JAPANESE TEXT PRESENTATION SYSTEM FOR PERSONS WITH READING DIFFICULTY
Design and Implementation

Kyota Aoki and Shinjiro Murayama

Graduate School of Engineering, Utsunomiya University, 7-1-2 yoto, 321-8585 Utsunomiya, Japan

Keywords: Reading Difficulty, Text Presentation, Highlighting/Masking, Dyslexia.

Abstract: There are many pupils with reading difficulty in Japanese schools. The dyslexia is the disability about reading and writing texts. We use Kanji, Hiragana, Katakana characters in Japanese sentences. We propose the Japanese text presentation system that eases the difficulties about reading Japanese texts with or without dyslexia. The kanji is an ideograph. The hiragana and the katakana are phonograms. The reading difficulties include 2 types. One is a difficulty about reading the kanji. Another is the difficulty about tracing the reading sequence. This paper proposes a system that presents the Japanese sentences with suitable presentation method for each pupil with reading difficulties. The main function of the proposed system is 3 levels of highlighting/masking that are independently controlled. The highlighting only is not enough to prevent the error about the reading sequence of character chunks. The 3 level highlighting/masking enables to adapt the presentation to wide varieties of reading difficulties. The proposed system also proposes the function to measure the reading ability of the pupil with recording their operations. This paper proposes the plan and the implementation of the Japanese text presentation system.

1 INTRODUCTION

There are many pupils with reading difficulty in Japanese elementary schools. There are many kinds of difficulties. The big and first one is reading Japanese characters. Japanese characters are the construction of hiragana (phonetic character), katakana (another type of a phonetic character), kanji (Semantic character) and other characters.

Figure 1 shows the examples of Japanese characters. The characters in the left group are hiragana. The characters in the center group are katakana. In those groups, the characters in the corresponding positions have same pronunciation. The characters in the right group are kanji. They have multiple pronunciations. However, one pronunciation is same as the other 2 groups. The first character in the right group means 'love'. The second character means 'up'. The third one does 'tail'.

The other characters include Arabic numbers, English alphabet and, etc. The Japanese sentence can include the words of other languages. The number of characters used in Japanese sentences is about 1000 in an elementary school. This hugeness is the first barrier.

あいうえお	アイウエオ	愛上尾
かきくけこ	カキクケコ	家木区家子
さしすせそ	サシスセソ	差市酢背素

Figure 1: Examples of Japanese characters.

The second barrier is the understanding of the sentence structure. There is no space between words in Japanese sentences as shown in the top of figure 2. In an elementary school, pupils learn hiragana at first, and then they do katakana. In the first stage in elementary schools, the Japanese text-book has a space between words for the ease of understanding the structures of the sentences. However, normal Japanese texts have no space between words.

The top of figure 2 means 'Cherry blossoms have opened'. The first character is a kanji character that means cherry. In Japanese, cherry also means cherry blossom. The third character is a kanji character that means 'open up'. The middle of figure 2 is hiragana writing of the sentence on the top. We can easily understand the structure of the sentence on the top of figure 2 with the help of character type changing and

123

the meaning of kanji characters. However, we have difficulties for understanding the structure of the middle sentence. The bottom in figure 2 is another writing that has same pronunciation as the middle. The meaning of the sentence differs from the top of figure 2.

桜が咲いた

さくらがさいた

柵等が裂いた

Figure 2: Writing examples.

Every pupil has those 2 difficulties at first. In the long school life, they studied the skill to conquer those difficulties. Anyway, those 2 difficulties are large barriers for reading and understanding Japanese sentences.

Every infant has no knowledge about the Japanese characters. Every pupil has no knowledge about the huge number of kanji letters at first. Then, they learn hiragana, katakana and kanji characters in a long elementary-school life. The pupils with a learning disability tendency have difficulty about reading Japanese sentences. Off cause, some pupils have difficulty about remembering kanji characters. Most of the pupils remember kanji characters gradually. However, pupils with a learning disability tendency have difficulty with reading Japanese sentences in the case that they can remember the kanji characters. In the case, they may have dyslexia.

There may be many causes of the difficulties on reading Japanese texts. We do not discuss the causes. We only pay attentions to the methods for easing their difficulties. We call their difficulties "reading difficulty" in this paper.

