and spoken COCA corpora (Iwaizumi & Webb, 2021). Furthermore, some derivatives at the 3000 and 4000 levels are highly frequent words (e.g., *difference*, *communication*), which may have contributed to the greater success with the 3000-5000 level words. Iwaizumi and Webb (2021) argued that "the presence of such frequent derivatives may have obscured the effect of word family frequency levels of the production of derivatives" (p. 14). However, these explanations are provisional and warrant additional research that analyses the derivatives produced and their frequencies to better understand L2 learners' derivative knowledge.

The results of the contextualized derivative recall test also revealed the highest success rates of verb and noun derivatives, which broadly corroborates the findings of prior studies counting headwords in a contextualized (Schmitt & Zimmerman, 2002) and decontextualized (Iwaizumi

& Webb, 2022; Nguyen & Nguyen, 2024) formats. The high percentage of accurate verb usage in the HC method in the present study could be explained by the fact that 14 out of the 30 headwords were verbs, facilitating identification of their POS to complete the gapped sentences. This was because recognizing the POS of the given headwords as verbs to complete the intended gaps is arguably less challenging than writing their derived forms. Indeed, the HNC results demonstrated a different trend with verbs being least successful, while nouns were still leading. So why do noun derivatives tend to be more successfully produced than others? Perhaps it could be that the target noun derivatives contain high-frequency affixes, which students might have drawn on to form derived words. Furthermore, as learners know more lexical items in a certain POS category (noun, adjective, verb, adverb), they are more likely to succeed in producing its derived forms (Nguyen et al., 2025). Therefore, the greater success with nouns could be because the learners might have had a larger noun-related vocabulary. Yet, this is only speculative and awaits further research to single out the effects of affix frequency and learners' vocabulary breadth on derivative production in a contextualized test.

Derivative errors in argumentative essays and their relationship with derivative knowledge measured by the contextualized derivative recall test

The findings revealed three main categories of derivative errors that the participants committed in their essays, namely misaffixation, misuse of POS, and mixed errors. Wrong affix usage in which suffix errors (86.44%) were more prevalent than erroneous use of prefixes (13.56%), generally corroborating previous research (e.g., Ali et al., 2021; Ibsheeyah et al., 2023; Kusumawardhani, 2018). The affixal errors show that these students were aware of the range of prefixes (e.g., *in-*, *un-*, *ir-*, *dis-*,...) and suffixes (e.g., *-ful*, *-ness*, *-tion*) to form new words. At the same time, they suggest intralingual factors such as overgeneralization or overinclusion (e.g., *joyness*, *participaters*) that lead to errors. The issue of suffixes could be attributed to a tendency to apply them based on phonetically or orthographically similar words, (e.g., *pedagogical*/*pedagogical*; *benefical*/*beneficial*), which perhaps resulted in formulations of non-existent words. Indeed, the complexity of the English derivative system renders it challenging for learners. The ability to recall the intended derivatives that align with the meanings being expressed could be taxing in a meaning-making task like writing an argumentative essay. This could be because the learners have to retrieve accurate lexical items to express intended messages under planning time constraints, a trade-off between accuracy and fluency prioritizing the latter might have occurred due to the L2 users' limited attentional capacity (Skehan, 2015), which possibly resulted in misuse.

The most prevalent was misuse of POS, accounting for more than 50% of the total derivational errors. Notably, adjective and noun misuse was more frequent than others, which echoed earlier studies (e.g., Johnson et al., 2016). On the one hand, the high occurrence of POS misuse could be attributed to the insufficient knowledge of word formation rules or a lack of clarity regarding the syntactic roles of different word parts (Leontjev et al., 2016). It is even more challenging to recall or create new word classes that could be used to express their intended meaning in essay writing. Furthermore, while system-based learning is more likely with higher proficiency learners (Iwaizumi & Webb, 2022, 2023), for the third-year English-major

students in the present study with quite a relatively high proficiency level (low B2 to low C1), it is quite surprising that POS misuse was common. It is likely that POS errors could be attributed to the influence of students' Vietnamese L1 as an isolating language with no derivative application, resulting in interlingual errors such as omission of affixes (e.g., "*can free develop their skill*" [freely]).

