

Text Mining of English Materials for Business Management

Hiromi Ban, Haruhiko Kimura, Takashi Oyabu

Abstract— Today, as globalization progresses, the economy and management of each country have become increasingly interdependent, and the knowledge of business management has become more important. Business management is a science that treats of management of business, which is one of the most important factors constituting modern society. It was born in the United States about 100 years ago, and its research has been prolific there ever since. Thus, reading materials in English are indispensable to study it. If we have beforehand enough knowledge of the features of English in the field, reading of the texts will become easier. In this paper, we metrically analyzed some famous English books on business management, comparing these with English journalism and a computer book. We used an approximate equation of an exponential function to extract the characteristics of each material using coefficients c and b of the equation. Moreover, we calculated the percentage of Japanese junior high school required vocabulary and American basic vocabulary to obtain the difficulty-level as well as the K -characteristic. As a result, English materials for management have the same tendency as English literature in the character-appearance. The values of the K -characteristic for the materials on management are high, compared with the journalism. Moreover, the books on management are easier to read than *BusinessWeek*. Besides, we inquired into the word-length distribution of the most frequently used 100 words. It has been cleared that while the distribution for journalism corresponds to the normal distribution, the distribution for the books on management corresponds to the Poisson distribution.

Index Terms—Business management, Computational linguistics, Statistical analysis, Text mining

I. INTRODUCTION

Today, as globalization progresses, the economy and management of each country have become increasingly interdependent, and the knowledge of business management has become more important. Business management is a science that treats of management of business, which is one of the most important factors constituting modern society. It was born in the United States about 100 years ago, and its research has been prolific there ever since. Thus, reading materials in English are indispensable to study it [1]. If we have beforehand enough knowledge of the features of English in the field, reading of the texts will become easier.

In this paper, we investigated several famous English books on business management, comparing them with English journalism and a computer book in terms of metrical

linguistics. As a result, it was clearly shown that English materials for management have some interesting characteristics about character- and word-appearance.

II. METHOD OF ANALYSIS AND MATERIALS

The materials analyzed here are as follows:

- Material 1: Thomas J. Peters and Robert H. Waterman, Jr., *In Search of Excellence*, HarperCollins, 1982
- Material 2: Michael E. Porter, *Competitive Strategy*, Free Press, 1998
- Material 3: Robert C. Higgins, *Analysis for Financial Management*, 5th ed., McGraw-Hill, 1998
- Material 4: Philip Kotler, *Marketing Management*, Millennium ed., Prentice-Hall, 2000

We examined the first three chapters of each material.

For comparison, we analyzed the famous economic magazines “The Economist” published on January 4-10 in 2003 and “BusinessWeek” published on January 13 in 2003, as well as the American popular news magazine “TIME” published on January 13 in 2003. In addition, we examined the introductory book to computers “Computing Essentials” written by Don Cassel issued from the Prentice-Hall in 1994, because the progress of management is closely related to the development of computers and network systems. Deleting pictures, headlines, etc., we used only the texts.

The computer program for this analysis is composed of C++. Besides the characteristics of character- and word-appearance for each piece of material, various information such as the “number of sentences,” the “number of paragraphs,” the “average of word length,” the “number of words per sentence,” etc. can be extracted by this program [2].

III. RESULTS

A. Characteristics of Character-appearance

First, the most frequently used characters in each material and their frequency were derived. The frequencies of the 50 most frequently used characters including the blanks, capitals, small letters, and punctuations were plotted on a descending scale. The vertical shaft shows the degree of the frequency and the horizontal shaft shows the order of character-appearance. The vertical shaft is scaled with a logarithm. This characteristic curve was approximated by the following exponential function:

$$y = c * \exp(-bx) \quad (1)$$

Hiromi Ban, Graduate School of Engineering, Nagaoka University of Technology, Nagaoka, Niigata, Japan, +81-80-5042-7437.

