ICOT Technical Report: TR-125

TR-125

Proofreading Japanese Word Usage and Proper Nouns by Computer

by
Satoru Ishii
(Toshiba Corp.)

July, 1985

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191～5
Telex ICOT J32964

Institute for New Generation Computer Technology

Proofreading Japanese Word Usage and Proper Nouns by Computer

Satoru Ishii

Information and Communication Systems Laboratory
TOSHIBA Corporation, Kawasaki, Japan

ABSTRACT

A Japanese Proofreader System for Japanese text has been constructed. This system proofreads word usage and proper nouns in text. Word usage is checked using knowledge mainly derived from a stylebook. For proofreading proper nouns, a company directorors database is used to check articles on personnel changes of company directors. Due to limitations of the processing system, experiments were made with about one tenth of the entire volume of data, and the system was proven effective. The system is written in Prolog. The system was used to evaluate Prolog and determine the desirable performance level. Findings obtained during evaluation are also given.

This work was done as a part of the Fifth Generation Computer Project while the author was at the Institute for New Generation Computer Technology (ICOT). The author wishes to thank ICOT for the opportunity of pursuing this research and the permission to present this report.

TABLE OF CONTENTS

1. Introduction 1

2. Correction of word usage 2

2.1 Configuration and knowledge in use 2

2.2 Results of proofreading 6

3. Correction of proper nouns 10

3.1 Configuration and knowledge in use 10

3.2 Results of proofreading 12

4. Evaluation of Prolog 15

4.1 Quantitative evaluation 15

4.2 Qualitative evaluation 16

5. Conclusion 18

References 19

(Final page 20)

1. INTRODUCTION

Computerization of proofreading is much desired at present since proofreading is considered a suitable computer application. In particular, proofreading newspaper articles seems to be a good candidate as the standards of expression newspapers must observe are clear. Accordingly, proofreading currently performed by humans at a newspaper company was studied. The study revealed that it is highly possible that present computer techniques can achieve two types of proofreading: Checking word usage using dictionaries and stylebooks, and checking proper nouns using directories.[1]

Knowledge required for proofreading word usage and proper nouns was investigated and collected [2] [3] to construct the Japanese Proofreader System. This report describes the quality and quantity of knowledge used for proofreading and the experimental results.

The system is written in DEC10 Prolog on the DEC2060. The results of the Prolog evaluation are also reported here.

2. CORRECTION OF WORD USAGE

Fig. 1 illustrates the simplified configuration of the Japanese Proofreader System. The system uses stored knowledge to point out which parts of input sentences are to be corrected and, if possible, to indicate how correction should be made. In actual operation, it was not possible to execute the entire system as a whole due to restrictions of the processing system. Accordingly, the system was divided into two parts to be executed separately: one for correction of word usage and the other for proper nouns. The following describes each type of proofreading. Different procedures are performed to input data in the knowledge base or to edit it (process shown to the left of Fig. 1). These procedures will be described in another report.

2.1 Configuration and knowledge in use

Proofreading to correct word usage was conducted in three steps. This was mainly because of restrictions of the processing system, but also to make implementation easier. The following explanation is based upon Fig. 2.

