International Journal of English Language Studies
ISSN: 2707-7578
DOI: 10.32996/ijels
Journal Homepage: www.al-kindipublisher.com/index.php/ijels

IJELS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Grammatical Deviations in Philippine Phishing Emails

Julysa C. Cardona
Assistant Professor, Language and Letters Department, Bukidnon State University, Malaybalay City, Bukidnon, Philippines
Corresponding Author: Julysa C. Cardona, **E-mail**: julysacardona@buksu.edu.ph

| ABSTRACT
Just as language is proven to be useful in making day-to-day life convenient, some situations have also demonstrated the possibility of language being used to deceive people. One instance is the proliferating use of phishing emails. Given this occurrence, a riveting endeavor in the concept of actual language utilization is the study of grammatical deviations in phishing emails. Employing error analysis, this study sought to determine how a grammar study can be helpful in the examination of the imposters' language, specifically in the case of phishing emails. The purpose of this research was to document the dominant errors that appeared in Philippine phishing emails and to explain how grammatical deviations can give away deception. The result of this research showed that of the fifteen collected phishing emails, all or 100% contained usage errors. This means that of the fifteen phishing emails in this study, each one has errors. The most frequent of these errors are errors in capitalization, punctuation, and word forms. This result implies that although imposters pretend to be legitimate, they cannot imitate and copy the language of authentic emails. This may be because when legitimate institutions and organizations like banks, schools, or establishments release official emails, they do so after thorough proofreading and editing. Phishers, however, may not have the same mechanisms to ensure grammar correctness and accuracy. Based on these findings, the researcher infers that with good grammar skills, one can have a larger inclination to distinguish phishing from genuine emails. Contrariwise, those who do not have profound knowledge of grammar conventions may have a larger possibility of falling into the phishing trap.

| KEYWORDS

Grammatical deviation, language correctness and errors, phishing emails, imposters.

| ARTICLE INFORMATION

ACCEPTED: 31 May 2024 **PUBLISHED:** 05 June 2024 **DOI:** 10.32996/ijels.2024.6.2.18

1. Introduction

Language plays an integral role in people's daily lives. With language, one can produce and receive messages and ultimately participate in communication situations. For official transactions, language also plays a crucial part in financial undertakings such as sales marketing, shopping, and banking. There is no doubt that language is a powerful tool for any type of human activity (Racoma, 2013). Just as language is undoubtedly proven to be powerful and useful in making life convenient, some instances, however, have also demonstrated the possible means by which language is used by imposters to deceive, steal, or commit a crime, and one widespread occurrence is that of phishing emails (Nlebedum, 2017).

Phishing emails are electronic messages sent to a large number of people from websites that pretend to be legitimate in an attempt to steal people's money, identity, or private and sensitive information such as credit card numbers, bank information, or passwords. It is among the goals of phishing attacks to look as if it is from an authentic organization in the hopes that someone will click on the link and provide their personal information or download malware (Giandomenico, 2020). Taking these definitions of phishing emails, it can be deduced that phishers are imposters who pretend to be someone they are not. They claim to be legitimate banks, organizations, or institutions. Although these types of deceit may be harmless for people who immediately ignore

these phishing emails, these dishonesties may harm other vulnerable victims because of money loss and identity theft (Baykara & Gürel, 2018).

Given the idea that language can be used by imposters in deceitful activities, one riveting field of study may be the examination of grammatical deviations in deceitful messages like phishing emails. Grammatical deviation refers to the breaking of the rules in the formation of words and sentences (Budiharto, 2016). This definition implies that grammatical deviations do not demonstrate language correctness when measured against the prescriptive rules of English. In the context of prescriptive grammar, language correctness refers to the notion that certain words, word forms, and syntactic structures meet the standards and conventions or the rules prescribed by traditional grammarians. When a text (either written or oral) displays an instance of faulty, unconventional, or controversial usage, like a misplaced modifier or an inappropriate verb tense, then such a case can be described as a usage error or a grammatical error (Nordquist, 2019; 2020). In this study, grammatical deviations in phishing emails were to be treated as errors since business emails are supposed to reflect a certain degree of grammatical accuracy. In light of this goal, the following were the objectives of this paper:

a) Document the grammatical deviations in the phishing emails in the Philippines
b) Explain how grammatical errors can give away deception in the context of phishing emails

