

RESEARCH

Open Access



# Effects of teacher, automated, and combined feedback on syntactic complexity in EFL students' writing

Nang Kham Thi<sup>1\*</sup> and Marianne Nikolov<sup>2</sup>

\*Correspondence:  
[nang.kham.thi@edu.u-szeged.hu](mailto:nang.kham.thi@edu.u-szeged.hu)

<sup>1</sup> University of Szeged, 32-34.  
Petőfi Sgt., Szeged 6722,  
Hungary

<sup>2</sup> University of Pécs, 6. Ifjúság  
Street., Pécs 7624, Hungary

## Abstract

Although studies on written feedback have confirmed the effectiveness of multiple sources of feedback in promoting learners' accuracy, much remains to be discovered about its impact on other aspects of language development. Concerns were raised with regard to the possible unfavourable impact of feedback on the complexity of students' writing which resulted from their attention to producing accurate texts. In response to this need for research, the study investigated the effects of teacher, automated, and combined feedback on students' syntactic complexity over a 13-week semester. Our data collection included 270 students' texts including their drafts and revised texts and pre-and post-test writing. Essays were analysed using the web-based interface of the L2 Syntactic Complexity Analyzer. Regardless of feedback from multiple sources, paired sample *t*-test results indicate no significant differences between initial and revised texts, resulting in minimal variance between comparison pairs. Moreover, no significant differences were found between the pre-and post-writing assessment in all complexity measures. These findings suggest that providing feedback on students' writing does not lead them to write less structurally complex texts. The syntactic complexity of their revised essays varied among high-, mid-, and low-achieving students. These variations could be attributed to proficiency levels, writing prompts, genre differences, and feedback sources. A discussion of pedagogical implications is provided.

**Keywords:** Written corrective feedback, Automated feedback, Syntactic complexity, Learner corpus analysis, Second language writing

## Introduction

The importance of written corrective feedback (WCF) has been recognized by educators and scholars as a means to improve student writing. Feedback on second language (L2) writing involves various responses to learner output ranging from attempts to rectify errors in writing (e.g., grammatical errors; Kang & Han, 2015) to written commentary on content and rhetorical concerns (Goldstein, 2004). Existing research, however, has primarily focused on the impact and benefits of WCF rather than on that of written commentary (Pearson, 2022). Particularly, studies in WCF research suggest that providing feedback is beneficial to improve students' grammatical accuracy (e.g., Bonilla Lopez

et al., 2018; Zhang, 2021). Although WCF has been the subject of debate in the literature (Truscott, 1996, 2007), writing instructors consider it to be a useful pedagogical practice for improving the writing performance. Also, students' willingness to receive feedback from teachers (Lee, 2008) and their positive attitudes toward feedback (McMartin-Miller, 2014) help teachers continue to provide feedback on L2 writing.

Given that students nowadays receive feedback from multiple sources, previous studies have investigated which feedback sources were more beneficial for student writing. Findings suggested large discrepancies among different feedback sources (e.g., teacher and automated feedback) and these variations relate to feedback areas, strategies, and accuracy of feedback (e.g., Dikli & Bleyle, 2014; Niu et al., 2021; Thi & Nikolov, 2021; Thi et al., 2022). Noting the affordances and constraints of teacher and automated feedback, recent studies recommended using automated feedback as an assistance tool to complement the traditional ways of providing feedback (Dikli & Bleyle, 2014; O'Neill & Russell, 2019; Thi & Nikolov, 2021). Within this research paradigm, the majority of studies compared students' initial drafts and revised texts and examined how they made use of feedback in their revisions through measuring the uptake rates or accuracy gains. Findings from such studies shed light on the complementation of multiple feedback sources (Ranalli, 2018) while exploiting the strengths of automated feedback.

In studies examining the effect of WCF, the primary aim is to develop accuracy with little consideration of the fact that an increase in accuracy might come at the cost of syntactic complexity. Specifically, although research on the impact of WCF on accuracy development has demonstrated that feedback on student writing is conducive to their accuracy development, analysing accuracy without regard for other dimensions of writing (e.g., complexity) would be meaningless. For example, as Truscott (1996, 2007) asserted, students' fear of making mistakes may actually lead them to limit the complexity of their writing. Also, Polio (2012a) argued that studies on error correction emphasized the importance of feedback on accuracy development, but a likely tendency is that "attention to accuracy could help their accuracy but harm the fluency or the complexity" (p. 147). In other words, the attention of L2 writers to accuracy is likely to divert their attention from other aspects of writing. Therefore, Polio (2012b) suggested that it would be beneficial for WCF studies to examine how feedback affects other aspects of language development, such as complexity and fluency.

Against this backdrop, the present study examined the influence of multiple feedback sources on EFL students' writing complexity and studied the effect of students' levels of proficiency (high-, mid-, and low- performing students) on the changes in their syntactic complexity during the course. The findings of this study are expected to have implications for the research and pedagogy of L2 writing. Theoretically, the findings, derived from 270 written texts, will add to the existing research on WCF, where the handful of studies addressed the impact of WCF on syntactic complexity (e.g., Hamano-bunce, 2022). From pedagogical perspectives, investigating students' syntactic complexity might help teachers gain a better understanding of which aspects of syntactic complexity could or could not be affected by feedback. Moreover, such awareness might provide some indication of whether feedback on L2 writing leads students to produce structurally less complex writing as a result of attempting to improve their linguistic accuracy.

### **Written corrective feedback as a means for accuracy development**

The effectiveness of WCF on student writing continues to be debated, however, several studies have examined the facilitative role of teacher feedback in students' writing development (Benson & DeKeyser, 2018; Frear & Chiu, 2015; Nicolas-Conesa et al., 2019; Shintani & Ellis, 2015). WCF, regardless of the feedback type, resulted in accuracy improvement in the majority of studies. Nicolas-Conesa et al., (2019) compared direct feedback with indirect feedback on students' writing and found that both conditions improved accuracy. Researchers found similar results in Benson and DeKeyser (2018) and Shintani and Ellis (2015): students who received direct or metalinguistic feedback performed better than those who did not.

