Journal of Namibian Studies, 33 S1 (2023): 1624-1645 ISSN: 2197-5523 (online)

Quality Assessment Of Automatic Paraphrasing Tool For English: An Analysis At Lexical Level

Zala Wahab[1] , Liaqat Iqbal[2] , Irfan Ullah[3] , Rafiq Nawab[4] , Ubaid Ullah Ubaid[5]

[1]MPhil Scholar, Department of English, AWKUM.
[2]Associate Professor, Department of English, AWKUM.
[3]Assistant Professor, Department of English, AWKUM.
[4]Assistant Professor, Department of English, AWKUM.
[5]The Oxford School Dubai, UAE.
Corresponding author email: liaqat@awkum.edu.pk

Abstract

Automated paraphrasing is considered an important and emerging educational phenomenon that has a key role in academia. Paraphrasing tools facilitate students as well as professionals' work. However, there is a need to assess and evaluate the quality of automated paraphrasing by such paraphrasing tools. This study aims at evaluating the quality of automated paraphrasing carried out by one of such tools, that is, QuillBot, by exploring the lexical relatedness and differences in the original and the paraphrased text. Hence, the limitations of automated paraphrasing tools are highlighted. In this regard, QuillBot is used for paraphrasing as it is considered a standard paraphrasing tool by many. Automated paraphrasing of literary and non-literary text carried out by QuillBot is used as data. AntConc, a corpus tool, is used for exploring the lexical features of the data. In addition, word relatedness is also computed from HSO measure through the WS4J demo, an online WordNet ontological tool. Differences were found at the lexical level between the original and paraphrased text. Automated paraphrasing carried out by QuillBot of the non-literary text is more aligned with the original text than that of the literary text. In the lexical category, mostly content words were substituted with synonymous words. The morphological structure of words was also modified. At times, these alterations would obscure the original meaning, while on other occasions, they would enhance and refine it.Therefore, automated paraphrasing should not be taken for granted instead, it should be subject to manual review and revisions. In addition to automated paraphrasing of literary text carried out by such tools, it should be rechecked and updated manually while standard tools, like QuillBot, could be relied upon for paraphrasing of non-literary text.

Keywords: Lexical Analysis, Automated Paraphrasing, QuillBot, HSO, WordNet.

Introduction

Paraphrasing is rewriting the original text. It is an act of rewriting a segment from a written text in the style of the paraphraser rather than that of the original author (Soles, 2003). Sometimes, paraphrasing is difficult for students as they are required to repeat the original text (O'Reilly, 2012). The limitations to reformulating the meaning content and changing the structure of the sentence

can encourage the use of paraphrasing techniques (Basori, 2017). Furthermore, advancement in technology has made this task easy for the students. Students can access automatic paraphrasing through various computer software.

Paraphrase generation has an important role in many downstream tasks, such as machine translation, question answering, and information retrieval [Hu et al., (2019), as cited in (Guo, Huang, ZHU, Chen, Zha, Chen, Huang, 2021)]. There are various paraphrasing tools available. Paraphrasing tools are rephrasing tools that use specific software for rewriting the original text. Paraphrasing tools are very useful, especially, for the students as it provides instant paraphrasing, save their time and energy as well as enhance their language competence. But, as every coin has two sides, the automated paraphrasing quality lag behind manual paraphrasing. Automated paraphrasing has also certain defects. It sometimes deviates in meaning from the original. Online paraphrasing tools are at risk in terms of the generated output quality (Rogerson, 2017). Some services are lacking in these tools, creating incoherent sentences that, while not directly matching, are almost unreadable (Bailey, 2017).

Various researchers have analysed texts lexically. Phillips (2001), for example, studied to refine, the sound modifications theory. Other researchers also worked with individual parts of speeches, for example, Rosen (1989), in her study, worked on Noun Incorporation (NI). In the same way, Savoy (2010) studied primary features and contrasted the speech's frequent use of popular English vocabulary with everyday language (Brown corpus). Additionally, he contrasted various metrics that could be used to extract words that best described a certain portion of the full text corpus. Moreover, the study of Zhang, Liu, Chen., Zhan., and Yu, (2003), uses a Hierarchical Hidden Markov Model (HHMM) to offer a unified strategy for Chinese lexical analysis, to incorporate segmentation of Chinese words, disambiguation, Speech-parts tagging, and unknown words identification into a single theoretical framework. In word segmentation, they used a class-based HMM, and treated unknown terms the same as frequent words in the lexicon at this level. In role-based HMM, unknown words are reliably identified.

