Vol. 1 · January 2013
Print ISSN 2244-1824 Online ISSN 2244-1816
doi: http://dx.doi.org/10.7719/irj.v1i1.207

JPAIR Institutional Research is produced
by PAIR, an ISO 9001:2008 QMS certified
by AJA Registrars, Inc.

Effectiveness of Coaching Interventions using Grammarly Software and Plagiarism Detection Software in Reducing Grammatical Errors and Plagiarism of Undergraduate Researches

GENARO V. JAPOS
ORCID No. 0000-0001-7627-0988
genarojapos@gmail.com
Research and Publication, Liceo de Cagayan University
Cagayan de Oro City

ABSTRACT

Grammar checker and plagiarism software are indispensable tools in aiding researchers to identify and correct their mistakes. The study examined the errors detected by grammarly software and plagiarism software and how these were reduced due to coaching interventions. The study used the descriptive and inferential research design involving 47 specimens of student research as undergraduate theses written as publishable paper. These papers were intended for the Advancing Research Journals by academic discipline available online, cross referenced and indexed. The study involved a statistical analysis of data derived from software reports tested through the use of t-test and Pearson correlation. The results showed significant reductions in the occurrence of the grammar errors indicating that the coaching interventions were effective. There is also an inverse, moderate and significant correlation between occurrence of grammatical score and plagiarism. The study concludes that those who commit higher incidents of plagiarism also have lower commission of grammatical errors. When students are poor in writing style and vocabulary use, their score for originality increases. The study recommends for adoption of academic integrity policies to curb grammar and plagiarism occurrences among student researchers in the university.

Keywords - Education, grammar checker, plagiarism detector accumulator, plagiarism detection, coaching, descriptive design, Cagayan de Oro City, Philippines

INTRODUCTION

My interviews with students revealed that they do not really know much about plagiarism since the term is not discussed in their academic classes. The majority came to know about plagiarism when they were required by their research teachers to have their manuscripts tested for plagiarism and grammar errors in the Research and Publication Office of Liceo de Cagayan University as requirement for final oral defense. The students said that the commission of plagiarism was not intentional.

Many of the teachers were educated prior to the internet era so their exposure was the print sources. They admitted that they had difficulty in documenting online sources in their manuscripts. The inadequacy to document online sources is passed on to the students who, despite their web literacy, cannot properly document their sources for lack of an effective pedagogy.

More often, students copy important parts of a text and graft them into their manuscript. This is plain copy paste. Few of them are aware that they need to restate in their own words the contents of the material. Students are also fond of copying parts of a text without paraphrasing the sentence being borrowed. This limitation has pedagogical implications. The absence of teaching modules on proper documentation has forced students to resort to plagiarism. For the teachers, the dependence on printed textbooks and the little efforts to use online journal sources have contributed to plagiarism of their students.

With today's technology accessible by most people and with online freely accessible sources within reach by the click of a mouse, it becomes imperative that schools resort to software technology to identify errors and correct them before manuscripts are considered for publication. To date, Liceo de Cagayan University is the only school in Mindanao that implements the mandatory testing of scientific articles for plagiarism and grammatical errors for theses and faculty research. Students are coached by the research center staff and by their research advisers on how to interpret and implement the corrections based on the generated reports. As practiced, a paper is tested only when the owner is present and witnesses the whole process. Immediately, the coaching starts and the students return for retesting, to meet the standard of at least 80 percent or even higher.

Initially I discovered that there seems to be a pattern that students who have high scores in plagiarism also turn out to have higher scores in grammarly. Meaning, the

reduction of errors in grammarly could be due to the copy pasting of materials in the text. Hence, this study attempts to look into the correlation between grammar errors and plagiarism.

FRAMEWORK

The use of grammar checker has been proven efficient and effective in detecting grammatical errors in various languages such as Czech (Holan, Kubon & Platek 1997), Norwegian (Johannessen, Hagen & Lane 2002), Swedish (Domeij, Knutsson, & Eklundh, 2002; Arppe 2000), Brazilian Portuguese (Martins et al., 1998), and English (Adriaens, Schreors, 1992; Bernth, 1997; Liou, 1991). Grammarly software has been used as indispensable tool in cleansing scientific manuscripts (Majhi et al., 2013; Clifton 2013; Narita, 2012; Kosta, 2011; Aldrete, 2011).

Grammar checker such as Granske, which is Swedish, combines probalistic and web-based methods to achieve high efficiency and robustness (Domeij et al., 2002). The system adopts grammar error rules to detect Swedish grammar problems and provides options for the corrections.