2.1.1 Conversion of input data format

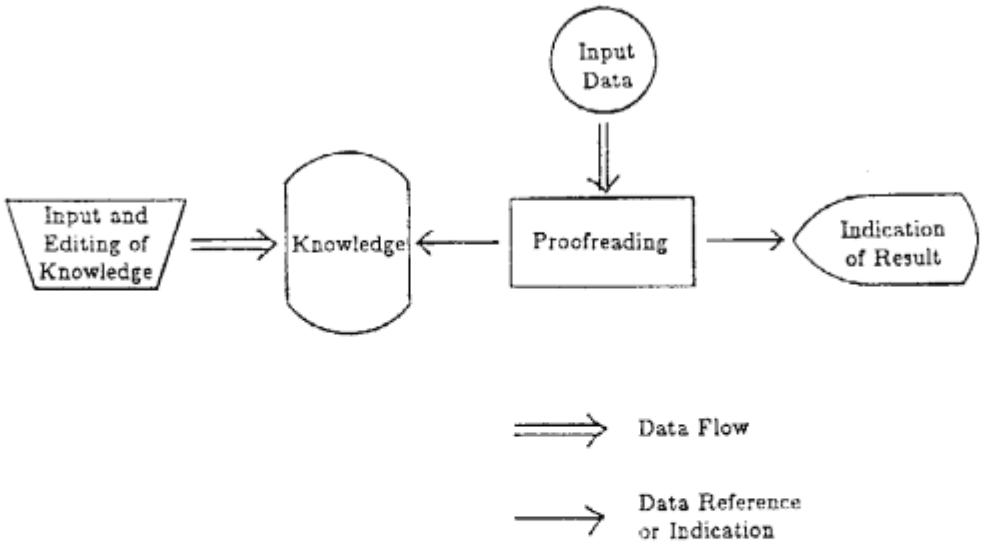

Fig. 1 Entire System Configuration

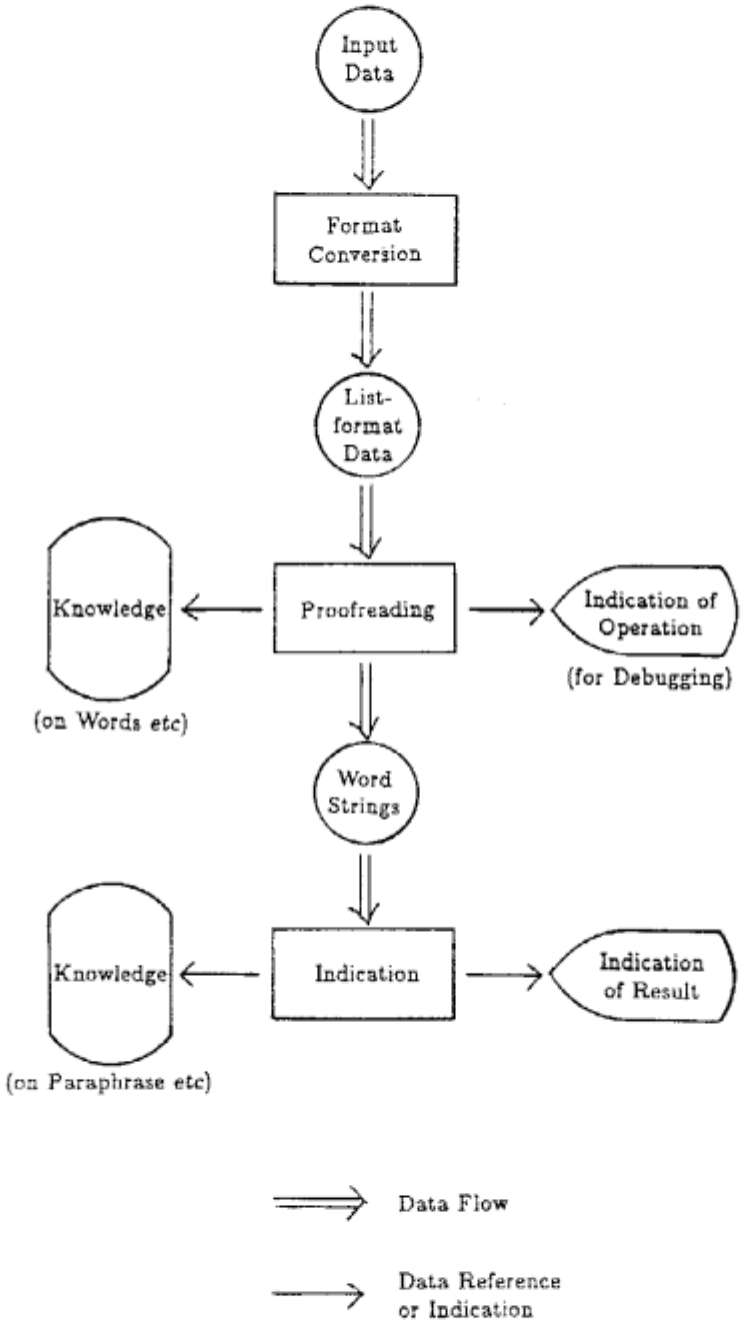

Fig. 2 Proofreading to Check Word Usage

In this step, input data is converted to a list format, in which each component of the list corresponds to one character of the JIS KANJI codes. However, two one-byte ASCII characters are in effect used in the system because DEC 10 Prolog is not provided with KANJI codes.

As procedural processing is performed in this step, details are omitted here.

2.1.2 Proofreading

Of the three steps, proofreading plays the most important role. In this step, a sentence input in the list format is divided into words which are output in list format. Each component of the output list is information on one word. This information is also in the form of a list.

Data output for each word is as follows:
- Word (form in the input sentence)
- Part of speech
- SHUSHI (sentence-final) form
- Inflectional form
- Whether the word is incorrect or not
- Whether the word should be changed or not
- Title number.

The title number is a number given to each word for easy identification. The following describes the knowledge and processing algorithm used in this step.

(1) Knowledge in use

Based upon dictionaries[4], Common KANJI Table [6], and the stylebook being used at the Asahi Shimbun Publishing Co. [5], the following knowledge was used in the system. As details of the knowledge have already been reported [3], only its items, sources, and volume in terms of Prolog are listed below.

- Knowledge on inflection[4]: Knowledge on inflection of verbs and adjectives. About 140 lines.
- Knowledge on postpositional words and auxiliary verbs [4]: Knowledge on the inflection of auxiliary verbs. Knowledge on conjunctions included. About 210 lines.
- Word dictionary [5] [6]: A dictionary containing about 27,000 words was compiled. The dictionary includes incorrect words and words to be replaced by others: These words are provided with information to tag them as such.

(2) Processing algorithm

The following briefly explains the processing algorithm using Fig. 3.

