Computer Assisted Language Learning Electronic Journal (CALL-EJ), 23(1), 150-167, 2022

The Effects of Language Proficiency and Online Translator Training on Second Language Writing Complexity, Accuracy, Fluency, and Lexical Complexity

Syed Muhammad Mujtaba* (mujtaba2313@gmail.com) Corresponding Author
IQRA University (North Campus), Karachi, Pakistan

Rakesh Parkash (rakesh.parkash@yahoo.com)
IQRA University (North Campus), Karachi, Pakistan

Barry Lee Reynolds (barryreynolds@um.edu.mo)
Faculty of Education & Centre for Cognitive and Brain Sciences, University of Macau, Macau SAR, China.

Abstract

Language classes have adopted technology to aid L2 learning for the last 30 years. One important technological adoption by L2 teachers is Google Translate (GT). While L2 scholars have demonstrated a positive effect of GT on L2 learners' writing, there are a few unexplored areas. For instance, previous studies have used holistic measures to assess writing, leaving an area open as to how the use of GT affects the complexity, accuracy, fluency, and lexical complexity (CAFL). Similarly, research scholars believe that when teachers provide training and apprise the learners of the limitations of using online translators (OTs), learners can use them more effectively. However, not much evidence is available to support this. The current study was therefore conducted to fill these gaps in the literature. Two hundred twenty-five first-semester university students were made part of the study. These students were first put into high proficiency (HP; n= 108) and low proficiency (LP; n=117) and later were divided randomly across the three groups: GT+ training, GT only, and control group. The writing was evaluated in terms of CAFL; the results demonstrated that HP learners exhibited higher syntactic and lexical complexity and fluency, while LP learners fared better on accuracy.

*Keywords***:** Google translate; L2 writing; writing performance; proficiency; online translators

Introduction

Language classes have drastically changed in the last three decades, with technological advancements replacing traditional teaching methods (Chung & Ahn, 2021). Education stakeholders have also realized the importance of such advancements and have started adopting them in language classes to supplement learning (Cancino & Panes, 2021). L2 learners have reported using online translators (OTs) in writing classes (Ducar & Schocket, 2018) because OTs increased metalinguistic awareness and allow learners

to rectify orthographic, lexical, and grammatical issues (Abraham, 2009; Correa, 2014). Among the different OTs, GT is regarded as the most common tool that assists learners in writing (Clifford et al., 2013; Garcia & Pena, 2011; Jolley & Maimone, 2015). Google introduced its OT tool GT in 2006; this OT contained a translation model centered on phrase-based algorithms that examined word pairs based on the "frequency of use across massive amounts of digitalized data" (King, 2019, p.2). After a decade, a new translation system was developed, "Google Neural Machine Translation (GNMT) – a system which brings improvements in the quality of translation (Wu et al., 2016). Despite the upgradation of its OT by Google, research scholars have argued that a few challenges arise from using OTs. For instance, OTs seem to provide wrong translation due to cultural references, pragmatic expressions, proverbs, idioms, and polysemic words (Abraham, 2009; Chung & Ahn, 2021; Correa, 2014; Ducar & Schocket, 2018). Besides, even after using OTs, learners have to involve in post-editing to make their work of sufficient quality because OTs seem to rely on text type, subject area, and language pair (Godwin-Jones, 2015).

Despite these challenges, past studies have demonstrated a positive effect of OTs in helping learners improve writing (Abraham,2009; Garcia & Pena, 2011; Lee, 2020; Stapleton & Kin, 2019). However, these studies have used holistic measures to examine learners' writing quality after using OTs, thereby leaving the area open as to how OTs' use systematically affects the CAFL of L2 learners. This is important to investigate as writing is multidimensional (Cumming 1990), hence learners' writing performance can comprehensively be captured by CAFL (Barrot & Gabinete, 2019; Lu & Ai, 2015; Skehan, 2009). In addition to the multidimensional nature of writing that requires learners to plan and revise their writing (Sokolik, 2003), there are an array of variables that seem to influence the CAFL, such as L1 interference, proficiency level, and prior experience (Bakry & Alsamadani, 2015; Chung & Ahn, 2021; Okasha & Hamdi, 2014). For instance, L2 writers are influenced by their mother tongue (L1), making them unable to write like native speakers (Tsai, 2019). For these learners, writing is difficult (Okasha & Hamdi, 2014; Salma, 2015) because L1 and L2 are naturally intertwined in learners' minds (Cook, 2010; Druce, 2012). However, this should be taken as an opportunity for harnessing rather than rejecting (Leonardi, 2010). Therefore, research scholars advocate using the translation from L1 to L2 in writing classes to enable learners to look for a range of lexical items and phrases, thus making them produce syntactically complex and cohesive texts (Cohen & Brooks- Carson, 2001; Tsai, 2019). Moreover, linguistic factors, such as L2 proficiency– knowledge of grammar and vocabulary– have received their due attention because of their direct impact on the L2 writing process (Kaplan, 1966; Kim & Pae, 2021). Research scholars argue that L2 writing is a complex process that entails interactions among different variables, including L2 proficiency level (Zabihi, 2018). While past studies have demonstrated that language proficiency accounts for a significant difference in L2 writing (Engber, 1995; Grabe & Kaplan, 1996; Laufer & Nation, 1995; Pae, 2018), how learners with different language proficiency (high vs. low) benefit from OT in terms of CAFL has not been explored systematically. Therefore, the current study adds to the body of literature in two ways: first, examining how the use of GT affects the CAFL of L2 writers; second, how learners with differential proficiency levels benefit from GT.