The studies relating to text relevance comparisons mainly focused on natural language processing. Wup (Wu & Palmer, 1994)

measure in WordNet had been used for semantic relatedness measure between texts. They do not use other WordNet measures such as Lesk (1986), HSO (Hirst & St-Onge, 1998), etc. Moreover, they also did not work with paraphrasing tools. Very few research studies have been conducted until now for qualitatively measuring the quality of paraphrasing tools at the word level. As automated paraphrasing is different lexically from the original text, this study aimed at exploring the performance of a standard paraphrasing tool, QuillBot, and the paraphrasing of literary and non-literary carried out by QuillBot. Here the focus is on evaluating the quality of automated paraphrasing carried out by paraphrasing tools where the relevance and disparity between the original and the paraphrased texts at the lexical level are studied.

The data sets used for this study consisted of a small selection taken from English literary and non-literary texts and their automated paraphrases carried out by Quillbot (a standardized paraphrasing tool). Through random sampling technique, for the literary text, the first two chapters of Emma (Jane Austen's novel) were used and for the non-literary text, an article from the DAWN newspaper was taken. These texts were paraphrased and the product was analyzed. The lexical similarity in the source text and the paraphrased text was focused. Moreover, the incongruity in the automated paraphrase was also explored.

The original and paraphrased texts were closely analysed both qualitatively through close observation as well as quantitatively through the online tool, WS4J and AntConc software. For the lexical relevance measuring of both the original and paraphrased texts, the word lexicons categories were found. The morphological construction of words was studied. The parts of speech of various words were calculated through AntConc and in this way, both the texts were compared. Then, how the grammatical categories of words were modified in automated paraphrasing by the tool was observed. The WordNet gave many knowledge based and corpus based measures values but the study focus was on considering only the HSO values for analysing the relatedness.

With the advent of this technological age, many software and tools are used for various purposes without knowing the pros and cons of such tools. In this regard, this study has significance both for the public as well as for the researchers. As far as the general

public is concerned, this study highlights the limitations of automatic paraphrasing tools for them. For this purpose, the quality of automated paraphrasing of literary and non-literary texts is explored adding to the debate that far automatic paraphrasing tools are reliable and what specific issues such tools may have. It also highlights the values of such tools where different text genres are involved, e.g., literary and non-literary text. For researchers, this study opens up new ways of research. Moreover, this study also tries to ensure the developers of various paraphrasing tools enhance the quality of such tools. As a result, these software could be upgraded and more advanced features could be added to them for the users. Further, this study also encourages research into other computer-assisted educational tools. The quality of other tools could be evaluated and hence, upgraded. It could inform students about the usability of various computer-assisted tools available online for them.

Literature Review

Various works have been carried out to compute the lexical relatedness and relevance between texts and various ways and techniques were adopted for finding such relatedness as well as relevance. They compared texts lexically from various dimensions and used various computer-assisted tools and software for text similarity measuring. Phillips (2001), for example, asserts that Bybee, in her book of Morphology, 1985, linked word frequency and lexical analysis. He quotes her saying that the distant lexical connections are formed mostly by High-frequency words than by low frequency words. High-frequency words are less analyzed and less dependent on their related base words (in morphologically complex words) than low-frequency words. He says that suppletive forms like golwent are the best example of this principle. However, the same concept can also be seen in the linguistic process, when common words and phrases like housewife become hussy and day's eye become daisy. Phillips (1998: 231, as cited in Phillips, 2001) used this fact for developing a theory about sound modification lexical diffusion. He says that the super-segmental modifications that necessitate analysis e.g., by part of speech or morphemic element etc. first affect the least frequent words, while changes that ignore grammatical information first affect those words that are most frequent. This remark was a tweak to the earlier Frequency-Actuation Hypothesis (Phillips 1984: 336, as

cited in Phillip 2001), but it kept the portion on segmental alterations intact: “Physiologically motivated sound changes affect the most frequent words first; other sound changes affect the least frequent words first" (Phillips, 2001, p. 123). The goal of his (Philips’) study was to refine this theory, namely, the sound modifications that need analysis that whether syntactic, morphological, or phonological, affects firstly the least common words.

Researchers also worked with individual parts of speeches., for example, Rosen (1989), in her work argued that Noun Incorporation (NI) results from head movement, in which a complex verb is created as the head of an object noun phrase shifts into the verb. Instead, she argues that NI derives, pre-syntactically, from rules of word formation apply in the lexicon. She argued that there are two procedures of word formation available, based on clusters of grammatical properties associated with NI. In one process, the argumental structure of the verb was modified such that a complex verb takes one less argument when a noun root gathers with a verb root. In the NI other form, when a noun root gathers with a verb root, the argument structure of the complex verb is unchanged.

