

Roxifyonline: Helping Students Improve Their Writing through Online Feedback

Roxanne Miller*
George Miller

University of Jyväskylä

DOI: 10.22236/JER_Vol3Issue2pp152-167

Today's students want immediate feedback on their writing which instructors often cannot provide. They embrace the internet and want to use it for learning. To keep pace with these desires, Roxifyonline was created. It is a new online Computer Assisted Language Learning (CALL) tool that provides automated essay feedback for use by students worldwide. The program was originally written for ESL learners in Hong Kong, but is now being used by students throughout the Asia Pacific region. It is based on both empirical research and that of other academics. The program is robust enough to be used for general academic writing from upper secondary through post-graduate levels. It is a free Independent Learning (IL) tool. The program utilizes vocabulary to identify possible errors and provides online IL links to aid in the correction of possible errors in writing. Unlike other programs, it does not provide suggested corrections. This can help students to become more autonomous in their learning. Comparisons can be made between drafts, and there is an annotation function which can allow for more individualized peer and teacher feedback.

Keywords: AWE (Automated Writing Evaluation), feedback, assessment, computer programming, Java, CALL (Computer Assisted Language Learning)

Saat ini siswa ingin cepat mendapatkan umpan balik terhadap tulisannya. Mereka mencari di internet dan ingin menggunakannya untuk belajar. Untuk memenuhi kebutuhan tersebut, Roxifyonline diciptakan. Ini Pembelajaran Bahasa Berbasis Komputer (Computer Assisted Language Learning - CALL) yang memberikan umpan balik terhadap esai yang dapat digunakan di seluruh dunia. Program ini awalnya dibuat untuk siswa bahasa Inggris di Hong Kong, tetapi sekarang telah digunakan di seluruh wilayah Asia Pasifik. Program ini didasarkan pada studi empirik dan studi yang dilakukan akademik. Program handal untuk digunakan pada menulis akademik mulai tingkat sekolah menengah hingga pascasarjana. Program ini alat yang gratis untuk belajar mandiri dan memberikan bantuan link untuk koreksi. Tidak seperti program lainnya, program ini tidak memberikan saran koreksi. Ini bisa membantu siswa lebih mandiri. Perbandingan dapat dilihat mulai dari draf, dan ada fungsi catatan yang memberikan umpan balik teman dan guru.

* Corresponding author. Email: katsmom01@gmail.com

INTRODUCTION

Feedback on writing is an important part of the learning process for students. They should be able to receive it soon after they have finished writing to be better able to improve their work. This is not always possible and for this reason, students turn to technology to obtain the immediate feedback they crave. Roxifyonline is an automated essay feedback system that was designed to provide students with immediate, informative, and useful feedback on their writing.

Research has shown that in order for feedback to be effective, it must be personalized and timely (Alderson, 2005) (Hyland K. &, 2006) (Chan, 2014) (Ferster, 2013). Roxify online does just that. Students are now given results within a matter of seconds. Once they have entered the source texts if used, and their essays, the results are shown on the webpage immediately. There is no need to wait for a teacher to respond. This instantaneous feedback is beneficial to both students and teachers. It saves time for both parties and allows for more time to improve the writing.

Students want immediate feedback and teachers want to provide it; however, they often do not have the time to give timely feedback (Wen, 2013) (Leki, 1991) (Radecki, 1988). When teachers have many classes that require writing feedback, they face marking overload and the possibility of early burnout is high (Lee, 2014). Automated feedback can help reduce the instructor marking load thus freeing them up to focus on the areas of organization (Ferris, 2003) (Hyland K. &, 2006) (Zamel, 1985) (Stevenson, 2016), and grammar (Chen, 2008) (Li, 2015), and hence decrease the pressure on the teachers.

Automated essay marking and feedback systems have been around since the 1960's when Ellis Page designed Project Essay Grader (PEG) (Page, 1968). In 1982 Writer's Workbench was developed for use on Unix systems (MacDonald, 1982). This was followed by Intellimetric in 1996 (Elliot, 2003), Intelligent Essay Assessor in 1997 (Relations, 1998), and E-rater in 1999 (Burnstein J. , 2003). These are commercial products that are unavailable to students without a paid subscription. As such, they are only useful a handful of students worldwide.

As it is important for me as a teacher to be able to give timely and useful feedback to my students, I scoured the internet for free resources to help them. After extensive searching, I found a small handful of programs that claim to provide free feedback for students on their writing. These include Marking Mate (Jordan, n.d.), After the Deadline (After the Deadline, n.d.), PaperRater (PaperRater, n.d.), Hemmingway App (Long, 2013), Grammarly (Grammarly, n.d.), Online Correction (OnlineCorrection, n.d.), SpellCheck Plus (Nadasdi, 2001), and The PolyU Common Error Detector (Morrall, 2000). While each of these programs is useful for checking grammar, style, and spelling; they do not help the student with correction. Nor do they give the student advice on where to find help.