L2 writing and online translators

OTs are tools that are free to use for translating from one language to another. Currently, the free and readily available OTs are Babylon, Bing, I'm translator, and Google Translate. The adoption of OTs in language classes has substituted textbooks and dictionaries (O'Neill, 2016). Despite this, translation of one language to another is not widely accepted because of its similarity with the grammar-translation method (GTM) (Cancino & Panes, 2021). For instance, Cook (2009) argues that translation in language classes focusing on form is still practiced globally. This method can be used in large classes and give learners a feeling of attainment and confidence, particularly low proficiency learners. There has been a debate in the literature concerning the effects of OTs on language learning, with some showing skepticism and others showing cautious optimism (Jolley & Maimone, 2015). Albeit all this, Lee (2020) believes that advancements in technology have given OTs a special place in language classes as OTs help learners improve their grammar and lexical resources accuracy. Lee, advocating the positive side of OTs, states that OTs can prove beneficial for learning a language provided that instructors are cognizant of their limitations and that learners are provided sufficient instruction. Niño (2009) states that learners embed OTs to complete written assignments as they provide instant translation with a variety of languages. Besides, a number of undergraduate learners perceive OTs as an important part of learning a second language. Echoing the same, Briggs (2018) reported that 48.8 % out of 80 Korean undergraduate learners perceived OTs as a valuable tool for learning a language. The study also concluded that undergraduate learners commonly use OTs, so their use should not be discouraged.

In addition to the studies that have advocated in favor of OTs in language classes, there are a few scholars who recommended against using OTs in language classes, citing reasons such as challenges with idioms, ambiguous lexical items, and undue reliance on OTs (Ducar & Schocket, 2018; Luton, 2003; Somers, 2011). One of the most common challenges is the quality of the linguistic form given by OTs, which may influence the quality of the writing (Chung & Ahn, 2021). This problem seems to be more challenging for low proficiency learners who cannot decide whether the translation done by OT is right or wrong (Cancino & Panes, 2021). Furthermore, Somers et al. (2006) stated that it is not ideal to compare a text that has been produced using OT with the one that has been produced without using OT because, in the latter case, a learner puts efforts to produce his writing without any means to access OT. The authors also argued that a text that is produced with the help of OT does not have any pedagogical value, and this wastes time and efforts of language teachers—rebutting these issues, Stapleton and Kim (2019) claim that these advancements in technology are now progressively providing their users with accurate translation, hereby enabling the users to improve their written output.

While a great strand of research has demonstrated a positive effect of OT on L2 learners' writing performance (see Kol et al., 2018; Lee, 2020; O'Neill, 2016; Tsai, 2019;), these studies have not used the CAFL to measure the writing performance of learners. For instance, O'Neill (2016) examined the effects of an OT on L2 learners' writing output. The participants were divided into three groups: group with access to an OT, group with no access to an OT, and group with access to an OT + training. The results demonstrated that OT + training group produced better output than the control group; the results concluded that learners who adopted OT exhibited greater grammar, content, and comprehensibility. The study stated that OTs are a valuable tool that aids learners' in writing, provided if the learners are given training before using them. Similarly, Tsai

(2019) conducted a study to examine the effects of GT on L2 learners' writing output. To this end, 124 adult Chinese EFL learners' writings were assessed using two automatic writing evaluation software: 1Checker (for grammar and spelling) and VocabProfiler (for lexical features). The study had learners write in their L1 first and later asked to write the text in English. In doing so, the learners were divided into two groups: self-written (those who had no access to GT) and GT. The results found that GT learners outperformed their counterparts in that they used greater words, higher academic words, and fewer grammar and spelling errors. Recently, Lee (2020) conducted a study to ascertain the influence of GT on L2 learners' writing quality. The study had 34 EFL intermediate and high intermediate proficiency levels, based on their TOEFL iBT score. They wrote a text in L1 and then wrote the same text in L2 using GT. The learners then edited and revised the final version of the text by referring to GT's version of their L1 text. The results unveiled that GT enabled learners to have fewer grammar and lexical errors than their initial L2 text; however, no statistical difference was observed between final and initial versions.

The above studies, albeit providing evidence of the positive effects of OTs on writing performance, have left much to be explored. Since writing is a multidimensional skill, many researchers have used the CAFL model to evaluate learners' writing performance (Abdi Tabari, 2020; Barrot & Gabinete, 2019; Cho, 2019; Lu & Ai, 2015; Skehan, 2009), citing the reason that the CAF model can comprehensively capture the writing performance of the learners. To the best of the researchers' knowledge, only two studies have examined the effects of GT on L2 learners' writing performance measured in terms of CAFL (Cancino & Panes, 2021; Chung & Ahn, 2021). Cancino and Panes (2021) examined the effects of OT on L2 learners' writing performance measured in terms of T-unit length, syntactic complexity, and accuracy. To this end, sixty-one 11 EFL grade learners were made part of the study who were assigned to three groups: GT, GT + training, and no GT. The learners' proficiency level, based on the Quick Oxford Placement Test, was A1/A2. The results unveiled that learners who used GT exhibited higher complexity, wrote more words, and produced more accurate texts than the learners who did not have access to GT. There was no significant difference between GT and GT + training groups, indicating that both groups aided learners. More recently, Chung and Ahn (2021) conducted a study assessing the effects of GT on L2 learners' CAFL. The study also investigated the effects of language proficiency and text genre on GT use in L2 writing. A total of 91 Korean adult EFL learners were recruited for the study; these learners were divided into high and low proficiency based on the English proficiency test scores. Without using GT, these learners were asked to write a narrative essay (my first day at college). A week later, they were given another topic (my last day at college) for which they had access to GT. A similar procedure with 31 learners was followed to assess the text genre (argumentative writing). The results concluded that L2 learners exhibited higher accuracy, but no concrete improvements were reported in terms of syntactic and lexical complexity. It was also reported that GT had differential effects subject to the proficiency level and text type (narrative vs. argumentative).

