Title	On a new method for content analysis of TV news: the 2010 house of councillors election as an example
Sub Title	
Author	河野, 武司(Kono, Takeshi)
Publisher	Global Center of Excellence Center of Governance for Civil Society, Keio University
Publication year	2011
Jtitle	Journal of political science and sociology No.15 (2011. 9) ,p.121- 131
JaLC DOI	
Abstract	The purpose of this paper is to introduce a new method for content analysis of TV news using TV coverage of the 22nd House of Councillors election as a case. Instead of human power, I make the most of computer power to analyze the content of TV news. I will show that the combination of TV news archive that has begun to appear in Japanese business scene and computer software for morphological analysis brings a constant result.
Notes	Research Notes
Genre	Journal article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AA12117871-20110900-0121

Research Notes

On a New Method for Content Analysis of TV News: The 2010 House of Councillors Election as an Example

Takeshi Kohno

Abstract

The purpose of this paper is to introduce a new method for content analysis of TV news using TV coverage of the 22nd House of Councillors election as a case. Instead of human power, I make the most of computer power to analyze the content of TV news. I will show that the combination of TV news archive that has begun to appear in Japanese business scene and computer software for morphological analysis brings a constant result.

I. Introduction

The 22nd House of Councillors election held on July 11, 2010 dealt a significant defeat to the Democratic Party of Japan (DPJ), which had only just wrested the government away from the Liberal Democratic Party (LDP) the previous year in the 45th House of Representatives election. Certainly there was much working against the DPJ. Immediately prior to the election, the Prime Minister talked about hiking the consumption tax, there were disputes within the party itself over politics and money, and there was also waffling on the relocation of the Futenma Base in Okinawa. The results were a vivid re-creation of a classic historical pattern last seen in the 15th House of Councillors election of 1989 when the ruling party was roundly defeated in an election fought over tax increases and political ethics.

This paper examines the coverage of the election by television. Even today, when a wide range of information is available on the Internet, television remains one of the largest sources of election information for most voters.[1] So it must be still very important for us to know what kind of information about election we can get from mass media. Scientific research, content analysis, is needed to study the nature of information provided by the mass

media. It has been fraught with difficulties in the past. This paper departs from conventional content analysis of television news coverage that relies primarily on human power. Our analysis uses computer power. We make use of television news database that appears newly in Japan and a computer software for morphological analysis, and the paper describes the procedures used to do so. In addition, we discuss some of the findings about the coverage on the election by television with these new methods.

II.   Development of New Methods for Content Analysis

In the past, content analysis was a time-consuming and expensive endeavor requiring the researcher to prepare a detailed coding manual and train large numbers of coders in its application. When the research based on content analysis begun to be conducted in Japan in the early 1960s,   Tsuneo Kano wrote at the beginning of his content analysis of newspaper articles on the 1960 security pact that the method "produced little reward for the amount of effort involved" (Kano 1961, p. 16).

Television content analysis requires even more time and effort than newspapers. Newspapers preserve reduced-size versions and the researcher can keep the research material on hand, so there is little need to worry about availability of data. Reduced-size versions also enable the analysis period to be determined after the fact. For television, however, the researcher is basically forced to establish a plan in advance of the event and record broadcasts himself or there is no material to analyze. One of the biggest hurdles to the completion of a television content analysis is whether the news programs are all recorded. It is not unheard of for there to be recording mistakes because of sudden changes in program content or broadcast hours. Another factor is simply failure to record because the researcher forgot to change the tape. The spread of hard disk recorders today has eliminated much of the concern with inability to record due to lack of storage capacity, but there is still potential for the hard disk itself to be damaged, and with it the recordings. Power outages and recording agreement failures can also be fatal.

Taking it a step further, with a newspaper it is clear where an article starts and stops, but with television, determining the beginning and end of a news story can be a challenge. No one has difficulty distinguishing between a feature on the legislature and the weather report. However, as one example, a story on party leaders' initial statements may run for nearly 20 minutes on the first day of the official campaign, and some coders will consider it to be a single chunk of news, while others will divided it up into separate news stories for each party. Strictly speaking, how news stories are divided up needs a reliability check for coders.

Today, however, the technical environment for content analysis is dramatically different than in the past. The first advances in content analysis were made for newspapers and other print media. In Japan, the turning point for newspaper content analysis was the creation of databases of newspaper articles[2] by the Nikkei and other publishers in the latter

half of the 1970s. This was further aided by the development of KHCoder and other software dedicated to text-mining and morphological analysis that enabled researchers to analyze the content of digitized text from their personal computers.[3] Today, at least for text in the form of digitized newspaper articles, most research is performed on personal computers as "text-mining" style quantitative analysis using materials extracted from databases rather than by human coding,[4] and many significant results have been built up this way.[5]

The technological advances described above have brought great benefits to the content analysis of print media, but it is television news and other forms of broadcast media that presumably have the most impact on voting behavior, and in recent years Japan has begun to use computers in this way to perform content analysis of broadcast media as well. Basically, however, the methodology follows tradition: the researcher records either the video or audio of the materials analyzed on his own and then uses a large number of workers to perform the coding at great time and labor. To this point there has been no database of television news, and given its nature—the fact that it was television and the impact of video—there has been no thought given to preserving it for analysis in the form of text.