Additionally, not all of these 'free' programs are actually free. PaperRater will check up to five pages of text before stopping. If a student wants more help, they must purchase a paid subscription. This, however, will only cover 20 pages rendering it not useful for students at higher academic levels. Grammarly will check for only grammar and spelling. For other error types, the student must again purchase a subscription. Hemmingway Editor is geared more towards younger writers and tags complex sentences as being wrong. Marking

Mate will not accept biographical information and is not conducive to academic writing similar to Hemmingway Editor; it marks complex sentences as incorrect. Online correction analyses spelling, spacing, typographical errors and grammar. The PolyU Common Error Detector has specific error types that are common among native Cantonese speakers of English. It, along with Grammarly, is the only program that suggests corrections.

As there are no online programs that link students to independent learning help, it was decided to develop one of my own. This paper is, in fact, two studies in one. The first is the design and implementation of the online system, and the second is the changes in student writing based on the usage of that program. As such, all sections of this paper will be divided into these two studies.

Description of the Program

Roxifyonline is an online automated essay feedback system that provides independent learning help through analysis of students' academic writing and provides directed links to areas needing improvement. Student papers are entered into the program and are then marked (but not scored). Potential problems are highlighted and links are provided for independent learning. This enables the student to receive immediate help on problem areas of the writing. A results section shows the percentage and number of times a student has made such possible errors, as well as an estimate of the readability of the paper. A comparison feature that provides students and teachers with a visual of the changes made between drafts is another function of the program. There is also an annotation function for teachers to make any other comments on the paper that they feel important. Additionally, this function can be used for peer-feedback. The program is robust enough that it can be used with students (and staff) writing at all levels (for a sample of the results of this paper (for a copy of this please go to <http://roxifyonline.com/Rox/htmls/ttoc.html>). On top of that, it includes a facility for teachers to gather statistics of each essay submitted through Roxifyonline for further analysis of student progress during the course of the semester in the form of a csv file that can be used with any spreadsheet program. Additionally, there are student and teacher manuals, as well as video instructions, lesson ideas, and justification documentation of the program itself embedded in the program.

The fact that it is web-based allows users to access Roxifyonline on virtually any browser and operating system on any internet connected computer. This program allows students to obtain help with their writing at any time of the day without the need for teacher intervention. It also enables teachers to witness and record progress made by students throughout the semester. Both teachers and students are able to see all drafts of the essays submitted. It is fast, robust, and highly secure. There is a built-in results page which is tailored to the needs of the individual student.

As program links to online independent language learning sites, the students are given immediate access to help. This enables them to immediately investigate methods of improving their writing without the need for teacher input. Ultimately, this will help to lighten the marking load as the students should have made improvements in the paper before teachers see it. As the program is based on vocabulary usage, it should free up teacher's time to mark higher level discourse or content errors. This program has been self-funded as a part

of my PhD and use of it has been provided free of charge to students and teachers worldwide.

METHODS

Computer program

Roxify was originally a standalone program that was written in Python. The results of the program looked confusing to the non-imitated user without providing any real feedback to the students. They received a color-coded copy of their essays with a number of links and some generic comments on their writing. This program was only available on the researcher's computer and was incredibly labor intensive. If a student wanted to have their paper Roxified, they needed to email it to the teacher who would run it through the program, save the results, and email it back to the student. As this was not ideal, it was decided to put the program onto the internet to enable students to have 24-hour access to it.

After the initial standalone program, which was written and designed by the researcher, was deemed unhelpful, the aid of a programmer was enlisted. A specification document was drawn up and the programming began. As with all contracted programming work, communication was a problem. The programmer was on the west coast of the United States, and the researcher was in Hong Kong; therefore, it was difficult to communicate in a timely manner. Email and texting were the primary forms of communication at this distance.

The first major change that was made to the program was the shift from Python to Java Script. The program also needed to include CSS (Cascading Style Sheets) and HTML 5 (HyperText Markup Language) to make it fully compatible with today's online environment. As these are two completely different programming languages, there was no usable code which could be taken from the first program. The programmer had to start over with only the specification document to guide him. The distance and time differences involved was the main deterrent to the program at this time.

One of the requirements of the program was that it be made modular. In other words, the code had to be written in such a way as to be able to change or add any part to the program without having to rewrite the entire program. Finally, the essays had to be available for the researcher to retrieve. Because of this, the system had to be cloud based. This required a further level of programming. We chose to use Google Firebase as the storage system.

Classroom usage

Once the program was written, the first study was conducted on its use. In January 2017, the program went live and students in the researcher's English for Academic Purposes (EAP) class were given the option to use it for their writing work. There were two studies done over the course of the 12-week semester involving two groups of students.

As this was a feasibility study, there was no control or experimental group. All students involved volunteered to participate. In the first pilot, students wrote the first draft of an IMRD paper (Introduction, Methods, Results, Discussion). There were a total of 49 non-native English speaking students from Hong Kong and China in the study. There were 28 students from Hong Kong (Cantonese L1), 19 From Mainland China (Putonghua L1) and one from Indonesia. The ratio of males to females was 26:22. There were 40 freshmen in the study and eight sophomores. The students ranged in age from 17-24 with the majority being

18 and 19. 21 of the students accessed the program through an Apple operating system while 27 used an android or windows based system.

