AJET | ASCILITE

Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly

Ruth O'Neill
Academic Learning Centre, School of Access Education, CQUniversity, Australia

Alex M. T. Russell
Experimental Gambling Research Laboratory, School of Health, Medical and Applied Sciences, CQUniversity, Australia

Universities are increasingly looking towards online technology to assist their students. Grammarly advertises itself as the world's most accurate online grammar checker, yet little research exists on its performance as a feedback tool in universities. This paper reports on a study investigating students' perceptions of Grammarly when used in conjunction with advice from an academic learning advisor. Using a mixed methods sequential explanatory design, the study compared one group of students' responses to the feedback they received from Grammarly (*n* = 54) with another group's responses to the traditional non-automated grammar feedback they received from the Academic Learning Centre at CQUniversity (*n* = 42). Results show that students receiving feedback from Grammarly responded more positively to 9 of the 15 survey items and were significantly more satisfied with the grammar advice that they received compared with non-Grammarly students. No significant differences were registered between cohorts or delivery mode, which suggests that Grammarly can be used effectively by academic learning advisors to provide grammar support for both international and domestic students, online and on campus. It is recommended, however, that the program be used in conjunction with academic learning advisor input as the program is currently not accurate enough for independent use to be justified.

Introduction

Many students in the Australian higher education system struggle with the demands of tertiary education because of language problems (Harris, 2016). This difficulty has been well documented with international students, who tend to exhibit more surface-level errors relating to accurate grammar, punctuation and spelling (Tynan & Johns, 2015). Traditionally, domestic students performed better than international students in relation to grammatical accuracy (Scouller, Bonanno, Smith, & Krass, 2008), but changes in the domestic student population have resulted in members of this cohort also exhibiting difficulties with sentence structure and form (Cocks & Stokes, 2013). Grammar feedback is, therefore, needed for different cohorts to achieve academic success. Responsibility for that feedback is often perceived to lie with learning support units, typically known as the Academic Learning Centre (ALC). This expectation can, however, conflict with the ALC's perception of their role, resulting in students leaving ALC sessions dissatisfied with the amount and quality of the grammar feedback they receive. Automated feedback programs have been advocated as tools to reconcile this tension. Grammarly is one such program and was chosen by CQUniversity as a potential grammar feedback tool based on its ability to reconcile contrasting theories on written corrective feedback. Building on the work of Caveleri and Dianati (2016), the primary purpose of this study was to analyse students' satisfaction with Grammarly as a feedback tool and compare this response with the manual, non-automated feedback traditionally offered at CQUniversity by the ALC. It also aimed to determine whether the program was more suitable for particular student cohorts based on language level, visa status and delivery method.

Literature review

The need for grammar and grammar input

Students need to use standard grammar and adhere to grammatical conventions to succeed in academic writing at universities (Caveleri & Dianati, 2016). Students cannot, however, improve their language skills without input. International students need external input to develop their linguistic skills (Murray, 2011), including the explicit teaching of grammar needed for academic writing (Knoch, Rouhshad, & Storch, 2014). Different groups of domestic students also need practical interventions with regards to punctuation,

grammar and spelling (McNaught & Shaw, 2016). This includes local students (defined as being born in Australia) who have not received formal grammar training as part of the school curriculum or who have accessed tertiary education through non-traditional means such as enabling programs (Cocks & Stokes, 2013). It also includes domestic non-English speaking background students (DNESB), defined by the Department of Education and Training (n.d.) as being domestic students who arrived in Australia less than 10 years before starting at university and who speak a language other than English at home.

Academic learning advisors' and students' attitudes to grammar intervention

The responsibility for grammar teaching in university settings is traditionally assigned to ALCs (Channock, D'Cruz, & Bisset, 2009) and the academic learning advisors (ALAs) who work there. ALAs at CQUniversity provide assignment feedback in online and face-to-face consultations. These consultations are 60–90 minutes, but ALAs accord a fraction of that time to grammar feedback because it is time-consuming (Dikli & Bleyle, 2014) and deemed less important than the long-term augmentation of higher order skills such as developing an argument (Winder, Kathpalia, & Koo, 2016). Students on the other hand value grammar feedback and identify immediate grammar correction as their number one concern (Huang, 2011). It was noted anecdotally that this conflict in priorities and the time constraints imposed on grammar feedback were distressing some students, because they felt they were not receiving the amount of grammar feedback they needed both for the assignment and their long-term development. They felt that if the ALC did not assist with developing their grammar, they had no other recourse, as lecturers focus on content, not the development of writing skills (Bacon & Anderson, 2004).

Automated feedback programs

Automated feedback programs could reconcile the differences between ALAs' and students' attitudes to grammar feedback, by providing students with the amount of grammar feedback they want and need without compromising ALAs' prioritisation of macro skills development. Research into online grammar checkers is limited (Caveleri & Dianati, 2016), particularly regarding students' responses to them, but studies have been undertaken on automated essay scoring systems which provide grammar feedback as part of their functionality (Ramineni & Williamson, 2012). Students' responses to the grammar feedback that these systems provide are mixed. Students using CorrectEnglish cited improvement in grammar and written accuracy as a positive facet of the program (Wang, Shang, & Briody, 2013), and those using My Editor similarly believed that it helped them to improve their grammar and spelling (Hoon, 2006). Participants in a study by Chen and Cheng (2006), however, were largely unimpressed with the grammar feedback they received. Student dissatisfaction in this study highlights the importance of the facilitator's role in the feedback process as the students' negative responses were attributed to the teachers' attitudes towards the program. One teacher was not conversant with the tool's editing functions and the other was not impressed by them, which meant that neither teacher communicated the best way for the students to use the tool.