Other studies examined how automated feedback helps to improve the quality of writing (Bai & Hu, 2016; El Ebyary & Windeatt, 2010; Huang & Renandya, 2020; Ranalli, 2018; Stevenson & Phakiti, 2014, 2019). Unlike studies on teacher feedback, those on automated feedback demonstrated conflicting findings. For example, Ranalli (2018) found that automated feedback is beneficial to student writing. Also, Bai and Hu (2016) noted that automated feedback can supplement teacher feedback in EFL writing classrooms while decreasing teachers' feedback burden (Ranalli, 2018). Despite this, Stevenson and Phakiti (2014) uncovered little evidence that automated feedback improves the quality of writing or that the effects of such feedback can be transferred to improvements in overall writing proficiency. Furthermore, Huang and Renandya (2020) found that integrating automated feedback did not necessarily lead to improvements in students' revisions.

### **Syntactic complexity as a complex construct**

L2 writing scholars commonly agree that complexity, accuracy, and fluency (CAF) measures best capture students' language development (e.g., Barrot & Gabinete, 2021; Housen et al., 2012; Skehan, 2009). As Barrot and Gabinete (2021) posited, complexity is characterized as "the ability to produce more advanced language", accuracy as "the ability to avoid errors in performances", and fluency as "the ability to produce written words and other structural units in a given time" (pp.1–2). These traits of language development are assessed to investigate the effects of instruction and individual differences (Housen et al., 2012). The CAF measures play significant roles in students' writing development, but the study focused solely on syntactic complexity, thus we will not examine that issue further.

As mentioned previously, syntactic complexity focuses on the sophistication of syntactic features that an L2 learner produces and the range or variety of those features (Ortega, 2003). Therefore, the assessment of syntactic complexity requires manual analysis to calculate the production units including phrases, clauses, and sentences. Though earlier studies employed a limited number of syntactic complexity measures (i.e., only two to five) (see Ortega, 2003), the use of online computational tools render the evaluation of syntactic complexity possible while overcoming the constraints of a labour-intensive nature of manual analysis (Petchprasert, 2021). As a result, recent studies have utilized automated tools to evaluate the syntactic complexity of students' writing including Coh-Metrix, and L2 Syntactic Complexity Analyzer (L2SCA).

### Written corrective feedback and its effects on syntactic complexity

Limited studies in WCF research examined whether and how the provision of WCF influences students' syntactic complexity in writing (Eckstein & Bell, 2021; Eckstein et al., 2020; Hartshorn & Evans, 2015; Van Beuningen et al., 2012; Xu & Zhang, 2021) (Table 1). Generally, findings from such studies have remained inconclusive: some studies (e.g., Van Beuningen et al., 2012) found that WCF did not cause students to produce structures that were linguistically simplified, whereas others (e.g., Hartshorn & Evans, 2015) stressed an adverse effect on writing complexity. As Van Beuningen et al. (2012)

**Table 1** Summary of empirical studies

Authors	Participants & context	Source of feedback	Complexity measure(s)	Key findings
Eckstein et al. (2020)	23 international graduate students; US	Teacher	Mean length of sentence, mean length of T-unit, and complex nominals per clause from L2SCA	Timely feedback supports syntactic complexity development
Eckstein and Bell (2021)	63 international FYC students; US	Teacher	Mean length of T-unit, ratio per T-unit of clauses, coordinate phrases and complex nominals from L2SCA	Students with dynamic WCF demonstrated significantly less syntactic complexity over time
Evans et al. (2011)	30 undergraduate students; US	Teacher	Mean length of T-unit	Negligible effect of dynamic WCF on syntactic complexity
Fazilatfar et al. (2014)	30 advanced students from an English institute; Yazd	Teacher	Mean length of sentence and a dependent clause ratio from L2SCA	Unfocused WCF led to gains in syntactic complexity
Hartshorn et al. (2010)	47 ESL students; US	Teacher	Mean length of T-unit	Dynamic WCF had slight unfavourable effect on writing complexity
Hartshorn and Evans (2015)	27 learners; US	Teacher	Mean length of T-unit Clause to T-unit ratio	No meaningful difference between the control and treatment groups in terms of two complexity indices
Li et al. (2020)	66 non-English major freshmen; China	AWE (Pigai)	All indices from L2SCA	Students' syntactic competences developed in some aspects of syntactic complexity
Van Beuningen et al. (2012)	268 students from secondary education; Dutch	Teacher	Number of subordinate clauses per clause	Positive effect of comprehensive error correction on students' structural complexity
Xu and Zhang (2021)	65 sophomores; China	AWE (Pigai)	Number of clauses per T-unit from L2SCA	Syntactic complexity remained unchanged
Zhang and Cheng (2021)	72 English major sophomores; China	Teacher	Mean length of T-unit and Ratio of clauses per T-unit from L2SCA	Comprehensive WCF showed no effects on syntactic complexity

L2SCAL2 syntactic complexity analyzer; FYC first-year composition; AWE automated writing evaluation

found, students who received feedback demonstrated higher syntactic complexity than those from the practice group. Along the same lines, Fazilatfar et al. (2014) also indicated significant complexity gains in the experimental group when comparing their first and final compositions. These findings were later reinforced by Li et al. (2020) in which the students' syntactic competence improved in some syntactic complexity measures following the feedback from an automatic writing evaluation tool.

Conversely, other studies reported that writing complexity was largely unaffected by the provision of feedback (Eckstein & Bell, 2021; Hartshorn & Evans, 2015; Hartshorn et al., 2010; Xu & Zhang, 2021; Zhang & Cheng, 2021). Hartshorn et al. (2010) reported that ESL learners' writing complexity was negatively affected by dynamic WCF. These results closely corresponded to those reported by Evans et al. (2011), as comparing the complexity of the treatment and control groups did not show any significant differences. Building on these studies, Hartshorn and Evans (2015) conducted a 30-week study and examined the effects of feedback on complexity. Similar results were reported and thus the authors postulated that a gain in one aspect of writing (accuracy) is offset by a loss in another (complexity). The same holds for Eckstein and Bell (2021): a significant reduction in syntactic complexity was observed among students with dynamic WCF over time compared to control group. Overall findings from these studies shed light on the fact that improvements in complexity tend to be at odds with accuracy. In other words, L2 writers might produce structurally less complex writing in an attempt to improve their linguistic accuracy.

