

Natural Language Processing: An Alternative Approach to Analysis of Student's Summarization Strategies based on Heuristic Rules

^{1,2}Jaafar Zubairu Maitama and ²Norisma Binti Idris

¹Department of Artificial Intelligence, University of Malaya, Kuala Lumpur Malaysia

²Faculty of Computer Science and Information Technology, Bayero University Kano, Nigeria

Abstract: The rapid increase in the electronic information due global technology advancement as well as manual text usage has resulted in text information overload problem. This has leads to significant changes on text processing activities from the traditional way. Due to continous increase in both online and offline text, demand for text summarization has become the focus of huge interest and investment in both reserarch and commercial sectors. This research is aim to investigate various summarizing strategies in order to identified the commonly used strategies as well as problems their limitations in students' summary sentences. In this paper, we used heuristic rules to analyzed students' summaries. The results indicated that deletion is the commonly used summarization strategy among student's while topic sentence selection is the least. Also, it was found that grammatical error is the common problem affecting students' summary sentences followed by spelling mistakes. In addition, processes of sentences combination were identified and it was found out that apart from improper application of conjunction word and conjunction symbol, irrelevancy between combined sentences is the most frequent problem associated with sentences combination strategies. The results of this research can be used by teachers and institutions of learning to adapt their automatic summarization systems to improve teaching strategies which in turn enhanced student's summarization skills.

Key words: Summary • Students • Sentences • Summarizing Strategies

INTRODUCTION

Natural Language Processing discipline concerns itself with the design and implementation of computational machinery that communicate with humans using natural language (Dale, Moisl, & Somers, 2000) [1]. There have been significant advances in the areas of language technology recently. These include text summarization, text categorization and text assessment. Research has shown that lack of strategic skill is the cause of students' difficulties in writing good summaries (Idris, Baba, & Abdullah, 2011) [2]. Most researchers only focused on the end product of summary writing instead of its production stages. Consequently, efforts to identify, assess and diagnose students' summarizing strategies have been very limited.

Summary writing is an important aspect of learning which contributes greatly in facilitating general understanding and remembering text by students. It is quite difficult for human beings to manually summarize

large documents of text. In view of this, previous study has suggested that writing summary in logical way of narrative structure and timely-based measure should be at the early stage of student's learning, right from junior high school levels to overcome the long time process of learning summary writing and difficulties (Hill, 1991) [3]. It was also discovered that students who received direct instruction in summary writing based on their developmental level are usually successful (Kintsch & Van Dijk, 1978)[4]. However, successful teachers are those who understand the developmental stages of the learner and task just to build the relationship between the two in order to teach summary writing across curriculum areas (Hill, 1991) [3].

Summarization is a process of condensing a source text into a shorter version preserving its reformation content and can serve several goals ranging from survey analysis of a scientific field to quick indicative notes on the general topic of a text (Barzilay & Elhadad, 1999). According to Mani & Maybury (1999) [5] text

summarization is the process of extracting the most important information from a text to produce an abridged version for a particular user and task. Summarization plays a significant role in many natural language processing applications; and these applications are quite useful to human daily activities. Idris, Baba, & Abdullah (2009) [6] described summarization strategies as a set of rules applied to produce summaries, such as deletion, paraphrase, sentences combination, generalization etc.

Currently, some efforts were made to address problems associated with summary writing among which are summary assessment and diagnosis systems. One of those systems used an algorithm to identifying students' summarization strategies (Idris, Baba, & Abdullah, 2011) [2]. However, there are limited works on investigating processes as well as difficulties involved in students' summaries which can consequently, hinders students' summary writing ability.

In this research, we intend to investigate various summarizing strategies in order to identified the commonly used strategies as well as problems associated with those strategies in students' summary sentences.