2. Literature Review

Frost (2012) has argued that many people misspell common words, confuse similarly spelled words like "it's" and "its," or sometimes simply use the incorrect forms of words such as "there, their, and they're." More so, she enumerated the common types of grammatical errors such as subject-verb disagreement, mixing up the past and present tenses, apostrophe errors, failure to put a proper ending on a past tense verb, misuse of commas, and misplaced modifiers. In a similar vein, Amiri and Puteh (2017) also presented the common errors in academic writing. These errors are classified under sentence structure, articles, punctuation, capitalization, word choice, prepositions, verb form, singular/plural noun ending, redundancy, word form, subject-verb agreement, word order, possessive, and verb tense.

Although grammar correctness is not treated as the primary indicator of excellence, people require grammar to communicate efficiently since language cannot function without it (Norquist, 2020). In the academe, grammar skills can be one way to show learning and progress, which is why some studies have focused on identifying the grammatical errors present in students' writing (Limengka & Kuntjara, 2013; Royani & Sadiah, 2019; Alghazo & Alshraideh, 2020). Their studies have shown that students' writing contained various grammatical errors such as verb agreement, capitalization, usage, sentence patterns, pronoun, spelling, addition, omission, misformation, misordering, blends, verb error, article errors, wrong word, noun ending errors, and sentence structures. Based on these results, it can be deduced that grammar remains a difficult thing to master even in the academe. This paper, however, attempts to bring a new insight into which grammar can be viewed. Since some authors have acknowledged that spelling or grammatical irregularities are among the cues that essentially distinguish between phishing and genuine emails (Parsons, Butavicius, Pattinson, Calic, Mccormac, & Jerram, 2016; Diaz, Sherman & Joshi, 2020; Irwin, 2022), the researcher of this present work contends that the study of grammatical errors is not only boxed in the academe but may also be useful in the examination of the legitimacy or falsity of a particular message. Hence, the goal of this present work is to bring insight into how grammar studies can be helpful in the examination of the language of imposters, specifically in the case of phishing emails.

The detection of phishing emails has received considerable research efforts over the last few years. In a more technical spectrum of phishing email studies, some authors have developed certain software to aid in the detection of fraud messages. In one study, Baykara and Gürel (2018) developed a software called the "Anti Phishing Simulator." With their software, phishing, and spam emails are detected by examining mail contents. It must be noted, however, that in the earlier work of Park, Stuart, Taylor, and Raskin (2014), the need for human and computer competencies to complement each other in the detection of phishing emails has long been indicated. Hence, although some computers and gadgets have been installed with spam email detectors as a product of modern-day technology and innovations, the need for human beings to develop the skill to manually detect deception and lies cannot be underestimated.

Brooks (2018) has previously recognized the importance of being able to manually recognize when an email is genuine or not. Hence, she explored the language of persuasion in phishing emails through the lens of speech acts. Eight years before this work of Brooks (2018), Chiluwa (2010) had already conducted the same study, which revealed that the commissive act is used as a persuasive strategy in hoax emails through unrealistic and suspicious promises, while the directive act is used to impart urgency in the receiver to act promptly. Both studies by Brooks (2018) and Chiluwa (2010) have accentuated the fact that phishing emails can be examined by looking at the discursive function of the language of imposters. The researcher of this current work, however, contends that other than looking at the pragmatics in the language of the phishers, the lies of the imposters behind the phishing emails can also be brought into the open through the use of error analysis.