The research about teachers shows that the pupils with ADSH tendency have difficulty about following the characters sequentially and recognizing the grammatical structures (Murayama, Aoki and Morioka, 2009). In the case, Murayama and Aoki (2010) showed that the restriction of the presented text helps to understand the text. Off cause, there are many types of reading difficulties. There are many causes about the reading difficulties. The resulting reading difficulties show the same symptom that is the difficulty about following the characters sequentially, recognizing grammatical structures and reading kanji characters. This paper proposes the presentation system that eases this reading difficulty with presenting the Japanese texts with a proper method for each pupil. The differences of pupils are ages and disabilities.

For normal non-Japanese students learning

Japanese, there are same problems as the Japanese pupils with reading difficulties. They also have difficulties for understanding grammatical structures and reading kanji characters. The proposed system helps the students for reading Japanese texts.

This paper describes the design and the implementation of a visual text presentation system for persons with reading difficulty in windows environments. First, this paper evaluates the DAISY (Digital Accessible Information SYstem) in primary schools that has the pupils with reading difficulty (DAISY consortium, 2012). Then, we discuss about the functions needed in the Japanese text presentation system. Next, we show the structure of the proposed Japanese sentence presentation system and its implementation. Then, this paper discusses the function of the implementation. And last, we conclude our work.

2 EVALUATIONS OF DAISY

The DAISY is one implementation of digital talking book. Figure 3 shows an example of visual presentation with DAISY. It includes highlighting function of talking chunk of characters. It reads out the DAISY contents. It has visual text presenting functions, also (DeMeglio, Hakkinen and Kawamura, 2002) (DAISY consortium, 2012). The DAISY started for helping the people with sight disability. The DAISY is helpful for pupils with reading difficulty. However, the DAISY is a digital talking book. The content must be carefully prepared. Furthermore, there is a difficulty to make DAISY contents from electrical texts. The advanced teachers prepare the DASY contents and use the DAISY in their courses. However, the preparations need a large amount of works. It is difficult to use emerging materials in their courses for the large preparation works. For example, it is difficult to use the morning news in the course on the same day. The interesting fresh materials can attract pupils' attentions more. The DAISY does not offer such functions.

The DAISY is a talking book. We can change the speed of talking. However, the visual presentation follows the talking. The user cannot control the visual presentation directly. The user cannot change the size of a chunk of uninterrupted talk. In the first stage of elementary schools, most of the pupils like the short chunk. With the progress of their abilities, they like longer chunks. We can control the speed of talking and start and stop of talking. We cannot change the visual chunk of characters. In figure 3,

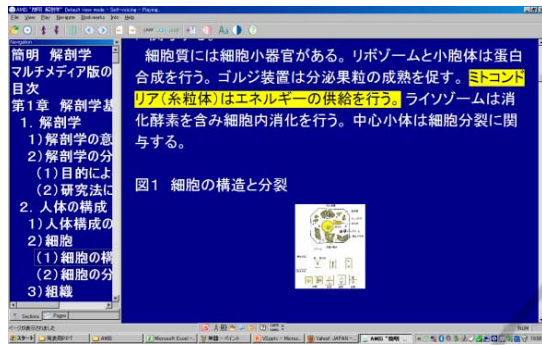the highlighted chunk is a whole sentence. There is no smaller chunk in the DAISY.

Figure 3: DAISY visual presentation example.

Another problem about DAISY is the difficulties of the adjustment for each disability of each pupil. Each pupil differs on their ages and abilities. For each pupil, there is the best fit visual presentation. However, the control about visual presentation is very limited in DAISY. The DAISY contents can have a variety of visual presentations. Each visual presentation needs individual DAISY contents. Each pupil differs on his ability with his progress. The DAISY cannot withstand the environment.

The proposed Japanese text visual presentation system overcomes such problems with a very small preparation work. The preparation costs almost nothing. The proposed presentation system offers the functions that are adapting the material for pupil's age and adapting it for pupil's difficulties. For adapting the material for pupil's age, our Japanese texts need to be constructed with the proper kanji characters fitting for the age. Adapting the proposed Japanese text presentation system for pupil's difficulties, the amount of information presented at a time is restricted properly. For the persons with a dyslexia tendency, the restriction of presented information is effective to improve the understandability. Our proposed presentation system provides the function for restricting presented information with 3 levels of masking and highlighting. The only highlighting is not enough for the person with reading difficulty.

3 PRESENTATION SYSTEM DESIGN

Our goal is to propose the method that helps pupils that have difficulties to read Japanese texts. There are many needs in Japanese elementary schools. However, the proposed system concentrates on the

difficulties about reading Japanese texts.

