Vietnamese EFL Students’ Use and Misconceptions of Cohesive Devices in Writing

Hung Phu Bui

Abstract
As part of a larger project, this study explores Vietnamese college students’ use and concepts of cohesive devices in writing. Cohesion is a crucial element that ties components of a text together. Although the use of cohesive devices in L2 writing has been investigated by a large body of research, there is no such study exploring the effects of L2 learners’ misconceptions of this word class. One hundred sixty-eight academic reports of totally 67,400 words written by Vietnamese final-year English-majored undergraduate students were collected for data analysis. An email invitation was sent, and 23 students accepted to participate in semi-structured interviews which were audio-recorded for analysis. The findings showed that the students employed references, conjunctions, and lexical items the most frequently in writing. Interestingly, the students’ cohesion use and errors had similar patterns. The transcript analysis showed that the students’ misconceptions of some language items and writing requirements affected their choice of cohesive devices. The current study gives implications for teaching and learning of this word class in L2 contexts.

Keywords
cohesion, discourse analysis, error analysis, L2 writing, misconception, second language acquisition

Introduction
Recent studies have provided results about Asian second language (L2) learners’ difficulties with using cohesive devices (Al-Jarf, 2001; Chiang, 1999; Crossley & McNamara, 2012; Crossley et al., 2007, 2016; Eng et al., 2020; Hammad, 2016; Lake, 2004; Liu & Braine, 2005; Olateju, 2006; Ong, 2011; Rahman, 2013; Tanskanen, 2006; Todd et al., 2007; Yang & Sun, 2012; Yule, 2008; Zhang, 2000). These studies have examined the relationship between L2 learners’ use of cohesive devices and such factors as writing quality (Liu & Braine, 2005; Todd et al., 2007), proficiency levels (Chiang, 1999; Zhang, 2000), syntactic and semantic repertoire (Al-Jarf, 2001; Crossley & McNamara, 2012), and first language (L1) interference (Blagoeva, 2004; Kang, 2005). Some controversial results can be detected from these studies due to their methodological flaws. For instance, while the study by Chiang (1999) found that there was a strong relationship between L2 learners’ use of cohesive devices and proficiency level, Crossley and McNamara (2012) concluded from their research that such correlation was insignificant. Although Ong (2011) attempted to total up cohesion errors made by a homogeneous group of Chinese English L2 learners, this study could not explain the causes of the identified errors. Rahman’s (2013) endeavor tried to relate L2 learners’ errors to their proficiency level, but the research results were mainly based on counting and contradicted with the studies by Al-Jarf (2001), Crossley and McNamara (2012), and Zhang (2000) regarding the relationship between L2 learners’ use of cohesive devices and proficiency levels.

Although cohesive devices used by L2 students were vastly explored, studies of this research strand were mainly interested in analyzing texts or delved into the relationship between students’ cohesion use and language competence. The literature in second language acquisition shows that it is L2 students’ conceptualizations that are responsible for their language use (Alonso et al., 2016; Ellis, 2015; Langacker, 2008) because the writer encodes his or her knowledge and socio-cultural background in their message (Edwards, 2013; Hung & Khoa, 2022; West & Turner, 2007). In particular, L2 students construe the configurations of a lexical item acquired and then store it in their memory in the form of knowledge or inventory, from which their conceptions are established. Accordingly, meanings are encoded and then

1University of Economics Ho Chi Minh City, Vietnam

Corresponding Author:
Hung Phu Bui, University of Economics Ho Chi Minh City, 279 Nguyen Tri Phuong Street, Dist. 10, 59C Nguyen Dinh Chieu Street, Dist. 3, Ho Chi Minh City 084, Vietnam.
Email: hungbp@ueh.edu.vn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
projected as mental representation “as construed by the human mind, mediated by our unique perceptual and conceptual systems” (Evans & Green, 2006, p. 7). In other words, it is the learner’s conceptualization that is attributed to their understanding and language use (Baddeley, 2010; Hung, 2017; Karim & Nassaji, 2013; Langacker, 2008; Tyler, 2012). L2 learners’ word choice reflects their understanding and conceptualization of the lexical and grammatical units they use (Nuyts & Pederson, 2012; Papafragou & Grigoroglou, 2019; Thierry, 2016). An investigation into L2 learners’ misconceptions of English cohesive devices would shed useful light on their difficulties using this word class and contribute to the literature in text cohesion and error analysis. Research results would give implications for learning and teaching this word class.

The current study aims to examine the relationship between Vietnamese English as a foreign language (EFL) learners’ choice of cohesive devices used in writing and their misconceptions of this word class. As such, the errors of cohesive devices used in writing by Vietnamese EFL learners were first analyzed and their misconceptions of the cohesive devices were subsequently identified through semi-structured interviews to find the interference of their misconceptions in their choice of cohesive devices.

**Literature Review**

**Cohesion and Coherence**

A text is a means of communication, spoken and written, which indicates the development, organization, and connection of ideas. In the context of the text, the speaker or writer incorporates their socio-cultural accounts, personal knowledge and experience, and coding of such information into a unified beyond-sentence communication, in which cohesion and coherence are interwoven (Collins & Hollo, 2010; Halliday & Matthiessen, 2014; Reynolds, 2001; Sanders & Maat, 2006). In other words, cohesion and coherence are text-dependent which are used to tie parts of a text and make it comprehensible. These two indispensable elements of a text are supposed to have an unclear boundary because they both rely on linguistic elements and together construct meaning. The absence of either of them can disrupt the intelligibility of the text (Cox et al., 1990; Malmkjaer, 2001). Thus, it is difficult to distinguish one element from the other as they are considered intersected concepts.