The aforementioned studies have provided empirical evidence regarding the effects of GT on L2 writing performance. However, most of the studies have not assessed the writing concerning learners' CAFL, and they also suffered from methodological limitations. For instance, Lee (2020) and Tsai (2019) had their students write a similar topic twice. Learners in both studies first wrote a text in L1 then were asked to write the same text in L2. Later on, the L1 text was translated using GT, and a comparison was

made. This consequently might have confounded the effects of GT as the changes in learners' writing may be because of the practice effect. The current study aims to answer the research question:

RQ. Does the use of GT promote CAFL for high and low proficiency learners across the three groups (GT+ training, GT only, and control)? If yes, does proficiency moderate the effect of the use of GT on CAFL?

Methods

Research Design

The current study adopted a quasi-experimental research design to examine the effectiveness of GT on L2 learners' writing performance measured in terms of CAFL. Such a design is suitable for intervention studies where researchers attempt to investigate how a particular method influences the performance of the learners (Phakiti, 2015). The current study has two treatment groups (GT+ training and GT only) and one control group. The GT+ training group was given training on the use of GT, while the GT-only group was not provided training. The control group did not have access to GT; they had to complete their writing without taking the help of GT.

Participants

240 first-semester university students from six- intact classes from a private sector university in Pakistan were initially invited to participate in the study. These participants had previously taken the Introduction to Computer Science course in their high school before admission to the university. The course required the use of computers and typing to complete their written assignments. It is reasonable to believe that they had sufficient knowledge of typing through completing this previous course. All participants also possessed personal computers (confirmed by the researchers), which increases the likelihood that they had experience using keyword and word processing programs. During the study, no participants showed any signs of inability to use the computer or keyword (e.g., two-finger typing, labored use of the software, inability to use a mouse). After considering the participants who were unable to appear on the day of the experiment, the sample size was adjusted to 225 students. The data was collected in the middle of the semester, ensuring that the students had learned about a paragraph and essay writing. The participants' English language proficiency was assessed using Cambridge Placement Test (CPT) and the write and improve (writeandimprove.com) developed by Cambridge English. The CPT had 25 items that examined knowledge of vocabulary and grammar. The writing task was taken from the certified website (writeandimprove.com) that asked the participants to write on the topic, *"Would you prefer an interesting job with low pay or a boring job with a high salary?"* The writing was evaluated using the same website (writeandimprove.com), which assessed the writing based on the Common European Framework of Reference (CEFR) scale ranging from A1 lowest (1 point) to C2 highest (7 points). The learners' final proficiency score was calculated by combining grammar and vocabulary, and writing task scores (e.g., A student obtained 20 points in CPT out of

25 and 4 points on a 7-point scale in CEFR then the student's proficiency scores is calculated as (20/25 + 4/7)/2) 68%. The participants were placed into high proficiency (HP) and low proficiency (LP) groups according to their final proficiency score. The median was used to classify the students into LP and HP because proficiency data is slightly skewed right. Therefore, it was deemed better to use the median to ascertain the central point of the data. Using median to classify learners into HP and LP has also been used in other studies (Chung & Ahn, 2021; Sarrett et al., 2021). 108 with a score greater than or equal to the median (42) were placed in the HP group, while 117 students with a score less than the median (42) were placed in the LP group. These students were then randomly divided into two treatment groups and a control group across the HP and LP (see Table 1).

Table 1

Number of Participants

High Proficiency (HP) n=108			Low Proficiency (LP) n=117		
GT + Training	GT Only	Control	GT + Training	GT Only	Control
n= 36	n=36	n=36	n=39	n=39	n=39

Writing Task

The argumentative genre of writing was selected primarily because of two reasons: first, it is a commonly tested genre at national and international levels (Huang & Zhang, 2020); therefore, the findings of the study would apply to a wider population of language learners, ensuring the external validity of the study (Mackey & Gas, 2005). Secondly, most studies on GT seem to have used narrative or descriptive genres of writing (Cancino & Panes, 2021; Lee, 2020; Tsai, 2019). Research scholars (Lu, 2011; Way et al., 2000) have argued that the writing genre affects L2 compositions. For instance, an argumentative genre of writing tends to have higher syntactic complexity and lexical sophistication than a narrative genre. In contrast, the narrative genre tends to have a higher score on lexical complexity (Lu, 2011; Yoon & Polio, 2017). Given these differences in the genre of writing, we decided to explore the effects of GT on argumentative writing to push forward this line of research.

Writing prompt

The participants of all three groups were asked to write an argumentative essay on whether students should come to the university in formal dress. Since the university has a policy on wearing a formal dress on Mondays and Wednesdays, we thought it would be an exciting and familiar topic for the students.

Some people believe that university students should come in formal dress, while others believe that it is not important. What is your opinion?
You should write at least 250 words in no more than 45 minutes.

The procedure of the study

The data was collected during a class in a single day. First, the GT+ training group was invited to a lab to attend a training session lasting for 60 minutes on the use of GT. The second author, who gave the training session, briefly explained the advantages and drawbacks of using GT. For training, the second author used a Microsoft PowerPoint presentation, explaining the features of GT. The learners were allowed to ask questions if needed. The second author briefed the learners about the number of strategies related to a web-based GT version as it offers more features than a mobile-based one. Firstly, the learners were informed that web-based GT offers translation history and definitions of a single word. They were also shown translation of short texts, phrases, and words. Secondly, they were shown how GT could perform reverse translation (translating from L1 to L2 and vice-versa); learners were recommended doing this while using GT as it would allow them to see whether the translated text was accurate or not. Finally, the learners were cautioned to edit and check the translation before using them in writing. In the end, the author allowed the learners to practice translating short pieces of text and advised them to use the strategies explained to them. After the training session, the GT-only group and control group were invited to the computer lab. Three computer labs were utilized for the study; the GT+ training and GT-only groups were placed in labs with access to the Internet, while the control group was placed in the lab without access to the Internet. The first author and two assistants monitored the three labs during the session. Before starting the task, the participants were told to save their writing task after completing it. All the groups were asked to write an argumentative essay of no less than 250 words in 45 minutes. The GT+ training and GT-only group were allowed to use GT during their writing, while the control group had to complete their writing without accessing GT.