The second study was conducted with 53 non-native English speaking students from Hong Kong and China, 49 from the first study and an additional four who wanted to use the program. This group wrote a second draft of the IMRD and final IMRD. All three drafts were run through Roxifyonline to check for errors. As I did not expect to have additional participants in the study, I did not ask demographic questions of this group. Upon analyzing the feedback, it has been determined that the additional four students were female. All of the new participants were freshmen.

Students used Roxifyonline in their EAP class on university provided computers or on their own laptops. All computers must have the latest version of the web browser being used. The program was used on Chrome, Mozilla, Safari, and Opera. It does not work on IE as that system has not been updated to accept HTML 5.

Students wrote the paper in their choice of a word processing program and brought it to class for a peer-feedback session. After receiving peer-feedback, they were requested to make any changes they wished and submit it to the teacher for feedback. After this second draft, students were again requested to submit the paper to Roxifyonline and make any final changes based on both feedback from the teacher and the program before submitting it for final grading.

Upon completion of the semester, students were asked to complete an online survey using google forms. This was done for ease of use by all participants. All students in the study completed the survey. After the survey was completed, eight students volunteered for one-on-one interviews with the researcher. For the convenience of the students, these were conducted in the English Language Centre of City University. They were recorded using AudioNote 2 software for Mac. This was done for ease of recording and note taking at the same time. All student participants were paid HK\$50 (approximately US\$6.50) each for a twenty-minute interview.

FINDINGS AND DISCUSSION

Program design

From November 2016, the programmer began writing an entirely new version of Roxifyonline. This one uses JavaScript, CSS and HTML 5. As the original program was written in a different language, it was deemed unusable and he had to start over from the beginning. A soft deadline was set for an initial rollout in the beginning of January of 2017. As stated earlier, it was difficult as there was a time difference and any changes made were delayed as a result of this. In January 2017, the programmer moved to Hong Kong and work began in earnest. The researcher and programmer worked together on a daily basis and it made for quicker changes. An added benefit of this move was the ability to make instantaneous updates based on comments made by the users.

Initially the program only had eight types of feedback provided for the students. These included: use of GSL words (General Service List), use of AWL (Academic Word List), number of discourse markers, number of duplicate words, number of key vocabulary (defined as any vocabulary taken from the source text that is unique to that text), number of misspelled words, number of reporting verbs used, and number of sources. The results were

given to the students in a color-coded page with along with a matrix that included the number of instances each of the previous items was used, the percentage of the overall paper that used these items, the word count of the essay, a general comment on the writing level and one or two links to an online resource to help the student make corrections.

As the program was to be put online, there had to be a method of separation of students, teachers, and topics. It was decided to make a 'safe page' for data entry. This comprises the essay topic, student name, the student number, and the teacher name. (For a copy of this please go to <http://roxifyonline.com/Rox/htmls/ttoc.html>). Upon clicking the safe handle for entry into the program, the students saw a blank box in which to paste their essays.

The first major change from the original program was to add additional error types. To do this, it was decided to incorporate buttons that would each be individually clicked so students could readily move between the different features of the program and not have an excessive number of comments on one page. This feature was added because the students felt that they were not really given enough information on each of the categories in the off-line version of Roxify. The program had five main sections of errors, a results page, and a button to enter a new essay.

The buttons constitute: clichés (clichés, plague words, idioms, and slang); warnings (misspellings, contractions, pronouns, and value-laden words); AWL/GSL (the AWL and the NGSL); vocabulary (duplicates, key vocabulary, discourse markers, hedging, and statistical meta language); grammar, and citations (reporting verbs, and in-text citations). Each of these buttons has the possible errors highlighted to make it simple for the students to see their problem areas. Each of the error types was then given between two and four online links that the students could click to go to independent language learning (for a copy of this please go to <http://roxifyonline.com/Rox/htmls/ttoc.html>).

The final page consisted of the results the students were given. They include AWL, GSL, clichés, idioms, slang, plague words, key vocabulary, discourse markers, and hedging. Not all features of the program are listed in the results as it again was of little use for student improvement (for a copy of this please go to <http://roxifyonline.com/Rox/htmls/ttoc.html>). Cloud storage was created using Google Firebase. It was however not incorporated into the main program in time for this study. This was unfortunate as the researcher was unable to determine the number of times each student used the program.

The program design was one of modularity. This means that each of the buttons has their own module. A module is all the code pertaining to a particular button. Thus, if we are having a problem with AWL/GSL we only need to investigate the AWL/GSL module. Each module has its own set of dictionaries. This button, for example, has two individual dictionaries consisting of all the words in the GSL, the other the AWL. In total, there were eleven different modules in the program.

Initially the program had a grammar function included but it was not useful for the students so it was taken out. The grammar function highlighted the different parts of speech used, but did not give any insight into the correctness of the parts of speech. It was decided to hide this function of the program until a future date when there is more time to work on the grammatical aspects to be included in Roxifyonline. Grammar checkers are dangerous in

themselves as students expect all of their errors to be noticed by the machine. Grammar checkers cannot find all errors (source). Quite often, these online programs give students misinformation. Rather than give the students a faulty product, it was deemed better to hide the feature until a later date.