Conflicting findings are also reported regarding the student language level that most benefits from automated feedback programs. Advanced students in Chen and Cheng's (2008) study found the focus on accuracy constraining and suggested that the automated feedback would be more useful for beginner and intermediate learners. Dikli (2010) contradicted this assertion by suggesting that automated feedback can overwhelm students with low English proficiency, and Liao (2016) agreed that automated feedback is preferable for higher performing writers because they possess the metacognitive strategies needed to develop grammatical accuracy long term. In their study on Grammarly, Caveleri and Dianati (2016) also stated that the program may benefit more-able writers because students need to have sufficient grammatical understanding to be able to filter suggestions that are incorrect, or that they deem unnecessary and not useful.

The argument as to whether automated feedback programs are more suitable for domestic or non-domestic students is similarly contentious. Hoang and Kunnan (2016) make the point that the feedback systems were not traditionally designed for non-native student writers of English and that international students do not have enough understanding of the target language to interpret and apply the suggestions the programs make. They also suggest that non-native students might lack the metalanguage to understand the suggestions provided by the programs. This argument, however, could just as easily be applied to local students who have acquired language with little to no instruction in metalanguage and rule-based principles.

A further disadvantage of automated feedback from the student perspective is the lack of human interaction (Dikli, 2010). According to Wang et al. (2013), this is a particular problem for international students and less experienced writers who need instruction, modelling and practice to develop their written English. A lack of dialogue is additionally detrimental to writing proficiency as it lowers student motivation and makes their writing more mechanical and less authentic (Wang et al., 2013). To counteract these issues and render automated feedback useful, Hoang and Kunnan (2016) suggest a combined pedagogy in which the writing instructor acts as an intermediary between the students' work and the feedback. Research suggests that feedback given in this way is most effective (Choi & Lee, 2010) because teachers can provide instruction on how best to use the program (Burston, 2008), provide clarification of errors (Hoang & Kunnan, 2016) and eliminate errors which the program has made.

Grammarly

In 2016, the ALC at CQUniversity piloted the use of Grammarly Premium, in conjunction with ALA advice. Grammarly Premium is an automated proofreading system which can identify errors related to 250 grammar rules (Grammarly Inc., 2017). Users upload their assignment and receive two scores. The first score is based on percentage accuracy, and the second is the total number of errors that the program has identified. Grammarly organises the errors according to six categories: contextual spelling, grammar, punctuation, sentence structure, style and vocabulary enhancement. The number of errors within these specific categories are counted and presented (see Figure 1).

Contextual Spelling	**10**
Misspelled Words	5
Mixed Dialects of English	4
Confused Words	1
Grammar	**17**
Determiner Use (a/an/the/this, etc.)	10
Faulty Subject-Verb Agreement	3
Incorrect Noun Number	2
Wrong or Missing Prepositions	2
Punctuation	**8**
Comma Misuse within Clauses	5
Misuse of Semicolons, Quotation Marks, etc.	2
Punctuation in Compound/Complex Sentences	1
Sentence Structure	**3**
Misplaced Words or Phrases	2
Incomplete Sentences	1
Style	**9**
Wordy Sentences	6
Improper Formatting	3
Vocabulary enhancement	✔ No errors

Figure 1. Categorised errors in Grammarly report

Feedback problems and Grammarly as the potential solution

Grammarly was chosen because of its potential to reconcile contrasting theories relating to written corrective feedback. Much has been written about the relative merits of direct and indirect feedback. Proponents of indirect feedback (Jamalinesari, Rahimi, Gowhary, & Azizifar, 2015) state that the error should only be highlighted for the student to independently review and correct, whereas advocates of direct feedback (Bitchener, 2008) maintain that explicit corrections need to be provided by the teacher. Grammarly reconciles these different theoretical approaches by providing teachers with the opportunity to use both. On the left-hand side of the screen, indirect feedback can be given using the underlined error. On the right side of the screen, direct feedback can be provided using the correction and explanation card (see Figure 2).
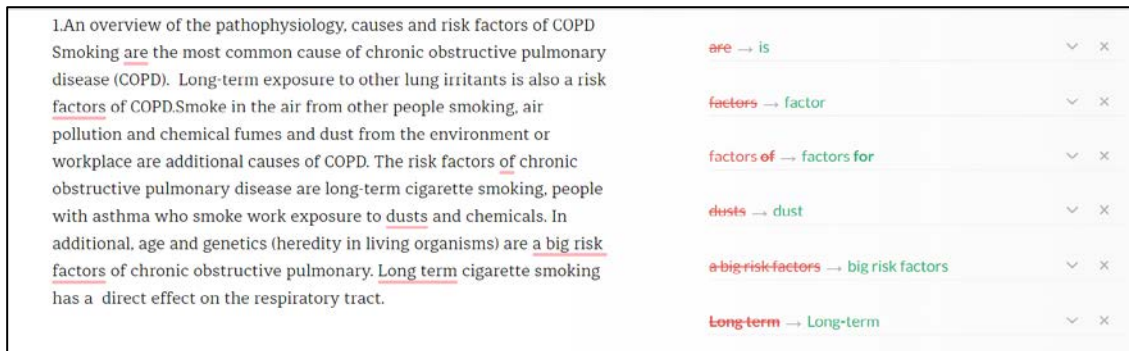
Figure 2: Direct and indirect feedback in Grammarly

Further debate exists over the amount of feedback which should be given. Students value a correct-all approach (Salehfard & Bagheri, 2014) and can become frustrated when their goals for line-by-line editing are not addressed (Linville, 2009). Research, on the other hand, has shown a need for more targeted error correction (Jamalinesari et al., 2015) as focused feedback on high-frequency errors improves students' writing (Linville, 2009). Grammarly again offers a middle ground solution, as its ability to identify 250 breaches in grammatical rules offers an all-inclusive approach, whereas the categorisation system provides the students with cues to address high-frequency errors.

