Metrical Analysis of English Textbooks in Finland

Hiromi Ban1 and Takashi Oyabu2

1Fukui University of Technology
3-6-1, Gakuen, Fukui-shi, Fukui, 930-8505, Japan, ban@fukui-ut.ac.jp
2Kanazawa Seiryo University
10-1, Ushi, Goshomachi, Kanazawa-shi, Ishikawa, 920-8620, Japan, oyabu@seiryo-u.ac.jp

Abstract: Finland, which topped in the world in a reading comprehension, mathematics, and scientific literacy in the Programme for International Student Assessment (PISA) by the Organisation for Economic Co-operation and Development (OECD), is also excellent in English language skills, and the 6th place in the world in the TOEFL iBT. The English language education starts formally from the 3rd grade in the elementary school there. In this paper, we investigated English textbooks for elementary school students in Finland, in terms of metrical linguistics. In short, frequency characteristics of character- and word-appearance were investigated. These characteristics were approximated by an exponential function. Furthermore, we calculated the percentage of Japanese junior high school required vocabulary and American basic vocabulary to obtain the difficulty-level. As a result, it was clearly shown that the English textbooks in Finland have a similar tendency to English journalism in the characteristics of character-appearance, and the difficulty level is almost the same as the English textbooks for Japanese junior high school students.

1. INTRODUCTION
Finland, which topped in the world in reading comprehension, mathematics, and scientific literacy in the Programme for International Student Assessment (PISA) by the Organisation for Economic Co-operation and Development (OECD), is also excellent in English language skills, holding 6th-place position in the world in terms of TOEFL iBT[1]. English language education in Finland starts formally from the 3rd grade in elementary school, and continues in a consistent manner to junior high school and high school[1].

In this paper, we conducted metrical linguistic analysis of English textbooks in Finland for the 3rd to 6th grade students in elementary school, so as to ascertain characteristics of writing styles used therein. We investigated types of characters and words and their use frequency in English textbooks issued by WSOY, the largest textbook publisher in Finland.

2. METHOD OF ANALYSIS & MATERIALS
The following materials were analyzed in this study.

Material 1: Wow! 3 (2002, WSOY) [for the 3rd grade]
Material 2: Wow! 4 (2003, WSOY) [for the 4th grade]
Material 3: Wow! 5 (2005, WSOY) [for the 5th grade]
Material 4: Wow! 6 (2006, WSOY) [for the 6th grade]

For comparison, we also analyzed the content of English textbooks used in Japan in junior high schools (NEW HORIZON English Course 1, 2, 3 (2010, Tokyo Shoseki Co., Ltd.) (hereinafter referred to as “JHS 1, 2, 3”) and high schools (UNICORN ENGLISH COURSE I, II, READING (2010, Bun-eido Publishing Co., Ltd.) (hereinafter referred to as “HS 1, 2, 3”).

3. RESULTS
Firstly, we investigated the types of characters frequently used and their frequency. We showed the top 50 characters most frequently used in each material in a single logarithmic plot, with the frequency on the vertical axis and the ranking on the horizontal axis. We obtained an exponential approximation for such frequency characteristics using the following formula.

\[ y = c \times \exp(-bx) \] (1)

Fig. 1 shows coefficients \( c \) and \( b \) thus obtained for respective materials. Mostly linear relations can be observed between coefficients \( c \) and \( b \) for all materials. With regard to textbooks in Finland, \( c \) is between 9.4042 and 9.9314, and \( b \) is between 0.0933 and 0.1023. In the former report, the authors introduced the results of the analysis of English sentences in various categories, indicating the fact that a positive correlation can be observed between coefficients \( c \) and \( b \) in these materials and that these values tend to be smaller for materials closer to journalistic or technical articles and larger for materials closer to literary works[2]. Therefore, it can be said that textbooks used in Finnish elementary schools are closer to journalistic or technical articles.
Just as in the case of characters, we showed the top 50 words most frequently used and obtained an exponential approximation using formula (1). Coefficient \( c \) for Material 1 is around 0.5 higher than that for the other 9 materials. The values of \( c \) and \( b \) for Materials 2 and 3 are similar to the textbooks for Japanese junior high schools and those for Material 4 are similar to the textbooks for Japanese high schools.

We investigated the difficulty level of each material based on the words used. As parameters for indicating difficulty, we considered difficulty level based on the number of types of words used \( (D_{nt}) \) and difficulty level based on the number of words used \( (D_{nw}) \). As standard vocabulary, we used the 508 words that are required in the Japanese junior high school curriculum and 798 of the words contained in the American Heritage Picture Dictionary (American Heritage Dictionary, Houghton Mifflin, 2003) (hereinafter the latter shall be referred to as “basic words”). Two types of difficulty levels are obtained as follows, with the total number of words used as \( n(i) \), the total number of types of words used as \( n_t \), the number of required/basic words as \( n_b \), and the number of required/basic words used as \( n_{rb} \).[2]

\[
D_{nw} = \left( 1 - \frac{n_b}{n} \right)
\]

(2)

\[
D_{nt} = \left\{ 1 - \left( \frac{1}{n_t} \ast \Sigma n(i) \right) \right\}
\]

(3)

Then, in order to obtain more appropriate indices, we conducted principal component analysis, considering \( D_{nw} \) and \( D_{nt} \) as variables. Fig. 2 shows obtained principal component scores.

![Difficulty](image)

Fig. 2 Principal component scores of difficulty.

Judging from required words, textbooks used in Finland are more difficult than those for Japanese junior high school students. Material 4 is the most difficult and Material 1 is the 2nd most difficult among Materials 1 to 4. With regard to basic words, for both Finnish materials and Japanese materials, difficulty level becomes higher for those for higher grades. Materials 1 to 3 are around the same level of textbooks for Japanese junior high school students, while Material 4 is more difficult than that for Japanese 3rd-grade junior high school students.

We also investigated other quantitative characteristics of each material. Table 2 shows the results thereof.

The “mean word length” for Materials 1 to 4 is between 5.119 (Material 2) and 5.285 (Material 4), slightly longer than that for Japanese materials for junior high school students (between 4.994 and 5.161), and shorter than that for Japanese materials for high school students (between 5.478 and 5.568).

The “number of words per sentence” increases for textbooks for higher grades. The number for Finnish materials is between 5.761 (Material 1) and 8.522 (Material 4), being almost the same as Japanese materials for junior high school students (between 5.335 and 8.183), and considerably smaller than those for high school students (between 12.769 and 15.778). This shows that English textbooks for Japanese high school students are rather difficult.

As for the “frequency of auxiliaries,” we ascertained that the percentage is highest, at 1.464%, for Material 1 and declines gradually for textbooks for higher grades, down to 0.690% for Material 4. It can be said that textbooks for lower grades try to express subtle nuance by using a larger number of auxiliaries and that assertive expressions are frequently used in textbooks for higher grades.

4. CONCLUSIONS

We investigated some characteristics of character- and word-appearance of English textbooks for elementary school students in Finland. The results show that frequency characteristics of character-appearance in Finnish materials have a tendency similar to that seen in English journalism, and their difficulty level is almost the same as or even more difficult than that of the those for Japanese junior high school students.

We will continue research on the characteristics of English textbooks used in foreign countries, and will consider how to apply the analysis results to English language education.

REFERENCES

[1] English textbooks used in Finland [<http://www.kknews.co.jp/developer/finland/index.html>]