In a qualitative study by Blythe, Petrie & Clark (2011), they posited that spelling and grammar could not be a reliable cue to give away phishing since 38% of the texts were spelled appropriately, and 68% looked convincing as they contained logos indistinguishable from the authentic online article. This is interesting because this opposed the contentions of Parsons, Butavicius, Pattinson, Calic, Mccormac, & Jerram (2016), Sherman & Joshi (2020), and Irwin (2022), who acknowledged that spelling or grammatical irregularities are among the cues that essentially distinguish between phishing and genuine emails. These differing views imply that the cues that give away a phishing attack may vary from one context to another. In the work of Blythe, Petrie & Clark (2011), the texts used were phishing emails in Canada, where English is the first language. Hence, the phishers may have proficiency in the language since it was their first language to begin with.

The review of the literature discloses that the examination of grammar issues has proven to be useful not only in the field of education but also in other fields. Since language is also used in acts of deception by imposters through phishing attacks, a study on language correctness and grammatical deviations may also be carried out to reveal the deception behind phishing emails. The review further reveals that there was a plethora of investigations on phishing emails through the lens of software development and textual and discourse analysis, but not many phishing email studies were done through the lens of error analysis. Moreover, the review also shows that authors have varying notions on whether or not grammatical deviations can be used to reveal a phishing attack, underscoring the need for a contextualized study to be conducted. Hence, this paper aims to bring the study of grammar errors outside the halls of the academe and instead identify the grammatical errors in phishing emails in the Philippines. More so, this paper aims to explain how these errors may give away deception, which can be a significant contribution to the efforts to detect phishing attacks.

3. Methodology

This study employed qualitative research, particularly error analysis. Ellis and Barkhuizen (2005) defined error analysis as a set of procedures used to identify, describe, and explain errors in a language. Simply put, error analysis deals not only with the identification and the detection of errors but also with explaining the reason for such error occurrences. Ellis and Barkhuizen (2005) further outlined the process of error analysis in four ways. These are the (a) collection of a sample, (b) identification of errors, (c) description of errors, and (d) explanation of errors. These steps are similar to the steps in error analysis laid by Corder (1974), which includes the (a) examination of the text word for word and sentence by sentence and (b) counting and converting the errors into a percentage to examine the occurrence.

To gather and analyze the data, a combination of the steps of error analysis specified by Corder (1974) and Ellis and Barkhuizen (2005) was followed in this study. First, the phishing emails were collected, and each email was examined word for word and sentence by sentence. Second, the errors were then identified, and the frequency of errors was counted and converted into a percentage to examine the occurrence. Lastly, the researcher then described and explained the possible reasons for such occurrences.

These steps of error analysis have been useful in this study because the texts used were phishing emails. Since the phishers were pretending to be legitimate banks in the Philippines, such as Landbank of the Philippines (LBP), Bank of the Philippine Islands (BPI), and Banco de Oro (BDO), the language in their emails needed to be free from errors to appear legitimate and convincing, otherwise such errors may give away their lies and deception. The phishing emails pretending to be from LBP, BPI, and BDO were chosen since these were the ones that appeared most frequently on social media platforms when looking up the terms "email scams," suggesting that these were the banks that phishers frequently use in their phishing emails. To ensure that these emails were indeed phishing or scams, the researcher gathered the emails that the victims, targets, or recipients of these scams published on their official social media platforms to warn the public about such deception. To abide by the ethical considerations in this study, the names of these individuals who posted the phishing emails on their public social media platforms were kept confidential and were then not disclosed in this work.

4. Results and Discussion

Fifteen (15) phishing emails were gathered in this study. Examination of these emails revealed that all or 100% of these emails contained usage errors. This means that of the fifteen phishing emails gathered, none of those emails were error-free. Table 1 below shows the types of deviations or errors found across the fifteen phishing emails, as well as the frequency distribution of these types of deviations and the corresponding percentage.