Haruhiko Kimura, Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa, Ishikawa, Japan, +81-76-234-4836.

Takashi Oyabu, Kokusai Business Gakuin College, Kanazawa, Ishikawa, Japan, +81-76-222-3196.

From this function, we are able to derive coefficients c and b [3]. The distribution of coefficients c and b extracted from each material is shown in Figure 1.

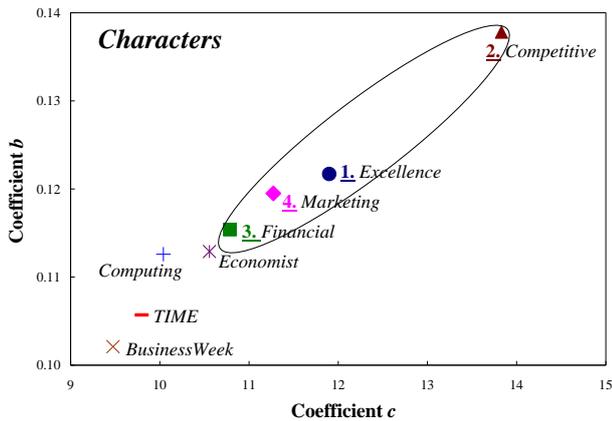


Figure 1: Dispersions of coefficients c and b for character-appearance.

There is a linear relationship between c and b for the eight materials. The values of coefficients c and b for Materials 1 to 4 are high: the value of c ranges from 10.786 (Material 3) to 13.830 (Material 2), and that of b is 0.1154 (Material 3) to 0.1378 (Material 4). On the other hand, in the case of the American economic magazine *BusinessWeek*, c is 9.4758 and b is 0.1021, both of which are lowest of the eight materials. Previously, we analyzed various English writings and reported that there is a positive correlation between the coefficients c and b , and that the more journalistic the material is, the lower the values of c and b are, and the more literary, the higher the values of c and b [4]. Thus, the materials on management have a similar tendency to literary writings.

B. Characteristics of Word-appearance

Next, the most frequently used words were derived. Just as in the case of characters, the frequencies of the 50 most frequently used words in each material were plotted. Each characteristic curve was approximated by the same exponential function. The distribution of c and b is shown in Figure 2.

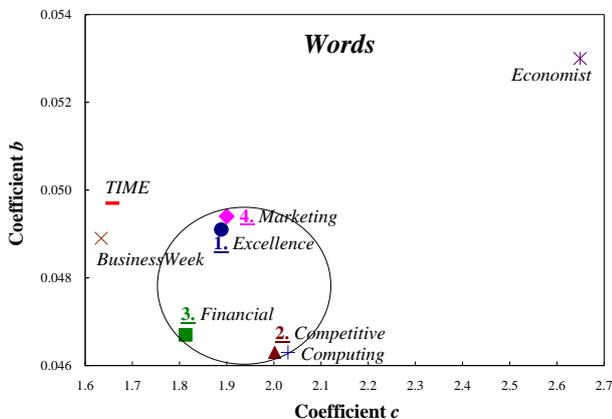


Figure 2: Dispersions of coefficients c and b for word-appearance.

While the values of c for Materials 1 to 4 are between *TIME* and *The Economist*, those of b are lower than *TIME*.

Although we cannot see a positive correlation between coefficients c and b such as in the case of character-appearance, the values for Materials 1 to 4 are relatively similar and we might be able to regard them as a cluster.

As a method of featuring words used in writing, a statistician named Udny Yule suggested an index called the “ K -characteristic” in 1944 [5]. This can express the richness of vocabulary in writings by measuring the probability of any randomly selected pair of words being identical. He tried to identify the author of *The Imitation of Christ* using this index. This K -characteristic is defined as follows:

$$K = 10^4 (S_2 / S_1^2 - 1 / S_1) \tag{2}$$

where if there are f_i words used x_i times in a writing, $S_1 = \sum x_i f_i$, $S_2 = \sum x_i^2 f_i$.

