INTRODUCING A READABILITY EVALUATION SYSTEM FOR JAPANESE LANGUAGE EDUCATION

Yoichiro Hasebe (長谷部陽一郎), Doshisha University
Jae-Ho Lee (李在鎬), University of Tsukuba

Abstract: This study introduces a readability evaluation system that was developed to support educators and learners of Japanese (available at http://jreadability.net). The system analyzes input text and estimates its readability, using a formula based on a regression analysis of data collected from 100 language textbooks and the balanced corpus of contemporary written Japanese (BCCWJ). In addition to scoring text in six-level categories, the system has rich functionalities that are implemented to support teachers and learners in carrying out various reading-related activities efficiently and effectively. Furthermore, feedback collected on an earlier version of the system is discussed, which confirms the usefulness of our method of evaluating text readability, while suggesting the necessity for further improvements.

Keyword: readability, text analysis, web application, Japanese education, BCCWJ

1. INTRODUCTION

This brief article introduces an online text readability evaluation system for Japanese educators and learners. In the following sections, we first illustrate how we obtained our readability evaluation formula. Second, we discuss the functionalities of our system that make it a useful tool for both educators and learners of the language. Finally, we examine feedback from users of an earlier version of the system that was requested in order to obtain suggestions for further improvement and development of the system.

2. READABILITY EVALUATION

The evaluation formula was developed through analyzing text from 100 Japanese language textbooks that were chosen carefully so that the levels of the titles were varied and balanced. In addition, text data from the "books-in-general" segment of the BCCWJ was added to the data as models of more advanced types of text. The resulting dataset was then processed in the following manner.

First, the overall dataset was separated into 958 plain text files, each having 1,000 Japanese characters. Then, human evaluators individually studied the files from the textbook data and classified them into five readability levels ("lower elementary," "upper elementary," "lower intermediate," "upper intermediate," and "lower advanced"), and the files from the books-in-general segment were categorized as a group titled "upper advanced."

Next, a morphological analysis of these six groups of text was carried out utilizing MeCab and UniDic to break down text strings into individual words and obtain textual information such as the mean number of words in a sentence and grammatical/categorical attributes of words. Among such word attributes were parts of speech, vocabulary levels, and origin types (i.e., *wago*, words of Japanese origin; *kango*, words of Chinese origin; and *gairaigo*, words of other origins that are primarily from the West).

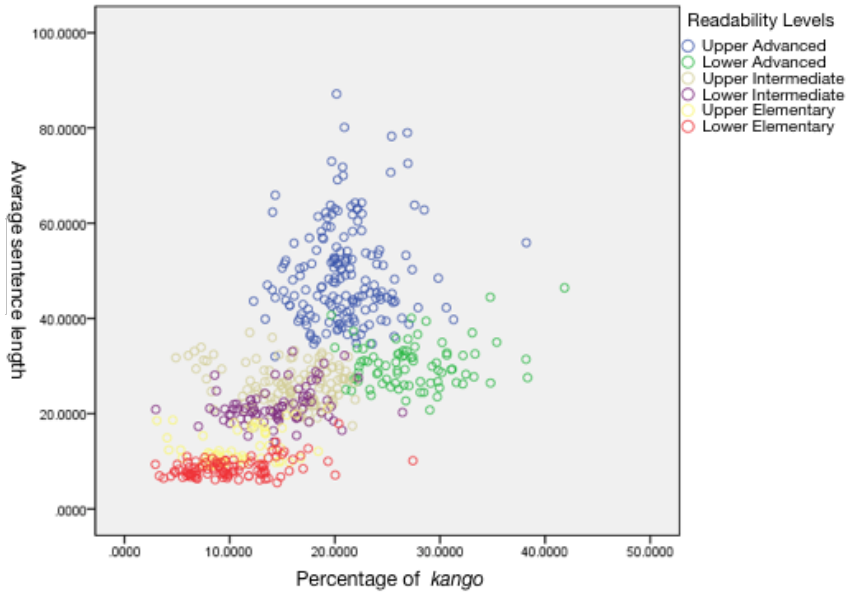

Figure 1: Text files of different word levels

A discriminant analysis was performed on the resulting dataset using SPSS so that we could classify text files that represented six different readability levels well from those that did not. In other words, through this process, we created optimized sets of files that could be regarded as fairly prototypical of the six different readability levels. As a final step, we then conducted a linear regression analysis (with the stepwise option enabled) on the optimized dataset and obtained the following regression formula (R^2 = .896):

$$\textit{Readability Score} = 11.724 - .056a - .126b - .042c - .145d - .044e$$

where a = average length of sentence, b = percentage of *kango,*
c = percentage of *wago*, d = percentage of verbs, and
e = percentage of particles

Our readability evaluation system applies this regression formula to input text and returns a score in six-level categories. In addition, it has several other functionalities that are implemented to aid both educators and learners of Japanese, as discussed in the next section.