An additional reason for choosing Grammarly was that little research exists on its performance within university settings. Japos (2013) conducted a study of undergraduate research students, finding that Grammarly improved their written accuracy. Qassemzadeh and Soleimani (2016) concluded that Grammarly performed better than traditional teacher intervention with regards to long-term retention of passive voice rules. Students in a study by Reis and Huijser (2016) preferred Grammarly to the alternative feedback system Marking Mate. Results from Caveleri and Dianati's (2016) investigation of students' perceptions of Grammarly within an Australian higher education setting were also largely positive, with students stating that Grammarly was useful and easy to use. They also maintained that it improved their written work and helped develop their understanding of grammatical rules.

This study builds on Caveleri and Dianati (2016) but offers essential points of difference. Firstly, students did not use Grammarly as an independent learning tool but in conjunction with an ALA. Secondly, this study compared Grammarly with the traditional ALC approach of using "Track Changes" in Microsoft Word. This study also had a larger sample size and investigated whether the software better benefits a particular student cohort based on language level, visa status and delivery method.

Research questions

(1) What are students' perceptions of Grammarly compared to perceptions of the traditional ALC grammar feedback approach?
(2) Does language level affect students' response to Grammarly?
(3) Does cohort (international, DNESB, local) affect students' response to Grammarly?
(4) Does delivery mode (face to face, online) affect students' response to Grammarly?

Methodology

Ethical clearance

Ethical clearance for this project was given by the Human Research Ethics Committee at CQUniversity on 12 February 2016.

Setting

Data was collected over one university term between February 2016 and July 2016. Six researchers collected data: five on CQUniversity's Sydney campus and one on CQUniversity's Mackay campus.

Participants

Researchers agree that a control group is integral for studies aiming to determine the efficacy of corrective feedback (Bitchener, 2008). Studies considering second language acquisition also need to be conducted under controlled experimental conditions using a control and treatment group (Ferris, 2010). As this study is driven by a consideration of best practice in corrective feedback and has a large second language cohort, an experimental group which received feedback from an ALA using Grammarly and a control group which did not use Grammarly were created.

Ninety-six students were involved in this study: 54 students were in the experimental group, and 42 were in the control group. Students in the experimental group were either enrolled in a degree at CQUniversity ($n = 48$) or preparing for university with the CQUniversity English Language Centre (ELC) ($n = 6$). Students in the control group were also either enrolled in a degree ($n = 33$) or in the ELC ($n = 9$).

Instruments

A mixed methods sequential explanatory design with follow-up variant was used (Creswell & Plano Clark, 2011). First, quantitative data was collected and analysed to reveal patterns in the students' responses to the program and identify deviations from those patterns for specific statements and cohorts. Qualitative recordings were designed to explain these patterns; to reveal why students responded in a particular way to the program. It was hoped that using both sets of data in this complementary way would provide greater insight into the program' performance and that this greater insight could be used to inform future use of the program as a feedback tool.

A student survey was, therefore, created and presented in 3 phases: personal details, quantitative responses and qualitative reactions. The survey was designed using the findings from the literature review and by adapting the research instrument employed by Caveleri and Dianati (2016). Three CQUniversity lecturers and Caveleri, co-author of the 2016 investigation into students' perceptions of Grammarly, reviewed the survey for relevance to the research questions and readability.

Personal details

The measures from the personal details phase were:

- Visa status: Students were asked to identify as local, DNESB or international. International students were subdivided into two groups: undergraduate and postgraduate students or those undertaking university preparatory courses in the ELC.
- Delivery method: Whether the students received feedback online or face-to face was recorded.
- Completion of language tests: Scores were recorded for respondents who had completed a language test such as the International English Language Test (IELTS), Pearson Test of English or Occupational English Test. Most had completed IELTS, and the scores for those who had not were converted into IELTS scores using an online test converter (Pearson Education Ltd, 2012).
- Previous assignments at the university: Students who had previously submitted an assignment were asked for the grade they had received on that assignment.

Quantitative phase

A survey using a 5-point Likert-scale measuring tool was completed by the students. Fifteen items were devised to test students' responses to the grammar feedback with scores for each statement ranging from a possible 1 (*strongly disagree*) to 5 (*strongly agree*). Items 1–4 were drawn from the literature review relating to students' beliefs about the importance of grammar feedback and their satisfaction with the grammar feedback received, both overall and in relation to the time accorded, amount provided and focus. All other statements were adapted from the survey designed by Caveleri and Dianati (2016). These aimed to determine the perceived usefulness of the grammar feedback by evaluating the ways in which the students believed the feedback was helpful to their writing and confidence in the short and long term. The ease with which the feedback could be assimilated into the writing was also considered, as was the propensity for this to happen.

Qualitative phase

Students answered three open questions:

- Q1: What did you like about the grammar feedback you received?
- Q2: What did you dislike about the grammar feedback you received?
- Q3: Any additional comments on the grammar feedback you received?

Thematic analysis was used to interpret the qualitative data. Initial themes were identified from the quantitative findings. These themes were coded and then categorised depending on whether they corroborated or refuted the quantitative findings. Additional themes, which did not arise from the statistical analyses, were then categorised from the qualitative comments using inductive analysis.

Procedure

All students had a Grammarly Premium account for the duration of the trial. Students in both the control and experimental groups received a maximum of 15 minutes grammar feedback.

Control group: face-to-face and online

ALAs identified grammar issues in the students' assignments. They discussed their findings and recorded their suggestions in the students' assignments using "Track Changes."