Each of the buttons was designed using empirical research on the needs of Hong Kong students writing. To begin, the students often use items under the clichés button in their writing. The teacher explained that the words highlighted on this button should be avoided in academic writing. They can sometimes be used to good effect in novels, or speeches. They are however seen as too informal for writing at the university level. Any use of these words should be done sparingly. Clichés, Idioms, and Slang are self-explanatory. Plague words are defined as language that has been over used in student writing. These words often have no meaning, are empty in substance, or are too vague. They are often used only as ‘fillers’ to ‘add meat’ to an essay. They should be avoided in academic writing.

The next button is warnings. These are defined as not necessarily wrong but they should be reviewed before submission. They may need changing, or at the very least, students may need to justify the usage of them to their instructor. Misspellings and contractions are again self-explanatory, but the other buttons are defined as follows.

Inclusive pronouns should be avoided in academic writing. These are pronouns that include the reader in the writer’s opinions or ideas. They include *we*, *us*, *our*, *you*, and *your*. *We* can be used if it is exclusive. While inclusive *we* refers to both the writer and reader, exclusive *we* only refers to the writer and associates (Harwood, 2005). At the same time, *us* and *our* can also be inclusive or exclusive (Quirk, 1985). As above, these can be used in quotations, idioms or clichés, or in comments.

Value Laden Words are defined as words that cause an emotional response in the reader, or try to bias the reader or change their opinion on an issue. They should be avoided but if used, students should be able to justify their reason for using these words. Some words and phrases reveal your own opinion or bias. For instance, if you state that something will *obviously* happen, you are actually indicating that you think the occurrence is obvious – not stating a fact. Expressing an opinion is usually only appropriate in certain sections of a paper (namely the preface, acknowledgements, discussion, and reflection), so be careful when using these words. (Vinz, 2017).

Academic writing is usually direct. Adverbs of frequency (*always* and *never*), superlatives (*the best*, *the worst*), and intensifiers (*very*, *really*) are often too strong to be used. If used, the claim made is that something is either perfect, or that it is terrible. This more often than not is not accurate.

Archaic Words are no longer in everyday use or have lost a particular meaning in current usage. They are sometimes used for a deliberately old-fashioned effect in historical novels, or in standard conversation or writing for a humorous effect. They are also frequently used in legal documents. Most academic writing will not use this type of language unless it is in the fields of literature, law, or history. While it is not necessarily wrong to use archaic words, if a more modern word can be found, it should be used.

The AWL/GSL button include only these two word lists. The General Service List (GSL) is a list of roughly 2,284 words published by Michael West in 1953 (West, 1953). The words were selected to represent the most frequent words of English and were

taken from a corpus of written English. The target audience was English language learners and ESL teachers. To maximize the utility of the list, some frequent words that overlapped broadly in meaning with words already on the list were removed.

These words selected were said to be of the greatest "general service" to learners of English at that time. The list is important because a person who knows all the words on the list and their related forms would understand approximately 90–95 percent of colloquial speech and 80–85 percent of common written texts. The list consists only of headwords, which means that the word "be" is high on the list, but assumes that the person is fluent in all forms of the word, e.g. am, is, are, was, were, being, and been.

The Academic Word List (AWL) was developed by Averil Coxhead at the School of Linguistics and Applied Language Studies at Victoria University of Wellington, New Zealand (Coxhead, 2000). The list contains 570 semantic fields which were selected because they appear with great frequency in a broad range of academic texts. The list does not include words that are in the most frequent 2000 words of English (the General Service List). The AWL was primarily made so that it could be used by teachers as part of a program preparing learners for tertiary level study or used by students working alone to learn the words most needed to study at colleges or universities. The 570 words are divided into 10 sublists. The sublists are ordered such that the words in the first sublist are the most frequent words and those in the last sublist are the least frequent.

The Vocabulary button highlights the vocabulary used in the text. It is simply here for students to notice their usage. It is not necessarily wrong, but students should make sure they are using the language they intended. Often students will choose a word that has a similar meaning to the one they should use because they have over used certain words. These frequently do not make sense within the text. Students will occasionally 'lift' vocabulary from the source text. They also have a tendency to not use a thesaurus and repeat individual vocabulary.

Key Vocabulary is vocabulary that has been taken from the source texts. It is the most important vocabulary based on frequency of use. If more than one source text is inserted, the vocabulary changes accordingly. By important we mean those frequently occurring words that are not naturally frequent. We have eliminated the NGSL, AWL, and UWL as these are commonly used in academic writing. While there is no problem with using these words in writing, students may want to find synonyms for them if they are paraphrasing or summarizing. Not only will it make the text more interesting for the reader, but it will also ensure that plagiarism has not occurred.

Discourse markers, 'linking words', 'linking phrases', 'cohesive devices', 'transitions' or 'sentence connectors' ('however', 'although' and 'therefore') are all highlighted in this section. They are used to show logical connections between the various parts of an essay. Without sufficient discourse markers in a piece of writing, a text would not seem logically constructed and the connections between the different sentences and paragraphs would not be obvious.