Given the conflicting results and paucity of studies that have solely focused on the effects of feedback on writing complexity, we examined the influence of feedback from multiple sources on the syntactic complexity of EFL students and the possible effect of students' proficiency levels on their changes in syntactic complexity. The following three research questions were addressed:

1. To what extent do teacher, automated, and combined feedback affect EFL students' syntactic complexity in their revised texts?
2. To what extent does the feedback provision affect EFL students' syntactic complexity over the semester?
3. What is the effect of students' levels of proficiency (high-, mid-, and low-performing students) on the changes in their syntactic complexity during the course?

## **Methodology**

### **Participants**

The study recruited 30 undergraduate students (11 males and 19 females) at a university in Myanmar. They majored in English and registered for a communicative skills module to improve their English language skills. In terms of the writing component, the students completed argumentative and narrative essays and revised them following the feedback. All participants were Burmese native speakers who began learning English as a foreign language at the age of 5. As determined by their scores on the National Matriculation Exam, the participants' level of language proficiency was low-intermediate, which corresponds to the CEFR B1 level. They were of typical university age, ranging from 17 to

18 years old. Before the study, students confirmed their willingness to participate voluntarily. They were informed about the research objectives and the data that would be collected. Moreover, they were told their anonymity would be maintained and they were able to withdraw from the study at any time. As a result of their failure to complete some essays during the intervention, three students were excluded from the study.

## Instruments and measures

### *Writing tasks*

We used six writing tasks (including pre-and post-tests) which were extracted from the curriculum prescribed by the Ministry of Education. As the tasks were based on the themes introduced in each unit of the curriculum, we reasoned that students were familiar with these topics and that they had fewer difficulties in generating ideas as they completed their task. Also, we provided some guiding prompts for each essay to elicit students' responses (Fig. 1). More specifically, these tasks required them to compose a four-paragraph guided essay (300-to 400-word essay) without separate introduction and conclusion paragraphs. The rationales for providing sub-topics were to help students generate ideas as well as to allow for an entire essay to stay on topic. Despite the fact that different writing topics can affect students' writing performance, these were free-constructed responses which were found to be valid measures in examining the efficacy of WCF on students' writing (Ellis, 2010; Li, 2010) as they enable students to produce the target language with meaningful communication.

### *Feedback treatment*

The whole program included four treatment sessions, but the feedback students received on each writing task differed. Specifically, they received teacher feedback on the first writing task (Week 4) and feedback from the free version of Grammarly on the second (Week 6). Contrary to this, combined feedback was provided on the third and fourth writing tasks (Week 8 and Week 10). Notwithstanding several differences in the aspects of writing teacher and Grammarly feedback focused on, we did not attempt to limit the scope of the feedback to reflect the feedback practices in a general English course. Particularly, teacher feedback addresses language-and content-related issues associated with

Write the essay according to its instructions. Your writing will be evaluated according to the following four criteria:

- the degree to which you have achieved the task;
- the organization and coherence of paragraphs;
- the range and accuracy of grammatical structures;
- the range and accuracy of vocabulary.

For the following essay, four prompts are given. Write about the four points in separate paragraphs. Write about 300-400 altogether.

***Learning Languages***

Write an essay on your ideas about learning languages. In the text,

- discuss the role of learning languages in people's lives;
- describe an important event where knowing languages matters;
- explain why you think people nowadays should learn more than two languages; give examples;
- discuss how knowing many languages can influence your future.

**Fig. 1** Sample writing task

idea development, supporting details, clarity of ideas, task achievement, coherence and cohesion, grammatical range and accuracy, and lexical range and accuracy. However, Grammarly feedback mainly focuses on language errors: article/determiner, preposition, and miscellaneous errors including conciseness and wordiness issues (for details, see Thi & Nikolov, 2021). Moreover, other differences were associated with how feedback was provided: the teacher used the “Track Changes” functionality of Microsoft Word and gave error feedback and comments pertaining to content-related issues in students’ writing. In contrast, Grammarly’s website allows students to upload their essays independently and receive instant feedback.

### **Syntactic complexity measures**

In our study, we used L2SCA (Lu, 2010, 2011), a free automated text analyser that can compute 14 indices of syntactic complexity. We included six measures which were previously used in the studies which looked into the effect of feedback on writing complexity (Table 1). Two of these measures tap length of production (mean length of T-unit [MLT] and mean length of sentence [MLS]), two measures reflect the degree of phrasal sophistication (complex nominals per clause [CN/C] and complex nominals per T-unit [CN/T]), and two measures gauge the amount of subordination (clauses per T-unit [C/T] and subordinate clauses per clause [DC/C]). The selection was informed by Ortega (2003) who reported that MLS, MLT, C/T, and DC/C were the most widely employed syntactic complexity measures across 21 studies of college-level L2 writing. She also noted three indices (except mean length of sentence) were the most satisfactory measures, as they were correlated linearly with programme, school, and holistic rating levels. Moreover, MLT and CN/C indices were found to be important indicators of English essay quality, as they indicated significant differences in essays written by non-native English students (Lu, 2011; Lu & Ai, 2015). The number of complex nominals per T-unit (CN/T) was also added because it was supposed to be the best indicator of writing quality and complexity in students’ writing (Eckstein & Bell, 2021). See Fig. 2 for definitions of the measures of syntactic complexity.

We used an automated approach for assessing linguistic complexity due to its free availability, speed, and reliability. Lu (2010) reported that the accuracy of structural unit identification ranged from 0.830 to 1.000 and reliability ranged from 0.834 to 1.000 when compared to hand-coding. In addition, reliability and validity were confirmed by Polio and Yoon (2018) in which a high degree of reliability was achieved with the syntactic complexity scores generated by the system, as all correlations were significant at the 0.01 level.

