

Exploring the impact of Grammarly on EFL students' writing

John Maurice Gayed, Tokyo Institute of Technology

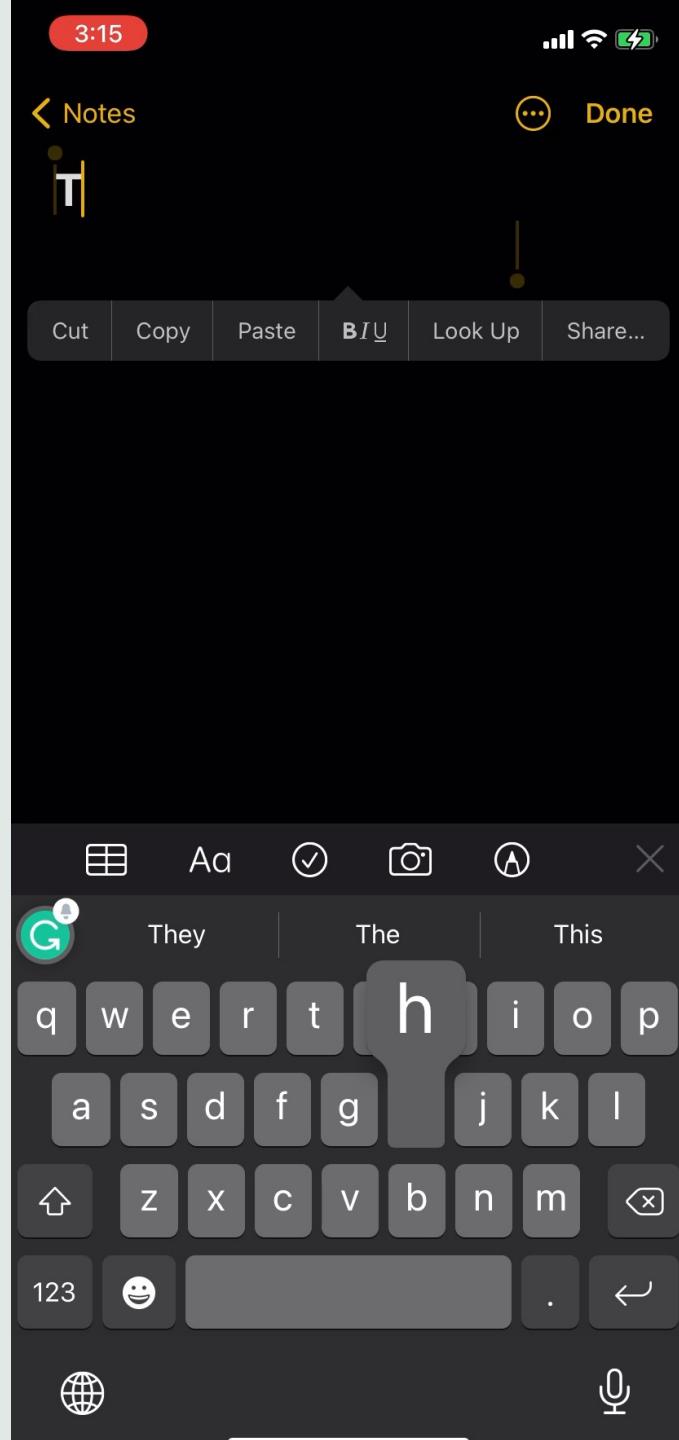
Gilbert Dizon, Himeji Dokkyo University

JALTCALL 2021



What is Grammarly?

- An intelligent writing assistant that incorporates predictive text and auto-correction features within its mobile keyboard app.
- Grammarly is also offered as a browser extension, web application, and Office 365 plug-in, but they lack the predictive text feature.



Literature Review – MALL Writing

- Li and Hegelheimer (2013) found that use of a web-based mobile application led to increased performance on a grammar test, more self-editing, and a reduction of errors among ESL learners
- Lee (2020) examined the difference between smartphone- and handwritten-based writing fluency and found that handwritten submissions were longer in length and sophistication
- Heil et al. (2016) identified significant limitations in the use of MALL for improving L2 learning, including an inability to adapt to the individual learner and limited corrective feedback.