Many students, however, over-use of discourse markers. Using too many of them, or using them unnecessarily, can make a piece of writing difficult to read. They are important, but should only be used when necessary. As a general rule, writing should not have one in

every sentence. Not all sentences should start with a discourse marker, they are also not needed at the beginning of every paragraph.

The use of discourse markers is set at a level of only 1 in every 10 sentences. This is to help make sure there are not too many used. It is not an absolute number that students should try to achieve. There is no correct number of usage times, but be sure they are used correctly, and with purpose.

One feature of academic writing is the use of caution to distinguish between facts and claims. This is known as ‘Hedging’. It was first defined by Lakoff (1973) as “words that make things fuzzy”. Hedging can be defined as the use of language to show hesitation or uncertainty and to display politeness and indirectness. Hedging is a necessary part of academic writing. It shows formal abstract writing, and the ability to show the how much belief the writer has in the claim that has been made. People also use hedging for several other purposes but perhaps the most fundamental are the following:

- to minimise the possibility of another reader opposing the claims you are making
- to be more precise when reporting results, e.g. you can show that something is not 100% proven, but rather that it is suggested and therefore assumed.
- To show a politeness strategy while acknowledging that there may be flaws in your claims
- To conform to an accepted practice of writing
- To reduce your commitment to the truthfulness of a statement.

(Hyland K. , 1998)

Hedging is highlighted in Roxifyonline to indicate where it has been used. There is no clear amount of times it should be used, but it should be used in the examples shown above.

Statistical Meta-language are the words used to discuss statistics. This has been included to raise awareness of word choice. Students should check all the highlights to make sure they are using the words correctly. This section may not have any highlighted words if a student is not analyzing data.

The citations button tells the reader to avoid the semblance of plagiarism, as all ideas and opinions taken from other sources must be cited. It gives credit to the original author for the ideas. This also allows the reader to find the sources used if they would like to use the same source later or if they have further questions about it. Finally, it gives the reader the understanding that the student has done research into the topic.

Roxifyonline uses APA format for its preferred citation style, but there are many others that may be used. Students should check with their instructors for the preferred citation style. All links on this page lead to APA referencing.

This button only works if a student has added a reference list. Roxifyonline will not highlight any sources within the text, or any reporting verbs if a student has not included the reference list with a proper heading. This may include the words bibliography, references, reference list, and works cited to name a few.

Roxifyonline also has the Flesch-Kincaid Reading Ease score (Flesch, n.d.) included on the results page. Students at City University are aware of this number from their readings and teachers use it to set exam readings and in the design of the textbook. It also provides

students with the grade level they are reading at and a generic statement about the difficulty level.

Student Usage

First trial

This semester there was no experimental group. As the program itself was new to the web, I decided to offer it to students to use as and when they wanted. I did not require any to use it as I had in the past. This allowed half of my students to opt-out of the program.

For those students (n=49) who chose to use Roxifyonline, I held a separate out-of-class training system on how to use the program. The students spent an hour in a computer lab learning to use the features of the system. They were then sent out on their own to use Roxify and make any corrections to their writing that they wanted to.

After the initial use of the system I gave the students a survey to fill out. All 49 students completed the survey. At this stage I was only interested in the type of feedback the students preferred to receive, the perceived usefulness of the program as a whole, the amount of times they used Roxifyonline as well as the number of links used.

In the first trial of the study, I discovered that although 43% of the students owned an Apple product, only 19% used the Safari web browser. The other 47% of students owned a desktop or PC laptop.

Only one student did not like the color coding. Only one student did not find the program easy to use, could not understand the feedback provided and found it difficult to enter text. Eight of the 49 students thought that Roxifyonline should not be compulsory for all students in the EAP course.

When asked about the type of feedback preferred, 12% opted for face-to-face only, 16% wanted computer only, 20% requested written feedback with a face-face-consultation, while 23% wanted written only. The remainder of the students wanted computer generated feedback followed by face-to-face consultations. When asked about the usefulness of the program, the majority at 78% found it to be highly useful, 12% were neutral, and 10% claimed it was not helpful in improving their writing.

The students were asked which buttons of the program were most useful. They were allowed to tick all that applied. Surprisingly the least developed aspect of the program, Grammar, topped the list in the most useful of the program. This was chosen by 71% of students. Not surprisingly, it was followed by vocabulary with 65% with warnings and the AWL/GSL coming in the third and fourth positions with 53 and 51% respectively. Another revelation was the results page. Only 30% of the overall students found this page to be useful.

When it came to the usefulness of the links, 18% were neutral, 8% did not find them helpful, and the majority (74%) found them very useful. Of these, 8% said they were better than using books. These are reasonable results given the number of links students reported using. 6% did not use any links at all, 40% used less than 3 links, 36% used 3 to 5 links, and the remaining 19% used between 5 and 8 links. No one claimed to use over 9 links. This is again, not a real shock to the researcher for two reasons. First, the students have come from an environment where the teachers correct all the errors, and the students simply need to

rewrite the essays, and second, this is the first time any of these students have encountered a program like this. Overall, I was pleased with the feedback provided and decided to continue the study for the rest of the semester.

