English Language Teaching; Vol. 17, No. 6; 2024
ISSN 1916-4742 E-ISSN 1916-4750
Published by Canadian Center of Science and Education

Exploring the Impact of Online Translation on Writing Revision among Chinese Non-English Major Students

Yan Zuo[1], Mengyu Liu[2] & Zhengwei Pei[1]

[1] College of Foreign Studies, Nanjing Agricultural University, China

[2] Dancheng NO.1 Senior High School, China

Correspondence: Zhengwei Pei, College of Foreign Studies, Nanjing Agricultural University, China. E-mail: wei@njau.edu.cn

Received: April 12, 2024 Accepted: May 11, 2024 Online Published: May 13, 2024

doi: 10.5539/elt.v17n6p1 URL: https://doi.org/10.5539/elt.v17n6p1

Abstract

The present study examined the efficacy of online translation as an auxiliary revision tool for Chinese non-English majors, assessing its influence on writing performance and the students' perceptions of its role in the revision process. This study employed writing task, questionnaire survey, and semi-structured interview to examine the efficacy of online translation as an auxiliary revision tool for Chinese non-English majors, assessing its influence on writing performance and the students' perceptions of its role in the revision process. A total of 94 Chinese non-English majors who learn English as a foreign language (EFL) participated in the study. The results indicate that compared with those who revise their texts independently, those Chinese non-English majors who refer to online translation in the process of revision, made significantly more changes at the below-clause and clause and above levels in terms of revision domain, with increased additions and substitutions. Additionally, having access to online translation in the process of revision, Chinese non-English majors significantly increased their text length in final drafts, with decreased error rate, more low-frequency and sophisticated words, and varied lexical choices. Besides, virtually most of the participants use online translation frequently in English writing activities with a generally positive attitude towards the integration of online translation into English writing revision. For English teachers, they can consider allowing students to use online translation in the process of revision and giving students sufficient guidance on how best to realize its fullest potential. For Chinese non-English majors, they'd better keep improving their self-learning ability and double-check the online translation output by using other resources.

Keywords: online translation, machine translation, computer-assisted language learning, writing revision, writing performance

1. Introduction

The rapid evolution of modern information technology, including the Internet, big data, and artificial intelligence, is reshaping traditional teaching paradigms, methods, and learning approaches. Free Online Translation (OT) tools such as Google Translate (GT) have become widely accessible, offering improved translation quality. ChatGPT, as an example of artificial intelligence, further contributes to the technological landscape, facilitating language learning through conversation and text generation. With advancements in artificial intelligence and natural language processing, OT tools like Google Translate have seen substantial improvements in translation quality, making them indispensable resources for language learners worldwide.

Nowadays, an increasing number of English learners are relying on freely available OT tools to support their language learning endeavors. These tools are commonly utilized for completing course assignments, grasping the main ideas of complex texts, and generating suitable expressions for foreign language writing. Numerous studies have highlighted the potential benefits of OT in language learning, including improved metalinguistic awareness, reduced cognitive load, and decreased language learning anxiety (Niño, 2009; Garcia & Pena, 2011; White & Heidrich, 2013; Stapleton & Kin, 2019; Tsai, 2020; Lee, 2019). However, concerns have been raised by some language instructors regarding the quality of OT translations, academic integrity, and its potential impact on language acquisition (Clifford et al., 2013; Jolley & Maimone, 2015).

Writing is always a weakness for Chinese students (Qiao & Wang, 2020), they encounter difficulties in both

content ideation and linguistic expression when composing in English. Consequently, enhancing writing quality poses an ongoing challenge for English learners. Writing revision is a crucial aspect of the writing process that plays a pivotal role in improving the overall quality of written work. It has been widely recognized as an indispensable part of foreign language writing (Barkaoui, 2016). In foreign language writing, some empirical studies have found that OT-assisted foreign language writing has a positive significance in enhancing the richness of vocabulary and the complexity of syntactic structures (Cancino & Panes, 2021; Chon, Shin & Kim, 2021), and improving the accuracy of texts (Stapleton & Kin, 2019; Tsai, 2019, 2020; Lee, 2019). However, due to the lack of relevant research in China, the impact of online translation systems on Chinese non-English majors' writing performance remains unknown. While English majors' writing performance is also important to consider, our study aimed to address the needs of a broader group of students with varying levels of English proficiency and assess the generalizability of findings across different academic backgrounds.

Therefore, the present study explores the possibility of OT represented by GT as an auxiliary revision tool, namely, the impact of OT-assisted revision on Chinese non-English majors' writing performance. Specifically, this study is dedicated to probing into the following questions:

(1) What is the influence of online translation as an auxiliary revision tool on the frequency and types of revisions made in different aspects of writing, including revision domain, revision action, and revision effect?

(2) What is the impact of online translation as an auxiliary revision tool on Chinese non-English majors' English writing performance in terms of complexity, accuracy, and fluency?

(3) How do Chinese non-English majors use and perceive OT in the process of writing revision?

2. Literature Review

2.1 Writing Revision

Zhao (2007, p. 42) held that "revision can be defined as something done (i.e., the addition or deletion of a word or word string), in order to reach a certain goal (i.e., improving style, modifying the content, etc.), at a text level (i.e., word, phrase, clause, sentence, paragraph, or the whole text), on a certain text (i.e., pretext in the mind, or already written text), at a certain point of writing (i.e., during the process of writing or after finishing the draft), with a certain effect (i.e., improvement, neutral, decreasing effect, etc.) and with a certain cognitive cost." This definition comprehensively reveals the essence of writing revision. Self-revision can train students to carefully examine their output, think about their mistakes, and find solutions to them, which is conducive to cultivating students' self-reflection ability, autonomous learning and lifelong learning ability.

According to Stevenson, Schoonen, and Glopper (2006, p. 206), revisions can be categorized into three domains: below-word, below-clause, and clause and above. Below-word entails changes involving one or more characters, less than a whole word (letter). Below-clause involves changes affecting one or more words, but less than a clause (word and phrase). Clause and above include revisions impacting one or more clauses (clause, sentence, and paragraph). Besides, Stevenson, Schoonen, and Glopper (2006: 206) outline four main categories of revision actions: addition, deletion, substitution, and others.

Regarding the revision effect, studies often align with Ferris's (1997, p. 324) and Zhong's (2015) categorization, which includes positive, negative, and neutral revisions. Positive revision involves changes aimed at enhancing the quality of the text, such as improving language form, enhancing logical structure, ensuring accuracy of meaning, and enriching content. Conversely, negative revision refers to changes that diminish the quality of the text, often by introducing new errors. Neutral revision neither improves nor undermines the quality of the article and fails to address the original issue.

Table 1. Multidimensional revision taxonomy

Revision dimension	Main categories
Domain	below-word; below-clause; clause and above.
Action	addition; deletion; substitution; other.
Effect	positive revision; negative revision; neutral revision.