Some linguists have attempted to make a distinction between coherence and cohesion. As Rahman (2013, p. 3) put it, these two elements demonstrate the “two faces of the same coin.” Oller and Jonz (1994) and Reynolds (2001) posit that cohesion does not always contribute to text coherence. While coherence is a means that the writer uses to persuade and interact with the reader and helps the reader interpret the text (Malmkjaer, 2001; Reynolds, 2001; Sanders & Maat, 2006), cohesion functions as glue that connects different parts of a text (Alarcón & Morales, 2011; Salkie, 1995). To illustrate this notion, Enkvist (1990, p. 12) gave the example, “My car is black. Black English was a controversial subject most people have retired. To retire means ‘to put new tires on a vehicle’. Some vehicles such as hovercraft have no wheels. Wheels go round.” Although this excerpt shows the use of cohesive devices, typically repetition of words, it does not make sense to the reader. As noted by Koshik (1999, p. 11), a coherent text can be comprehensible to readers even when it is lacking in cohesion as in the example, “Someone came my house. Says give me money. Husband take gun shoot. Go outside die. Call police. Emergency 911. Policeman come. Take black man go hospital die.” Although cohesion is not detected in this example, it is coherent with a series of sequential incidents. To further provide a clear-cut distinction between cohesion and coherence, Malmkjaer (2001) and Sanders and Maat (2006) argued that cohesion refers to the use of cohesive devices or cohesion ties, which make a text become a unified structure. In general, cohesion mainly deals with the utilization of lexical syntax and grammar, whereas coherence focuses more on the logical organization of ideas to form meaningful and comprehensible content. The aforementioned argument sets the grounding for distinguishing cohesion from the coherence in text analysis.

**Taxonomy for Analyzing Cohesive Devices**

The lexico-grammatical system of cohesion outlined by Halliday and Matthiessen (2014) provides a taxonomy for analyzing cohesive devices. It classifies cohesive devices into five main types with distinct functions: (1) reference, (2) substitution, (3) ellipsis, (4) conjuncts, and (5) lexical cohesion (see Figure 1). From the critical discussion above, this study considers cohesive devices as indications of cohesion, which a text requires.

Reference is defined as the use of a linguistic item to refer to a word that precedes or follows it and is mainly subcategorized as exophora and endophora. Exophora occurs when a word refers to an out-of-text item. A speaker, for example, says “this” or “it” when he or she points at an object as in “I like this.” or “I like it.” The listener needs to understand the situation in which the utterance takes place. In contrast, endophoria refers to an item within the text. Endophora can be anaphora or cataphora. While anaphora refers to a preceding item, cataphora refers to the following item. Anaphora makes up a major part of endophora. As exophora is a nonlinguistic element, it is not usually considered in text analyses.

Substitution is conceptualized as the use of a word or phrase to substitute a linguistic item. The substitution can take place at the nominal, verbal, or clausal level. Like substitution, ellipsis consists of three main types, including nominal, verbal, and clausal. Although substitution and ellipsis are similar in terms of classification, their distinction is
that ellipsis involves omission rather than substitution of linguistic items.

Conjunction refers to the use of adverb conjuncts or connectives to connect sentences as in “There is no easy road to success. Individuals usually encounter obstacles. Therefore, we must act together.” The adverb conjunct “therefore” shows a causal relation. There are four types of adverb conjuncts: additive, adversative, causal, and temporal.

Lexical cohesion is the repetition of words in a previous sentence or the use of synonyms or words of the same category or class. Lexical cohesion is of four types: synonymy, repetition, hyponymy, and meronymy. Figure 1 provides detailed descriptions and examples of each type and sub-type of cohesive devices.

**L2 Learners’ Use of Cohesive Devices in Writing**

The use of cohesive indices in writing by L2 learners has intrigued a large body of research. These endeavors have attempted to investigate the relationships of this word class and other factors. These studies can be classified into five main research strands.

A survey of literature shows that two studies were interested in exploring the use of connectives by a homogeneous group of L2 learners. Liu and Braine (2005) discovered that 96 Chinese first-year college students of English were generally incapable of using cohesive instances properly in their argumentative essays. Similarly, Olateju (2006) analyzed essays written by 70 senior high school students in Nigeria and found that their writings exposed cohesion problems despite their 6 years of learning English. These two studies recommended raising learners’ awareness of the importance of cohesion in the written text.