Second trial survey results

The second half of the semester saw the students finishing their IMRD papers, having teacher feedback on them, and they had a final exam. I expected to have a number of drop outs from the program so I was pleased and shocked to see that not only did I have the 49 students stay in the study, but I also had another four students join the study. When asked, those four had spoken to their classmates and at first found it intriguing, and later discovered it was a useful tool.

In January, in addition to the one hour of training on the use of Roxifyonline, the students were given a printed instruction manual. This was designed so students would be able to leave the classroom and have any questions they may have answered for them without having to wait for the teacher to reply to emails or phone calls. 2/3 of the students still preferred the hour-long lesson, but the others wanted the manual to fall back on.

The other big change was the removal of the grammar button. I did not find it useful and felt it only hindered the students. Because it was not fully developed, I decided to take it out of the system. That did not stop the students from commenting on it though. Three felt the grammar checking function was useful and two wrote that Grammarly was an equally helpful and free program.

This semester, 77% still used chrome, but now that the Apple users realized that Roxifyonline works on Safari; 25% were using it compared to 19% in the beginning of the semester. IE users increased to 6% and this time around 2% of the students were using Mozilla.

As for the type of feedback, in semester B of the 2016-2017 academic year, 6% of the participants did not want any. Only 2% wanted written only and 8% wanted face-to-face only. The other 83% wanted some kind of computer-generated feedback. For 19% it was computer feedback with teacher written feedback included, 28% wanted a consultation after having had time to read the computer-generated feedback and the other 45% wanted only computerized feedback.

Based on the results of the type of feedback the students wanted, it came as no surprise that 96% of the students found Roxifyonline to be useful. When asked to explain, the most frequent comments related to the use of the grammar checker and the ability to see their level of AWL usage. The grammar, as stated above, only showed the parts of speech but this obviously was all the students needed to make corrections. The use of AWL was new to these students so it was expected that they would find it useful. Other important comments include “I can check the essay in anytime and do not need to wait someone to correct”, “I can refer to the feedback when I want to make some changes on my term-paper because it’s easy to forget some of the feedback after face-to-face feedback.”, “It imitates a teacher to help me.”, and “The computer can point out the problem I may not realize. I can learn from this feedback.”

Again, when the students were asked which aspects of the program they found most useful and allowed more than one choice, the majority liked the inclusion of the AWL with

66% choosing this option. This was followed by duplicate words at 51% and key vocabulary at 49%. The least useful features were contractions (11%) and Statistical Metalanguage (15%). This was the same order found when given only one choice.

This is probable as the students have not been taught the AWL before they entered university. Instructors in the ELC provide them with a bare-bones lesson on what it is, and where to find exercises. No real instruction is provided on how to improve writing through the use of the AWL. Students found duplicate words to be useful as they have been told up until this time to use a thesaurus. While many said they did, they also admitted that they did not notice the over use of some words and Roxifyonline helped them notice this problem. They would have liked the program to take them to a thesaurus to see other available vocabulary choices. The key vocabulary was a new concept for them as they had often been told in their secondary schools to use the 'big words' from any text they were citing. They did not realize that over use of this vocabulary is a precursor to plagiarism and should be used more sparingly.

The most useful button was grammar, followed by warnings and AWL. Interestingly, the students did not comment on the grammar when asked why they chose this button. Instead they wrote about the vocabulary they were using. "It is because it could be able to find out which of my word isn't academic enough and can to change it immediately.", "My English level was just passed the exam standard with some lucky, so the natural English writing and the vocabulary choice would be the most difficulty for me. And NGSL really help me to enhance the paper professional.", "It can show what my essay whether achieve university level and which words are too simple.", and finally, "It helps me for paraphrasing as well as learning wide range of vocal." The least useful buttons were citations, vocabulary, and clichés.

For me as a teacher and Language Learning Advisor, one of the most rewarding results was that 90% of students felt the highlighted areas motivated them to do more self-study. As the designer of the program, it was also pleasing to find that 78% of students used the automated feedback to improve their writing. 89% of the students found the links provided to be useful and used them. Some of the more interesting comments from the students were "automated feedback help me to find out my mistake then I can correct it", "Find more detail from other online sources after received the feedback", and "I use it when I have something not sure. I treat it as a teacher. It can help me to learn independently." This is encouraging as students need that immediate feedback which I am unable to provide.

Only 4 students said they would not use Roxifyonline again, and 96% said they would recommend it to others. 80% of students in this group used the program more than three times for each section of their IMRD. Finally, a resounding 95% believed the use of the program helped them to improve their English writing. All students were told at the end of the semester that Roxifyonline is available for the students to use if they wish.