Based on the data the table presents, 60 errors were found across the fifteen phishing emails pretending to be Landbank, BPI, and BDO. These errors were then distributed to nine (9) types, and they were related to capitalization, punctuation, word form, preposition, verb form, spelling, articles or determiners, sentence fragments, and run-on sentences. Of the nine types of errors identified, the three dominant ones are errors in capitalization at 26.67%, punctuation at 25%, and errors in word form at 11.67%. These three categories of errors were discussed one after the other, being the top three dominant errors found in phishing emails.

Table 1: Frequency distribution of the grammatical deviations in phishing emails

Types of Errors	Frequency	Percentage
Capitalization	16	26.67%
Punctuation	15	25%
Word form	7	11.67%
Preposition	5	8.33%
Verb form	5	8.33%
Spelling	4	6.67%
Articles/Determiners	4	6.67%
Sentence Fragment	2	3.33%
Run-on Sentence	2	3.33%
Total	60	100%

The first dominant type of error found in phishing emails is capitalization. It was the most dominant in the sense that more than one-fourth of the errors across the fifteen phishing emails were related to capitalization issues. Frame 1 shows the samples of capitalization errors found in phishing emails.

"Greetings, our valued **C**ustomer,"
"Dear **V**alued **C**lient,"
"You can verify your **A**ccount at…"
"This is to inform you that you have (1) **P**ending transaction…"

Frame 1

Nordquist (2018) explicitly discussed the rules of capitalization in the English language. According to him, to capitalize means to use an upper case for the first letter of a word. There is a need to capitalize the first word in a sentence, the pronoun "I," and the names and nicknames of particular persons and characters. In the shown samples from the phishing emails, however, some words that do not need to be capitalized were capitalized, such as the "c" in the *"Greetings our valued Customer,"* the "v" and "c" in the *"Dear Valued Client,"* the "a" in the *"You can verify your Account at…"* and the "p" in the "*This is to inform you that you have (1) Pending transaction…"*. The words *customer, valued, client, account,* and *pending* did not appear at the beginning of the sentence, and neither are they proper nouns. Hence, capitalizing them in the sentence, as shown in frame 1, makes them deviate from the standard norm in English. One possible reason for this error could be that the phishers were not aware of these rules and conventions concerning proper capitalization.

The second dominant type of error found in phishing emails has something to do with the proper punctuation in English. It was the top two most dominant errors in the sense that one-fourth of the errors across the fifteen phishing emails are related to punctuation issues. Frame 2 shows the samples of punctuation errors found in phishing emails.

"Greetings from BDO Unibank"
"To verify your account, please login in our app."
"Your account must be verified/"

Frame 2

According to Nordquist (2016), punctuation is the set of marks such as ampersands, apostrophes, asterisks, brackets, bullets, colons, commas, dashes, diacritic marks, ellipsis, exclamation points, hyphens, paragraph breaks, parentheses, periods, question marks, quotation marks, semicolons, slashes, spacing, and strike-throughs which used to regulate texts and clarify their meanings, mainly by separating or linking the words, phrases, and clauses. One basic rule involving punctuation is to end a sentence with a period (.), a question mark (?), or an exclamation point (!). In the shown samples from the phishing emails, however, the sentences *"Greetings from BDO Unibank,"* "*To verify your account, please login in our app,"* and "*Your account must be verified/"* did not end with any of these marks. The first sentence, *"Greetings from BDO Unibank,"* should have ended with an exclamation point (!), while the next two sentences should have ended with a full stop or a period. Thus, the lack of proper punctuation in the sentences, as shown in frame 1, makes them erroneous. One possible reason for this error could be that the phishers were not aware of these rules and conventions concerning proper punctuation or that they simply did not pay enough attention to the importance of punctuation in the sentences.

The third dominant type of error found in phishing emails is categorized under word from. Based on the numerical data presented, 11.67% of the total errors across the fifteen phishing emails are related to issues with the word form. Frame 3 shows the samples of word form errors found in phishing emails.

"We urge you to update your record to keep your record **update.**" "…to shield your information and to **Security** your account."