Experimental group: face-to-face

The ALA uploaded each student's assignment into Grammarly Premium. They then downloaded the detailed Grammarly report with feedback across the six categories: contextual spelling, grammar, punctuation, sentence structure, style and vocabulary enhancement. To make the error correction more manageable, the five most frequently occurring errors from the report were dealt with in descending order. Indirect error correction feedback was employed first with the ALA locating an underlined example of the error in the Grammarly report. To ascertain the student's current understanding of the error and encourage self- correction, the student was asked if the underlined section was accurate or not. If the student deemed the section inaccurate, they were asked to write it a different way. If an appropriate correction was made, an extra example of the error was reviewed to confirm competency with the language item. If the ALA felt that further confirmation of understanding was required, then additional error correction checks were employed using the most appropriate technique for the language point. These techniques vary; for example, concept checking questions are used to check meaning and function, visuals such as timelines are used to check verb tenses and highlighting can clarify errors in form. All techniques, however, had the same objective: to encourage the students to think about their choices, articulate their choices and not just passively accept the program's suggestions.

If the student was unsure or unable to articulate their answer, then direct feedback was employed. The student was shown the suggested correction in Grammarly and asked whether this suggestion was appropriate. If the student answered correctly, then another example of this error in the student's work was reviewed. However, if the student was unsure whether the correction was appropriate or not, they were asked to read the Grammarly information card and examples. The ALA used the explanation and examples on the card to unpack the grammar rule and illustrate its use. If the Grammarly information was not clear or detailed enough, the ALA provided additional instruction by modelling and providing examples themselves. In this way gaps in the Grammarly feedback were filled by the ALA, so the students were provided with the instruction and clarification they needed. In addition to clarifying language issues, identifying erroneous and unnecessary suggestions, this procedure also encouraged students to ask questions about the Grammarly feedback process and provided a model of use for the students to replicate.

Experimental group: online

A Grammarly Premium account was created for the student and the student's assignment uploaded. The five most frequently committed errors were identified and examples in the assignment located. Any errors that the program had made in relation to these suggestions were deleted. The assignment was uploaded into

PDF Annotator, and any errors that Grammarly had missed were highlighted with feedback from the ALA. An email was sent to the student with details of their Grammarly account, a PDF copy of their annotated assignment and advice about which errors to target.

To encourage survey completion, on campus students recorded their responses in a hardcopy survey and the data was manually input into SurveyMonkey. Online students completed the same survey in an online form in SurveyMonkey.

Analyses

The dependent variables were the level of agreement with the 15 statements about the grammar feedback. Independent variables were whether the student received feedback via Grammarly or not; their English proficiency (IELTS score and previous grades; cohort (local students, DNESB, international students, and international students from the ELC) and delivery mode (face-to-face or online). The latter variables were included to determine which specific aspects of Grammarly (or non-Grammarly) appealed to which types of students. Comparisons between those who received feedback via Grammarly and those who did not were conducted using inferential statistics, which are independent samples t-tests. For some statements, variances between the groups were significantly different, violating an assumption of independent samples t-tests. In those cases the more robust Welch t-test was used instead, indicated by degrees of freedom with decimal places. Effect sizes are reported for significant results using Cohen's *d*.

Relationships between reported agreement with statements and English proficiency, as well as previous grades, were conducted using non-parametric (Spearman's) correlations as the agreement with statements variables were skewed, and because the previous grades variable was ordinal in nature. Comparisons between students from different cohorts were conducted using ANOVA with post-hoc Tukey pairwise comparisons. Alpha was set at .05 throughout.

Results

Students' perceptions of Grammarly feedback as compared to perceptions of non-Grammarly feedback

Those who received ALA advice with Grammarly feedback were generally very positive about the experience, as indicated by the high mean in Table 1. In general, the respondents understood the importance of getting grammar feedback on assignments, believed that they had received enough assistance from Grammarly and found the suggestions useful and easy to understand. They believed that the feedback improved their assignment and grade. It also improved their confidence in the assignment and developed their confidence and language skills beyond the assignment. The only mean below 4.4 was for the statement "I made additional grammar corrections not recommended by the ALC") (*M* = 3.72, *SD* = 1.47). This suggests that although the students still agreed with the statement, there was some hesitance to do more than that recommended by the ALA. Qualitative responses largely substantiate the positive findings illustrated in Table 1, with six students stating that they would like the university to provide the software on a permanent basis, 11 students claiming that the program helped them, and 14 students describing the feedback using positive verbs and adjectives; for example, "I liked it. Really good. Often use it. If I could give 6/5, I would. Really happy."

Table 1
Mean and SD ratings for students' statements about grammar feedback

Statement	Grammarly (*n* = 54)		Non-Grammarly (*n* = 42)		Inferential statistics			
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
It was important for me to get grammar feedback on my assignment	4.93	.26	4.93	.26	-.05	94	.961	-
The advisor spent enough time on grammar feedback to help me improve my assignment	4.74	.62	4.38	.96	2.11	66.43	.039	.45
I got enough grammar feedback on my assignment	4.57	.69	4.33	.87	1.51	94	.135	-
The grammar feedback really focused on my main errors – i.e. those that were made a lot	4.54	.77	4.14	.93	2.23	79.24	.029	.47
It was easy to understand the errors because the explanations given were clear	4.65	.76	4.36	1.01	1.56	73.83	.123	-
I got a lot of useful suggestions about how to improve my grammar in the assignment	4.67	.70	4.29	.97	2.15	71.86	.035	.45
It was easy to make grammatical changes to my work using the feedback	4.65	.68	4.20	.98	2.54	67.55	.014	.53
I made the grammar corrections the ALC advisor recommended during the consultation	4.54	.86	4.38	.94	.85	94	.399	-
I made additional grammar corrections not recommended by the ALC	3.72	1.47	3.12	1.33	2.08	94	.041	.43
The grammar feedback improved my assignment	4.80	.49	4.55	.92	1.59	59.06	.117	-
The grammar feedback made me feel more confident about handing the assignment in	4.76	.58	4.38	.99	2.21	62.56	.031	.47
The grammar feedback improved my grade	4.45	.85	4.17	1.01	1.50	93	.136	-
The grammar feedback developed my language long term (not just for this assignment) as I could understand the grammatical rules more	4.48	.72	4.07	1.07	2.14	68.48	.036	.45
The grammar feedback developed my confidence in my language use long term (not just for this assignment) as I could understand the grammatical rules more	4.46	.75	3.98	1.09	2.47	69.02	.016	.51
I was satisfied with the grammar advice I received	4.75	.55	4.43	.83	2.19	68.09	.032	.45

Note: df = degrees of freedom, with decimal places indicating Welch t-tests was used. *d* refers to Cohen's *d* (effect size).

