International Journal of Research in English Education (IJREE)

**Original Article** **Published online: 20 December 2021.**

# Error Analysis of Taiwanese University Students' English Essay Writing: A Longitudinal Corpus Study

Abolfazl Shirban Sasi[1] & Jiin Chyuan Mark Lai[2*]

* Correspondence:
Pictologics@gmail.com
1. Department of Foreign Languages (DFL), National Chiayi University, Taiwan
2. Department of Applied Foreign Languages (DAFL), TransWorld University, Taiwan

Received: 26 July 2021
Revision: 14 October 2021
Accepted: 5 November 2021
Published online: 20 December 2021

**Abstract**

Writing is considered one of the most difficult skills in EFL/ESL. Thus, meticulous recognition and classification of students' errors in certain contexts is a worthwhile endeavor which provides us with both diagnostic and prognostic power. Accordingly, a total of 430 students in 15 English writing classes held during 12 consecutive semesters in a private university in central Taiwan were the subjects of this study. They composed 5703 essays which were rated and coded by the authors. Adopting and modifying the error taxonomy proposed by Zheng and Park (2013), the authors classified a sum of 63460 errors into four main groups with their subsequent subcategories. This study revealed that the highest problematic areas for Taiwanese university students were 'misformation' with 51.55% of the whole including errors in tenses, parts of speech, prepositions, subject/verb agreements, and run-on sentences. Then, 'omission' errors ranked second with 21.30% including errors in articles, plural suffix-s, and relative pronouns. Finally, the third and fourth error types were 'others' with 15.13% including spelling, capitalization, and wrong vocabulary, and 'addition' with 12.01% containing errors in articles, unnecessary words, and conjunctions. This study provides numerous genuine samples from the students' compositions being annotated based on the applied error taxonomy. Thus, the data presented in this study can provide researchers with a practical framework for future studies in error analysis, as well as pedagogical implications in the field.

**Keywords:** addition, error, misformation, mistake, omission, taxonomy



[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]

[ DOI: 10.52547/ijree.6.4.57 ]

## 1. Introduction

### *1.1 Preview*

By essence, writing is a complex task, and writing in a foreign/second language makes it even more challenging. We might rightfully assume that recognition, understanding, and evaluating students' errors in any of the four language skills can benefit both teachers and learners. Accordingly, the authors of the current paper postulate that language learners' errors can be depicted and studied more efficiently in written forms because of reasons like more durability, ease of access, analysis speed, and so on.

### *1.2 Statement of the Problem*

The authors believe that there are three problems or deficiencies in the relevant literature that the current study addresses: 1) sample size, 2) study duration, and 3) a moderate approach towards both explaining and correcting students' errors/mistakes.

Firstly, although even a limited number of samples might statistically be sufficient, when we claim of studying, analyzing, and reporting a corpus, then it should contain a substantial number of samples. The valid sample size of any given corpus under scrutiny in social sciences is a controversial topic. However, there are many scholars who vote for bigger corpora. For example, Conrad (1999) asserts that the corpus size is an important principle to consider because a small corpus will not include representative samples of the subject under investigation. Likewise, according to Baker (2006), corpora are generally large representative samples of a specific type of language to be used as a standard reference to make measurements. To our best knowledge, large corpora of second language (L2) English essay samples have not been collected or studied. Table 1 summarizes some relevant researches and the number of essays used in them.

Table 1. Number of essays used in L2 English error analysis corpus studies

| Author(s) | Publication Year | Number of Essays |
|---|---|---|
| Amoakohene | 2017 | 50 |
| Bao | 2015 | 134 |
| Bond | 2016 | 36 |
| Chan | 2010 | 696 |
| Chen | 2002 | 28 |
| Darus & Khor | 2009 | 70 |
| Darus & Subramaniam | 2009 | 72 |
| Divsar & Heydari | 2017 | 70 |
| Firas | 2012 | 350 |
| Hart | 2017 | 600 |
| Hou | 2016 | 112 |
| Hu | 2014 | 200 |
| Lan, Lucas, & Sun | 2019 | 100 |
| Mohamed, Goh, & Wan | 2004 | 15 |
| Pimpisa & Norma | 2015 | 40 |
| Quibol-Catabay | 2016 | 30 |
| Ridha | 2012 | 80 |
| Zhan | 2015 | 177 |
| Zheng & Park | 2013 | 168 |

As evident in Table 1, the two largest corpora found in literature was a study by Chan (2010) in which 696 essays composed by Hong Kong Cantonese English as a Second Language (ESL) learners were investigated, as well as a study by Hart (2017) in which the researcher had analyzed 600 writing samples produced by Chinese undergraduates,

graduates, and researchers. The other studies have used far fewer samples, even 15 in Mohamed, Goh, and Wan (2004)'s.

The second reason, we believe, is the element of study duration. The current research was conducted during a time span of 12 consecutive semesters (six years). This means that the possible fluctuation in the subjects' English proficiency level (which is presumably supposed to grow exponentially during the undergraduate period) has been accounted for. In other words, we are dealing with a wider range of proficiency level among the subjects of the study here; hence, more population representativeness. In contrast, most of the other studies found in the literature gathered the samples in one or a few sessions only.

Finally, the most important property of the present study is the moderate approach towards analysis and correction of the errors/mistakes. As discussed in the following sections, the authors of the current study do not see, in many cases, an absolute one-to-one correspondence between an erroneous structure and the way(s) it can be fixed. This flexibility gives both English instructors and learners an ease of mind to tackle the problem(s). Thus, in this process, two abilities of 1) recognizing different causes to a problem; and 2) selecting among alternatives with the use of a methodic reasoning will constantly be applied. These two are among the critical thinking skills advocated by many scholars (Angeli & Valanides, 2009; Glassner & Schwarz, 2007; Halpern & Sanders, 2004; Nakagawa, 2011; Popil, 2011). We believe that the L2 practitioners can make the best of this moderate outlook towards language learners' errors or mistakes, particularly when categorizing and correcting them.