The presentation system has functions for selecting font types, font size, character spacing and line spacing as NotePad has. For helping the pupil with ADSH tendency to read, the proposed system has 3 levels highlitings/maskings. For some types of reading difficulties, it is effective to decrease the amount of the presented information at one time. The DAISY has only the highlighting function. Figure 4 shows the examples of combinations of highlighting and masking. The B in figure 4 is a normal presentation. The E is an example of DAISY type highlighting. The F is a classical masking. With the masking method, there is no information about the masked area. The G is an example of weak masking. With the masking, we can read the part masked.
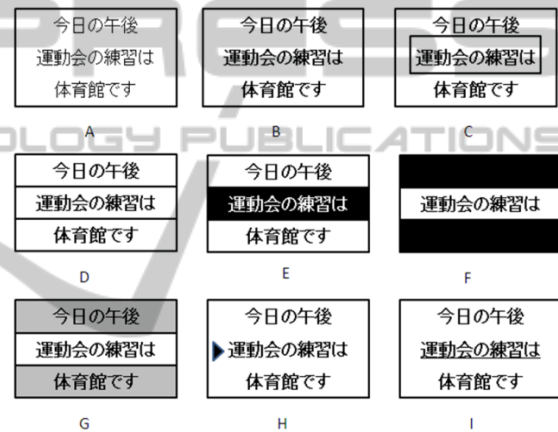

Figure 4: Highlighting and masking examples.

Figure 5 is an example of 1 level highlighting of DAISY type. The figure 6 is an example of 3 levels presentation with the combination of a highlighting and a masking. In figure 6, we easily find the sentence, comparing with Figure 5. The un-attention part is less attractive than the attention part. This type of visual presentation may help to the pupil with ADSH tendency.
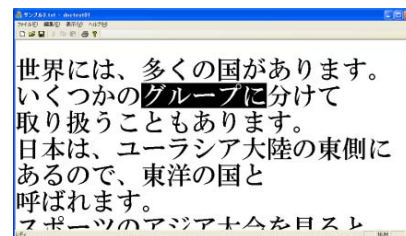

Figure 5: 1 level highlighting.

There are many problems for utilizing the ITC technology in Japanese elementary schools

(Murayama et al., 2009). The problems are listed in Table 1. For solving the problems, the proposed text presentation system treats only the electronic text. In Japan, a law forces to prepare the electronic readable text of text books (Law 2009). Furthermore, there are many documents accessible through the internet. There is no paper document for an input in the proposed system.
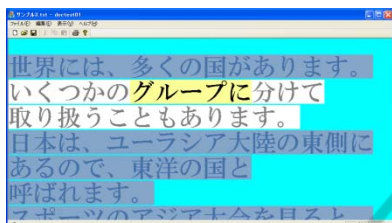

Figure 6: 1 level highlighting and 1 level masking with background coloring.

Table 1: Problems about ICT usability in a special aid school in Japan.

There is a load concentration into the teacher, who is good at ICT.
There are a few educational materials for the DAISY.
They do not use the SAVE AS DAISY.
Using OCR for preparing educational materials for pupils with a learning disability tendency, the recognition errors make a large check and correct work.
There are large works for replacing difficult kanji characters with hiragana.
Using classical ICT tools as the DAISY, we need to prepare educational materials for each pupil who has a different age and a different disability.
It is difficult to evaluate the performance.

The proposed Japanese text presentation system has only 2 functions. As the first row of Table 2, we restrict the functions of the proposed system. The proposed system must have some performance measurements function about pupil. However, performance measurement about pupil is another big theme. The teachers around the pupils with reading difficulties need the objective measurements about the performance about the pupils. For this purpose, the proposed system provides the operation logging function. The operation logs describe the reading speed at each meaningful chunk of characters.

For decreasing the preparation work about education materials, the proposed system provides the online presentation generation from a plain text. This is the main and most important function. With this online presentation generation, there is no need to prepare the educational materials beforehand. This enables to use the very new materials in a class.

This enables to use one-time materials for measuring the performance of a pupil. Many pupils may remember the full text of the many times used materials as text books. Those remembered materials cannot be used for evaluating the reading performance of a pupil. The online presentation generation enables any new plain text materials at any time.

Table 2: The plan for covering the problems.