Proofreading is performed by recursively calling the predicate for distinguishing words. In this process, the type of each word (KANJI, HIRAGANA, KATAKANA, alphabet, symbol, numeric, punctuation mark) is judged first, then the appropriate dictionary is retrieved by means of the longest-match method. When the word cannot be found in the dictionary, it is treated as an undefined word up to a certain point of the process. Upon being identified, the word undergoes inflection processing. The proofreading result is output at the end.

2.1.3 Indication of proofreading result

Knowledge and the processing algorithm used in this step in which corrections are indicated are described in the following.

(1) Knowledge in use

The following knowledge based upon the Common KANJI Table and the stylebook used at the Asahi Shimbun Publishing Co. was used. As details of the knowledge have already been reported [3], only its items and volume are listed below.

- Knowledge on sources: Information on the source of each word in the above word dictionary. About 35,000 lines.

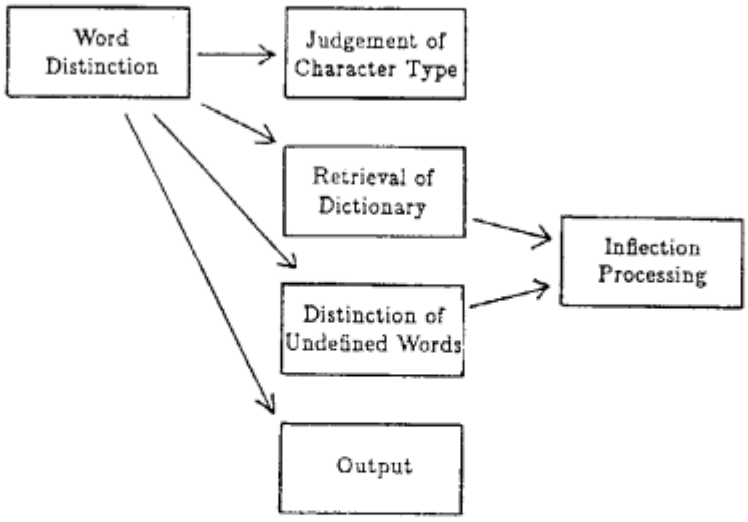

Fig. 3 Proofreading Algorithm

- Knowledge on paraphrase: Knowledge on words to replace incorrect words or those to be changed. About 2900 lines.

(2) Processing algorithm

This step uses a simpler processing algorithm than the previous step. Basically, data output in the previous step is printed as it is. In this step, information on the source of the word is added. When the word is incorrect or should be replaced, it is indicated as such, and the replacement is shown at the same time.

2.2 Result of proofreading

The result of executing the program is as follows. Sentence examples were chosen as much as possible from data on the newspaper articles during the period from July 1 to 10, 1984, received from the Asahi Shimbun Publishing Co. However, they were not the same as the actual sentence which appeared in the paper because it was necessary to omit words and insert mistakes for the purpose of our experiments.

In this experiment to evaluate the system, the entire volume of the word dictionary, knowledge on sources, and knowledge on paraphrases could not be used because of limitations of the processing system. Therefore, only part of the knowledge was used.

- Word dictionary: Knowledge on words with no inflection and those starting from symbols were omitted, which reduced the number of lines to about 3300. Also the reading of each word (KANJI combination) was omitted.

- Knowledge on sources: Only knowledge on the sources of incorrect words or words to be replaced by others was used. As a result, the number of lines became about 6200.

- Knowledge on paraphrase: Only 500 lines were used.

During program execution, the system responded in the cases described below. Whenever the system responded, some kind of problem existed to indicate the necessity of correction. This seems to prove the effectiveness of the system.

2.2.1 Undefined words

Fig. 4 shows the result of analyzing the sentence "足元から雨が降る。"(ASHIMOTO KARA AME GA FURU: It rains from the footstep.) In this example, a noun "雨"(AME: rain) is handled as an undefined word, because words without inflection were omitted from the dictionary. We realize that the word dictionary must be improved as one of the first priorities to reduce the number of undefined words.

Dealing with proper nouns is another problem. This research did cover knowledge on

proper nouns, but this aspect of the system should be improved with respect to both volume and quality. An important object of this kind of system is to collect not only a greater number of people's names but also a greater variety of nouns such as place names and names of organizations.

2.2.2 Incorrect words and words to be changed

Fig. 5 shows the result of analyzing the sentence "足下から雨が降る。"(ASHIMOTO KARA AME GA FURU: It rains from the footstep.) The difference between the example in Fig. 5 and that in Fig. 4 is the expression of the word "ASHIMOTO". The system indicates that "足元" instead of "足下" should be used: This is based upon the knowledge on page 245 of the stylebook used at the the Asahi Shimbun Publishing Co. The system can indicate sources of correct words as well. However, they are not indicated in this experiment as knowledge on sources of correct words was not stored due to restrictions on the processing system.

2.2.3 Mixed use of HIRAGANA and KATAKANA

Fig. 6 shows the result of analyzing the sentence "再検討すべきだと思う。" (SAIKENTOU SU BEKI DA TO OMOU: Reconsideration seems to be necessary.) In this example, "べ"("BE") is written in KATAKANA, which has caused incorrect word distinction. However, even when it is written in HIRAGANA, words still cannot be distinguished correctly as the literary expression "べき"("BEKI") is not included in the system. The problem of how to handle literary expressions has to be solved in the future.