Frame 3

Word form errors happen when the correct word is selected, but an incorrect form of the word is utilized in the sentence. An example of this is the sentence, *"Young people can be independence in the United States."* The correct one should have been, "*Young people can be independent in the United States."* In the incorrect sentence, the word "independence" is a noun, and it cannot be used to describe young people. To solve this, the adjective form "independent" was used instead. The same is true for the sentences in frame 3. In the first sentence, *"We urge you to update your record to keep your record* ***update",*** the word update is in the form of a verb. However, the word "record" is the one being described. Hence, the adjective *"updated"* should have been used to modify the noun *"record"* instead. In the second phrase, "…to shield your information and to **Security** your account", the word *"security"* is a noun, but since it follows the infinitive *"to",* the verb form *"secure"* must have been used instead. One possible reason for this error could be that the phishers were not aware of these rules and conventions concerning proper word form or that they simply did not know that there were errors in the word form they used.

Being a country that treats English as a second language, it is common for people in the Philippines to commit errors in their actual language use, especially when speaking. Emails, however, are a different story, especially those that are supposed to come from legitimate institutions and organizations like banks. Unlike oral speeches, written texts like emails can undergo editing, proofreading, and finalizing before they are released to the intended targets. Some institutions like banks may have already prepared generic email templates that can be easily filled with personalized details to suit the intended recipient. This means that the emails of authentic and legitimate banks have a few errors if not completely none.

Phishing emails have to keep up with the quality of authentic emails. Since phishers are pretending to be legitimate, the quality of their emails has to reflect professionalism and authority; otherwise, they will appear as a copycat. Considering that there are sixty errors found across the fifteen phishing emails gathered, the result of this study implies that the errors in the phishing emails are not mere honest mistakes that the authors have overlooked while writing; instead, these errors were a reflection of the phishers' linguistic capacity. This goes to imply that although imposters pretend to be legitimate, they cannot imitate and copy the language of authentic emails. Unlike authentic emails, phishing emails contain an apparent amount of grammar deviations or errors. These errors then give away signs of deception. These findings concurred with the previous arguments of Parsons, Butavicius, Pattinson, Calic, Mccormac, and Jerram (2016), Sherman and Joshi (2020), and Irwin (2022), who acknowledged that spelling or grammatical irregularities are among the cues that essentially distinguish between phishing and genuine emails.

5. Conclusion

Generally, this study revealed that, unlike genuine emails, phishing emails contain grammatical deviations or errors, and the most frequent ones are related to capitalization, punctuation, and word forms. This result implies that although imposters pretend to be legitimate, they cannot imitate and copy the language of authentic emails. This may be because when legitimate institutions and organizations like banks or institutions send out emails, they do so after thorough proofreading and editing to make sure their emails are flawless or at least close to perfection. Phishers, however, may not have the same mechanisms to ensure grammar correctness and accuracy. Hence, the grammatical errors in their phishing emails may become cues that give away their intent to deceive.

Based on these findings, the researcher of this study infers that the Philippine phishers lack grammar proficiency, considering that the capitalization and punctuation errors in their phishing emails are basic concepts taught spirally across Philippine education systems. Additionally, the researcher also infers that with good grammar skills, one can have a larger chance of distinguishing phishing from authentic emails, subsequently lowering the chance of getting deceived. On the contrary, those who do not have profound knowledge of grammar conventions may have a larger possibility of falling into the phishing trap. However, there is a need to point out that this study is limited to analyzing the grammar errors found in phishing emails. The texts used were the ones posted on the social media platforms by the recipients of these emails, and these email recipients were not interviewed to verify the role of grammar errors in their detection of these phishing emails. Considering this limitation, the study recommends a