Related Work: It is obvious that text summarization has been in existence for a very long period of time and this made it possible for human to invent different kind of processes for summarizing text. The aim of accomplishing the task of text summarization by human, is simply to obtain an abridge version of the origination text with the most important information. Brandow, Mitze, & Rau (1995) [7] explained that production of text summarization involves the use of the following steps: (1) Understanding the content of the document; (2) identification of most important pieces of information contained in the document; and finally (3) write up this information, whereas Endres-Niggemeyer & Neugebauer (1998) [8] described 3 fundamental tasks involved in summarization processes. There are two types of summaries: (1) indicative summary and (2) informative summary. The description of these summaries by Mani (2001) [9] are as follows: indicative summary provides a concise idea about the content of the document that serve as tips to readers to decide whether to read the entire text or otherwise. It can be applied in documents with limited structures such as stories, opinions and essays. However, informative summary focus on collecting the most important information which provides details of the stuffs contained in the document in a short form. In the presence of informative summary, readers do not need to read the full document. It is commonly applied in science and

technology documents experimental study. Researchers contribute in defining as well as describing some important features of these summaries that claim indicative summary as a summary consisting of about 5-10% of the original text retained sufficient material from a particular document to indicate its content (Strzalkowski, Wang, & Wise, 1998) [10]. On the other hand, informative summary retains the entire relevant information a user need from the original document with about 20-30%. In addition, indicative summary is not meant to capture all the relevant details of a summary, but only for the users to judge whether the entire text is relevant, while informative summary is otherwise (Firmin & Chrzanowski, 1999 [11]; Barzilay & Elhadad, 1999 [12]).

Furthermore, the rapid increase in the internet usage as well as huge amount of electronic information services in this contemporary era has resulted in information overload problem. Electronic information are said to be produced and accumulated continuously overtime (Mani & Maybury, 1999) [13] and there is no enough time to read and process everything and make something reasonable out of it. Therefore, this recent surge in accumulation of information, warrant the urgent need for researchers to consider summary related issues such as summary process, as well as problems involves in writing summaries with the intension of improving summarization activities around the globe.

In recognition of the summarization demands, there so many efforts to contain the present situation among which are, (He, Hui, & Quan, 2009) [14] proposed an integration of LSA and n-gram co-occurrence summarization techniques in order to circumvent the limitations associated with each of the individual techniques so as to improve summarization performance. The proposed approach achieved an accuracy of 96% whereas the optimal performance of the existing technique which is BLEU was 87%. However, the study focuses only on content without considering strategies.

Student assessment is also a vital issue and need serious attention from researchers. Chen [15] (2011) lack of conceptualization for student's knowledge organization by existing student assessments resulted in misunderstanding of desirable method for improving future learning progress by both teachers and students. The research was aimed at providing remedial learning avenues for individual learners based on their knowledge structure. An experiment was conducted on 145 sample of students offering introductory JAVA programming language course at Central Taiwan technology University and the result shows that the system improve

effectiveness of students learning skill based on their knowledge structure. It also demonstrated that weak students benefits from the system more than strong students academically.

The fact that summary writing is an important skill taught in school, strategic approaches and tools that assist students in improving their summarizing skill is also important. Currently, an algorithm for identifying summarizing strategies used by students was developed due to the large amount of time spends as well as difficulties faced by teachers in examining their students' summarizing strategies (Idris, Baba, & Abdullah, 2009). In the research, heuristic rules were generated through careful observations of expert's summaries which lead to development of summary sentence decomposition algorithm. The research was able to achieve the target of identifying student's summarizing strategies automatically and it is quite effective in determining source of the summary sentences and word positions of the sentences. It considers only syntax level summarizing strategies.

In addition, Idris, Baba, & Abdullah (2011) introduced the use of summary sentence decomposition algorithm to identify syntactic level summarization strategies to resolve the difficulties student undergoes in writing good summaries. The approach performed appreciably and was able to identify the five syntactic levels of summarization strategies with 94% accuracy when compared to human. However, this is the most closest to our proposed work, but it does not give much consideration on problems associated on the summarizing strategies identified despite its significance. In addition, non among the related literatures give much concern on investigating the summarizing strategies and identify their associated problems in students' summaries.

Summarization Strategies Based on Heuristic Rules:

Three studies were conducted with focus on investigating students summarizing strategies and common problems associated with some of these strategies. Emphasis is given to a selected strategy known as sentence combination strategy so as to give students more knowledge on how it can be properly used, because it is regularly employed by students but has less concern from researchers.

Case Study 1: Investigation of Students' Summarizing Strategies:

An analysis was conducted on 30 students' summary (see few samples in Appendix A) collected from an article adopted from Reader's Digest August 2005 entitled "Floating Boat".

