

# **Chinese University Students' Engagement with Automated Feedback on English Writing**

by  
**Yun Ji**

Submitted to the  
**Department of English and German Studies**

in partial fulfillment of the requirements for the degree of

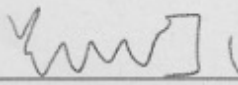
**Master's in Teaching and Learning English as a Foreign /  
Second Language**

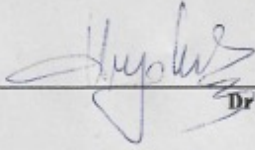
at the

**UNIVERSITAT ROVIRA I VIRGILI**

June 07, 2022



Signature of Author 

Certified by   
Dr / Ms / Mr Olga Hernández de Apollón  
Master Thesis Supervisor

Accepted by \_\_\_\_\_  
Dr Mar Gutiérrez-Colon  
Master's Coordinator



Master in Teaching and Learning English as a Foreign/Second Language  
2021-2022

**Master's Final Project**  
**ORIGINAL WORK FORM**

STUDENT

Last Name(s) Ji

First Name Yun

I hereby state that all the work presented as part of my Final Master's Project is original with no exception and that I have not, voluntarily or otherwise, misused or misreported any previously published information. I am aware that any failures to comply with these statements will automatically result in disqualification of my final paper and that I will not be able to obtain any credits for it.

Signature of Student



In Tarragona, 05,31, 2022.

### **Abstract**

Recent studies have produced contrasting results on how language proficiency affects students' engagement with automated feedback in second language writing. Some studies concluded that low language proficiency could prevent students from understanding the feedback and result in low engagement. Other studies have found the opposite and considered the issue of trust in automated tools to play a crucial role in the revision process. This study undertakes a quasi-experimental research design and examines nine Chinese university students' engagement with Grammarly in revising their texts to determine whether language proficiency level and students' engagement are correlated and how the issue of trust interacts with engagement. Results suggest that language proficiency level is not correlated with students' engagement with automated feedback, and trust plays a crucial role but not in a linear manner. Other internal factors such as learners' beliefs and motivation interact with trust in a seemingly dynamic and complex way.

*Keywords:* Grammarly, writing, engagement, EFL, AWE

## ACKNOWLEDGMENTS

First and foremost, I would like to thank all the participants in this study. This project would not have been possible without the generous help from Mr. Andrew Cardoso and his English creative writing class students at Beijing Normal University. Thank you all for your interest in the study and your time and attention spent participating in the study.

I benefited from my thesis tutor Olga Hryckiewicz who guided me from the very beginning. She read my various drafts and provided clear and actionable comments. I am also grateful for all the help and guidance I received from other professors in this program. A special thanks to Prof. Joaquín Romero, who kindly pointed me in the right direction when I was lost in data analysis. I would also like to thank all the incredible colleagues I studied with. I have learned so much from my fellow students, and I am forever grateful for the help and support I received. Finally, I must express my gratitude to my family for the unconditional support and love from the other side of the world and numerous friends that encouraged me along the way. Thank you.

## Table of Contents

<b><i>Table of Contents</i></b> .....	<b>5</b>
<b><i>List of Tables</i></b> .....	<b>7</b>
<b><i>List of Figures</i></b> .....	<b>8</b>
<b><i>List of Abbreviations</i></b> .....	<b>9</b>
<b><i>Glossary</i></b> .....	<b>10</b>
<b><i>I. INTRODUCTION</i></b> .....	<b>1</b>
<b><i>II. LITERATURE REVIEW</i></b> .....	<b>2</b>
<b>2.1. Teaching English Writing as a Foreign Language in China</b> .....	<b>2</b>
<b>2.2. Providing Feedback on L2 English Writing in China</b> .....	<b>3</b>
<b>2.3. Dimensions of L2 Student Engagement with AWE Feedback</b> .....	<b>5</b>
<b>2.4. Previous Research on Student Engagement with AWE Feedback</b> .....	<b>5</b>
<b>2.5. Factors that influence L2 Student Engagement with AWE Feedback</b> .....	<b>10</b>
L2 Proficiency .....	10
Motivation to improve writing skills .....	10
Human-automation trust.....	10
Accuracy, explicitness, and focus of the feedback.....	10
Other factors.....	11
<b>2.6. The Present Study</b> .....	<b>11</b>
<b><i>III. MATERIALS AND METHODS</i></b> .....	<b>12</b>
<b>3.1. Context and Participants</b> .....	<b>12</b>
<b>3.2. Data Collection Instruments</b> .....	<b>15</b>
Screencast.....	15
Stimulated recall.....	15
Semi-structured interviews.....	16
<b>3.3. Data analysis</b> .....	<b>16</b>
Analysis of screencasts .....	16
Analysis of recall and interview data.....	20
<b><i>IV. Findings</i></b> .....	<b>20</b>
<b>4.1. Grammarly Feedback</b> .....	<b>20</b>
<b>4.2. Students' Engagement</b> .....	<b>25</b>
<b>4.3. Factors that Influence Students' Engagement</b> .....	<b>27</b>
Language Proficiency Level (Quantitative).....	27
The Issue of Trust and Other Factors (Qualitative).....	28
<b><i>V. Discussion</i></b> .....	<b>36</b>
<b><i>VI. Conclusion</i></b> .....	<b>38</b>
<b>5.1. Conclusion</b> .....	<b>38</b>
<b>5.2. Implications</b> .....	<b>39</b>
<b>5.3. Limitations and Future Recommendations</b> .....	<b>40</b>
<b><i>VII. References</i></b> .....	<b>41</b>
<b><i>Appendix A</i></b> .....	<b>44</b>

<b><i>Stimulated recall script and questions</i></b> .....	<b>44</b>
<b>A.1. Stimulated recall script</b> .....	<b>44</b>
<b>A.2. Stimulated recall guiding questions</b> .....	<b>44</b>
<b><i>Appendix B</i></b> .....	<b>46</b>
<b><i>Semi-structured retrospective interview questions</i></b> .....	<b>46</b>
<b><i>Appendix C</i></b> .....	<b>47</b>
<b><i>A Sample of Teacher Feedback</i></b> .....	<b>47</b>
<b><i>Appendix D</i></b> .....	<b>49</b>
<b><i>A Sample of Grammarly Feedback</i></b> .....	<b>49</b>

**List of Tables**

Table III-1 .....	13
Table III-2 .....	14
Table III-3 .....	16
Table III-4 .....	17
Table III-5 .....	18
Table IV-1 .....	21
Table IV-2 .....	26
Table IV-3 .....	26
Table IV-4 .....	27
Table IV-5 .....	27
Table IV-6 .....	28
Table IV-7 .....	29
Table IV-8 .....	35

**List of Figures**

Figure III-1 .....	19
Figure IV-1 .....	22
Figure IV-2 .....	22
Figure IV-3 .....	24

**List of Abbreviations**

<b>Abbreviation</b>	<b>Definition</b>
AI	Artificial intelligence
AWCF	Automated written corrective feedback
AWE	Automated written feedback
CET	College English Test
EFL	English as a foreign language
ESL	English as a second language
L2	Second language
RQ	Research question
TEM4	Test for English Major band 4
TOELF	Test of English as a Foreign Language
iBT	Internet-based test

## Glossary

<b>Term</b>	<b>Definition</b>
Grammarly	Grammarly is an English writing assistant that provides feedback on spelling, grammar, punctuation, clarity, engagement, and delivery. It is available across all writing environments, e.g., Microsoft Word, Google Docs, smartphone keyboards, Email, etc.
WeChat	The most widely used instant messaging application in China with 1.2 billion monthly active users.

## I. INTRODUCTION

The current global pandemic has moved much learning and social interaction online, increasing the need for writing in digital environments (Hegelheimer & Ranalli, 2020). For language learners, writing is not only an essential skill to acquire but also crucial for personal and academic advancement. Studies consistently found that EFL/ESL students consider error corrections necessary to improve their written English. They want and expect all errors to be corrected (e.g., Corpuz, 2011; Leki, 1991). Given teachers' heavy workload, scholars in the second language(L2) acquisition field have been interested in the potential benefits of using automated writing evaluation (AWE) tools in students' L2 writing practice.

Extensive studies have investigated the effectiveness of AWE in improving L2 writing from draft to revised tests. However, not much evidence shows the improvement across texts and whether AWE leads to better writing and/or more independent revisions (Leng, 2014; Stevenson & Phakiti, 2019). For learning to happen in L2 writing, engagement with corrective feedback is considered to be crucial. Zhang(2020) states that it is the student engagement with AWE feedback rather than AWE itself promoting language learning (Z. (Victor) Zhang, 2020). However, we know little about this complex, multi-faced and dynamic process(Z. (Victor) Zhang & Hyland, 2022a). Consequently, in recent years, the factors that might encourage engagement to facilitate learning have attracted a growth of research interests (Koltovskaia, 2020; MA, 2019; Ranalli, 2021; Z. V. Zhang, 2017; Z. (Victor) Zhang & Hyland, 2018, 2022b). To delve into this critical question, Koltovskaia(2020) conducted a naturalistic case study exploring two college students' engagement with AWE provided by Grammarly. She found that students had different levels of engagement, and linguistic competence appeared to impact their ability to effectively process feedback—this finding aligned with the conclusion in Zheng&Yu(2018)(Zheng & Yu, 2018a). However, a recent multiple case study found conflicting results. Ranalli(2021) investigated L2 student writers' engagement with automated

feedback and uncovered that the student's proficiency level did not determine their engagement level but their trust in the tool. Motivated by the inconsistent findings in different studies, the present study intended to employ a quasi-experimental approach to provide an additional lens for exploring this complex process. The present study adopted the data collection methods used in the previous studies with intentional modifications. Given the nature of two different research approaches, this study selected samples randomly, while the last two case studies chose appropriate samples for the study purpose.

## **II. LITERATURE REVIEW**

### **2.1. Teaching English Writing as a Foreign Language in China**

In China, English as a foreign language teaching holds a crucial role in higher education, where writing is a necessary skill for both English and non-English majors across all colleges and universities (Mo, 2012). Even though English teaching in this massive, fast-changing country is challenging to generalize, general tendencies, such as centralized national syllabuses and exams, broadly used textbooks, and common beliefs and practices among Chinese English teachers are present (Cortazzi & Jin, 1996). These tendencies resulted in persistent problems in teaching English writing—a heavy emphasis on linguistic accuracy and the final writing product, inadequate input of genre knowledge and assessment variety, and discounting of students' writing development, to name a few (Gao J., 2007). A survey study collected data from five universities in China and found that the teaching of English writing is “unsatisfactory and problematic” (Mo, 2012, p.1). The researcher attributed this situation to both teachers and students, among which are inadequate teaching methodologies, students' insufficient linguistic competence, and the absence of cultural content of the target language (Mo, 2012). Furthermore, contextual factors such as the pressure of College English Test (CET 4 and 6) for non-English majors and Test for English Majors (TEM 4 and 8), together with the curriculum, constrained

English teachers to mainly focus on grammatical correctness and the result in teaching English writing. The intimidating nature of teaching second language writing and the lack of writing teaching training further exacerbated the issue (Dai, 2010).