### *1.3 Research Objectives*

The current study was conducted in order to firstly recognize and categorize English lexicogrammatical errors/mistakes typically made by Taiwanese EFL college students; and at the same time, suggest a flexible approach to help both teachers and students be aware of them and, subsequently, try to remove or at least minimize them. The authors' assumption was that knowing the problem precisely would be the precursor of solving the problem permanently.

## 2. Review of the Literature

### *2.1 Probable Sources of Errors*

According to Richards and Schmidt (2002), error is the use of a linguistic item in a way that a fluent or native speaker of that language regards as showing faulty or incomplete learning. They classify errors as vocabulary (lexical errors), misunderstanding of a speaker's intention or meaning (interpretive errors), and wrong communicative effect production (pragmatic errors). Another way to conceive sources of errors is to dichotomize errors into two categories of interlingual and intralingual errors. Interlingual (or interference) errors are those which can be traceable in learners' first language. That is, these errors can be attributed to negative interlingual transfer. According to Selinker (1972), the term "interlingual" was first introduced by Weinreich in 1953. On the other hand, Intralingual (or developmental) errors are those which result from faulty or partial learning of the target language rather than language transfer (Fang & Jiang, 2007).

Similarly, according to Richards (1992), intralingual errors reflect not the mother tongue structures, but generalizations based on partial exposure to the target language. Erdogan (2005) postulates that intralingual errors happen as a result of learners' attempt to build up concepts and hypotheses about the target language via their limited experience. According to Brown (1994), particularly in the early stages of learning and before the system of the L2 is familiar, the first language (L1) is the only previous linguistic system that the learner can depend on. Thus, errors occur as a result of familiarity with the L1 and by being transferred either directly or indirectly from the L1 to the new language. From another perspective, Leki (1992) asserts that English as a Second Language (ESL) students' language proficiency does not unambiguously correlate to their writing skills, and that "Fluency in language may be obscuring lack of experience with writing and even lack of cognitive academic development" (p.87). Mitchell and Myles (2004) believe that language errors are normal and inevitable features of learning. They assert that the study of errors can reveal a developing system of the learners' language which is supposed to be dynamic and open to changes.

### *2.2 Taxonomy of Errors*

According to Ellis (1997), once a writing sample has been assessed as incorrect, it is still sometimes difficult to determine whether an error or a mistake has been produced. Corder (1974) suggests a five-stage procedure which has become known as "The Error Analysis Method." The complete procedure includes: 1) choosing the language corpus,


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

2) identifying errors in the corpus, 3) classifying errors, 4) explaining errors, and 5) evaluating errors. Furthermore, Dulay, Burt, and Krashen (1982, p.150) suggested a "surface strategy taxonomy" which highlights the ways surface structures are altered. This taxonomy consists of four elements, namely: 1) misformation, 2) omission, 3) addition, and 4) misordering. Adopted by many researchers in the field, a modified version of this taxonomy (to be explained further in the forthcoming sections) constitutes the base of the current article. In the following lines, we give a short description of the four types of errors in this taxonomy.

Errors of misformation are characterized by the use of wrong morphemes and structures in which the user is using one grammatical form in place of another grammatical form. That is, he/she provides a sentence which is grammatically incorrect in English. For example: '*Mr. Smith has going to the hospital.' instead of 'Mr. Smith has gone to the hospital.', or 'Mr. Smith is going to the hospital.' in which the verb in the erroneous structure has been conjugated wrongly. Omission errors, on the other hand, are categorized as structures that do not contain some segments which must exist in them. For example, the sentence '*Mr. Smith a doctor.' instead of 'Mr. Smith is a doctor.' that agrees with the structure of good English grammar by having the copula 'is'. Then, error of addition is the opposite of omission. This error happens when we include an element that does not need to exist. For example, the sentence '*Mr. Smith he is a doctor.' instead of 'Mr. Smith is a doctor.' Here the subject pronoun "he" has been added to the first structure unnecessarily. Finally, error of misordering is marked by misplacing a morpheme or a group of morphemes in the structure. For example: '*I don't know where is Mr. Smith from.' instead of 'I don't know where Mr. Smith is from.'; in which bringing the verb "is" after the relative pronoun in the indirect interrogative clause does not comply with the correct English grammar.

It is also important to notice that in L2 acquisition (SLA) literature the word "mistake" frequently refers to incorrect language forms due to carelessness, whereas the word "error" refers to incorrect language use due to inadequate knowledge of grammar. This division, however, is not unanimously agreed upon or required among the practitioners in the field. For example, according to Richards and Schmidt (2010, p.201) this distinction is only "sometimes" made between error and mistake. Thus, in the current study (and for the sake of laconic arguments) these two terms are used interchangeably.

### *2.3 Benefits of Studying Error Corpus*

Since the mid-1980s, corpus linguistics has been progressively recognized and practiced as an influential methodology in language teaching and learning (Conrad, 2000; Ro¨mer, 2011; Sinclair, 2004). There have also been various studies on the language learners' error corpus in the literature such as (Chuang & Nesi, 2006; Granger, 2003; Han, Chodorow, & Leacock, 2006; Hou, 2016; Myles, 2005; Pravec, 2002). According to Kotsyuk (2015), by studying learners' errors, we can predict the difficulties involved in learning a foreign language. Also Ellis (2008) postulates that classification of errors helps teachers in recognizing and analyzing learners' language problems at any phase of their L2 development. Likewise, Hou (2016) asserts that by identifying and classifying errors in large scales, it is possible to design instructional materials that are focused more locally in order to better suit learners of a specific mother tongue in a specific context. Moreover, Keshavarz (2008) affirms that by describing and classifying language learners' errors in linguistic terms, we can build up a picture of the features of language which are causing their learning problems.