Another research strand investigated whether L2 learners’ proficiency levels had impacts on the cohesion of their texts. Xuefan (2007) analyzed essays written by two groups of L2 learners at different levels of proficiency and found that proficiency levels did not have a strong impact of their use of

---

**Figure 1.** Cohesive devices (Rahman, 2013, p. 5).

| 1. Reference | Exophora | Reference to a nonlinguistic element, e.g., (a teenager listening to loud music) Father: Stop doing that, I want to read. |
| 2. Substitution | The use of pro-forms and pro-VPs to establish cohesion, e.g., the morning paper didn’t carry a story about the robbery, but the evening paper has one. |
| 3. Ellipsis | Cohesion can be established through ellipsis, which is to leave out a word or phrase rather than repeat it, e.g., I’ve drunk a lot of coffee in my time, but this is the worst I’ve ever tasted. |
| 4. Connectives | Additive | and, furthermore, for instance, likewise, etc. |
| 5. Lexical cohesion | Repetition | yet, in fact, however, on the other hand, instead, etc. |
| | Synonyms | so, therefore, as a result, because, etc. |
| | Superordinates | then, first, second, third, finally, in conclusion, etc. |
| | Antonyms | To achieve cohesion in texts, writers tend to use the same words. Specifically content words, such as nouns, adjectives, adverbs, and main verbs which are used by writers to help a text to be coherent. For instance, the people of this country aren’t stupid. They know when politicians are lying to them. They know when newspapers are not giving them the full picture. |
| | | This method is used to avoid repeating exactly the same word by using a word that is very close in meaning, for example. “The doctor told me I’d been working too hard and I needed at least six weeks off work to get my strength back.” Amanda’s employer, however, was less sympathetic. “My boss told me it was redundancy money—two weeks’ pay—$280. I was shocked.” |
| | | Another way of creating cohesion is to refer back to a word by using its superordinate. General words are referred to as “superordinates” and the more specific ones are called “hyponyms.” Such as, I love all dogs, especially Collies. |
| | | Antonyms, according to O’Grady, Dobrovolsky, and Katamba (1996), are “words or phrases that are opposites with respect to some components of their meanings” (p. 705). For instance, that car is big, compared to the small one I have.
lexical indices of cohesion. Crossley and McNamara (2012) conducted a larger project in which the use of cohesion and linguistic sophistication in 514 essays by L2 learners were analyzed. The results showed that students’ proficiency levels did not directly influence their text cohesion. This study analyzed a larger scale of cohesion devices and confirmed the finding found by Xuefan (2007).

Intrigued by the concept that the proficiency level is too broad to describe writing competencies, several other studies traced further to the relationship between L2 learners’ cohesion use and writing quality. Chiang (1999) analyzed 172 essays written by heterogeneous L2 English college students. These participants were at the beginning and intermediate proficiency levels. The findings demonstrated that the more cohesive the compositions were, the higher they were rated. This finding was confirmed by Yang and Sun (2012) that L2 learners’ text cohesion was not directly correlated with their proficiency level but writing quality. However, the study by Todd et al. (2007) analyzed essays by Thai L2 English postgraduate students and concluded that there was a weak relationship between writing achievement and the use of connectives based on the tutors’ comments. This result was aligned with the study by Zhang (2000). The conflicting results on the relationship between L2 learners’ use of cohesion and writing quality in the aforesaid studies was due to their employment of different rating criteria.

To discover the difficulty encountered by L2 English learners in using cohesion ties, several studies conducted in different contexts around the world centered on error analysis. Al-Jarf (2001) analyzed cohesive devices used in academic writings by 59 EFL students at an Arabian university. The findings indicated that the biggest problem coped by the students was substitution, followed by reference and ellipsis. The researcher argued that students’ awareness of syntactic and semantic features of cohesive devices rather than levels of proficiency helps constitute text cohesion in writing. However, Ong (2011) concluded that Chinese EFL students’ use of references was the biggest problem. In a similar vein, Rahman (2013) concluded that Omani EFL first- and third-year students employed lexical cohesion the most frequently, followed by references and conjunctions; native speakers, in contrast, used substitution and ellipsis more frequently than Omani EFL students. Also, although the students had received extensive English instruction for several years, they misused and overused cohesive devices.

A limited body of research investigated the interference of the first language (L1) in the use of L2 English cohesive instances. Blagoeva (2004) concluded that overuse of reference was the main problem in essays written Bulgarian L2 English students due to L1 interference. Kang (2005) compared the use of cohesive devices by Korean L2 English students and American students. The findings also revealed that Korean overused reference in their writing. Although Blagoeva (2004) and Kang (2005) conducted their studies in different contexts, they both concluded that reference overuse was one of the biggest problems among L2 English learners. However, like the other previous studies, these two studies did not explore the causes of writing problems.

Previous research on the use of cohesive devices mainly relied on discourse analysis. Although several studies attempted to tackle the problems encountered by L2 English learners in using cohesion ties in writing, they only totaled up the errors identified in written texts. As such, they failed to satisfactorily interpret the causes of identified errors. Recent research in second language acquisition has shown that L2 learners’ choice of semantic and grammatical units at word, phrasal, and discourse levels is regulated by human cognition (Leacock et al., 2010; White, 2010). L2 speakers can select from their accumulated repertoire of the linguistic items that they think best fit the context.