The findings revealed that there was a positive relationship between students' derivative recall test scores and the derivative errors they made in their essays only in the HC method, but the effect size was very small. Together with the non-significant correlation between the overall derivative recall test scores and the derivative errors in essays in the HNC method, the findings suggest that the contextualized recall test and use of derivatives in essay writing might measure different constructs. These findings are both surprising and interesting and they could be explained by several reasons. Firstly, the contextualized recall test assessed students' ability to recall and generate the derivational forms within the immediate, local context of simple sentences. In contrast, essays represent larger discourse which requires students to make their own meanings and this might account for errors that occurred. The different linguistic demands of essay writing bounded by the given topic and the argumentative genre may necessitate different ways of accessing and mobilizing one's lexical resource to express personal viewpoints. In contrast, in the contextualized recall test, the meaning of each target derived form is predetermined by the given sentential context that elicits one single correct form or a very limited number of acceptable forms. The derivative recall test could thus be viewed as a form-focused activity that aims at the accuracy of derived forms whereas essay writing, a type of meaning-focused activity (Nation, 2020), allows students their free choice to use any derivatives at their disposal to express their own meanings, thus increasing the likelihood of committing errors. To further complicate matters, learners' English proficiency and their idiosyncratic ways of vocabulary use might have widened the gap between the ability to produce target derivatives in a contextualized derivative recall test and their use in written communication. In addition, even though the students were allowed to write their essays and complete the derivative recall test in as much time as they wished, the untimed recall test that targeted short and simple sentences to maximize understanding and derived forms related to given headwords might have aided the accurate production of the intended derivatives to a greater extent than the untimed writing performance. In essay writing, students' use of derivatives might be influenced by a wide range of factors including their own meanings to express. As such, the derived forms elicited by the contextualized test were not necessarily the derivatives that students actually needed to convey their ideas in the essays. In addition, the peer pressure created by early finishers, together with the meaning-driven act of argumentative writing might have led to a trade-off between form accuracy and meaning conveyance (Skehan, 2015), prioritizing the latter, thus prompting errors. Further research is clearly worthy to obtain more insights into the relationship between use of derivatives in meaning-focused tasks such as essay writing and their productive derivative knowledge measured via derivative recall tests.

IMPLICATIONS AND CONCLUSIONS

The present study i) measured Vietnamese EFL learners' productive derivative knowledge via the contextualized derivative recall test (Iwaizumi & Webb, 2021) and ii) explored its relationship with misuse of derivatives in written argumentative essays. The findings revealed that the Vietnamese EFL learners were successful at about 52% and 43% of the time in the HC and HNC methods respectively. The low production rates suggest that producing derived forms of the given headwords could be challenging for these learners and the success rates were dependent upon the scoring methods (HC vs. HNC). The results additionally indicated that the learners committed a wide range of derivative errors in their argumentative essays, of which, misuse of affixes and POS were most common. Above all, the higher the contextualized derivative test scores in the HC method, the more derivational errors made in essays. However, no such significant relationship was found in the HNC approach.