### *2.4 Related Studies*

Jiang (1995) inspected Taiwanese EFL learners' errors in English prepositions and found that a large number of errors were due to language transfer. The postulated reason was that compared to English speakers, Mandarin speakers use fewer prepositions for more concepts. In another study, Darus and Khor (2009) investigated Chinese students in a public school in Malaysia. They found out that four most common errors of their subjects were mechanics, tenses, preposition, and subject/verb agreement. They concluded that intralingual transfer of Malay and developmental errors were the main causes observed in their study. Chen (2002) investigated the characteristics and problems of university English as a Foreign Language (EFL) writing in Taiwan. The most frequent errors the participants made in the rank order were: word usage, tense, definite article usage, prepositions, verbs, singular or plural, relative clauses, and redundant usage. In another study, common error types observed among Hong Kong Cantonese ESL students were detected and categorized by Chan (2010). Chan's proposed error taxonomy views the errors in a continuum from shorter to longer structures. Thus, four main levels of morphological, lexical, syntactic, and discourse are identified and explained in her study. As for the frequency of the errors, calquing, word class confusion, concord problems, and omission of copulas scored the highest respectively. Chan postulates two main reasons for the errors: L1 influence, and inadequate mastery of appropriate English usage.


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

Hart (2017)'s work seems more like a practical manual for the language teachers and learners. He does not engage too much with the technical terms. Instead, he uses three generic categories of "classes", "choices", and "components." By "classes" he means grammatical errors which must be recognized; "choices" then deals with confusion about word choice, such as related or unrelated terms; and finally, "components" means different elements that make up an essay or paper such as tables, punctuation marks, etc. Similarly, Divsar and Heydari (2017) developed an error-coding scheme which ultimately categorized the errors into 13 subgroups. They found out that the two most frequent errors made by the International English Language Testing System (IELTS) candidates under their study were word choice and verb forms (tenses and inflections). Generally, it seems those researchers who modified existing taxonomies or even developed their own cataloging system, performed a more rigorous job in terms of precision in error classification. Some of those studies are Hart (2017), Divsar and Heydari (2017), Zheng and Park (2013), and, we assume, the present study.

Then Bao (2015)'s study is also similar to the current study in that it considers "part of speech" as a good type indicator of errors. We also believe this way of looking at the problem provides us with more flexibility without compromising accuracy. Investigating lexical errors, Bao spotted four main types including part of speech, substitution, absence, and redundancy errors. The sources of errors were attributed to mother tongue interference and inadequate learning and understanding of the L2 rules. In another study, Darus and Subramaniam (2009) used Markin software to analyze the errors, and found out that learners' six most common errors were singular/plural form, verb tense, word choice, preposition, subject-verb agreement and word order. Using computer software is doubtlessly a practical way to enhance speed and accuracy in research. However, due to the nature of this kind of research, we believe typing students' essays in a word processor to be analyzed later is time-consuming. In addition, several errors such as spelling, punctuations, or even word choice might automatically be corrected by the software without the researcher's knowing.

Next in an interesting study, Lan, Lucas, and Sun (2019) investigated the probable impact of L2 writing proficiency level on the complexity of noun phrases used by the students. Their study subjects involved an equal dichotomy of high-proficiency and low-proficiency level students. Their study illustrated a positive correlation between the use of 11 noun modifiers and L2 writing proficiency of the first-year Chinese students. Then, in another research by Hou (2016), insufficient mastery of vocabulary was claimed to be the causes of the lexical errors made by the Taiwanese subjects in the study. Hou also concludes that the top three grammatical errors were verb forms, article errors, and prepositional errors. Finally, similar to the current study, Zhan (2015)'s study ranked errors in tenses as the most frequent found in Chinese EFL students' writing. Among others were spelling, vocabulary use, singular and plural forms of nouns, parts of speech, non-finite verbs, run-on sentences, and pronouns. Again, this study looked at parts of speech as a different error type.

## 3. Methodology

### *3.1 Participants*

A total of 430 students enrolled in the writing classes offered during the six years of sample collection in this study. Subsequently, a sum of 5703 essays were collected and coded by the authors. Table 2 shows the details.



[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

Table 2. Number of students and their essays during the sample collection period

| Year (Semester) | Course Title | Number of Students | Number of Essays |
|---|---|---|---|
| 2012-13 (1) | Eng. Writing 3 | 33 | 503 |
| 2012-13 (2) | Eng. Writing 4 | 34 | 524 |
| 2013-14 (1) | Eng. Writing 3 (*2) | 25 + 19 | 628 |
| 2013-14 (2) | Eng. Writing 4 (*2) | 21 + 21 | 593 |
| 2014-15 (1) | Eng. Writing 3 (*2) | 27 +24 | 613 |
| 2014-15 (2) | Eng. Writing 4 | 36 | 524 |
| 2015-16 (1) | Eng. Writing 1 | 51 | 669 |
| 2015-16 (2) | Eng. Writing 2 | 43 | 607 |
| 2016-17 (1) | Eng. Writing 1 | 34 | 351 |
| 2016-17 (2) | Eng. Writing 2 | 24 | 276 |
| 2017-18 (1) | Eng. Writing 1 | 19 | 211 |
| 2017-18 (2) | Eng. Writing 2 | 19 | 204 |
| Sum | 15 classes | 430 | 5703 |

*3.2 Design of the Study*

The present study applied a qualitative research design because the analysis of the data was mostly descriptive in nature. We believe that even the frequency counts in the process of rating students' essays essentially tried to describe the area(s) of mistakes or errors, more than merely quantifying a phenomenon.