Another researcher, Savoy (2010), described the 245 speeches delivered by Senators John McCain and Barack Obama between the 2007 and 2008 calendar years. He presented the collection's primary features and contrasted the speeches' frequent use of popular English vocabulary with everyday language (Brown corpus). Additionally, he contrasted various metrics that could be used to extract words that best described a certain portion of the full-text corpus. He identified and examined the terms over and under-used by both candidates during the most recent US presidential election from a statistical and dynamic perspective.

Various methods and techniques were developed for lexical analysis by various researchers. For instance, LOCkey, a fully automated technique developed by Nair & Rost in 2002, analyses SWISS-PROT keywords to determine sub-cellular localization. Their approach served as a helpful tool for enhancing functional information that is already available. The SWISS-PROT sequence database mostly consists of functional annotations for diverse

proteins in the form of keywords. Only a few proteins, however, have information accessible on their sub-cellular localizations.

In a thorough cross-validation test, the Nair and Ross (2002), technique achieved a level of more than 82 ways. Less than half of all the proteins in SWISS PROT could have their localisation deduced because there were no functional antonations. They used LOCkey to annotate five fully sequenced proteomes, including a portion of human proteins, Drosophila melanogaster (a fly), Arabidopsis thaliana (a plant), and Saccharomyces cerevisiae (yeast). They discovered 5000 annotations of these prokaryotes' sub-cellular localization using LOCkey.

Computational tools for investigating corpora (machine-readable) are currently still old. There are concordance programmes in more advanced lexicographic organisations, that are primarily KWIC (keyword in context) indexes with more features like the ability to extend the context, rightwards along with sort leftwards, and so on. Interactive software is not enough to meet the demand. According to Church, Gale, Hanks, & Hindle, (1991), this may be one of the reasons why American dictionaries keep machine-readable corpora less important and instead rely on collections of the chosen citations supplemented by introspection and not by full-text analysis. They assert that in Britain, things are a little different. They say that machine-readable corpora are becoming increasingly important to UK's lexicographers, particularly those concerned with dictionaries for foreign learners. They employed these corpora, as well as the above-mentioned fundamental concordance method, to write precise syntactic descriptions (triggering a move, that will possibly dominate lexicography in the 1990s, towards thorough descriptions of lexical syntax).

Moreover, the study of Zhang H., Liu Q., Cheng X., Zhang H., and Yu H., (2003), uses a hierarchical hidden Markov model (HHMM) to offer a unified strategy for Chinese lexical analysis, to incorporate segmentation of Chinese word, disambiguation, Speech-parts tagging, and unknown words identification into a single theoretical framework. In word segmentation, they used a class-based HMM, and treated unknown terms the same as frequent words in the lexicon at this level. In role-based HMM, unknown words are reliably identified.

For disambiguation, the researchers proposed an n-shortest-path technique that saves top N-segmentation results as candidates and covers more uncertainty in the early stages. Several investigations have already shown that each HHMM level helps with lexical processing. ICTCLAS, an HHMM-based system, was developed. They asserted that, according to a recent government assessment, ICTCLAS is one of the top Chinese lexical analyzers. In short, HHMM is useful for Chinese lexical analysis.

Although they worked with Chinese, the same could be applied to English or any other language. There are various lexical analyzers available for English as well as other languages. Parts of speech tagging, disambiguation, N- segmentation, lexical strings comparisons, etc. could easily be found through various tools available online.

Although, in the above section discussion was made on the various works carried out for finding the quality of automated paraphrasing and the similarity and relatedness between the texts, the area left uncovered is assessing the quality of a paraphrasing tool at the lexical level. Therefore, the aim of this study is to explore the performance of paraphrasing tools on the lexical level and hence, study the quality of the tool used for paraphrasing.

Research Questions:

This study answered important questions regarding the quality of paraphrasing tools and the relatedness of the target text to the original. The questions answered through this study are:

1. How much is automatic paraphrasing lexically relevant to the original text?

2. What various modifications are made at the word level by the paraphrasing tool?

The lexical relevance of the paraphrase to the original text has been evaluated. By lexical relevance, it is meant how much the words of the automated paraphrased text are relevant to that of the original text in terms of morphological changes. How a word is related to its alternate word in the original text in the HSO measure in the WS4J tool has been studied. The limitations of the paraphrasing tool have been highlighted. In evaluating the lexical, the focus remained on the paraphrased sentences and its alternative in the original text.