Literature Review – AWE

- Automated writing evaluation (AWE) has been shown to significantly impact grammatical accuracy in L2 writing (Liao, 2016; Wang et al., 2013)
- Studies have shown that AWE also has an influence on overall writing quality (Cheng, 2017; Cotos, 2011; Li et al., 2015)
- Accuracy of feedback may be an issue (Ranalli et al., 2017)

Literature Review – Grammar Checkers

- Chen (2009) found that NTNU statistical grammar checker was superior to Microsoft ESL Assistant
- Chodorow et al. (2007) reported that their grammar checker could detect preposition errors produced by L2 English students at a 0.8 precision rate with a recall rate (percentage of valid errors detected) of 0.3
- Gomon et al. (2009) found Microsoft Research ESL Assistant was able to correctly recall 37% of article errors, 27% of the noun-related errors, and 18% of preposition errors
- John and Woll (2020) concluded that neither Grammarly nor Virtual Writing Assistant can replace teacher feedback at their current stage of development based on their level of coverage

Research Question

RSQ: Is there a significant difference between EFL students who used Grammarly and those who did not use the writing aid as it pertains to:

- grammatical accuracy?
- lexical richness?
- semantic complexity?
- writing fluency?

Methodology – Design

Counter-balanced research design – 8 weeks; 248 writing samples

	Class A & C	Class B & D
Week 1	health (Grammarly)	university life (control)
Week 2	winter vacation (Grammarly)	work (control)
Week 3	hometown (Grammarly)	holidays (control)
Week 4	travel (Grammarly)	hobbies (control)
Week 5	university life (control)	health (Grammarly)
Week 6	work (control)	winter vacation (Grammarly)
Week 7	holidays (control)	hometown (Grammarly)
Week 8	hobbies (control)	travel (Grammarly)

Methodology – Participants

N = 31 Japanese university students

Participants' Eiken test levels with CEFR equivalent

	# of students	Grade	Eiken level	CEFR level
Class A	5	1 st -year	4 & 5	A1
Class B	8	1 st -year	3	A1
Class C	10	2 nd -year	Pre-2	A2
Class D	8	2 nd -year	Pre-2	A2

Methodology – Data Analysis

- Human raters → • 1 measure
 - grammatical accuracy
- Software analysis → • 3 measures
 - lexical richness, writing fluency, semantic complexity

Methodology – Data Analysis – Human

Inter-rater agreement was 90.8%.

Error types identified by human assessment	
verb tense	When I was a kid I chase dragonflies in Kochi.
subject-verb agreement	We loves Korea.
verb omission	I learn many countries culture in there. And I (had) free time in the night.
singular/plural noun	But this university force student to study something we don't want to study, or don't need.
possessive noun	I can go to back parents home during winter vacation.
incorrect article	The festival is held in my hometown. It is very interesting.
article omission	I have (a) part-time job in LAWSON.
unnecessary article	But the winter vacation ends shortly.

Methodology – Data Analysis – Software

LFP (BNC-COCA 1-25k):

<https://www.lextutor.ca/vp/comp/>

Note: Proper nouns and Japanese words were removed prior to lexical analysis

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token (%)
K-1 :	55 (67.9)	58 (65.91)	71 (67.0)	67.0
K-2 :	12 (14.8)	12 (13.64)	13 (12.3)	79.3
K-3 :	9 (11.1)	10 (11.36)	14 (13.2)	92.5
K-4 :	1 (1.2)	1 (1.14)	1 (0.9)	93.4
K-5 :	1 (1.2)	1 (1.14)	1 (0.9)	94.3
K-6 :	1 (1.2)	1 (1.14)	1 (0.9)	95.2
Coverage 95 [?]				
K-7 :	1 (1.2)	1 (1.14)	1 (0.9)	96.1
K-8 :				
K-9 :	1 (1.2)	1 (1.14)	1 (0.9)	97.0
K-10 :				

Methodology – Data Analysis – Software

Semantic complexity:

<https://aihaiyang.com/software/l2sca/single/>

Haiyang Ai

Login | Register

Home LCA L2SCA

Web-based L2 Syntactic Complexity Analyzer – Single Mode

The single mode of the web-based L2 Syntactical Complexity Analyzer takes up to 2 samples of English text and generates both numeric and graphical results of any or all 14 indices covering (1) length of production units, (2) amounts of coordination, (3) amounts of subordination, and (4) phrasal sophistication and overall sentence complexity. Please note that each text should have a **maximum of 1000 words**. If you have multiple files to be analyzed, please use the **Batch Mode**. By accessing and using the Lexical Complexity Analyzer, you are acknowledging that you agree to be legally bound and to abide by the **L2SCA Terms of Service**. Please cite:

- Lu, Xiaofei (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4):474–496.
- Lu, Xiaofei (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers's language development. *TESOL Quarterly*, 45(1):36–62.
- Ai, Haiyang & Lu, Xiaofei (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In Ana Díaz-Negrillo, Nicolas Ballier, and Paul Thompson (eds.), *Automatic Treatment and Analysis of Learner Corpus Data*, pp. 249–264. Amsterdam/Philadelphia: John Benjamins.
- Lu, Xiaofei & Ai, Haiyang. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, 29, 16–27.

Step 1: Enter text #1

INSTRUCTIONS: Type or paste your text here and click the yellow SUBMIT_window button. VocabProfile will tell you how many words the text contains from frequency bands as determined by analysing research corpora. For a demonstration, enter this text, or one of the sample texts below. TEXT SET-UP General: Include an empty space after every comma or full stop. Research: Deal with spelling errors and proper nouns. SIZE LIMITS: Web form input is currently max about 25,000 words - use UPLOAD method below for larger files (must be ~.txt; send in straight from your own hard drive). Text is NOT stored on Web VP's server

Methodology – Data Analysis – Software

Lexical richness:

<https://aihaiyang.com/software/lca/single/>

Haiyang Ai [Login | Register](#)

Home LCA L2SCA

Web-based Lexical complexity analyzer – Single Mode

The Single Mode of the web-based Lexical Complexity Analyzer takes an English text as input and computes 25 indices of lexical complexity of the text. You may choose to see the results of any or all of the 25 indices, and the system will create a graphical representation to visualize the results. Additionally, you may enter another text in order to compare their lexical complexity. Please note that each text should have a minimum of 50 words and a maximum of 10,000 words. If you have multiple files to be analyzed, please use the [Batch Mode](#). By accessing and using the Lexical Complexity Analyzer, you are acknowledging that you agree to be legally bound and to abide by the [LCA Terms of Service](#). If you intend to publish a paper that used the web-based interface to the LCA software, please cite:

- Ai, Haiyang and Lu, Xiaofei (2010). *A web-based system for automatic measurement of lexical complexity*. Paper presented at the 27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10). Amherst, MA. June 8–12.
- Lu, Xiaofei (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Language Journal*, 96(2):190–208.

Step 1: Enter text #1

Results – Grammatical Accuracy

- Errors/ total words written

Grammatical accuracy results			
	M	SD	Total
Non-Grammarly	0.06	0.04	453.5*
Grammarly	0.04	0.03	310

- $t(60) = 2.32, p = 0.02$; effect size = 0.56

Positive impact on the students' ability to write with appropriate grammar in the L2.

Results – Lexical Richness

Lexical richness results

	M	SD	1K		2K		3K		4K		5K		Off-list	
			#	%	#	%	#	%	#	%	#	%	#	%
Non-Grammarly	0.15	0.03	6912	91.5	331	4.4	97	1.3	54	0.7	35	0.5	81	1.07
Grammarly	0.18	0.04	6800	90.1	335	4.4	78	1.0	65	0.9	84	1.1	120	1.59

- $t(60) = 3.83$, $p = 0.0003$; effect size = 0.84

Positive effect on lexical richness

Results – Writing fluency

Writing fluency results

	Writing fluency		
	M	SD	Total
Non-Grammarly	243.68	97.69	7554
Grammarly	245.48	97.75	7610