*3.3 Instruments*

Basically, essays written by the students were the instruments of the current study. Totally, more than 60 various topics were used in this study. They covered a vast range of subjects and themes, as well as many of the most popular writing styles such as descriptive, narrative, argumentative, cause and effect, problem and solution, comparison or contrast, and so on. Since students varied in different semesters and courses, many topics were used several times. Necessary precautions were taken so that a given student would not write on the same topic twice. In retrospect, however, this would not have been a big problem as, presumably, the students' English proficiency level, their opinions as well as their learning attitude would also vary in the course of time. As part of the teaching program, each essay was marked, scanned, and then analyzed on the classroom projection screen the following session. Apart from


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

the scanned essays, students in each class also did other writing exercises either in the class or as homework. Those are not covered in the present paper. The following is a list of some of the topics used in this study:

- A day without my smart phone
- Benefits of exercising
- Earthquakes and potential dangers they bring
- Fast foods, to eat or not to eat?
- Global Warming
- How to improve your teamwork with your classmates
- Human rights in today's society
- I like/don't like shopping.
- Role of music in our life
- The best place you've been to. Why?
- The future of AI (Artificial Intelligence)
- To smoke or not to smoke?
- What is true happiness?
- Watch the video clip carefully and write a summary of the information provided in the film (you may add your personal impression as well).
- Using your imagination, describe the following picture.

*3.4 Data Collection*

As mentioned before, a total of 430 students in 15 composition classes were the subjects of this study. During 12 semesters, they produced 5703 essays altogether (an approximate average of 13 essays for each student during each semester). On the whole, the authors have recognized and coded a sum of 63460 errors (an average of approximately 11 errors per essay). These errors have been categorized in four main groups with their subsequent subcategories which will be further discussed and explained in the following sections.

**4. Data Analysis and Results**

*4.1 Preview*

It is very important to mention here that the authors believe that in many cases, classifying errors, by essence, is a rather controversial and hence, dubious task. Take the following short sentence for example:

* She been teached English.

Since we do not have access to the mind of the producer of this structure, also because we see this "sentence" in isolation, then we cannot be completely sure what he or she wanted to state. The intended correct message/form might have been one of the followings, among some other possibilities (thus, the tagged errors):

She has been teaching English. (tense)

She has taught English. (tense)

She has been taught English. (passive voice construction)

She has been being taught English. (passive voice construction)

She is/has been an English teacher. (tense, auxiliary, article, noun form, etc.)



[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]

[ DOI: 10.52547/ijree.6.4.57 ]

Thus, many options being considered, the present study has adopted the error classification used by Zheng and Park (2013), a practical taxonomy which was itself based on the works of Dulay, Burt, and Krashen (1982), and Kim (2009). When allocating the students' errors to each individual subcategory, the authors of the current study tried to zoom in the closest possible entry. In doing so, we have modified some subcategories used in Zheng and Park (2013) in order to avoid probable overlaps explained above. Furthermore, it should be emphasized here that in all depicted erroneous structures in this study, only one error type has been tagged to each individual error in any given structure.

Accordingly, based on Zheng and Park (2013), errors can be categorized into four main groups of ***misformation***, ***omission***, ***addition***, and ***others***. Each error category will be discussed and explained along with original examples for each pertinent subcategory in the order of the highest to the lowest observed frequencies. Also, it should be noted that in the current article, the erroneous sample structures are shown either in annotated tables, or with an initial asterisk.

*4.2 Misformation*

According to Dulay, Burt, and Krashen (1982), misformation refers to errors characterized by "the use of the wrong form of the morpheme or structure" (p.158). In the present study, misformation errors were recognized and inspected to be a total of 32715 cases which is equal to 51.55 % of all the errors identified in this study. Table 3 shows the type, rank and number, as well as percentage of each of these misformation errors.

Table 3. Details of the misformation errors observed in this study

| Rank | Number of Errors | % | Categories |
|---|---|---|---|
| 1 | 7451 | 22.78 | Tenses |
| 2 | 3105 | 9.49 | Parts of speech |
| 3 | 2650 | 8.10 | Prepositions |
| 4 | 2407 | 7.36 | Subject/Verb agreement |
| 5 | 2394 | 7.32 | Run-on sentences |
| 6 | 1856 | 5.67 | Adjective/Adverb forms |
| 7 | 1808 | 5.53 | Auxiliaries |
| 8 | 1691 | 5.17 | Gerunds |
| 9 | 1558 | 4.76 | Articles |
| 10 | 1416 | 4.33 | Possessives Adj. / 's |
| 11 | 1391 | 4.25 | Passive voice constructions |
| 12 | 1374 | 4.20 | Relative pronouns |
| 13 | 1093 | 3.34 | Pronouns |
| 14 | 1041 | 3.18 | Conjunctions |
| 15 | 524 | 1.60 | Infinitives |
| 16 | 375 | 1.15 | Demonstratives |
| 17 | 319 | 0.97 | Quantifiers |
| 18 | 262 | 0.80 | Single/Plural forms |
| Sum | 32715 | 100 | |


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

As can be seen in Table 3, misformation errors have been subcategorized into 18 subgroups of errors with ***tenses*** being the most frequent of all. Consequently, in Table 4, we have given an example for each misformation error subcategory taken from the students' essays. It should be emphasized here that, mostly, the errors did not occur in isolation. That is, in the corpus which is the base of analysis of the present study, we often observed multiple error types in even short structures/sentences. However, for the sake of brevity and clarity, the forthcoming examples (for misformation errors, as well as the other three types) are those with only one error in each structure. For more examples of the multiple error types co-occurring in the students' compositions, please see the Appendix.