A collection of simple software is better than complex multi-functional software.
Avoid the usage of OCR.
An educational material presentation system does not need the special material preparations.
We need an evaluation method/function for evaluating the performance of a pupil.

This online presentation generation enables to adapt the presentation for each pupil with different reading difficulties. DAISY has no function about adaptation for each pupil.

For adapting the variety of pupils' ages and difficulties, the presentation system has the function to replace the un-studied kanji characters with hiragana characters. The phonic hiragana character is first studied character. There is a little difficulty about reading hiragana.

The operations on the presentation system have the information about a subject. The proposed system logs every operation with the time. This log represents the fluency of the reader.

The presentation system has the highlighting/masking methods. Many presentation systems have highlighting methods for distinguish the point where a reader must look at and pay attention. Some persons with reading difficulty have difficulty following the characters in proper sequence. The highlighting method helps to find the place to look, but it does not help to prevent to look at un-proper place.

The proposed system has rich unified highlighting/masking methods. This enables to help to look at a proper place and not to look at un-proper places.

For controlling the amount of information at each chunk, the proposed system has the function controlling the size of the chunk of characters in semantic senses. Japanese text has no word spacing between semantic chunks of characters. For automatically making proper highlighting/masking, there must be understanding of the semantic structure of the texts. The proposed system uses MeCab for understanding the semantic structure of the texts. MeCab is the free Japanese grammatical parser (Kudo, 2012). With the help of MeCab, the

proposed system enables to highlighting/masking the text with the proper size of chunks.

The proposed system has the features listed in Table 3. The network problem is important Japanese schools. There is a large limitation about the internet access. As a result, some cloud based implementation cannot work. The proposed system must work without the internet access.

Table 3: The additional plan for adapting the school environments.

| The system must work without network connections. |
| The function replacing unlearned kanji characters with hiragana characters. |
| A semantic collection of characters must keep their connectivity in a line. |
| Line breaks must be between the semantic collections. |

4 PRESENTATION SYSTEM IMPLEMENTATION AND ITS WORKING EXAMPLES

The implementation has 2 inputs and 3 outputs. The 2 inputs are a plain text file and user operations for controlling the reading places. The 3 outputs are the text presentation on a visual display, logs about the operations and the reading out sound. We implement the proposed Japanese text presentation system on windows environments.

For using the presentation system, the user must assign the text file. This is the only normal preparation to start reading text. Then, the system displays the first part of the text with the pre-set font-size, the size of a chunk of character and, etc. If the sound output is enabled, the system talks the part.

If there are un-studied kanji characters, the continuous sequence of kanji characters including the un-studied kanji characters are replaced hiragana.

The top in figure 7 is kanji character sequence represents the name of city. If the second character is an un-studied kanji, the system replaces the continuous sequence of kanji character with the sequence of hiragana as the bottom in figure 7. If we replace the second character only, we have the middle in figure 7. It is difficult to understand. Because, the continuous sequence of kanji character represent the meaningful chunk. The middle in figure 7 breaks this rule.

On a visual display, the first meaningful chunk of character is highlighted at first. The sentence including the chunk is also highlighted with another method. The other sentences on a display are masked as a pre-decided method. Figure 6 is an example of our standard highlighting/masking setup. For reading the text, the user only uses the arrow keys. With typing keys, the highlighted chunk moves the next chunk. With the move of the highlighted chunk, the system keeps the display properly. The system records the key typing with the time. Afterward, the record shows the reading performance of the user.

大阪
大さか
おおさか

Figure 7: Kanji replacement with hiragana.

The proposed system has a general highlighting/masking function. At each level, we can select a free combination of font color, background color and the non-character region color. The best-fit presentation differs for each pupil. The user can search the best-fit presentation easily. The 3 level highlighting/masking help the user to find the place to read and understand the structure of the sentence.

Figure 8 shows the layout change with the presentation character size changes. The reading chunk stays at the same place. There is no line break in a semantic chunk of characters.
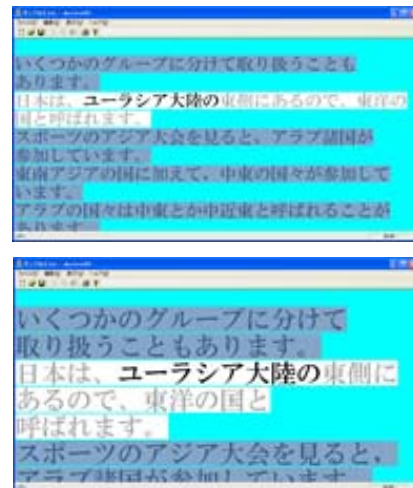

Figure 8: Layout change with the change of the character size.