2.2 Measures of Writing Performance

It is generally believed that the language performance of second language learners can be measured from three dimensions: complexity, accuracy and fluency (CAF) (Polio, 1997; Wolfe-Quintero et al., 1998). Wolfe-Quintero et al. (1998) took the lead in proposing CAF as criteria for evaluating language performance and expression. In the field of L2 acquisition, fluency, accuracy, syntactic complexity and lexical complexity measures have been

widely used to measure the quality of L2 performance.

Complexity can be further divided into syntactic complexity and lexical complexity. The present study measures syntactic complexity in terms of mean length of T-unit (MLT), dependent clause per T-unit (DC/T) and complex nominal per T-unit (CN/T). These three indicators represent the three dimensions, respectively sentences, clauses and phrases, in line with Bulté and Housen's (2012, p. 23) classification of L2 syntactic complexity. Laufer and Nation (1995, p. 307-322), systematically studied lexical complexity and its measurement indexes, believing that lexical complexity mainly consisted of four predictors: lexical uniqueness, lexical density (LD), lexical sophistication (LS) and lexical variation (LV). Read (2000) argued that in the context of L2 writing, lexical uniqueness was not suitable for evaluating learners' lexical development. Moreover, due to the instability of lexical uniqueness, "there are few studies on lexical uniqueness" (Bao, 2010). Most studies mainly measure LD, LS and LV (Bao, 2010). Therefore, this study follows the same practice.

According to Skehen, language accuracy refers to the quality of the language produced by learners in terms of the rule system of the target language (quoted from Ellis & Barkhuizen, 2012, p. 139). Many studies exploring the effect of GT on L2 writing also used the error rate (Fredholm, 2015; Tsai, 2020; Chon, Shin & Kim, 2021). As multiple errors were frequently found in a single T-unit that would skew the results, the error rate will be adopted in this study to measure accuracy.

Wolfe-Quintero et al. (1998) define fluency as the ability to produce written language quickly, coherently, appropriately, and creatively, which means to produce as many words and structures as possible in a limited time. Wolf-Quintero et al. (1998) thought that total word count and mean length of T-unit reflect writing fluency. But mean length of T-unit is generally used to measure syntactic complexity (Ortega, 2003). Qin and Bi (2012) also pointed out that composition length can predict writing quality. Therefore, the total number of words will be counted in the present study to measure fluency.

Over the past 20 years, numerous studies have employed the CAF framework to assess language proficiency, particularly in EFL writing. While researchers widely accept the validity of this framework, there is variability in the specific measures utilized across studies. Ultimately, the present study established a framework encompassing syntactic complexity, lexical complexity, accuracy, and fluency to provide a comprehensive overview of foreign language performance. The framework is presented in Table 2 below.

Table 1. The framework of CAF measures

Category	Measure	Formula
Fluency	Words (W)	the number of total words
Accuracy	Error rate	the number of errors/the number of words
Syntactic Complexity	Mean length of T-unit (MLT)	the number of words/the number of T-units
	Dependent clause per T-unit (DC/T)	the number of dependent clauses/the number of T-units
	Complex nominal per T-unit (CN/T)	the number of complex nominals/the number of T-units
Lexical Complexity	LV (LV)	the number of lexical word types/the number of lexical words
	Lexical word variation (LWV)	
	Uber index (Uber)	(logTokens)2/(logTokens-logTypes)
	Number of different words (NDW)	the number of word types
	LS (LS)	the proportion of academic words and off-list words
	LD (LD)	the number of content words/the number of total words

2.3 Related Studies

Although there is a lack of relevant research on the impact of online translation systems on Chinese non-English majors' writing revision, many researchers have studied the influence of OT use on foreign language writing output with different research designs.

Several researchers have explored the potential of OT as a supplementary tool for foreign language writing revision, akin to a dictionary (Godwin-Jones, 2015). Lee (2019) observed that OT provided personalized feedback and lexical alternatives, aiding students in identifying and rectifying mistakes. Jiménez-Crespo (2018, p. 184) noted that free OT has become "another de facto dictionary for language learners," widely utilized, requiring a deeper understanding of its utility.

Lee (2019) compared the initial English compositions of 34 Korean English majors with their final edited versions using OT. While the edited versions exhibited a notable reduction in lexical and grammatical errors compared to the originals, there was no statistical difference in lexical and syntactic complexity. However, the study's limitations include a small sample size and confounding variables, such as variations in OT systems and the use of other online resources, necessitating careful control of confounders in research design.

Tsai (2020) investigated Chinese students' perceptions of using GT as a translingual CALL tool in EFL writing, observing significant improvement in the revised English text over the original, particularly for non-English majors. Students who employed GT for revision demonstrated enhanced written language and content, with more academic words and fewer errors. Wang and Ke (2022) focused on low and intermediate Chinese EFL learners with weak vocabulary and grammar skills, finding that the experimental class, which revised English drafts with OT, outperformed the control class following traditional revision methods. Lee and Briggs (2021) explored the effects of GT on the revision process of Korean university students' academic writing, noting a significant reduction in errors across all error types in the final revised writing, particularly among higher proficiency students.

As evident from the aforementioned studies, these researchers employed a method where students first wrote in their native language (L1) and then translated it into their target language (L2). However, this approach has several drawbacks. Firstly, writing in native language will lead to separation from the cognitive process of foreign language writing, which is not conducive to the cultivation of thinking in English (O' Brien, 2018). Secondly, using the mother tongue during foreign language writing is typically discouraged due to the potential for negative transfer caused by cross-language differences (Göpferich, 2017). Lastly, these researchers make inaccurate presumptions about the L2 writing process, as they involve writing the L2 text equivalent to translating the L1 text (Chung & Ahn, 2022). Therefore, it remains to be seen whether OT can be used as a revision tool.

As indicated in the literature review, researchers commonly employed a method wherein students initially composed their texts in their native language before translating them into the target language, a practice with several limitations. Consequently, the present study endeavors to utilize OT at the post-writing stage. This approach integrates writing tasks, questionnaires, and interviews to examine the use of OT in the revision process and to gauge participants' perceptions of its role in writing revision.

3. Method

3.1 Participants

As for the subjects of the writing task, a total of 100 freshmen from two intact classes participated in this study. In order to assess the efficacy of OT as an auxiliary writing revision tool, one of the two classes was randomly selected as the experimental class and the other as the control class. Students in the experimental group employed GT to refine their first drafts into final versions, whereas those in the control group revised and improved their drafts on their own without utilizing GT. On average, they were 19 years old. During the data collection, they were in the second semester of their first year in university. They have been learning English as a foreign language for 9 to 10 years, mainly in English classes.

In order to examine participants' writing ability, the researcher and two graduate students majoring in foreign languages and literature graded the 100 participants' first drafts which were completed without referring to any resources in accordance with rating rubrics in CET-4 writing. The overall score for writing is 15. The average score given by two raters was assigned as the final score if less than 3 points lie between. Otherwise, the essay had to be rated again by the third rater, and the final score was the average of the two closer scores. When grading their compositions, the researcher found that 4 compositions in the experimental class and 2 compositions in the control class digressed from the subject. Consequently, these 6 compositions were excluded from the research. Therefore, in the end, 94 students were identified as participants. Besides, 21 participants whose writing scores were above the average or equal to the average were labeled as high-achievers in the experimental class, and 27 participants with scores less than the mean were regarded as low-achievers in the experimental class.