However, whilst students were largely positive about the program, they did identify six areas of concern. The main area of concern was the accuracy of the feedback, with nine students citing it as a problem: "Not all grammar errors were valid or correct." Students also stated that the pre-set to American English caused the program to notice invalid spelling errors. They identified issues with passive voice, complex sentences and vocabulary choices, which Grammarly suggested students changed based on style and not accuracy.

Students also noted that the program missed errors and that the software suggestions were at times difficult to understand. Further concerns about the software related to technical problems and included issues with uploading documents and compatibility. For two students, the program's faults clearly outweighed the positives and they explicitly expressed a preference for the traditional ALC feedback method, with one of the students writing, "I liked getting feedback via the Word document. Rather than Grammarly. I hated that."

Table 1 also provides the mean and standard deviations for the 15 statements for students who received feedback via traditional, non-Grammarly feedback mechanisms. In general, the pattern of results was similar, in that the non-Grammarly students indicated that they found the feedback useful, that they received enough, and that it assisted their grammar, grades and confidence. However, the means for Grammarly students were significantly higher for some of these statements. Grammarly students were significantly more likely to state that an ALA spent enough time on their grammar feedback to improve their assignment compared to non-Grammarly students, and that the feedback via Grammarly focused on errors that they make a lot. Grammarly students were significantly more likely to agree that they received a lot of useful feedback via Grammarly, and that it was significantly easier to make grammatical changes based on this feedback. Grammarly students reported being significantly more likely to make additional corrections that were not recommended by Grammarly, and were significantly more confident about handing the assignment in. Grammarly students agreed significantly more strongly that the feedback had developed their language, as well as their confidence in their language, outside of the assignment, compared to non-Grammarly students. In general, Grammarly students were significantly more satisfied with the grammar advice that they received compared to non-Grammarly students.

Students' perceptions of Grammarly based on language level

IELTS (and IELTS indexed) scores were captured for 35 Grammarly students and 25 non-Grammarly students, while previous grades were captured for 36 Grammarly students and 25 non-Grammarly students. For Grammarly students, those with lower IELTS scores tended to be more satisfied with the grammar advice they had received (Spearman's rho = -.41, $p < .05$), and those with higher previous grades were significantly more likely to agree that it was easy to make grammatical changes (Spearman's rho = .38, $p < .05$). These results were not statistically significant for non-Grammarly students, and no significant difference was observed between the results. For non-Grammarly students, those with lower IELTS scores were significantly more likely to agree that the grammar feedback improved their grade (Spearman's rho = -.40, $p < .05$). No other correlations were statistically significant, although it is important to acknowledge that due to missing data, the small sample and the conservative nature of Spearman's correlations, these results are most likely underpowered, and thus conclusions from these results are tentative.

Qualitative data supports the finding that graduate students with a lower IELTS score (6 or less) were satisfied with the Grammarly feedback, because students with a score of 6 or less were the group least likely to post negative comments about the program. Students from the ELC, however, who had the lowest language proficiency of all groups, were the most vocal about Grammarly's limitations, with all six students registering concerns. Four of these related to the program's inaccuracies, whilst the other two stated that it "doesn't really make sense" and that it "gives wrong suggestions." These findings suggest that Grammarly can be used with students of different language levels but a different approach may need to be taken when using it with ELC students.

Students' perceptions based on visa status

As the Grammarly and non-Grammarly students were found to differ in their agreement with some statements, differences between cohorts were conducted as a factorial ANOVA, with Grammarly (2 levels: yes vs no), Cohort (4 levels: local, DNESB, international and international ELC) as the main effects, and an interaction effect for each statement. These analyses should be regarded with caution for two reasons. Firstly, some cells were constant for some statements (i.e., all respondents selected the same answer); and secondly, cell sizes were relatively small for this type of analysis. Where significant effects were found, post-hoc Tukey pairwise comparisons were conducted.

Differences between the cohorts were observed for three statements. International students were generally significantly less satisfied with the grammar feedback they received and significantly less likely than

DNESB students to agree that the grammar feedback developed their confidence in their language long term. Local students were significantly less likely to state that the grammar feedback had developed their language long term compared to DNESB and international ELC students. These results, however, are main effects for cohort, which means that they are averaged across whether or not they received Grammarly feedback and there was no significant interaction for any of these statements, indicating that the results did not differ significantly depending on whether or not they received Grammarly feedback.

Qualitative data supports the finding that the DNESB cohort responded positively to Grammarly. Question 2 asked students to identify what they disliked about the program, and the DNESB cohort posted the fewest negative responses to this question (21% reply rate compared to 66% by both local and international students). This group was also the only one to explicitly identify that Grammarly was useful for them, with one student stating, "I think it is a great tool to assist non-English speaking background students." The high response rate to question 2 from international students, coupled with the statistical finding that they were less satisfied with the grammar feedback, suggests that, as with the ELC students, a more tailored approach may need to be taken when using Grammarly with them.