These findings have several important pedagogical implications for L2 derivative instruction and derivative use in extended discourse. The first implication could be to engage EFL learners in systematically learning derivational affixes and their meanings to develop a more profound comprehension of word formation patterns. The Word Part Levels Test (Sasao & Webb, 2017) which measures L2 learners' ability to recognize existing affixes and their meanings could be a useful platform to increase their awareness of derivational affixes and aid their production of derived words. In addition, as instruction has a significant impact upon derivative acquisition (Matwangsang & Sukying, 2023; Sumalee & Sukying, 2024), appropriate derivative teaching with word formation activities would allow learners to learn derived forms. Besides, extensive reading could be useful for students to encounter derivatives in contexts and increase their heightened awareness of how words are formed and used. Furthermore, corpus-based learning, data-driven learning from large corpora, for example, learning lexical collocations or word association through *Concordancers* (<https://www.lex tutor.ca/conc/eng/>) would help increase exposure to L2 input in authentic channels. It has been documented that exposure to meaning-focused input enhances morphological knowledge (Bauer & Nation, 2020), which in turn aids the development of language proficiency (Kieffer & Lesaux, 2012; Leontjev et al., 2016). In addition, it is pedagogically important that teachers identify the specific POS-related errors students make, and address them in subsequent lessons via post-writing correction/feedback that explains why, for example, a noun rather than other parts of speech is needed in that particular piece of text. It might be useful to teach students the syntactic functions of affixes so that they can avoid using incorrect derived forms in their writing. A further implication is related to the use of the contextualized derivative recall test. Since the knowledge of derived forms elicited through the contextualized recall test is not equated with their use in extended discourse, such as essay writing, the contextualized recall test should be interpreted with caution in order to draw valid conclusions about learners' derivative knowledge. While deliberate learning from form-focused activities via the decontextualized/contextualized derivative test might help expand L2 learners' vocabulary knowledge, opportunities to use derived forms in meaning-focused output activities (Nation, 2020) are clearly needed to develop automaticity and fluency of use, enabling 'learning- to-use' abilities (Suzuki, 2024) to achieve communicative purposes. Finally, the different results associated with the different scoring methods (HC and HNC) in the present research accentuate the necessity for instructors and researchers to

employ both methods to score productive derivative recall tests to better measure L2 learners' ability to produce derived forms of the target headwords.

Despite its significant contributions, the study has some limitations that are worth acknowledging. Firstly, the sample size of 88 English-major students from one Vietnamese university may restrict the generalizability of the findings. Future research could thus employ a larger sample, considering variables such as proficiency levels and educational background in other Vietnamese contexts and beyond. Next, a follow-up analysis of the derivatives of specific headwords produced correctly and whether they are associated with higher frequency levels would provide more explanatory power to interpret the derivative recall test outcomes at the different word frequency levels. In addition, future research could track cases of unsuccessful production in the derivative recall test and explore the link between them and erroneous derivative use in essay writing. Exploring the relationship between receptive derivative knowledge and productive derivative knowledge as measured via a contextualized format would be a further research direction (cf. Iwaizumi & Webb, 2023; Nguyen et al., 2025). Furthermore, fine-grained qualitative analysis of derivative use in written texts produced by L2 learners would further our understanding of how L2 derivative knowledge is acquired. Thirdly, the findings regarding the correlation between the derivative test scores and derivational errors were only exploratory as a first study. It focused on only derivative errors, not in light of the wider lexical use and derivative distribution in argumentative essays. Future research with additional foci and other writing genres would provide a more subtle understanding of how the use of derived forms in written texts is linked to their productive derivative knowledge as measured via derivative recall tests.

THE AUTHORS

Phuong Nhi Tran earned her BA in Teaching English as a Foreign Language (TEFL) from the University of Foreign Languages and International Studies, Hue University, Vietnam. Her enthusiasm lies in teaching English to young students and investigating learners, language production, as well as EFL language teaching and learning methods.
phuongnhi036@gmail.com

Pham Thi Hong Nhung obtained her MA and PhD in Applied Linguistics from the University of Queensland, Australia. She has taught various postgraduate courses for the Faculty of English at the University of Foreign Languages and International Studies, Hue University, Vietnam, and has published in the field of applied linguistics and language education.
n.pham@hueuni.edu.vn

Huy Van Nguyen is a lecturer in Applied Linguistics and TESOL at the University of Foreign Languages and International Studies, Hue University, Vietnam. He earned his PhD in Language Education from the University of Queensland, Brisbane, Australia. His main research interests lie in language planning and policy (LPP) with a special focus on language-in-education policy, English as a Medium of Instruction (EMI), and educational leadership in developing countries.
nguyenvanhuy@hueuni.edu.vn

Bao Trang Thi Nguyen is a lecturer and researcher in TESOL and Applied Linguistics at Hue University, Vietnam. Her research interests include task-based language teaching and learning, teacher cognition, vocabulary acquisition, language and culture, learner proficiency and SLA. She has also written book chapters published by John Benjamins, Bloomsbury, Springer and Routledge. Her research articles have also been published in journals such as *Language Teaching Research*, *System*, *TESOL Journal*, *RELC Journal*, *Asia Pacific Journal of Education*, *International Journal of Comparative Education and Development*, *Language Related Research and Language Testing in Asia*.
ntbtrang@hueuni.edu.vn