Table 4. Samples of students' misformation errors

| Structure | Error Type |
|---|---|
| When I was little, I go to temples near my house very often.<br>----------, I used to go ----------. | (tense) |
| Everyone has a different hobby, even me is no except.<br>----------, even I am no exception. | (part of speech) |
| They bullied me and laughed for me.<br>---------- at -----. | (preposition) |
| And if my friends comes to my house, they won't feel uncomfortable.<br>---------- come ----------. | (sub./v. agreement) |
| When I cleaning sweating all over my body.<br>When I am cleaning, I sweat all ----------. | (run-on sentence) |
| India is South America gooder.<br>India is better than South America. | (adjective form) |
| You must to obey the traffic laws.<br>---------- obey ----------. | (auxiliary) |
| I think live in a big house is safer.<br>-------- living ----------. | (gerund) |
| I'd like a ice-cream cone.<br>-------- an ----------. | (article) |
| I have an own garage.<br>------- my ----------. | (possessive adj.) |
| The Chang Gung Memorial Medical System found by Wang Yongging.<br>---------- was founded ----------. | (passive voice cons.) |
| This is the place when I saw my girlfriend.<br>---------- where ----------. | (relative pronoun) |
| Temples are truly places full of culture. You should visit it!<br>---------- them! | (pronoun) |
| I usually go shopping with my sister, but we both like shopping. | (conjunction) |


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

---------- , because ----------.

| | |
|---|---|
| Last Saturday I invited some of my friends accompany me to cinema.<br>---------- to accompany ----------. | (infinitive) |
| I'm talking to those people over here at the school now.<br>---------- these ----------. | (demonstrative) |
| There are a little people left in the future.<br>---------- a few ----------. | (quantifier) |
| In the night market, I like to eat fishes.<br>---------- fish. | (plural form) |

*Note: The first line is the erroneous structure, and the second line is a/the corrected one. Also, a dashed line means the segment is copied without any change.*

The first example in Table 4 shows an error in the correct usage of tense. This type is by far the most frequent error observed during this study. This finding is in line with most of the previous researchers for it seems errors in English tenses or verb formations almost always rank among the top three errors/mistakes in L2 students' writing. Moreover, rarely did the authors see students use more complex English tenses such as future perfect, or present perfect progressive tenses. It seems Taiwanese students have a strong tendency to use simple present tense for almost all other time frames. According to Chen (2006), verbs in Chinese remain unchanged regardless of the tenses. This may be the most prominent reason that the use of tenses in English is challenging for Taiwanese EFL students. This detected phenomenon of language transfer is consistent with the reports of L1 interference from Chinese in Jiang (1995), from Malay in Darus and Khor (2009), from Thai in Pimpisa and Normah (2015), and from Persian in Divsar and Heydari (2017). Furthermore, according to Jiang (1995), Taiwanese EFL learners make many errors when using English prepositions because of the same reason, and that compared to English speakers, Chinese speakers use fewer prepositions for more concepts. Likewise, in the current study, wrong usage of prepositions was the third highest frequency with 8.10% of all misformation cases. Then, as for subject/verb agreement, according to Chen (2006), although when conjugating English verbs, we must apply subject/verb agreement, in Chinese verbs remain unchanged regardless of person. In the current study, this type of error ranked the fourth with 7.36% of all misformation cases.

*4.3 Omission*

According to Dulay, Burt, and Krashen (1982), omissions are errors characterized by "the absence of an item that must appear in a well-formed utterance" (p.154). In the present study, omission errors were recognized and inspected to be a total of 13519 cases which is equal to 21.30 % of all the errors identified in this study. Table 5 shows the type, rank, and number, as well as percentage of each of these omission errors.


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

Table 5. Details of the omission errors observed in this study

| Rank | Number of Errors | % | Categories |
|---|---|---|---|
| 1 | 3083 | 22.80 | Articles |
| 2 | 2631 | 19.46 | Plural suffix-s |
| 3 | 1481 | 10.96 | Relative pronouns |
| 4 | 1152 | 8.52 | Punctuation marks |
| 5 | 799 | 5.92 | Prepositions |
| 6 | 637 | 4.71 | 'to be' |
| 7 | 624 | 4.62 | Auxiliaries |
| 8 | 607 | 4.49 | Possessive adj. / 's |
| 9 | 501 | 3.70 | Objects |
| 10 | 452 | 3.34 | Gerunds |
| 11 | 436 | 3.22 | Subjects |
| 12 | 414 | 3.07 | Verbs |
| 13 | 406 | 3.00 | Nouns |
| 14 | 296 | 2.19 | Pronouns |
| Sum | 13519 | 100 | |

*Note: The omissions of conjunctions have been diagnosed as run-on sentences.*

As illustrated in Table 5, omission errors have been subcategorized into 14 subgroups of errors with ***articles*** being the most frequent of all. Subsequently, in Table 6, we have given an example for each omission error subcategory taken from the students' essays.

Table 6. Samples of students' omission errors

| Structure | Error Type |
|---|---|
| If today is rainy day, it will cause everyone to drive.<br>---------- a rainy day, ----------. | (article) |
| Temples also attract a lot of foreign tourist.<br>---------- tourists. | (plural suffix-s) |
| I'm sure he'll tell you he knows.<br>---------- what he knows. | (relative pronoun) |
| Before I booked a table at the restaurant I messaged my brother.<br>---------- restaurant, I ----------. | (punctuation mark) |
| If somebody goes your house, it will be bad.<br>---------- to your house, ----------. | (preposition) |
| I go to the gym twice a week, so my body still good. | ('to be') |


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

---------- is still good.

| | |
|---|---|
| I not really like playing computer games.<br>I really don't like playing ----------. | (auxiliary) |
| I gave Steve own scooter, but he never said thank you.<br>I gave Steve my own scooter, ----------. | (possessive adj.) |
| You don't have to buy for me.<br>---------- buy (object/object pron.) for me. | (object) |
| So by bus is the best way in the crowded city.<br>So traveling/going by bus ----------. | (gerund) |
| Asked me to tell my mother, and I said OK.<br>(Someone) asked me ----------. | (subject) |
| That's why female employment rate will up.<br>---------- go up. | (verb) |
| My grandfather couldn't drink in the glass.<br>---------- water/drink in the glass. | (noun) |
| Think cleaning the house is very important.<br>I think ----------. | (pronoun) |

