A Comparative Study of Summarizing Strategies

between Japanese and English Speakers

異言語母語話者の文章要約ストラテジーに関する比較研究

KATORI Mari
香取　真理

Abstract: 我々は長い話を聞いた後、それをかいつまんで他の人に話すことができる。長い文章や講義も 1 枚のノートにまとめることができる。人は無意識にも様々な場面で「要約」を行っている。近年、文章や発話の理解過程の解明は、様々な分野で重要なテーマとなっている。本研究では、人の「要約ストラテジー」を解明する手がかりの一端として、日本語母語話者 20 名による日本文要約のストラテジー、及び英語母語話者 20 名による英文要約のストラテジーを調査の対象とし、各グループの要約文を分析すると同時に、被験者へのインタビューを行い、要約のプロセスを、定性的・定量的、両側面から調査・解析した。定性的解析では、被験者へのインタビューと要約文分析の結果から、日本語母語話者及び英語母語話者の要約ストラテジーの基礎的モデルをデザインした。更に、定量的解析では、ノンパラメトリック検定により、2 つのグループの被験者が、要約する際に使用した IU（Idea Unit）使用率に関し検証を行った結果、有意な差が認められた。日本語母語話者被験者は、オリジナルの文章を使用しながら「大意」型の要約を行う傾向にあるが、英語母語話者被験者は、オリジナルの文章を他の言葉で言い換える「要旨」型の要約を行う傾向がわかった。

Keywords: summarizing strategy, idea unit, text-depended, event-depended

1. Introduction

1.1. Background of the study

In the previous few decades some studies have been made to elucidate the human system of summarizing a text. According to Sakuma (1994), "summarizing" means to express the content of an original text with less language without changing the overall content of it. Generally speaking, there are two kinds of "summarizing". One is called an "outline" or "大意" which means to shorten the original text with less language without changing the overall content. And the other is called "summary" or "要旨" which means to shorten the original text with paraphrasing.

In recent years, "summarizing" has been regarded as one of the human cognitive

behaviors and has emerged as a field to be investigated. Van Dijk & Kintch (1983) analyzed how micro-parts and macro-parts influenced the methods of summarizing. Lehnert (1982) divided a sentence into some units which are called "plot unit" and tried to elucidate the causation of every plot unit in the summarized text. In Japan, Sakuma (1989) prescribed a "Z unit" and took a quantitative approach to elucidate the summarizing process. In addition, Muramoto (1998) defined twelve IUs and tried to extract the universality from the methods of the human system of summarizing a text.

    As I mentioned above, some studies have been conducted to understand the human process of summarizing, however, no research has reached the complete elucidation of the human system of summarizing a text or proposed a universal model for the summarizing process.

1.2. Purpose of the study

        The aim of this research is not to score any summarized text but to analyze each summarized text both quantitatively and qualitatively to understand the differences in the methods of summarizing between Japanese and English speakers. Additionally, I tried to extract the flow of judgment of subjects on their consciousness in summarizing a text by interviewing.   The main purpose of this study is to propose two types of basic models of summarizing process and compare their features according to the results.

        In terms of some studies to elucidate the human summarizing system in Japan, Japanese subjects were required to summarize a text written in Japanese or Japanese learners (non-Japanese speakers) were required to summarize a text written in Japanese. Or some studies tried to investigate the relevance of summarizing method and reading ability of subjects. (Katori, 2009) However, in this paper, I focused on understanding the summarizing strategies of two groups. Therefore I prepared two types of original texts, one is in Japanese and the other is in English. Japanese speakers were required to summarize a Japanese text and English speakers were required to summarize an English text. Then two types of the summarized texts were examined by the researcher.

2. Method

2.1. Participants

    20 Japanese speakers and 20 English speakers took part in this study. All of them were in their late-teens to 60s. As for Japanese speakers, fifteen were enrolled in or graduated universities and five were enrolled in or finished graduate schools. On the other hand, sixteen English speakers were enrolled in or graduated universities and four were enrolled in or finished graduate schools. Therefore, it was presumed

that there was no difference in the academic careers between Japanese and English speakers.

2.2. Materials

Two original texts were adopted for this study. "戦争は解決でなかった", Asahi Shinbun editorial of September 11, 2003 (1,079 characters) was presented to Japanese speakers and from the International Herald Tribune, VIEWS / EDITORIALS & COMMENTARY "The Perils of Fighting 'Terror'-The Boston Globe" from September 12, 2003 (484 words) for English speakers. To confirm the readability level, the English text was examined by readability calculation software (Readability Calculations version 6). After the analysis, the English text was admitted as "college level" by "The Flesch Reading Ease Formula". Regarding the Japanese Material, any reliable readability calculation software to measure Japanese readability level was not found. However, both were parts of newspaper articles (editorial parts) and were expected that there were no distinct differences in terms of readability. Furthermore, since the main theme of both texts was about "Synchronized terrorist attacks" on September 11th, 2001, Japanese speakers as well as English speakers were expected to have equivalent background knowledge about the content. Additionally, since both original texts were written on an A4-sized sheet of paper, the quantities of both texts were almost the same.