Samples: To accomplish this study, 30 students' summaries were collected from form 5 level students (aged 16 to 17) at Technique Secondary School (ERT) Dato' Undang Hj. Muhamad Sharin, Rembau, Negeri Sembilan. The students that participated in the study were thought how to write a summary.

Procedures: Firstly, students were given an article with a titled Floating Boat which was adapted from Reader's Digest August 2005. The students were then instructed to read the article carefully and provide a summary of the passage which was used as samples throughout our analysis. Document analysis method is used to collect the qualitative data that enable the research to gather all the necessary information require for the research. The students employed basic summarization strategies.

Strategies and Analysis: Seven summarizing strategies were considered: deletion, sentence combination, paraphrasing, generalization, topic sentence selection, invention and copy-paste. The summarizing strategies were identified using the heuristic summarization rules proposed by Idris, Baba, & Abdullah (2009):

- Deletion – Elimination of trivial or redundant information from original sentences.
- Sentence combination- if two or more sentences from original text are combined to produce summary sentence.
- Generalization – if a list of things or items replaced by a more general word in the same class.
- Paraphrase – if similar word or phrase is use to replaced the original word or phrase.
- Copy-paste – if the words in the summary sentence are exactly the same as those in the original text.
- Topic sentence selection – representing main idea of the entire paragraph with one important sentence.
- Invention - if the original sentences are replaced by a new and more concise sentences conveying similar meaning.

Table 1 presents the results obtain from 30 students' summary samples. In this analysis, we focus on 7 strategies in which all were applied alternatively except for generalization strategy with no frequency. The results showed that deletion, paraphrasing and sentence combination are used more frequently than others, which occur in almost all the summary samples with 49.9%, 24.4% and 13.1% respectively. Though, sentence combination strategy is not the top most frequently used but it typically depends on other independent strategies

Table 1: Number of strategies use by students

Strategies	Frequency	Percentage (%)
Deletion	270	49.9
Sentence Combination	71	13.1
Paraphrase	132	24.4
Generalization	0	0
Copy-paste	53	9.8
Invention	13	2.4
Topic Sentence Selection	2	0.4

like deletion, paraphrase etc. It is observed that it is consistently used with deletion strategy which happens to be the most commonly used strategy. Copy-paste is the fourth most employed strategy in the series and the most easy to use strategy, but students are not encouraged to use it due to its close relation to plagiarism. The results showed that topic sentence selection is second to the list strategy applied by students, followed by generalization which does not appear completely. Invention has 2.8% frequency but very difficult to apply, as it requires students to produce the summary sentences in their own works. The result has unanimously showed that the deletion is the most commonly used by students because almost all summary sentences produced includes deletion except copy-paste strategy and some instances of sentence combination strategy. Example of deletion strategy application is shown in Table 2 with removal of 6 words from the original text sentence:

Paraphrasing depends on the existence of other strategies in some cases. This indicates that it can be both dependent and independent strategy. For example, consider the instances in Table 3 showing some summary sentences produce using paraphrasing by students:

Example of Copy-paste strategy application is in Table 4 where the summary sentence is exactly similar to the original text sentence:

However, among the most commonly use summary strategies by students, the student employ independently includes sentence combination strategy and is in the third position. In some cases, sentences combination strategy is considered as second most frequently used strategy due to the fact that it is the strategy which followed deletion strategy in terms of independent application. In view of the fact that the deletion is one of the most commonly used summary strategies used by students, it needs attention by researcher because it deserves proper exploration. Some examples of summary sentences produced using sentence combination independently are shown in Table 5.

Case Study 2: Problems Associated with Most of the Students' Summary Sentences:

This stage investigates general problems associated with most of the students' summary sentences. The relationships between the problems and combined sentences were studied.

Samples: Twenty summary samples were selected from the 30 samples discussed in Study-1. These samples were collected in the same manner with the samples used in study-1.

Procedures: We consider three criteria which are improper use of grammar, spelling mistakes and incomplete sentences in the summary sentences. For the purpose of this research, we provide a definition to each of these criteria. Consider the definitions below:

Improper Use of Grammar: Is a situation when rules governing the form and meaning of sentences are not properly used.

Spelling Mistakes: Is a situation when a word is misspelled either mistakenly or ignorantly.