Nonetheless, in the meantime, important exceptions and varieties were found under these prevailing circumstances. Over the years, both universities and English teachers have sought new approaches to tackle this issue. Sun Yat-sen University incorporated English language creative writing for its university students and found that it was an effective way of teaching English and promoting students' development (Dai, 2010). The 7<sup>th</sup> international conference on English language teaching reported the ongoing trends in China—different pedagogical approaches in EFL writing classrooms, a variety of feedback in EFL writing instructions (teacher feedback, peer feedback, self-editing feedback), and the integration of new technology in teaching writing (C. Zhang et al., 2015). The last-mentioned trend has evolved considerably with China's immense information technology advancement over the past ten years. Applying technology in teaching English writing has been widely adopted; for instance, AI-based writing evaluation systems are commonly used in Chinese higher education (Lu, 2019).

## **2.2. Providing Feedback on L2 English Writing in China**

Studies on feedback in writing have taken a significant part in teaching English writing research. Broadly, it can be divided into three categories based on the source of the feedback: peer feedback, teacher feedback, and automated feedback. Peer and teacher feedback can be further divided into face-to-face and online feedback. Research on how different feedback impacted students' writing motivation and engagement produced inconsistent results. One large-scale study found expressive feedback to be the most used type, and written corrective feedback was the least frequent one; process-oriented feedback was found to discourage students' motivation and engagement in writing while scoring, peer and self-feedback, and expressive feedback had

a positive effect on writing motivation and engagement(Yu et al., 2020). A naturalistic case study revealed a dynamic and mutual engagement with different sources of feedback as well as exposed a mix of individual and contextual factors that interacted with learners' responses to the feedback (Tian & Zhou, 2020).

Studies on teacher corrective written feedback exposed a disparity between students' expectations and teachers' practice. While teachers mainly provided feedback on language-related errors, learners favored extended error corrections on both content and grammar level from their teachers and instructors (Chen et al., 2016; Cheng et al., 2021). Regarding engagement with the feedback, students' low language proficiency negatively affected their cognitive and behavioral engagement with the feedback(Zheng & Yu, 2018b).

Ample studies investigated the effectiveness of peer feedback. Some studies found that peer feedback was reliable and performed almost as well as teachers' feedback in making comments on content level issues. In addition, peer feedback promoted students' autonomy in learning(Wu et al., 2022; M. Yang et al., 2006). Despite the abundant research, peer feedback has not been widely used in teaching English writing in China. This is partially due to the lack of knowledge about peer feedback in English teachers and is further hindered by the examination-orientated education system, the teacher-driven learning culture, and learners' low language proficiency(Zhao, 2018).

Consequently, researchers turned their attention to computer-mediated, or technology-supported peer feedback in recent years. A meta-analysis study collected data from 1568 Chinese L2 writers and found that online feedback from teachers/instructors was more effective than online peer feedback and automated feedback. It also revealed that educational levels and task genre impacted the result of the feedback(Lv et al., 2021). As technology advances, another surging wave of research concerns online feedback not provided by humans but by programs, automated written feedback (AWE).

### **2.3. Dimensions of L2 Student Engagement with AWE Feedback**

This study investigated the aspect of student engagement with AWE feedback. Previous studies revealed that student engagement is a complex process (Z. (Victor) Zhang & Hyland, 2022a). Ellis (2010) used the term ‘engagement’ to describe how students respond to oral and written feedback and suggested it can be studied from three perspectives: behavioral, cognitive, and effective (Sheen & Ellis, 2005). Built on the previous studies, Zhang and Hyland (2018) proposed a model to investigate student engagement with feedback in L2 writing and was used by numerous researchers addressing the issue. They applied a three-dimensional construct as follows:

- Behavioral engagement refers to students’ physical reactions to feedback, including revision actions and time spent on revision.
- Affective engagement comprises students’ emotional responses and attitudinal reactions to feedback, as evidenced in retrospective interviews.
- Cognitive engagement is concerned with how students respond to feedback through revision operations (such as rewriting and reorganization) and cognitive strategies (such as evaluating and monitoring).

(Z. (Victor) Zhang & Hyland, 2018)

Given that the previous studies applied this model to measure engagement level, this study used the same model to maintain consistency for comparing results.

### **2.4. Previous Research on Student Engagement with AWE Feedback**

Research about the effect of AWE feedback on student writing can be classified into three dimensions: the impact on students’ written production, the effect on students’ writing processes, and students’ perception(Lai, 2010). A larger part of the studies was concerned with

written production and students' perception; until recently, little attention was paid to the writing and revising processes.

Researchers have found some evidence that AWE feedback can reduce errors and improve across drafts with the same texts; not much evidence demonstrates long-term improvement in writing across texts. As Hyland and Hyland (2006) noted, evidence that a writer manages to correct errors of a specific text tells us little about whether the writer has developed metalinguistic skills to identify and correct similar errors in different texts (Hyland & Hyland, 2006). Steveson and Phakiti (2019) suggest that behaviorally students may edit their texts after receiving AWE feedback. Still, cognitively they may not have engaged with the input that could facilitate learning (Stevenson & Phakiti, 2019).

Studies that have examined the effect of AWE feedback in the light of student engagement supported this claim, meanwhile showing mixed results in factors that influence successful engagement. Zhang and Hyland (2018) compared how two Chinese students of English engaged with teacher feedback and the AWE feedback tool *Pigai* on their writing over a 16-week semester. Both participants showed behavior engagement with AWCF, with Flora, the highly engaged learner showing considerable emotional investment and cognitive engagement. This particular participant not only corrected most low-level language errors but improved the high-level details such as the style, wording, and explicitness of the text (Z. (Victor) Zhang & Hyland, 2018). In Zhang's follow-up multiple case study, his in-depth analysis of students' revision processes unveiled a similar result. One of the three participants was able to go beyond error correction and evaluate the content and logical flow of his writing (Z. (Victor) Zhang, 2020).

Koltovskais(2020) attempted to address these questions while shifting the focus to an ESL context and a different AWE tool, Grammarly. Two students from a large southcentral university in the U.S, Alex, and Kelsey, completed all aspects of the study among eight

volunteered participants in the English as a second language writing course and became the research subjects. Alex was a 21-years-old junior student who majored in Journalism, and his first language is Cantonese and Mandarin. Kelsey was a 22-years-old native Arabic speaker who was a freshman in the computer science major. To assess the language proficiency level and writing skills of the two participants, the researcher and two of her colleagues independently employed the modified TOELF iBT Test on the participants' written texts. Alex's score(25 out of 30) makes him an advanced L2 writer, while Kelsey received 14 out of 30, which determines her as a low-intermediate L2 writer. For this study, the students wrote literature review assignments and went through three stages of revisions: 1) initial submission for teacher feedback;2) revisions for higher-order concerns; 3) revisions for low-order concerns with Grammarly. A free version of Grammarly was used for this study. It provides feedback on spelling, punctuation, grammar, and conventions (e.g., spacing, capitalization, and dialect-specific spelling). The students were asked to screen record their revision processes with the QuickTime player. Within 48 hours of the revision, the students were shown their screencast videos on a TV monitor and were instructed to recall their thoughts when they made each correction. The stimulated recall method was used to explore participants' cognitive, affective and behavioral engagement with AWE feedback. After the recall, the students took part in a semi-structured interview consisting of ten questions to investigate their affective engagement with Grammarly feedback. The screencasts were used to analyze students' behavioral engagements with five phrases: 1) reclassify the error types according to Han and Hyland(2015)'s taxonomy of error categories; 2) Diagnose the accuracy of the feedback; 3) Identify revision operations in three categories-accepted, rejected and substituted;4) Identify revision strategies(e.g., consulting the Internet or check the dictionary to verify the correction).5) Calculate the time spent, including the total time of revision and time spent on each error. The recall and interview data portrayed their three-dimensional engagement with AWE feedback.

Alex, the high-proficiency learner, received 14 errors in his 914-word draft. All were accurate. He spent around five minutes revising the draft, accepting eight mistakes correctly and rejecting six incorrectly; in total, he fixed 57% of his errors. He used the Google search engine to verify the accuracy once. He appeared to notice the majority of the feedback and understand the eight successful revisions (left out six unsuccessful ones). He often returned to the sentence where the error occurred to determine the appropriate revision operation. He frequently showed distrust in the feedback, which enhanced his cognitive engagement in the revision process. However, instead of looking for additional information, he relied on his linguistic knowledge and intuition to modify his text. Overall, Alex felt positive about Grammarly and found it helpful in ways such as finding mistakes he could not find on his own.

The low-proficiency level student Kelsey received 26 errors in her 810-word draft, of which 69% were accurate. Kelsey spent a similar time to Alex making 15 correct revisions and ten incorrect ones. Unlike Alex, she did not refer to external resources to verify the corrections. In the case of article usage, Kelsey could not articulate the rules and admitted it was her weakness, and thus thought Grammarly was more likely to be correct. This explained her acceptance of all but one article error correctly and incorrectly. At the beginning of the revision, she tended to read the sentences where the errors were flagged, but at the end, she blindly accepted all corrections. Compared to Alex, Kelsey showed a lower level of cognitive engagement. She responded positively emotionally when AWE feedback confirmed her earlier doubts in writing but negatively when she found the feedback inaccurate.

Behaviorally, both students only made moderate changes to their texts on a surface level. Cognitively, the high-proficiency level student was more engaged than the low-proficiency participant. The researcher suggested that Kelsey's insufficient linguistic knowledge affected her ability to be cognitively engaged with the feedback. Affectively, the low-proficiency participant appeared to rely excessively on AWE feedback, especially when she lacked

knowledge about the error, and the high-proficiency student questioned every correction provided by Grammarly (Koltovskaia, 2020).

On the contrary, a recent multiple case study did not find students' language proficiency level related to their engagement with AWE tools but rather the issue of trust. The researcher recruited six Chinese L1 students from a Midwest American university. Each of them provided a text they had already drafted, submitted it to Grammarly Premium, and revised the text based on the feedback provided by Grammarly. Then the researcher carried out a stimulated recall and a semi-structured interview to collect data. One of the high language proficiency participants, Na, who worked as a part-time tutor in an undergraduate English course, was the least engaged student among all the participants. She rarely took time to read the context surrounding the error before accepting Grammarly's correction. She was a Grammarly user but has stopped using it since she deemed Grammarly not useful in helping her or other L2 users learn the target language. She considered Grammarly a labor-saving device and said, "It's like, 70% Grammarly let me know things, like, I already know, and 30% is the stuff I didn't know but rarely use in my daily life" (Ranalli, 2021, p. 31).

On the other hand, a low-level proficiency participant Yong appeared to be the most engaged one in his revision process. He has purchased Grammarly premium and was a frequent Grammarly user. He considered Grammarly helpful in letting him pay attention to the form of the language. Despite his diligent effort in treating Grammarly's feedback, he felt the tool unlikely would improve his grammar. The study concluded that L2 writers adopt a proofreading approach instead of a learning orientation toward AWE feedback; proficiency is not determinative of success engagement; human-automation trust and the characteristics of AWE feedback may play a crucial role (Ranalli, 2021). These contrasting findings speak to the complex and multi-facet nature of engagement with feedback.