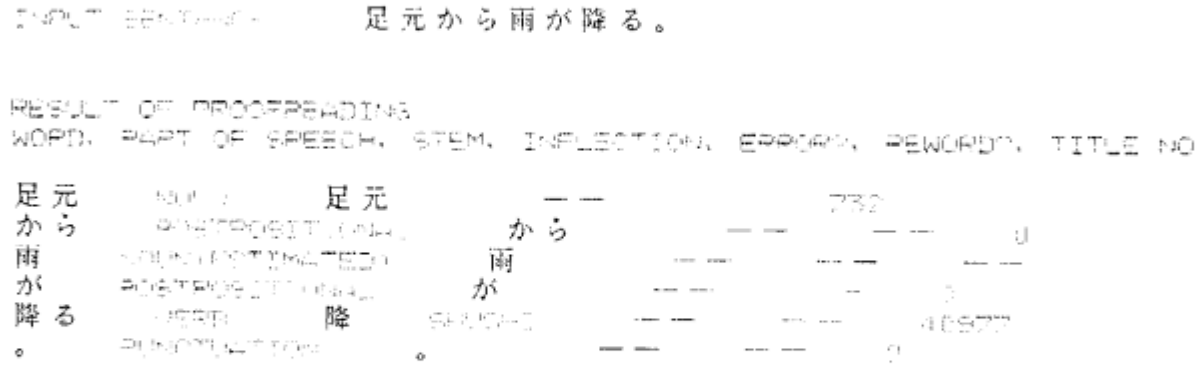

Fig. 4 Proofreading for Word Usage (No. 1)

2.2.4 Words written in KANA

At the end, Fig. 7 shows the result of analyzing the sentence " ゆるやかになった。" (YURUYAKA NI NATTA: It has become gentle.) This is an example of a word commonly written in KANJI designated as Common KANJI, but often written in KANA too. A processing algorithm may be designed to handle such cases. However, the example seems to suggest that knowledge other than the Common KANJI Table is necessary to differentiate the use of KANJI and KANA.

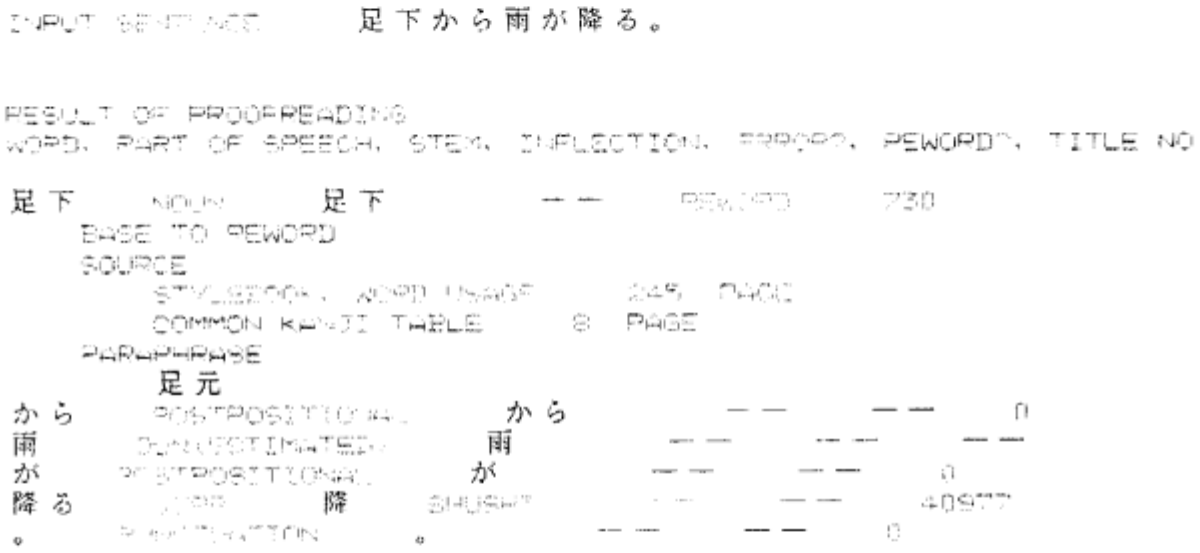

Fig. 5 Proofreading for Word Usage (No. 2)