Incomplete Summary Sentences: Is a situation where a summary sentence does not convey complete information which tends not to be meaningful.

Analysis: We analyze the students summaries based on the three criteria briefly described in section 3.2 (ii) above. Table 6 shows the results of the analysis:

Figure 1 has revealed that grammatical problem is the most common problem in the students' summary sentences consisting of 83.3% frequency. The results also showed that spelling mistakes and incomplete sentences are considered to be challenging to students' summary sentences with relatively low frequencies of 11.8% and 4.9% respectively. The probable reason why grammatical error was found to be the most common is because the students which the study was conducted in their school are not native speakers of English language. Possibly if the study is replicated in native English language speaker students, the encountered problem could probably be different from the present research.

We carefully investigate the problems and study their relationship with sentence combination strategy as shown below:

Table 2: Deletion strategy application

Summary sentence	Original text sentence	Application of Deletion
However the water just too deep and my body was too weak to fight.	(13) However, the water was just too deep, the current was too strong and mybody was too weak to fight.	■ Deletion

Table 3: Paraphrase strategy application

Summary sentence	Original text sentences	Application of Paraphrase
I was enjoyed.	(34) I was overjoyed.	■ Paraphrase
Terror overcame me and try to gasped for air in desperation but only salt water filled.	(14) Terror overcame me as I felt myself sinking to the bottom. (15) I gasped for air in desperation, but only salty water filled my throat and nostrils.	■ Paraphrase ■ Deletion ■ Sentence combination/(“and”)

Table 4: Copy-paste strategy application

Summary sentence	Original text sentence	Application of Copy-paste
I was overjoyed.	(34) I was overjoyed.	■ Copy-paste

Table 5: Sentence Combination Strategy Application

Summary sentence	Original text sentences	Application of Sentence Combination
Then his hand brushed my hair and pulled me to shore and started mouth-to-mouth resuscitation.	(20) Then his hand brushed against my hair. (22) He pulled me to shore and started mouth-to-mouth resuscitation.	■ Deletion ■ Sentence combination/(“and”)
When I was confident enough, I was overjoyed.	(30) When I was confident enough, he took me to a jetty and challenged me to jump into the water. (34) I was overjoyed.	■ Deletion ■ Sentence combination (“;”)

Table 6: Common problems associated students' summary sentences

Summaries	Improper use of grammar in summary sentences	Spelling mistakes in summary sentences	Incomplete summary sentence
Summary1	4	0	0
Summary2	2	0	1
Summary3	9	1	0
Summary4	6	2	0
Summary5	4	0	0
Summary6	6	1	0
Summary7	2	1	1
Summary8	5	0	0
Summary9	6	1	0
Summary10	3	1	0
Summary11	2	0	1
Summary12	3	0	0
Summary13	2	0	0
Summary14	5	1	0
Summary15	4	1	0
Summary16	5	2	0
Summary17	1	0	0
Summary18	3	0	2
Summary19	6	1	0
Summary20	7	0	0
Overall Total	85	12	5

Table 7: Improper grammar problem

Summary sentence	Original text sentence	Grammar problem
My father dived and swarm as hard as he could to the spot where i had gone under.	(18) He dived in and swam as hard as he could to the spot where I had gone under.	Grammar problem

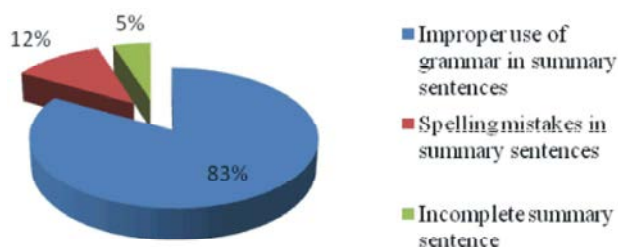


Fig. 1: Frequencies of problems associated with summary sentences

Improper Use of Grammar: A sentence is a group of words which express a complete thought and consist of subject and predicate (Sargeant 2007). Since our emphasis is on summary sentences, there are Basic English grammar rules that are directly related to sentence structure. So, the remaining rules for sentence structure including basic part of speech as well as punctuations. However, the result shows that majority of the students face great challenges of improper use of grammar because it is difficult for the student to master all the grammar rules at their current age and they are not native English speakers. Table 7 shows a summary sentence produced with grammatical problem.