## **2.5. Factors that influence L2 Student Engagement with AWE Feedback**

### ***L2 Proficiency***

Zhang and Hyland(2018), Zhang(2020), and Koltovskia(2020) all reported more successful engagement with higher-proficiency students. Students with a lower L2 proficiency were more likely to adopt a form-focused approach and make surface-level corrections. In contrast, the higher-level proficiency student tended to focus on both language and content in their revision process (Z. (Victor) Zhang, 2020). Ranalli(2021) did not find this to be the case. It was the lower-proficiency student who engaged the most with the feedback.

### ***Motivation to improve writing skills***

Previous research suggested that motivation increased student engagement with AWE feedback in the revision process (Z. V. Zhang, 2017; Z. (Victor) Zhang, 2020; Z. (Victor) Zhang & Hyland, 2018).

### ***Human-automation trust***

While trust is known to influence engagement with teacher feedback, Ranalli(2021) suggested that it may play a much more crucial role in engagement with AWE feedback. Compared to teacher feedback, students rely more on themselves to determine how to respond to each piece of feedback, while both their linguistic limitations and limited understanding of the tool itself further complicate the issue.

### ***Accuracy, explicitness, and focus of the feedback***

To extend the human-automation trust factor, researchers suggest feedback accuracy and explicitness could be explanatory factors in understanding student engagement with AWE tools. Studies indicate that explicitness is a decisive factor in students' willingness to utilize the feedback.

### *Other factors*

Zhang(2019) suggested other contextual and individual factors that might affect student engagement, including learner beliefs(e.g., the importance of collocations, the role of revision, and attitude toward punctuation), teaching and learning contexts(teacher stance, workload).

## **2.6. The Present Study**

The foregoing review shows a need for additional research to investigate student engagement with AWE feedback, particularly the factors that influence this dynamic and complex process. The present study is built on the initial Koltovskaia( 2020) and Ranalli( 2021) studies but took a quasi-experimental research design. While the data-elicitation instruments remained as close to the previous studies as possible, this study randomly selected a larger sample size. Some intentional changes were made to further explore the research questions.

The present study took place in an EFL context which differs from the previous two studies, and used Grammarly premium (the paid version) instead of the free version. The paid version provided more comprehensive feedback and was also the choice in Ranalli's (2021) study that produced contrasting results with the initial Koltovskaia(2020) paper. Regarding experimental methods, the stimulated recall sessions were conveyed immediately after the revision process instead of within 48 hours in the initial study.

Informed by the research questions of the previous studies, the study was guided by the following research questions and hypotheses:

### **RQ#1**

What types of feedback do the participants receive from Grammarly, and in what ways do they respond to the feedback?

### **RQ#2**

How do students engage with AWE feedback provided by Grammarly when revising their draft?

**RQ#3**

What factors influence the participants' engagement with AWE feedback, and how do language proficiency and trust affect the participant's engagement?

**Hypothesis 1.** Grammarly will provide mainly low-level language-related feedback, and the students will make moderate revisions according to the feedback.

**Hypothesis 2.** Overall, the participants will adopt a proofreading approach. They will show behavioral and emotional engagement but little cognitive engagement.

**Hypothesis 3.** Issue of trust will appear to be a more robust indicator of engagement level than language.

### III. MATERIALS AND METHODS

#### 3.1. Context and Participants

The students were third-year undergraduate English majors in a public research university in Beijing, China. In the first semester of their third-year study, they had a compulsory English creative writing class for 16 weeks after a year-long elementary English writing course. The teacher was a native English speaker with more than ten years of teaching experience. They met up once a week for 90 minutes. According to the teacher, this class lies somewhere between a typical western creative writing class and rudimentary academic writing, where the focus is on the writing process (A. Cardoso, personal communication, January 11, 2022). The course had three writing assignments that came in the form of writing goals: 1) use implication in writing; 2) use characterization in writing; 3) explore concepts through non-fiction writing of personal experience. The third assignment also acted as the course's final project. The students had to decide what they wanted to write about for each task. The average length of the paper is 1500 words.

The teacher helped the researcher recruit participants through the class' WeChat group, which is the preferred way of communication in China. All 42 students of the class received the recruiting message, of which 12 responded and volunteered to participate in the study. Nine students provided texts and completed the revision process. Table III-1 shows the participants' biodata collected with a questionnaire. Their TEM4 scores from the second year were used to indicate their language proficiency level. TEM4 is the standardized test for English majors in China.

**Table III-1**

*Participant Biodata*

Name*	Sex	Age	TEM4	Type of Grammarly User
Chen YL	F	21	81	First time
Guo XY	F	20	84	Inactive(free)
Wang YX	F	20	64	Inactive(free)
Chen ZT	F	21	83	Inactive(free)
Li J	F	20	86	First time
Luo XY	F	21	84	First time
Gao CX	F	21	64	First time
Li TR	F	21	84	First time
Li YY	F	21	87	First time

*Note.* TEM4=Test for English Major band 4, *free*= free Grammarly account

\*All names are pseudonyms

The students were asked to provide one text they had written for the course that had not yet been checked with any AWE tools. Ranalli (2021) argued, to select text in this manner in hoping that the students would be more willing to improve their drafts with the AWE feedback. In addition, this approach brought the study to a more natural setting. Though it lacked control of text type, I agree that the benefit outweighed the drawback. The submitted texts were assignments from a Creative writing class A, a creative writing class B, and a 19 century English novel class; most of them have already received three rounds of feedback as follows:

1. First-round feedback was peer feedback. Students were organized into groups and read each other's pieces with optional guiding questions.

2. Then the revised texts were handed to the teacher. The teacher provided written feedback along with the grades.
3. Optionally the students could meet up with the teacher and receive oral feedback.

Word counts ranged from 411 to 2058. Frequencies of Grammarly floggings can be seen in Table III-2.

**Table III-2**

*Participant-contributed Texts and Feedback Flagging Frequencies*

Proficiency	Participant	Text Type	Length (words)	Total Flaggings	Flaggings/100 words
Lower	Wang YX	Essay for creative writing A course	2058	108	5.2
	Gao CX	Essay for creative writing A course	1496	108	7.2
Higher	Li J	Essay for creative writing A course	554	30	5.4
	Chen ZT	Essay for creative writing A course	1783	127	7.1
	Chen YL	Essay for creative writing A course	1790	111	6.2
	Li TR	Essay for creative writing A course	1894	104	5.5
	Li YY	Essay for creative writing A course	1482	78	5.3
	Guo XY	Essay for creative writing A course	1285	69	5.4
	Luo XY	Essay for creative writing A course	411	26	6.3

Grammarly was chosen for this study because it was the choice of the previous studies. In addition, Grammarly is one of the most widely used AWE tools, with 30 million daily active users (Grammarly, 2021). Moreover, it has a high error detection and accuracy rate. A recent study evaluated the accuracy of corrective feedback from Grammarly and found that this online grammar checker has a 65% error detection rate with a strong performance at catching articles and preposition errors. Furthermore, Grammarly provides mostly accurate suggestions and rarely makes false alarms (Moon & Prof, 2021).

### **3.2. Data Collection Instruments**

Adapted from previous studies, three instruments were used for data collection: screencasts, stimulated recall, and semi-structured interviews. Whereas the initial study analyzed engagement days after the event, the present study collected data immediately after the revision process to capture the cognitive and affective elements stored in short-term memory as accurately as possible.

#### ***Screencast***

The participants were asked to provide a text that they had already written and revised based on teacher feedback for higher-order concerns but not yet checked with an AWE tool. For each student, their text was then uploaded to Grammarly premium, and the researcher downloaded a Grammarly report that contains a full flagging record. Each flagging is labeled with the corresponding error type. Then the researcher initiated an online video call where the revision process was screen recorded using the video conference tool Tencent Meeting. First, the student was facilitated with Grammarly premium account access and the uploaded text. Then, the student was instructed to revise their text based on Grammarly suggestions and be as natural as possible. The recordings were later coded for analysis to address RQ1 and RQ2, as well as contribute quantitative data to explore RQ3.

#### ***Stimulated recall***

Right after the revision, the student then was shown the recording and asked to recall their thoughts and emotions in the event of each revision operation. The recall script and guiding questions were provided by the researcher of the initial study, which can be seen in Appendix A and B. This method explored students' engagement with AWE in all three dimensions.

### *Semi-structured interviews*

Subsequently, the researcher had a semi-structured interview with the student using the same ten questions of the initial study(See Appendix C). Follow-up questions were asked to delve deeper into the complex picture.

The recall and interview sessions were conducted in Chinese/English and were audio-recorded. In total, it lasted from 30 to 90 minutes per student. Later, based on the quantitative revision data, five students were selected for the case-study narrative to provide qualitative data addressing RQ3.

**Table III-3**

*Data Sources and Analyses Used to Address Each Research Question*

Research Question	Data Source	Analyses
RQ1: What types of feedback do the participants receive from Grammarly, and in what ways do they respond to this feedback?	Grammarly Report, Screencast	Statistics and visual analysis of flagging category, flagging accuracy, and revision operation data
RQ2:How do students engage with AWE feedback provided by Grammarly when revising their draft?	Screencast(quantitative) Stimulated recalls/semi-structured interviews(qualitative)	Statistics of participants' viewing data, average engagement time per flagging data, demonstration of revision strategy, and success inaccurate flagging revision rate.
RQ3:What factors influence the participants' engagement with AWE feedback?	Screencast(quantitative) Stimulated recalls/semi-structured interviews(qualitative)	Engagement level Case-study narrative

### **3.3. Data analysis**

#### *Analysis of screencasts*

The researcher examined the screencasts in six phases. Firstly, the flaggings were classified by category of error where Grammarly's own classification system was used, e.g., *correctness*, *clarity*, *engagement*, and *delivery*. The category *correctness* was further classified into three divisions: *spelling*, *grammar*, and *punctuation*. Table III-4 shows the classification.

**Table III-4***Category of Error*

Categories	Flagging type
Grammar	Faulty tense sequence Faulty subject-verb agreement Wrong or missing prepositions Text inconsistencies Pronoun use Determiner use Conjunction use Incorrect verb forms Incorrect noun number Misuse of modifiers Modal verbs Misplaced words or phrases Incomplete sentences Incorrect phrasing
Punctuation	Closing punctuation Punctuation in compound/complex sentences Comma misuse within clauses Misuse of semicolons, quotation marks, etc.
Spelling	Confused words Misspelled words Improper formatting Commonly confused words Mixed dialects of english Unknown words
Clarity	Intricate text Unclear sentences Wordy sentences Hard-to-read text Passive voice misuse
Engagement	Word choice, monotonous sentences
Delivery	Tone suggestions Inappropriate colloquialisms

Grammarly's *clarity* suggestions are based on the idea that "The most effective writing is concise—free of filler words, with just the right tone to get your point across"(Grammarly, 2020, para.3). It focuses on three aspects: cut unnecessary words and phrases, identify passive voice, and rephrase wordiness. *Engagement* in Grammarly's term refers to avoiding using "bland words", "overused words," and "monotonous passages"(Grammarly,2019, para.5). *Delivery* concerns with pragmatics, including politeness, formality level, inclusive language, confident language, etc. Wang YX received the most *delivery* flagging as her text was the only academic paper among all texts. Table III-5 presents examples of each category.