The validity to use editorial parts for the study of methods of the human system of summarizing was described by Sakuma (1989). According to Sakuma, since the editorial part is rather easy to have the theme grasped, it may be said that the editorial part was suitable to examine the influence of the macro-level structure in the summarized text.

As for the structures of languages, Yamanaka (1998) and Torikai (2001) mentioned translated texts are influenced by the structure of the target language. Therefore, in this study, subjects were requested to summarize an original text according to their mother tongue to avoid the influence of translation.

2.3. Procedures

First, subjects were requested to read an A4-sized paper. The Japanese speakers were asked to read an original Japanese text. Likewise, English speakers were requested to read an original English text. After reading the original text, both subjects were asked to summarize them in around 50% of the original material. However, they were also told that the quantity need not be strictly adhered to. The required quantity was decided by the results of the pilot study which was carried out beforehand. In addition, according to Muramoto (1998), most of subjects generate from 20% to 50% length of summarized text when they do not have any limitations on quantity.

In this paper, "summarizing rate" is defined as the number of characters in a summarized text divided by the characters in the original text (characters in the summarized text ÷characters in the original text) or (words in the summarized text ÷ words in the original text). The former is for Japanese texts and the latter is for English texts.

As for the examining period, subjects were told that the tentative time limit to generate their summarized texts was around 30-40 minutes. However, at the same time, they were told that the tentative time limit was not absolute. This tentative time limit was also decided because of the result of the pilot study. Even if subjects spent much longer or shorter time to summarize the text, there was no restraint from the examiner. Although the subjects could summarize by taking memos if they wish, they were requested not to consult a dictionary even if they found some words that they did not understand or did not know.

As I mentioned above, in general there are two kinds of "summarizing." One is called "outline" and the other is called "summary". In this research I did not specify which type of "summarizing" should be followed because one of the purposes of this study is to investigate which type of "summarizing" was adopted by the subjects.

To compare quantitative features of both groups, "idea unit analysis" was carried out. Over the past few decades, a considerable number of criteria to define IU were decided and adopted by researchers (Lehnert, 1982; Carrell, 1985; Ikeno, 1996; Kimura, 1999). In the present study, the criterion of Ikeno (1996) which was adopted by Kimura (1999) was referred to for the English text. In terms of Japanese text, the guideline made by Muramoto (1998, p.104) was referred to. According to those criteria, each original text was segmented into IUs. (See appendix 1.)

After generating a summary, the number of IUs taken from the original text and the rate of using IUs in the generated text were examined. In this research, "rate of using IUs" is defined as the number of IUs in a summarized text divided by the IUs in the original text (IUs in the summarized text ÷IUs in the original text). Then, the result of the "idea unit analysis" was examined by χ^2 test to investigate the difference of distribution by mother tongues.

As a qualitative approach, each generated text was analyzed to investigate the features of the method of summarizing. Additionally, each subject was interviewed after that or in a few days later to understand and confirm their summarizing strategy. During the interview, each subject was asked mainly about the process of summarizing, that is, reasons for sentence reduction, lexical paraphrasing and so on. Subjects, who found it impossible to have an interview in a few days, answered the same questions in an email or over the phone. Even if the subjects did not answer the questions nor had no idea, the examiner did not oblige them to answer.

Finally, both results were combined and each (Japanese-speakers and English-speakers) summarizing model was proposed.

3. Results

3. 1. Results of quantitative approach

　　　　　The following *Tables* show the results of idea unit analysis. "J" stands for "Japanese speaker" and "S" means "Subject". Likewise, "E" stands for "English speaker" and "S" means "Subject".

Table 1. Result of Japanese Speakers

Subject	JS1	JS2	JS3	JS4	JS5	JS6	JS7	JS8	JS9	JS10
Number of IUs	38	24	27	25	6	20	9	12	11	7
Rate of using IUs (%)	48	30	34	32	8	25	11	15	14	9
Summarizing rate (%)	50	52	50	52	46	46	42	43	42	49
Subject	JS11	JS12	JS13	JS14	JS15	JS16	JS17	JS18	JS19	JS20
Number of IUs	19	25	16	23	8	34	12	6	5	13
Rate of using IUs (%)	24	32	20	29	10	43	15	8	6	16
Summarizing rate (%)	42	46	44	38	37	47	46	45	48	44