Student responses by delivery mode

Comparisons were also made between those who had received feedback online compared to on campus. Twenty-seven of the 54 Grammarly students and 13 of the 42 non-Grammarly students received feedback online, while the remainder received feedback on campus. No significant differences were found between those who had received feedback online compared to on campus for any of the statements (largest $F(1,90) = 2.58$, $p = .112$). Interaction effects were also calculated comparing differences between online vs on campus feedback for Grammarly and non-Grammarly students. No significant interaction effects were found (largest $F(1,90) = 3.32$, $p = .072$.

Discussion

The amount of time and detail accorded to grammar feedback was identified in the literature review as a potential cause of dissatisfaction with students visiting ALCs. Quantitative data showed that Grammarly students were significantly more likely to believe that enough time had been spent on grammar feedback. Qualitative data supported these findings as the issue of time was not mentioned by any of the students who received feedback using Grammarly, whereas four students who received feedback in the traditional manner bemoaned the lack of time accorded to grammar feedback. An example of this is from a student who wrote that they "would like more time spent on grammar and language."

Students who received Grammarly advice were also significantly more likely to state that they had received a lot of useful feedback. There were no responses from students in the Grammarly group complaining that they had not received enough grammar feedback. Six students in the non-Grammarly group did, however, express dissatisfaction.

> I haven't received enough grammar feedback. I want to fix my grammar but the advisor didn't pay attention to my grammar mistakes.

The tendency amongst ALAs to prioritise macro skills over grammar was flagged as a potential issue in the literature review. This issue was not referred to by any of the students in the Grammarly group, but students in the control group did believe that ALAs sacrificed grammar feedback to focus on other issues.

> We could not correct the whole assignment as to grammar issues since we should consider and correct other factors such as structure.

The shortfall of grammar feedback was identified in the literature review as a potential hindrance to language development, and some students in the control group did believe that their long-term development was inhibited.

> They just fix the grammar but I don't understand why they did it that way. It's just for the assignment not for overall English development.

Students' responses to Grammarly also identified features of effective written corrective feedback outlined in the literature review and support the proposition that the program can reconcile different theoretical approaches. Firstly, Grammarly does not just highlight one type of error as students identified a range: "repeated", "serious", "basic"; "errors that I did not know I had made" and mistakes that "I had overlooked." The issue of how much to correct was addressed, with some students expressing satisfaction that the feedback was "detailed", "thorough" and "line by line", whereas other students commented on the targeted nature of the feedback and the fact that it identified their "areas of weakness" and highlighted specific issues they had; for example, article use, sentence structure, commas, spelling, prepositions and word choices.

In addition to identifying positive feedback practices from the literature review, students' responses revealed that Grammarly adheres to other best practice elements in written corrective feedback. Students liked the fact that the feedback targeted their errors. This personalisation is a key factor in feedback assimilation, with students more likely to value feedback that takes into account their individual needs (Hyland & Hyland, 2006). Several students commented on the quickness of the program stating that the feedback was "prompt." Such timely feedback is identified by McGregor, Merchant, and Butler (2008) as crucial to its usefulness as advice is more likely to be acted on whilst fresh in the student's mind. Channock et al. (2009) further qualify the link between time and feedback by stating that students reap the most benefits if they make the changes to the assignment after the review of the first draft, which is what the students in this study did. Several students also made the link between the Grammarly feedback and their grade, which according to Bacon and Anderson (2004) is necessary to stimulate improvements in grammar and punctuation.

Acting on the literature review findings, ALAs were positioned as an intermediary between the Grammarly feedback and the student. This approach reduced negative responses about the program. Firstly, the ALAs managed students' expectations about the feedback by making them explicitly aware that it was not infallible. This understanding of the program's limitations is reflected in the following student's comment: "Some places it could not point out so it is not 100% [sic]. Ok I know it is hard and it's not perfect." Secondly, ALAs helped make the feedback more reliable by identifying and deleting erroneous suggestions and misrepresentations. Inaccuracies are acknowledged to be an issue with AFPs (Hoang & Kunnan, 2013; Liu & Kunnan, 2016) but ALAs can make them less of an issue by reducing their number. They can also use Grammarly's errors as a learning tool so that the student understands why they are inaccurate: "The Grammarly suggestion was wrong but the instructor did a pretty awesome job to clear why [sic]." Finally, and most importantly, ALAs can provide additional instruction and explanation about the students' errors. This is achieved by explaining the metalanguage used by the program, by grading explanations for lower level students, or by providing alternative examples or models. Irrespective of the technique, the outcome remains the same: a clearer explanation of the problem so that students have a better understanding of how they can improve their writing. None of the Grammarly students posted negatively about the explanations they received, whereas four non-Grammarly students stated that they needed "clearer explanations." Similarly, no Grammarly students asked for additional support, whereas non-Grammarly students mentioned a need for further clarification through additional grammar classes, sessions, materials and websites. It is therefore recommended that Grammarly continue to be used in conjunction with an ALA until comparable results are collated, proving its effectiveness as an independent tool.