Measuring writing performance

The writing performance of the learners was assessed by employing multiple measures of complexity, accuracy, fluency, and lexical complexity because such measures allow to comprehensively and precisely capture the writing performance instead of holistic measures of writing (Abdi Tabari, 2020; Johnson, 2017; Yoon & Polio, 2017). The syntactic complexity, lexical complexity, and fluency were computed via online software: syntactic complexity analyzer (Lu, 2010) and Coh- Metrix (McCarthy & Jarvis, 2010). The complexity in the study was measured by using two metrics (Lu, 2010): clauses per T –unit (C/T) and Mean length of clause (MLC), the number of words to clauses in learner's writing. The accuracy in the study was computed using two metrics: 1) Error-free clauses, the number of clauses without any errors (EFC/C). Following Rostamian et al. (2017) and Cho (2019), lexical, morphological, word order, and syntactic errors were included in the analysis, while spelling and punctuation errors were not; 2) Correct verb forms: the % of all correctly written verbs concerning the subject-verb agreement, tense, aspect and modal verbs (Abdi Tabari, 2020; Rostamian et al., 2017;). Lastly, fluency was measured using a number of words produced within a minute (Wolfe-Cho, 2019; Quintero et al., 1998;). The lexical complexity was computed by measuring lexical diversity. To measure lexical diversity, we employed a measure of textual and lexical diversity (MTLD)– the number of different words learners have used in a text (McCarthy & Jarvis, 2010).

To ensure the reliability of the two metrics of accuracy (EFC and CVF) for both HP and LP learners, the essays were double coded by the first and the second author. For HP learners, the inter-rater reliability for both EFC and CVF was found to be good. EFC: GT + training (ICC= .881, 95% CI from .766 to .939); GT only (ICC= .902, 95% CI from .811 to .951); control (ICC= .898, 95% CI from.800 to .948); similarly, for CVF: GT + Training (ICC= .900, 95 % CI from .803 to .949); GT only (ICC=.860, 95% CI from .725 to .929); control (ICC= .860, 95% CI from .725 to .929). Similarly, the inter-rater reliability of EFC and CVF was found to be good for the LP learners. EFC: GT+ training (ICC=.920, 95% CI from .844 to .959); GT only (ICC= .895, 95% CI from.794 to .947); control (ICC= .775, 95% from .558 to .885); similarly, for CVF: GT + training (.829, 95% CI from .673 to .910); GT only (.975, 95% CI from .953 to .987).

Data Analysis

The data were analyzed using both descriptive and inferential statistics. The descriptive statistics were used to get the mean and standard deviation values for the CAFL across the three groups for both HP and LP. Subsequently, inferential statistics the analysis of variance (ANOVA test) was run to ascertain the difference across the three groups for HP and LP learners. The Multiple Analysis of Variance test (MANOVA) was also applied to examine the interaction effect of proficiency and GT on CAFL.

Results

RQ1: Does the use of GT promote CAFL for high and low proficiency learners across the three groups? If yes, does proficiency moderate the effect of the use of GT on CAFL?

To answer the first research question, both descriptive and inferential statistics were run to examine the effect of GT on HP and LP learners on CAFL across the three groups: GT+ training, GT only, and control group. For HP learners, the GT+ training group exhibited the highest mean values, followed by the GT-only group and the control group for all the measures of CAFL (see Table 2 for M and SD values). An ANOVA test was run to examine whether the difference of HP learners across the three groups is statistically significant. The results indicated a statistically significant difference of HP learners across the three groups on all the constructs of CAFL: [CT: $F(2,222) = 50.18$, $p = .000$; MLC: $F(2, 222) = 12.8$, $p = .000$; EFC: $F(2,222) = 168.4$, $p = .000$; CVF: $F(2,222) = 70.7$, $p = .000$; COMP: $F(2,222) = 44.4$, $p = .000$; MLTD: $F(2,222) = 152.5$, $p = .000$]. Similarly, for LP learners, the GT+ training group exhibited the highest mean values, followed by the GT-only group and control group for all the measures of CAFL (see table 2 for M and SD values). An ANOVA test was run to examine whether the difference of LP across the three groups is statistically significant. The results confirmed a statistically significant difference in LP across the three groups on all the constructs of CAFL: [CT: $F(2,114) = 38.7$, $p = .000$; MLC: $F(2,114) = 192$, $p = .000$; EFC: $F(2,114) = 200$, $p = .000$; CVF: $F(2,114) = 59$, $p = .000$; COMP: $F(2,114) = 29$, $p = .000$; MLTD: $F(2,114) = 5658$, $p = .000$].