Table 2. Result of English Speakers

Subject	ES1	ES2	ES3	ES4	ES5	ES6	ES7	ES8	ES9	ES10
Number of IUs	0	2	0	1	0	7	1	6	0	0
Rate of using IUs (%)	0	2	0	1	0	7	1	6	0	0
Summarizing rate(%)	26	50	40	50	19	43	45	58	22	52
Subject	ES11	ES12	ES13	ES14	ES15	ES16	ES17	ES18	ES19	ES20
Number of IUs	0	6	36	22	1	2	3	7	0	5
Rate of using IUs (%)	0	6	37	22	1	2	3	7	0	5
Summarizing rate(%)	54	50	50	52	52	37	49	52	15	52

　　　After the idea unit analysis, the results were examined by χ^2 test to investigate the difference of distribution by mother tongues. Japanese Speakers were defined as the expectation frequency.

　　　Null hypothesis: The distribution of the rate of using IUs of the English speaker fits the distribution of the rate of using IUs of the Japanese speaker.

　　　Significance level: 1%

　　　Degree of freedom:　6

Table 3. shows the rate of using IUs by mother tongues.

Table 3. The rate of using IUs by mother tongue.

Rate of using IUs	X≦10%	15%	20%	25%	30%	35%	35%＜X
English speaker（population）	18	0	0	0	1	0	1
Japanese speaker（population）	5	4	2	2	2	3	2

X= Rate of using IU

The result of the *Table* 3. was examined in terms of goodness of fit by χ^2 test. χ^2=45.8

When the degree of freedom is 6, the limit value for 1% of significance level is 16.8117 （16.8117≪45.8）. Therefore, the null hypothesis is denied. Furthermore, according to *Table* 3, the possibility to have "χ^2=45.8" is $3.2×10^{-8}$.

Thus, clear difference is admitted between distributions of the rate of using IUs by mother tongues. When the adopted criteria to divine IUs in this study are assumed to be appropriate, it can be presumed that the distribution of the rate of using IUs may depend on mother tongues. In other words, limited as this research was, both language groups had tendencies concerning the rate of using IUs. Japanese speakers have a tendency to try to use original phrases frequently, on the other hand, English speakers are presumed that they try not to use original expressions.

3. 2.　Results of qualitative analysis
3.2.1. Analysis of summarized texts

As a consequence of the analysis of the summarized text, some clear features were found in terms of the summarizing process between Japanese and English speakers.

The majority of Japanese speakers summarized the text closely following the original text from the top to the bottom.　For this reason, their summarized texts closely followed the consecutive order of the original text. Even if they eliminated some parts or paragraphs, it was rather easy for the researcher to identify the eliminated parts.　On the other hand, English speakers seemed to have focused on the main idea of the original text. Most of them constructed the summarized text by paraphrasing with their own words and phrases. Since their summarized texts did not follow the consecutive order of the original text, it was often difficult to identify the paraphrased or eliminated parts.

In addition, Japanese speakers had a tendency to use the parts which were put in a bracket such as「憎悪の思想」→「憎悪の思想」.　However, English speakers were apt to paraphrase the parts even if they were put in quotation marks such as "crusader America" → "Justify US action".

Another feature of English speakers were "generalization" and "specification", they are also called "free enrichment" according to Higashimori, Yoshimura (2003).

For example, though three subjects added the phrase "September 11" in their summarized text, the concrete date was not seen in the original text. Similarly, another subject added an actual proper noun that was not seen in the original text. Those tendencies were not seen in the Japanese speakers.

　　　Typical examples summarized by both subjects are shown in the "Appendix 2". Sampled Japanese summarized text was generated closely following the original's consecutive order and featuring the use of assertive expressions eliminating the inference and suggestion. Italicized characters are the sentences and expressions adopted from the original text as they were.　In contrast, sampled English speaker tried to express the main idea of the original text with his own words. Since only one part was almost identical to the original as it was (the last sentence of the summarized text), the examiner asked the reason in the interview. Then the subject answered "This is partly due to my laziness…."

3. 2. 2. Results of interview

　　After generating a summary, each subject was interviewed after that or a few days later. Even if the subjects did not answer the questions nor had no idea, the examiner did not oblige them to answer.

　　The result of the interview concluded that most of the Japanese subjects adapted a similar procedure to summarize the original text.　That is, Japanese speakers summarized a text from the top to the bottom deleting the unimportant parts and leaving the important parts according to the consecutive order of the original text. Few Japanese speakers tried to paraphrase the original phrases and sentences. On the other hand, the majority of English speakers tried to identify the main idea or the theme of the original text then paraphrased the text with their own words and phrases. In other words, Japanese speakers showed a tendency to try to use the phrases or sentences from the original text as they were. English speakers, in contrast, showed a tendency to avoid using the same expression as the original text. That was one of the most contrastive points in terms of the summarizing method between Japanese and English speakers.