We examined the K -characteristic for each material. The results are shown in Figure 3.

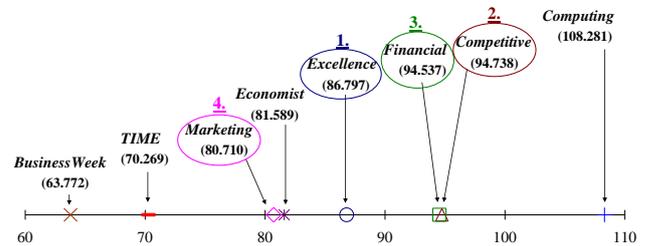


Figure 3: K -characteristic for each material.

According to the figure, Material 3 ($K = 94.537$) and Material 2 (94.738), and Material 4 (80.710) and *The Economist* (81.589) have almost the same values respectively. As for the four materials for business management, the values for them are higher than *TIME* and *BusinessWeek*, and lower than *COMPUTING ESSENTIALS*, and the value gradually increases in the order of Material 4, Material 1, Material 3 and Material 2. This order corresponds with the coefficient b for word-appearance in reversed order. We would like to investigate the relationship between K -characteristic and the coefficients for word-appearance in the future.

C. Degree of Difficulty

In order to show how difficult the materials for readers are, we derived the degree of difficulty for each material through the variety of words and their frequency [6, 7]. That is, we came up with two parameters to measure difficulty; one is for word-type or word-sort (D_{ws}), and the other is for the frequency or the number of words (D_{wn}). The equation for each parameter is as follows:

$$D_{ws} = (1 - n_{rs} / n_s) \tag{3}$$

$$D_{wn} = \{ 1 - (1 / n_t * \sum n(i)) \} \tag{4}$$

where n_t means the total number of words, n_s means the total number of word-sort, n_{rs} means the required English vocabulary in Japanese junior high schools or American basic vocabulary by *The American Heritage Picture Dictionary* (American Heritage Dictionary, Houghton Mifflin, 2003), and $n(i)$ means the respective number of each required or

basic word. Thus, we can calculate how many required or basic words are not contained in each piece of material in terms of word-sort and frequency.

Thus, we calculated the values of both D_{ws} and D_{wn} to show how difficult the materials are for readers, and to show at which level of English the materials are compared with other materials. In order to make the judgments of difficulty easier for the general public, we derived one difficulty parameter from D_{ws} and D_{wn} using the following principal component analysis:

$$z = a_1 * D_{ws} + a_2 * D_{wn} \quad (5)$$

where a_1 and a_2 are the weights used to combine D_{ws} and D_{wn} . Using the variance-covariance matrix, the 1st principal component z was extracted: $z = (0.5672 * D_{ws} + 0.8236 * D_{wn})$ for the required vocabulary, and $z = (0.4636 * D_{ws} + 0.8861 * D_{wn})$ for the basic vocabulary, from which we calculated the principal component scores. The results are shown in Figure 4.

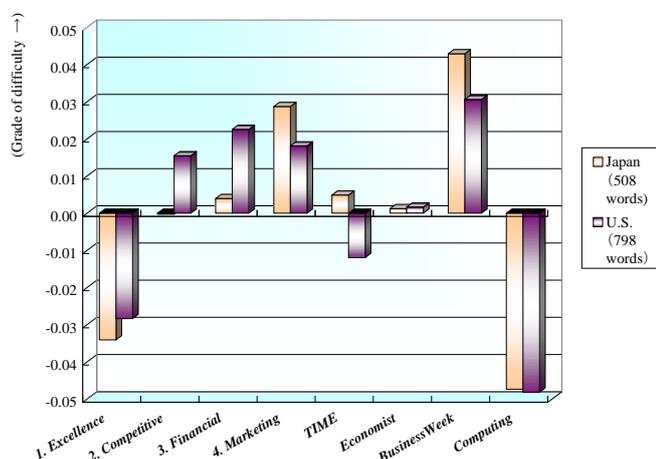


Figure 4: Principal component scores of difficulty shown in one-dimension.