A further study was conducted to compare the relationship between improper use of grammar with the used of sentence combination strategy in students' summary. The result of the comparison is presented in Table 8.

As we can see, it is now clear that out of 85 sentences with improper grammar in the study, 28 sentences were in combined sentences. This shows that 25% of improper grammars occur within combined sentences, while only 75% involves the remaining six strategies based on our study. In addition, 35% of the total combined sentences contain improper grammar.

Spelling Mistakes: This type of mistake happens very frequent among students, due to the fact that some students rush to finish their summary on time which result to common writing mistake. On the other hand, some students make spelling mistakes simply because they do not know the right spellings of some particular words. Though, spelling mistakes happen to be one of the problems affecting students' summary writing, but based on the result in Figure 1 the rate is low. It is commonly occurred because of typo errors as shown in Table 9.

The study compared spelling mistakes in relation to combined sentences. The results are shown in Table 10.

Table 10 shows that 25% of the total spelling mistakes in the 20 summary samples analyzed was within combined sentences. This indicates that spelling mistake is one of the problems affecting sentence combination application in students' summaries.

Incomplete Summary Sentences: This is the least problem among the three common problems affecting summary production. It usually occurs when some important words in a sentence are deleted or ignored by students. We observed in most cases, students believed that the aim of summarization is just production of smaller version of a particular text, rather than smaller and meaningful version. Thus, they end up ignoring some important words which leads to production of incomplete summary sentences and meaningless. Example of incomplete summary sentence problem is shown in Table 11.

Comparison between incomplete sentence problem and combined sentences was also conducted and the result is shown in Table 12.

The results of the study show that incomplete sentence problem does not commonly occur in the combined sentences whereby, only 17% out of 83% of the total incomplete sentences problem in this study.

Case Study 3: Identification of Problems Associated to Combined Sentences: This step identified the common problems associated to sentence combination strategy. In this stage, we use summary samples used in study-2 with focus on sentence combination strategy.

Samples: Twenty summary samples used in study 2 were also employed and are collected in a similar way.

Procedures: We repeat the same initial procedure used in study 2. The procedure continues by investigating the sentence combination strategy and its formation. In this stage, three elements were identified, namely conjunction

Table 8: Improper grammar problem Vs sentence combination strategy

Criteria	Frequencies	Percentage(%)
Improper grammar used in sentences	85	75
Improper grammar within combination sentences	28	25
Total Combined sentences	53	

Table 9: Spelling mistake problem

Summary sentence	Original text sentences	Spelling mistake problem
I felt myself sinking to the bottom only salty water filled my throatnad nostrils.	(14) Terror overcame me as I felt myself sinking to the bottom. (15) I gasped for air in desperation, but only salty water filled my throat and nostrils.	Spelling mistake

Table 10: Spelling mistakes problem Vs combined sentences

Criteria	Frequencies	Percentage(%)
Total spelling mistakes in the summary samples	12	75
Spelling mistakes within combined sentences	4	25
Total Combined sentences in the summary samples	53	

Table 11: Incomplete summary sentence problem

Summary sentence	Original text sentence	Incomplete Sentence problem
Father was horrified to see me struggling.	(17) When my father emerged from the hut, he was horrified to see me struggling in the water some distance from the shore.	Convey incomplete meaning of original sentence

Table 12: Incomplete summary sentence problem Vs combined sentences

Criteria	Frequencies	Percentage(%)
Total incomplete sentences in the summary samples	5	83
Incomplete sentences within combined sentences	1	17
Total Combined sentences in the summary samples	53	

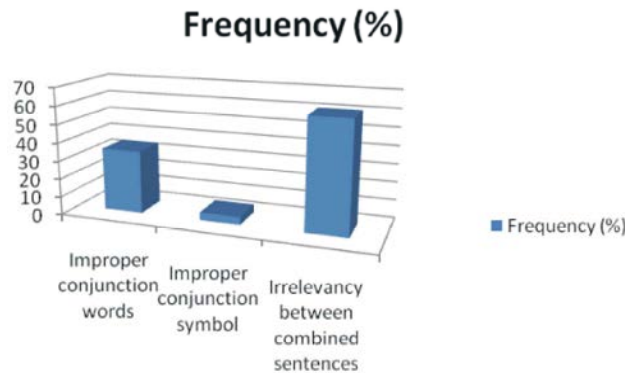


Fig. 2: Analysis of combined sentences problems

words, conjunction symbols and relevancy between combined sentences. These elements were believed to be significant combined sentences due to their frequent appearance. However, it is observed that improper application of any of these elements affect the meaning of the combined sentences, which make the sentence a common challenge in combined sentences. For the purpose of this research, we provide a simple definition to each of these elements.