　　According to the interview, some of the English speakers mentioned that they considered "paraphrasing" as an important process of summarizing texts because "understanding" is necessary for summarizing. Additionally some of them mentioned that "summarizing" is the result of "understanding and paraphrasing". On the other hand, it was not likely that the Japanese subject considered "summarizing" was the process of understanding texts. However, it doesn't mean Japanese subject didn't understand the original text. According to the interview, the majority of Japanese subjects considered "summarizing" as the result of reduction and extraction from the original. Therefore, they thought they should not paraphrase important parts of the original text.　Although some Japanese speakers paraphrased some phrases from the

original text, the cases were very limited.

4. Summarizing models

According to the result of the analyses and the interview, I would like to design two basic models in summarizing a text. I should point out that these proposed models are able to cover the summarizing process of the subjects in this research; however, it might not be applicable as a universal process of human system of summarizing for all texts.

4.1. Summarizing strategy (Japanese speaker)

The following diagram shows the flow of the summarizing process taken by the majority of Japanese speakers. First, subjects read an original text (1. text inputting). Next, subjects elaborated by choosing the part or sentence to extract (2. text elaborating). During (2), subjects were supposed to adopt the following five procedures: a) exclusion of ambiguity, b) lexical paraphrase, c) screening examples, d) sentence combination, e) exclusion of inference and suggestion.

Explanations of details are shown in the following diagram:
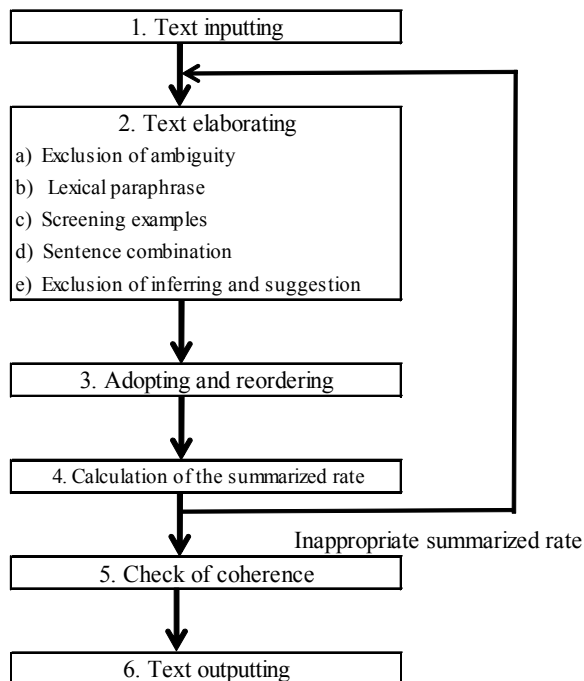
Fig.1 Basic summarizing strategy model of Japanese speakers

a) **Exclusion of ambiguity:** Subjects (Japanese Speakers) tended to remove

ambiguous or incomprehensible parts in the original text. For example, a subject who considered the phrase ”民衆の生活もままならない” ambiguous, paraphrased the part to ”民衆の生活は混乱している”. Another subject also paraphrased the part to ”民衆の生活は安定しない”.

　　　b) **Lexical paraphrase:** According to the interviews, subjects also tended to change some words and phrases from unfamiliar expressions to more familiar ways. “温床” → ”背景”, “撲滅”→“根絶”, “立ちすくんでいる” →“足踏み状態” are examples. In the interview, subjects who changed some expressions mentioned that they would not like to use unfamiliar expressions in their summarized text. Only one subject answered that he changed a part even though the original expression was more familiar than the paraphrased text. He mentioned he was obliged to change that part because of the summarized rate.

　　　c) **Screening examples:** When an original text contains some concrete examples, subjects are apt to eliminate all or some of them. The original Japanese text contains concrete proper nouns in its first part. Three subjects adopted the original sentences with all concrete nouns. However, five subjects adopted some of them and twelve eliminated the whole sentence.