Conclusion

This study shows that students using the ALC want assistance with their grammar to improve their assignments and develop their long-term writing performance. Grammarly was chosen as a means of providing the grammar feedback students want without compromising the development of their higher order skills. A systematic approach to using Grammarly was devised with supplementary ALA advice provided if students did not understand the program's response to the language item. Results between the students who received Grammarly and non-Grammarly advice were compared. Students who received grammar advice from ALAs in the traditional manner were largely satisfied with the feedback they received, but students who received feedback from Grammarly in conjunction with an ALA were more satisfied in relation to the amount of time spent, the amount of feedback received, the short-term benefits to their grade, the long-term benefits to their writing, their confidence in their writing and their ability to proofread independently. Students also identified several features of best practice in written corrective feedback in the Grammarly approach. This included the speed of response, range of errors addressed and

the targeted, personalised nature of the feedback. Concerns were, however, expressed with regards to the program's accuracy, its propensity to miss errors and correct those that were not wrong and technological glitches. These concerns would have been more prevalent had an ALA not engaged with the feedback first, as the ALA was able to manage students' expectations, reduce erroneous feedback and use the inaccuracies as a learning tool. No significant differences were identified between delivery mode, suggesting that the program can be used with online and on-campus students. Similarly, the results showed that Grammarly can be used with students of different language levels and visa status, but greater discretion or a revised approach may be needed with ELC and international students. Based on these findings, it is therefore recommended that Grammarly be added to the ALC's toolbox, but it should continue to be used in conjunction with an ALA.

Limitations to the study

Statistical power issues exist within this study because the overall sample size is small. A larger sample might find significant differences, so any future research into student perceptions' on Grammarly should use a larger sample. A number of student variables, which could have impacted on responses, were also not considered; for example, the length of time the students had spent learning English and the length of the assignment. Data connected to relational variables was also not collected. Such variables as the student-ALA relationship and the student's preferences in an ALA (e.g., gender) could have affected the students' responses. Data relating to language level used IELTS or IELTS equivalent writing scores, but these are not always a reliable indication of language performance. Asking students to self-report on these scores and university grades is also problematic as students may not report accurately and may enhance performance levels.

Future research

This study focused on students' perceptions of Grammarly. To build a more comprehensive picture of Grammarly as a feedback tool additional research is needed. Firstly, there is a need to consider ALA perceptions of the program and compare these findings with ALA attitudes to the traditional feedback approach. Secondly, as inaccuracy was cited as an issue with the program, further investigations are needed to identify the errors which Grammarly most frequently misses or misidentifies. This will help inform future feedback sessions and render them more useful.

It is also necessary to determine the degree to which Grammarly enhances students' grammatical accuracy. Short-term improvements can be identified by comparing students' grammatical accuracy in the first draft of an assignment (pre-Grammarly feedback) with their grammatical accuracy in their final draft (post-Grammarly feedback). Similar analysis should also be undertaken on the assignments of students who received feedback in the traditional manner so that it can be seen whether accuracy in the short term is more improved in students using Grammarly than in those receiving conventional ALC advice. Long-term improvement in students' grammatical accuracy could be determined through longitudinal studies in which students submit an assignment (post-Grammarly feedback) each term during their university studies. The percentage accuracy of these assignments could be compared with their first submission to determine whether improvements have been made in relation to high-frequency errors and overall written accuracy. These students could also be asked to take a grammar test each term in relation to their high-frequency errors to determine whether their accuracy using the item has improved when the program is not being used.

References

Bacon, D. R., & Anderson, E. S. (2004). Assessing and enhancing the basic writing skills of marketing students. *Business Communication Quarterly, 67*(4), 443–454. https://doi.org/10.1177/1080569904271083

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102–118. http://dx.doi.org/10.1016/j.jslw.2007.11.004

Burston, J. (2008). BonPatron: An online spelling, grammar, and expression checker. *CALICO Journal, 25*(2), 337–347. Retrieved from http://journals.equinoxpub.com/index.php/CALICO/article/view/23116

Caveleri, M., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students, *Journal of Academic Language and Learning, 10*(1), 223–236. Retrieved from http://www.journal.aall.org.au/index.php/jall/article/view/393/246

Channock, K., D'Cruz, C., & Bisset, D. (2009). Would you like grammar with that? *Journal of Academic Language and Learning, 3*(2), 1–12. Retrieved from http://www.journal.aall.org.au/index.php/jall/article/view/70/63

Chen, C. E., & Cheng, W. (2006, May). *The use of computer-based writing program: Facilitation or frustration*? Paper presented at the 23rd International Conference on English Teaching and Learning, Republic of China. Retrieved from ERIC database. (ED492078)

Chen, C. E., & Cheng, W. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language, Learning & Technology*, *12*(2), 94–112. Retrieved from http://llt.msu.edu/vol12num2/chencheng.pdf

Choi, J., & Lee, Y. (2010). The use of feedback in the ESL writing class integrating automated essay scoring (AES). In D. Gibson & B. Dodge (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 3008–3012). Chesapeake, VA: Association for the Advancement of Computing in Education. Retrieved from https://www.learntechlib.org/p/33826/

Cocks, T., & Stokes, J. (2013). Policy into practice: A case study of widening participation in Australian higher education. *Widening Participation and Lifelong Learning, 15*(1), 22–38. http://dx.doi.org/10.5456/WPLL.15.1.22

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Department of Education and Training. (n.d,). *Non-English speaking background*. Canberra: Australian Government. Retrieved from http://heimshelp.education.gov.au/sites/heimshelp/resources/glossary/pages/glossaryterm?title=Non-English%20speaking%20background

Dikli, S. (2010). The nature of automated essay feedback. *CALICO Journal, 28*(1) 99–134. Retrieved from https://calico.org/page.php?id=5

Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing, 22*, 1–17. http://dx.doi.org/10.1016/j.asw.2014.03.006

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition, 32*(2), 181–201. https://doi.org/10.1017/S0272263109990490

Grammarly Inc. (2017). *Grammarly Premium: Keep it clean and write your heart out*. Retrieved from https://www.grammarly.com/premium

Harris, A. (2016). Integrating written communication skills: Working towards a whole of course approach. *Teaching in Higher Education, 21*(3), 287–300. http://dx.doi.org/10.1080/13562517.2016.1138456

Hoang, G. T. L., & Kunnan, A. J. (2016). Automated essay evaluation for English language learners: A case study of MY Access. *Language Assessment Quarterly, 13*(4), 359–376. http://dx.doi.org/10.1080/15434303.2016.1230121