REFERENCES

- Ali, M., Hamid, A., & Rahman, G. (2021). Acquisition of English derivational morphemes by students at selected universities in Khyber Pakhtunkhwa, Pakistan. *Liberal Arts and Social Sciences International Journal*, 5(1), 338–355. <https://doi.org/10.47264/idea.lassij/5.1.22>
- Al-Khresheh, M. H. (2016). A review of error analysis study. *International Journal of Humanities and Social Sciences Research*, 2, 49–59. <https://doi.org/10.6000/2371-1655.2016.02.05>
- Bauer, L., & Nation, I. S. P. (2020). *English morphology for the language teaching profession*. Routledge. <https://doi.org/10.4324/9780367855222>
- Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: A synthesis of the research on teaching words, phrases, and patterns*. Routledge. <https://doi.org/10.4324/9781003005605>
- Chan, A. Y. (2010). Toward a taxonomy of written errors: Investigation into the written errors of Hong Kong Cantonese ESL learners. *TESOL Quarterly*, 44(2), 295–319. <https://doi.org/10.5054/tq.2010.219941>
- Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*. Oxford University Press.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. Sage.
- Graves, M. F., Elmore, J., & Fitzgerald, J. (2019). The vocabulary of core reading programs. *The Elementary School Journal*, 119(3), 386–416. <https://doi.org/10.1086/701653>
- Hemchua, S., & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect*, 21(3), 3–25.
- Heydari, P., & Bagheri, M. S. (2012). Error analysis: Sources of L2 learners' errors. *Theory and Practice in Language Studies*, 2(8), 1583–1589. <https://doi.org/10.4304/tpls.2.8.1583-1589>
- Ibsheeyah, F. F., Senowarsito, S., & Nur'Aini, S. (2023). Derivational morpheme errors in English: A study of Libyan students at Al-Mergib University. *Voices of English Language Education Society*, 7(2), 361–372. <http://dx.doi.org/10.29408/veles.v7i2.21305>
- Ishii, T., & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC Journal*, 40(1), 5–22. <https://doi.org/10.1177/0033688208101452>
- Iwaizumi, E., & Webb, S. (2021). To what extent does productive derivational knowledge of adult L1 speakers and L2 learners at two educational levels differ? *TESOL Journal*, 12(4), 1–22. <https://doi.org/10.1002/tesj.640>
- Iwaizumi, E., & Webb, S. (2022). Measuring L1 and L2 productive derivational knowledge: How many derivatives can L1 and L2 learners with differing vocabulary levels produce? *TESOL Quarterly*, 56(1), 100–129. <https://doi.org/10.1002/tesq.3035>
- Iwaizumi, E., & Webb, S. (2023). To what extent do learner- and word-related variables affect production of derivatives? *Language Learning*, 73(1), 301–336. <https://doi.org/10.1111/lang.12524>
- Iwaizumi, E., & Webb, S. (2024). The effect of test format on productive recall of derivatives. *ITL-International Journal of Applied Linguistics*, 175(2), 242–270. <https://doi.org/10.1075/itl.23002.iwa>