　The 13th paragraph also contains some concrete name of places such as “パレスチナやインドネシア”, “チェチェン”. Twelve subjects used no concrete examples, seven used parts of them and one used all of them.
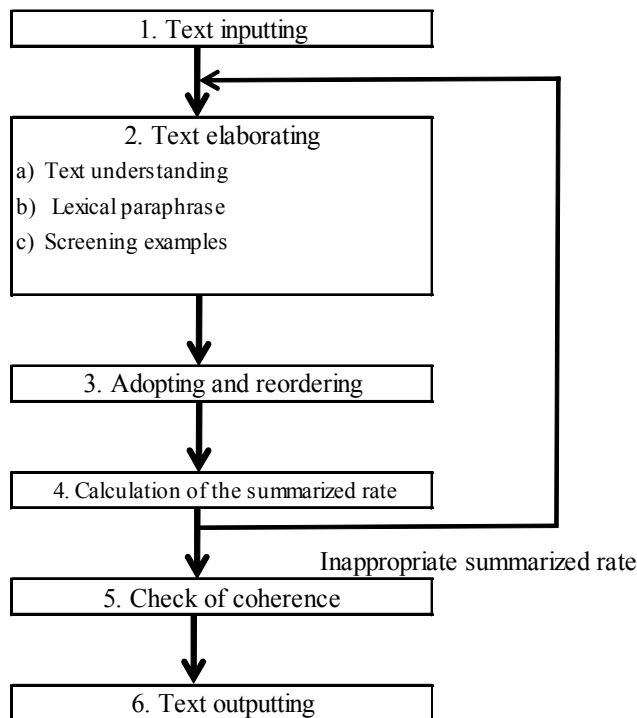
　　　d) **Sentence combination:** When subjects found sentences or phrases which have rather strong cohesion, they tended to combine them. From the 4th paragraph to the 6th paragraph of the original text, some words were used with repetition such as “アフガニスタン”(twice), “イラク”(twice), “テロ”(three times), “テロリスト”(three times), “政権”(three times) and so on. In those three original paragraphs, 20% out of all characters (248 characters) were occupied by those words and phrases. Although methods were diverse, all twenty subjects combined those parts somehow and reconstructed new sentences. In the interview, 16 subjects out of 20 mentioned that they thought of those three paragraphs “as one part” and combined sentences on purpose. Others answered they did not do that on purpose, however, the result was they combined those parts.

　　　e) **Exclusion of inference and suggestion:** Subjects were apt to exclude inferring expression or suggestion during their summarization. “～だろう”→“～ではないか”, “～である” are the examples. According to the interview, they mentioned that they would like to make the author's opinion clear. After the elaboration noted above, the subjects adopt sentences which will be used in their summarized text, then they calculate the summarized rate. The average of summarized rate was 45.5%. If the rate is inappropriate, they go back to “2” (text elaboration) or go to “5” (check of coherence). On “5” (check of coherence), subjects review conjunctions or particles to improve their sentences and the coherence. Finally, subjects output their summarized text as the

result of their elaboration.

4.1. Summarizing strategy (English speaker)

Likewise, Fig.2 shows the flow of the summarizing process used by the majority of English speakers. The distinct tendency of English speakers was "paraphrasing". Different from Japanese speakers, English speakers tended to pay attention to three points specifically, "text understanding", "paraphrasing" and "screening examples".

```
┌─────────────────────────────┐
│      1. Text inputting       │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     2. Text elaborating      │
│ a) Text understanding        │
│ b)  Lexical paraphrase       │
│ c) Screening examples        │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│   3. Adopting and reordering │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ 4. Calculation of the summarized rate │
└─────────────────────────────┘
              │          Inappropriate summarized rate
              ▼
┌─────────────────────────────┐
│    5. Check of coherence     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      6. Text outputting      │
└─────────────────────────────┘
```

Fig.2 Basic summarizing strategy model of English speakers

As for their summarizing flow, first, subjects read an original text (1. text inputting); next, subjects elaborated by choosing the part or sentence to extract (2. text elaborating). According to the interview, the majority of the English speakers answered that they focused on "text understanding", "paraphrasing" and "screening examples" during (2). Similar to Japanese speakers, when an original text contains some concrete examples such as proper nouns, English speakers were apt to eliminate all or select some of them. *Table* 4. shows the actual examples paraphrased by the subjects.

Table 4.　Examples paraphrased by English speakers

original expression	paraphrased expression
provoke	welcome
domestic struggle	local struggle, domestic strife
standing behind	backbone, back fire, supporting
repressive regimes	oppressive regimes
turn up	arrive
United States	America

According to the interviews, subjects also tended to paraphrase some words and phrases from unfamiliar expressions to more familiar ones. The distinct features of the English speakers' were "understanding" and "paraphrasing". Seven out of twenty subjects mentioned that they tried to find the "main idea" of the text before they began to summarize it. This means that English speakers tried to understand the original text first then they paraphrased sentences or general idea with their expressions. In terms of "Screening examples," English speakers have a tendency to eliminate all or select some of them from the original text. Although the English original text contains some concrete proper nouns such as "Niger, Chad, Mali, Djibouti or Uzbekistan", few subjects adopted all of them in their summarized text.

After the elaboration, they calculated the summarized rate as well as Japanese speakers. The average of summarized rate was 43.4%. If the rate is inappropriate, they go back to "2" (text elaboration) or go to "5" (check of coherence). On "5" (check of coherence), subjects review conjunctions or particles to improve the coherence of their generated text. Finally, subjects output their summarized text.