According to Figure 4, the difficulty level increases in the order of Material 1, Material 2, Material 3 and Material 4. The difficulty of these four materials much varies: while the easiest Material 1 is a little more difficult than *Computing Essentials*, which is the easiest of the eight materials, because it is an introductory book, the most difficult Material 4 is more difficult than *TIME* and *The Economist*. On the other hand, in the case of the basic vocabulary, Material 3 is a little more difficult than Material 4. We can judge that the three materials for business management, that is, Materials 2, 3 and 4 are more difficult than *TIME* and *The Economist*, and easier than *BusinessWeek*, which is the most difficult of the eight materials.

D. Other Characteristics

Other metrical characteristics of each material were compared. The results of the “average of word length,” the “number of words per sentence,” etc. are shown together in Table 1. Although we counted the “frequency of relatives,” the “frequency of modal auxiliaries,” etc., some of the words counted might be used as other parts of speech because we didn’t check the meaning of each word.

1) Average of word length

As for the “average of word length” for the four materials for business management, it varies from 6.071 letters for Material 1 to 6.378 letters for Material 4. They are a little longer than *Computing Essentials* (5.808 letters) and journalism (5.853 to 5.980 letters). It seems that this is because the materials for business management contain many long-length technical terms for management such as *MARKETING* and *ACCOUNTING*.

2) Number of words per sentence

The “number of words per sentence” for Material 2 is 27.096 words, which is the most of the eight materials, and approximately 10 words more than *BusinessWeek* (17.878 words), which is the fewest. From this point of view, the Material 2 seems to be rather difficult to read. In the case of other three materials for business management, it is 19.002 (Material 4) to 22.537 (Material 3) words, which are a little fewer than *TIME* (24.931 words) and almost the same as *Computing Essentials* (19.546 words) and *The Economist* (21.682 words).

3) Number of commas per sentence

The “number of commas per sentence” for Materials 1 to 4 is from 1.062 (Material 4) to 1.376 (Material 1), which is almost the same as the three journalism (1.122 to 1.389).

4) Frequency of auxiliaries

There are two kinds of auxiliaries in a broad sense. One expresses the tense and voice, such as *BE* which makes up the progressive form and the passive form, the perfect tense *HAVE*, and *DO* in interrogative sentences or negative sentences. The other is a modal auxiliary, such as *WILL* or *CAN* which expresses the mood or attitude of the speaker [8]. In this study, we targeted only modal auxiliaries. As for the result, the “frequency of auxiliaries” is highest in Material 2 (2.438%), which is more than three times of Material 1 (0.801%) and twice of *TIME* (1.125%). Therefore, it might be said that while the writer of Material 2 tends to communicate his subtle thoughts and feelings with auxiliary verbs, the style of Material 1 and *TIME* can be called more assertive.

E. Characteristics of Preposition, Relative, Auxiliary, and Personal Pronoun Appearance

Next, we examined in detail the “prepositions,” “relatives,” “modal auxiliaries,” and “personal pronouns” of each material. We valued each part of speech used in each material at 100%, and checked the kind of words and its frequency. As for Relatives, *WHICH* and *HOW* are frequently used for Materials 1 to 4: *WHICH* is the 2nd to 5th, and *HOW* is the 4th to 6th most frequently used. These days, *THAT* has been taking place of *WHICH* [9]. Therefore, the literally style of the materials for business management might be older. *HOW* is also frequently used. This seems to be because the contents of these materials are mainly about consideration of some methods for solving a problem. In the case of Auxiliaries, the frequency of *CAN*, which often means possibility of

Table 1: Metrical data for each material.