- $t(60) = 0.07, p = 0.94$

No marked difference in writing fluency between the two writing conditions

Results – Semantic Complexity

Semantic complexity results			
	Semantic complexity		
	M	SD	Total
Non-Grammarly	1.31	0.45	1091
Grammarly	1.23	0.24	1134

- $t(60) = 0.07, p = 0.94$

No marked difference in semantic complexity between the two writing conditions

Results – Student Writing Sample

Student 1, Sample A (Non-Grammarly) – Topic: Hobbies

My hobby is playing (the) guitar. Because I like listening to music. When I *am* (an) elementary school student, I *listening* to rock music. So I like to play (the) guitar. I *join* a band club. In the club, I play *a* guitar and sing *a song*. But I can't sing *a song* well. I want to sing very well. But this is too difficult for me. Holiday I practice *a* guitar (a) long time. It's a chance to practice *a* guitar. But I don't do *a* voice training. Because my family *get* angry to me about my loud voice. So my voice training room is a Karaoke.

Writing sample details according to variables studied

Grammatical errors marked: 16

Number of words written beyond 2K word-frequency level: 3

Words written: 105

Number of clauses: 13

Results – Student Writing Sample

Student 1, Sample B (Grammarly) – Topic: Travel

I went to Okinawa for a high school trip at the end of summer. This trip was so amazing. First, we went to Tyura aquarium. That place is very huge. So we can see a lot of time there. And there are so many kinds of fish and sea creatures. I couldn't all spot. Because we *don't* have too much free time. Next, we went to *the* camp place. In there we had a bbq. And then we had some events. For example, (there was) a bingo game, a birthday party, and a Karaoke tournament. Those events were

Writing sample details according to variables studied

Grammatical errors: 3 (decrease from 16)

Number of words written beyond 2K word-frequency level: 6 (increase from 3)

Words written: 96 (decrease from 105)

Number of clauses: 14 (increase from 13)

Summary

RSQ: Is there a significant difference between EFL students who used Grammarly and those who did not use the writing aid as it pertains to:

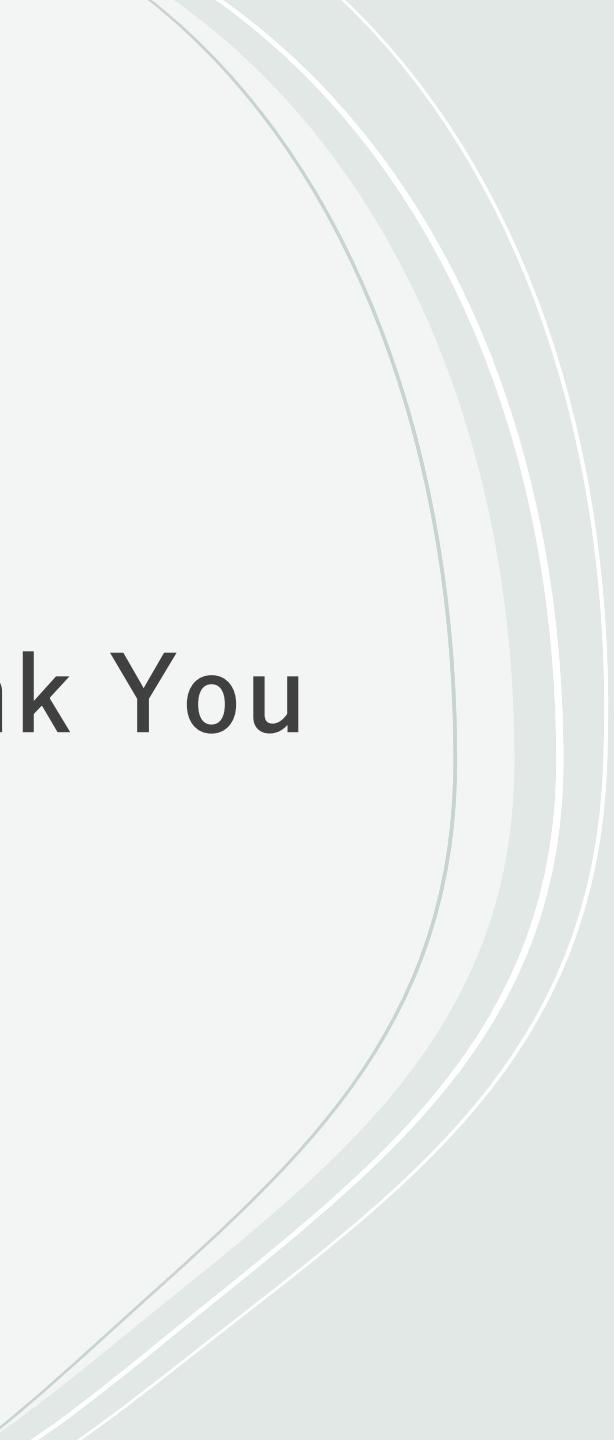
- grammatical accuracy? → Yes
- lexical richness? → Yes
- semantic complexity? → No
- writing fluency? → No