In Norway, Johannessen, Hagen and Lane (2002) explained that the Norwegian Grammar Checker (NGC) uses a Constraint Grammar with three main parts: 1) a morphological analyzer, which provides each word form with all its lexically possible readings; 2) a morphological CG drawn disambiguator which eliminates incorrect tags according to the grammatical context; and 3) an error detection that identifies different kinds of grammatical errors.

The grammar checker for Czech (Holan, Kubon & Platek, 1997) has three features of its architecture system: (1) morphological and lexical analysis; (2) grammar checking (extended variant of syntactic passing), and (3) evaluation. The core of the system is the second, grammar checking phase.

In Taiwan, a computer software package, Grammatik IV (1989) was used for English as Foreign Language (EFL) students (Liou, 1991). The system works using these components: (1) error analysis and categorization; (2) electronic dictionary; (3) passing and error detection; and (4) feedback.

The Arabic Gram Check (Shaalan, 2005) is a grammar checker program for Modern Standard Arabic designed to help the average user by checking his/her writing for particular common grammatical errors. It describes the problem and gives suggestions for improvement. The use of this checker increases productivity and improves the quality of the text for anyone who writes in Arabic.

Grammarly is known as the world's most accurate American English grammar checker. It has the power to detect 250 grammatical errors and provides superscript

numbers to guide the reader in finding the error annotations after the paragraph. It provides suggestions. It gives data on the overall grammatical competence score and counts of occurrences of each of the grammatical errors detected. After the corrections are entered, another testing is required to show increase in the grammar score and decreases in the error counts as detected. A user may adopt a standard cut-off score for competence. In Liceo de Cagayan University, a score of 80 is required or even higher.

This study posits that grammatical errors significantly influence the occurrence of plagiarism. Writers who have good grammatical competence are more inclined to commit plagiarism.

Mason (2009) wrote an editorial article on plagiarism in scientific publications. He invokes the US Federal Government definition for research misconduct as the "fabrication, falsification or plagiarism in proposing, performing, in receiving research or in reporting research results". Plagiarism is the deliberate or reckless use of someone else's thoughts, words or ideas as one's own, without clear attribution of their source. He argues that occasional use of a phrase or sentence that repeats word for word, a phrase or sentence published elsewhere is the main problem of plagiarism. He states that editors must determine if the copying of a material constitutes intentional and serious breach of the validity of the paper.

To help curb plagiarism as a form of academic dishonesty, Smith, Dupre and Mackery (2007) suggested using a database to deposit all student papers. It shall be easy to detect copying from any paper found in the database.

Thompson and Stobart (2002) suggest three approaches to counter plagiarism in academic institutions: 1) educate the students to avoid plagiarism; 2) have staff introduce anti-plagiarism strategies with regard to arguments, and 3) deter students from plagiarism by staff being effective in its detection. They further present approaches to combat plagiarism: 1) give lecture sessions on how to cheat and get caught; and 2) use of peer groups reviewing process to help deter and detect plagiarism. They concluded these approaches really worked and gave significant results in reducing plagiarism.

The use of technology on plagiarism prevention and detection has some limitations as found by Beasley (2004). He says that detection technology based on matching text from previous papers or publishing in accessible Internet has limitations: (1) web sources may have disappeared between where they were cited in the paper and where the pages was checked; (2) a web source used in a paper is a new addition to the web or in part of the web not searched by the algorithm; (3) Algorithm cannot access all subscription sources on the web. Only a small fraction of that tool in the Internet

content can be analyzed. The deep web is so much more (450 times larger) than the surface web.

OBJECTIVES OF THE STUDY

The study aimed to 1) determine the effectiveness of Grammarly software in reducing the occurrence of grammatical errors in research; 2) determine the effectiveness of plagiarism detector accumulator software in reducing the occurrence of plagiarism in research; and 3) determine the relationship between grammar errors and occurrence of plagiarism.

METHODOLOGY

The study used the purposive sampling design involving 10 nursing, 12 Medtech, 10 IT and 15 Pharmacy with complete test results or a total of 47 as subjects of the study.

As part of the school requirement, students submitted their work for tests using plagiarism detector accumulator and grammarly software during their final oral defense.

The plagiarism software measures originality, referenced and plagiarism and provides a lead towards finding out the sources of plagiarism.