<b>Category</b>	<b>Code</b>	<b>Description</b>
1. Mean length of T-unit	MLT	Number of words per T-unit
2. Mean length of sentence	MLS	Number of words per sentence
3. T-unit complexity ratio	C/T	Number of clauses per T-unit
4. Dependent clause ratio	DC/C	Number of dependent clauses per clause
5. Complex nominals per clause	CN/C	Number of complex nominals per clause
6. Complex nominals per T-unit	CN/T	Number of complex nominals per T-unit

**Fig. 2** Measurement variables for syntactic complexity

### Data collection

All students completed six writing tasks over a semester (August to October 2020) including their pre-and post-tests and the writing tasks during the treatment sessions. For Weeks 1 and 2, the project was introduced, and the students were initially given a writing task (i.e., a pre-test) which was later assessed by two authors using an adapted B1 analytical rating scale (Euroexam International, 2019). From Week 3 onward, the students composed four assigned essays in Microsoft Word and submitted their draft writings to the teacher through emails on a weekly basis. Following the submission of their first drafts, students received feedback from multiple sources (i.e., teacher, Grammarly, or combined feedback). Following the feedback, the students revised their essays and resubmitted them the following week. The process of providing feedback and revision continued weekly until the 10<sup>th</sup> week when they submitted their revised essays of the fourth writing task. The post-tests took place in Week 13. In total, the complete dataset comprised 270 essays including 108 preliminary drafts and their corresponding revised texts.

### Data analysis

Statistical analysis, including descriptive statistics and paired sample *t*-tests, was used to scrutinize the influence of feedback on syntactic complexity of students' writing. Particularly, students' first drafts and revised essays were compared to examine the revision effects. Also, comparing pre- and post-tests completed on Week 3 and Week 13 allowed us to assess the effects of feedback over the semester. The analysis for RQ3 was conducted in two stages. Students were initially classified into three groups: high-, mid-, and low-performers according to their scores in the pre-test using a tripartite split (Cardelle & Corno, 1981). Here, the mean scores (i.e., of the total scores of the two assessors divided by two) were calculated. For the pre- and post-tests, the inter-rater reliability coefficients (Pearson's *r*) were 0.92 and 0.94, respectively. Following this, we compared the changes in students' writing complexity in four revised texts across the three groups.

## Results

### Effect of feedback from multiple sources on syntactic complexity of EFL students' revised texts

Table 2 presents how the students' initial and revised essays differed in terms of syntactic complexity in response to feedback from teacher, Grammarly, and combined feedback. Overall, comparing most comparison pairs revealed minimal differences, indicating no significant effects of feedback on writing complexity in their revised essays. A few exceptions to this pattern were evident, however: the revised Essays 1 and 4 showed a reduction in some complexity indices. Particularly, in Essay 1, students' decline in three T-unit measures (i.e., mean length of T-unit, T-unit complexity ratio, and complex nominals per T-unit) indicates that they applied fewer words, clauses, and complex nominals per T-units in their revised texts. For example, complex nominals per T-unit significantly decreased from initial drafts ( $M = 1.64$ ) to revised essays ( $M = 1.58$ ,  $t(26) = 2.22$ ,  $p = 0.04$ ). This was also true in the case of Essay 4 in which they showed a significant decrease in dependent clause ratio and complex nominals per T-unit. They produced shorter complex nominals per T-unit when a comparison was made between first drafts



**Table 2** Paired sample *t*-tests of syntactic complexity gains between the initial and revised essays

Index	Essay 1 (Teacher)			Essay 2 (Grammarly)			Essay 3 (Combined)			Essay 4 (Combined)		
	$\Delta M$	<i>t</i>	<i>p</i>	$\Delta M$	<i>t</i>	<i>p</i>	$\Delta M$	<i>t</i>	<i>p</i>	$\Delta M$	<i>t</i>	<i>p</i>
MLT	.73	2.86	.00*	.01	0.18	.86	-.03	-0.25	.81	.11	1.31	.20
MLS	.44	1.57	.13	.05	1.02	.32	-.08	-0.64	.53	.06	0.88	.39
C/T	.06	2.06	.04*	.00	0.36	.72	-.02	-1.56	.13	.02	1.78	.09
DC/C	.02	1.72	.09	.00	0.65	.52	-.01	-1.78	.09	.01	2.24	.03*
CN/C	.02	0.80	.43	.00	0.58	.57	-.01	-0.77	.45	.01	1.44	.16
CN/T	.11	2.22	.04*	.01	0.72	.48	-.02	-1.37	.18	.03	2.45	.02*

$\Delta M$  Mean difference between revised essay and its initial draft

\* $p \leq .05$

**Table 3** Comparisons of syntactic complexity measures in the pre-and post-tests

Index	Pre-test		Post-test		Paired sample <i>t</i> tests		
	Mean	SD	Mean	SD	<i>t</i>	df	<i>p</i>
MLT	14.63	2.69	14.64	2.61	0.01	26	0.99
MLS	15.84	3.13	16.12	3.16	0.87	26	0.38
C/T	1.62	0.26	1.64	0.25	0.50	26	0.61
DC/C	0.36	0.08	0.36	0.09	-0.38	26	0.70
CN/C	0.85	0.27	0.84	0.25	-0.22	26	0.82
CN/T	1.34	0.37	1.39	0.5	0.64	26	0.52

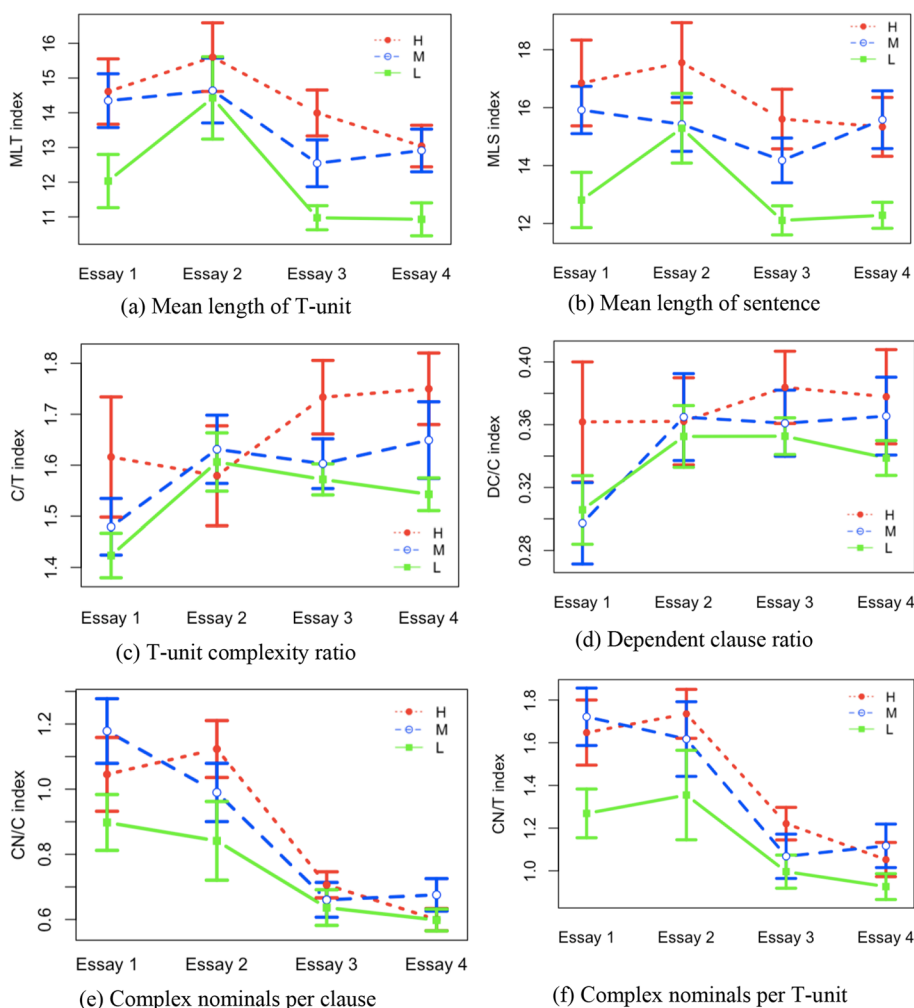
( $M = 1.07$ ) and their revised versions ( $M = 1.05$ ,  $t(26) = 2.45$ ,  $p = 0.02$ ). No other results exhibited significant differences.