Table 2

Mean and standard deviation of high proficiency and low proficiency groups

		HP Mean (SD) n=108			LP Mean (SD), n=117		
DV	Measure	Control	GT only	GT+T	Control	GT only	GT+T
Complexity	CT	1.82 (.13)	1.91 (.23)	2.35 (.27)	1.14 (.40)	1.54 (.30)	1.93 (.46)
	MLC	7.23 (.59)	7.72 (1.44)	8.94 (1.17)	5.21 (2.18)	6.63 (1.76)	8.45 (2.12)
Accuracy	EFC	.51 (.14)	.77 (.20)	1.47 (.29)	.93 (.07)	1.06 (.30)	1.66 (.39)
	CVF	.67 (.14)	.76 (.23)	1.1 (.37)	.43 (.18)	.84 (.27)	1.02 (.30)
Fluency	COMP	12.25 (3.30)	13.92 (.72)	16.61 (2.97)	8.26 (4.0)	11.11 (2.81)	14.5 (3.85)
Lex. Complexity	MLTD	69.88 (.25)	75.25 (.16)	80.33 (.22)	52.57 (.32)	67.14 (1.57)	74.91 (.27)

Having established a significant difference across the three groups for both HP and LP learners, an LSD post-hoc test was run to isolate the group differences (see table 3). For HP learners, there exists a statistically significant difference between control and GT + training and GT + training and GT only for all the constructs of CAFL. However, for the GT only and control group, a significant difference was found for all the constructs, excluding one construct of accuracy– EFC and one construct of complexity – CT (see Table 3 for p- values).

Table 3

Least Significant Difference (ANOVA)

Prof.	DV	Measures	Control vs. GT only		Control vs. GT+T		GT only vs. GT+T	
			SE	d	SE	d	SE	d
HIGH	Complexity	CT	0.05	.48	0.05	2.5***	0.05	1.7***
		MLC	0.21	.44**	0.21	1.8***	0.21	0.9***
	Accuracy	EFC	0.07	1.5	0.06	4.2***	0.07	2.8***
		CVF	0.05	0.47**	0.05	1.5***	0.05	1.1***
	Fluency	COMP	0.61	0.7**	0.61	1.4***	0.61	1.2***
	Lex. Complexity	MLTD	0.3	2.5***	0.3	4.4***	0.3	2.6***
LOW	Complexity	CT	0.09	1.1***	0.09	1.8***	0.08	1.0***
		MLC	0.46	0.8***	0.46	2.1***	0.46	0.9***
	Accuracy	EFC	0.05	2.3***	0.05	4.0***	0.05	1.7***
		CVF	0.06	1.7***	0.06	2.4***	0.06	0.6***

Fluency	COMP	0.81	0.8***	0.81	1.6***	0.81	1.0***
Lex. Complexity	MLTD	0.21	4.69***	0.21	7.5***	0.21	3.5***

Note. *, ** & *** means p value < .05, .01 & .001 respectively.

Similarly, for LP, a statistically significant difference was found for all the constructs of CAFL across the three groups (see table 2 for p values). Finally, a MANOVA test was run to examine the interaction effect of proficiency and GT on CAFL. The results of the MANOVA test (Wilk's Lambda=.077, p=.000, and partial eta square= .723) confirmed that proficiency moderates the effect of GT. According to table 3, there exists a statistically significant interaction between GT and proficiency for all constructs except COMP (see Table 4).

Table 4

Interaction Effect

IV	DV	SS III	df	MS	F	Sig.	η^2
Use of GT	CT	16.548	2	8.274	79.617	.000	.421
	MLC	227.569	2	113.785	44.774	.000	.290
	EFC	29.474	2	14.737	229.916	.000	.677
	CVF	9.531	2	4.766	74.889	.000	.406
	COMP	1060.057	2	530.028	53.242	.000	.327
	MLTD	10218.959	2	5109.479	4068.008	.000	.974
Proficiency	CT	13.353	1	13.353	128.494	.000	.370
	MLC	116.930	1	116.930	46.011	.000	.174
	EFC	5.125	1	5.125	79.958	.000	.267
	CVF	.407	1	.407	6.403	.012	.028
	COMP	496.699	1	496.699	49.894	.000	.186
	MLTD	5937.120	1	5937.120	4726.949	.000	.956
Use of GT*Proficiency	CT	1.034	2	.517	4.973	.008	.043
	MLC	1018.151	2	509.076	200.318	.000	.647
	EFC	.469	2	.235	3.661	.027	.032
	CVF	1.035	2	.518	8.134	.000	.069
	COMP	33.845	2	16.922	1.700	.185	.015
	MLTD	1456.770	2	728.385	579.917	.000	.841

Discussion

The current study was conducted to fulfill two aims: first, to investigate the effect of GT on both HP and LP learners across the three groups on CAFL; second, to investigate whether the learners' proficiency level moderates the effects of GT on CAFL. The study concluded that the use of GT promoted CAFL, particularly the GT + training group, benefited the most from GT. The study also concluded that the learners' proficiency level

moderates the effect of GT. This result is in line with previous studies, which reported that giving training to the learners on the use of OTs (Cancino & Panes, 2021; O'Neill, 2016) is more effective than not. The current study also highlighted that learners with different proficiency levels (high vs. low) benefitted differently from the GT, suggesting an interaction effect of proficiency and GT use. This is in line with the findings of Chung and Ahn (2021). This result is not surprising, given that the plethora of research has demonstrated that learners with HP formulated more complex texts with advanced lexical complexity (Crossley & McNamara, 2012; Jarvis et al., 2003; Shin & Kim, 2014). Since HP learners may seem to have a greater general understanding of L2 (Guo, 2015), this may have given them leverage and allowed them to use GT more effectively than their counterparts. Another possible explanation of the interaction between proficiency and the use of GT on CAFL could be explained from the learners' working memory capacity (WMC). Writing in L2 is a complex process that requires the interaction of different factors, such as L2 proficiency and WMC (Zabihi, 2018). WMC is a person's ability to encode, process, and retrieve information (Choi, 2017). Learners with different proficiency levels may have different WMC (Guo, 2015); HP learners are believed to have a higher WMC span, and this gives them an advantage as they can process, encode and retrieve information more quickly as compared to LP learners (DeKeyser, 2007; Guo, 2015). The HP learners of the current study may have higher WMC, and this may have helped them free some of their WMC and allowed them to focus more on the features of GT.