Hoon, T. B. (2006). Online automated essay assessment: Potentials for writing development. In A. Treloar & A. Ellis (Eds.), *Making a difference with Web technologies. Proceedings of AusWeb 2006, The Twelfth Australasian World Wide Web Conference* (pp. 230–234). Lismore: Southern Cross University. Retrieved November 1, 2016, from http://ausweb.scu.edu.au/aw06/papers/refereed/tan3/paper.html

Huang, L. S. (2011). Are we having the effect we want? Implementing outcomes assessment in an academic English language-support unit using a multi-component approach. *Writing Program Administration, 35*(1), 1–11. Retrieved from http://wpacouncil.org/archives/35n1/35n1huang.pdf

Hyland, K., & Hyland, F. (2006). Interpersonal aspects of response: constructing and interpreting teacher written feedback. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 206–224). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139524742.013

Jamalinesari, A., Rahimi, F., Gowhary, H., & Azizifar, A. (2015). The effects of teacher-written direct vs. indirect feedback on students' writing. *Procedia - Social and Behavioral Sciences, 192*, 116–123. http://dx.doi.org/10.1016/j.sbspro.2015.06.018

Japos, V. (2013). Effectiveness of coaching interventions using Grammarly software and detection software in reducing grammatical errors and plagiarism of undergraduate researchers. *JPAIR Institutional Research, 1*(1), 97–109. http://dx.doi.org/10.7719/irj.v1i1.207

Knoch, U., Rouhshad, A., & Storch, N. (2014). Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? *Assessing Writing, 21*, 1–17. http://dx.doi.org/10.1016/j.asw.2014.01.001

Liao, H. (2016). Enhancing the grammatical accuracy of EFL writing by using an AWE-assisted process approach. *System, 62*, 77–92. http://dx.doi.org/10.1016/j.system.2016.02.007

Linville, C. (2009). Editing line by line. In S. Bruce & B. Rafoth (Eds.), *ESL writers: A guide for writing center tutors* (2nd ed., pp. 116–131). Portsmouth, NH: Heinemann.

Liu, S., & Kunnan, A. J. (2016). Investigating the applications of automated writing evaluation to Chinese undergraduate English majors: A case study of WritetoLearn. *CALICO Journal*, *33(1)*, 71–91. http://dx.doi.org/10.1558/cj.v33i1.26380

McGregor, K. M., Merchant, A. R., & Butler, M. (2008). Making feedback more immediate. *CALlaborate International, 16*(1) 8–11. Retrieved from https://openjournals.library.sydney.edu.au/index.php/CAL/article/view/6020/6671

McNaught, K., & Shaw, G. (2016). Preparing undergraduate students to be successful writers: Exploring the spelling, grammar and punctuation skills of students identified as potentially 'at risk'. *Journal of Academic Language and Learning, 10*(2), 11–19. Retrieved from http://www.journal.aall.org.au/index.php/jall/article/view/406/253

Murray, N. (2011). Ten 'Good Practice Principles' … ten key questions: Considerations in addressing the English language needs of higher education students. *Higher Education Research & Development, 31*(2), 233–246. http://dx.doi.org/10.1080/07294360.2011.555389

Pearson Education Ltd. (2012). *PTE academic score guide, version 4*. Retrieved from http://pearsonpte.com/wp-content/uploads/2014/07/PTEA_Score_Guide.pdf

Qassemzadeh, A., & Soleimani, H. (2016). The impact of feedback provision by Grammarly software and teachers on learning passive structures by Iranian EFL Learners. *Theory and Practice in Language Studies, 6*(9), 1884–1894. http://dx.doi.org/10.17507/tpls.0609.23

Ramineni, C., & Williamson, D. M. (2012). Automated essay scoring: Psychometric guidelines and practices. *Assessing Writing*, *18*(1), 25–39. http://dx.doi.org/10.1016/j.asw.2012.10.004

Reis, C., & Huijser, H. (2016). Correcting tool or learning tool? Student perceptions of an online essay writing support tool at Xi'an Jiaotong-Liverpool University. In S. Barker, S. Dawson, A. Pardo, & C. Colvin (Eds.), *Show Me The Learning. Proceedings of ASCILITE 2016, 33rd International Conference of Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education* (pp. 529–533). Adelaide: University of South Australia). Retrieved from http://2016conference.ascilite.org/wp-content/uploads/ascilite2016_reis_concise.pdf

Salehfard, E., & Bagheri, M. S. (2014). Preferences of EFL teachers and students towards corrective feedback in writing. *Modern Journal of Language Teaching Methods, 4(*4), 308–319. Retrieved from http://mjltm.org/en/archive.php?rid=18

Scouller, K., Bonanno, H., Smith, L., & Krass, I. (2008). Student experience and tertiary expectations: Factors predicting academic literacy amongst first year pharmacy students. *Studies in Higher Education, 33*(2),167–178. http://dx.doi.org/10.1080/03075070801916047

Tynan, L., & Johns, K. (2015). Piloting the post-entry language assessment: Outcomes from a new system for supporting research candidates with English as an additional language. *Quality in Higher Education, 21*(1), 66–78. http://dx.doi.org/10.1080/13538322.2015.1049442

Wang, Y. J., Shang H. F., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer Assisted Language Learning, 26*(3), 234–257. http://dx.doi.org/10.1080/09588221.2012.655300

Winder, R., Kathpalia, S., & Koo, S. (2016). Writing centre tutoring sessions: Addressing students' concerns. *Educational Studies, 42*(4), 323–339. http://dx.doi.org/10.1080/03055698.2016.1193476

56

Corresponding author: Ruth O'Neill, r.oneill@cqu.edu.au