As for the participants of the Questionnaire, a total of 100 participants from the two classes completed the survey about their daily use of OT. They are freshmen from different majors. There are 37 boys and 63 girls, with an average age of 19.

As for the respondents of interviews, 14 experimental participants from different English proficiency groups were selected by convenience sampling to answer 5 open-ended questions.

In Table 3, an independent sample t-test analysis of their first draft score indicated that there were no significant differences in English writing proficiency between the experimental class and control class (P=0.520). According to Table 4, the difference in the scores of the two proficiency groups was statistically significant (P=0.000). The writing proficiency of the high-achiever group was significantly higher than that of the low-achiever group.

Table 2. Independent samples t-test of the first draft score of two classes

	Class	N	Mean	SD	t	p
first draft score	control class	46	9.78	2.01	0.645	0.520
	experimental class	48	9.52	1.92		

* $p<0.05$ ** $p<0.01$*** $p<0.001$

Table 3. Independent samples t-test of the first draft score of high-achiever group and low-achiever group in the experimental class

	Group	N	Mean	SD	t	p
first draft score	high-achiever	21	11.33	1.24	10.067	0.000**
	low-achiever	27	8.11	0.89		

* $p<0.05$ ** $p<0.01$*** $p<0.001$

3.2 Instruments

3.2.1 Writing Task

Following the completion of Unit 2, which explores art and nature themes including renowned Western and Classical Chinese paintings, participants were assigned the task of composing an English essay titled “A Comparison between Chinese and Western Paintings.”

3.2.2 Questionnaire

The Questionnaire was designed with reference to the questionnaires of Giannetti (2016), Ahn and Chung (2020), and Tsai (2020). It’s made up of three parts. The first part gathers background information from students, including their name, gender, grade, major, age, CET-4 score, and self-perceived English proficiency. The second part consists of five multiple-choice questions aimed at assessing students’ prior experiences with OT in English writing, specifically, how often they use OT in daily EFL writing, why and how they use OT in English writing as well as how they use OT in revision. These questions explore the frequency, purpose, and methods of OT use in their EFL writing, as well as its role in revision. The third part employs a five-point Likert Scale (ranging from 1 to 5) to gauge students’ attitudes toward OT’s efficacy in English writing revision. It contains 17 sub-questions focusing on 5 dimensions, namely, students’ English proficiency, error correction, expression search, emotional adjustment, and overall evaluation.

Only students from the experimental class, who utilized OT during the revision process, completed the third part of the questionnaire. The reliability analysis yielded a Cronbach’s alpha value of 0.96, indicating high reliability. Additionally, the validity assessment, with a KMO value of 0.93 and a p-value less than 0.001, confirmed the questionnaire’s high validity. In conclusion, the questionnaire demonstrates satisfactory reliability and validity, ensuring the collection of effective data for the study.

3.2.3 Semi-structured Interview

The semi-structured interview mainly focuses on 5 open-ended questions in the interview outline covering the advantages and disadvantages of using OT in revision, the procedures for correcting errors during revision, and the reasons underlying the changes. The interview was conducted in the form of voice calls. With the consent of the participants, the interview was recorded. To facilitate communication, the interview was conducted in Chinese. Students’ answers were transcribed and translated into English by the researcher.

3.3 Data Collection

The experiment was conducted during regular class time. In Step 1, participants complete their writing task. Each participant was asked to produce a minimum of 150 words within 30 minutes, using either a laptop or tablet with Microsoft Word. Both the experimental and control groups were required to independently complete their writing without accessing any reference materials, telephones, or any other electronic devices. Participants saved their initial drafts after writing.

In Step 2, following the completion of their writing tasks, participants were allocated 20 minutes for revision. Both the experimental and control groups received revision guidance. The revision criteria were quoted from Farzi (2016). Crucially, while revising, the experimental group exclusively utilized GT to refine their initial drafts into final versions. They received a brief tutorial on how to leverage GT for editing purposes. Conversely, students in the control group copied their first drafts to the next page and revised and improved their drafts by themselves within 20 minutes. They were strictly prohibited from consulting any external resources. This ensured the control of extraneous variables, maintaining a sole independent variable in the experiment.

In Step 3, participants saved their Word documents using their Chinese name, serial number and CET-4 score, and emailed both their initial drafts and final revisions to the researcher.

In Step 4, students from both classes completed the questionnaire.

In Step 5, 14 experimental participants underwent interviews. Subsequently, the first drafts and final drafts uploaded by experimental students were analyzed by three types of online computational assessments.

Overall, the questionnaire was completed by 100 subjects from both classes. A total of 94 first drafts and 94 final drafts were subjected to comparative analysis. Following this, semi-structured interviews were conducted with 14 participants from the experimental class. Recording software was employed to capture their responses, with each interview lasting approximately 5 to 8 minutes. Transcriptions of the students' recordings yielded texts ranging from 152 to 507 words in length.

3.4 Data Analysis

3.4.1 Data Coding

This study employed a mixed-method approach, integrating quantitative text analysis of students' initial and revised drafts, along with questionnaires and qualitative analysis of their interviews. Regarding the text analysis, several steps were undertaken, beginning with data coding.

Initially, revisions were coded based on three main categories: domain, action, and effect. Below-word revisions were denoted as 1D, below-clause as 2D, and clause and above as 3D. Positive revisions were marked with a check (√), negative revisions with a cross (×), and neutral revisions with an asterisk (*). Additionally, revisions were identified and reviewed by another graduate student specializing in foreign languages and literature, ensuring consistency. Any discrepancies were resolved through discussion until a consensus was reached. The frequency of each level and type of revision was then calculated and recorded in Excel.

Furthermore, to assess the accuracy of the first and final drafts, an error analysis of the written texts was conducted. Lexical, syntactic, and orthographical errors in the participants' drafts were meticulously examined. Ferris' (2011, p. 102) error taxonomy framework was consulted to establish error categories. Each piece of writing was reviewed, and errors were identified and categorized. Through this iterative process, 11 distinct error categories were identified: word choice, verb form, word form, subject-verb agreement, articles, noun endings, pronouns, prepositions, sentence structures, mechanics, and mistranslations.

Finally, errors were analyzed quantitatively and categorized. The researcher identified errors in participants' first and final drafts based on the error taxonomy. Another researcher also reviewed and verified the errors, with any discrepancies resolved through negotiation. This collaborative approach ensured the identification and resolution of differences during the analysis process.

3.4.2 Analytical Tools

To analyze participants' linguistic features in their first and final drafts, three analyzer tools were employed. The Second Language Syntactic Complexity Analyzer (L2SCA), developed by Lu (2010), was utilized to assess the syntactic complexity of written English samples, calculating indices such as MLT, DC/T, and CN/T. VocabProfiler was utilized to determine the number of words from specific lists in the text, including K1 (the most frequent 1000 word families), K2 (the second most frequent 1000 word families), AWL (the Academic Word List), and Off-list Words (words not appearing on these lists). Notably, spelling errors needed correction before pasting the texts into VocabProfiler for accurate analysis. The Lexical Complexity Analyzer (LCA),

developed by Ai Haiyang and Lu Xiaofei, is utilized to compute indices such as LWV, Uber, and NDW.