**Effect of WCF on students’ syntactic complexity over the semester**

The comparison of the means of syntactic complexity between pre- and post-tests showed little variation over a semester of WCF intervention with no significant differences in the complexity measures (Table 3). Specifically, the post-tests showed increases in the means of MLT, MLS, C/T, and CN/T, but none of these gains reached statistical significance. Furthermore, the means of subordinate clauses per clause remained unchanged from pre- ( $M = 0.36$ ) to post-test ( $M = 0.36$ ). In addition to these results, the students showed a reduction in the measure of complex nominals per clause, suggesting that the students produced fewer complex nominals per clause in the post-tests as opposed to pre-tests. Based on these findings, it would be reasonable to suggest that WCF does not show any effects on students’ syntactic complexity development.

**Effect of students’ levels of proficiency (high-, mid-, and low-performing students) on the changes in their syntactic complexity**

A comparison of students’ syntactic complexity in their revised writing showed variations among the high-, mid-, and low-performing students (Fig. 3). Overall, students from all three groups exhibited progress in the indices of T-unit complexity ratio and dependent clause ratio with certain levels of decline in the remaining complexity indices. Specifically, the T-unit complexity ratio increased, but the degree of improvement



**Fig. 3** Changes in the students' writing complexity over the semester. *Note* Level of performance was defined as follows: High = scores of 11–12 on the averaged pre-test; Medium = 8–10.5; Low = 5.5–7.5

varied among the three groups. The high achievers' performance declined from Essays 1 to 2, but Essays 2, 3, and 4 showed consistent development. In contrast, the mid- and low-performing students achieved a significant improvement from Essays 1 to 2 with minor fluctuations from Essays 2 to 4. From Essay 1 to 4, we noted that the number of dependent clauses per clause increased, although there was some variation among the groups. However, a major difference between the highest achievers and the other groups concerned the degree to which they made improvements: a certain level of development was found between Essays 1 and 2 among the mid- and low-performers, whereas slow and steady growth was observed in the dependent clause ratio of the highest-achievers throughout the course.

Unlike these two indices, all remaining syntactic complexity measures exhibited a decline during the course. Although the mean length of T-unit increased from Essays 1 to 2, the results showed a reduction from Essays 2 to 4 regardless of their levels of proficiency. This was not the case with the mean length of sentences where mid-performing students experienced a gradual reduction from Essays 1 to 3 which was followed

by a noticeable improvement in Essay 4. As for the high-and low-performers, the mean length of sentences fluctuated dramatically from Essays 1 to 3 which was levelled off or stabilized in Essay 4. As for the indices of complex nominals per clause and T-unit, the results showed reduction among all groups throughout the course. The only exception included the high-achieving students, who showed increases in these two indices from Essays 1 to 2.

## Discussion

The study investigated how teacher, automated, and combined feedback influenced EFL students' syntactic complexity in their revisions and over the semester. In addition, the potential effect of students' levels of proficiency on the changes in their syntactic complexity during the course was also examined. We discuss our findings in light of previous research that studied the impact of feedback on students' writing complexity. Overall, this study demonstrates that writing complexity was unaffected by feedback, as reflected in the minimal variations between the students' initial drafts and revised texts. Similarly, no significant differences were found in the syntactic complexity of students' writing between pre- and post-tests. Our findings concur with that of Evans et al. (2011) and Zhang and Cheng (2021) who found that the provision of WCF did not enhance students' syntactic complexity, with that of Hartshorn and Evans (2015) who discovered no meaningful differences between the treatment and control groups for the measures of syntactic complexity, and with that of Xu and Zhang (2021) who contended that students' syntactic complexity remained unchanged following the automated feedback.

Looking at the T-unit complexity measures (MLT, C/T, and CN/T), all three indices showed a pattern of reduction in the students' revisions after the provision of teacher feedback. This finding corresponded to the previous studies in which the students with dynamic WCF exhibited a decrease in MLT, C/T, and CN/T from pre- to post-tests (Eckstein & Bell, 2021; Hartshorn et al., 2010). One explanation outlined by previous research is that students' attempts to improve accuracy may hinder the development of their syntactic complexity (Eckstein et al., 2020; Hartshorn et al., 2010). In Eckstein et al.'s (2020) study, it was argued that L2 writers might employ linguistically simplified structures to improve their linguistic accuracy. Similarly, Hartshorn et al. (2010) explained that, as students strive to improve their accuracy, complexity may be inhibited slightly by the careful monitoring of their writing.