*Note: The first line is the erroneous structure, and the second line is a/the corrected one.*

*Also, a dashed line means the segment is copied without any change.*

As illustrated in the Tables 5 and 6, with 22.80% of all relevant type cases in the category, omission of articles constitutes the highest frequent errors. According to Chen (2006), nouns in Chinese do not need articles, yet are sometimes marked through measure words that are particular words to identify units of nouns. This language transfer might be the most important source of this repeatedly-occurring type of errors. Moreover, the omission of plural suffix-s (ranking the second with 19.46% of all cases in this category) might as well be due to the same reason, as according to Chen (2006), unlike nouns in English, nouns in Chinese remain unchanged when pluralized. The frequent occurrence of omission of prepositions has also been reported in the studies by Jiang (1995), Darus and Khor (2009), and Hou (2016). However, unlike the present study, the omission of articles, and plural suffix-s have not been reported as highly frequent errors in the other works in the literature.

*4.4 Addition*

According to Dulay, Burt, and Krashen (1982), addition errors are defined as "the presence of an item which must not appear in a well-formed utterance" (p.156). In the current study, addition errors were identified to be a total of 7625 cases which is equal to 12.01 % of all the errors found in this study. Table 7 shows the type, rank, and number, as well as percentage of each of these addition errors.



[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]

[ DOI: 10.52547/ijree.6.4.57 ]

Table 7. Details of the addition errors observed in this study

| Rank | Number of Errors | % | Categories |
|---|---|---|---|
| 1 | 2059 | 27.00 | Articles |
| 2 | 1547 | 20.29 | Unnecessary words |
| 3 | 1288 | 16.89 | Conjunctions |
| 4 | 857 | 11.24 | Prepositions |
| 5 | 571 | 7.48 | Auxiliaries |
| 6 | 540 | 7.09 | Pronouns |
| 7 | 518 | 6.80 | 'to be' |
| 8 | 245 | 3.21 | Possessive Adj. / 's |
| Sum | 7625 | 100 | |

As shown in Table 7, addition errors have been subcategorized into eight subgroups of errors with ***articles*** being the most frequent of all. Then, in Table 8, we have brought an example for each addition error subcategory taken from the students' compositions. In order to locate the error in these samples, the unnecessarily-added segment has been shown within [ ].

Table 8. Samples of students' addition errors

| Structure | Error Type |
|---|---|
| I don't want to leave [the] Taiwan. | (article) |
| It can make me feel happy [and happiness]. | (unnecessary words) |
| Though I never learned cooking, [but] it was my habit to watch the ways of my mother. | (conjunction) |
| We went to Kenting National Park last [on] weekend. | (preposition) |
| My sister likes to [can] cook for other people. | (auxiliary) |
| Success is not what you want [it]. | (pronoun) |
| In fact, it [is] obviously breaks the old thinking. | ('to be') |
| I think everybody's father is also a hero['s]. | ('s) |

*4.5 Others*

Further to the three highly-frequent main categories of errors explained before, and also in addition to the fourth group of "misordering" suggested by Dulay, Burt, and Krashen (1982), in the present study the authors came along other errors/mistakes which did not fit into any of the subcategories already mentioned. Nevertheless, it should be highlighted here that we postulate that we are again mostly dealing with rater's subjective point of view. In other words, given enough explanations (and imagination!), perhaps all errors can be categorized into already well-established taxonomies.

Let's have a look at the following sample:

* I will through action to makes comfort.


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

The structure above has been tagged as others (unclear structure). We assume this can be rightfully accepted and left there as it is, and that is because we cannot be sure of what the writer actually wanted to convey. However, giving it more time, we might also analyze and correct this structure as the following (and this is only one correction, as we might be able to think of many more possible corrections).

Correction: I will take action so that it will make me comfortable.

Another possibility might be that the student who made this error apparently translated Chinese into English directly, taking "through" as a verb, which is often translated in Chinese as "**通過**".

Error types: omission (verb); omission (article); misformation (run-on sentence); omission (subject/pronoun); misformtion (verb); omission (object/pronoun); misformation (part of speech).

A more confusing situation might be the "wrong vocabulary" errors which have been the second highest frequent errors among the "others" in the current study. Let's take a look at the following example:

* So many people are so happy to be his foe.

This structure was made in the context of describing a hero in life. Can we be sure that the word "foe" has been used in its literal meaning? Has the writer been ironic, or even sarcastic here? Or has he/she been just careless with the word choice? It is really difficult to be sure of either case.

Then, Table 9 lists the type, rank, and number, as well as percentage of each of these other errors as well as the misordering errors.

Table 9. Details of the "other" errors observed in this study

| Rank | Number of Errors | % | Categories |
|---|---|---|---|
| 1 | 3117 | 32.47 | Spelling and Capitalization errors |
| 2 | 2435 | 25.36 | Wrong vocabulary |
| 3 | 2198 | 22.89 | Misordering |
| 4 | 1851 | 19.28 | Awkward/Unclear structures |
| Sum | 9601 | 100 | |

As can be seen in Table 9, there are four main groups of errors with spelling and capitalization errors (sometimes referred to as *'mechanical errors'*) being the most frequent of all. Other errors include wrong vocabulary, misordering, and awkward/unclear structures, as shown in the examples taken from the students' writings below:

* I need music when my mind is [complexity]. (wrong vocabulary)

I need music when my mind is confused/tangled/disordered. (corrected structure)

* Most people prefer life city rather than life village. (misordering)

Most people prefer city life rather than village life. (corrected structure)

* I think gender without to success. (awkward/unclear structure)

The reason students made a lot of spelling mistakes/errors might have been due to several factors such as being careless, not knowing the true spelling, and so on. One interesting observation was that many of the spelling errors were due to wrong choice between or among homophonic words or minimal pairs or sets. Some examples were: whole/hole, see/sea, rite/right, seen/scene, site/sight, lung/long, will/wheel, sit/seat, rid/reed, and sweet/sweat. In addition, so many students made errors not capitalizing proper nouns, words at the beginning of sentences, and even subject pronoun "I." Word separation cases like 'can not' instead of cannot, or 'all ways' instead of always were also observed quite frequently. Moreover, misordering cases occurred mostly when connecting interrogatory structures, using adverbs of frequency, and using two-word verbs or idiomatic expressions.