**Table III-5***Examples of flagging category*

Category	Sentence	Grammarly Suggestion
Clarity	The three minutes <u>in which</u> he stands by me and looks at my computer screen feel like a decade.	<del>in which</del> It appears that <i>in which</i> may be unnecessary in this sentence. Consider removing it.
Engagement	[...] Basil, as a painter who always pursues beauty, they interpret Wilde's unique and <u>strong</u> aestheticism in different ways and from different perspective.	<del>strong</del> → robust The word <i>strong</i> is often overused. Consider using a more specific synonym to improve the sharpness of your writing.
Delivery	<u>But</u> Dorian was bewitched by Henry's hedonism, jealous that the portrait could stay young forever, ...	<del>But</del> → However, Nevertheless, <i>But</i> may not be the best choice here. Consider replacing it with another word or removing it.

Next, under each category, different error labels were identified and coded; for instance, *Misspelling words* in the spelling category and *Determiner use* in the grammar category.

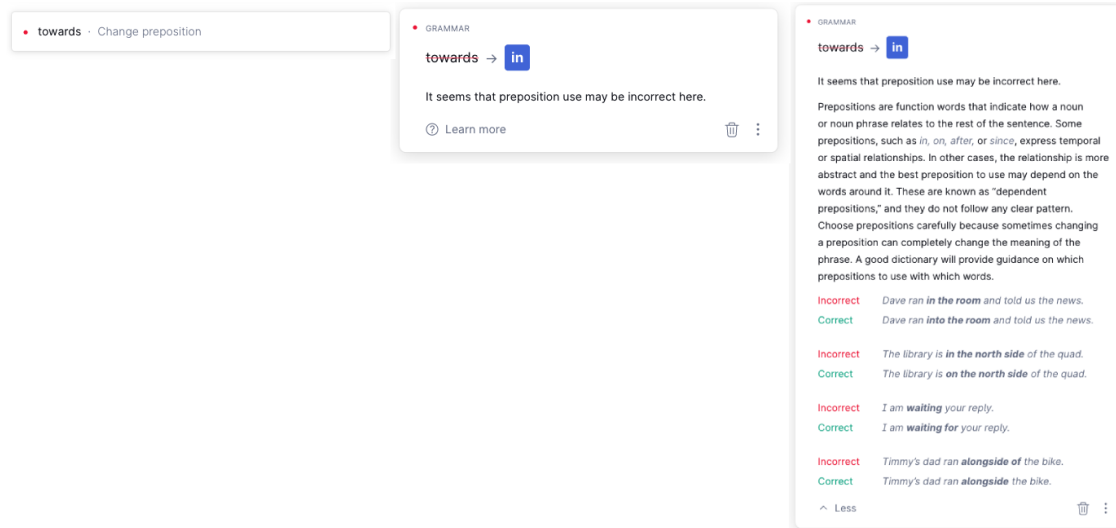
Thirdly, to examine students' engagement in relation to the accuracy of the feedback, each flagging was further classified by its correctness using three codes: *accurate*, *inaccurate*, and *indeterminate*. The last code is used for arguable flaggings where it cannot be determined whether the suggestion can improve the text quality; it appeared most frequently in clarity and engagement categories.

Later, revision operations were identified and classified into three divisions: *accept* (the participant used Grammarly's suggestion), *ignore* (the participant dismissed Grammarly's feedback), and *modify* (the participant applied a change that was not Grammarly's proposed option). Meanwhile, viewing options of Grammarly's flagging were assigned to each revision action: *collapsed*, *expanded*, and *learn more*. Additional revision strategies such as consulting

search engines, and looking up online dictionaries were also considered and were categorized as *additional*.

### Figure III-1

Three viewing options for Grammarly feedback: collapsed(left), expanded(middle), and more(right).



The fifth phase concerns the demonstration of cognitive strategies, for instance, saving the unsure flaggings to the end, going back to the sentences where the flagging occurred, evaluating and monitoring, and revising the text in multiple stages.

The last phase was the amount of time spent in revision. The total time of working on the revision and the average time spent on each error were calculated.

Subsequently, to answer research question 3, a two-dimensional engagement scale was constructed based on Zhang&Hyland's engagement model. To measure behavioral engagement, average engagement time per flagging, and viewing modes were considered. To assess cognitive engagement, *modify* revision response rate (including rewriting and reorganization) was evaluated, as well as a demonstration of cognitive strategies and success rate in rejecting inaccurate flaggings. The students were ranked on each dimension and were given a number rating from 1 to 9. The final rating was then calculated by multiplying the scores. The bigger the number, the higher the engagement level.

### *Analysis of recall and interview data*

Five students' recall and interview data were chosen to be analyzed based on the representativity of their revision operations. The researcher transcribed the recall and interview audio recordings, and only language-related episodes were selected. Then the researcher's questions and the participants' responses were separated. The students' responses were further coded with Atlas software. For interviews conducted in Chinese, the chosen data was then translated into English by the researcher.

## **IV. Findings**

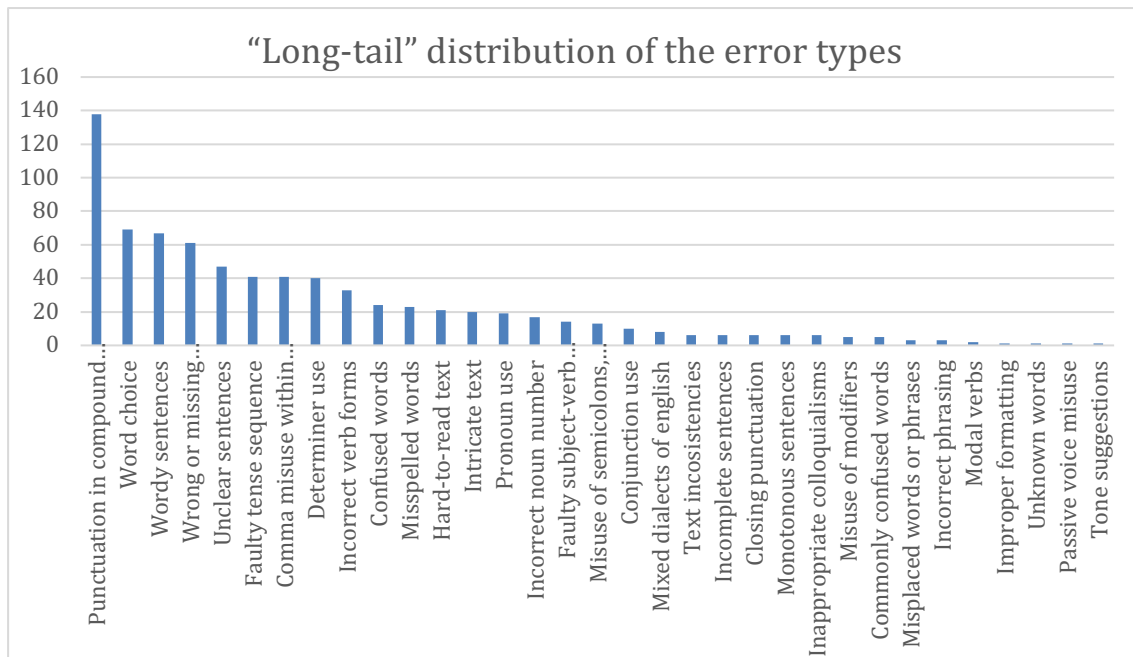
### **4.1. Grammarly Feedback**

RQ1 asked about the types of feedback the participants received from Grammarly and how the students responded. Table IV-1 shows the frequency of 758 total flaggings sorted by feedback categories. In total, 33 different error types were detected across all the texts supplied by the participants. For all participants, Grammar, punctuation, and clarity accounted for the most frequent feedback, while *delivery* was the least frequent. Li J and Wang YX were the only two students who received *delivery* feedback. Wang YX was the only student who provided an academic paper and received six *delivery* feedback. Low-level language-related feedback (*Grammar, Spelling, Punctuation*) accounted for 68% of the feedback, the rest concerned with high-level content issues.

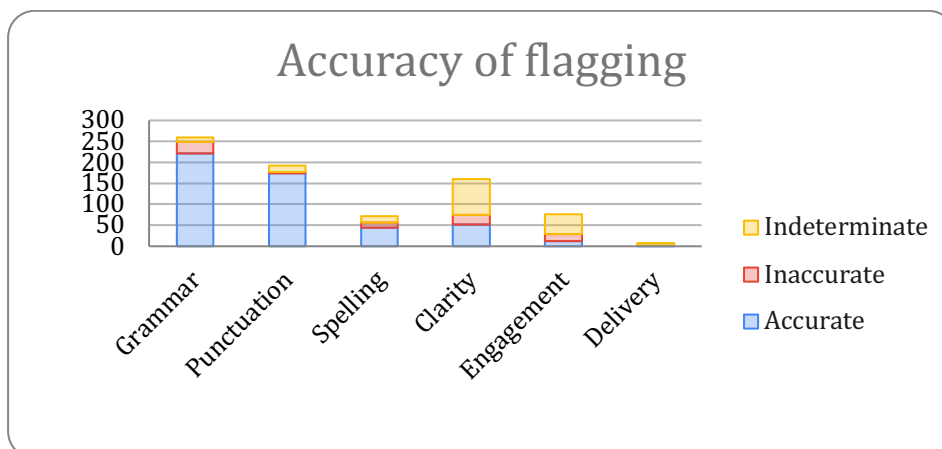
**Table IV-1***Frequency of flaggings for different feedback categories*

	Lower-Proficiency		Higher-Proficiency							Total
	Wang YX	Gao CX	Chen ZT	Guo XY	Li YY	Chen YL	Li J	Luo XY	Li TR	
Grammar	47	40	44	43	27	36	3	7	34	274
Spelling	4	7	10	3	11	3	10	4	5	53
Punctuation	19	36	40	8	14	33	6	12	17	173
Clarity	22	15	29	8	16	28	9	3	26	153
Engagement	10	10	4	7	10	11	1	-	22	75
Delivery	6	-	-	-	-	-	1	-	-	7

Overall, the most frequent error type was *Punctuation in compound/complex sentences* which includes *Joining two independent clauses improperly with a comma* and *Missing a comma before the coordinating conjunction and/or in a compound sentence*. The second frequent error that occurred was *Word choice* in the engagement category. *Wordy sentences* ranked the third most frequent error type where words such as *really, kind of, around, and so* were often considered unnecessary; phrases like *more and more fast* were suggested to change to *faster*. The most frequent grammar error was *Wrong or missing prepositions* and it was followed by *Faulty tense sequence, Determiner use, and Incorrect verb forms*. The grammar mistake distribution matches with common grammar mistakes made by L1 Chinese speakers that were identified by researchers (e.g., Swan & Smith, 2001). Figure IV-1 shows the complete frequency count of each error type.

**Figure IV-1***“Long-tail” distribution of error types*

Regarding the accuracy of flagging, the overall accuracy rate was 66%, 10% of the flaggings were inaccurate, and 24% were indeterminate. This result confirmed the accuracy rate found in recent studies (Moon & Prof, 2021). Figure IV-2 shows the distribution of the accuracy rate down by each feedback category.

