

Chapter 177

An Application of Bibliometrics to Marketing Planning of Election Campaigns

Kuo-Tai Tang and Chun-Yi Liu

Abstract In the election campaigns in Taiwan, a candidate's White Book, a book that outlines his or her political views, is the most important source of information for the voters. Using bibliometrics to analyze a White Book, this study identifies key political terms and simplifies the information in the White Book to facilitate the campaign planning in promoting voter preference for candidates and generating votes, with the ultimate aim of assisting a candidate to win election. This chapter investigated Taiwanese President Ma Ying-jeou's 21 major political views in his White Book during the 2008 election in Taiwan, a text totaling 60,571 vocabularies. The key vocabulary distribution in the White Book was: Taiwan, development, government, culture, industry, education, policy, international, society, tourism, economy, nation, ocean, and environment.

Keywords Bibliometrics · Zipf Laws · White Books · Marketing planning · Election · Vocabulary · Political views

Introduction

In Taiwan's election campaigns, the White Book is the most important source of information for the voters on the candidates' political views. For this reason, simplifying the message of the White Book should help voters form deep impressions and goodwill toward candidates, establish voter preferences, generate votes, and ultimately, win the election. Accordingly, in addition to acting as a reference for political slogans, the simplified White Book could be an important element of successful campaign communication and promotion.

In this study, bibliometrics was used to analyze Taiwanese President Ma Ying-jeou's 2008 election White Book [1]. Zipf Laws illuminated the relevance of

C.-Y. Liu (✉)

Department of Information Management, Yuanpei University,
30015 Hsinchu City, Taiwan, China
e-mail: tojo@ms38.hinet.net

vocabulary frequency, identifying the distribution of the most thematically significant vocabularies. Next, through the distribution of these vocabularies, we extracted the primary political content and terms from the 60,571 vocabularies White Book and simplified the political content therein to facilitate campaign planning.

Literature Review

Bibliometrics includes three basic laws: Bradford Laws, Lotka Laws, and Zipf Laws. George Kingsley Zipf found that vocabulary simplicity has a high correlation with frequency of use, while it appears to have an inverse relationship with the number of occurrences [2]. In 1965, he presented the Principle of Least Effort, which claimed that human beings tend to use the simplest and most efficient vocabularies, and that the relationship between a vocabulary frequency and its rank is defined as the two multiplied together equaling a constant (i.e., $r \times f = c$). This is known as Zipf Laws [3], where r is the rank of vocabulary, f is the frequency of vocabulary, c is the constant, and rf is the product of rank and frequency. Zipf Laws primarily investigate vocabulary distribution, studying the relationship between the distribution of vocabularies, and their frequency.

Goffman [4] asserts that a transition point should exist between high- and low-frequency vocabularies, and that the vocabularies in the transition point are the most significant vocabularies in the literature. He also proposes a formula, $n = (-1 + \sqrt{1 + 8I_1})/2$, where n is the transition value between high and low frequency, I_1 are vocabularies that only occur once in the literature, and I_n is the total number of vocabularies that appear n times. The formula shows that the transition point found in the literature by removing vocabularies with a value greater than and less than n , and the remaining vocabularies are the most significant vocabularies in the literature.

Methodology

This chapter investigates Ma Ying-jeou's 21 major political views in his White Book during the 2008 election in Taiwan, a text totaling 60,571 vocabularies. We first applied Zipf Laws to discover the distribution of vocabularies, then used Goffman's theory to find the distribution of key vocabularies, and finally confirmed compliance with Zipf Laws' inverse square distribution. Using the results of the bibliometrics analysis, we identified the key vocabularies.

The White Book's statistics tool of vocabulary segmentation and frequency used the "Chinese vocabulary segmentation system" developed by the Academia Sinica's Institute of Information Sciences [5].