3.4.3 Statistical Analysis

Quantitative analysis involved several steps using SPSS version 25.0. Initially, descriptive statistics were employed to determine the frequencies and means of each error and revision type. Then, paired sample t-tests were conducted to assess the effect of OT use on writing revision, comparing error rates and linguistic complexity of L2 texts between first drafts and revised texts. Paired sample t-tests were also conducted to examine the main effect of OT use on students with varying proficiency levels. Independent t-tests were utilized to compare differences between the experimental and control classes. Besides, survey responses were descriptively analyzed by calculating the percentages of multiple-choice question answers.

4. Results

4.1 Influence of Online Translation as an Auxiliary Revision Tool on Frequency and Types of Chinese Non-English Majors' Writing Revision

The frequency of revision is calculated by the average number of revisions per 100 words (Stevenson et al., 2006). Experimental students made 7.77 revisions per 100 words on average in the final texts, while for students in the control class, the mean of the total frequency of revisions in the final drafts was 3.70 per 100 words.

4.1.1 Revision Domain

The results of the independent sample t-test reveal that the experimental class demonstrated a notably higher frequency of revisions at the below-clause level compared to the control class (M=5.17, M=2.39, p=0.000). Furthermore, experimental writers exhibited a higher frequency of revising large chunks of text, such as clauses and sentences, in comparison to the control class (M=1.49, M=0.55, p=0.000). However, no statistical difference was observed in the revision of below-word between the two classes (p=0.099>0.05).

Table 4. Independent samples t-test of revision domain in two classes

Domain	Class	N	Mean	SD	t	p
below-word	control class	46	0.76	1.01	-1.667	0.099
	experimental class	48	1.11	1.02		
below-clause	control class	46	2.39	1.97	-6.360	0.000***
	experimental class	48	5.17	2.26		
clause and above	control class	46	0.55	0.87	-3.775	0.000***
	experimental class	48	1.49	1.47		

* $p<0.05$ ** $p<0.01$*** $p<0.001$

4.1.2 Revision Action

Results reveal significant differences in addition and substitution between the two classes (p=0.042, p=0.000). Remarkably, the experimental class exhibited a notably higher frequency of substitutions (M=6.53, M=2.78) and additions (M=1.04, M=0.69) compared to the control class. However, there is no significant difference in deletion and other revision action between the two classes (p=0.493, 0.086 > 0.05).

Table 5. Independent samples t-test of revision action in two classes

Action	Class	N	Mean	SD	t	p
Addition	control class	46	0.69	0.79	-2.067	0.042*
	experimental class	48	1.04	0.86		
Deletion	control class	46	0.24	0.61	0.688	0.493
	experimental class	48	0.17	0.31		
Substitution	control class	46	2.78	2.32	-7.622	0.000***
	experimental class	48	6.53	2.45		
Other	control class	46	0.00	0.00	-1.752	0.086
	experimental class	48	0.03	0.11		

* $p<0.05$ ** $p<0.01$*** $p<0.001$

4.1.3 Revision Effect

Results show that the experimental classes made significantly more positive revisions than the control class (M=5.88, M=2.62), as well as neutral revisions (M=1,05, M=0.36). No significant difference in the frequency of negative revision between the two classes was found (p=0.353 > 0.05).

Table 6. Independent samples t-test of revision effect in two classes

Effect	Class	N	Mean	SD	t	p
positive revision	control class	46	2.62	1.80	-7.555	0.000***
	experimental class	48	5.88	2.36		
negative revision	control class	46	0.72	0.65	-0.933	0.353
	experimental class	48	0.84	0.63		
neutral revision	control class	46	0.36	0.55	-5.055	0.000***
	experimental class	48	1.05	0.75		

* $p<0.05$ ** $p<0.01$*** $p<0.001$

4.2 Impact of Online Translation as an Auxiliary Revision Tool on Chinese Non-English Majors' English Writing Performance in Terms of Complexity, Accuracy and Fluency

4.2.1 Fluency

Given that all participants were allotted 30 minutes for writing and 20 minutes for revision simultaneously, text length serves as a suitable gauge of fluency. The independent samples t-test of writing fluency of final drafts in two classes indicated that the total number of words in the final drafts of the experimental class (M=209.81) was significantly higher than that of the control class (M=193.43).

Besides, although the participants were asked to write a composition in no less than 200 words, some students fell short of the requirement in final drafts (Experimental class: M=209.81, max=307, min=140; Control class: M=193.43, max=272, min=112). Especially, the control class wrote less than 200 words on average in final drafts.

4.2.2 Accuracy

The accuracy level was evaluated by the error rate (the number of errors divided by the number of words per text). That is to say, a high level of error rate indicates a low quality of accuracy.

A total of 2267 errors were identified across the students' first and final drafts in both classes. In the experimental class, errors decreased from 31.72% in initial drafts to 13.19% in final drafts. Similarly, the control class observed a decline from 30.08% to 25.01% in errors between initial and final drafts. The most common errors comprised word choice (n=427, 18.84%), mistranslation (n=326, 14.38%), mechanics (n=295, 13.01%), sentence structure (n=230, 10.15%), word form (n=207, 9.13%), and verb form (n=205, 9.04%), amounting to 74.55% of all errors. Additional error types included noun ending (n=185, 8.16%), preposition (n=157, 6.93%), article (n=76, 3.35%), subject-verb agreement (n=70, 3.08%), and pronoun (n=40, 1.76%).

Furthermore, as depicted in Table 3, the error rates in the final drafts of the experimental class are notably lower than those of the control class, except for articles and sentence structure (p=0.321, p=0.467). No significant differences were observed in the error rates of articles and sentence structure.

Table 7. Independent samples t-test of accuracy of final drafts in two classes

Measure	Class	N	Mean	SD	t	p
Error Rate	Control class	46	0.0636	0.0240	8.284	0.000***
	Experimental class	48	0.0294	0.0146		
Word Choice	Control class	46	0.0129	0.0078	4.624	0.000***
	Experimental class	48	0.0063	0.0059		
Verb Form	Control class	46	0.0056	0.0065	2.012	0.047*
	Experimental class	48	0.0032	0.0045		
Word Form	Control class	46	0.0059	0.0065	3.208	0.002**
	Experimental class	48	0.0023	0.0039		
S-V	Control class	46	0.0023	0.0035	2.789	0.007**
	Experimental class	48	0.0007	0.0016		
Article	Control class	46	0.0017	0.0035	0.998	0.321
	Experimental class	48	0.0011	0.0021		
Noun Ending	Control class	46	0.0056	0.0054	5.300	0.000***
	Experimental class	48	0.0010	0.0025		
Pronoun	Control class	46	0.0013	0.0027	2.409	0.019*
	Experimental class	48	0.0003	0.0011		
Preposition	Control class	46	0.0046	0.0057	3.285	0.002**
	Experimental class	48	0.0015	0.0030		
SS	Control class	46	0.0067	0.0075	0.731	0.467
	Experimental class	48	0.0057	0.0052		
Mechanics	Control class	46	0.0068	0.0076	3.860	0.000***
	Experimental class	48	0.0019	0.0038		
Mistranslation	Control class	46	0.0103	0.0069	3.627	0.000***
	Experimental class	48	0.0054	0.0061		

note. $p<0.05$ ** $p<0.01$*** $p<0.001$; S-V= Subject-verb agreement; SS = Sentence structure.