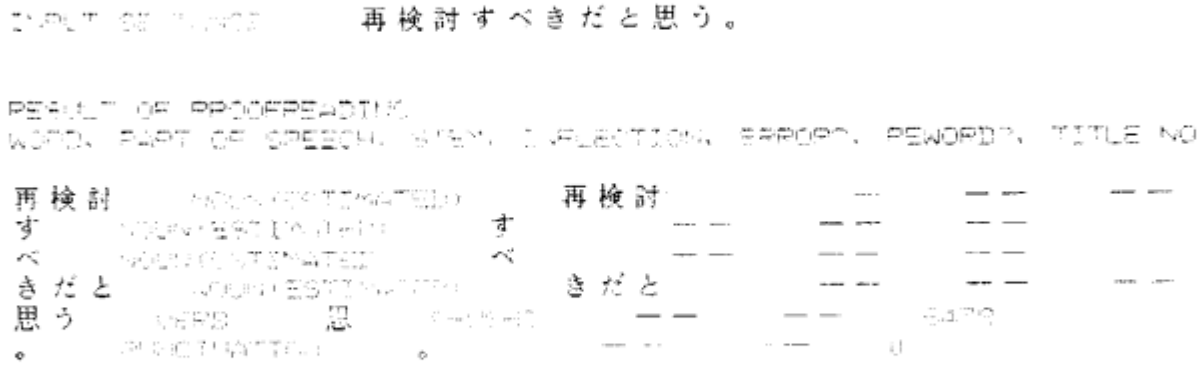

Fig. 6 Proofreading for Word Usage (No. 3)

INPUT SENTENCE　　ゆるやかになった。

RESULT OF PROOFREADING
WORD, PART OF SPEECH, STEM, INFLECTION, ERROR?, REWORD?, TITLE NO

ゆるやかになった	NOUN(ESTIMATED)	ゆるやかになった			
。	PUNCTUATION	。			9

Fig. 7　Proofreading for Word Usage (No. 4)

3. CORRECTION OF PROPER NOUNS

The experiment of proofreading text with respect to proper nouns is described below. To correct proper nouns, in particular, a people's names, the organization he belongs to, and the post he holds, an appropriate directory has to be selected. Criteria for such material are that it be readily available and that it contains names which frequently appear in the text. [2]

The following describes the example of proofreading articles to report personnel changes of directors working at companies listed on the stock market. Although such knowledge is not very often given in articles, data has already been organized as a database, so system application seems effective.[2] Also, an example of handling KANJI characters which are not designated as Common KANJI is described. This experiment was made because KANJI characters not designated as Common KANJI appear in proper nouns rather frequently.

3.1 Configuration and knowledge in use

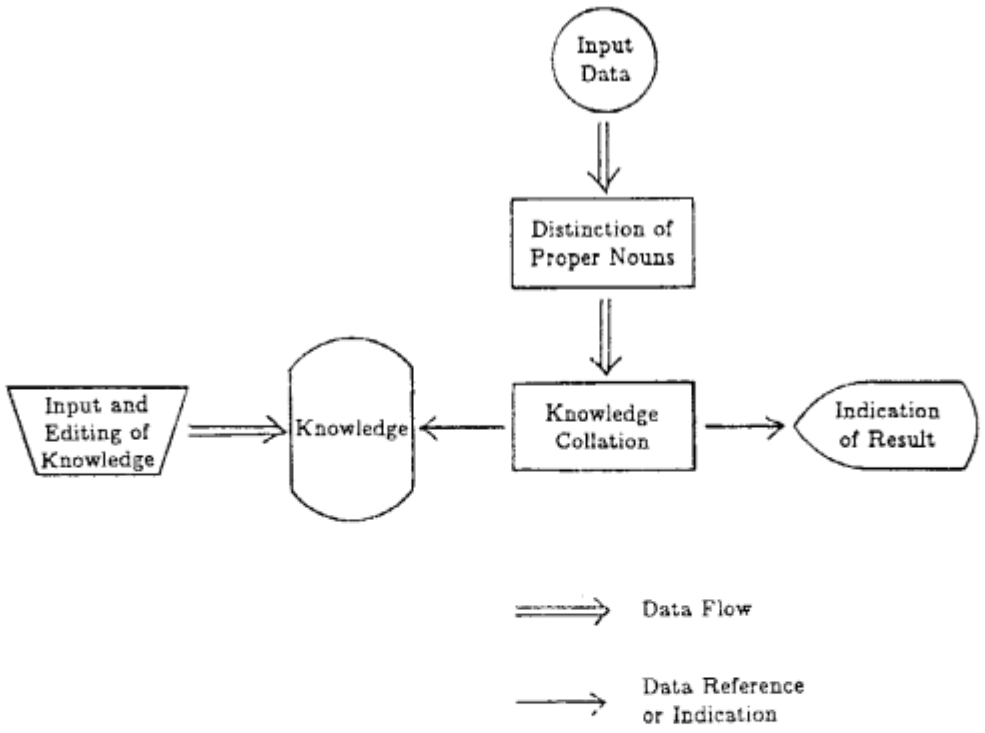

Fig. 8 Proofreading to Check Proper Nouns

Proofreading to check proper nouns is performed as shown in Fig. 8. Format rules can often be used to identify proper nouns as they are commonly found in articles having certain formats[2] and they are also used in this experiment. Basically, however, processing to distinguish between individual words is similar to that in checking word usage (described earlier). Consequently, details are not explained here.

The following knowledge was used in the experiment.