Methodology

This study applied an exploratory qualitative method. Generally, the main objective of qualitative research is description, understanding and interpretation (Lichtman, 2010). An exploratory design is applied when there are no earlier studies to refer to or rely upon to predict an outcome. According to Swedberg, exploratory research aims at the discovery of something new and important (Stebbins, 2001). This research has used automated paraphrasing of a literary (Emma) as well as a non-literary (DAWN article) text as data. According to Given (2015) documents can be utilized as the only focus for data gathering and analysis. Therefore, the document used was the original texts and the automated paraphrases. QuillBot is used for paraphrasing. It is selected because it is considered as one of the standard tools and paraphrasing mistakes expectancy is less in it. According to the reviewers, QuillBot is easy to be used and it works on a variety of programs such as Chrome Extension, text editing, etc. It produces presentable writing in minimum time. It is convenient to be used for it rewrites the text components by altering the structure of the sentences and rewriting words with synonyms preserving the meaning of the original content (Fitria, 2021). Moreover, AntConc software was used for computing the frequency of various parts of speech, keyword comparison and for finding lexical relevance between the original and paraphrased texts. The AntConc "hosts a comprehensive set of tools including a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a word distribution plot" (Anthony, 2004, p.7). The text supplied could be compared with the reference text provided to AntConc. Moreover, WS4J was used for finding lexico-semantic relevance through the HSO measure. Words were compared through WS4J and the HSO measure values were calculated for the lexico-semantic relevance of data.

The automated paraphrasing was analyzed lexically. The grammatical categories and morphological construction of words in the source text and its alternative in the paraphrased text are studied showing word relatedness and grammatical changes, if any. Various measures, categorized mainly into corpus-based measures and knowledge-based measures are available for analysing relatedness and similarity between words in texts but this study utilised the knowledge-based, HSO measure as it yielded a reliable estimate of four parts of speech which is the maximum

part of speech analysis in WS4J. An online tool, WordNet Similarity for Java (WS4J) was used for obtaining HSO values of relatedness. WS4J is a JAVA API online tool for finding concept similarity using published measures of semantic similarity. It depends on relations between concepts in WordNet and applies the measures of semantic similarity. It gives the similarity between concepts. The online demo provides two options: the first option is to find out the similarity and relatedness between the two words while the second option allows measuring the match between multiple words at a similar time. The first option was used.

This research study with reference to methodology has certain limitations. It has not assessed the quality of all paraphrasing tools available; it only evaluated the paraphrasing carried out by QuillBot. The sampling is also limited to a few chapters/sections of both literary and non-literary texts. In lexical analysis, the strings, tokenization, etc. were not considered. While measuring semantic relatedness and similarity, WS4J was used but all the measures obtained at WS4J were not considered in this study, only the HSO measures were analyzed and calculated, which gave relatedness between the concepts by measuring the path distance between them. Moreover, one of the problems in using WordNet is that certain words are not present in WordNet whose HSO values cannot be calculated. The returning HSO measure value is -1 for such words.

Analysis and Discussion

The building blocks of a text are words. As a part of a text, words are morphologically complex, for text-oriented applications, it is possible to register the structure of a word (Hippisley, 2010). To find out the lexical relevance and match between the original and the paraphrased text, the morphological structures of words were compared. The grammatical category of paraphrased text words was also compared to that of the original texts. It is found through QuillBot paraphrasing that the grammatical categories of certain words are modified while mostly it is left unmodified by the tool. Therefore, words having an important role in text building should be studied while analyzing larger texts.

Lexical relevance of the literary text:

Literary text paraphrased by QuillBot is found modified at the lexical level: the grammatical categories, inflections, and

vocabulary items were modified. Certain words were added while others were dropped out. such as, it is said in the original text that, Emma 'has her own way', but it is paraphrased as 'she has little too much of her own way'. Hence, the word little is added. Similarly, the function words such as prepositions were often dropped out, for example, in the sentence, '…of moments of regret' the word of has been dropped out. Moreover, the lexical words such as, nouns were also dropped, the name of Mr Woodhouse was omitted in one place, for instance.

In addition, other types of modifications were also found in words and on a lexical level, for example, words replaced with idioms or verb phrases, etc. the lexicon marry+past participle, -ed (married) is replaced with the idiom 'tied the knot'. Another example is the change of verb phrase that is put into is changed into go+present participle (going).