5. Conclusion

In this paper, I tried to identify the differences in the process of summarizing texts between Japanese speakers and English speakers. The aim of this research is not to score any summarized texts but to analyze each summarized text both quantitatively and qualitatively to identify the differences in the methods of summarizing between Japanese and English speakers.

As a consequence of the quantitative approach, idea unit analysis was carried out and a clear difference is admitted between distributions of the rate of using IUs by mother tongues. Therefore, in this research, it can be presumed that the distribution of the rate of using IUs may depend on mother tongues.

Regarding the results of interviews and qualitative analysis of generated texts, most of all Japanese speakers summarized the text closely following the consecutive order of the original text.　Furthermore, Japanese speakers had a tendency to use the

phrases or expressions from the original.　And this is supported by the result of qualitative analysis in this research. On the other hand, English speakers reconstructed the text by paraphrasing with their own words and phrases.

　　It follows from what has been said that Japanese subjects in this research considered "summarizing" to "shorten a text" without changing the content. In contrast, English speakers in this research considered "summarizing" to "understand the context and to restate the main idea" with paraphrasing.

　　From the above-mentioned reasons, it would be concluded that Japanese speakers tend to generate "outline-type" summarization. On the other hand, English speakers tend to generate "summary-type" summarization. In other words, Japanese speakers are apt to adopt the "text-depended method" (Muramoto, 1998), whereas, English speakers adopt the "event-depended method" (Muramoto, 1998).

　　　The point that requires clarification is that the result of this study was limited by those subjects, and it may not be clear whether it may be presumed that those features represent the universal method of summarizing of Japanese and English speakers. To elucidate the methods of the human system of summarizing a text, it is undeniable that further studies are needed.

References

Alderson, J.C. (1984). Reading in a foreign language: a reading problem or a language problem? In J.C. Alderson and A.H.Urquhart(eds.) *Reading in a Foreign Language.*London:Longman.

Carrell, P. L.(1985). Facilitating ESL reading comprehension by teaching text structure. *TESOL Quarterly, 19*, 727-752.

Carrell, P. L.(1991).Second language reading: Reading ability or language proficiency ? *Applied linguistics,12*, 159-179.

Clark, M.A.(1980). The short circuit hypothesis of FEL reading—Or when language competence interferes with reading performance. *Modern Language Journal,64*, 203-309.

van Dijk, T.　A.,　& Kintch, W. (1983). *Strategies of discourse comprehensi*on. New York: Academic Press.

Endres,B.N(1998). *Summarozong Information*. Berlin/Heidelberg:Springer-Verlag

東森勲, 吉村あき子（2003）『関連性理論の新展開□認知とコミュニケーション□』研究社

Ikeno, O. (1996). The effects of text –structure-guiding question on comprehension of texts with varying linguistic difficulties. *JACET Bulletin,27*, 51-68.

Inderjeet, M. (2001). *Automatic Summarization,* Amsterdam / Philadelphia: John Benjamins.（＝奥村学, 難波英嗣, 上田禎子訳 2003 『自動要約』共立出版）

門田修平, 野呂忠司（2001）『英語リーディングの認知メカニズム』くろしお出版

Katori, M. (2009). Relevance-First Language and Foreign Language Summarizing Strategy:Pertaining to Reading ability－Case Studies－. *Journal of Aomori Public College,15(2),* 17-25.

Kimura, Y. (1999). Metacognitive awareness training and reading comprehension of Japanese EFL learners. *JACET Bulletin,30*, 45-58.

Lee, J.-W. and Schallert, D. L.(1997) The relative contribution of L2 lamguage proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly,31*, 713-739

Lehnert, W. G. (1982). Plot units: A narrative summarization strategy. In W. G. Lehnert & M. H. Ringle (eds.), *Strategies for natural language processing* , (pp. 375-412). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

益岡隆志，仁田義雄，郡司隆男，金水敏（1997）『岩波講座 言語の科学 5 文法』岩波書店

増山繁，山本和英（2002）「テキスト自動要約における新たな展開と展望－統計的方法，換言処理，そして・・・－」情報処理学会論文誌『情報処理』43 (12), 1310-1316

邑本俊亮（1998）『文章理解についての認知心理学的研究－記憶と要約に関する実験と理解過程のモデル化－』風間書房

長尾真，黒橋禎夫，佐藤理史，池原悟，中野洋（1998）『岩波講座 言語の科学 9 言語情報処理』岩波書店

佐久間まゆみ 編(1989)『文章構造と要約文の諸相』くろしお出版

佐久間まゆみ 編（1994）『要約文の表現類型』ひつじ書房

鳥飼玖美子（2001）『歴史を変えた誤訳』新潮社

山中桂一(1998) 『－対照言語学からのアプローチ－日本語のかたち』東京大学出版会

Software

Readability Calculations Set version 6.(preferred edition for Windows)[computer software].(2004) Texas :Micro Power & Light Co.