Table 177.1 Key vocabulary distribution in the transition point

Rank (<i>r</i>)	Frequency (<i>f</i>)	<i>r</i> <i>f</i>	Vocabulary	Rank (<i>r</i>)	Frequency (<i>f</i>)	<i>r</i> <i>f</i>	Vocabulary
4	284	1,136	Taiwan	41	71	2,911	Democracy
7	211	1,477	Development	43	70	3,010	Youth
10	179	1,790	Government	44	69	3,036	Aboriginal
12	172	2,064	Culture	45	68	3,060	System
15	152	2,280	Industry	46	68	3,128	Family
16	150	2,400	Education	48	67	3,216	Employment
18	138	2,484	Policy	56	62	3,472	Resources
22	128	2,816	International	61	58	3,538	Services
24	125	3,000	Society	62	58	3,596	Implementation
25	113	2,825	Tourism	63	56	3,528	Hakka
30	97	2,910	Economy	64	55	3,520	Security
33	84	2,772	Nation	66	55	3,630	Reign
34	83	2,822	Ocean	67	55	3,685	Women
36	80	2,880	Environment	68	55	3,740	Job

Analysis and Results

Bibliometrics Analysis

Vocabulary Distribution in the White Book

Based on Zipf Laws ($r \times f = c$), as the vocabularies descended in rank-order according to the number of occurrences, under the confidence interval with a 95% confidence level, the standard deviation was 804.02 and the average, 4,636, was similar to the median, 4,822, indicating that the vocabulary distribution for the White Book appeared to be normally distributed. With a kurtosis, -0.29 , slightly lower than the normal distribution and a skewness, -0.7 , less than 0, there appeared to be a relatively large negative deviation value, creating a skew to the left. The White Book contained 60,571 different vocabularies appearing a total of 37,799 times; 2,817 vocabularies occurred only once, accounting for 7.45% of all occurrences, while 39 vocabularies occurred more than 100 times for a total of 12,536 times, or 33.16% of the total number of occurrences.

Key Vocabulary Distribution

After excluding synonyms, articles, and prepositions, we calculated the transition point to be 74.55, meaning that vocabularies, which occurred 75 times represented the point when vocabulary distribution shifted from high frequency to low frequency. As Table 177.1 shows, a total of 28 vocabularies center around the transition point (14 before and 14 after) and were therefore located within the transition point, making them the vocabularies most representative of the White Book's message.

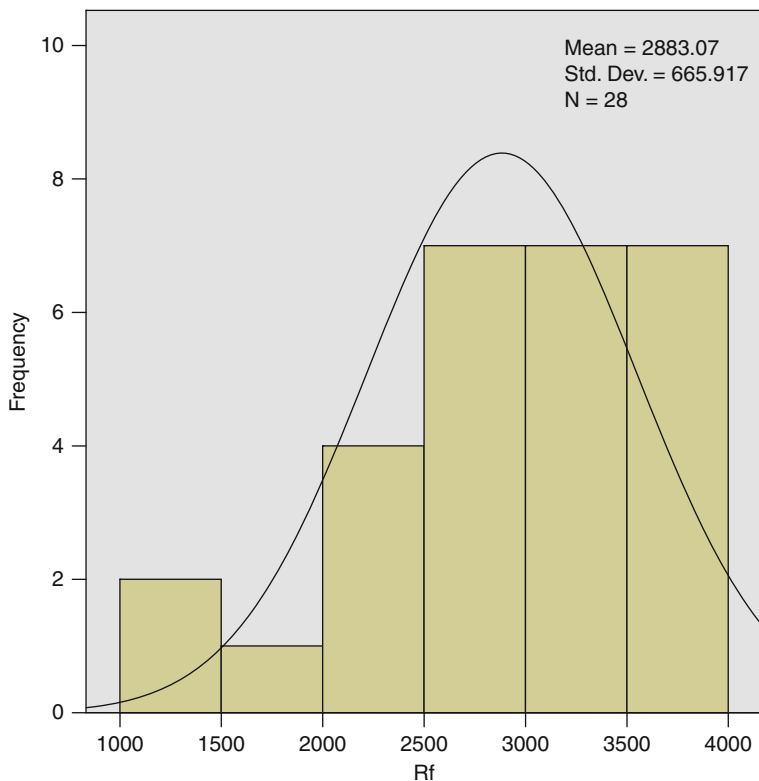


Fig. 177.1 Ma Ying-jeou’s White Book—Normal distribution of transition point key vocabularies

According to Table 177.1 (key vocabulary distribution in the transition point), under the confidence interval with a 95% confidence level, the standard deviation was 665.91, and the average, 2,883, and median, 2,955, were nearly equivalent, indicating that the key vocabularies in the transition point were normally distributed. With a kurtosis of 0.752, and a skewness, at -0.99 , less than 0, the distribution of keywords in the transition point appeared normally distributed and skewed to the left, as shown in Fig. 177.1 (Ma Ying-jeou’s White Book—Normal distribution of transition point key vocabularies). The average ($2,883 \pm 665$) indicates that 71% of the vocabulary distribution was within the standard deviation, signifying that the key vocabulary distribution was extremely concentrated.