Conjunction Word: Are word used to join two clauses.

Conjunction Symbol: Are symbols used to join two clauses.

Relevancy of Sentences: Is a situation where the combined sentences are either related or similar in original text.

Table 13: Analysis on how sentences are combined

S/N	Summary Sentences	Original Text sentences used	Strategies Used/(Conjunction words/Symbol used)	Remarks
1	When I tried to get my sailboat from the water, I missed it and fell into the ocean.	(5) But when I reached out to get my boat, a soft wind caught its tail and pushed it away from me. (10) Unfortunately, I missed and fell into the ocean, not realizing how deep the water was and forgetting that I could not swim.	<ul style="list-style-type: none"> ▪ Sentence combination/(" ; ") ▪ Deletion ▪ Paraphrase 	" ; " is properly used because both combined sentences are meaningful
2	I could not swim then I was enveloped in seawater.	(13) However, the water was just too deep, the current was too strong and my body was too weak to fight.	<ul style="list-style-type: none"> ▪ Sentence combination/("then") ▪ Deletion 	"then" is not properly used, grammar problem
3	I trying to remain above the surface with the current was to strong and my body was too weak to fight.	(12) Panic-stricken, I paddled and kicked hard, trying to remain above the surface. (13) However, the water was just too deep, the current was too strong and my body was too weak to fight.	<ul style="list-style-type: none"> ▪ Sentence combination/("with") ▪ Paraphrase ▪ Deletion 	"with" is not properly used, Grammar problem, Affect content meaning, spelling mistake.
4	Then I felt myself sinking to the bottom and I passed out.	(14) Terror overcame me as I felt myself sinking to the bottom. (16) Then I passed out.	<ul style="list-style-type: none"> ▪ Sentence combination/("and") ▪ Deletion 	"and" is properly used, "then" is more proper conjunction.
6	He grabbed my hair and yanked my head out of the water.	(20) Then his hand brushed against my hair. (21) He grabbed it and yanked my head out of the water.	<ul style="list-style-type: none"> ▪ Sentence combination/("and") ▪ Paraphrase ▪ Deletion 	"and" is properly used
12	I plunged into the ocean and swam back to shore with overjoyed.	(33) I plunged into the ocean and swam back to shore. (34) I was overjoyed.	<ul style="list-style-type: none"> ▪ Deletion ▪ Sentence combination/("with") 	Grammar problem; "with" is not properly used

Analysis of the Combined Sentences Problems: The result of the analysis is presented in Figure 2. It is found that improper use of conjunction symbols is not the most common problem found in sentence combination with only 4.3%. Follow by improper conjunction words and irrelevancy between combined sentences with 34.8% and 60.9% respectively. In addition, irrelevancy between combined sentences problem happen to be the most challenging problem among the combined sentences. This is simply because, improper application of any of the other two elements contribute irrelevancy between combined sentences. In this case, our research on text similarity measure between combined sentences is significant in students' summary sentences. The logic is that, knowing the similarity will help in understanding the relationship between sentences and leads to a proper combination of related sentences, hence boost students summarization skills.

How Sentences Are Combined: To reveal how sentences are combined in students' summaries, we consider the summary #16 (see Appendix B) for an investigation. The results are shown in Table 13.

We begin by decomposing one summary sample (summary#16), which was chosen randomly from 30 summaries. Summarization rules discussed in study- 1 were used to identify the summary sentences produced using sentence combination. The same rules were also used to identify the remaining summarization strategies involved in this research.