Figure 9 shows the replacing un-studied kanji characters with hiragana characters. With replacing kanji characters with hiragana characters, the number of characters increases.

Figure 10 shows the scrolling example with the move of reading chunk of characters. The reading chunk moves from the last chunk in a sentence to the first chunk in the next sentence. There is the previous sentence on the top. On a display, there is 1 line of the previous sentence, the reading sentence and the following sentences.
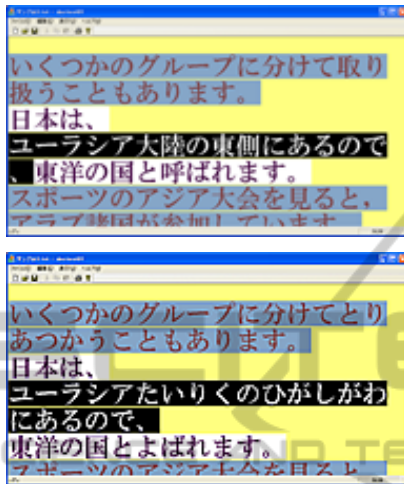

Figure 9: Replacing un-studied Kanji characters with Hiragana characters.
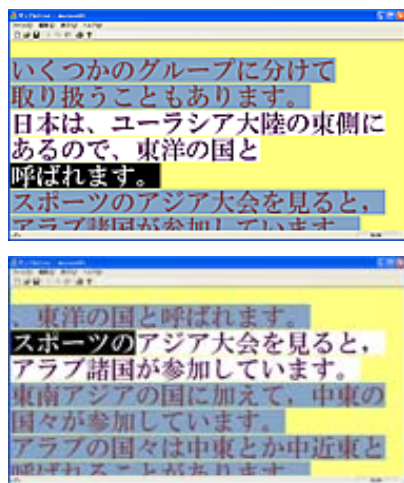

Figure 10: Scrolling example.

We have the experiments in a Japanese elementary school. The subjects are a 10 years-old girl and 2 boys. They attend a special aid class. They have some reading difficulties. One boy has used DAISY in a long term. All of them like the 3 level highlighting/masking. They have no difficulties about using the text presentation system. Teachers have some difficulties about using new ITC methods. The proposed text presentation system enables to use a same article in a wide range of

pupils. This helps to ease the emotional estrangement of the pupils with reading difficulties.

5 CONCLUSIONS

This paper proposes the Japanese text presentation system for pupils with reading difficulties. The proposed Japanese text presentation system makes reading Japanese texts to be easy. We need little preparation works beforehand. The proposed Japanese text presentation system implements all needs of teachers in Japanese elementary schools. The small number of experimental uses shows that the presentation system is a helpful tool in elementary-school environments for decreasing the reading difficulties of a pupil and estimating the performance of the reading ability of the pupil.

The next step of this research is a wide distribution of the software. However, this needs not only the software distribution, but also the distribution of the method to use the software effectively.

REFERENCES

DAISY consortium. (2012). *DAISY Multimedia*. Retrieved from: http://www.daisy.org/mutimedia

DeMeglio, M., Hakkinen, M., Kawamura, H. (2002). *Accessible Interface Design: Adaptive Multimedia Information System (AMIS)*. COMPUTERS HELPING PEOPLE WITH SPECIAL NEEDS, Lecture Notes in Computer Science. Springer, 2398, 319-330.

Kudo, T. (2012). *MeCab: Yet Another Part-of-Speech and Morphological Analyzer*. Retrieved from: http://mecab.sourceforge.net/

Law. (2009). *Overview of the Act on Promotion of Distribution of Specified Books, etc. Used as Textbooks for Disabled Children and Students*. Retrieved from: http://www.bunka.go.jp/chosakuken/pdf/tokuteitosyo_fukyu_gaiyo.pdf

Murayama, S., Aoki, K. (2010). *Real Time Image Presentation System for Persons with a Learning-Disabled Tendency*. IEICE-ET, 109(387), ET2009-96. 25-29.

Murayama, S., Aoki, K., Morioka, N. (2009). *Image processing to make teaching aids for learning disability persons*. IEICE-108, IEICE-WIT-2009-79, IEICE, 61-66.