The aforementioned results show little difference between the standard deviation for the “vocabulary distribution in the White Book” and the standard deviation for “key vocabulary distribution.” The median for the key vocabulary distribution after removing unnecessary vocabularies was significantly less than the median before the vocabularies were removed. This difference indicates that, due to the deconstruction of the 21 major themed political views in the White Book, the vocabulary distribution for the latter was effective, while the former tended toward vocabulary concentration.

Table 177.2 The key vocabularies of Ma Ying-jeou's White Book

Number	Rank(<i>r</i>)	Percentage (%)	Frequency(<i>f</i>)	Vocabulary
1	4	14.22%	284	Taiwan
2	7	10.57%	211	Development
3	10	8.96%	179	Government
4	12	8.62%	172	Culture
5	15	7.62%	152	Industry
6	16	7.52%	150	Education
7	18	6.91%	138	Policy
8	22	6.41%	128	International
9	24	6.26%	125	Society
10	25	5.66%	113	Tourism
11	30	4.86%	97	Economy
12	33	4.21%	84	Nation
13	34	4.16%	83	Ocean
14	36	4.01%	80	Environment

Inverse Square Distribution

According to Zipf Laws, if, for the number of vocabularies, which appeared once in the sample, $x = 2,817$, then the number of vocabularies, which appear two, three, four ... n times should be the "inverse square" of x . The frequency of vocabularies appearing 1–16 times estimated by their inverse squares was approximately equal to the actual frequency found (57%). The results, therefore, comply with Zipf Laws.

Critical Political Content Analysis

Based on the above analysis, we know that 39 vocabularies appeared more than 100 times with a total of 12,536 appearances, 33.16% of all vocabulary appearances. Also, according to Table 177.1, for the 36 key vocabularies ranked from high frequency to low frequency, the top 14 shows a high frequency of appearance and growth can be regarded as the critical political views. These 14 vocabularies are displayed in Table 177.2.

In regards to number of occurrences, the top vocabulary was "Taiwan," then "development," followed by "government." These vocabularies coincide with Ma Ying-jeou's campaign slogans during the election, "A forward-marching Taiwan is a winning Taiwan" and "Love Taiwan, revitalize the economy." These vocabularies also coincide to cross-strait policies emphasizing economic and trade exchange and opening of Taiwan to Mainland tourists.

Conclusions and Recommendations

The results of this study can assist candidates in election campaigns by serving as a reference in the simplification of their political views as outlined in White Books and in the formulation of campaign slogans, and in the development of a marketing plan for a successful election by increasing voters' (consumers') knowledge of the candidate (the product). The key vocabulary distribution, in line with the often-quoted critical political views of winning candidates, was: Taiwan, development, government, culture, industry, education, policy, international, society, tourism, economy, nation, ocean, and environment.

The most frequently occurring vocabulary was "Taiwan," followed by "development" and "government," showing the high-profile nature of "Taiwan's development." Similarly, "economic and industrial development policy" and "tourism and economic development" are the primary directions of candidates' political views, while cultural policy, education policy, and the marine environment are only a focus after a candidate gets elected.

References

1. Ma Ying-Jeou Election Campaign Website, <http://www.ma19.net>
2. Robert M. Losee: Term Dependence: A Basis for Luhn and Zipf Models. *Journal of the American Society for Information Science and Technology*, vol. 52(12), pp. 1019–1025, (2001)
3. Zipf, George Kingsley: *Human Behavior and Principle of Least Effort*. Cambridge, Mass: Addison-Wesley, 1949, reprint ed., Hafner, pp. 24, New York (1965)
4. Goffman, William, Morris, Thomas G.: Bradford's 'Law' Applied to the Maintenance of Library Collection. In: *Introduction to Information Science*, edited by Tefko Saracevic, R. R. Bowker, pp. 200–203, New York (1970)
5. Feng-Tyan Lin, Ging-Yuan Zeng: Chinese vocabulary segmentation system. Thesaurus Group, Institute of Information Science, Academia Sinica, <http://ckipsvr.iis.sinica.edu.tw>