Conceptualizations first take place in L1 and then transfer to L2 (Alonso et al., 2016; Croft & Cruse, 2004; Ellis, 2015; Evans & Green, 2006; Langacker, 2008; Radden & Dirven, 2007; Robinson & Ellis, 2008) because the writer encodes his or her knowledge and socio-cultural background in their message (Edwards, 2013; Sanders & Maat, 2006; West & Turner, 2007). From their experiences, people construe the configurations of a lexical item obtained and then store it in their memory in the form of knowledge or inventory, from which learners’ perception and subsequentially conception are established (see Figure 2). Accordingly, meanings are encoded and then projected as mental representation (Hung, 2017) “as construed by the human mind, mediated by our unique perceptual and conceptual systems” (Evans & Green, 2006, p. 7). In other words, it is the learner’s conceptualization that is attributed to their understanding and language use (Baddeley, 2010; Hung, 2019; Karim & Nassaji, 2013; Langacker, 2008; Tyler, 2012). In particular, L2 learners’ word choice reflects their understanding and conceptualization of the lexical and grammatical units they use. In short, language allows humans to acquire language and convey their meanings through the construction of conceptualizations (Evans & Green, 2006; Fitch, 2011; Hung et al., 2018; Nyuts & Pederson, 2012; Papafragou & Grigoroglou, 2019; Thierry, 2016).

As cohesive devices serve to tie components of a text, they represent the users’ understanding of the relationship between these components (Halliday & Matthiessen, 2014; Malmkjæra, 2001; Rahman, 2013; Reynolds, 2001; Sanders & Maat, 2006). This means that the use of cohesion ties indicates the user’s conception (Sanders & Maat, 2006) at both the lexical and discourse levels (Connor, 1984; Cox et al., 1990; Halliday & Matthiessen, 2014). Understanding L2 learners’ use and misconceptions of cohesive devices will help suggest implications for teaching and learning of this word class (Ellis, 2015; Langacker, 2008). Cohesive devices used by L2 learners can be intervened by their L1 (Hung et al., 2021; Rahman, 2013).
Research Aims and Questions

This study seeks to address Vietnamese final-year English-major college students’ difficulty in using cohesive devices in writing. It attempts to address the following research questions (RQs):

1. What cohesion error patterns do Vietnamese final-year English majors often commit in writing?
2. What misconceptions lead Vietnamese final-year English majors to difficulty in using those cohesive devices?

Research Methods

Research Approach and Design

The current study was both qualitative and quantitative. It analyzed Vietnamese final-year English-major college students’ use and conception of cohesive devices in writing as analyzing language use helps collect and interpret data in natural settings (Dezin & Lincoln, 2000; Dunne et al., 2005; Mackey & Gass, 2005). The study comprised of three main stages. First, the cohesive devices used in academic reports written by Vietnamese final-year English-major college students were collected and analyzed. Second, errors of cohesive devices were identified and then classified into types based on the taxonomy (see Figure 1). Finally, semi-structured interviews were administered with the participants selected from the first stage to delve into their misconceptions of the cohesive devices underlying the errors identified.

Participants, Materials, and Settings

This study employed the convenience sampling strategy to select participants. It collected 168 academic reports written by Vietnamese L2 English college seniors who were spending their internships at five different language centers in Ho Chi Minh City, Vietnam. The internship advisors and supervisors required the students to submit their reports which compared their practices to the perspectives they had learned at college. All participants had the C1 level certificate (Cambridge Advanced English) on Common European Framework References (CEFR). They had spent 4 years receiving extensive English instruction, including approximately 100 hours of grammar and 168 hours of academic writing (67 hours of paragraph writing, 67 hours of essay writing, and 34 hours of research writing). In both grammar and writing courses, the students were instructed in cohesion at the lexical and discourse levels. On implementing the current study, the management boards of the language centers and individual participants were contacted and consented. They were informed of the main purpose of the study and ethical considerations. Their identity was kept confidential. They were named with pseudonyms S1 to S23 in the data report.

Text Collection and Analysis

The data collection was interrupted by the outbreak of the Covid-19 pandemic in the place. The first half of students’ reports were collected at the beginning of 2020, and the second half was collected in October 2020. The text collection procedure consisted of three main stages.

The collected data comprised approximately 67,400 words of the reports written and submitted periodically by Vietnamese L2 English seniors during their internships. The main objective of the analysis was to point out the cohesion errors which were presumably influenced by the students’ L1 by contrasting their misused instances with those in L1. As such, it consisted of three main phases.

First, the cohesive devices used in the students’ writings were read through and detected. Afterward, the errors
detected in the texts were categorized into the types and sub-types in the taxonomy (see Figure 1). The inventory of errors by cohesion type and sub-type were documented.

Second, the errors were classified into types (see Table 1). The error categorization consisted of four main groups: misuse, redundancy, omission, and overabundance (Ong, 2011; Rahman, 2013) to figure out the error patterns committed by the participants.

Finally, the errors identified in the first step were contrasted to the students’ L1 equivalents. This step was to trace the participants’ possible misconceptions of the cohesive devices used in writing. Given the particular function of each type of cohesive instances and writer’s intention in their word choice, the percentage of each type was calculated to point out what cohesive devices constituted the highest rates of errors.

**Interviews**

This study employed the prompt interview strategy. An invitation email was sent to the participants. Twenty three out of 32 students accepted my invitation to participate in individual interviews. All semi-structured interviews, administered in December 2020, were conducted in the students’ L1 (Vietnamese), each of which lasted for 25 to 30 minutes. In the interviews, the participants were asked to explain in detail their choice of each cohesive device identified as problematic. They were shown the cohesive devices they used in their texts so that they could explain their choices. They were also asked to confirm and clarify their responses occasionally. All interviews were audio-recorded and subsequently transcribed and coded for analysis. Their coded answers were then sent back to them for confirmation.