4.2.3 Complexity

Syntactic Complexity. The statistics of independent samples t-tests for syntactic complexity between the final drafts of the two classes showed no significant differences for all three measures of lexical complexity ($p=0.088$, 0.521, $0.259>0.05$).

Lexical Complexity. The independent samples t-test examining the lexical complexity of final drafts in both classes showed that there is no significant difference in LD ($p=0.641>0.05$). Regarding LS, the AWL words and Off-List words in final drafts were used more frequently in the experimental ($M=13.84$) and control classes ($M=12.09$). In terms of LV, the number of different words (NDW) used by the experimental group in the final drafts ($M=106.50$) is much higher compared to the control group ($M=97.74$). However, there was no significant difference in LWV and Uber index ($p=0.182$, $0.163>0.05$).

4.3 Chinese Non-English Majors' Use and Perception of Online Translation in Writing Revision

4.3.1 Use of Online Translation in the Process of Revision

According to the questionnaire survey, 38 students indicated that they regularly use OT to revise their compositions. For instance, 4 participants mentioned using OT to "translate sentences or phrases to find better expressions." Furthermore, 2 students wrote that "comparing my sentence with OT output, I found the grammatical errors in my text and then I tried to correct them."

In addition, interviews were conducted with 14 experimental participants to complement the quantitative findings. More than half of the respondents mentioned using OT to assist in correcting errors, such as spelling,

lexical, and grammatical errors. Furthermore, 8 participants reported using OT to search for expressions. For example:

I also prefer to look up phrases, which can make sentences more concise and understandable (No.11).

While I was editing my compositions, I used OT to find out the correct fixed collocations and phrases (No.20).

What's more, participants were required to tell the reason why they accepted or rejected an OT output in their revision and how they decided. Some of them answered this question by giving examples. For example:

I intended to write "Chinese paintings pay attention to artistic conception," but I didn't know how to express "artistic conception." In the process of revision, I used GT to search for the expression. Its output was "mood." I was skeptical about that, because in my opinion, "mood" is used to express emotion. Hence, I used Baidu Translation whose output is "artistic conception". I thought this one was better, so I adopted it (No.16).

Interviewees also discussed their criteria for rejecting OT output. Four interviewees mentioned that they would accept words and phrases they had previously learned but since forgotten. However, if provided advanced words and expressions unfamiliar to them, they would choose not to use them:

I rarely accepted those long and complex sentences. On the one hand, I can't write such complex sentence patterns by myself. On the other hand, I can't master the grammatical knowledge underlying the sentence (No.20).

4.3.2 Perception of the Effects of Online Translation in Writing Revision

To survey students' perceptions of the effects of OT-assisted revision, they were acquired to complete the five-point Likert scale with 17 sub-questions. The results of the third part of the questionnaire are set out in Table 4.

Table 8. Descriptive Statistics for the perception questionnaire

Dimension	Item	Min.	Max.	Mean	SD	Overall attitude
Students' ability of integrating OT into the revision process	1	2.00	5.00	3.87	0.86	4.00
	2	2.00	5.00	3.22	0.76	3.00
Students' attitudes towards the effect of OT-assisted revision on correcting errors	3	1.00	5.00	3.64	1.06	4.00
	4	3.00	5.00	4.15	0.78	4.00
	5	2.00	5.00	3.60	0.87	4.00
	6	2.00	5.00	3.91	0.75	4.00
Students' attitudes towards the effect of OT-assisted revision on enriching expressions	7	2.00	5.00	4.15	0.73	4.00
	8	2.00	5.00	3.58	0.92	4.00
	9	3.00	5.00	4.09	0.59	4.00
	10	2.00	5.00	3.76	0.79	4.00
Students' attitudes towards the effect of OT-assisted revision on adjusting emotion	11	1.00	5.00	3.87	0.94	4.00
	12	1.00	5.00	3.80	0.85	4.00
Students' overall perception of OT-assisted revision	13	1.00	4.00	2.04	0.82	2.00
	14	1.00	5.00	2.40	0.91	2.00
	15	3.00	5.00	4.09	0.73	4.00
	16	2.00	5.00	3.69	0.72	4.00
	17	3.00	5.00	3.82	0.64	4.00

Question 1 aimed to assess students' ability to identify errors in OT output. The findings indicated that 70.91% of students agreed or strongly agreed that they could identify these errors. In Question 2, students were asked about their confidence in fixing errors produced by OT. The results showed that only 38.18% of students felt confident in correcting these errors.

Questions 3 to 6 focus on the impact of OT-assisted revision on error correction, encompassing mechanics errors, lexical errors, grammatical errors, and sentence meaning verification. Results indicated that 61.82% of

participants acknowledged that utilizing OT aided in identifying spelling, punctuation, and capitalization errors. 76.36% of participants hold a positive view on OT's effectiveness in detecting and rectifying lexical errors. However, only 56.37% of participants believed that OT was helpful in identifying and correcting grammatical errors. Nonetheless, 74.55% of respondents recognized OT's positive impact on verifying sentence meaning.

Questions 7 to 10 explore students' attitudes towards the effect of OT-assisted revision on enriching expressions. 83.64% of respondents reported that utilizing OT facilitated the discovery and utilization of low-frequency and academic vocabulary. Similarly, 87.27% of students expressed that OT aided in finding suitable alternative expressions during composition revision. Moreover, 69.1% of respondents viewed OT as a beneficial tool for enhancing essay content and expression of ideas. However, fewer students (54.54%) believed that OT-assisted revision contributed to the utilization of more complex and varied sentence structures.

Questions 11 and 12 delve into the emotional impact of OT-assisted revision, focusing on whether it alleviates anxiety related to English writing. Around 70% of participants expressed the belief that utilizing OT during revision could mitigate anxiety and pressure, thereby boosting self-confidence and self-efficacy in English writing.

Questions 13 to 17 aimed to gauge participants' overall evaluation of OT-assisted revision. Items 13 and 14 were reverse questions, assessing perceptions of OT accuracy and its contribution to English learning. Results indicated that over 70% of participants found OT tools helpful for English writing. A significant majority expressed a positive attitude toward OT-assisted revision (75.55%), expressed a desire to learn effective OT use (78.18%), and indicated intent to continue using OT for writing revision (69.09%).

5. Discussion

5.1 What is the Influence of OT as an Auxiliary Revision Tool on the Frequency and Types of Revisions Made in Different Aspects of Writing, Including Revision Domain, Revision Action, and Revision Effect?