	<u>1. Search of Excellence</u>	<u>2. Competitive Strategy</u>	<u>3. Financial Management</u>	<u>4. Marketing Management</u>	<i>TIME 2003</i>	<i>Economist 2003</i>	<i>BusinessWeek 2003</i>	<i>Computing Essentials</i>
Total num. of characters	165,785	140,494	161,076	258,199	163,880	297,739	272,309	80,602
Total num. of character-type	80	75	84	84	82	80	82	78
Total num. of words	27,309	22,029	26,368	40,569	27,998	50,150	45,534	13,878
Total num. of word-type	5,050	3,286	3,573	5,586	7,083	8,665	8,053	2,224
Total num. of sentences	1,325	813	1,170	2,135	1,123	2,313	2,547	710
Total num. of paragraphs	238	253	256	401	284	599	573	194
Average of word length	6.071	6.378	6.109	6.364	5.853	5.937	5.980	5.808
Words/sentence	20.611	27.096	22.537	19.002	24.931	21.682	17.878	19.546
Repetition of a word	5.408	6.704	7.380	7.263	3.952	5.937	5.654	6.240
Commas/sentence	1.376	1.224	1.187	1.062	1.389	1.271	1.122	0.785
Sentences/paragraph	5.567	3.213	4.570	5.324	3.954	3.861	4.445	3.660
Freq. of prepositions	14.899	15.189	14.517	12.606	14.641	16.006	15.265	14.246
Freq. of relatives	2.878	2.260	2.049	2.059	2.404	2.341	1.857	2.514
Freq. of auxiliaries	0.801	2.438	1.482	1.716	1.125	1.404	1.430	1.484
Freq. of personal pronouns	5.759	2.324	2.662	3.177	5.375	3.496	3.075	1.708

something, is high: it is the 1st or 2nd in the four materials for management. As for Personal Pronouns, *ITS* and *WE* are used frequently: while *ITS* is the most or the 2nd most frequently used in Materials 2 to 4, *WE* is the 1st to 6th in the four materials for management.

Next, the frequencies of the most frequently used words, that is, the top 44 for Prepositions, 9 for Relatives, 8 for Auxiliaries, and 14 for Personal Pronouns in each material were plotted on a descending scale. The vertical shaft was scaled with a logarithm. Each characteristic curve was approximated by the exponential function: $[y = c * \exp(-bx)]$. We derived coefficients c and b for each part of speech. The results are shown in Table 2. As a result, in the case of Relatives, the value of c is high for the four materials on management as a whole: it is 23.809 (Material 3) to 52.564 (Material 2). On the other hand, in the case of Auxiliaries, as for the three materials for management except for Material 2, the value of c is 30.643 (Material 4) to 32.581 (Material 3) and b is 0.2349 (Material 4) to 0.2638 (Material 1), both of which are lower than other materials. This means that more kinds of auxiliaries are used in the materials for management.

F. Word-length Distribution of the Top 100 Words

We examined the word-length distribution of the most frequently used 100 words of each material. Then, we calculated the variance, standard deviation and coefficient of variation for the distribution. The results are shown in Table 3. As a result, the coefficients of variation for the four materials for management are 49.065 (Material 1) to 55.333 (Material 2), which are higher than three journalism materials, which are 31.582 (*TIME*) to 42.257 (*The Economist*).

Therefore, we can say that the variation of the word-length for the materials on management is bigger than that for journalism.

Table 3: Coefficients of variation for word-length distribution of the top 100 words.