It is discovered that sentence combination strategy is typically used with other summary strategies such as deletion and paraphrasing strategies. Table 13 clearly demonstrated that all the six sentence combination strategies are used together with either deletion, paraphrasing of both strategies. However, existence of deletion strategy leads to some improper outcomes of

Table 14: Combined Sentences with no Conjunction

Summary Sample	Summary sentence #	Combined sentences with no-conjunctions	Original Text sentences used
Summary #3	9	He grabbed mouth-to-mouth resuscitation.	(21) He grabbed it and yanked my head out of the water. (22) He pulled me to shore and started mouth-to-mouth resuscitation.
Summary #4	4	I felt myself sinking to the bottom only salty water filled my throatnad nostrils.	(14) Terror overcame me as I felt myself sinking to the bottom. (15) I gasped for air in desperation, but only salty water filled my throat and nostrils.
Summary #11	1	When I tried to get my sailboat from the water, but the currents kept pushing the boat further and further away.	(5) But when I reached out to get my boat, a soft wind caught its tail and pushed it away from me. (6) I took a couple of steps towards it, but the currents kept pushing the boat further and further away.
Summary #19	1	When I tried to get my sailboat from the water in horror as my floated out of reach in stretched my arm as far as it could go and tried to grab the boat.	(5) But when I reached out to get my boat, a soft wind caught its tail and pushed it away from me. (7) I watched in horror as my boat floated out of reach. (9) In one huge step, I stretched my arm as far as it could go and tried to grab the boat
Summary #21	7	My father dived in and swam as hard under the water frantically searching for my body.	(18) He dived in and swam as hard as he could to the spot where I had gone under. (19) He repeatedly dived under the water, frantically searching for my body.
Summary #26	12	My father taught me swim I was confident.	(29) In the following weeks, my father taught me how to swim. (30) When I was confident enough, he took me to a jetty and challenged me to jump into the water.
Summary #29	8	One day my father taught me how to swim.	(27) One day, he took me to the edge of the water and urged me to go in with him.(29) In the following weeks, my father taught me how to swim.
Summary #30	6	My father worried he taught me how to swim for a week.	(26) My father was worried that the incident would scar me for life. (29) In the following weeks, my father taught me how to swim.

sentences in most cases, as far as sentence combination strategy is concerned. In this study, we found out that these improper outcomes includes incomplete sentences, grammar problem, improper use of conjunctions as well as changing meaning of original text, as shown in Table 13.

Three key items are considered which are typically involved in the production of combined sentences in summary sentences include:

Conjunctions Appearance: Conjunctions are words or symbols used to join two clauses. In our analysis, the two commonly used conjunctions involved in students' summary sentences are: conjunction words "and, but, or, with, etc." and conjunction symbol, which is ",". For example, referring to Table 13, summary sentences 2, 3, 4, 6 and 12 apply conjunction words with "then", "with", "and, then", "and" and "with" respectively. On the other hand, summary sentence 1 applies conjunction symbol ",", to combine the sentences.

Absence of Conjunction (No Conjunction): There are quite a number of combined sentences which does not employ any kind of conjunction. Though, most of the summary sentences that have no conjunctions are

exposed to grammatical problems danger that can temper with original content meaning. Some students still used it as a method of combing sentences in there summary. Table 14 show 8 summary sentences generated from sentence combination strategy with no conjunction.

Selection of Sentence Partner: In sentence combination, there is individuality in the process of selecting desired sentences to be combined among the sentences. Based on our analysis, two or three sentences were usually combined to produce a summary sentence. For example, see Table 14 sentence#7 of summary#21 was produced by combining two sentences while sentence#1 of summary#19 was produced by combining three sentences. Some of these sentences are in adjacent position, while some are very far in the original text. In addition, many sentences were combined because they are related, similar, or opposite to each other in the original content.

DISCUSSION

In study-1, 7 different summarizing strategies were identified. The study starts by investigating 30 samples

produce by students and compared each with original text. Then, we identify these strategies using summarization heuristic rules proposed by Idris, Baba, & Abdullah (2009). However, the results of the study justify the hierarchical sequence of the commonly used strategies based on their frequency such as deletion, paraphrase, sentence combination, copy-paste, invention, topic sentence selection and generalization respectively. It was found that deletion is the most frequently used strategy while generalization appeared as the least frequently employed summary strategy by students.