The first question is dedicated to reporting a comparative study between student independent revision and OT-assisted revision, specifically analyzing revision domain, action, and effect. To standardize the revision data and mitigate the influence of composition lengths, the frequency of revision is calculated by the average number of revisions per 100 words (Stevenson et al., 2006). Experimental students made an average of 7.77 revisions per 100 words in the final texts, whereas students in the control class averaged 3.70 revisions per 100 words in the final drafts.

The analysis of revision domains revealed that the experimental class predominantly engaged in below-clause revisions in their final drafts, aligning with findings from Lee (2019) and Lee and Briggs (2021). Additionally, there was substantial revision at the word and phrase level, indicating that students used OT not merely for direct translation but to seek context-appropriate words and authentic expressions. Notably, they did so thoughtfully and critically, rather than uncritically accepting OT suggestions. However, fewer revisions were observed at the clause and above levels, indicating a limitation of OT in facilitating macro-revisions. OT's inability to provide feedback beyond the sentence level, as noted by Lee (2019), hinders its usefulness for making discourse-level or content-related revisions. This finding also supports prior research by Zhong (2015) and Bai and Wang (2018), suggesting that EFL learners tend to focus on surface-level errors rather than deeper textual or content-related revisions.

Regarding actions of revision, both classes exhibited a preference for substituting words and phrases and adding new ideas to their texts, with fewer instances of deletions and other revisions. This finding aligns with previous research by Stevenson et al. (2006) and Bai and Wang (2018), highlighting substitution as the most common revision action in foreign language writing. Deletions were infrequent, likely due to students' efforts to enrich their compositions. Moreover, the experimental class demonstrated significantly more substitutions and additions than the control class, illustrating the scaffolding role of OT in the revision process. Access to OT facilitated expression retrieval and language verification for students.

With respect to the effects of revision, the experimental classes demonstrated a higher frequency of both positive and negative revisions compared to the control class. This underscores two key points. Firstly, it's essential to recognize that OT is not infallible; it may produce grammatical errors and awkward literal sentences. Additionally, the revisions made by experimental students were not always accurate, particularly challenging for non-English majors with limited language proficiency. Their compositions often contained fragments and run-on sentences, less accurately translated by OT (Lee, 2019). As noted by Chen (2019), students lacked objective evaluation criteria for OT output, and some students with low English proficiency were unable to make correct judgments on the quality of translation.

Chinese non-English majors who referred to OT in the process of revision made significantly more positive changes than those who did not, which shows the supporting role of OT as an external resource. However, since the experimental class made remarkably more revisions in total than the control class, they also had a higher frequency of negative revisions, which indicates that OT is not perfectly accurate, and it is difficult for Chinese non-English majors with low language proficiency to choose the proper language to use. As for the revision domain and action, experimental students made changes most frequently at the level of below-clause as well as made the most substitution changes, through which it can be inferred that students did not simply adopt the clauses and sentences provided by OT, mainly using OT to search for words and phrases. Access to OT provides Chinese non-English majors with chances to retrieve certain expressions and verify their language, so as to make more revisions.

5.2 What is the Impact of OT as an Auxiliary Revision Tool on Chinese Non-English Majors' English Writing Performance in Terms of Complexity, Accuracy and Fluency?

The second question investigates the influence of OT use on non-English majors' writing performance in terms of CAF. The statistical analysis revealed a significant improvement in writing fluency when OT was employed during revision. This led to an increased total word count among non-English majors, aligning with prior research (Garcia & Pena, 2011; Kol et al., 2018; Tsai, 2020; Chon, Shin & Kim, 2021; Cancino & Panes, 2021). However, experimental students, initially falling short of the 200-word requirement in first drafts, met this criterion in final drafts by utilizing OT to search for specific words and expressions. OT serves as an external vocabulary source, enabling learners to enrich their written output. By referencing OT, students are encouraged to select words familiar to them, facilitating detailed elaboration of their intended information (Lee, 2019; Chon, Shin & Kim, 2021).

Regarding accuracy, results revealed that final drafts written by experimental learners had significantly lower error rates compared to those by the control class. These findings align with previous studies (Giannetti, 2016; Lee, 2019; Lee & Briggs, 2021; Chung & Ahn, 2022; Wang & Ke, 2022), underscoring the positive impact of OT on revision for EFL learners. Overall, OT serves as a helpful tool for students, reminding them to do error detection and correction.

In terms of complexity, results showed no significant difference between the final drafts of the two classes. Prior researches present conflicting findings regarding the impact of OT on syntactic complexity. Some studies suggest a positive effect on utilizing varied and sophisticated grammatical structures (Cancino & Panes, 2021; Chon, Shin & Kim, 2021), while others question its advantage in this regard (Fredholm, 2015; O' Neill, 2016; Lee, 2019; Chung & Ahn, 2022; Wang & Ke, 2022). Given the lower rate of revisions in clauses and above compared to below-clause, it appears that OT may not significantly contribute to making sentence structures more intricate.

Firstly, the statistical analysis of lexical complexity revealed varied results across three aspects. Regarding LD, there was no statistically significant difference observed between the first drafts and final drafts within the experimental class, nor between the final drafts of the two classes. This indicates that the students' LD in the final drafts remained unchanged after revision with OT (Tsai, 2019; Lee, 2019; Chung & Ahn, 2022). It's worth noting that while the total word count significantly increased in the experimental class, so did the number of content words, which may results in unchanged LD values.

Secondly, concerning LS, there was a significant difference between the final drafts of the experimental class and the control class. This suggests that access to OT during revision aided students in improving lexical sophistication. This finding aligns with previous studies indicating the significant benefit of OT for LS (Kol, 2018; Chon, Shin & Kim, 2021; Wang & Ke, 2022). The notable increase in the proportion of low-frequency words indicates that OT served as a reference source, facilitating the substitution of common and basic words with more advanced ones. Although these advanced words may not have been previously used by students, they were likely exposed to them (Kol, 2018). The increased exposure to advanced words suggested by OT may prompt students to incorporate them into their active vocabulary, thus broadening their productive vocabulary.

Thirdly, regarding LV, the mean LWV, NDW, and Uber values of the final drafts of the experimental class were significantly higher than those of the first drafts. Additionally, a significant difference in the number of different words between the two classes was identified. Therefore, it can be argued that allowing the use of OT in the revision process is likely to support students in varying their lexical choices. Although lexical diversity was calculated using different measures in some studies, such as textual lexical diversity and vocabulary diversity, they yielded similar findings (Tsai, 2019; Chung & Ahn, 2022). Similarly, Farzi (2016) observed the writing behavior of university non-English majors and found that learners' writing showed increased vocabulary variety

approximately 67% of the time when they utilized OT tools.

As a whole, OT-assisted revision has a positive effect on Chinese non-English majors' writing performance. Specifically, having access to OT in the process of revision, experimental students significantly increased their text length in final drafts, decreased the error rate of first drafts, utilized more low-frequency and sophisticated words and varied their lexical choices to use an increasing variety of distinct words. However, the benefits of using OT in syntactic complexity and lexical density were not clear, since there were no statistical differences in the number of subordinate clauses, complex nominals and longer sentences, as well as LD between two classes. Overall, getting OT involved in the revision, learners achieved better English writing performance in final drafts compared with first drafts.