The Grammarly software is touted as the most accurate grammar checker in the world today. Grammarly measures four components: contextual spelling check (spelling, and commonly confused words), grammar (use of articles, incorrect use of prepositions, pronoun agreement, comparing two or more things, faulty parallelism, confusing modifiers, subject and verb agreement, verb agreement, verb form use, use of adjectives and adverbs, conditional sentences, sentence structure, incorrect use of negatives, modal verb, use of conjunction, use of nouns, use of qualifiers and quantifiers, wordiness, passive voice use, others) , punctuation (punctuation's within a sentence, capitalization, and clarity in punctuation), and style and word choice (vocabulary use and writing style).

The study used primary data from the test results of the online software. The two tests are standardized and were purchased online. The statistical techniques used were the frequency count, mean, percentage, mean, standard deviation, t test and the use of Pearson Product Moment Coefficient of Correlation. The data were processed by the University Data Analyst using Minitab Software.

RESULTS AND DISCUSSION

Table 1. Mean and standard deviation of grammatical errors of students detected by grammarly software before and after grammar interventions

Grammar Mean Scores	BEFORE		AFTER		T	P-value	DF
	Mean	Sd	Mean	Sd			
Commonly confused words	5.45	6.18	2.74	4.36	2.45	0.016	82
Use of articles	13.9	11.4	2.09	3.97	6.69	0.000	56
Pronoun agreement	1.21	1.43	0.426	0.617	3.47	0.001	62
Use of adjectives and adverbs	1.55	1.64	0.81	1.35	2.41	0.018	88
Faulty parallelism	1.34	1.58	0.596	0.901	2.81	0.006	73
Confusing modifiers	4.23	3.63	1.91	2.61	3.56	0.001	83
Subject and verb agreement	4.34	3.96	1.17	1.88	4.95	0.000	65
Conditional sentences	0.255	0.441	.128	0.337	1.58	0.118	86
Sentence Structure	4.55	7.05	3.23	6.27	0.96	0.340	90
Incorrect use of prepositions	0.96	1.25	0.362	0.819	2.73	0.008	79
Comparing two or more things	1.40	1.39	0.74	1.19	2.47	0.015	89
Verb form use	1.55	2.07	0.49	1.21	3.04	0.003	74
Verb agreement	.489	.997	.149	.360	2.20	0.032	57
Incorrect use of negatives	0.085	.458	.043	.292	0.54	0.593	.78
Modal Verbs	.362	.870	.191	.537	1.14	0.258	.76
Use of Conjunction	0.38	1.15	0.064	0.438	1.77	0.081	58
Use of Nouns	0.128	.337	0.021	0.146	1.98	0.052	62
Use of Qualifiers and Quantifiers	0.191	.398	0.064	0.247	1.87	0.65	76

Wordiness	6.3	11.1	4.53	8.29	0.88	0.379	85
Passive Voice Use	31.8	65.5	29.6	53.7	0.18	0.858	88
Punctuation within a sentence	21.3	15.6	9.2	12.8	4.10	0.000	88
Capitalization	4.36	5.12	1.57	3.85	2.980	0.004	85
Vocabulary use	6.13	8.47	3.62	5.89	1.67	0.099	82
Writing style	2.30	4.67	1.45	3.46	1.00	0.319	84
Spelling	54.2	59.8	10.9	31.8	4.39	0.000	70
Mean Grammar Score	65.7	11.3	86.3	13.8	-7.93	0.000	88

Data show that majority of the grammatical errors were greatly reduced such as spelling, commonly confused words, use of Articles, pronoun agreement, use of adjectives and adverbs, faulty parallelism, confusing modifiers, subject and verb agreement, conditional sentences, sentence structure, incorrect use of prepositions, comparing two or more things, verb form use, verb agreement, punctuation within a sentence, and capitalization. This happened because the students entered the right corrections to the requirements of grammarly. In grammarly, superscript numbers are placed in every error committed and a notation is given at the end of the paragraph. This allows the coach and the student to agree on the best answer required by grammarly.

On the other hand, incorrect use of negatives, modal verbs, use of conjunctions, use of nouns, wordiness, passive use of voice, vocabulary use and writing style did not reduce significantly.