Material	Total words	Average of word length	Variance	Standard Deviation	cv (%) ($\sigma / \bar{x} * 100$)
<u>1. Search of Excellence</u>	7,692	3.905	3.669	1.916	49.065
<u>2. Competitive Strategy</u>	7,502	4.753	6.918	2.630	55.333
<u>3. Financial Management</u>	8,095	4.636	5.888	2.427	52.351
<u>4. Marketing Management</u>	12,062	4.798	5.794	2.407	50.167
<i>TIME</i>	6,844	3.426	1.171	1.082	31.582
<i>Economist</i>	12,556	3.687	2.427	1.558	42.257
<i>BusinessWeek</i>	10,768	3.935	2.532	1.591	40.432
<i>Computing Essentials</i>	4,686	4.547	5.153	2.270	49.065

Next, the results of the word-length distribution of the most frequently used 100 words of Material 2, Material 4, *TIME* and *The Economist* are shown in Figure 5. As a result, we can see that while the distribution for journalism such as *TIME* and *The Economist* corresponds to the normal distribution, the distribution for the books on management such as Materials 2 and 4 corresponds to the Poisson distribution.

Moreover, we inquired into the coefficient of variation for the word-length distribution of the most frequently used 100 words except for articles and prepositions. The results are shown in Table 4. In this case, the coefficients of variation for the four materials for management are 32.512 (Material 2) to 36.125 (Material 3), which are lower than three journalism materials, which are 36.886 (*The Economist*) to 40.532

Table 2: Coefficients c and b of each part of speech for each material.

Material	Prepositions (top 44 words)		Relatives (top 9 words)		Auxiliaries (top 8 words)		Personal pronouns (top 14 words)	
	c	b	c	b	c	b	c	b
<u>1. Search of Excellence</u>	10.1680	0.1277	38.2800	0.3898	32.3570	0.2638	27.5320	0.2395
<u>2. Competitive Strategy</u>	9.8237	0.1313	61.3060	0.4527	72.6150	0.5534	52.5640	0.4634
<u>3. Financial Management</u>	7.9657	0.1157	41.0530	0.4007	32.5810	0.2588	23.8090	0.2273
<u>4. Marketing Management</u>	9.5009	0.1293	36.7370	0.3332	30.6430	0.2349	31.3820	0.2909
<i>TIME</i>	9.5259	0.1153	40.3220	0.3583	39.1550	0.3022	15.9560	0.1396
<i>Economist</i>	9.0504	0.1135	35.5230	0.3473	45.3260	0.3543	31.9980	0.2808
<i>BusinessWeek</i>	9.6760	0.1170	39.2510	0.3847	47.1920	0.3742	31.0620	0.2735
<i>Computing Essentials</i>	9.7093	0.1383	62.8310	0.5098	51.2740	0.4201	23.9830	0.2345

(BusinessWeek). This means that the variation of the word-length for the materials on management is less than that for journalism.

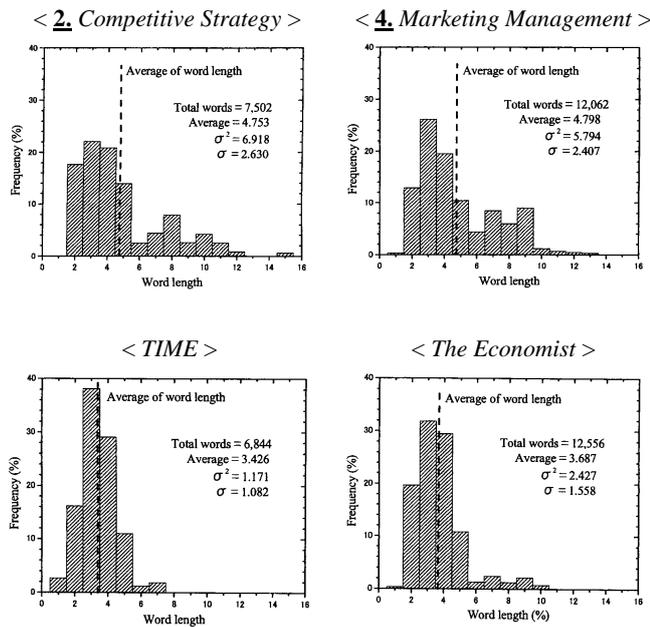


Figure 5: Word-length distribution of the top 100 words.