5.3 Chinese Non-English Majors' Use and Perception of Online Translation in Writing Revision

The third question aimed to elucidate the frequency, reasons, and use of OT utilization in English writing tasks. The findings revealed a high prevalence of OT usage, with a significant proportion of participants reporting occasional or frequent reliance on this resource, consistent with existing literature (White & Heidrich, 2013; Clifford et al., 2013; Jolley & Maimone, 2015; Alhaisoni & Alhaysony, 2017; O' Neill, 2019). For instance, Clifford et al. (2013) surveyed 905 Romanian students, of whom 71% acknowledged sometimes (39%) or often (32%) using OT. Similarly, in a study involving 128 Spanish-speaking students (Jolley & Maimone, 2015), 74.22% of respondents reported regular (35.92%) or occasional (38.28%) OT use.

Students cited various reasons for utilizing OT in writing tasks. Notably, 89.90% of respondents cited "low English proficiency" as a primary reason, suggesting that students turned to OT due to their perceived inadequacies in completing writing tasks proficiently. This underscores both the reliance on OT as a supportive tool in English writing and students' lack of confidence in their writing abilities. Murtisari et al. (2019) similarly observed high response rates for "text difficulty" (32.1%) and "low confidence in L2 writing ability" (35.7%). Psycholinguistically, Selcuk et al. (2019) proposed that low confidence and heightened anxiety regarding foreign language writing could drive the use of MT.

Regarding the use of OT in English writing activities, many participants reported utilizing it primarily during the pre-writing and while-writing stages. At the pre-writing stage, most users searched for English equivalents of vocabulary related to their writing topics, aiming to overcome vocabulary barriers before composing. Similarly, during the writing process, students heavily relied on OT to find single words or phrases they struggled to express in English. This aligns with previous research indicating that students predominantly use MT for translating individual words or short phrases rather than longer segments (Jolley & Maimone, 2022). OT's accessibility, speed, convenience, and currency make it a preferred tool over traditional paper dictionaries.

Based on the empirical data collected through questionnaires, despite being aware of its limitations, virtually most non-English majors use OT frequently in English writing activities, primarily due to their limited English proficiency and OT's feature of being fast and convenient. As for the purpose of using OT in English writing, an overwhelming majority of Chinese non-English majors employ OT to look up words, phrases and idiomatic expressions. Similarly, in terms of segment length, students use OT most frequently to translate individual words or short phrases compared to paragraph-length or longer segments. More than half of the participants also used OT to verify whether their ideas were successfully conveyed. With regard to the writing phase using the OT tools, a large number of students referred to OT while they were writing, with a 34.48% minority using them at the revision stage. Regarding how learners use OT in the process of revision, the answers to open-ended questions and interviews displayed that they used OT to search for advanced words and set phrases, verify the meaning of their English sentences, and detect and correct errors.

Chinese non-English majors' perceptions of the effects of OT in revision are also examined, which indicates that experimental students hold a generally positive attitude towards the integration of OT into EFL writing revision. They believed that OT-assisted revision could improve their writing performance, and especially help correct spelling, lexical and grammatical errors, substitute common words with varied and advanced words, enrich the content, and alleviate anxiety and tension in English writing. However, it's also found that a majority of students were able to recognize incorrect output, but were not able to correct the errors, which is a proof of the finding that the experimental class made significantly more negative revisions than the control class. Therefore, Chinese EFL learners must constantly improve their English ability and have a better understanding of OT's benefits, drawbacks and strategies to minimize the errors.

6. Conclusion

This study investigates the utility of online translation (OT) as an auxiliary revision tool among Chinese non-English majors, exploring its impact on writing performance and students' perceptions of its role in revision.

While the impact of OT on syntactic complexity and lexical density remains inconclusive, its overall influence on writing performance is positive. Chinese non-English majors who utilized OT during revision made significantly more positive changes compared to those who did not, indicating its supportive role as an external resource. However, they also exhibited a higher frequency of negative revisions, suggesting that OT is not entirely accurate. Most revisions occurred at the below-clause level, predominantly through substitution, indicating that students primarily used OT to search for words and phrases rather than adopting its provided clauses and sentences. Access to OT enabled students to retrieve expressions and verify language, facilitating more revisions. Survey and interview data also revealed that many students were able to identify incorrect output, highlighting the need for continuous improvement in English proficiency and a better understanding of OT's benefits and limitations.

For EFL learners in China, continuous improvement in language skills is essential. While OT can be a valuable resource, it's important to recognize that it may contain grammatical errors. Therefore, a solid grasp of English is necessary to comprehend OT output and make necessary revisions and corrections effectively. Additionally, learners who rely on OT for vocabulary and expressions can also utilize additional resources such as online dictionaries, corpus tools, etc.

For English teachers, they should consider incorporating OT as a CALL tool during the revision process, acknowledging its increasing relevance in the digital age. Rather than disregarding OT entirely, educators should familiarize themselves with its capabilities and limitations to effectively guide students in its usage. Teachers should educate students on OT's strengths and weaknesses, offering guidance on maximizing its benefits and advising them to cross-reference OT output with other resources for accuracy. Additionally, educators should emphasize macro-level revisions, focusing on discourse and content coherence, while also encouraging students to assess the content, layout, and coherence of their texts.

Nevertheless, the study has limitations. It focused solely on evaluating written products in terms of CAF, without considering broader aspects of writing such as cohesion and organization typically assessed by human raters. Future research could involve human raters to evaluate these aspects and explore how language teachers perceive the impact of OT-assisted revision. Additionally, the study relied on self-reported frequency and process of OT use, in which students may conceal their improper use of it. Future research could employ methods such as screen casts and stimulated recalls to examine learners' actual use and integration of OT in revision.

7. Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Alhaisoni, E., & Alhaysony, M. (2017). An investigation of Saudi EFL university students' attitudes towards the use of Google Translate. *International Journal of English Language Education, 5*(1), 72-82. https://doi.org/10.5296/ijele.v5i1.10696

Bai, Y., & Wang, J. J. (2018). Effects of feedback modes on EFL students' revision process and text quality. *Journal of Northeast Normal University (Philosophy and Social Sciences), 292*(02), 134-139.

Bao, G. (2010). A comparative study of EFL learners' linguistic complexity. *Modern Foreign Languages (Quarterly), 33*(02), 166-176+219.

Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal, 100*(1), 320-340. https://doi.org/10.1111/modl.12316

Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA, 32*, 21. https://doi.org/10.1075/lllt.32.02bul

Cancino, M., & Panes, J. (2021). The impact of Google Translate on L2 writing quality measures: Evidence from Chilean EFL high school learners. *System, 98*, 102464. https://doi.org/10.1016/j.system.2021.102464

Chen, Z. Y. (2019). Research on the strategies of integrating machine translation into college English teaching. *Journal of Yancheng Institute of Technology (Social Science Edition), 32*(01), 74-77.

Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System, 96*, 102408. https://doi.org/10.1016/j.system.2020.102408

Chung, C. S. (2015). Research on independent writing revision based on online automatic feedback. *Journal of the PLA Foreign Languages Institute, 38*(04), 81-88.

Chung, E. S., & Ahn, S. (2022). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning, 35*(9), 2239-2264. https://doi.org/10.1080/09588221.2020.1871029

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning?. *@ tic. revista d'innovació educativa,* (10), 108-121. https://doi.org/10.7203/attic.10.2228

Ellis, R., & Barkhuizen, G. P. (2012). *Analysing learner language.* Shanghai Foreign Language Education Press.

Farzi, R. (2016). *Taming Translation Technology for l2 Writing: Documenting the Use of Free Online Translation Tools by Esl Students in a Writing Course* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *Tesol Quarterly, 31*(2), 315-339. https://doi.org/10.2307/3588049

Ferris, D. (2011). *Treatment of error in second language student writing.* University of Michigan Press. https://doi.org/10.3998/mpub.2173290

Fredholm, K. (2015). Online translation use in Spanish as a foreign language essay writing: Effects on fluency, complexity and accuracy. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas,* (18), 7-24. https://doi.org/10.26378/rnlael918248

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning, 24*(5), 471-487. https://doi.org/10.1080/09588221.2011.582687

Giannetti, T. R. (2016). Google translate as a resource for writing. *Education Masters,* p. 353.

Godwin-Jones, R. (2015). Contributing, creating, curating: Digital literacies for language learners. *Language Learning & Technology, 19*(3), 8-20. Retrieved from http://llt.msu.edu/issues/october2015/emerging.pdf

Göpferich, S. (2017). Cognitive functions of translation in L2 writing. *The handbook of translation and cognition,* 402-422. https://doi.org/10.1002/9781119241485.ch22

Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. *Learn languages, explore cultures, transform lives,* 181-200.

Jolley, J. R., & Maimone, L. (2022). Thirty years of machine translation in language teaching and learning: A review of the literature. *L2 Journal, 14*(1). https://doi.org/10.5070/L214151760

Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in academic writing courses?. *The EuroCALL Review, 26*(2), 50-57. https://doi.org/10.4995/eurocall.2018.10140

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics, 16*(3), 307-322. https://doi.org/10.1093/applin/16.3.307

Lee, G., & Lee, B. (2021). Characteristics of eight Korean EFL young adults' L2 writing using machine translation programs. *Multimedia-Assisted Language Learning, 24*(2). https://doi.org/10.15702/mall.2021.24.2.34

Lee, S. M. (2019). The impact of using machine translation on EFL students' writing. *Computer assisted language learning, 33*(3), 157-175. https://doi.org/10.1080/09588221.2018.1553186

Lee, S. M., & Briggs, N. (2021). Effects of using machine translation to mediate the revision process of Korean university students' academic writing. *ReCALL, 33*(1), 18-33. https://doi.org/10.1017/S0958344020000191

Murtisari, E. T., Widiningrum, R., Branata, J., & Susanto, R. D. (2019). Google Translate in language learning: Indonesian EFL students' attitudes. *Journal of Asia TEFL, 16*(3), 978. https://doi.org/10.18823/asiatefl.2019.16.3.14.978

Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL, 21*(2), 241-258. https://doi.org/10.1017/S0958344009000172

O'Brien, S., Simard, M., & Goulet, M. J. (2018). Machine translation and self-post-editing for academic writing support: Quality explorations. *Translation quality assessment: From principles to practice,* 237-262. https://doi.org/10.1007/978-3-319-91241-7_11

O'Neill, E. M. (2016). Measuring the impact of online translation on FL writing scores. *IALLT Journal of Language Learning Technologies, 46*(2), 1-39. https://doi.org/10.17161/iallt.v46i2.8560

O'Neill, E. M. (2019). Online translator, dictionary, and search engine use among L2 students. *CALL-EJ: Computer-Assisted Language Learning–Electronic Journal, 20*(1), 154-177.

Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language learning, 47*(1), 101-143. https://doi.org/10.1111/0023-8333.31997003

Qiao, L. T., & Wang W. Y. (2020). An empirical study of the impact of dictionary use on learner's lexical performance in L2 writing: Chinese L2 learners as an example. *Foreign Language Research, 37*(05), 36-42+112.

Qin, X. Q., & Bi, J. (2012). Validity of L2 writing fluency measures: A study based on textual feature. *Foreign Language Teaching and Research (bimonthly), 44*(06), 899-911+961.

Read, J. A. (2000). *Assessing vocabulary*. Cambridge university press. https://doi.org/10.1017/CBO9780511732942

Selcuk, H., Jones, J., & Vonkova, H. (2019). The use of Google Translate as an ICT tool in web-based collaborative writing: Self-reported accounts of EFL learners. *Proceedings of IAC,* 74-80.

Stapleton, P., & Kin, B. L. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes, 56*, 18-34. https://doi.org/10.1016/j.esp.2019.07.001

Stevenson, M., Schoonen, R., & De Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing, 15*(3), 201-233. https://doi.org/10.1016/j.jslw.2006.06.002

Tsai, S. C. (2019). Using google translate in EFL drafts: a preliminary investigation. *Computer Assisted Language Learning, 32*(5-6), 510-526. https://doi.org/10.1080/09588221.2018.1527361

Tsai, S. C. (2022). Chinese students' perceptions of using Google Translate as a translingual CALL tool in EFL writing. *Computer assisted language learning, 35*(5-6), 1250-1272. https://doi.org/10.1080/09588221.2020.1799412

Wang, J., & Ke, X. (2022). Integrating machine translation into EFL writing instruction: Process, product and perception. *Journal of Language Teaching and Research, 13*(1), 125-137. https://doi.org/10.17507/jltr.1301.15

White, K. D., & Heidrich, E. (2013). Our policies, their text: German language students' strategies with and beliefs about web-based machine translation. *Die Unterrichtspraxis/Teaching German, 46*(2), 230-250. https://doi.org/10.1111/tger.10143

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). L2 development in writing: measures of fluency, accuracy, & complexity. *Studies in Second Language Acquisition, 23*(3), 423-425. https://doi.org/10.1017/S0272263101263050

Zhao, W. B. (2007). *English writing revision model*. Henan University Press.

Appendix A

List of Abbreviations

English Abbreviation	Full Title
MT	Machine Translation
OT	Online Translation
GT	Google Translate
CALL	Computer Assisted Language Learning
MLT	mean length of T-unit
L2SCA	Second Language Syntactic Complexity Analyzer
DC/T	dependent clause per T-unit
CN/T	complex nominal per T-unit
LD	lexical density
LS	lexical sophistication
LV	lexical variation
Uber	Uber Index
NDW	number of different words
LWV	Lexical word variation
LCA	Lexical Complexity Analyzer
CAF	complexity, accuracy and fluency
K1	the most frequent 1000 word families
K2	the second most frequent 1000 word families
AWL	academic words list
Off-list	off-list words