Table 2. Mean and standard deviation of originality and plagiarism of students detected by the software's before and after coaching interventions

Mean Scores	Before		After		T value	P value	DF
	Mean	SD	Mean	SD			
Originality	93.86	8.84	95.89	7.01	-1.23	0.220	87
Plagiarism	6.07	8.87	4.11	7.01	1.19	0.236	87

In Table 2, occurrence of plagiarism reduced but not statistically significant after the coaching interventions. It was found out that the sources of plagiarism were the

links coming from yahoo and Google which constituted gray literature. It was only during the coaching sessions that the students knew of scientific sources such as Google Scholar, Directory of Open Access Journals, Philippines Electronic Journals, Thomson Reuters, Scopus Elsevier, Asia Journals Online, iamure.com, asianscientific journals, and Philippine Journals Online, among others. Before, students search in non-scientific websites. Now, they search in research databases. In Google Scholar, the students were taught to click "cite" at the bottom right of every source to get the bibliographic list. They were instructed to key in the topic, geographic location and year of publication in Google Scholar and to select at least 10 sources from every continent.

Another reason for plagiarism is the failure to paraphrase or rephrase a statement but sticking to the same idea. Students simply copy pasted the text. The detector sees this as plagiarism. Copying in excess of three lines is also plagiarism. Using previous work by the author is self-plagiarism. The limitation of the detection is that it cannot detect plagiarism of printed sources which do not have online access. The technology compares the text with some 20 billion information in the web and searches for a match of phrases, sentences and paragraphs. The detection provides the links to find the sources.

Some 26 grammatical errors were detected by Grammarly software. The overall picture shows that the grammar mean score (65.7) rose up to 86.3 in the final score after appropriate coaching interventions were done by the Research and Publication Staff supported by the teacher advisers. The difference of 20.60 is significant at .001 levels with a t value of -11.94. The increase is 31.35 percent, which is close to 30 percent average. Thus, without Grammarly testing, these errors could not have been detected and could have been carried out to the print version of the journal.

Table 3. Relationship between grammar indicators and originality

Grammar Indicators	Originality	
	Correlation	P value
Contextual spelling check	-0.055	0.714
Grammar	-0.271	0.065
Punctuation	-0.214	0.148
Writing Style and Word Choice	-0.334	0.022

In Table 3, writing style and word choice have a low, inverse and significant relationship to originality, indicating that those who write well tended to have low originality.

Table 4. Relationship between grammar indicators and plagiarism

Grammar Indicators	Plagiarism	
	Correlation	P value
Contextual	0.037	0.085
Grammar	0.261	0.077
Punctuation	0.215	0.147
Writing Style and Word Choice	0.331	0.023

In Table 4, writing style and word choice has a low, direct and significant relationship with plagiarism. This means that students who have very good writing style and word choice tended to commit high level of plagiarism.

The findings imply that when the students' writing is poor, they are using their original ideas. Conversely, writing style and vocabulary use are directly related to plagiarism. This means students who write well and show good vocabulary use are likely to commit plagiarism. This makes sense because when students copy and paste their sources, they carry over also the good writing style and vocabulary use of the authors they plagiarized.

CONCLUSIONS

There is sufficient evidence to show that the coaching interventions reduced the occurrence of grammatical errors by 40 percent as measured twice by grammarly software. Indeed, grammarly is an effective tool to measure grammatical errors.

The coaching interventions also reduced the occurrence of plagiarism but the difference is not significant. Thus, the use of plagiarism and grammar software is effective in helping coaches assist the students in reducing the errors and increasing quality of academic writing.

Occurrence of plagiarism is a critical function of good grammatical skills of the students. Students with good grammatical competence are more likely to commit plagiarism.

RECOMMENDATIONS

An academic integrity program needs to be put in place that focuses on policies and programs that deter plagiarism and improve grammatical skills of university students.

LITERATURE CITED

Adriaens, G., & Schreors, D.
(1992, August). From COGRAM to ALCOGRAM: Toward a controlled English grammar checker. In *Proceedings of the 14th conference on Computational linguistics-Volume 2* (pp. 595-601). Association for Computational Linguistics. Retrieved August 24, 2013 from http://dl.acm.org/citation.cfm?id=992163

Aldrete, J. A.
2011 Plagiarism and Other Literary Missteps in Medicine and Particularly in Anesthesia. *Revista Colombiana de Anestesiología*, *39*(2), 217-229. Retrieved August 24, 2013 from http://www.scielo.org.co/scielo.php?pid=S0120-33472011000200006&script=sci_arttext&tlng=en

Arppe, A.
(2000, December). Developing a grammar checker for Swedish. In *The 12th Nordic Conference of Computational Linguistics* (pp. 13-27). Retrieved August 24, 2013 from Arppe, A. (2000, December). Developing a grammar checker for Swedish. In *The 12th Nordic Conference of Computational Linguistics* (pp. 13-27).