Given that WCF does not support the development of syntactic complexity, we studied the degree of students' feedback acceptance, as we reasoned that their unsuccessful utilisation of feedback could be an underlying reason for the non-significant impact of feedback on syntactic complexity. However, this was not the case in our study. We found that students utilised feedback effectively in their revisions, resulting in 71.0% (teacher feedback), 76.2% (Grammarly feedback), and 61.8% (combined feedback) of correct revision. Moreover, they made notable improvement in their writing performance when the comparison was made from pre- ( $M=8.98$ ,  $SD=2.09$ ) to post-writing assessment ( $M=10.46$ ,  $SD=1.96$ ,  $p=0.006$ ) (for details, see Thi & Nikolov, 2021). These findings reflect how students utilised feedback in their revision and the general impact of feedback after a semester-long feedback treatment.

Although no significant improvements were found in students' syntactic complexity following the WCF intervention over a semester, this finding shed light on the fact that WCF did not result in structurally less complex writing. This is an important observation, as Truscott (2007) reasoned that WCF has a negative influence on syntactic complexity by causing them to avoid complex structures with the fear of making mistakes. Also, Polio (2012b) contended that students may ignore complexity in pursuit of accuracy. To put in another way, as teachers provide feedback on L2 writing as a means of improving writing accuracy, students are likely to fasten their attentional focus on how to rectify grammatical errors and produce more accurate revisions, probably resulting in structurally less complex writing. However, the findings from our study contradicted these assertions established in previous studies. Rather, our findings are partially in agreement with those of previous studies (e.g., Hartshorn et al., 2010; Van Beuningen et al., 2012; Zhang & Cheng, 2021) which found that WCF did not lead to participants using less complex syntactic structures. As Zhang and Cheng (2021) explicitly stated, the result that WCF does not show any effects on students' syntactic complexity does not support the contention that it negatively affects syntactic complexity, as asserted by Truscott (1996, 2007). Taken together, it would be reasonable to conclude that WCF did not negatively affect students' writing complexity despite not resulting in complexity gains.

The above results have implications for L2 writing and pedagogy. Based on the findings that students did not exhibit any significant improvements in most measures of syntactic complexity, WCF has a negligible effect on students' writing complexity. Understanding these negligible effects of WCF might inform writing teachers that students' focus on producing accurate texts would not deviate their attention from complexity. This reassures L2 writing teachers that gains in one aspect of writing (i.e., accuracy) do not tend to be conflicted with another aspect of writing development. Thus, it is advisable for teachers to continue their feedback practices in L2 writing classrooms, as providing feedback does not lead students to produce less complex writing.

However, these findings might be mediated by some feedback-related and task-related factors such as feedback sources, topic familiarity, and/ or genres of writing. When students received Grammarly feedback on their writing (Essay 2), there was no difference between their original and revised writing in the indices of C/T, DC/C, and CN/C. A likely explanation is that feedback scope of Grammarly feedback is limited to accuracy issues which in turn limits students' attention to the complexity of their writing. Also, non-significant differences in the draft and revised texts provide some indication of the potential influence of feedback sources on the complexity measures. Other points of discussion concern the influence of topic familiarity on syntactic complexity. As Abdi Tabari and Wang (2022) suggested, topic familiarity had a positive effect on syntactic complexity in students' writing. In light of this, the authors conclude that L2 learners tend to deploy their subject-matter knowledge quickly when dealing with a familiar writing task and focus more on generating ideas and producing structurally more complex texts. In our study, though the writing tasks were taken from the curriculum, these tasks were somewhat different from one another regarding the degree of topic familiarity. For example, the writing topic "The best teacher who inspired me" would probably be more familiar to the students compared to the topic "The worst teacher who discouraged me";

as they had experience in writing about a person they admire in their secondary schools. However, we could not draw any valid conclusions about how topic familiarity supports syntactic complexity, as it is beyond the scope of our study.

While not central to the purpose of the study, we reasoned that it is important to consider the impact of genre differences on syntactic complexity, as different genres tend to have different communicative and functional requirements which might result in different linguistic features (Lu, 2011; Yoon & Polio, 2017). In our study, students completed argumentative essays (Essays 1 and 2) and narrative essays (Essays 3 and 4) during the treatment sessions. Appendix 1 provides a visual representation of the syntactic complexity in students' writing across two genres of writing. Overall, the findings show higher complexity measures in argumentative texts than in narrative essays for all measures, except the two subordination measures. Our results were similar to those in Yoon and Polio's (2017) who found that students' language was more complex in argumentative essays than in narrative essays based on length of production units and phrase-level complexity measures. Interestingly, the study found little genre effect on subordination measures, which is also true in our study.

Research is yet to be conducted on how multiple feedback sources affect students' syntactic complexity in writing, where different groups receive feedback from different sources (e.g., teacher or peer). This may reveal how multiple feedback sources affect syntactic complexity over time. Moreover, future research that investigates the impact of WCF on syntactic complexity, especially a study using multiple drafts would yield useful insights into how feedback affects the sub-constructs of syntactic complexity in multiple rounds of feedback.

## Conclusion

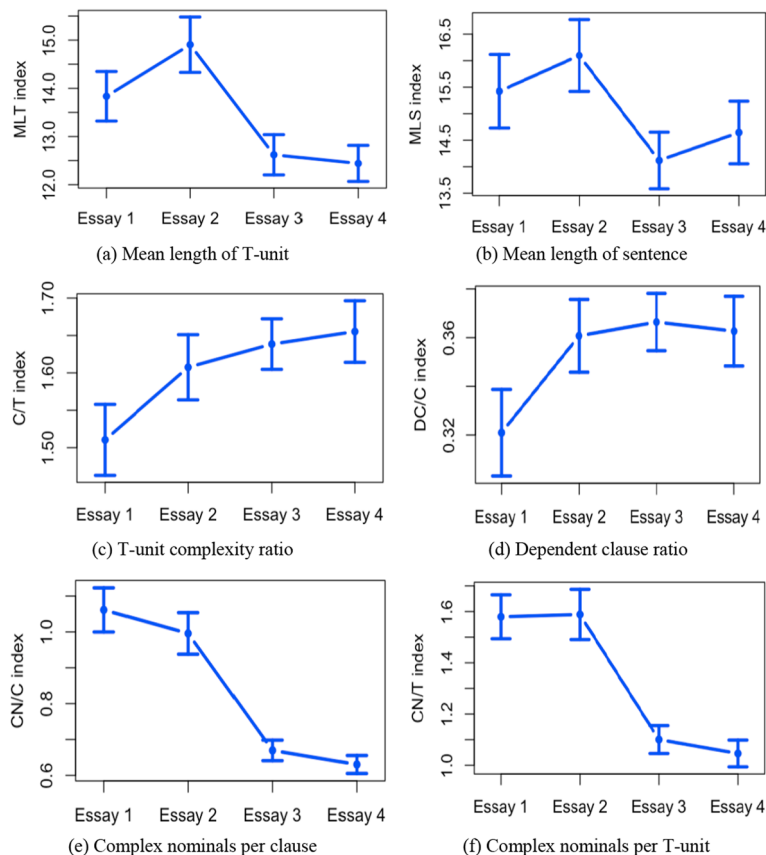
This study examined the effects of teacher, automated, and combined feedback on syntactic complexity in EFL students' writing. Overall findings revealed limited changes in the comparison pairs between the initial and revised texts and no significant differences between the pre- and post-tests. Specifically, length of production unit indices (MLT and MLS), C/T, and CN/T increased in the post-tests, but these improvements did not reach statistical significance. We also discussed the variations among the high-, mid-, and low-performing students when the comparison of their syntactic complexity was made in their revised writing throughout the course. The results suggest that the mere exposure to feedback would not be sufficient to enhance the complexity of students' writing. Therefore, future research should take a more interventionist approach in which students are exposed to texts with higher syntactic features (e.g., model texts) and explore the impact on syntactic complexity.