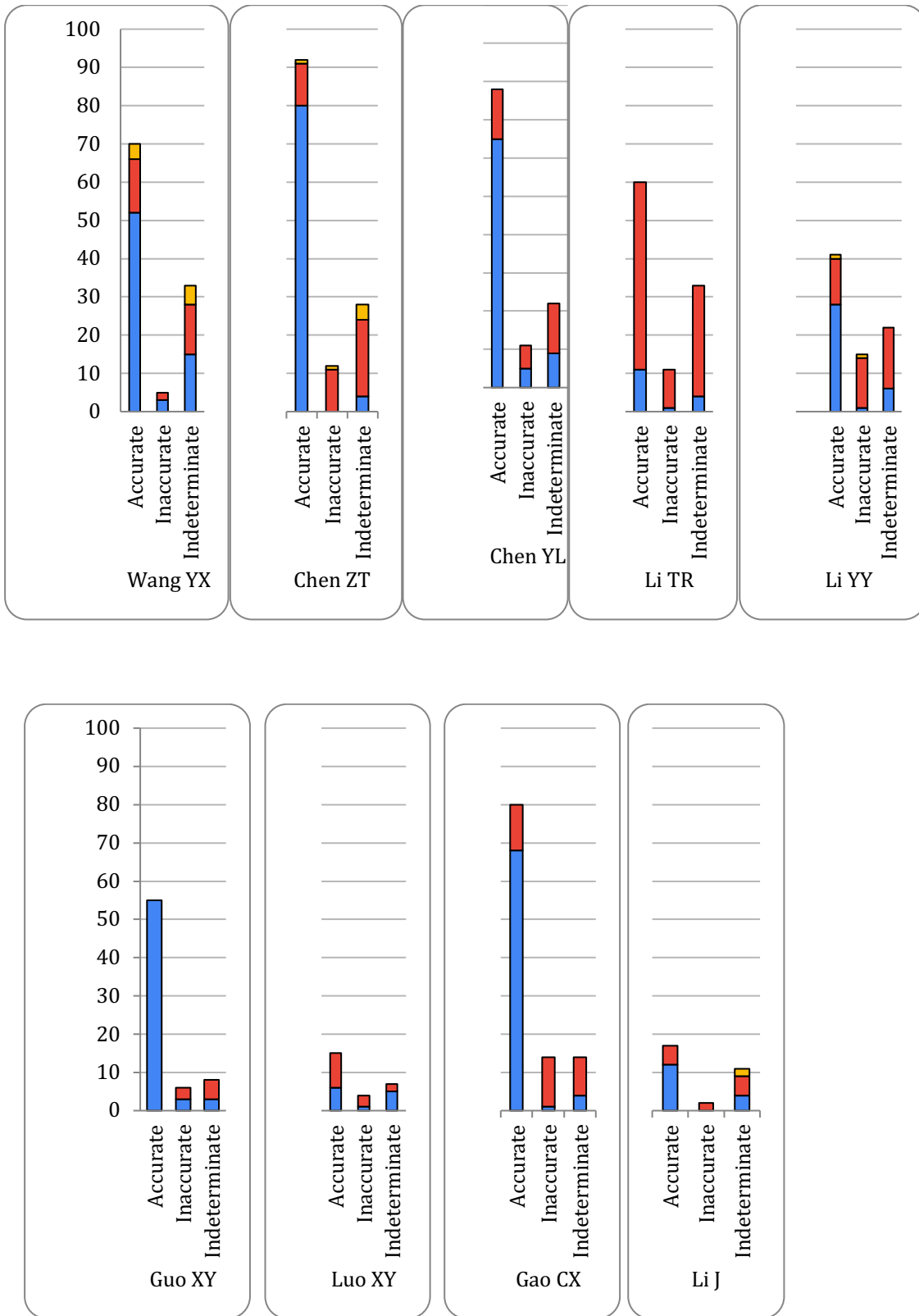
**Figure IV-2***Accuracy rate sorted by feedback category*

*Punctuation* mistakes had the highest accuracy and the lowest inaccuracy rates, making it the most accurate feedback category. The highest inaccuracy (21%) and indeterminate (62%) rates appeared in the *engagement* category. In general, *grammar* feedbacks were accurate (86%). This unveils the nature of Grammarly's feedback that it is more accurate in lower-level language-related issues but less accurate in higher-level content concerns. Nonetheless, 32% of *clarity* feedback, 25% of *delivery* suggestions, and 17% of *engagement* suggestions were correct.

In relation to how flagging accuracy influenced students' responses, cross-referencing data can be found in Figure IV-3.

**Figure IV-3**

*Frequencies of response types according to the accuracy of flagging sorted by feedback category*



In general, feedback accuracy influenced students' responses. Accurate feedback had a higher acceptance rate; all participants were selective in accepting indeterminate feedback. However, significant variation can be seen in responding to accurate feedback. Guo XY accepted all accurate suggestions, while Li TR dismissed most of them and left 71 suggestions unresolved. While the accuracy of the feedback suggested a correlation to students' feedback response, individual factors seemed to play a crucial role. On the whole, students who had a high acceptance rate tended to do so across all feedback categories and feedback accuracy rates. Guo XY falsely accepted numerous inaccurate flaggings, but Li TR rejected most of them. Overall, students made moderate to high degrees of modification to their text. Chen ZT had the highest success revision rate with inaccurate feedback and provided the most modifications(6). The other three students that made additional modifications were Wang YX(5), Li YY(2), and Li J(1).

#### **4.2. Students' Engagement**

RQ2 was concerned with how the students engage with Grammarly feedback. Since all participants showed emotional engagement (e.g., surprised to see the amount of flagging she received, like or dislike of the feedback and Grammarly, etc.), a two-dimensional engagement scale that investigates students' behavioral and cognitive engagement was employed to answer this question. Feedback viewing data was used to measure students' behavioral engagement. The data were first sorted by the *Additional* category, from smallest to biggest, as it was the option that required the highest engagement. Then it was sorted by the *Learn more* category from low to high. Lastly, it was sorted by the *Collapsed* category from biggest to smallest with the current result. Table IV-2 showed the feedback viewing data sorted and the score each student received correspondently.

**Table IV-2***Sorted feedback viewing data*

Participant	Collapsed	Expanded	Learn		Score
			More	Additional	
Gao CX	71	120	0	0	1
Guo XY	44	123	0	0	2
Chen YL	7	183	0	0	3
Li TR	5	155	0	0	4
Wang YX	1	178	0	0	5
Luo XY	0	43	0	0	6
Li J	0	49	2	1	7
Li YY	0	130	8	4	8
Chen ZT	0	221	11	10	9

Engagement time measurements concerned with average time spent per flagging, as well as average time spent in relation to text length, and it was sorted in the mentioned order.

**Table IV-3***Sorted engagement time data*

Participant	Time	Time in seconds	Text length	Flaggings	Time/flagging	Time/length	Score
Guo XY	00:09:35	575	1285	69	8,33	0,45	1
Gao CX	0:20:26	1226	1496	108	11,35	0,82	2
Li TR	00:23:28	1408	1894	104	13,54	0,74	3
Li YY	00:17:40	1060	1482	78	13,59	0,72	4
Chen YL	00:26:27	1587	1790	111	14,30	0,89	5
Wang YX	00:39:11	2351	2058	108	21,77	1,14	6
Luo XY	00:09:26	566	411	26	21,77	1,38	7
Chen ZT	00:48:18	2898	1783	127	22,82	1,63	8
Li J	00:17:13	1033	554	30	34,43	1,86	9

The cognitive engagement was assessed firstly with the total amount of *modify* revision responses, then the demonstration of cognitive strategies (1 = positive, 0 = negative), and the success rate in rejecting inaccurate flaggings. Table IV-4 showed the sorted cognitive engagement data.

**Table IV-4***Cognitive engagement ranking*

Participant	Modify	Cognitive strategies	Inaccurate success rate	Score
Chen YL	0	0	0.545	1
Luo XY	0	0	0.750	2
Gao CX	0	0	0.929	3
Li TR	0	0	0.909	4
Guo XY	0	1	0.500	5
Li J	2	1	1.000	6
Li YY	2	0	0.933	7
Chen ZT	6	1	1.000	8
Wang YX	9	1	0.400	9

A total engagement score was calculated by multiplying each score and can be found in Table IV-5.

**Table IV-5***Total engagement score*

Participant	Feedback Viewing	Engagement Time	Cognitive Engagement	Total
Chen ZT	9	8	8	576
Li J	7	9	6	378
Wang YX	5	6	9	270
Li YY	8	4	7	224
Luo XY	6	7	2	84
Li TR	4	3	4	48
Chen YL	3	5	1	15
Guo XY	2	1	5	10
Gao CX	1	2	3	6

### 4.3. Factors that Influence Students' Engagement

#### *Language Proficiency Level (Quantitative)*

RQ3 asked about the factors that influence students' engagement levels, e.g., language proficiency level and the issue of trust.

Regarding proficiency level, a correlation coefficient analysis was conducted, engagement score and TEM4 score were not correlated (0.12), see Table IV-6. This further proved Ranalli

(2021) study's finding that efficiency level does not appear to be a factor that influences engagement level. The following case study narrative further verified this result.

**Table IV-6**

*Correlation coefficient test of Engagement score and TEM4 score*

	<i>Engagement score</i>	<i>TEM4 score</i>
Engagement score	1	
TEM4 score	0,123444144	1

To investigate the issue of trust, a case-study approach was used. Two high engagement students (Chen ZT, Li J) and three low engagement participants' (Gao CX, Li TR, and Guo XY) stimulated recall, and interview data were used for analysis.

***The Issue of Trust and Other Factors (Qualitative)***

*Chen ZT: High proficiency, moderate trust, high engagement student*

Chen ZT's creative writing piece received 126 suggestions from Grammarly. Later in the interview session, she expressed being shocked by the high number of flaggings she saw. While commenting on her revision operation, the most frequent emotions she vocalized were hesitation, curiosity, uncertainty, feeling strange, and suspicious. Her effective engagement was reflected in her revision operations. She sought additional information ten times and looked at *learn more* options 11 times during her revision process. Both actions appeared more times than all participants combined. She also showed high cognitive engagement by providing six modifications, demonstrating a conscious revision strategy, and having a 100% success rate in revising inaccurate flaggings. Overall, she received the highest engagement score. In the stimulated recall, Chen ZT demonstrated cognitive engagement that could potentially lead to learning.

Among the 126 suggestions, four were related to misuse of the period within quotation marks.

Table 1IV-7 shows two of the four flaggings.

**Table IV-7**

*Two examples of punctuation flaggings Chen ZT received*

Sentence	Grammarly suggestion
"I was wondering maybe we can buy directly from them." I ask Ling.	them. → them, You have used a period in the middle of a sentence. Consider replacing it with a comma.
[...]I don't mind paying them a few dozen kuai for their time." Ling responds.	time. → time, You have used a period in the middle of a sentence. Consider replacing it with a comma.

The first time she encountered this flagging, she paused and clicked on *learn more* option, read it for about a minute but neither accepted nor dismissed the suggestion, and continued with the next flagging. She did the same with the other ones. After she reached the end of the text, she returned to the beginning and revised it for the second time. When she saw the flagging, she typed "*end a quote with a comma when the sentence is finished*" in the Google search box. Not satisfied with the search result, she then tried other keywords. In the end, she typed an example sentence, "*good luck,*" she said to see if such use exists. After confirming the usage, she went back to Grammarly and accepted all suggestions of this type.