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

## 5. Discussion

The findings of the present study confirm those of Chen (2002), Darus and Khor (2009), Darus and Subramanian (2009), and Zhan (2015) in that among Chinese students, errors of English tenses rank either the first or second highest. This observation on tense errors is also consistent with the results of the study by Divsar and Heydari (2017) conducted with Iranian participants. In this regard, however, the current study does not support that of Chan (2010)'s in which tense errors ranked among the lowest of all observed errors. Then, erroneous usage, omission, or addition of propositions were frequently detected and reported in the present study supporting those by Chen (2002), Darus and Khor (2009), Darus and Subramaniam (2009), Hou (2016), and Jiang (1995).

Although discovering the potential cause(s) of errors was not an objective of the current paper, based on several similarities or commonalities of the findings among the present study and those in the related literature, we might be able to postulate that L1 transfer has a significant impact on the kinds and frequency of errors. This assumption is in tandem with the findings of Bao (2015), Chan (2010), Darus and Khor (2009), Divsar and Heydari (2017), Jiang (1995), Lan, Lucas, and Sun (2019), and Pimpisa and Normah (2015).

## 6. Conclusion

The authors assume that there is no practical way to eliminate or avoid all possible errors made by English language learners. Actually, there is no need for this highly-idealistic goal, as even native speakers make language mistakes or errors. What the current study tried to do was to carefully collect, code, classify, and rank a limited sample of Taiwanese EFL university students' written errors in the course of 12 consecutive semesters of essay writing. We strongly believe that error analyses can be used to determine what a language learner still needs to think about, practice, and learn. Hence, we recommend English language learners and teachers to be aware of the common errors and try to remove them in their writing (and perhaps speaking). One practical way to do so might be to keep track of one's errors in the course of a few months of deliberately practicing writing. We suggest making and using a detailed individual error checklist based on the error taxonomy discussed in this article with dates to function as the learners' guide.

### *6.1 Implications of the Study*

The findings of the present study shed light on the types and frequency of English lexical and syntactic errors in writing made by Taiwanese EFL university students. As a result, and based on the proposed detailed and "user-friendly" taxonomy of the errors, as well as the suggested flexible approach towards removing or fixing these errors or mistakes, obvious pedagogical implications emerge. The current study might help language teachers, syllabus designers, and curriculum decision makers proceed with higher degrees of certainty and precision. The study can particularly enable them with a prognostic ability to predict (and hence, program for) typical areas of errors along with their gravity/severeness. Another pedagogical implication addresses language learning and/or proficiency assessment in that EFL test designers in high schools or universities can focus more on the problematic areas of the language, and so reduce the redundant or low-discriminatory test items which would otherwise be added. This very usage can also increase the validity of relevant English language aptitude or speed tests used in high school and university entrance exams.

### *6.2 Limitations/Delimitations of the Study*

- This study did not focus on tracking each individual language learner's probable progress throughout the course of time,
- The study was conducted in just one university in central Taiwan. Including other participants from other universities in the country would have produced some different results, and
- The coding and rating of the essays were predominantly done by the main author (who taught all the composition classes). The second author acted more as consultant and proofreader. Having more raters who could mark and code the essays discretely would have yielded more reliable results.

### *6.3 Recommendations for Further Research*

Based on the procedure and data presented in the current article, the authors wish to suggest the following topics for further research:


[ DOR: 20.1001.1.25384015.2021.6.4.5.3 ]
[ DOI: 10.52547/ijree.6.4.57 ]

- Investigating and tracking each individual EFL learner's writing proficiency variation and change through long periods of instruction time,
- Exploring potential causes of errors in detail,
- Examining differences in making errors among EFL learners with different demographic backgrounds, and
- Studying EFL learners' attitudes and feelings towards recognition of errors they make, as well as the remedies they think of to overcome them.

**References**

Angeli, C., & Valanides, N. (2009). Instructional effects on critical thinking: Performance on ill-defined issues. *Learning and Instruction, 19*(4), 322-334. https://doi.org/10.1016/j.learninstruc.2008.06.010

Baker, P. (2006). *Using corpora in discourse analysis*. London/ New York: Continuum.

Bao, X. (2015). Analysis on lexical errors in college English writing. *Canadian Social Science, 11*(12), 127-130. doi:10.3968/7775

Brown, D. (1994). *Principles of language learning & teaching*. Eaglewood Cliffs, New Jersey: Prentice Hall.

Chan, A. Y. W. (2010). Toward a taxonomy of written errors: Investigation into the written errors of Hong Kong Cantonese ESL learners. *TESOL Quarterly, 44*(2), 295-319. doi: 10.2307/27896726

Chen, L. L. (2006). The effect of L1 and CAI on grammar learning: An error analysis of Taiwanese beginning EFL learners' English essays. *Asian EFL Journal, 9*(2), 1-19.

Chen, Y. M. (2002). The problems of university EFL writing in Taiwan. *The Korea TESOL Journal, 5*(1), 59-79. https://koreatesol.org/content/korea-tesol-journal-5

Chuang, F.Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of L2 English produced by Chinese students. *Corpora, 1*(2), 251-271. doi:10.3366/cor.2006.1.2.251

Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly, 34*(3), 548–560. https://doi.org/10.2307/3587743 https://www.jstor.org/stable/3587743

Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System, 27*(1), 1-18. https://doi.org/10.1016/S0346-251X(98)00046-3

Corder, S. P. (1974). The significance of learner's errors. In Richards, JC (Ed) *Error analysis* (pp. 19-27). Singapore: Longman.