- Johnson, M. D., Acevedo, A., & Mercado, L. (2016). Vocabulary knowledge and vocabulary use in second language writing. *TESOL Journal*, 7(3), 700–715. <https://doi.org/10.1002/tesj.238>
- Kieffer, M., & Lesaux, N. (2012). Direct and indirect roles of morphological awareness in the English reading comprehension of native English, Spanish, Filipino, and Vietnamese speakers. *Language Learning*, 62(4), 1170–1204. <https://doi.org/10.1111/j.1467-9922.2012.00722.x>
- Kusumawardhani, P. (2018). The error analysis of derivational morphology in EFL's English narrative composition. *International Journal of Language Education*, 2(1), 22–30.
- Laufer, B. (1991). Some properties of the foreign language learner's lexicon as evidenced by lexical confusions. *International Review of Applied Linguistics in Language Teaching*, 29(4), 317–348. <https://doi.org/10.1515/iral.1991.29.4.317>
- Laufer, B. (2017). From word parts to full texts: Searching for effective methods of vocabulary learning. *Language Teaching Research*, 21(1), 5–11. <https://doi.org/10.1177/1362168816683118>
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322. <https://doi.org/10.1093/applin/16.3.307>
- Laufer, B., Webb, S., Kim, S. K., & Yohanan, B. (2021). How well do learners know derived words in a second language? *ITL-International Journal of Applied Linguistics*, 172(2), 229–258. <https://doi.org/10.1075/itl.20020.lau>
- Leontjev, D., Huhta, A., & Mantyla, K. (2016). Word derivational knowledge and writing proficiency: How do they link? *System*, 59, 73–89. <https://doi.org/10.1016/j.system.2016.03.013>
- Leontjev, D., Huhta, A., & Tolvanen, A. (2023). L2 English vocabulary breadth and knowledge of derivational morphology: One or two constructs? *Language Testing*, 40(2), 300–324. <https://doi.org/10.1177/0265532221114015>
- Llach, M. P. A. (2011). *Lexical errors and accuracy in foreign language writing*. Multilingual Matters.
- Llach, M. P. A. (2017). Vocabulary teaching: Insights from lexical errors. *TESOL International Journal*, 12(1), 63–74.
- Li, X. (2015). Analysis on lexical errors in writings of Mongolian English majors. *Theory and Practice in Language Studies*, 5(12), 2565–2570. <http://dx.doi.org/10.17507/tpls.0512.18>
- Matthews, J., Milliner, B., & McLean, S. (2023). Can learners understand words with derivational affixes and does presence of context make a difference? *RELJ Journal*. Advance online publication. <https://doi.org/10.1177/00336882231222034>
- McCutchen, D., & Stull, S. (2015). Morphological awareness and children's writing: accuracy, error, and invention. *Reading and Writing*, 28(2), 271–289. <https://doi.org/10.1007/s11145-014-9524-1>
- Matwangsang, R., & Sukying, A. (2023). The effects of morphological awareness on L2 vocabulary knowledge of Thai EFL young learners. *World Journal of English Language*, 13(2), 51–63. <https://doi.org/10.5430/wjel.v13n2p51>
- Nation, I. S. P. (2012). The BNC/COCA word family lists. <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139858656>
- Nation, I. S. P. (2020). Is it worth teaching vocabulary? *TESOL Journal*, 12(4), Article e564. <https://doi.org/10.1002/tesj.564>
- Nation, I. S. P. (2021). Thoughts on word families. *Studies in Second Language Acquisition*, 43(5), 969–972. <https://doi.org/10.1017/S027226312100067X>
- Nguyen, B. T. T., & Nguyen, L. Q. (2024). Measuring productive derivative knowledge of Vietnamese EFL learners: The role of headword scoring, vocabulary breadth, and headword familiarity. *System*, 123, Article 103322. <https://doi.org/10.1016/j.system.2024.103322>

- Nguyen, B. T. T., Nguyen, T. H. H., Nguyen, L. Q., Vo, L.-H., & Dung, D. T. X. (2025). Vietnamese EFL learners' productive derivative knowledge: The role of part of speech, receptive derivative knowledge, and derivative-formation strategies. *Journal of Research in Applied Linguistics*, 16(1), 3–18. <https://doi.org/10.22055/RALS.2024.46383.3252>
- Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, 1(1), 12–30. <https://doi.org/10.1177/1362168815586083>
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, 19(1), 17–36. <https://doi.org/10.1017/S0272263197001022>
- Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, 36(2), 145–171. <https://doi.org/10.2307/3588328>
- Skehan, P. (2015). Limited attention capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 123–156). John Benjamins.
- Snoder, P., & Laufer, B. (2022). EFL learners' receptive knowledge of derived words: The case of Swedish adolescents. *TESOL Quarterly*, 56(4), 1242–1265. <https://doi.org/10.1002/tesq.3101>
- Sukying, A. (2018). Investigating receptive and productive affix knowledge in EFL learners. In D. Hirsh (Ed.), *Explorations in second language vocabulary research* (pp. 183–218). Peter Lang.
- Suzuki, Y. (2024). Skill acquisition theory: Learning-to-use and usage-for-learning in SLA. In K. McManus (Ed.), *Usage in second language acquisition: Critical reflections and future directions* (pp. 147–168). Routledge.
- Sumalee, P., & Sukying, A. (2024). The effects of derivational suffix instruction on English vocabulary knowledge in Thai high school learners. *rEFlections*, 31(2), 335–352. <https://doi.org/10.61508/refl.v31i2.272906>
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Erlbaum.
- Yin, R. K. (2015). *Qualitative research from start to finish* (2nd ed.). The Guilford Press.