3.1.1 Knowledge on listed companies

This is knowledge on the names of all companies listed on the stock market. The predicate shown below was created using the database of the Toyo Keizai Shinposha. The contents of the database are really the same as that of the Yakuin Shikihou (Japan Company Director Handbook).[7]

```
kaisya(1301 , [ ゛極 ゛ , ゛洋 ゛ ] ) .
```

The security code and name of each company are used as arguments. The predicate was made not only for official names of companies but also for abbreviated names. The database used in the experiment includes about 100 unlisted companies and other organizations (economic organizations, government institutions, etc.) besides listed companies. In total, it contains knowledge on about 2200 organizations.

The number of lines is about 4400. KANJI characters are in the JIS KANJI codes. However, two one-byte ASCII characters are in effect used in the system because DEC 10 Prolog is not provided with KANJI codes. The situation is same for the knowledge described in the following **3.1.2**.

3.1.2 Directors of listed companies

This is knowledge on all directors of all listed companies (including the directors of the other organizations mentioned in **3.1.1**). The following predicate was created based on the database of Toyo Keizai Shinposha.

```
yakuin(1301 , ゛D ゛ , ゛Y ゛ , [ ゛志 ゛ , ゛水 ゛ ] , [ ゛廣 ゛ , ゛典 ゛ ] ,
       [ ゛し ゛ , ゛み ゛ , ゛ず ゛ ] , [ ゛ひ ゛ , ゛ろ ゛ , ゛の ゛ , ゛り ゛ ] ) .
```

The security code, post code, representation right (whether the person holds the right or not), last name, first name, reading of the last name, and reading of the first name are used as arguments.

The number of lines is about 32,000.

3.1.3 Knowledge on KANJI

This is knowledge on Common KANJI and their readings. Two kinds of knowledge were prepared: one based on the Common KANJI Table, and the other based on the KANJI table in the stylebook used at the Asahi Simbum Publishing Co. Details have already been described[3], so they are omitted here.

The number of lines is about 4100.

3.2 Results of proofreading

The results of system execution are described in the following. Owing to restrictions of the processing system, only data items listed below were used in the system evaluation.

- Knowledge on company names: Knowledge on a total 139 companies with security codes 1950 through 3011 and 3863 was used. The number of lines is about 280.
- Knowledge on names of directors: Knowledge on the directors of the 139 companies with the security codes mentioned above was used. The number of lines is 3100. Readings of the first and last names were omitted.
- Knowledge on KANJI: Readings of KANJI were omitted and only one line was assigned to each KANJI character. The number of lines is about 1940.
- Program: A short program to analyze only articles in the specific format used for personnel changes was used.

Several examples are used to describe the results of proofreading. The example sentences were extracted from newspaper articles mentioned in 2. However, as noted before, they are not exactly the same as actual sentences in the paper.

3.2.1 Checking articles

Fig. 9 shows an example of data checking. At first, "candidates" of proper nouns such as company names are distinguished in the input sentence. The word candidate is used to imply that distinction is performed mechanically by simply checking characters. Then candidate words are compared to stored knowledge. In the example of Fig. 9, no problem was found with the article in the proofreading. "post OK" at the end means that knowledge indicates that this person holds the old post (director).

3.2.2 Update lag

The example of Fig. 10 is the same as that of Fig. 9 except that the new post this

person holds (executive director) is already included in the knowledge. This is because the database was not originally designed for this system. Data on this person had already been updated before proofreading was performed.

3.2.3 Handling of KANJI other than Common KANJI

The last example is shown in Fig 11. As shown in the figure, "十条製紙", the name of the company cannot be found when the system is executed normally. This is because the name of the company is included in the knowledge only as "十條製紙" and "十条紙", which cannot be found by the normal searching procedure. Here, the KANJI "條" is the old, formal form of the KANJI "条".

```
INPUT SENTENCE      ◇片倉工業（６月１８日）退任（取締役）守田健二

COMPANY CANDIDATE     片倉工業
DATE CANDIDATE      ６月１８日
NEW POST CANDIDATE      退任
OLD POST CANDIDATE      取締役
NAME CANDIDATE      守田健二

COMPANY NAME EXISTS      CODE NO.   3001
NEW POST EXISTS
OLD POST EXISTS
NAME EXISTS
POST OK
```

Fig. 9 Proofreading for Proper Nouns (No. 1)

```
INPUT SENTENCE      ◇三和大栄電気興業（６月２９日）専務（取締役）鳥山好三

COMPANY CANDIDATE     三和大栄電気興業
DATE CANDIDATE      ６月２９日
NEW POST CANDIDATE      専務
OLD POST CANDIDATE      取締役
NAME CANDIDATE      鳥山好三

COMPANY NAME EXISTS      CODE NO.   1958
NEW POST EXISTS
OLD POST EXISTS
NAME EXISTS
POST   SAME AS NEW POST
```

Fig. 10 Proofreading for Proper Nouns (No. 2)