Darus, S., & Khor, H. C. (2009). Common errors in written English essays of form one Chinese students: A case study. *European Journal of Social Sciences, 10*(2), 242-253. https://www.researchgate.net/publication/235772389_Common_errors_in_written_English_essays_of_form_one_Chinese_students_A_case_study

Darus, S., & Subramaniam, K. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European Journal of Social Sciences, 8*(3), 483-495. https://www.researchgate.net/publication/235772401_Error_analysis_of_the_written_english_essays_of_secondary_school_students_in_Malaysia_A_case_study

Divsar, H., & Heydari, R. (2017). A corpus-based study of EFL learners' errors in IELTS essay writing. *International Journal of Applied Linguistics & English Literature, 6(3),* 143-149. http://dx.doi.org/10.7575/aiac.ijalel.v.6n.3p.143

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford: Oxford University Press.

Edrogan, V. (2005). Contribution of error analysis to foreign language teaching. *Mersin University Journal of the Faulty of Education, 1*(2), 261- 270. https://dergipark.org.tr/en/download/article-file/161018

Ellis, R. (1997). *Second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.

Fang, X., & Jiang, X. M. (2007). Error analysis and the EFL classroom teaching. *US-China Education Revie, 4*(9), 10-14. https://files.eric.ed.gov/fulltext/ED502653.pdf

Glassner, A., & Schwarz, B. B. (2007). What stands and develops between creative and critical thinking? Argumentation? *Thinking Skills and Creativity, 2*(1), 10-18. doi: 10.1016/j.tsc.2006.10.001

Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal, 20*(3), 465-480. https://www.jstor.org/stable/24157525

Halpern, D. F., & Sanders, B. R. (2004). *Teaching critical thinking skills across the curriculum.* [Participant Packet: December 1, 2004; 1:30 – 3:0 PM CT: STARLINK.]

Han, N. R., & Chodorow, M., & Leacock, C. (2006). Detecting errors in English article usage by non-native speakers. *Natural Language Engineering*, *12*(2), 115–129. doi:10.1017/S1351324906004190

Hart, S. (2017). *English exposed: Common mistakes by Chinese speakers*. Hong Kong University Press, Hong Kong.

Hou, H. I. (2016). Learner corpus and academic writing: Identifying the error patterns of Taiwanese EFL students. *Journal for the Study of English Linguistics, 4*(1), 19-30. doi: 10.5296/jsel.v4i1.9193

Jiang, M. C. (1995). An analysis of Chinese learners' errors in prepositions. *Journal of National Chiayi Institute of Agriculture, 41*, 187-201.

Keshavarz, M. H. (2008). *Contrastive analysis and error analysis*. Tahran: Rahnama Press.

Kim, M. S. (2009). *An error analysis of a learner corpus of written and spoken English produced by Korean university students* (Unpublished doctoral dissertation). Korea University, South Korea.

Kotsyuk, M. L. (2015). English language error analysis of the written texts produced by Ukraine learners: Data collection. *Cognitive Studies*, *15*, 389–395. https://doi.org/10.11649/cs.2015.027

Lan, G., & Lucas, K., & Sun, Y. (2019). Does L2 writing proficiency influence noun phrase complexity? A case analysis of argumentative essays written by Chinese students in a first-year composition course. *System*, *85*, 1-13. doi.org/10.1016/j.system.2019.102116

Leki, I. (1992). *Understanding ESL writers: A guide for teachers*. Portsmouth, NH: Boynton/Cook Publishers.

Mitchell, R., & Myles, M. (2004). *Second language learning theories*. New York: Hodder Arnold.

Mohamed, A. R., & Goh, L., & Wan, R. E. (2004). English errors and Chinese learners. *Sunway College Journal, 1*, 83–97.

Myles, F. (2005). Interlanguage corpora and SLA research. *Second Language Research, 21*(4), 373-391. https://www.jstor.org/stable/43103691

Nakagawa, T. (2011). Education and training of creative problem-solving thinking with TRIZ/USIT. *Procedia Engineering, 9*, 582-595. doi: 10.1016/j.proeng.2011.03.144

Pimpisa, R. N. P., & Normah, O. (2015). Understanding EFL students' errors in writing. *Journal of Education and Practice*, *6*(32), 99-106. https://files.eric.ed.gov/fulltext/EJ1083531.pdf

Popil, I. (2011). Promotion of critical thinking by using case studies as teaching method. *Nurse Education Today*, *31*(2), 204-207. https://doi.org/10.1016/j.nedt.2010.06.002

Pravec, N. (2002). Survey of learner corpora. *ICAME Journal, 26*, 81-114. http://korpus.uib.no/icame/ij26/pravec.pdf

Richards, J. C. (1992). *Error analysis: Perspectives on second language acquisition*. London: Longman.

Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of teaching and applied linguistics*. Edinburgh: Pearson Education.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). GB: Pearson Education Limited.

Ro¨mer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics, 31,* 205–225. doi: https://doi.org/10.1017/S0267190511000055

Selinker, L. (1972). INTERLANGUAGE. *International Review of Applied Linguistics in Language Teaching, 10* (3), 209-231. https://doi.org/10.1515/iral.1972.10.1-4.209

Sinclair, J. (2004). *How to use corpora in language teaching*. Amsterdam: John Benjamins.

Zhan, H. (2015). Frequent errors in Chinese EFL learners' topic-based writings. *English Language Teaching, 8*(5), 72-81. doi:10.5539/elt.v8n5p72 http://dx.doi.org/10.5539/elt.v8n5p72

Zheng, C., & Park, T. J. (2013). An analysis of errors in English writing made by Chinese and Korean university students. *Theory and Practice in Language Studies*, *3*(8), 1342-1351. doi:10.4304/tpls.3.8.1342-1351