Appendix A

Task prompts

Task 1:

*'Salary is the most important thing to consider in choosing a job.' Discuss.
Write a for and against essay. You should write at least 250 words.*

Task 2:

*'Young children should not participate in TV game shows.' Discuss.
Write a for and against essay. You should write at least 250 words.*

Appendix B
Target headwords of the contextualized derivative recall test
(Iwaizumi & Webb, 2021, p. 8)

1000 (k = 10)	2000 (k = 10)	3000-5000 (k = 10)		
		3000	4000	5000
Art	Tradition	Formal	Communicate	Interrogate
Depend	Theater	Differ	Obsess	Anatomy
Arrange	Identity	Publish	Sincere	Plausible
Girl	Alter	Persuade		
Forget	Measure			
Protect	Indicate			
Child	Active			
History	Intense			
Health	Vary			
Possible	Polite			

Note. The total number of possible derivatives at the 1000, 2000 and 3000-5000 levels is 41, 37 and 40 respectively. The total of all is 118.

Appendix C

Correlations between the derivative errors made by students and their productive derivative knowledge in the HC method (N = 88)

Spearman's rho correlation coefficient		Noun	Adjective	Verb	Adverb	Total derivative test score	Total derivative errors in essays
Noun	r_s	1.000	.829**	.720**	.777**	.921**	.218*
	Sig. (2-tailed)	.	.000	.000	.000	.000	.042
Adjective	r_s	.829**	1.000	.737**	.853**	.956**	.178
	Sig. (2-tailed)	.000	.	.000	.000	.000	.098
Verb	r_s	.720**	.737**	1.000	.727**	.831**	.348**
	Sig. (2-tailed)	.000	.000	.	.000	.000	.001
Adverb	r_s	.777**	.853**	.727**	1.000	.899**	.217*
	Sig. (2-tailed)	.000	.000	.000	.	.000	.042
Total derivative test score	r_s	.921**	.956**	.831**	.899**	1.000	.225*
	Sig. (2-tailed)	.000	.000	.000	.000	.	.035
Total derivative errors in essays	r_s	.218*	.178	.348**	.217*	.225*	1.000
	Sig. (2-tailed)	.042	.098	.001	.042	.035	.

**Correlation is significant at the .01 level.

Appendix D

Correlations between the derivative errors made by students and their productive derivative knowledge in the HNC method (N = 88)

Spearman's rho correlation coefficient (r_s)		Noun	Adjective	Verb	Adverb	Total derivative test score	Total derivative errors in essays
Noun	r_s	1.000	.802**	.496**	.765**	.914**	.199
	Sig. (2-tailed)	.	.000	.000	.000	.000	.063
Adjective	r_s	.802**	1.000	.606**	.818**	.943**	.135
	Sig. (2-tailed)	.000	.	.000	.000	.000	.209
Verb	r_s	.496**	.606**	1.000	.519**	.607**	.098
	Sig. (2-tailed)	.000	.000	.	.000	.000	.364
Adverb	r_s	.765**	.818**	.519**	1.000	.910**	.217*
	Sig. (2-tailed)	.000	.000	.000	.	.000	.042
Total derivative test score	r_s	.914**	.943**	.607**	.910**	1.000	.183
	Sig. (2-tailed)	.000	.000	.000	.000	.	.089
Total derivative errors in essays	r_s	.199	.135	.098	.217*	.183	1.000
	Sig. (2-tailed)	.063	.209	.364	.042	.089	.

**Correlation is significant at the .01 level.