Discussion

- Findings show that intelligent corrective feedback and predictive text can ease the cognitive burden of L2 students and help them write more accurately and with greater lexical variety
- Implications
 - Recommend the use of Grammarly among beginner L2 students, especially those who struggle with grammatical accuracy
 - Grammarly could be recommended to L2 learners as a means to increase lexical variation in their writing

Limitations & Future Directions

- Limitations
 - Small sample size
 - Short intervention
 - Examined only four aspects of L2 writing
 - Not known if L2 students can become more effective English writers by using Grammarly
- Future Directions
 - Incorporate larger sample size and/or conduct a longer intervention
 - Examine overall writing quality through a holistic scale
 - Utilize pre-posttest design
 - Investigate student perceptions of Grammarly



Thank You

John M. Gayed: himejijohn@gmail.com

Gilbert Dizon: gilbert.dizon.jr@gmail.com

References

- Ai, H., & Lu, X. (2010). A web-based system for automatic measurement of lexical complexity. Paper presented at the 27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10). Amherst, MA. June 8-12
- Chen, H.-J. H. (2009). Evaluating two web-based grammar checkers – Microsoft ESL Assistant and NTNU statistical grammar checker. *Computational Linguistics and Chinese Language Processing*, 14(2), 161-180.
- Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education*, 34, 18-27. <https://doi.org/10.1016/j.iheduc.2017.04.002>
- Chodorow, M., Tetreault, J. R., & Han, N.-R. (2007). Detection of grammatical errors involving prepositions. In F. Costello, J. Kelleher and M. Volk (Eds.), *Proceedings of the 4th ACL-SIGSEM Workshop on Prepositions* (pp. 25–30). <https://doi.org/10.3115/1654629.1654635>
- Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal*, 28(2), 420-459. <https://doi.org/10.11139/cj.28.2.420-459>
- Frankenberg-Garcia, A. (2019). Combining user needs, lexicographic data and digital writing Environments. *Language Teaching*, 53(1), 29-43. <https://doi.org/10.1017/S0261444818000277>
- Gamon, M., Leacock, C., Brockett, C., Dolan, W. B., Gao, J., Belenko, D., & Klementiev, A. (2009). Using statistical techniques and web search to correct ESL errors. *CALICO Journal*, 26(3), 491–511. <https://doi.org/10.1558/cj.v26i3.491-511>
- Lee, B. (2020). Smartphone tapping vs. handwriting: A comparison of writing medium. *The EuroCALL Review*, 28(1), 15-25. <https://doi.org/10.4995/eurocall.2020.120>
- Li, Z., & Hegelheimer, V. (2013). Mobile-assisted grammar exercises: Effects on self-editing in L2 writing. *Language Learning & Technology*, 17(3), 135–156. <http://dx.doi.org/10125/44343>
- Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, 27, 1-18. <https://doi.org/10.1016/j.jslw.2014.10.004>
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 1, 8-25. <https://doi.org/10.1080/01443410.2015.1136407>
- Shang, H.-F. (2019). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments*, 1-13. <https://doi.org/10.1080/10494820.2019.1629601>