Morphological Modifications:

The tool morphologically modifies texts. Changing the word's morphological construction differentiates both texts. The following heads and tables highlight some of the lexical changes in the paraphrased text from the original text:

Changes in lemma: Lemmatization as well as other changes in lemmas were observed in automated paraphrase. Lemma is the dictionary variant of the word, while lemmatization is the process of reducing the word into its base form. The following table lists a few such modifications:

Table 1 Changes in Lemma in Automated Paraphrasing

Original words	Paraphrased words	Inflectional change	Type of change
owing	owed	Base+ing to -ed	Present participle changed into past participle
going	go	Base-ing to zero inflection.	The present participle changed into the base form
coming	come	Base+ing to zero inflection.	The present participle changed into the base form
Put into	going	base+ prep to-ing.	present participle is used in place of the infinitive

paying	pay	Base+ing to zero inflection	The present participle is changed base form

It is shown in the table how the words changed in automated paraphrasing. In the original text it was said, 'A large debt of gratitude was owed here', it was paraphrased as 'A great debt of gratitude was owed here'. Similarly, the word going is lemmatized to go, coming to come, paying pay, and the word phrase put into is paraphrased into going. These changes affect the meaning, for example, 'put into' changing into 'going' completely changes the word form and sense as 'put into… stable' and 'going into… stable' give different senses. Changing the lemma definitely changes the sense as from one tense the utterance changes into another.

Changes in Number:

The distinction in number is expressed morphologically, by adding suffixes to nouns or other types of changes in its form in order to indicate whether it refers to one or more than one (Baker, 2006). It was also noted that the tool singularised certain plural lexemes. For example, 'disparity in their ages' was substituted with **'**their actual age difference, and the word ages changed into age. Moreover, the singular words were also pluralised, such as, in the original text it is said, 'every expression of congratulation which her marriage had already secured' is replaced by the **'**expression of congratulations that her marriage had already attracted', the word congratulation is pluralised here. Therefore, the words were changed taking into account the changes that occurred in the surrounding words.

Parts of Speech Modification:

The tool changes the speech parts of the texts while paraphrasing. The grammatical categories of the words are often modified. The QuillBot mostly changes the adjectives. The verbs were also replaced with synonyms. The following table shows the POS changed words.

Table 2 Parts of Speech Changes

Source Text	Paraphrased Text	POS changes	Lexical analysis

Sister	Sisterly	the adjective used instead of a noun	Base form+ adjective element 'ly'
Self-denying and Generous	Selflessly and Generously	adjectives substituted by adverbs	Adjective to adverb element 'ly'
Happy	Pleased	The adjective changed to a verb	Base form + past participle element 'ed'
Lose	Loss	Verb to noun	Stem word to stem word

The word sister is changed into sisterly, changing the POS category of the word. Similarly, in the next example, the adjectives, Self-denying and generous, were changed into adverbs, selflessly and generously. Happy is changed into pleased, adjective is changed to a verb. In addition, the Verb (lose) to the noun (loss). These changes in word status were made according to the word order changed by the tool.

Other changes:

The lexicons are often replaced with words that fall short of the original word, for instance, in the original text it was said that 'Emma was aware of…', but it was paraphrased as 'Emma realized…', which was out of context here. Similarly, in another place, it is said about Emma that her father's presence made it necessary for her to be cheerful but this was paraphrased as "forced her to smile" She was not forced but she herself did not want her father to be said, and tried to make him cheerful.

Moreover, changes like changing a word to an idiom, phrase, and/or vice versa, tense changing, etc. are also evident in tool paraphrasing. The following table lists these types of changes:

Table 3 Various Types of Modifications Brought by Tool

Original words	**Paraphrased words**	**Change**
Married	Tied the knot	From simple past participle to Idiom
Put into	Going	From verb phrases to Simple present participle

Bear	Going to	Verb phrase to simple/base verb
Was aware	Realized	Past participle to the base form

The above table shows how QuillBot changed the words from one grammatical category into another. The simple past participle is converted into the idiom, the verb phrase to a simple present participle, the verb phrase to a simple/base verb, etc. Though all these changes are contextually unfit in the paraphrased text, some of these have completely different senses/functions in the paraphrased text such as the word realized used in place of aware changes the sense, as Emma was aware carries a different sense than Emma realized.