The HP learners of all three groups exhibited higher complexity, greater lexical complexity, and fluency than the LP learners of all three groups. In contrast, the LP learners across the three groups exhibited higher accuracy than the HP learners across the three groups. A number of possible explanations could be offered to justify such a result. First, the LP learners may not have reached a maximum level of threshold on their L2 proficiency; therefore, the use of GT may have enabled them to refine their expressions more effectively than the HP learners (Chon et al., 2021). The LP learners also seemed to be more careful in writing simple and acceptable sentences, which made them exhibit higher accuracy than the HP learners. For instance, *students should not come in formal dress. This is because some cannot afford it.* This sentence is accurate but lacks complexity. Since LP learners were more careful in writing simple sentences to ensure accuracy, this may have declined their syntactic and lexical complexity. This is in line with Chung and Ahn (2021), who reported that LP learners used OTs to write grammatically correct sentences. This result echoed the notion that OTs help LP learners with grammatical problems during the writing process (Kol et al., 2018; Lee, 2020; Tsai, 2019). L2 learners use L1 while writing in L2 (Cook, 2010; Laufer & Girsai, 2008; Wang & Wen, 2002; Weijen et al., 2009; Woodall, 2002). Using L1 during writing in L2 enables learners to have access to a broader repertoire of vocabulary and phrases (Cohen & Brooks-Carson, 2001; Tsai, 2019). The participants in the current study used GT to translate their ideas from L1 to L2, thereby producing more complex and lexically rich texts. In contrast, HP learners exhibited higher fluency, syntactic and lexical complexity than LP learners across the three groups. However, a decline in accuracy was observed for HP learners. For instance, *although students should come to the university in formal dress, but it should not be made compulsory on them.* The sentence is nevertheless complex; however, it lacks accuracy as the learner has used *but* after the subordinate clause. A number of possible explanations could be given to justify such a finding. First,

as Chung and Ahn (2021) state that L2 learners may seem more confident in their ability to write accurate sentences, which may have led to not focusing on accuracy, consequently allowing them to make less accurate sentences (Cho, 2019). Another possible explanation could be given through Skehan's (1998) trade-off hypothesis, which states that improvement in one aspect of writing comes at the expense of another aspect (Abdi Tabari, 2020; Cho, 2019; Kormos, 2012). The HP learners' decline in accuracy can be explained by the additional time spent in improving the complexity and fluency of their writing.

Similarly, the GT + training of HP learners produced more fluent and lexically rich text within the group and across the group based on the mean value. This result explains that giving training to learners before using OTs would yield more improvement. This result echoes Chung and Ahn's (2021) and Cancino and Pane's (2021) findings. Garcia and Pena (2011) state that OTs are an external source of enriching vocabulary in L2 writing; therefore, the HP learners may seem to have access to a broader repertoire of vocabulary while translating from L1 to L2 (Cohen & Brooks-Carson, 2001; Tsai, 2019). There is a possibility that the HP learners tend to have sufficient knowledge of the target lexical structure that led them to express their ideas more vividly using rich lexical resources (Ferris & Hedgcock, 2005; Hyland, 2003; Kim & Pae, 2021). Skehan (2009) also states that L2 learners are equipped with sufficient L2 lexical knowledge to ascertain acceptable lemma for the concept.

Another reason for higher lexical complexity in the writing of HP learners is their ability and confidence in using complex sentences, thereby enabling them to focus more on the lexical side in their writing (Chung & 2021). Regarding fluency, the HP learners, particularly the GT+ training, exhibited greater fluency than the LP learners across the three groups. One possible reason for such a finding is that the HP learners in the GT+ training group might have aided their working memory (WM) due to the use of GT more effectively than the LP learners; the HP learners seem to have influenced the mental operations in WM during writing. Another possibility is that the HP learners may seem to have more ideas to support their argument, and with the provision of GT, they were able to convert those ideas from their L1 to L2 more effectively than the LP learners.

Conclusion

The current study was conducted to examine how GT influences the writing performance of L2 learners measured in terms of CAFL and whether the learners' proficiency level moderates the effect of GT. The study results indicated that GT helped learners improve their writing performance, with the HP learners faring better in terms of syntactic complexity, lexical complexity, and fluency. In contrast, the LP learners fared better on accuracy. The study highlighted that the GT + training group benefitted the most from GT, suggesting that 60 minutes of training given to the GT+ training group proved to be sufficient in helping this group understand the features of GT. The study further illuminated that proficiency moderated the effect of GT on CAFL as HP and LP learners benefited differently from GT. The HP learners exhibited higher syntactic and lexical complexity, with the GT+ training group scoring the highest, followed by GT only

and the control group. In contrast, the LP learners fared better on the accuracy, with the GT+ training group scoring the highest, followed by GT only and the control group. The study offers a couple of pedagogical implications for language teachers. First, language teachers can use GT to teach learners sentence structures. For instance, language teachers could design language chunk activities to raise awareness of the difference between L1 and L2 (Reynolds, 2015). Secondly, GT can function as peer feedback or teacher's correction (Chon et al., 2021). Since peer feedback is often considered unsatisfactory by both teachers and students (Hyland & Hyland, 2019; Paulus, 1999; Rollinson, 2005), GT can provide learners with synchronous feedback and help learners improve writing accuracy. Therefore, language teachers can use GT to provide learners with corrections, particularly in large classes where it may seem not plausible for teachers to provide individual feedback to every learner. Third, since the study has highlighted the positive side of GT, educational institutions should invest and adopt OTs to facilitate the teaching and learning process. These institutions should also train their learners on the use of GT as the study has concluded that providing training to the learners is more effective than not. In this regard, Tsai (2019) claims that with the advancement made in technology, it is important that language teachers consider adopting OTs in classes to better cater to the learners' needs. Further, the use of technology seems to expose learners to authentic language use, which will help improve their language skills. In this regard,

Golonka et al. (2014) and Pérez (2018) stress that language teachers must change their style of teaching and inculcate technology in their lesson plans to maximize the teaching and learning process. Due to time constraints, it is particularly important in EFL settings where learners do not have access to L2 beyond the classroom and do not interact much with their teachers.