Data collected from interviews were subject to thematical analysis. This study employed the qualitative data analysis procedure proposed by Creswell and Creswell (2018). This content-based iterative inductive procedure was composed of five main steps: transcribing, coding, rereading, refining, and interpreting. Emerging themes were identified and scrutinized.

**Results**

**RQ1: What Cohesion Error Patterns Do Vietnamese Last-Year English Majors Often Commit in Writing?**

Data analyses showed the cohesive devices and errors detected in writing by Vietnamese English-major seniors (see Table 2). They used lexical cohesion the most frequently, constituting 39.68% of all cohesive instances, followed by reference (37.48%) and conjunction (15.52%). The least frequently used type was ellipsis (3.31%). The cohesion use and errors had a relatively similar pattern. Surprisingly, substitution was employed quite marginally (4.01%) by the students, and its errors coincidentally accounted for a small percentage, at approximately over 1.5%. Overall, the cohesive devices used by the Vietnamese final-year English majors made up a small proportion of slightly less than 7.90% of the data set word count, and its errors amounted to nearly 10.05% of all cohesive devices used in their reports.

As substitution and ellipsis constituted a small percentage each, further analyses focused on the three groups with the highest rate of errors. An examination into each sub-type of cohesion demonstrated the error patterns in writing by the Vietnamese final-year English majors (see Table 3). Accordingly, they made more errors with definite articles, pronominal sub-type, and comparative sub-type of the reference group; additives, adversatives, and casual sub-types of conjunctions or connectives; and synonyms and repetition sub-types of lexical cohesion. These sub-types of cohesion constituted over 20% each. In general, the students showed their problems using particular types and sub-types of cohesion.
An investigation into types of error demonstrated the categories with high ratios or percentages of errors (see Table 4). In general, cohesion errors mainly fell into three categories: misuse (MU), redundancy (RD), and omission (OM) (see Table 4). Misuse and omission of necessary devices were the most common types of errors, followed by redundancy. In particular, a major percentage of cohesive instances were misused and omitted. Similarly, many conjunction errors were attributed to the students’ omission of cohesive indices required in the context. However, misuse, redundancy, and overabundance were the main problems using synonyms and lexical overlaps.

In summary, cohesion use and errors identified in writing by the Vietnamese final-year English majors had a similar pattern. From the data, it is important to note that the more cohesive devices the students used, the more errors they made. The errors mainly fell into particular sub-types of each group. As the errors found were not caused by one type but included misuse, redundancy, omission, and overabundance, it was crucial to further investigate the students’ conceptions through interviews to tackle the students’ main problems with cohesion use.

RQ2: What Misconceptions Lead Vietnamese Last-Year English Majors to Difficulty in Using Those Cohesive Devices?

Before in-depth interviews, a further examination showed that the students frequently made errors with some words of each sub-type. For reference, many students interchangeably used the words “it” and “this.” For conjunctions, the detected errors were found with the use of the items “on the other hand” (adversative) and “so” (causal). The students also overused synonyms and repeated words to lexically maintain cohesion in their writings. These were the focus of the interviews.

Analyses of the interview transcripts showed some misconceptions of students about the meanings and structure of the aforesaid cohesive devices. When asked about inappropriate choices, some students expressed their misunderstanding of the meanings and structure. First, the students conceptualized the words “it” and “this” as synonyms; therefore, they thought these words could be used interchangeably. They were first confused by the interviewer’s question, but they showed their misconception when their writing instances were shown to them. In Excerpt 1, the word “it” should be used because S12 meant the noun “regulation,” but S17, in Excerpt 2, misused the word “it.” Instead, S17 should have used the word “this” as she referred to an activity.

Excerpt 1 (S12)

Interviewer: Could you tell me how to use “it” and “this”? Student: . . . (embarrassed). I used “it” and “this” to mean something singular.

Interviewer: What is the difference in meaning and use between these two words? Student: . . . (smiled with no answer)

Interviewer: For example, in your report, you wrote here “One regulation of the center does not allow the office staff to have refreshment during the working hours. This requires the staff to focus on their work.” Why did you use “this” here? (pointing at the word “this”)

Student: The word “this” here refers to the “regulation”.

Interviewer: How about “it”? Student: I think they are interchangeable.

Excerpt 2 (S17)

Interviewer: May I ask you about your word choice now? Student: Yes.

Interviewer: Could you tell me why you used the word “it” in this sentence? (pointing at the word “it” in the second sentence “Some students cheated on the test. It was unacceptable.”)

Student: I used “it” to mean “cheating”.

Interviewer: How about “this”? Can “this” be used here instead? Student: I think both are correct.

Many students also inappropriately conceptualized the adversative “on the other hand” as an additive to provide further information. They even referred this item to a Vietnamese equivalent meaning “in addition.” This misconception of the functional meaning of the term “on the other hand” as an additive rather than an adversative mainly led to their misuse.

In the case of the causal connective “so,” the students made an inappropriate analogy between this term and “therefore.” Their misconception was that the terms “so” and
Therefore had the same meaning and function in a text; thus, they used these two terms interchangeably. However, although the terms are both classified into causal connectives, they apply different grammatical structures. While “so” is a coordinator which combines two clauses into a sentence, “therefore” is an adverb conjunct. The distinction between these two words should not be made by using their lexical meanings but structure or function in a text.