3. A TOOL FOR EDUCATORS AND LEARNERS

The web system at jreadability.net accepts Japanese text of up to 20,000 characters at once and processes it using MeCab and UniDic, which were also used to obtain the input data for the discriminant and regression analyses that produced the formula. Besides evaluating text and returning the results in six-level categories, the system offers various functionalities that will be helpful both for teachers and learners of Japanese in their reading-related practices.

(a) Left: Main result screen

(b) Right (Top): Line-by-line representation
(c) Right (Middle): Word list
(d) Right (Bottom): Dictionary lookup

Figure 2: Screenshots of jreadability.net

First, as partially indicated in Figure 2a, the system shows the readability score on its main panel ("text info" panel). Also shown on this panel is graphically represented information about the text as well as a button in the upper-right corner that plays a computationally synthesized voice reading the text. In addition, the system builds two other panels that display the results of the analysis from different perspectives. One of the perspectives is the "text details" panel that presents the input text in a line-by-line fashion with every component word colorfully highlighted according to its vocabulary level (Figure 2b). The other one is the "vocab list" panel that shows, as the name suggests, the list of vocabulary items with a detailed description of the grammatical properties alongside the number of its occurrence in the text (Figure 2c). Last, vocabulary items shown on these two panels are given a clickable link that opens a little pop-up window presenting dictionary definitions and example sentences for the word (Figure 2d).

With this system, we hope that teachers and learners of Japanese will be able to ensure that the readability of text on hand corresponds with the targeted level as well as

be equipped with tools that further boost the effectiveness and efficiency of their teaching/learning.

4. DISCUSSION/CONCLUSION

Since its alpha release in October 2013, the system has earned approximately 10,000 page views and garnered 205 reviews from users via its online questionnaire. The responses to the questionnaire, which inquires whether users consider the returned score to be reasonable based on a four-point Likert scale (with 1 representing the most negative response), were generally positive (Table 1). However, approximately 21 percent of respondents (44 out of 205) assigned low scores (1 or 2). Hence, we examined the mean readability score for each of the 1 to 4 answer groups, obtaining the results in Table 2. The figures do not necessarily show a salient difference between the mean readability scores of the lower points and the higher ones. A close look at the input text of each group, however, reveals a few likely reasons for the negative responses. First, as shown in Table 3, the length of the text that received lower scores is relatively shorter compared to the text that received higher scores. Moreover, it was discovered that some of the textual passages in the latter group deal with highly abstract or technical topics.

This questionnaire is admittedly rather informal, and the two findings above obviously suggest a limitation of our readability evaluation system. Yet, the limitation does not necessarily apply only to our system; rather, it is considered to imply the necessity for further discussion to achieve a better definition/understanding of the concept of "readability," and this will be reflected somehow in future versions of the system.

Table 1: Evaluation of the readability score

Likert scale of 1 - 4	1	2	3	4
Responces (N=205)	13	31	102	59

Table 2: Mean readability scores

Likert scale of 1 - 4	1	2	3	4
Mean readability score	3.76	2.48	2.43	3.08

Table 3: Mean text lengths (characters)

Likert scale of 1 - 4	1	2	3	4
Mean text size	806	403	1711	1101

ACKNOWLEDGMENTS

This work was supported by JSPS Grant-in-Aid for Scientific Research 25370573.

REFERENCES

Lee, J.-H. 2011. The Utility of Corpora for Composing Reading Comprehension Questions for Large-Scale Tests. *Nihongo Kyouiku*. 148, 84–98.

Lee, J.-H. and H. Shibasaki. 2012. Relationship between Readability and Vocabulary: What Lexical Properties Determine the Characteristics of Text for Different School Grades. *Corpus and Text Mining*. Tokyo: Kyoritsu Shuppan, 181–192.