Despite the pedagogical implications, the current study is not without limitations. First, the current study participants were recruited from a single institution in Pakistan. Future replication studies recruiting participants from different institutions are needed to draw more definite conclusions. Secondly, the current study did not employ qualitative data collection methods, such as interviews or open-ended questionnaires. Without such data, it was impossible to examine how the participants engaged with GT and what they thought about its usage in the writing classes. Therefore, future studies should consider this and use a mixed-method approach to supplement the quantitative data.

References

Abdi Tabari, M. (2020). Differential effects of strategic planning and task structure on L2 writing outcomes. *Reading & Writing Quarterly*, *36*(4), 320-338. https://doi.org/10.1080/10573569.2019.1637310

Abraham, L.B. (2009). Web-based translation for promoting language awareness: Evidence from Spanish. In L. B. Abraham & L. Williams (Eds.), *Electronic discourse in language learning and language teaching* (pp.65–84). Amsterdam: John Benjamins.

Bakry, M. S., & Alsamadani, H. A. (2015). Improving the persuasive essay writing of students of Arabic as a foreign language (AFL): Effects of self-regulated strategy

development. *Procedia-Social and Behavioral Sciences*, 182, 89-97. https://doi.org/10.1016/j.sbspro.2015.04.742

Barrot, J., & Gabinete, M. K. (2019). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. IRAL, International Review of Applied Linguistics in Language Teaching. https://doi.org/10.1515/iral-2017-0012. Advance online publication

Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal, 14*(1), 3-24. https://doi.org/10.29140/jaltcall.v14n1.221

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

Cho, M. (2019). The effects of prompts on L2 writing performance and engagement. *Foreign Language Annals*, *52*(3), 576-594. https://doi.org/10.1111/flan.12411

Choi, Y. H. (2017). *Roles of Receptive and Productive Vocabulary Knowledge in L2 Writing through the Mediation of L2 Reading Ability*. *English Teaching*, *72*(1). http://journal.kate.or.kr/wp-content/uploads/2017/04/01yeonhee-choi.pdf

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, E. S., & Ahn, S. (2021). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning*, 1-26. https://doi.org/10.1080/09588221.2020. 1871029

Clifford, J., Merschel, L., & Munn_e, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *Revista de Innovaci_on Educativa*, (10), 108-121. https://doi.org/10.7203/attic.10.2228

Cohen, A. D., & Brooks-Carson, A. (2001). Research on direct versus translated writing: Students' strategies and their results. *The Modern Language Journal, 85*(2), 169-188. https://doi.org/10.1111/0026-7902.00103

Cook, G. (2009). Use of translation in language teaching. In M. Baker, & G. Saldanha (Eds.), *Routledge encyclopedia of translation studies* (pp. 117-120). Routledge.

Cook, V. (2010). *Prolegomena to second language learning. In Conceptualising 'learning' in applied linguistics* (pp. 6-22). Palgrave Macmillan, London.

Correa, M. (2014). Leaving the "peer" out of peer-editing: Online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin American Journal of Content and Language Integrated Learning*, *7*(1), 1–20. https://doi.org/10.5294/3568

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, *7*(4), 482-511. https://doi.org/10.1177%2F0741088390007004003

Dekeyser, R. (2007). Skill acquisition theory. In B. VanPatten, & J. Williams (Eds.), Theories in second language acquisition: An introduction (pp.97–113). Lawrence Erlbaum.

Druce, P. M. (2012). Attitude to the use of L1 and translation in second language teaching and learning. *Journal of Second Language Teaching and Research, 2*(1), 60–86. https://pops.uclan.ac.uk/index.php/jsltr/article/view/82

Ducar, C., & Schocket, D.-H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google translate. *Foreign Language Annals, 51*(4), 779–795. doi:10.1111/flan.12366

Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of second language writing*, *4*(2), 139-155. https://doi.org/10.1016/1060-3743(95)90004-7

Ferris, D., & Hedgcock, J. (2005). *Teaching ESL composition: Purpose, process, and practice* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum

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

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

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer assisted language learning*, *27*(1), 70-105. https://doi.org/10.1080/09588221.2012.700315

Guo, Q. (2015). *The effectiveness of written CF for L2 development: A mixed-method study of written CF types, error categories and proficiency levels* (Doctoral dissertation, Auckland University of Technology). https://openrepository.aut.ac.nz/bitstream/handle/10292/9628/GuoQ.pdf?sequence=6&isAllowed=y

Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing.* London: Longman.

Hyland, K., & Hyland, F. (Eds.). (2019). *Feedback in second language writing: Contexts and issues.* Cambridge university press.

Huang, Y., & Jun Zhang, L. (2020). Does a process-genre approach help improve students' argumentative writing in English as a foreign language? Findings from an intervention study. *Reading & Writing Quarterly*, *36*(4), 339-364. https://doi.org/10.1080/10573569.2019.1649223

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of second language writing*, *12*(1), 17-29. https://doi.org/10.1016/S1060-3743(02)00124-8

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, lexical complexity, accuracy, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, 37,13–38. doi:10.1016/j.jslw. 2017.06.001

Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.), *Learn languages, explore cultures, transform lives* (pp. 181-200). Minneapolis: 2015 Central States Conference on the Teaching of Foreign Languages.

Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language learning*, *16*(1), 1-20. http://ksuweb.kennesaw.edu/~djohnson/6750/kaplan.pdf

Kim, K. J., & Pae, T. I. (2021). Examining the Simultaneous Effects of L1 Writing, L2 Reading, L2 Proficiency, and Affective Factors on Different Task Types of L2 Writing. *Discourse Processes*, 1-19. https://doi.org/10.1080/0163853X.2021.1872989

King, K. (2019). Can Google translate be taught to translate literature? A case for humanists to collaborate in the future of machine translation. *Translation Review, 105*(1), 76-92. https://doi.org/10.1080/07374836.2019.1673268

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

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of second language writing, 21*(4), 390-403. https://doi.org/10.1016/j.jslw.2012.09.003

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied linguistics, 29*(4), 694- 716. https://doi.org/10.1093/applin/amn018

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, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning, 33*(3), 157–175. doi:10.1080/09588221.2018.1553186

Leonardi, V. (2010). *The role of pedagogical translation in second language acquisition.* Bern: Peter Lang.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474–496. doi:10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36-62. https://onlinelibrary.wiley.com/pb-assets/journal-banners/15457249-1501384741387.jpg

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16-27. https://doi.org/10.1016/j.jslw.2015.06.003

Luton, L. (2003). If the computer did my homework, How come I didn't get an "A"? French Review, 76(4), 766-770. https://doi.org/10.2307/3133085

Mackey, A., & Gass, S. (2005). Common data collection measures. *Second language research: methodology and design.* Mahwah: Lawrence Erlbaum, 43-99.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42*(2), 381–392. doi:10.3758/BRM.42.2.381

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

Okasha, M. A., & Hamdi, S. A. (2014). Using strategic writing techniques for promoting EFL writing skills and attitudes. *Journal of Language Teaching and Research, 5*(3), 674. https://doi.org/10.4304/jltr.5.3.674-681

Pae, T. I. (2018). Effects of task type and L2 proficiency on the relationship between L1 and L2 in reading and writing. *Studies in Second Language Acquisition*, *40*(1), 63-90. https://doi.org/10.1017/S0272263116000462

Pérez, M. (2018). Technology for teaching English as a foreign language (EFL) writing. *The TESOL Encyclopedia of English Language Teaching*,1-12. https://doi.org/10.1002/9781118784235.eelt0439

Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of second language writing*, *8*(3), 265-289. https://doi.org/10.1016/S1060-3743(99)80117-9

Phakiti, A. (2015). *Experimental research methods in language learning*. Bloomsbury Publishing.

Reynolds, B. L. (2015). Helping Taiwanese graduate students help themselves: Applying corpora to industrial management English as a foreign language academic reading and writing. *Computers in the Schools*, *32*(3-4), 300-317. https://doi.org/10.1080/07380569.2015.1096643

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT journal*, *59*(1), 23-30. https://doi.org/10.1093/elt/cci003

Rostamian, M., Fazilatfar, A. M., & Jabbari, A. A. (2017). The effect of planning time on cognitive processes, monitoring behavior, and quality of L2 writing. *Language Teaching Research*, *22*(4), 418-438. https://doi.org/10.1177%2F1362168817699239

Salma, U. (2015). Problems and practical needs of writing skill in EFL context: An analysis of Iranian students of Aligarh Muslim University. *Journal of Humanities and Social Science*, *20*(11), 74-76. https://www.iosrjournals.org/iosr-jhss/papers/Vol20-issue11/Version-3/L0201137476.pdf

Sarrett, M. E., Shea, C., & McMurray, B. (2021). Within-and between-language competition in adult second language learners: implications for language proficiency. *Language, Cognition and Neuroscience*, 1-17. https://doi.org/10.1080/23273798.2021.1952283

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, *30*(4), 510-532. https://doi.org/10.1093/applin/amp047

Sokolik, M. (2003). Writing. In D. Nunan (Eds), *Practical English language teaching (PELT)*, (pp. 87-88). New York: McGraw Hill.

Somers, H. (2011). Machine translation: History, development, and limitations. In K. Malmkjær, & K. Windle (Eds.), *The Oxford handbook of translation studies* (pp. 427- 440). Oxford University Press.

Somers, H., Gaspari, F., & Niño, A. (2006). Detecting inappropriate use of free online machine translation by language students–A special case of plagiarism detection. In *Proceedings of the eleventh annual conference of the European association for machine translation* (pp. 41-48). https://aclanthology.org/2006.eamt-1.6.pdf

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

Tsai, S. (2019). Using Google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 1-17. https://doi.org/10.1080/09588221.2018.1527361

Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of second language writing*, *11*(3), 225-246. https://doi.org/10.1016/S1060-3743(02)00084-X

Way, D. P., Joiner, E. G., & Seaman, M. A. (2000). Writing in the secondary foreign language classroom: The effects of prompts and tasks on novice learners of French. *The Modern Language Journal, 84*(2), 171–184. doi:10.1111/0026-7902.00060

Weijen, D., Bergh, H., Rijlaarsdam, G., & Sanders, T. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing, 18*(4), 235–250. https://doi.org/10.1016/j.jslw.2009.06.003

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (No. 17). University of Hawaii Press.

Woodall, B. R. (2002). Language-switching: Using the first language while writing in a second language. *Journal of Second Language Writing*, *11*(1), 7-28. https://doi.org/10.1016/S1060-3743(01)00051-0

Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *Tesol Quarterly*, *51*(2), 275-301. https://doi.org/10.1002/tesq.296

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. https://arxiv.org/abs/1609.08144

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication*, *35*(1), 32-57. https://doi.org/10.1177%2F0741088317735836