The students’ misconception of the countability of nouns was found to have causal effects on their misuse and redundancy of the definite article. This misunderstanding was traced to the students’ L1 interference. The students provided explanations of their choice by translating, accompanied by follow-up questions denoted by the researcher (interviewer). They mismatched some uncountable nouns, such as “advice,” “equipment,” “information,” “news,” and “software,” to countable equivalents in Vietnamese. In fact, the use of the definite article should depend on the countability of the noun with which it goes to form a noun phrase.

Some misconceptions about cohesion use were also found from the transcript analysis. First, some students believed that the use of cohesive instances was optional. It only showcased the writer’s writing ability to make their writings considered “excellent” (S18) or “native-like” (S6). Also, lexical cohesion was deemed by the students to be given more credit than the other types. Some students even thought that the use of lexical cohesion indicated the writer’s large repertoire of vocabulary.

Another misconception among the students was that the use of ellipsis and substitution could result in low writing quality. Some students even thought that using these two cohesion types was risk-taking because the use of ellipsis could make a sentence “incomplete” (S4) or “fragmental” (S10). As such, it could be deemed “unclear in meaning” (S20) if readers did not make sense of it. In a writing test, “I was advised to make sentences as clear as possible.” (S22).

There were found from the transcript analysis several other causes of the students’ errors. First, heavy workload did not allow them to proofread and edit their reports sufficiently before submission. Second, it was the students’ carelessness that made their texts erroneous. Finally, many students admitted they forgot the meanings, structure, and functions of some cohesive devices, which resulted in errors. Although these causes detected from the semi-structured interviews did not directly reveal the students’ misconceptions, they could generally explain what the errors were attributed to.

In general, the students resorted to L1 and L2 usages to maintain text cohesion. Errors occurred because they misconceived the lexical and grammatical properties of the cohesive devices, differences between their L1 and L2 English, and inadequate proofreading and editing. It might be necessary to interpret the findings and give implications for teaching and learning cohesive devices.

**Discussion**

The research questions investigated the cohesion errors made by Vietnamese final-year English majors who were spending their internships as teachers at English language centers in Vietnam and their misconceptions that led to their inappropriate choices of cohesive devices in their reports. As the first attempt to attribute these errors to the students’ misconceptions, the current study has provided empirical results about the students’ difficulties maintaining text cohesion and interpreted the causes of these problems.

The first research question delved into the students’ difficulties using cohesive devices. The findings indicated that
the students’ cohesion use and errors had a relatively similar pattern. In general, the number of cohesive instances constituted approximately 7.90% of all word count in the data set. This finding was aligned with Rahman’s (2013) and Yang and Sun’s (2012) studies about the percentage of cohesive devices that L2 English students at the advanced proficiency level used in writing. Rahman (2013) explained that L2 English students tended to use more references, connectives, and lexical devices more often than the other two types. This result provides a reference about cohesive ties used by Vietnamese English majors because discourse is context-bound, and papers need to expose their good control of cohesion to be highly rated (Crossley & McNamara, 2012). In this study, the students’ use of more references, conjunctions, and lexical devices indicated their familiarity with these devices. Mojica (2006) and Rahman (2013) emphasized that L2 English learners tend to use lexical devices more frequently than the other types. Crossley and McNamara (2012), Xuefan (2007), and Yang and Sun (2012) explained that the use of one type of cohesion more than others does not indicate their more proficiency at that type because writers can manage to choose the devices with which they are familiar. Interestingly, the errors detected also focused on specific sub-types of cohesion. However, this does not indicate that the students were better at using the other sub-types because they could opt to use the cohesive ties in which they were confident (Mojica, 2006; Rahman, 2013). While the students misused and omitted a high ratio of references and conjunctions, errors of many synonyms and repeated words in their writings were found redundant and overabundant. It is detected from the error analysis that the students made more errors with the devices which they used more frequently.

The second research question mainly explored the students’ misconceptions of cohesive devices and cohesion use that resulted in the detected errors in response to the first research question. First, the words “it” and “this” were misconceived as synonyms. Although these words both refer to a singular unit, they are used to denote different types of referents in the context of a text. The word “it” is a third-person singular pronoun denoting a person, animal, or thing, whereas the word “this” can be used to indicate an idea or substitute a clause. The misconception occurred because of unclear distinction between these words when they are both used to refer to an object. The same problem was detected from the students’ use and understanding of the adversative “on the other hand.” As the students mainly used translation in a vocabulary learning strategy, they mismatched this item with a Vietnamese equivalent denoting an additive. Distinguishing these two words by lexically translating from L2 into L1 may be jeopardized. A possible explanation is that the cohesive devices were intervened by their L2 (Hung et al., 2021; Rahman, 2013) as conceptualizations first take place in L1 and then transfer to L2 (Alonso et al., 2016; Ellis, 2015) because L2 users employ their L1 and socio-cultural backgrounds in encoding their messages. The distinction between these words should be presented in a natural discourse or a text (McCarten, 2007) because these words carry an array of meanings that are subject to change in different contexts (Chung & Pennebaker, 2007; Halliday & Matthiessen, 2014). Using translation as a learning strategy may put the students’ understanding of this word group at risk. Also, the students misunderstood the word “so.” They simply categorized this word and the word “therefore” as denotations of a result. This finding provides an interesting example of the students’ overgeneralization. To mediate their memorization of these two words, the students simply mapped words that were considered synonyms. Although these two words expose to have similar functional meanings, they apply different grammatical structures because “so” is a coordinator, but “therefore” is an adverb conjunct. From a usage-based perspective, students can learn such structures through texts because a text can provide the context of situation, which helps students make sense of individual items in relation to the other elements (Derewianka & Jones, 2016; Macken-Horark et al., 2017; Thornbury, 2002).