Table 4: Coefficients of variation for word-length distribution of the top 100 words except for articles and prepositions.

Material	Total words	Average of word length	Variance	Standard Deviation	cv (%) ($\sigma/\bar{x} * 100$)
1. Search of Excellence	8,005	2.333	0.624	0.789	33.819
2. Competitive Strategy	4,862	2.353	0.585	0.765	32.512
3. Financial Management	5,848	2.364	0.729	0.854	36.125
4. Marketing Management	5,881	2.392	0.612	0.783	32.734
TIME	5,921	2.401	0.814	0.902	37.568
Economist	11,273	2.402	0.785	0.886	36.886
BusinessWeek	9,761	2.482	1.013	1.006	40.532
Computing Essentials	3,292	2.272	0.612	0.782	34.419

IV. APPLICATION TO EDUCATION

Using the three dictionaries of accounting terms, we checked what technical terms for management are included in each material. The top 20 nouns and their percentages for Material 2 and Material 3 are shown in Table 5. While the frequencies of INDUSTRY, COST and FIRM, including both singular and plural forms, are 1.058%, 0.940% and 0.881% respectively of all the words used in Material 2, the frequencies of CASH, COMPANY and ASSET are 0.747%, 0.971% and 0.729% respectively in Material 3.

As for Materials 2 and 3, the top 20 technical terms occupy as much as 6.897% and 6.786% respectively of all words. In the case of Material 1 and 4, the percentage is 3.039% and 7.602% respectively. If we teach beforehand these technical terms for management to students, reading of the texts will become easier.

Table 5: High-frequency technical terms for management and their percentages for each material.

	2. Competitive Strategy		3. Financial Management	
	Word	%	Word	%
1	INDUSTRY	1.058	CASH	0.747
2	COST	0.545	COMPANY	0.656
3	FIRMS	0.468	ASSETS	0.501
4	FIRM	0.413	VALUE	0.425
5	COSTS	0.395	SALES	0.391
6	STRATEGY	0.386	INCOME	0.368
7	ENTRY	0.377	MILLION	0.330
8	PRODUCT	0.363	COMPANIES	0.315
9	MARKET	0.340	EQUITY	0.315
10	POSITION	0.309	PERCENT	0.307
11	BUSINESS	0.295	RATIO	0.303
12	ANALYSIS	0.259	ACCOUNTING	0.296
13	GOALS	0.259	INTEREST	0.258
14	SCALE	0.236	RATIOS	0.235
15	BARRIERS	0.209	COST	0.231
16	DIFFERENTIATION	0.209	STATEMENT	0.231
17	SHARE	0.204	ASSET	0.228
18	EXPERIENCE	0.200	PERFORMANCE	0.224
19	COMPANY	0.186	BALANCE	0.216
20	MOVES	0.186	STATEMENTS	0.209
Total		6.897		6.786

V. CONCLUSIONS

We investigated some characteristics of character- and word-appearance of some famous English books on management, comparing these with English journalism and a computer book. In this analysis, we used an approximate equation of an exponential function to extract the characteristics of each material using coefficients c and b of the equation. Moreover, we calculated the percentage of Japanese junior high school required vocabulary and American basic vocabulary to obtain the difficulty-level as well as the K -characteristic. As a result, English materials for management have the same tendency as English literature in the character-appearance. The values of the K -characteristic for the materials on management are high, compared with the journalism. Moreover, the books on management are easier to read than BusinessWeek. Besides, we inquired into the word-length distribution of the most frequently used 100 words. It has been cleared that while the distribution for journalism corresponds to the normal distribution, the distribution for the books on management corresponds to the Poisson distribution.

In the future, we plan to apply these results to education. For example, we would like to measure the effectiveness of teaching the 100 most frequently used words in a certain material beforehand.