Beasley, J. D.
(2004, June). The impact of technology on plagiarism prevention and detection: Research process automation, a new approach for prevention. In *Proc." Plagiarism: Prevention, Practice and Policies 2004": Joint Information Systems Committee Plagiarism Advisory Service Conference* (pp. 28-30). Retrieved August 24, 2013 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.134.2382&rep=rep1&type=pdf

Bernth, A.
1997 Easy English: A Tool for Improving Document Quality. The Proceedings of the Fifth Conference on Applied Language Processing, Washington, 159-165. Retrieved August 24, 2013 from http://dl.acm.org/citation.cfm?id=974581

Clifton, S.
Integrating Technology in the Classroom. Retrieved August 24, 2013 from http://cliftonwowlearning2.pbworks.com/w/file/fetch/60779540/Integrating%20Technology%20in%20the%20Classroom%20workbook.pdf

Domeij, R., Knutsson, O., & Eklundh, K. S.
2002 Different Ways of Evaluating a Swedish Grammar Checker. In *LREC*. Retrieved August 24, 2013 from http://gandalf.aksis.uib.no/lrec2002/pdf/180.pdf

Holan, T., Kuboň, V., & Plátek, M.
(1997, March). A prototype of a grammar checker for Czech. In *Proceedings of the fifth conference on Applied natural language processing* (pp. 147-154). Association for Computational Linguistics. Retrieved August 24, 2013 from http://dl.acm.org/citation.cfm?id=974579

Johannessen, J. B., Hagen, K., & Lane, P.
(2002, August). The performance of a grammar checker with deviant language input. In *Proceedings of the 19th international conference on Computational linguistics-Volume 2* (pp. 1-8). Association for Computational Linguistics. Retrieved August 24, 2013 from http://dl.acm.org/citation.cfm?id=1071894

Kosta, M. G.
2011 Computer Assisted Language Learning. Retrieved August 24, 2013 from http://works.bepress.com/marygkosta/1/

Liou, H. C.
1991 Development of an English grammar checker: A progress report. *CALICO journal*, *9*(1), 57-70. Retrieved August 24, 2013 from http://journals.sfu.ca/CALICO/index.php/calico/article/view/424

Majhi, A., Sharma, Y. K., & Naik, D. V.
2012 Blending optimization of Hempel distilled bio-oil with commercial diesel. *Fuel*, *96*, 264-269. Retrieved August 24, 2013 from http://www.sciencedirect.com/science/article/pii/S0016236112000671

Martins, Ronaldo Teixeira, et al.
"Linguistic issues in the de elopment of ReGra: A grammar checker for Brazilian Portuguese." *Natural Language Engineering*4.4 (1998): 287-307. Retrieved August 24, 2013 from http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=48459

Mason, P. R.
2009 Plagiarism in scientific publications. *J Infect Dev Ctries*, *3*(1), 1-4. Retrieved August 24, 2013 from http://scholar.google.com.ph/scholar?hl=en&q=plagiarism+in+scientific+publication&btnG=&as_sdt=1%2C5&as_sdtp=#

Naber, D.
2003 A rule-based style and grammar checker. Retrieved August 24, 2013 from http://www.grin.com/en/e-book/108379/a-rule-based-style-and-grammar-checker

Narita, M.
2012 Developing a corpus-based online grammar tutorial prototype. *LANGUAGE TEACHER*, *36*, 4. Retrieved August 24, 2013 from http://scholar.google.com.ph/scholar?q=narita+grammar+checker&btnG=&hl=en&as_sdt=0%2C5#

Shaalan, K. F.
2005 Arabic GramCheck: A grammar checker for Arabic.*Software: Practice and Experience*, *35*(7), 643-665. Retrieved August 24, 2013 from http://onlinelibrary.wiley.com/doi/10.1002/spe.653/abstract

Smith*, M. W., Dupre, M. E., & Mackey, D. A.
2005 Deterring research paper plagiarism with technology: Establishing a department-level electronic research paper database with e-mail. *Journal of Criminal Justice Education*, *16*(1), 193-204. Retrieved August 24, 2013 from http://www.tandfonline.com/doi/full/10.1080/1051125042000333

550#.UhgRttJHKdc

Thompson, J., & Stobart, S.
2002 University research, plagiarism and the internet: problems and possible solutions. *Ethicomp 2002*. Retrieved August 24, 2013 from http://archive.plagiarismadvice.org/documents/resources/ethicomp2002paper.pdf