qualitative endeavor where the actual recipients of phishing emails were interviewed and asked about the factors that caused them to point out the deceit in the phishing emails they received.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Alghazo, K.M & Alshraideh, M.K. (2020). Grammatical errors found in English writing: A study from Al-Hussein Bin Talal University. Research Gate.
[2] Amiri, F. & Puteh, M. (2017). Error analysis in academic writing: a case of international postgraduate students in Malaysia. *Advances in Language and Literary Studie*s. ISSN: 2203-4714
[3] Baykara, M. & Gürel, Z. Z. (2018). Detection of phishing attacks. 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 2018, pp. 1-5, doi:10.1109/ISDFS.2018.8355389.
[4] Blythe, M., Petrie, H., & Clark, J. A. (2011, May). F for fake: Four studies on how we fall for phish. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 3469-3478).
[5] Brooks, H. S. (2018). Linguistic persuasion techniques in phishing emails: A corpus and critical discourse analysis. Hofstra University ProQuest Dissertations Publishing
[6] Budiharto, R. A. (2016, January). Language deviations in a popular novel: An alternative way to teach morphology and phonology for English Department Students of Madura University. In Proceeding of International Conference on Teacher Training and Education (Vol. 1, No. 1).
[7] Corder, S. P. (1974). Error analysis. In J. P. B. Allen and S. P. Corder (eds.) Techniques in Applied Linguistics (The Edinburgh Course in Applied Linguistics: 3). London: Oxford University Press (Language and Language Learning), pp 122-154.
[8] Diaz, A., Sherman, A.T. and Joshi, A. (2020). Phishing in an academic community: A study of user susceptibility and behavior. arXiv: 1811.06078v1 [cs.CR]
[9] Ellis, R. & Barkhuizen, G. B. (2005). Analysing learner language. Oxford University Press. UK.
[10] Frost, J. (2012). The 12 most common grammar errors. Retrieved from https://www.grammarcheck.net/the-12-twelve-most-common-grammar-errors/
[11] Giandomenico, N. (2020). What is spear-phishing? Defining and differentiating spear-phishing from phishing. https://owasp.org/www-chapter-ghana/assets/slides/OWASP_Presentation_FINAL.pdf
[12] Irwin, L. (2022). 5 ways to detect a phishing email – with examples. IT Governance. https://www.itgovernance.co.uk/blog/5-ways-to-detect-a-phishing-email
[13] Limengka, P.E & Kuntjara, E. (2013). Types of grammatical errors in the essays written by fourth-semester students of English Department, Petra Christian University. Siwalankerto 121-131, Surabaya, East Java, Indonesia
[14] Nlebedum, C.J. (2017). Dear valued customer: A forensic-linguistic analysis of scam texts. Academia. hhttps://www.academia.edu/37276641/Dear_Valued_Customer_A_Forensic_Linguistic_Analysis_of_Scam_Texts
[15] Nordquist, R. (2019). Definition and examples of correctness in language. Retrieved from https://www.thoughtco.com/correctness-grammar-and-usage-1689807
[16] Nordquist, R. (2020). English grammar: Discussions, definitions, and examples. Retrieved from https://www.thoughtco.com/what-is-grammar-1690909
[17] Nordquist, R. (2020). What is a grammatical error? Retrieved from https://www.thoughtco.com/grammatical-error-usage-1690911
[18] Park, G., Stuart, L.M., Taylor, G.M. & Raskin, V. (2014). Comparing machine and human ability to detect phishing emails. IEEE International Conference on Systems, Man, and Cybernetics (SMC), San Diego, CA, USA, pp. 2322-2327, doi: 10.1109/SMC.2014.6974273.
[19] Parsons, K., Butavicius, M., Pattinson, M., Calic, D., Mccormac, A. & Jerram, C. (2016). Do users focus on the correct cues to differentiate between phishing and genuine emails? Cornell University. https://doi.org/10.48550/arXiv.1605.04717
[20] Racoma, B. (2013). Language as a powerful tool continues to grow. https://www.daytranslations.com/blog/power-of-language/
[21] Royani, S. & Sadiah, S. (2019). An analysis of grammatical errors in students' writing descriptive text. Research Gate.