Later in the stimulated recall, she commented:

Here it said I used a period in the middle of a sentence, I didn't know about it at all. Then...I thought, isn't it that when the sentence is not finished only a comma is used? Here...I seemed to hesitate for a long time. Then...in the end, I Googled a bit, and decided to accept its suggestions for this type of error.

Similar punctuation mistakes were common in other students' texts, but Chen ZT appeared to be one of the few students who understood the suggestion and potentially learned the linguistic knowledge. Similar engagement could also be seen in the use of the present simple tense in telling stories, the different usage of the preposition "of/from," etc.

Her response to the comma misuse error demonstrated a clear, conscious revision strategy. First, she was unsure about the suggestion, so she hesitated and clicked on the “*learn more*” option. Then if she was not satisfied with the explanation, she left the suggestion aside and went back to it later. This strategy was consistent in cases of other suggestions that she was unsure of and eventually showed engagement that could potentially lead to learning. She was skeptical about *clarity* and *engagement* suggestions and was selective in her responses. For obvious incorrect suggestions, she found them confusing. She explained: “Like here, I don’t know why it suggests me to change it. I am like...ok, I didn’t accept it at all. Yeah, I think it doesn’t understand what I am saying here.” However, this does not make her distrust all suggestions in the same category. After modifying her sentence based on the Grammarly suggestion hard-to-read, she commented:

[...]it said my sentence is hard-to-read to a general audience, it made me think if it’s true. Why did it make this suggestion? ...I know that as a writer, sometimes it could be difficult to know if it is hard to read. Maybe Grammarly could really see it from the perspective of a general audience.

In total, she accepted four and modified 4 out of 28 indeterminate *clarity* and *engagement* suggestions.

She responded differently when she was confident, for example, adding a comma before *and*, or the use of prepositions. She reasoned, “[...] here it wanted me to add a comma before *and*, almost all suggestions like this I accepted unless I think it would affect the flow of the sentence. Yeah yeah, I accepted most of them.”

Later in the semi-structured interview, she considered Grammarly useful and explained her trust in Grammarly:

I think about...complicated sentences, for example, *a clerk of bakery to bakery clerk*, when it wants me to trim my writing, I think ok, I could accept. In terms of grammar, including punctuation, I think it’s quite...before I really didn’t know that you should

end a sentence with a comma. Prepositions, I think...I know it's quite important, I trust it quite a lot. But there were places the prepositions were not appropriate. I think all these things, like low-level grammatical errors, such as subject-verb agreement, I quite trust it, and accepted all. I don't trust its suggestions in word choice, for example, change *big* to *huge* or *massive*. Moreover, it considered some of my sentences unnecessary, I think it's important for conveying the meaning.

She pointed out that since her text was already revised regarding high-level writing issues, she mainly expected Grammarly to correct lower-level grammatical errors, thus, she was satisfied with the feedback. If the text had not been revised before, she thought that Grammarly might not have been useful in detecting logic, cohesion, and meaning issues.

*Li J: High proficiency, moderate trust, high engagement student*

Li J's text had the fewest grammar mistakes and an overall low flagging rate, but she spent the most time on each flagging and per word on average. For her 30 flaggings, she checked the *learn more* option twice and consulted additional information once. Her success rate of responding to inaccurate suggestions was also 100%, and provided modifications outside of Grammarly's suggestions twice. Later in the stimulated-recall session, she said she was unaware of the *learn more* option until later in the revision process, which could explain her low number in this viewing option. She demonstrated a similar revision strategy as Chen ZT: saving issues she was unsure of for later and revising the text in multiple rounds. Li J was the other student who noticed her mistake in comma misuse and understood the problem. She often put commas outside of quotation marks in her text, and Grammarly suggested moving the comma inside. After encountering the same flagging many times, she realized this was an issue and clicked on *learn more* to understand the linguistic knowledge behind it. Later she recalled:

I have some, some spelling mistakes, very stupid, and also standard usage, about punctuation, sometimes I didn't notice, and don't know the criteria, it's outrageous that

I didn't unify it. The whole article has punctuation problems because I did not know how to use punctuation marks. After seeing it a lot, I realized this is a problem...there are rules. It helped me to correct, I also learned a bit.

Apart from punctuation, she also found Grammarly helpful in correcting preposition mistakes and detecting clarity issues. Of the seven clarity flaggings she received, two were inaccurate, and five were marked as indeterminate. She modified two out of the five indeterminate suggestions. She said Grammarly made her realize that her sentences were "too long, too scattered." Overall, she said, "Grammarly is very good at the mechanical part and helpful in clarity." She considered Grammarly not only helpful in improving the text quality but also in facilitating learning.

Above all, Li J showed a greater understanding of how Grammarly works. She explained:

Machines have a keener sense than humans, for example, spelling mistakes. When professors see this, they would mind it too. Also, some of my expressions were not authentic, it might base its suggestions on database, might know more than I do. It at least could help me notice some problems, and I can learn more and change them.

*Gao CX: Low proficiency, low trust, and low engagement student*

Gao CX and Guo XY were the two participants who viewed many suggestions in *collapsed* mode and spent the least amount of average time per flagging and per word in their revision processes. Guo XY showed moderate cognitive engagement, while Gao CX's cognitive engagement appeared to be low.

Gao CX has a TEM4 score of 64 and is considered a low-proficiency participant in this study. Her text had the highest flagging per length number (every 7.21 words a flagging), as well as a high grammar and punctuation flagging number. She took a proofreading approach to Grammarly feedback, only made changes at the surface level, and ignored most suggestions relating to clarity and engagement.

In the stimulated-recall and interview session, she appeared to be a confident L2 writer and stressed the importance of the writer's choice in writing, and did not consider grammatical correctness important in L2 writing. For instance, she ignored many punctuation suggestions, and reasoned: "many suggestions about commas were not correct because punctuation was a good way to convey my emotion." The only type of suggestion she found useful was about verb tenses, she explained:

When we write, we just think about story, we didn't pay much attention in something like tense. I think part of my article talks about the past, but I used the wrong tense in this part. When I write it, I ignored the tense, and just keep writing.

She distinguished creative writing from academic writing and argued that L2 writers were exempted from grammatical rules for expressing meaning.

Stories you write is for your reader, not for a university or a course, if it's for a course, it must be correct. Also like Chinese writer, sometimes they use sentence or something in the wrong way, because they meant it. Because they want to communicate in this way.

She suggested that Grammarly could only help with basic grammar errors, and did not consider Grammarly useful for language learning. She explained:

I think it's good for revising article, when we write article, we first think about the story. Like me, I seldom think about Grammar. I just write it. When I finish, I use Grammarly to see if there are mistakes. It will not also give me some lessons about grammar. That is enough for this kind of website, it is not for teaching you how to learn English grammar. Sometime maybe when I feel confused about my mistake about my article, I will search on the internet about the meaning of the word...knowledge about the grammar in other websites.

Her low language proficiency level did not lead her to trust Grammarly blindly like other low proficiency students in the previous studies. Her confidence as a writer, as well as her beliefs about writing, made her dismiss grammatical correctness as essential and necessary. She considered Grammarly as merely a proofreading tool, and neither could it understand her writing nor could it help her gain linguistic knowledge.

*Li TR: High proficiency, low trust, low engagement student*

Li TR had one of the lowest acceptance rates of suggestions among all participants and left 71 suggestions behind. Unlike Gao CX, she is a high proficiency level student but expressed a similar attitude towards Grammarly. She used “*mechanical*” to describe Grammarly, and considered it only useful in specific grammar rules, and not helpful in detecting *clarity* and *engagement* issues; therefore she ignored most of them.

*Guo XY: High proficiency, high trust, low engagement student*

Guo XY had seen the advertisement of Grammarly before and downloaded a desktop version to try. However, she did not find out how to use it. Her 1285 words story received 69 flaggings. Her average flagging rate was in the lower end among all participants. Grammar mistakes accounted for 62% of all suggestions, among which *wrong or missing prepositions*, *determiner use*, and *incorrect verb forms* were the top 3 most frequent error labels. She accepted 61 out of 69 suggestions, ignored the rest, and made no modifications. Averagely, she spent the least amount of time on each flagging (8.33 seconds) which was less than half of the mean (17.99 seconds). She also had a high inaccurate flagging acceptance rate (50%).

In the stimulated recall session, she used “*authentic*” and “*more authoritative*” to describe Grammarly. When she was unsure, she chose to trust Grammarly across all categories. For Grammar errors, she usually accepted immediately. She explained, “[...] here is a grammar error, it’s probably right, so I accepted [...]it wants me to add a comma, I think it should be

right.” She also tended to trust Grammarly's suggestions for clarity and engagement suggestions.

Table IV-8 shows a flagging coded as indeterminate; Guo XY accepted the suggestion.

**Table IV-8**

*An indeterminate flagging that Guo XY received*

Sentence	Grammarly suggestion
Li Ping exclaimed with a sparkle of excitement in her bleared eyes, "Eh... old things, not worth mentioning... matchmaking corner!"	Your sentence may be unclear or hard to follow. Consider rephrasing. <del>Li Ping exclaimed with</del> <i>With</i> a sparkle of excitement in her bleared eyes, <i>Li Ping exclaimed,</i> "Eh... old things, not worth mentioning... matchmaking corner!"

She explained: “It says it’s unclear, I think both are correct, but it suggested to change it, maybe its suggestion is more authentic, so I decided to change it.” In another similar instance, she commented,

[...]maybe what I wrote doesn’t comply with the norms of English, and the suggestion is strange... After finishing a sentence, should be followed by who said it, then when it was said...But I think Grammarly is more authoritative, so I accepted it.

In the sentence “*Oh! The corner must have some magic! Ah! Then why you still come here? For another child?*” Said Li Ping with a little smile. Grammarly suggested changing *little* to a *bit*, reasoning that *little* is often overused. She first accepted the suggestion, then changed it back to her original wording, and then deleted the adverb. She explained: “I don’t understand so I changed it back, but because I don’t know, so I simply deleted the whole thing, just say ‘a smile’.”

Overall, she expressed her lack of confidence in determiner, preposition, and verb tenses usage (“I am not good at all these”), as well as pragmatics (“maybe it is not in line with foreigners’ usage habits”). Despite having a high language proficiency level, her lack of confidence in her linguistic knowledge contributed to her high trust in Grammarly. It prevented her from engaging more with the feedback in a way that could potentially facilitate learning.

## V. Discussion

Hypothesis 1 thought Grammarly would provide mainly low-level language-related feedback, and the students will make moderate revisions according to the feedback. The findings agreed with the first statement but showed mixed results with the second one. In this study, Grammarly provided mainly low-level language-related feedback (68%) where the distribution of *grammar* mistakes resembled common grammar mistakes made by L1 Chinese speakers. Overall, the feedback was accurate(66%), only 10% of flagging was deemed inaccurate, and the rest(23%) were classified as indeterminate. *Punctuation* and *grammar* feedback had a high accuracy rate, indicating that Grammarly excelled at identifying and correcting mechanical issues in writing. Although it was less accurate in providing content level feedback, 32% of *clarity* feedback, 25% of *delivery* suggestions, and 17% of *engagement* flaggings were accurate. Accurate feedback received the most acceptance from the participants, while indeterminate feedback was responded to more selectively. Great variation can be seen in the acceptance rate and the degree of revisions made by the students. Guo XY accepted most of the feedback across categories, while Li TR dismissed most of them.