Unlike the other types of cohesion, the incorrect choice of the definite article occurred not because the students misconceived this word class, but because they misconceived the countability of the nouns that the definite article accompanies. The students’ misclassification of some uncountable nouns as countable resulted in the misuse and redundancy of this type of determiner. Doetjes (2017) raises an awareness that the distinction between some countable and uncountable nouns is conceptual because they expose semantic properties similar to those of countable nouns. For instance, as an uncountable or a mass noun, “furniture” denotes heterogeneous individual objects. Doetjes (2017) argues that the con-strual of the concept is responsible for word choice as the students do not conceptualize the individualization of entities. Also, the unclear boundary between countable and uncountable nouns can lead to students’ overgeneralization. The fact that some nouns are countable in one sense but uncountable in another confuses the L2 English speakers (Han et al., 2006; Leacock et al., 2010). This finding gives implications for teaching and learning the definite article and countability of nouns in ESL and EFL contexts. Ellis and Cadierno (2009) stresses that the acquisition of L2 involves both L2 forms and meanings. In the L2 learning process, the teacher can direct the students to “construe the world like natives of the L2” (p. 111). From the students’ performance and interaction, the teacher may detect students’ inappropriate choice of the definite article and nouns. They can unearth the students’ conceptual pitfalls and solicit their misconceptions interactively or help re-construe their understanding of the countability of such nouns regarding their relationship with the definite article.

The final misconception was about the employment of substitution, ellipsis, and lexical cohesion. The students’ interview responses showed that they wanted to utilize more lexical cohesion to be considered at the advanced level of
writing proficiency. They also disregarded the use of substitution and ellipsis in writing. This result explained the statistics from the text analysis that the students used lexical cohesion the most frequently, but substitution and ellipsis instances constituted a small percentage. The errors in synonyms and repeated words were mainly categorized as redundancy and overabundance. This misconception of students about the use of substitution, ellipsis, and lexical cohesion in writing suggests implications for L2 writing teaching and learning. Instead, students should be directed to employ a syntactic and lexical variety in writing. In general, the rating of L2 writing is partly based on the diversity and complexity of vocabulary use (Crossley et al., 2014). Syntactic and lexical overlaps are negative indicators of writing quality. Native-English speakers generally use a wider range of cohesion indices than L2 English speakers, even at the advanced proficiency level (Mojica, 2006; Rahman, 2013).

Other causes of errors in cohesion use were detected from the interviews. Some students were careless of word choice in writing, and others suffered time constraints from their heavy learning schedules. Errors are inevitable in L2 learning; forgetting should not be a shame. Students can relearn what they have learned (Baker & Greif, 2020; Makombe et al., 2020). Given the importance of text cohesion, it might be necessary for L2 students to spend time proofreading and editing their papers before submission. It should be an indispensible part of the writing process.

Cohesion use is not optional in writing but contributes to discourse unity (Fitch, 2011; Yang & Sun, 2012). Cohesive papers are usually rated high due to the connection of ideas identified by the examiner (Crossley & McNamara, 2012). Writing teachers are advised to explicitly stress this rating criterion to students to help them know the crucial role of text ties and make their written ideas adhere connected.

This study contributed to the literature on cohesion and error analysis. Theoretically, as the first attempt to investigate the effects of L2 students’ misconception on their written text cohesion, the current study introduced a model of analyzing students’ errors accompanied by interviews to explore the causes of the detected errors. This study revealed many factors in L2 students’ written errors. It also traced L2 students’ misconceptions as the main cause of the inappropriate choice of cohesive devices in writing. In practice, the current study sheds light on the errors and students’ misconceptions of cohesive devices and cohesion use in writing.

This study has several limitations that future research should consider. First, the sample size was relatively small. The limited scope and scale only provided novel insights into Vietnamese L2 students’ misconceptions of the focused target items. Future research can recruit a larger participant pool and analyze more written papers to generalize the result.

Conclusion
As part of a larger project, this study explored final-year English majors’ use of cohesive devices in writing and their misconceptions that influenced their use of this word class in a particular EFL Asian context. Through error analysis and prompt interviews, the study highlights some relationship between the students’ misconceptions and errors of cohesive devices. Given error making as an inevitable phenomenon in the L2 learning process, this study has given some insights into patterns of cohesion errors and use; meanwhile, it has revealed some sources of confusion and problems that EFL students encounter in writing. The findings show that the potential of this approach could assist L2 students in overcoming the target language errors. This study was driven by the ultimate goal to contribute to both the local and international literature of the field.

Author Note
This study is part of a larger project on English language teaching and learning in Vietnam.

Acknowledgments
I would like to send great thanks to the editor and reviewers for their constructive feedback. Sincere thanks should also go to the participants involved in this study.

Declaration of Conflicting Interests
The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by University of Economics Ho Chi Minh City, Vietnam.