Words perfectly substituted:

Certain words were replaced by more authentic words by the tool according to the context. For example, consider the Italicised words used in the following sentences taken from the original and also from the paraphrased text:

Example 1:

'…he had received a good education'. (Original text)

'…He had obtained a solid education'. (Paraphrased text)

Example 2:

'…the sale of Randalls was long looked forward to'. (Original text)

'…the sale of Randalls had been long anticipated'. (Paraphrased text)

In the first example, the word good has been used which is replaced with solid which is a good collocate for education. Similarly, in the next example, the phrase (looked forward to) is substituted with anticipated which is more suitable to the context.

Thus, it was noted that in the literary texts automated paraphrasing that words belonging to any part of speech were modified but mainly the nouns, adjectives, verbs, and the tool-modified adverbs that is the lexicons or content words. While the

function words in the literary texts have mostly remained untouched. Idioms were also used by the tool. The grammatical category of words was also often modified which sometimes keeps the original taste while other times distorts it. Moreover, certain words were substituted with words more suitable to the context.

Lexical Relevance of the Non-literary Text:

Several modifications were also found at the lexical level in non-literary text. These are:

Morphological modification

The tool often modifies vocabulary. The lexicons were mostly substituted by synonyms while the function words were often dropped out which may or may not have affected the meaning. For example, the word seeks has been used for aims that do not affect the sense. But the word administration used in place of government does affect the meaning. The inflections attached to the main morphemes are mainly changed, for example, the 's' has been replaced by '-ing' or "-ed", etc.

Table 4 Morphological Changes in Non-literary Text

Lexicons used in the original text	Lexicons replaced in the paraphrased text	Inflections replaced
Increase	Increased	'ed' added, (tense)
Melting	Melt	The inflection is dropped out, (tense)
Warming	Warms	'ing' with 's' (tense)
Threats	Threatened	's' with ed',(tense)
Recognized	Recognises	'd' with s', (tense)
Rising	Risen	'Ing' with 'en', (tense)
Reading	Readings	's'(plurality) is added

In the above table, it is shown how the inflectional modifications were carried out in automated paraphrasing. The lexemes were inflected differently in the target text from the original. In the original text, it is mentioned that, thanks to the increase in greenhouse gas emissions which is paraphrased as 'Because of increased greenhouse gas emissions'. Similarly, the 'crop yields are facing growing threats from floods...' is substituted by 'agricultural harvests are increasingly threatened by floods…' The inflections attached to the lexicon have been changed. The 'government has recognized…' is paraphrased as 'administration recognizes…'. Similarly, 'had been rising into gentility' is replaced with 'that had risen through the ranks of gentility'.

Modifications in Various Parts of Speech:

Sometimes, the prepositions are wrongly used in the paraphrased text. Such as the sentence, 'there are numerous advantages to including the corporate sector in the fight against climate change' shows the preposition 'to' have been used instead of 'of', which best illustrates the point.

Adverbs are sometimes dropped out in tool paraphrasing. For example, the adverb, "also" is noted in several places in the paraphrased text is dropped out. Similarly, it is said in the original text that, 'the government has also taken various initiatives' which indicates that various other efforts were also done but this word is missing in the target text. Moreover, the sentence, 'the government can make use of the private sector's technical experience and capabilities', is substituted for the sentence that 'the government can along with other efforts also make use of technical expertise and capabilities of private business'.

It was found that mainly the adjectives and verbs were substituted with relevant terms. But as we know that no word can substitute the other, for no two words senses are exactly alike so the word replaced lag behind the original word sense. The following table sums up some of the nouns, adjectives, adverbs and verbs modified:

Table 5 Parts of speech modified in Non-Literary Text

Words taken from original text	Words taken from the paraphrased text	HSO values	Grammatical status
Arguably	Undoubtedly	0	Adverb
Biggest	Greatest	-1	Adjective
Threat	Challenge	3	Noun
Now	Currently	0	Adverb
Effects	Consequences	16	Noun
Are	Found	3	Verb
Anthropomorphic	Human	16	Adjective
Increase	Rise	5	Noun
Last	Previous	6	Adjective
Large	Huge	16	Adjective
chunk	Portion	3	Noun
Getting	Becoming	16	Verb
Effects	Consequences	16	Noun
Life	Existence	4	Noun
Factors	Issues	3	Noun
Metric	Measure	0	Noun
Occasionally	Periodically	0	Adverb
Alarming	Worrisome	0	Adjective
Smog	Pollution	6	Noun
Facing	Dealing	4	Verb

Difficult	Tough	16	Adjective
Range	Limit	5	Noun
Problems	Ailments	2	Noun

The table shows that each adverb replaced with another adverb has zero HSO value indicating that both the terms do not have any relatedness. The highest degree of relatedness has been found in adjectives. The nouns and verbs also have some degree of relatedness. The HSO-relatedness values were obtained from WS4J. Certain words have the highest value (16), which means the words are related 100% and would replace each other. Certain word pair give a -1 value which means there is no concerned word found in the WordNet hierarchy. The zero value indicates that there is no path distance between the concerned words.