Appendix 1

Idea unit (English text)

1	Two years after	51	But a promiscuous entanglement
2	a gang of reactionary religious zealots flew commercial airplanes	52	in the internecine
3	into the World Trade Center and the Pentagon	53	conflicts of countries ruled by vicious dictators
4	President George W. Bush commonly speaks	54	risks a strategic blunder
5	about a diffuse global war on terrorism	55	Pursuing Bush's indiscriminate war on terrorism
6	and a former CIA chief	56	U.S. military and intelligence personnel are currently
7	James Woolsey	57	on the ground
8	has even described the conflict as	58	cooperating with the corrupt and repressive regimes

	World War IV		
9	as World War IV	59	of among others
10	There is a danger	60	Islam Karimov in Uzbekistan and Maaouyah Ould Sid Ahmed Taya of Mauritania
11	in this imprecise way	61	These undemocratic rulers exaggerate the threat
12	of naming an enemy	62	of small Islamist groups
13	and invoking a worldwide struggle	63	in their lands
14	of indefinite duration	64	to excuse thuggish suppression
15	The danger is that repetition	65	of all dissent
16	of these misleading definitions will be used	66	Washington allows itself
17	to rationalize an antiterrorist strategy	67	to be identified with these and other despotic clients
18	that embroils Americans unnecessarily	68	in the war on terrorism
19	To fall into this trap is	69	opposition movements
20	to play into the hands of Osama bin Laden and his associates	70	against despised local rules
21	who would like	71	could be transformed
22	to provoke the global holy war	72	into anti-imperialist struggles against America
23	they preach	73	as the reviled foreign power
24	Instead of loose talk	74	behind the local tyrant and his torturers
25	that acts like America's moral enemy is an abstract noun-	75	This is precisely
26	terrorism	76	what bin Laden and his Egyptian partner Ayman al-Zawahiri allege
27	in all its forms and manifestations-	77	in their statements
28	Bush would be wise	78	for public consumption
29	to distinguish Al Qaeda and the groups affiliated with it	79	-that
30	from Islamist movements	80	crusader America
31	that may be trying	81	is the
32	to overthrow regimes	82	far power
33	in their own countries	83	standing behind collaborationist regimes
34	but have not declared war against the United States	84	in Muslim countries
35	When U.S. intelligence agents	85	that the Al Qaeda ideologues define
36	and armed forces turn up in countries such as Mauritania	86	as the
37	Niger	87	near power
38	Chad	88	Nothing would better suit Al Qaeda's recruiting tactics
39	Mali	89	than to be able to point to U.S
40	Djibouti	90	backing for the Karimovs
41	they appear	91	of the Muslim world as proof
42	to be waging the global war	92	that America is at war with all Muslims
43	against all Islamists	93	not merely the fanatics of Al Qaeda
44	that bin Laden invokes	94	who target Americans
45	and Washington appears	95	Bush must not nourish that fantasy
46	to be validating the grandiloquent ideological claims of Al Qaeda	96	by confusing a campaign
47	There is a genuine	97	against one aberrant Islamist faction
48	need for intelligence and law-enforcement cooperation	98	with a world war against all terrorists sts
49	against Al Qaeda and its affiliates		
50	－American's declared enemies		

Idea unit (Japanese text)