However, when KANJI other than Common KANJI are handled by a special procedure, the name of the company can b retrieved as shown in Fig. 12. Thus, it is indicated that a different KANJI character is used in the material. This operation is valid for symbols other than KANJI such as GETA symbols. The input operator uses a terminal with the ability to input only limited types of KANJI characters and inputs GETA symbols for the other type of KANJI in the manuscript.

```
INPUT SENTENCE      ◇十条製紙（６月３０日）専務（常務）遠藤健一郎

COMPANY CANDIDATE      十条製紙
DATE CANDIDATE      ６月３０日
NEW POST CANDIDATE      専務
OLD POST CANDIDATE      常務
NAME CANDIDATE      遠藤健一郎

COMPANY NAME CANNOT BE FOUND      THE FOLLOWING DATA IS OMITTED
```

Fig. 11 Proofreading for Proper Nouns (No. 3)

```
INPUT SENTENCE      ◇十条製紙（６月３０日）専務（常務）遠藤健一郎

COMPANY CANDIDATE      十条製紙
DATE CANDIDATE      ６月３０日
NEW POST CANDIDATE      専務
OLD POST CANDIDATE      常務
NAME CANDIDATE      遠藤健一郎

COMPANY NAME EXISTS      ON THE MATERIAL      十條製紙      CODE NO.      3963
NEW POST EXISTS
OLD POST EXISTS
NAME EXISTS
POST OK
```

Fig. 12 Proofreading for Proper Nouns (No. 4)

4. EVALUATION OF PROLOG

As mentioned earlier, the Proofreader System is implemented in DEC-10 Prolog on the DEC2060. Prolog was evaluated both quantitatively and qualitatively. In the following, the processing time was recorded when the system was used by one user.

4.1 Quantitative evaluation

The size and processing speed of programs to be handled can be considered in the quantitative evaluation. Of course, these factors are used more appropriately to evaluate the processing system rather than language specifications. However, although these factors may not be appropriate for purely logical research of languages, they were used in the following evaluation as the language and its processing system are closely related to each other.

The size and processing speed of programs which this system can handle have been mentioned in sections concerning the system configuration and functional evaluation. They are summarized in the following. As most of the program is data, the program size in indicated by the number of lines.

(1) Word distinction

In this evaluation, 12% of the entire volume of the word dictionary was used. Readings of words (KANJI combination) were omitted from the contents of the dictionary. Only short sentences such as those of the examples could be analyzed by the system due to insufficient storage capacity. The processing speed is 10 to 15 seconds for example sentences shown in Figs. 4 to 6. Those made by simply coding the longest-match algorithm appear to require more time.

(2) Source and change of wording

Only 18% of knowledge on sources and 7% of knowledge on word change was used. The processing speed is so fast that we have no problems with it such as those in item (1).

(3) Names of listed companies and directors

Only 6% of knowledge on company names was used. For director names, 10% of knowledge was used and their readings were omitted. Also, knowledge on the readings of words was eliminated from the KANJI knowledge store, which reduced the knowledge to 47% of the entire volume. The processing speed is about 20 seconds when knowledge on KANJI is used as shown in Fig. 12. In other cases, the speed is high enough to cause no particular problems. Most (80 to 90%) of the difference in speed seems to be accounted for by the time required for the retrieval of the Common KANJI Table and matching of incompletely matched lists.

Consequently, when processing is to be performed in several steps as in this experiment, it is necessary to develop methods to handle 10 to 20 times as much data and implement easy communications among processing steps. If processing cannot be performed in steps, at least 100 times as much data must be handled. The processing speed, on the other hand, has to be about 1000 times faster in practical system operation when knowledge is increased ten- to twentyfold in the future. Of course, the above estimate takes the effect of compilation and improvement in coding into consideration. Moreover, modification of the language specification and linkage to procedural languages were also considered.

The processing system used in this experiment is expected to handle a storage capacity of 800K bytes at a processing speed of about 2K lips.

4.2 Qualitative evaluation

Ease of writing, reading, debugging, and modification may be considered to evaluate a language qualitatively. However, methods for evaluating such attributes objectively have not been established yet. Consequently, to evaluate Prolog, a program with functions at a similar level was created using a conventional procedural language. Then, the total number of steps, numbers of modules, and number of steps per module were compared between Prolog and the procedural language to make an objective evaluation.

(1) Comparison with the procedural language

Two simple programs to perform word distinction, a basic operation of proofreading, were created using Prolog and Fortran (at the level of ANSI Fortran 66) and compared. These are very short programs. The program outline is as follows:

- One noun, verb, auxiliary verb, and postpositional word
- Knowledge on the inflection of the four words listed above
- No handling of undefined words

(i) Number of steps

Prolog	643 steps
Fortran	834 steps (including the COMMON statement) 1097 steps (INCLUDE statement unused)

The ratio of Prolog to Fortran is 0.77 to 1 or 0.59 to 1.

(ii) Number of modules

Prolog	80 (types of predicate)
Fortran	17 (subroutines)

The average number of steps

Prolog	8
Fortran	50 (including the COMMON statement) 65 (INCLUDE statement unused)

(2) Miscellaneous findings

(i) Advantages of Prolog

- The module is easy to understand. This is because the number of steps within the module is small and the processing flow is continuous without labels.
- It is easy to handle character strings with random length and data (lists) with random numbers of items.
- There is no COMMON statement, which reduces side effects.
- Long parameter names are possible.