Other changes:

Other types of modifications such as changing the number (singular-plural), changing the POS category of words, etc. could also be noted in automated paraphrasing. Consider Italicized words in the following sentences:

Example 1

'…the metric occasionally rises to alarming highs of more than 500'.

'…the measure periodically reaches dangerously high levels of more over 500'

Example 2

'In Lahore, the metric occasionally rises to alarming highs of more than 500'.

'In Lahore, the measure periodically reaches dangerously high levels of more than 500.'

In the first example, it is seen that the word highs have been replaced by high, thus the number has been changed and singularity inserted. In the next example, the adjective, alarming has been changed into adverb, dangerously.

Substitution with perfect words (sense-wise):

In an automated paraphrase of the literary text, certain words are found to be more suitable to the context of the sentence as a whole than the original words used. For example, the original text words: '…access to clean drinking water', are changed into '…safe drinking water', the word safe suits the sentence more perfectly. Similarly, in the sentence from the original text the words: 'crop yield' were replaced with the words 'agriculture harvests', which suits the sentence.

While comparing both texts on the lexical level, it is found that lexicons are mainly replaced with relevant words. These replacements often affect the sense and quality of automated paraphrasing as no two words could be 100% alike and one word cannot cover all the dimensions of another word, no word could fill the gap left by another one. Therefore, replacing one word with another places the sense at risk.

At the lexical level, the relevance of QuillBot paraphrased text with the original text was found high as the lexicons were mostly replaced with highly relevant words. The inflections modification was also mainly authentic. The grammatical category of certain words was also modified by causing lemmatical changes. The number that is singularity/plurality of words was also affected. These words were changed according to their position and function in the sentence. In lexical modifications, the whole sense and function of the whole sentence were considered. Such changes sometimes distorted the function of the word while at other times perfecting it. Mostly, the functions were not disturbed in QuillBot as it is a standard tool and it chooses often the right alternatives for words.

Various modifications, as well as relevance, were found at the level discussed above. In the lexical category, mostly the content words were substituted with synonymous words. The

morphological structure and grammatical category of words were changed. The relatedness of the words is also considered.

Summary of Findings

Modifications found at the lexical level in the literary text are in the word order, speech parts, addition and omission of words in sentences, etc. Word order was kept almost the same in the paraphrasing but the word order of certain parts of long sentences could be found changed. Grammatical ambiguities were also found in the automated paraphrasing. Certain words were substituted with wrong words which alters the function of the whole sentence/ the whole sense suffers. Various words were added in various sentences which were not found in the original text. Certain words of the original text were removed in paraphrasing. In the automated paraphrasing of the literary text, several words and particularly phrases of the original text were found missing. In the automated paraphrasing of the literary text, several lexicons and particularly original text phrases were found missing. The omission of certain words and phrases did not affect the flow, sense, and function of the complete sentence, while that of others did change it. Many sentences of the literary text were left un-paraphrased by the tool. While paraphrasing chapters from Emma through QuillBot, it was observed word order of certain sentences was changed in such a way that it seems incorrect, though syntactically they are allowable. Certain sentences were completely wrong paraphrased by changing the words with un-proper words like the word loss paraphrased into death so the cause of loss is not always death. The non-literary text was also found modified at the lexical level. No words or phrases were found omitted in the non-literary paraphrased text. The word order was modified and words were substituted but the sense was kept the same mostly. Comparing tool paraphrasing of literary and non-literary text, it was found that there were fewer ambiguities and differences in the literary text as compared to the non-literary text. Tense changes were there it was rare. The synonyms used were mostly authentic and according to the context.

Conclusion

This study analyzed and compared QuillBot's paraphrased text to the original text. The lexical relevance and deviation of the tool paraphrased text to the original texts are explored. The literary and non-literary texts were taken as the data sample, whose results

could be generalized to tool paraphrasing of all types of literary and non-literary texts. The morphological construction of the words is studied. In addition, the grammatical categories and word order modifications are explored. Therefore, the words as well as sentences of both texts are compared and their relevance as well as deviations in structure and meaning is explored. It is concluded, that tool paraphrasing could not be taken for granted. Although the standard tools such as QuillBot, paraphrasing could be trusted up to a certain extent but should not completely rely upon. Furthermore, this study could be developed by exploring texts at the semantic and syntactic levels. Other paraphrasing tools and measures could also be employed for evaluating automated paraphrasing.