#		#	
1	ビンラディン様。	41	事態が出現している。
2	私の息子は大きな夢を持ち、	42	力で世界を牛耳ろうとする
3	アメリカで懸命に生きていました。	43	米国への憎しみ。
4	なぜ、	44	その背景にある貧困。
5	あなたは罪もない人々を巻き添えにしたのか。	45	泥沼のパレスチナ問題。
6	世界貿易センタービルに対するテロ攻撃で	46	同時多発テロの温床として、
7	一人息子を失った白鳥晴弘さんは、	47	多くのことが語られてきた。
8	首謀者とされるビンラディンあての手紙にそう書いた。	48	だが、
9	パキスタンに駐在するアラビア語放送の記者に	49	国際社会は、
10	「本人は無理でも、できるだけ近い人に届けてほしい」	50	テロをどう撲滅できるのか
11	と託した。	51	今でも立ちすくんでいる。
12	深い憤りと悲しみと	52	米国がその軍事力に物を言わせて
13	無数の「なぜ」を残したあの日から、	53	戦争に勝っても、
14	もう2年である。	54	テロをなくせるわけではなかった。
15	世界は変わった。	55	様々な手段でテロを抑え込みながら
16	アフガニスタンでアルカイダの拠点をつぶし、	56	温床をなくしていくという、
17	タリバーン政権を打倒したブッシュ米政権は、	57	長くたゆまぬ努力が基本ということではあるまいか。
18	次の目標としてイラクを攻撃し、	58	白鳥さんは、
19	政権を転覆させた。	59	子息の遺産や補償金を基に
20	大量破壊兵器をテロリストに渡さないための	60	アフガニスタンの子どもたちを助ける
21	「対テロ戦」	61	プロジェクトを立ち上げようとしている。
22	とされた。	62	「テロをなくすには、そこから始めるしかない」
23	テロとの関連が明確で、	63	と考えたからだ。
24	国際社会の広い支持を得たアフガニスタンでは	64	国際社会もそんな思いに応えてほしい。
25	戦後いち早く暫定政権ができ、	65	最も大事なのは緊密な協調態勢だろう。
26	復興が動き出した。	66	米国はこの原点に返るべきだ。
27	ところが、	67	欧州とも、
28	イラク戦争の状況は悪化するばかりだ。	68	アラブ・イスラム諸国とも連携して
29	国連安保理と欧州の主要同盟国を振り切って開戦し、	69	テロ撲滅のための戦略を立て直すことだ。
30	何とか制圧したものの、	70	パレスチナやインドネシア、
31	米英軍への攻撃がやまない。	71	チェチェンと、
32	国連事務所さえテロの標的となった。	72	それぞれ異なる政治的な背景を持つテロでも、
33	大量破壊兵器は見つからず、	73	一つのテロが別な地のテロを呼ぶ。
34	民衆の生活もままならない。	74	とくにパレスチナ紛争の政治解決は
35	フセイン政権からの解放という	75	テロ撲滅の視点からも急務である。
36	成果を帳消しにしかねない混乱の中で、	76	テロリストの根を絶つために、
37	イラクが	77	イスラム諸国自身の民主化努力も求められる。
38	「憎悪の思想」	78	テロのない世界への新たな出発点として、
39	に支えられた	79	この9・11を位置づけたい。
40	新たなテロの巣窟になったかのような		

Appendix 2

Sampled generated text by a Japanese speaker（JS2）

　　　　　米国の世界貿易センタービルへのテロ攻撃から２年が過ぎた今、*世界は変わった。*
アフガニスタンでアルカイダの拠点をつぶし、タリバーン政権を打倒したブッシュ米政
権は次の目標としてイラクを攻撃し、政権を転覆させた。大量破壊兵器をテロリストに
*渡さないための「対テロ戦」*とされたのである。

　*テロとの関連が明確*で、国際社会の広い支持を得たアフガニスタンの戦後とは異なり、
イラクの戦後の状況は悪化しつづけている。

　力で世界を牛耳ろうとする米国への憎み。その背景にある貧困。泥沼のパレスチナ問
*題*など*同時多発テロの温床として多くのことが語られてきた。*

　軍事力に物を言わせて*戦争に勝っても*テロはなくせなかった。*様々な手段でテロを抑*
*え込みながら、*テロの温床をなくしていく長くたゆまぬ努力が基本なのではないか。

　最も大事なのは世界各国の緊密な協調体制である。米国は欧州、*アラブ・イスラム諸*
*国と連携して*テロ撲滅のための戦略を立て直すべきである。

　異なる政治的背景を持つテロでも*一つのテロが別の地のテロを呼ぶ。*

　パレスチナ紛争の政治解決は*テロ撲滅の視点からも急務である。テロリストの根を絶*
つためにイスラム諸国自身の民主化努力も求められる。

　テロのない世界への新たな出発点として、この９・11 を位置づけたい。（529 字）

Sampled generated text by an English speaker（ES6）

　　The prosecution of the global "War on Terror", declared by President George W. Bush in the wake of the attack on the World Trade Center, has created a powerful dynamic in the U.S.'s relationship with various countries around the world which might ultimately prove damaging to U.S. interests. Of particular concern in the prosecution of this war has been the failure to draw a clear distinction between Al Qaeda and affiliated groups, and other Islamist movements whose stated goals relate exclusively to local regimes or matters specific to their geographic vicinity.　Al Qaeda and its affiliated international Islamist movements have sought to galvanize sympathizers across the Muslim world by claiming that "crusader America" is the "far power" standing behind despotic regimes of local tyrants, the "near power". The presence of U.S. Intelligence agents and armed forces in countries such as Mauritania, Niger, Mali, Djibouti or Uzbekistan, is often interpreted as a nod of approval of their repressive regimes, under the guise of countering the threat from Islamist groups. Such actions fit neatly into Al Qaeda's image of an American puppet master waging war on all Muslims. *Bush must not nourish that fantasy by confusing a campaign against one Islamist, anti-American faction with a world war against all terrorists.* (208　words)

　　　　　　　　　　　　　　　* 　Italicized: close expression to original text