(ii) Possible improvements to Prolog

- There are differences with normal human thinking. For example, recursion is used instead of an iteration. In cases when a procedural language version is easier to understand, it would be a good idea to link Prolog with a procedural language.
- Errors cause unexpected operation. For instance, when a parameter name is spelled incorrectly due to a typing error, the system itself could be used to check this and thereby become even more useful to the human user.

5. CONCLUSION

Knowledge used in the proofreading texts to check word usage and proper nouns and the results have been described. Also, Prolog in which the system is described has been evaluated.

To check word usage, knowledge on inflection, postpositional words, and auxiliary verbs derived from dictionaries was used. Also, a word dictionary, knowledge on source, and knowledge on change of wording, derived principally from the stylebook used at the Asahi Shimbun Publishing Co., as well as the Common KANJI Table were used. Due to limitations of the processing system, about 10,000 lines of knowledge, that is, about one tenth of the entire volume of knowledge was used in operation for system evaluation. As a result, the Proofreader System pointed out undefined words, incorrect words, words to be changed to other words, mixed use of HIRAGANA and KATAKANA, and words written in KANA.

Proofreading of articles on personnel changes of directors working for companies listed on the stock market was used as an example of proofreading to check proper nouns. Knowledge was created from the database of the Toyo Keizai Shinposha. One tenth of the knowledge was used in the experiment to collate articles with the database. As proper nouns rather often include KANJI other than Common KANJI, knowledge on Common KANJI derived from the stylebook used at the Asahi Shimbun Publishing Co. was combined in the system. An example of proofreading using this knowledge was also described.

When Prolog was evaluated, it was noted that 10 to 20 times as much storage capability and smoother communication between processing steps will be necessary when processing is performed in steps as in this experiment. Also the processing speed will need to be about 1000 times as fast as that of the current interpretive execution. Finally, Prolog was compared to Fortran and programs written in Prolog were found to be easy to understand. Some other findings of the comparison were also described.

In this study, as noted previously, data on newspaper articles from the Asahi Shimbun Publishing Co. and data on directors of listed companies from the Toyo Keizai Shinposha were used. In addition, the Common KANJI Table and a part of the stylebook used at the Asahi Shimbun Publishing Co. were input into the computer. I would like to express my sincere gratitude to the Asahi Shimbun Publishing Co. and the Toyo Keizai Shinposha who kindly offered their data and approved of the use of their materials. Also, I would like to express my thanks to the Sharp Corp. for their help in inputting and editing the knowledge.

REFERENCES

[１] Ishii,S.：“新聞における校正・校閲の実データによる調査”、ＴＲ--０３９、Institute for New Generation Computer Technology(1983)
The abstract appeared as:
Ishii,S.：“Study of Proofreading Techniques Used at a Japanese Newspaper”、Proceedings of the 28th Annual Convention IPS Japan 、（Ⅱ）2M-7、pp.1205 ～1206、Information Processing Society of Japan （1984）

[２] Ishii,S.：“新聞における人名の使い方の調査”、ＴＭ－００６２、Institute for New Generation Computer Technology（1984）
The abstract appeared as:
Ishii,S.：“Study of a Newspaper's Treatment of Proper Names”、Proceedings of the 29th Annual Convention IPS Japan 、（Ⅱ）4J-5、pp.1449 ～1450、Information Processing Society of Japan(1984)

[３] Ishii,S.：“日本語の漢字・用語の校正のための知識”、ＴＭ－００９２、Institute for New Generation Computer Technology(1985)
The abstract appeared as:
Ishii,S.：“Knowledge for Proofreading Japanese Word Usage”、Proceedings of the 30th Annual Convention IPS Japan 、（Ⅱ）3G-3、pp.1635 ～1636、Information Processing Society of Japan(1985)

[４] 久松他：“角川国語辞典”６３版、角川書店（1982）

[５] 朝日新聞社用語幹事編：“朝日新聞の用語の手びき”第１８版、朝日新聞社（1984）

[６] 大蔵省印刷局編：“常用漢字表”、大蔵省印刷局（1982）

[７] 東洋経済新報社編：“役員四季報１９８５年版”、東洋経済新報社（1984）

All references are in Japanese and published in Japan. From [1] to [3], an abstract appeared in English for each, as mentioned. From [4] to [7], titles etc. in English are as follows. The translation is tentative and not authorized.

[4] Hisamatsu *et al.*: "Kadokawa Japanese Dictionary" 63rd Edition, Kadokawa Publishing Co. (1982)

[5] Secretary for Word Usage, Asahi Shimbun Publishing Co. (ed.): "Guide to Word Usage for the Asahi Shimbun" 18th Edition, Asahi Shimbun Publishing Co. (1984)

[6] Printing Bureau, Ministry of Finance (ed.): "Table of Chinese Characters in Common Use", Printing Bureau, Ministry of Finance (1982)

[7] Toyo Keizai Shinposha (ed.): "Japanese Company Directors Handbook 1985 Edition", Toyo Keizai Shinposha (1984)