Second study interviews

This semester eight students agreed to interviews. All subjects were paid volunteers. With one of the students (S1-male), the monetary reward caused bias. He felt he had to say positive things about Roxifyonline until I realized what he was doing and asked him to be

honest. He believed the highlighting made it easy to spot the mistakes and there was good separation between colors. He admitted that the feedback links were useful and that there were enough examples provided. However, he felt that it was too time consuming to go into the links and do the activities. He believed that it would be better if the program provided correct answers for the mistakes made. Other negative aspects of the program included thinking aspects of his writing was good even though it was not simply because something had not been picked up on. The grammar function only tells him the parts of speech and is useless. He thought it was wasting his time to click that button. Finally, he did not trust the program because it did not find all of his mistakes.

He said he was not a motivated student and he was not sure if the feedback in the form of links was the best option for students like him. He did think it was fun to see the mistakes in color though. He has never received this type of feedback before and would like to continue to use the program in his other classes and throughout his academic career. Looking back at the students who have used the program since September 2017, this student's name is not on the list. He was also the only student who realized that Roxifyonline not only helps the students save time, but also the teacher. He realized that if he made the effort to make changes that it would make marking the final papers easier and faster for the teacher.

He also compared it to other programs. He felt the turn-it-in feedback function (used by other subject teachers) was pointless. It was hard to see the feedback given, and when he moved the feedback button to read his writing under it, he had to try to remember where the button was to begin with. It did not stay in one place. Roxifyonline, by comparison, is better as everything is highlighted and not movable. He thought Grammarly was better for picking up grammar mistakes and would rather use that program. He also believed that Microsoft Word is better at finding spelling errors when he remembered to use that function.

Student 2 (S2), also male, was only positive about the program. This student was keen to improve his English and claimed he used the program 4 times for each draft of his paper. He preferred the computer feedback over any other but still wanted to have at least one face-to-face consultation for problems with style and organization. He had an interesting method of using Roxifyonline. First, he would use the program to check for vocabulary usage; key vocabulary, duplicate vocabulary, AWL/GSL. Then he would check for discourse markers and pronoun usage. Next, he would check the clichés button and finally, he would check it for general meaning. When asked how he did this with Roxify, he said it was easy with the color coding; he could see the discourse markers and the use of hedging. This showed him if he was organizing the paper logically.

S2 believed that Roxifyonline helped him obtain a better grade in the course as he was able to help himself find his errors. He was able to notice with the help of the program that he was using too many compound sentences and was falling into the habit of using a discourse marker at the beginning of each sentence again. Roxifyonline allowed him to change the sentence structure before turning it in to the teacher for grading.

He would like the program to analyze more grammatical aspects such as pronoun referencing and give examples on how to fix the grammatical errors he made. He would rather have corrections than links as self-learning takes too much time. Overall though he is of the opinion that Roxify online has helped him to become a better writer. He now realizes

that he has been writing at too low a level for university and will find ways to continue to improve this.

All of the students interviewed provided useful insight into the ways they used the program and many provided feedback on how to further improve it. These comments were taken into consideration and many will be implemented in the updated version of Roxifyonline.

Limitations

As this study was conducted using only the researcher's classes the results cannot be extrapolated to other populations. There is also the issue of researcher bias as the participants were in my classes and were by virtue of volunteering the 'better' students in the class. In addition, I have a vested interest in the program and my enthusiasm for it was easily seen by my students. The participants were all Chinese speaking L1 students (both Cantonese and Putonghua) in Hong Kong so this cannot be generalized to other first language of students. The feedback given to the students was specific to the EL0200B course taught at the English Language Center of City University of Hong Kong and may not reflect the needs of other student populations.

Another large limitation was the lack of cloud storage. The researcher had to rely on the survey conducted to determine the usage of the program by students. There was also lost opportunity in that I was unable to capture the various changes made on each version of the paper written. Furthermore, the inability to print the Roxified papers, students may have benefited from having a hard copy of the Roxified paper in front of them. Instead, they could only work on their essays while connected to the internet.

Internet connection is the final limitation. Students must be connected to a web browser that runs HTML 5. If they do not have access to a web browser dated before 2015, they will be unable to access this program. Many students said they were unable to access Roxifyonline in Internet Explorer (IE) and felt disadvantaged by this. Microsoft has not maintained IE and has not updated it to run programs using HTML 5.

CONCLUSION

The recruitment of a programmer to develop the program was shown to be a positive move. This enabled a more professional looking design, and the user interface was much friendlier than the original Python version of the program. By having the program modularized, any changes that needed to be made could be done quickly and more easily than rewriting the entire program.

Roxifyonline has been shown to be useful to the convenience sample discussed here. The students used the program and this, in turn, helped them to improve their writing. The teacher was able to save time on marking the first drafts of the IMRD as the students used Roxifyonline to make changes to the language features of their writing. This allowed for more in-depth feedback on content and organization.

In the next update of the program, there needs to be an annotation function so teacher and students can provide additional feedback that is not included in the automated program. We will also be adding scoring rubrics and most importantly, the cloud storage system must

be fully functional. Finally, it is important that the program have an email function so the teacher can notify the students when the essay has been annotated and given teacher feedback.

In the first semester of the 2017-2018 academic year, more students will be recruited to use the program. Again, a classical study cannot be performed as all the students in the department will all be required to use the program for their EAP course. This will enable the researcher to collect more data and survey information to better inform future changes to the program.