References

Anthony, L. (2004). AntConc: A learner and classroom-friendly, multi-platform corpus analysis toolkit. proceedings of IWLeL, An Interactive Workshop on Language e-Learning, 7-13.

Bailey, (2017, March 23). Automatic paraphrasing: a problem for academia. In PT Plagiarism Today. https://www.plagiarismtoday.com/2017/03/23/automatic-paraphrasing-a-problem-for-academia/

Basori, M. A. (2017, December 30). Strategi dan Teknis paraphrase dalam academic writing: reformulasi isi tanpa reduksi [Seminar and Workshop]. Seminar and Workshop presented at the Workshop Plagiarisme dan Pencegahannya. Retrieved June 19, 2021, from Workshop Plagiarisme dan Pencegahannya website: http://repository.uin-malang.ac.id/2272/.

Church, K., Gale, W., Hanks, P., & Hindle, D. (1991). Using statistics in lexical analysis. In Lexical Acquisition: Exploiting on-line Resources to Build a Lexicon (pp. 115-164). Psychology Press.

Fitria, T. N. (2021). QuillBot as an online tool: Students' alternative in paraphrasing and rewriting of English writing. Englisia: Journal of Language, Education, and Humanities, 9(1), 183-196. https://doi.org/10.22373/ej.v9i1.10233.

Guo, Z., Huang, Z., Zhu, K. Q., Chen, G., Zhang, K., Chen, B., & Huang, F. (2021). Automatically Paraphrasing via Sentence Reconstruction and Round-trip Translation. In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21) (pp. 3815-3821).

Hippisley, Andrew R., (2010). Lexical analysis. Handbook of Natural Language Processing, Second Edition, Nitin Indurkhya & Fred J. Damerau (Eds.), pp. 31-58. Linguistics Faculty Publications. https://uknowledge.uky.edu/lin_facpub/66

Hirst, G., and St-Onge, D. 1998. Lexical chains as representations of context for the detection and correction of malapropisms. In Fellbaum, C., ed., WordNet: An electronic lexical database. MIT Press. 305–332.

Leacock, C., and Chodorow, M. 1998. Combining local context and WordNet similarity for word sense identification. In Fellbaum, C., ed., WordNet: An electronic lexical database. MIT Press. 265–283.

Lin, D. (1998). An information-theoretic definition of similarity. In Proceedings of the International Conference on Machine Learning.

Litchman, M. (2010). Understanding and evaluating qualitative educational research.US: Sage Publications.

Martins, H. F. (2004). Basic Synonyms in English You Should Know. Lisban accounting and business school-IPL.

Nair R. Rost B. (2002) Inferring sub-cellular localization through automated lexical analysis. Bioinformatics 18 (suppl_1), S78-S86,

O'Reilly, K. (2012). Strategic reading level 1: Teacher's manual. Cambridge, UK: Cambridge University Press.

Phillips B. S., (2001). Lexical diffusion, lexical frequency, and lexical analysis. Typological Studies in Language. 45, 123-136

Resnik, P. 1995. Using information content to evaluate semantic similarity in a taxonomy. In Proceedings of the 14th International Joint Conference on Artificial Intelligence, 448–453.

Rogerson, A. M., & McCarthy, G. (2017). Using Internet-based paraphrasing tools: Original work, patchwriting or facilitated plagiarism? International Journal for Educational Integrity, 13(1), 1-15.

Rosen S. T., (1989). Two types of noun incorporation: A lexical analysis. Language, 294-317

Savoy J., (2010). Lexical analysis of US political speeches. Journal of Quantitative Linguistics, 17 (2), 123-141,

Soles, D. (2003). The essentials of academic writing. Boston, MA: Houghton Mifflin Company.

Stebbins, R. A. (2001). Exploratory Research in the Social Sciences. Sage University Papers Series on Qualitative Research Methods, vol. 48. Thousand Oaks, CA: Sage.

Stebbins, R. A. (2001). Exploratory research in the social sciences. Sage.

Wu, Z., and Palmer, M. (1994). Verb semantics and lexical selection. In 32nd Annual Meeting of the Association for Computational Linguistics, 133–138.

Zhang H., Liu Q., Cheng X., Zhang H., Yu H., (2003). Chinese lexical analysis using hierarchical hidden Markov model. In Proceedings of the second SIGHAN workshop on Chinese language processing, 63-70.