ORCID iD
Hung Phu Bui https://orcid.org/0000-0003-3468-4837

References
Alarcon, J. B., & Morales, K. N. (2011). Grammatical cohesion in students’ argumentative essay. *Journal of English and Literature*, 2, 114–127.
Al-Jarf, R. S. (2001). Processing of cohesive ties by EFL Arab College. *Foreign Language Annals*, 34, 141–151. https://doi.org/10.1111/j.1944-9720.2001.tbo2819.x
Alonso, A. R., Cadierno, T., & Jarvis, S. (2016). Crosslinguistic influence in the acquisition of spatial prepositions in English.
Oller, J. W., & Jonz, J. (1994). Cohesion in ESL classroom written texts. *Olateju, M. (2006). Cloze and coherence. Asian EFL Journal, 13(3), 42–65. Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: Evidence from event encoding. *Koshik, I. (1999). A preliminary investigation into the effect of grammatical cohesive devices: Their absence and their misuse on the comprehension of non-native speaker speech and writing. Issues in Applied Linguistics, 10(1), 3–25. Lake, J. (2004). Using “on the contrary”: The conceptual problem for EAP students. *ELT Journal, 58(2), 137–144. https://doi.org/10.1093/elt/58.2.137 Langacker, R. W. (2008). Cognitive grammar and language instruction. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 66–88). Routledge. Leacock, C., Chodorow, M., Gamon, M., & Tetreault, J. (2010). *Automatic grammatical error detection for language learners*. Morgan & Claypool Publishers. Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing. *System, 33*, 623–636. https://doi.org/10.1016/J.SYSTM.2005.02.002 Macken-Horarik, M., Love, K., Sandiford, C., & Unsworth, L. (2017). *Functional grammatics: Re-conceptualizing knowledge about language and image for school English*. Routledge. Mackey, A., & Gass, S. M. (2005). Second language research: Methodology and design. Lawrence Erlbaum Associates Publishers. Makombe, T., Dube, S. S., & Morare, M. (2020). Learn, unlearn, relearn: How to live the life you desire through the transformation of your beliefs and behaviours. The Golden Goose Institute. Malmkjær, K. (2001). *The linguistic encyclopedia*. Routledge. McCarten, J. (2007). Teaching vocabulary: Lessons from the corpus, lessons for the classroom. Cambridge University Press. Mojica, L. A. (2006). Reiterations in EFL learners’ academic papers: Do they contribute to lexical cohesiveness? *The Asia-Pacific Education Researcher, 15*, 105–125. Nuyts, J., & Pederson, E. (2012). *Language and conceptualization*. Cambridge University Press. https://doi.org/10.1017/CBO9781139086677 Olateju, M. (2006). Cohesion in ESL classroom written texts. *Nordic Journal of African Studies, 15*(3), 314–331. Oller, J. W., & Jonz, J. (1994). Cloze and coherence. Associated University Presses. Ong, J. (2011). Investigating the use of cohesive devices by Chinese EFL learners. *Asian EFL Journal, 13*(3), 42–65. Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: Evidence from event encoding. *Language, Cognition and Neuroscience, 34*(9), 1117–1128. https://doi.org/10.1080/23273798.2019.1589540 Radden, G., & Dirven, R. (2007). *Cognitive English grammar*. John Benjamins Publishing Company. Rahman, Z. A. (2013). The use of cohesive devices in descriptive writing by Omani student teachers. *Sage Open, 3*(4), 1–10. https://doi.org/10.1177/2158244013506715 Reynolds, D. W. (2001). Language in the balance: Lexical repetition as a function topic, cultural background, and writing development. *Language Learning, 51*(3), 437–476. https://doi.org/10.1111/0023-8333.00161 Robinson, P., & Ellis, N. C. (2008). *Handbook of cognitive linguistics and second language acquisition*. Routledge. Salkie, R. (1995). *Text and discourse analysis*. Routledge. Sanders, T., & Maat, H. (2006). Cohesion and coherence: Linguistic approaches. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 39–54). Elsevier. Tanskanen, S. K. (2006). *Collaborating towards coherence: Lexical*. John Benjamins Publishing. Thierry, G. (2016). Neurolinguistic relativity: How language flexes human perception and cognition. *Language Learning, 66*(3), 690–713. https://doi.org/10.1111/1467-1770.12186 Thornbury, S. (2002). *How to teach grammar*. Pearson Education. Todd, R. W., Khongput, S., & Darasawang, P. (2007). Coherence, cohesion and comments on students’ academic essays. *Assessing Writing, 12*(1), 10–25. https://doi.org/10.1016/i.asw/2007.02.002 Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge. Xuefan, C. (2007). Lexical cohesion in Chinese college EFL writing. *CELEA Journal, 30*, 46–57. Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education, 23*(1), 31–48. https://doi.org/10.1016/j.linged.2011.09.004 Yule, G. (2008). *The study of language* (3rd ed.). Cambridge University Press. West, R., & Turner, L. H. (2007). *Introducing communication theories: Analysis and application* (3rd ed.). McGraw-Hill. White, B. J. (2010). *In search of systematicity: A conceptual framework for the English article system* [Unpublished doctoral thesis, Michigan State University]. Zhang, M. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *REL C Journal, 31*(1), 61–95. https://doi.org/10.1177/003368820003100104s