In study-2, it was found out that improper use of grammar, spelling mistakes as well as incomplete sentence problems are the major problems affecting students' summary sentences. In addition, we also discover that these three problems can be considered as the external problems associated with sentence combination strategy. This is because, the results of the comparisons conducted between the problems and combined sentences shows that most of these problems occur at a point where sentence combination strategy is employed. Out of the remaining six strategies identified in study-1, sentence combination obtains the frequencies of 25%, 25% and 17% for improper use of grammar, spelling mistakes and incomplete sentence respectively.

Our findings in study- 3 revealed that the common challenges in sentence combination strategies are conjunction words; conjunction symbols as well as irrelevancy between the combined sentences of the original text. These are considered as internal problems associated with sentence combination strategy. The processes followed by students to combined sentences in their summaries were identified. In addition, the results shows that sentence combination strategy is observed to be one of the most commonly used strategies but most researcher are more concern with deletion and other syntactic level strategies such as generalization, paraphrase invention etc despite the fact that it also play a significant role in students' understanding of summary.

CONCLUSION

This study analysis students summarization strategies based on heuristic rules. The study investigates commonly used summarizing strategies, common problems associated with summarizing strategies as well as internal problems associated with sentence combination strategies. In case study-1, 7 different summarizing strategies were identified. It was found that deletion attracted attention from students more than other strategies whereas generalization is the least patronized

summary strategy among students. In case study-2, it was found out that improper use of grammar, spelling mistakes as well as incomplete sentence problems are the major problems affecting students' summary sentences. However, 17% for improper use of grammar, spelling mistakes and incomplete sentences, respectively. Case study- 3 showed that the common challenges in sentence combination strategies are conjunction words, conjunction symbols as well as irrelevancy between the combined sentences of the original text.

The results of this research can be used by teachers and institutions of learning to adapt their automatic summarization systems to enhance teaching strategies in terms of summarization skills. This could improve students understanding and production of summary. For future work, we intend to recommend individual investigation of the strategies and it processes.

REFERENCES

1. Dale, R., H. Moisl and H. Somers, 2000. Handbook of natural language processing: CRC Press.
2. Idris, N., S. Baba and R. Abdullah, 2011. Identifying Students' Summary Writing Strategies Using Summary Sentence Decomposition Algorithm, Malaysian Journal of Computer Sci., 24(4): 180-194.
3. Hill, M., 1991. Writing Summaries Promotes Thinking and Learning across the Curriculum: But Why Are They so Difficult to Write? Journal of Reading, pp: 536-539.
4. Kintsch, W. and T.A. Van Dijk, 1978. Toward a model of text comprehension and production, Psychological Review, 85(5): 363.
5. Mani, I. and M.T. Maybury, 1999. Advances in automatic text summarization: the MIT Press.
6. Idris, N., S. Baba and R. Abdullah, 2009. A Summary Sentence Decomposition Algorithm for Summarizing Strategies Identification, Computer and Information Science, 2(1): 200.
7. Brandow, R., K. Mitze and L.F. Rau, 1995. Automatic condensation of electronic publications by sentence selection, Information Processing & Management, 31(5): 675-685.
8. Endres-Niggemeyer, B. and E. Neugebauer, 1998. Professional summarizing: no cognitive simulation without observation, Journal of the American Society for Information Science, 49(6): 486-506.
9. Mani, I., 2001. Recent developments in text summarization. Paper presented at the Proceedings of the tenth international conference on Information and knowledge management.

10. Strzalkowski, T., J. Wang and B. Wise, 1998. A robust practical text summarization, Paper presented at the Proceedings of the AAAI Symposium on Intelligent Text Summarization.
11. Firmin, T. and M. J. Chrzanowski, 1999. An evaluation of automatic text summarization systems, *Advances in Automatic Text Summarization*, pp: 325-336.
12. Barzilay, R. and M. Elhadad, 1999. Using lexical chains for text summarization, *Advances in Automatic Text Summarization*, pp: 111-121.
13. Mani, I. and M.T. Maybury, 1999. *Advances in automatic text summarization: the MIT Press*.
14. He, Y., S.C. Hui and T.T. Quan, 2009. Automatic summary assessment for intelligent tutoring systems, *Computers & Education*, 53(3): 890-899.
15. Chen, L.H., 2011. Enhancement of student learning performance using personalized diagnosis and remedial learning system, *Computers & Education*, 56(1): 289-299.