The study showed mixed evidence of potential language development in using Grammarly. Overall, all participants made surface-level changes to their texts and found Grammarly useful in these errors. Only some high-engagement students made high-level language-related and content-level modifications. Chen ZT and Li J both demonstrated learning effects and admitted that Grammarly could facilitate learning, at least at the noticing level. All participants felt positive about Grammarly's feedback and admitted they would consider using it in the future. Hypothesis 2 claimed that the participants would adopt a proofreading approach, but this was not the case for all participants. It was true that some low-engagement students only made surface-level changes to their texts; high engagement students such as Chen ZT and Li J made not only surface-level changes but also improved high-level details such as style, sentence structure, wording, etc. Both behavioral and emotional engagement was seen during the

revision process but with great variations. The high engagement students spent almost twice as much time revising and sought additional information more often. Explicit cognitive engagement can be seen among four participants. Several participants acknowledged the language learning potential of using Grammarly (e.g., Li J. Chen ZT) and demonstrated learning-oriented behavior.

Hypothesis 3 considered the issue of trust as a more robust indicator of engagement level than language proficiency level, and this was found true. Language proficiency level alone did not appear to be correlated to students' engagement with AWE feedback. Two high language proficient students, Chen ZT and Guo XY, showed contrasting results in engagement. This could explain the mixed results produced by the previous case studies. Koltovskais (2020) contributed proficiency to the ability to understand feedback, therefore, promoting engagement. This was not found in this study. Trust in AWE seemed to play an important role, as suggested by Ranalli (2022), and it was influenced by the student's confidence in a profound way. Low confidence student Guo showed high trust in the tool, while high confidence student Li TR and Gao CX showed low trust regardless of language proficiency level. Moreover, the relationship between trust and engagement did not appear to be linear. This study found low engagement in cases of both low and high trust. Distrust in AWE could be the starting point of cognitive engagement, as shown by Alex in Koltovskais's (2020) study, and the two high engagement students in this study, but it would not work alone.

Learners' beliefs (e.g., the role of revision, and attitude toward correcting grammatical errors) and motivation to improve writing skills through AWE feedback appeared to be important factors. The two low-engagement students, Li TR and Gao XY did not consider correcting grammatical errors important, while Chen ZT and Li J considered it crucial and had higher engagement. This finding resonates with previous research that suggested motivation increased

student engagement with AWE feedback in the revision process (Z. V. Zhang, 2017; Z. (Victor) Zhang, 2020; Z. (Victor) Zhang & Hyland, 2018).

It is worth noting that chosen texts could influence students' engagement, which was voiced by participants in this study. The students distinguished creative writing from academic writing and expressed their concern that Grammarly would be more helpful in correcting academic writing. Another variable regarding text is in which revision stage AWE feedback comes into play. This was proposed by Li J and Chen ZT. Chen ZT attributed her attitude towards Grammarly to the fact that she had already revised her text on higher-level concerns; thus, she only expected Grammarly to correct low-level content issues. Li J asked whether her response would be the same if her article had not gone through three stages of revision already.

## VI. Conclusion

### 5.1. Conclusion

To conclude, Grammarly largely provided language-related feedback (68%) and was more accurate in language-related issues than in content-level concerns. Its *punctuation* feedback had the highest accuracy rate; few mistakes were found in the *grammar* category. Despite being less accurate in content-level concerns, 32% of *clarity* feedback, 25% of *delivery* suggestions, and 17% of *engagement* flaggings were accurate. In general, accurate feedback had a higher acceptance rate, and all participants were selective in accepting indeterminate feedback, though great variation can be seen. While the accuracy of the feedback suggested a correlation to students' feedback response, learners' internal factors seemed to play a critical role. On the whole, students who had a high acceptance rate tended to do so across all feedback categories and feedback accuracy rates.

The study showed mixed evidence of potential language development in using Grammarly. Overall, all participants made surface-level changes to their texts and found Grammarly useful in these errors. In addition, some high-engagement students made content-level modifications.

Chen ZT and Li J both demonstrated learning effects and recognized Grammarly's potential in facilitating learning. All participants felt positive about Grammarly's feedback and admitted they would consider using it in the future.

This study further exposed the dynamic and complex nature of students' engagement with automated written feedback. Language proficiency level did not appear to be the factor that determines students' engagement. The issue of trust was found to play a crucial role but did not present a linear correlation with engagement. Other factors such as learners' beliefs, linguistic confidence, and motivation turned out to interact with trust/distrust of the tool and influenced students' responses. The two high-engagement students in this study shared a moderate and selective trust in Grammarly and were aware of its strength and weakness. They both appeared to be confident with their linguistic knowledge while allowing themselves to be challenged by Grammarly suggestions and looking for additional information. Additionally, unlike the low-engagement students, they regarded correcting grammatical errors as important in improving their texts.

## **5.2. Implications**

This study brings several implications. Second language educators interested in using AWE feedback to foster learning should be aware of the factors influencing students' engagement. Basic instructions, such as how to use the different view options of AWE and the nature of AWE feedback, could guide the student's attitude in a direction that promotes moderate trust and critical thinking of AWE feedback. Furthermore, discussing the importance of correcting language-related errors and the potential ways of learning in this process could direct learners' beliefs to encourage engagement.

Second language learners could make a better judgment with AWE feedback by learning about the strength and weaknesses of the tool and forming discriminated trust on different kinds of feedback.

Grammarly could also adjust to support learning. To begin with, it could better inform the *Learn more* option to the users. It could also be more transparent about its accuracy rate in feedback and advise the users of its strength and weaknesses. Currently, Grammarly is a commercial product that targets the general population; as more and more second language writers use the tool, perhaps Grammarly could provide a learning mode for second language learners.

### **5.3. Limitations and Future Recommendations**

The participants in this study were recruited in a way that students who were interested in improving their texts and were confident in their pieces were more likely to participate. The texts used in this study lacked variety due to the research design. Though the sample size was bigger than that of the previous studies, it was still a considerably small sample. In addition, due to the research design, the distribution of language proficiency levels of the participants was not even. Further research could use a larger and more diverse sample size, as well as a variety of texts. Researchers could also investigate how students' engagement with AWE differs in different revision stages.

## VII. References

- Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners' perceptions and preferences of written corrective feedback: a case study of university students from Mainland China. *Asian-Pacific Journal of Second and Foreign Language Education*, 1(1), 1–17. <https://doi.org/10.1186/S40862-016-0010-Y/TABLES/10>
- Corpuz, V. A. F. (2011). *Error correction in second language writing : teachers' beliefs, practices, and students' preferences*.
- Cortazzi, M., & Jin, L. (1996). English teaching and learning in China. *Language Teaching*, 29(2), 61–80. <https://doi.org/10.1017/S0261444800008351>
- Dai, F. (2010). English-language creative writing in mainland China. *World Englishes*, 29(4), 546–556. <https://doi.org/10.1111/J.1467-971X.2010.01681.X>
- Gao J. (2007). *Teaching Writing in Chinese Universities: Finding an Eclectic Approach - Asian EFL Journal : Asian EFL Journal*. <https://www.asian-efl-journal.com/monthly-editions-new/teaching-writing-in-chinese-universities-finding-an-eclectic-approach/>
- Grammarly. (n.d.). *How to Ensure Your Writing Is Concise and Clear | Grammarly Spotlight*. 2020. Retrieved May 13, 2022, from <https://www.grammarly.com/blog/grammarly-conciseness-suggestions/>
- Grammarly. (2019, December 27). *How Grammarly Can Make Your Writing More Engaging | Grammarly*. <https://www.grammarly.com/blog/engagement-dimension/>
- Grammarly. (2021, March 9). *A History of Innovation at Grammarly | Grammarly Blog*. <https://www.grammarly.com/blog/grammarly-12-year-history/>
- Hegelheimer, V., & Ranalli, J. (2020). Call for papers for a special issue on Automated Writing Evaluation: Impacting Classrooms, Supporting Learners. *Language Teaching & Technology*, 24(2), 35–36. <https://doi.org/10.1191/1362168806lr190oa>
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83–101. <https://doi.org/10.1017/S0261444806003399>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44. <https://doi.org/10.1016/j.asw.2020.100450>
- Lai, Y. H. (2010). Which do students prefer to evaluate their essays: Peers or computer program. *British Journal of Educational Technology*, 41(3), 432–454. <https://doi.org/10.1111/J.1467-8535.2009.00959.X>
- Leki, I. (1991). The Preferences of ESL Students for Error Correction in College-Level Writing Classes. *Foreign Language Annals*, 24(3), 203–218. <https://doi.org/10.1111/J.1944-9720.1991.TB00464.X>
- Leng, K. T. P. (2014). An Analysis of Written Feedback on ESL Students' Writing. *Procedia - Social and Behavioral Sciences*, 123, 389–397. <https://doi.org/10.1016/J.SBSPRO.2014.01.1437>
- Lu, X. (2019). An Empirical Study on the Artificial Intelligence Writing Evaluation System in China CET. *Https://Home.Liebertpub.Com/Big*, 7(2), 121–129. <https://doi.org/10.1089/BIG.2018.0151>
- Lv, X., Ren, W., & Xie, Y. (2021). The Effects of Online Feedback on ESL/EFL Writing: A Meta-Analysis. *Asia-Pacific Education Researcher*, 30(6), 643–653. <https://doi.org/10.1007/S40299-021-00594-6/TABLES/2>
- MA, J. J. (2019). L2 Students' Engagement with Written Corrective Feedback. *The TESOL Encyclopedia of English Language Teaching*, 1–7. <https://doi.org/10.1002/9781118784235.EELT0966>
- Mo, H. (2012). A Study of the Teaching of ESL Writing in Colleges in China. *International Journal of English Linguistics*, 2(1). <https://doi.org/10.5539/ijel.v2n1p118>