REFERENCES

- [1] H. Ban, T. Dederick, H. Nambo, and T. Oyabu, "Metrical Comparison of English Materials for Business Management and Information Technology," *Proceedings of the 5th Asia-Pacific Industrial Engineering and Management Systems Conference 2004*, pp.33.4.1-33.4.10, 2004.
- [2] H. Ban, T. Dederick, and T. Oyabu, "Linguistical Characteristics of Eliyahu M. Goldratt's "The Goal"," *Proceedings of the 4th Asia-Pacific Conference on Industrial Engineering and Management Systems*, pp.1221-1225, 2002.
- [3] H. Ban, T. Dederick, and T. Oyabu, "Metrical Comparison of Singapore English Newspapers and Other English Journalism," *Proceedings of the 6th International Conference on Engineering Design and Automation*, pp.717-722, 2002.

- [4] H. Ban, T. Sugata, T. Dederick, and T. Oyabu, "Metrical Comparison of English Columns with Other Genres," *Proceedings of the 5th International Conference on Engineering Design and Automation*, pp.912-917, 2001.
- [5] G. U. Yule, *The Statistical Study of Literary Vocabulary*, Cambridge University Press, 1944.
- [6] H. Ban, T. Dederick, H. Nambo, and T. Oyabu, "Relative Difficulty of Various English Writings by Fuzzy Inference and Its Application to Selecting Teaching Materials," *An International Journal of Industrial Engineering & Management Systems*, 3(1), pp.85-91, 2004.
- [7] H. Ban, T. Dederick, and T. Oyabu, "Metrical Comparison of English Textbooks in East Asian Countries, the U.S.A. and U.K.," *Proceedings of the 4th International Symposium on Advanced Intelligent Systems*, pp.508-512, 2003.
- [8] H. Ban, T. Dederick, H. Nambo, and T. Oyabu, "Stylistic Characteristics of English News," *Proceedings of the 2004 Japan-Korea Joint Symposium on Emotion and Sensibility*, 4 pages, 2004.
- [9] H. Ban, T. Sugata, T. Dederick, and T. Oyabu, "Linguistical Analysis of American Presidents' Inaugural Addresses," *Proceedings of the 3rd Asia-Pacific Conference on Industrial Engineering and Management Systems*, pp.47-54, 2000.

Hiromi Ban received her B.A. and M.A. from Japan Women's University. She joined Toyama University of International Studies, Toyama, Japan as a lecturer in 1993. Currently, she is a professor at Graduate School of Engineering, Nagaoka University of Technology. She is engaged in researches on computational linguistics and text mining. She is a member of IEICE, SOFT (Japan), and JSKE.



Haruhiko Kimura graduated in 1974 Tokyo Denki Univ. In 1979, he received his Ph.D. degree from Tohoku Univ. Also in 1979, he joined in Fujitsu Corporation. In 1980, he became a lecturer in Kanazawa Women's Junior College. In 1984, he became an assistant professor in Kanazawa Univ. Currently, he is a professor in Kanazawa Univ. He is engaged in researches on optimal code conversion and the acceleration of production systems. He is a member of IEICE, ISAI (Japan), and IPSJ.



Takashi Oyabu was born in 1949. He received his B.E. and M.E. in 1971 and 1973 from Kogakuin University and also received B.A. in 1975 from Waseda University, Tokyo, Japan. In 1984, he received the Dr of Engineering from Kogakuin University. He joined the Electronics Research Laboratory of Denki Onkyo Co. Ltd., Tokyo in 1973. From 1991-1998 he was a Professor of Toyama University of International Studies. He joined Kanazawa Seiryu University, Kanazawa, Japan as a Professor of Graduate School of Regional Economic Systems in 1998. He is the President of Society for Tourism and Informatics of Japan and the academic President of Kanazawa Institute of Tourism. His current research activities are on tourism with advanced information technology and its applications to welfare fields.