Though the study provided insights into syntactic complexity and feedback, some limitations must be acknowledged. Language development requires a longer observation period compared with that in the study. The findings may indicate a limit to the benefits of WCF on several subcomponents of syntactic complexity. Therefore, the need emerges for longitudinal research that investigates how WCF affects syntactic complexity in the long run. Future research could examine patterns of differences of students with high and low proficiency levels to provide a clearer picture of the impact of WCF on the construct. Moreover, previous research selected a limited range of complexity measures (typically one to four). However, assessing different aspects of complexity would be beneficial for capturing a comprehensive picture of L2 writing development (or lack thereof) due to the multi-dimensional nature of the construct.

Despite the small sample size, the study contributed important findings to the research on L2 writing and the practice of L2 writing pedagogy. Analysis of 270 texts added to the growing body of research on WCF that focuses on syntactic complexity and informed scholars in this field on considering different aspects of writing in assessing learners' development. Moreover, the current findings provide pedagogical implications that contribute to the understanding of the influence of providing WCF on the writing complexity of students over the semester. From writing assessment perspectives, the study stressed the role of automated tools in assessing EFL learners' writing development. For L2 writing teachers, a computational system for automatic analysis of syntactic complexity could facilitate the comparison of linguistic complexity of writing samples, assessing changes in the linguistic complexity of writing after a particular pedagogical intervention, or monitoring students' linguistic development over a certain period. In the same manner, using computational tools can help L2 writing researchers understand the holistic aspects of the syntactic development of students with varying proficiency levels and evaluate the effectiveness of pedagogical interventions that aim to promote syntactic complexity development.

**Appendix 1**

See Fig. 4.



**Fig. 4** Syntactic complexity in students' writing across two different genres. Note Students completed argumentative writing in Essays 1 and 2, and narrative writing in Essays 3 and 4

### Acknowledgements

We would like to thank the teacher and students for their voluntary participation in this project. We are grateful for the helpful comments provided by the anonymous reviewers and the editor's guidance. Moreover, we are also thankful for a doctoral scholarship from the Stipendium Hungaricum scholarship programme and the University of Szeged, which enabled the first author to pursue her PhD in Hungary.

### Author contributions

NKT designed and performed the experiments; analysed and interpreted the data; and was a major contributor in writing the manuscript. MN supervised and wrote the paper—review & editing. All authors read and approved the final manuscript.

### Funding

Open access funding provided by University of Szeged with grant number: 5900.

### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### Declarations

#### Consent to participate

All participants voluntarily took part in this project and provided written informed consent prior to participating.

#### Competing interests

The author(s) have stated no potential conflict of interest.

Received: 3 October 2022 Accepted: 1 December 2022

Published online: 01 March 2023

### References

- Abdi Tabari, M., & Wang, Y. (2022). Assessing linguistic complexity features in L2 writing: Understanding effects of topic familiarity and strategic planning within the realm of task readiness. *Assessing Writing*, 52, 100605. <https://doi.org/10.1016/j.asw.2022.100605>
- Bai, L., & Hu, G. (2016). In the face of fallible AWE feedback: How do students respond? *Educational Psychology*, 37(1), 67–81. <https://doi.org/10.1080/01443410.2016.1223275>
- Barrot, J., & Gabinete, M. K. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. *International Review of Applied Linguistics in Language Teaching*, 59(2), 209–232. <https://doi.org/10.1515/iral-2017-0012>
- Benson, S., & DeKeyser, R. (2018). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*, 23(6), 702–726. <https://doi.org/10.1177/1362168818770921>
- Bonilla Lopez, M., Van Steendam, E., Speelman, D., & Buyse, K. (2018). The differential effects of comprehensive feedback forms in the second language writing class. *Language Learning*, 68(3), 813–850. <https://doi.org/10.1111/lang.12295>
- Cardelle, M., & Corno, L. (1981). Effects on second language learning of variations in written feedback on homework assignments. *TESOL Quarterly*, 15(3), 251–261. <https://doi.org/10.2307/3586751>
- Dikli, S., & Bley, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing*, 22, 1–17. <https://doi.org/10.1016/j.asw.2014.03.006>
- Eckstein, G., & Bell, L. (2021). Dynamic written corrective feedback in first-year composition: Accuracy and lexical and syntactic complexity. *RELC Journal*. <https://doi.org/10.1177/00336882211061624>
- Eckstein, G., Sims, M., & Rohm, L. (2020). Dynamic written corrective feedback among graduate students: The effects of feedback timing. *TESL Canada Journal*, 37(2), 78–102. <https://doi.org/10.18806/tesl.v37i2.1339>
- El Ebyary, K., & Wendeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, 10(2), 121–142. <https://doi.org/10.6018/ijes/2010/2/119231>
- Ellis, R. (2010). EPILOGUE: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, 32, 335–349. <https://doi.org/10.1017/S0272263109990544>
- Euroexam International. (2019). *Euroexam Detailed Specifications*.
- Evans, N. W., Hartshorn, J. K., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System*, 39(2), 229–239. <https://doi.org/10.1016/j.system.2011.04.012>
- Fazilatfar, A. M., Fallah, N., Hamavandi, M., & Rostamian, M. (2014). The effect of unfocused written corrective feedback on syntactic and lexical complexity of L2 writing. *Procedia-Social and Behavioral Sciences*, 98, 482–488. <https://doi.org/10.1016/j.sbspro.2014.03.443>
- Frear, D., & Chiu, Y. H. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System*, 53, 24–34. <https://doi.org/10.1016/j.system.2015.06.006>
- Goldstein, L. M. (2004). Questions and answers about teacher written commentary and student revision: Teachers and students working together. *Journal of Second Language Writing*, 13, 63–80. <https://doi.org/10.1016/j.jslw.2004.04.006>
- Hamano-bunce, D. (2022). The effects of direct written corrective feedback and comparator texts on the complexity and accuracy of revisions and new pieces of writing. *Language Teaching Research*. <https://doi.org/10.1177/13621688221127643>
- Hartshorn, K. J., & Evans, N. W. (2015). The effects of dynamic written corrective feedback: A 30-week study. *Journal of Response to Writing*, 1(2), 6–34.

- Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, 44(1), 84–109. <https://doi.org/10.5054/tq.2010.213781>
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). John Benjamins.
- Huang, S., & Renandya, W. A. (2020). Exploring the integration of automated feedback among lower-proficiency EFL learners. *Innovation in Language Learning and Teaching*, 14(1), 15–26. <https://doi.org/10.1080/17501229.2018.1471083>
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, 99(1), 1–18. <https://doi.org/10.1111/modl.12189>
- Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17(3), 144–164. <https://doi.org/10.1016/j.jslw.2007.12.001>
- Li, W., Lu, Z., & Liu, Q. (2020). Syntactic complexity development in college students' essay writing based on AWE. In K.-M. Frederiksen, S. Larsen, L. Bradley, & S. Thouesny (Eds.), *CALL for widening participation: Short papers from EUROCALL 2020* (pp. 190–194). Research-publishing.net. <https://doi.org/10.14705/rpnet2020.48.1187>
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, 29, 16–27. <https://doi.org/10.1016/j.jslw.2015.06.003>
- McMartin-Miller, C. (2014). How much feedback is enough?: Instructor practices and student attitudes toward error treatment in second language writing. *Assessing Writing*, 19, 24–35. <https://doi.org/10.1016/j.asw.2013.11.003>
- Nicolas-Conesa, F., Manchon, R. M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new texts: Looking into feedback for accuracy and feedback for acquisition. *The Modern Language Journal*, 103(4), 848–873. <https://doi.org/10.1111/modl.12592>
- Niu, R., Shan, P., & You, X. (2021). Complementation of multiple sources of feedback in EFL learners' writing. *Assessing Writing*, 49(January), 100549. <https://doi.org/10.1016/j.asw.2021.100549>
- O'Neill, R., & Russell, A. M. T. (2019). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology*, 35(1), 42–56. <https://doi.org/10.14742/ajet.3795>
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518.
- Pearson, W. S. (2022). Response to written commentary in preparation for high-stakes second language writing assessment. *Asian-Pacific Journal of Second and Foreign Language Education*. <https://doi.org/10.1186/s40862-022-00145-6>
- Petchprasert, A. (2021). Utilizing an automated tool analysis to evaluate EFL students' writing performances. *Asian-Pacific Journal of Second and Foreign Language Education*, 6(1), 1–16. <https://doi.org/10.1186/s40862-020-00107-w>
- Polio, C. (2012a). How to research second language writing. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 139–157). <https://doi.org/10.1002/9781444347340>
- Polio, C. (2012b). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21(4), 375–389. <https://doi.org/10.1016/j.jslw.2012.09.004>
- Polio, C., & Yoon, H. J. (2018). The reliability and validity of automated tools for examining variation in syntactic complexity across genres. *International Journal of Applied Linguistics (United Kingdom)*, 28(1), 165–188. <https://doi.org/10.1111/ijal.12200>
- Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning*, 31(7), 653–674. <https://doi.org/10.1080/09588221.2018.1428994>
- Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System*, 49, 110–119. <https://doi.org/10.1016/j.system.2015.01.006>
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532. <https://doi.org/10.1093/applin/amp047>
- Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19, 51–65. <https://doi.org/10.1016/j.asw.2013.11.007>
- Stevenson, M., & Phakiti, A. (2019). Automated feedback and second language writing. *Feedback in Second Language Writing*. <https://doi.org/10.1017/9781108635547.009>
- Thi, N. K., & Nikolov, M. (2021). How teacher and grammarly feedback complement one another in Myanmar EFL students' writing. *Asia-Pacific Education Researcher*, 31(6), 767–779. <https://doi.org/10.1007/s40299-021-00625-2>
- Thi, N. K., Nikolov, M., & Simon, K. (2022). Higher-proficiency students' engagement with and uptake of teacher and Grammarly feedback in an EFL writing course. *Innovation in Language Learning and Teaching*. <https://doi.org/10.1080/17501229.2022.2122476>
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327–369.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16(4), 255–272. <https://doi.org/10.1016/j.jslw.2007.06.003>
- Van Beuningen, C., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1–41. <https://doi.org/10.1111/j.1467-9922.2011.00674.x>
- Xu, J., & Zhang, S. (2021). Understanding AWE feedback and English writing of learners with different proficiency levels in an EFL classroom: a sociocultural perspective. *Asia-Pacific Education Researcher*. <https://doi.org/10.1007/s40299-021-00577-7>
- Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51(2), 275–301. <https://doi.org/10.1002/tesq.296>



Zhang, L. J., & Cheng, X. (2021). Examining the effects of comprehensive written corrective feedback on L2 EAP students' linguistic performance: A mixed-methods study. *Journal of English for Academic Purposes*, 54, 101043. <https://doi.org/10.1016/j.jeap.2021.101043>

Zhang, T. (2021). The effect of highly focused versus mid-focused written corrective feedback on EFL learners' explicit and implicit knowledge development. *System*, 99, 102493. <https://doi.org/10.1016/j.system.2021.102493>

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---