- Moon, D., & Prof, A. (2021). Evaluating Corrective Feedback Generated by an AI-Powered Online Grammar Checker. *International Journal of Internet, Broadcasting and Communication*, 13(4), 22–29. <https://doi.org/10.7236/IJIBC.2021.13.4.22>
- Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing*, 52. <https://doi.org/10.1016/j.jslw.2021.100816>
- Sheen, Y., & Ellis, R. (2005). *Handbook of Research in Second Language Teaching and Learning*.
- Stevenson, M., & Phakiti, A. (2019). Automated feedback and second language writing. In *Feedback in Second Language Writing: Contexts and Issues* (pp. 125–142). Cambridge University Press. <https://doi.org/10.1017/9781108635547.009>
- Swan, Michael., & Smith, B. (2001). *Learner English : a teacher's guide to interference and other problems*. Cambridge University Press.
- Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System*, 91, 102247. <https://doi.org/10.1016/J.SYSTEM.2020.102247>
- Wu, W., Huang, J., Han, C., & Zhang, J. (2022). Evaluating peer feedback as a reliable and valid complementary aid to teacher feedback in EFL writing classrooms: A feedback giver perspective. *Studies in Educational Evaluation*, 73, 101140. <https://doi.org/10.1016/J.STUEDUC.2022.101140>
- Yu, S., Jiang, L., & Zhou, N. (2020). Investigating what feedback practices contribute to students' writing motivation and engagement in Chinese EFL context: A large scale study. *Assessing Writing*, 44, 100451. <https://doi.org/10.1016/J.ASW.2020.100451>
- Zhang, C., Yan, X., & Liu, X. (2015). The development of EFL writing instruction and research in China: An update from the International Conference on English Language Teaching. *Journal of Second Language Writing*, 30, 14–18. <https://doi.org/10.1016/j.jslw.2015.06.009>
- Zhang, Z. V. (2017). Student engagement with computer-generated feedback: A case study. *ELT Journal*, 71(3), 317–328. <https://doi.org/10.1093/elt/ccw089>
- Zhang, Z. (Victor). (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, 43. <https://doi.org/10.1016/j.asw.2019.100439>
- Zhang, Z. (Victor), & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>
- Zhang, Z. (Victor), & Hyland, K. (2022a). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, 51. <https://doi.org/10.1016/j.asw.2021.100586>
- Zhang, Z. (Victor), & Hyland, K. (2022b). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, 51. <https://doi.org/10.1016/J.ASW.2021.100586>
- Zhao, H. (2018). Exploring tertiary English as a Foreign Language writing tutors' perceptions of the appropriateness of peer assessment for writing. <https://doi.org/10.1080/02602938.2018.1434610>, 43(7), 1133–1145. <https://doi.org/10.1080/02602938.2018.1434610>
- Zheng, Y., & Yu, S. (2018a). Student engagement with teacher written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing*, 37, 13–24. <https://doi.org/10.1016/j.asw.2018.03.001>
- Zheng, Y., & Yu, S. (2018b). Student engagement with teacher written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing*, 37, 13–24. <https://doi.org/10.1016/J.ASW.2018.03.001>



## Appendix A <sup>1</sup>

### Stimulated recall script and questions

#### A.1. Stimulated recall script

We are going to watch the video of your error correction process (editing process with *Grammarly*). As we watch the video, I will be asking you questions about what you were thinking. As you watch your error correction process with *Grammarly*, try to recall what you were thinking at the time of error correction. Try to put your mind back into the task. Anytime you remember something, say it, interrupt me, ask me to stop the video if you want.

I am interested in finding out what you were thinking when you were correcting each error identified by *Grammarly* and why you accepted/rejected/ignored *Grammarly* feedback. It does not matter at all to me if those thoughts were silly or profound. I will audio-record our conversation so I do not have to divide my attention by taking notes. At the end of our stimulated recall, I will ask you a few questions about your opinion regarding *Grammarly*.

I am going to put the computer mouse on the table here and you can pause the video any time you want. So, if you want to tell me something about what you were thinking, you can click on the mouse to pause the video. If I have a question about what you were thinking, then I will click on the mouse to pause and ask you to talk about that part of the video. Is everything clear? Are you ready? Let's get started!

#### A.2. Stimulated recall guiding questions

- (1) What were you thinking when you saw this number of alerts/ this many highlights?
- (2) What were you thinking right then when you were reading the feedback/ when you paused after reading the feedback/ when you were correcting your error?
- (3) Why did you reject/accept/ignore feedback provided by *Grammarly*?

---

<sup>1</sup> Provided by Koltovskaia S (2020) to use for this research.

- (4) What did you think of the feedback provided by *Grammarly*?
- (5) How did you arrive at accepting/rejecting/ignoring the feedback?
- (6) Did you always understand the feedback provided by *Grammarly*? Why or Why not?

## Appendix B <sup>2</sup>

### Semi-structured retrospective interview questions.

- (1) Is this is your first time using *Grammarly*? If yes, what is your overall impression of *Grammarly*? If not, how long have you been using Grammarly? What do you like about it? What don't you like about it?
- (2) In general, what do you think of *Grammarly's* feedback on the errors you made?
- (3) Were you satisfied with the feedback provided? Why or why not?
- (4) How do you think *Grammarly* helped you produce text with fewer errors?
- (5) To what extent did *Grammarly* help you understand why you made errors?
- (6) Do you think *Grammarly's* feedback is similar to human's feedback? Why or why not?
- (7) Can you tell me a little bit about your proofreading/editing strategies? Did your strategy change when you used *Grammarly*? How much time do you usually spend on proofreading your paper? Did this time change with *Grammarly*? Why or why not?
- (8) Will you consider using *Grammarly* in the future? Why or why not?
- (9) What do you think of the usability of *Grammarly*? Did you encounter any problems when using *Grammarly*? Can you identify the strengths and weaknesses of *Grammarly* and its feedback?
- (10) Is there anything else you have noticed about *Grammarly* that you would like to say?

---

<sup>2</sup> Provided by Koltovskaia S (2020) to use for this research.

## Appendix C

### A Sample of Teacher Feedback

#### 0/10 Originality

The writing approaches its subject or idea in a novel or unique way or incorporates unique plot elements or details. This is a score that marks how much the piece stands out from other peer writing.

#### 20/20 Experience

The piece uses narrative writing filled with concrete descriptions of an event or experience. The writing absorbs the reader in the event and conveys to them what it was like for the author to have experienced it.

#### 10/10 Accuracy

This is a holistic score to indicate a writing's technical aspects such as grammar, vocabulary, and sentence accuracy.

#### 5/5 Thesis

The piece has a concept, idea, or thesis it wants to convey to the reader. Through the piece's narrative structure, a larger idea is explored through both concrete and abstract discussion.

#### 5/5 Graders Discretion

This is the most subjective part of the rubric and indicates the grader's overall impression of the paper's success (execution) as a piece of writing. This is independent of the specific assignment's objective.

#### Total

50/50

Notes: It's your best piece yet. It's dramatic, hilarious, absurd, and highly expressionistic. It is a brief look into the inanities of modern life in China. So much is explored in this highly focused and descriptive piece that comments on social class, wanhonity, and consumer capitalism. It is my favorite piece I've read this year, and there have been some good ones.

Years ago when I started teaching, I didn't imagine students could write like this. A lot of that had to do with the fact that my assignments were much more typical of an ESL writing course. In the years since, I have moved away from that and more towards a communicative approach to writing that mirrors the actual world of professional writing. Sometimes I worry it's not as helpful to many students who will go on to teach high school students gaokao writing. When I get an essay like this though, I feel it is all worth it. I'm proud to have served as an editor of this piece. You perfectly satisfy the assignment requirements, but the piece goes much further than being a high-scoring essay. I hope you have come to realize you have a lot of talent in writing and I hope you will always consider writing a potential tool you have for your goals in life. The world is desperately lacking in Chinese voices that can write in English. Please find every opportunity you can to write. You are in a unique position as a person fluent in multiple languages in a world where communication seems to be eroding quickly. It's been a pleasure to work with you on all your pieces. Please keep in touch.

Goodluck.

## Appendix D

### A Sample of Grammarly Feedback

#### A sample of Grammarly feedback

1

"Let's go to Sanlitun after the last exam!" Ling came to me and invited.

"Um ... No."

"Come on, we'll be free then!"

"Not really, I still have two essays to go."

"Nobody is writing essay right after her last exam. There is a fantastic bakery opened recently and I tell you what, it sells Wellington croissants. Didn't you say that you want to have the Wellington steak of Pizza Hut after the exam? The croissant costs only half of the price and its steak is thicker!"

"..."

"I'm going back home the day after the exam and you are not spending the last time with me." She started to sob in an annoying way.

"OK, fine. I'll go with you."

"Sweet! I'll send you the official account. You can have a look first."

As we arrive at Sanlitun at 4:20 pm after the exam, the first thing I see is not the bakery with its iconic "&" sign, but the unrealistically long line.

"We are not wasting our time here, don't we?" I ask Ling as I am about to head to somewhere else.

"It's gonna be fast. And don't worry, I'm with you!"

The line is divided into three parts with each part allowing a limited amount of people to wait there, which costs us some time to find the real end of it. On our way to find the right place to start, I don't see many young and fashionable faces as I expected. In fact, there are so many grannies with their sisters, and some even have with them a folding stool. Apart from them are some middle-age men in helmets and jackets, saying "Meituan errands—delivery in 1 hour." And the next second I hear one of

▲ Formatting tools are not available. 1,777 words +

#### 126 All suggestions

✦ Accept 20 suggestions at once

is writing an essay

~~stand still~~ standstill

secrets from other others

before I come came

Creamorous, who which

Accept all 20 Dismiss

- invited · Correct pronoun usage
- , I · Correct the comma splice
- essay · Add an article
- opened · Correct pronoun usage
- and · Add a comma
- of · Change preposition
- and · Add a comma
- and · Add a comma
- in an annoying way · Change the wording
- OK · Use consistent spelling

#### HIDE ASSISTANT >>

**58**  
Overall score  
See performance >

Goals  
4 of 4 set >

All suggestions

Correctness  
93 alerts

Clarity  
Mostly clear

Engagement  
Very engaging

Delivery   
Just right

Get Expert Writing Help

Plagiarism

#### A sample of Grammarly feedback

Meituan errands—delivery in 1 hour." And the next second I hear one of them comforting his Eleme friend, "I waited for 4 hours yesterday."

Right after we stand still, a handsome clerk of the bakery approaches us saying, "Hi ladies, you have around two and a half hours to wait. We are sorry about that."

Seeing my hesitation, he adds enthusiastically, "You're lucky. It's the best time of the day to come. Guests coming at other times wait for far more longer than this."

"It's Okay, thank you!" Ling laughs. It seems that she kind of buys the rhetoric. "What I'm afraid of is that there wouldn't be a line." Ling whispers to me as if she is keeping some big secrets from other.

"What do you mean?"

"You see, the bakery has opened for a few days and there are still people queuing here. Doesn't it say something about the bread?"

"Let's see, then."

Standing in the line, I constantly see grannies holding a shiny green bag passing me. It is this bag I saw before I come that went viral on Little Red Book. It's said that the bakery Butterful & Creamorous, who calls itself an entry-level luxury brand, deliberately takes away the handle of their packing bag and makes its icon big and bold, so that young and beautiful customers can always hold it in arms and display the logo. Talking about the packing bags, right now there are some middle-aged men clumsily pinching several full bags with their fingers and walking around.

"I was wondering maybe we can buy directly from them." I ask Ling.

"I won't do that if I were you." says a primary school boy standing behind me voluntarily.

"I remember a post saying that he spent extra 200 kuai to have another man queue for him, I don't mind paying them a few dozen kuai for their

▲ Formatting tools are not available. 1,777 words +

#### 126 All suggestions

✦ OK · Use consistent spelling

- sign · Remove the comma
- to · Change preposition
- with · Add a comma
- real · Choose a different word
- The line is divided into t... · Rephrase sentence
- don't · Wrong verb tense
- expected · Wrong verb tense
- middle-age · Correct your spelling
- saying · Add a comma
- I · Add a comma
- stand still · Correct your spelling
- clerk of the bakery · Change the wording
- saying · Add a comma
- more · Change the word

#### HIDE ASSISTANT >>

**58**  
Overall score  
See performance >

Goals  
4 of 4 set >

All suggestions

Correctness  
93 alerts

Clarity  
Mostly clear

Engagement  
Very engaging

Delivery   
Just right

Get Expert Writing Help

Plagiarism