ACKNOWLEDEMENTS

I would like to thank my programmer George A Miller for his continued support and tolerance throughout the design and development of this program. Without him, this would not be possible.

REFERENCES

- After the Deadline*. (n.d.). Retrieved October 21, 2017, from After the Deadline: <http://www.polishmywriting.com/>
- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Burnstein, J. (2003). The E-rater(R) Scoring Engine: Automated Essay Scoring with Natural Language Processing. In M. D. Shermis (Ed.), *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 107-115). Mahwah, New Jersey, Usa: Lawrence Erlbaum Associates.
- Burnstein, J. C. (2003). Criterion Online Essay Evaluation: An Application for Automated Evaluation of Student Essays. *American Association for Artificial Intelligence*. AAAI.
- Chan, P. E. (2014). The critical role of feedback in formative instructional practices. *Intervention in School and Clinic*, 50(2), 96-104.
- Chen, C. a. (2008). Beyond the design of automated writing and evaluation: pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning and Technologies*, 12(2), 94-112.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213-238.
- Elliot, S. (2003). Intellimetric TM: From Here to Validity. In M. a. Shermis (Ed.), *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 67-81). Mahwah, New Jersey, USA: Lawrence Erlbaum Associates.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum.
- Ferster, B. H. (2013). Implementation of applicatins of hte intelligent essay assessor. In M. D. Shermis, & M. D. Shermis (Ed.), *Handbook of automated essay evaluation* (pp. 68-88). New York, NY: Routledge.
- Flesch, R. (n.d.). *How to Write Plain English*. Retrieved 6 30, 2017, from Management, Marketing and Entrepreneurship: http://www.mang.canterbury.ac.nz/writing_guide/writing/flesch.shtml
- Grammarly*. (n.d.). Retrieved from Grammarly: <https://www.grammarly.com>
- Harwood, N. (2005). "We Do Not Seem to Have a Theory...The Theory I Present Here Attempts to Fill This Gap': Inclusive and Exclusive Pronouns in Academic Writing. *Applied Linguistics*, 26(3), 343-375.
- Hyland, K. &. (2006). State-of-the-art review on "Feedback in second language students' writing". *Language Teaching*, 39(2), 83-101.

- Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamin Publishing.
- Jordan, E. &. (n.d.). *MarkingMate*. Retrieved October 21, 2017, from MarkingMate: <http://writingtools.xjtlu.edu.cn:8080/mm/markingmate.html>
- Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2, 458-508.
- Lee, I. (2014). Editorial: Feedback in Writing: Issues and Challenges. *Assessing Writing*, 19, 1-5.
- Leki, I. (1991). The preference of ESL Students for error correction in college-level writing classes. *Foreign Language Annals*, 24, 203-218.
- Li, J. L. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, 27, 1-18.
- Long, B. &. (2013). *Hemingway App*. Retrieved October 21, 2017, from Hemmingway App: <http://www.hemingwayapp.com/>
- Lynch, J. (2005). How Johnson's Dictionary Became the First Dictionary . *Johnson and the English Language conference*. Birmingham.
- MacDonald, N. F. (1982). The Writers Workbench: Computer Aids for Text Analysis. *IEEE Transactions on Communications*, 3(1), 105-110.
- Morrall, A. (2000, November 28). *Common Error Detector*. Retrieved October 21, 2017, from Centre for Independent Language Learning : <http://www2.elc.polyu.edu.hk/CILL/errordetector.htm#egessay>
- Nadasdi, T. &. (2001). *SpellCheckPlus*. Retrieved October 21, 2017, from SpellCheckPlus: <http://pro.spellcheckplus.com/>
- OnlineCorrection*. (n.d.). Retrieved from OnlineCorrection.com: <http://www.onlinecorrection.com/>
- Page, E. B. (1968). The Use of the Computer in Analyzing Student Essays. *International Review of Education*, 14(3), 253-263.
- PaperRater*. (n.d.). Retrieved October 21, 2017, from PaperRater: <https://www.paperrater.com/>
- Quirk, R. G. (1985). *A Comprehensive Grammar of the English Language*. Harlow: Longman.
- Radecki, P. &. (1988). ESL student reaction to written comments on their written work. *System*, 16, 355-365.
- Relations, U. o. (1998, April 16). Retrieved October 21, 2017, from http://lsa.colorado.edu/essay_press.html
- Stevenson, M. (2016). A critical interpretative synthesis: The integration of Automated Writing Evaluation into classroom writing instruction. *Computers and Composition*, 42, 1-16.
- Vinz, S. (2017, March 27). *Taboo words in academic writing*. Retrieved July 18, 2017, from Scribbr: <https://www.scribbr.com/academic-writing/taboo-words/>
- Wen, Y. (2013, March). Teacher Written Feedback on L2 Student Writings. *Journal of Language Teaching and Research*, 4(2), 427-431.
- West, M. (1953). *A General Servie List of English Words*. London: Longman, Green, and Co.
- Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, 21, 697-715.