However, in recent years commercial databases of Japanese television news have begun to appear. Examples include those run by JCC Corp. (http://www.jcc.co.jp/index.html) and Project Inc. (http://www.project-inc.co.jp/). Similar to newspaper databases, news can be searched with keywords, and hits can not only be viewed, but basic information on the broadcasting station, broadcasting date and broadcasting time is available, together with text on the content (title and summary). Commercial services are designed for corporate customers. News coverage is usually "sent and forgotten," but these services offer corporations basic information on which stations have covered them and in what ways.[6] For companies, how they are covered on television is important information that feeds into future policies and guidelines.

Regardless of the purpose for which they were created, databases of television news can also be used in the academic world to perform more efficient content analysis of television news. Prior research in this area has already been done in Japan. Examples include Inamasu and Ikeda (2009), Kohno (2010), and Inamasu (2010). However, neither Inamasu nor Kohno comments in detail on the television news databases described above. The biggest merits to computer coding are labor savings, higher degrees of reliability and shorter time spent on analysis. In other words, it is able to overcome the biggest problems with human coding. On the other hand, the question of applicability is not automatically solved just because computers are used in the analysis. Validity—whether the subject to be analyzed is analyzed correctly—is always an issue regardless of whether computers or humans are involved. It is an issue that the researcher himself must confront.

The problem of applicability in databases of television news arises primarily from the following circumstances. Databases of newspapers and similar media provide the entire piece while commercial television news databases at the current point in time only provide summaries. This is the equivalent of providing the title and lead of a newspaper article. What

is more, the title and lead are not from the media itself, but provided by a third party. In short, the question is how faithful the summary re-creates the reality of the coverage.[7] There is also the need to verify how individual news stories are divided up. Nonetheless, inasmuch as these products are offered on a commercial basis, the reliability is probably higher than that achieved with student coders.

This study uses the database provided by JCC Corp. (referred to as "JCC" for the remainder of this study)[8] to perform a basic analysis comparing different stations' volume of coverage of the 22nd ordinary House of Councillors election, which was held on July 11, 2010.

III. The Information Provided by the Television News Database

According to the JCC web site (http://www.jcc.co.jp/database/tvdb.html), JCC "comprehensively records television news, information and lifestyle programming, summarizing each article in approximately 400 characters and affixing time codes that are accurate to the second to create a real time database." Users "can search for articles from a special screen provided on the web using keywords, for example, the name of an incident." This is basically the same service that has been provided by the Vanderbilt Television News Archive (referred to as "VTNA" for the remainder of this study) in the United States since 1968 (Uozumi 1999). The VTNA is accessible over the Internet free of charge. It has been used in the United States to perform a large amount of content analysis (for example, Iyenger and Simon 1993). Japan does not, unfortunately, have a free television news archive like VTNA available, and we are forced to pay a considerable amount to make use of commercial databases like JCC.

The author used the JCC database to create the following data set in Excel. This data set organizes data in a manner suitable for the analysis.

Table 1: Data Set Structure

A	B	C	D	E	F	G	H	I	J	K	L
ID	Station	Year	Month	Day	Start	End	Length	Program	Title	Theme	Summary

For the information provided by JCC, there is no doubt as to the accuracy of the eight items from Station (Column B) to Program (Column I). However, the title, theme and summary were recorded by JCC coders. While they may be professionals, and while the summary of up to 400 characters provides a greater degree of detail than the VTNA index, there is probably room to verify how faithfully they reflect the content of the broadcast. There is also the problem of being unable to analyze the video, which is the chief characteristic of television.

Nonetheless, there are two points to make on the lack of video analysis and why it may be reasonable. The first is that there is no established methodology for appropriately and comprehensively analyzing video. Even if the video is watched, the coder must still code it as "favorable/unfavorable," which means relying mainly on information from the audio portion. The second is that domestic political news usually does not involve graphic video footage that plays on people's' emotions. Indeed, there is basically no broadcasting of shocking video or even a video that concentrates exclusively on specific candidates because of the need to ensure the fairness of coverage during the official election period.

IV.   Methodology

The period for which television news coverage of the 22nd House of Councillors election was analyzed was a total of 33 days, extending from June 8, 2010, the date on which the Naoto Kan Cabinet replaced the Yukio Hatoyama Cabinet, to July 10, the date immediately prior to election day. The analysis covered the key television stations in Tokyo: NHK, NTV, TBS, Fuji Television Network, TV Asahi and TV Tokyo. We examined all coverage of the House of Councillors election broadcast by the stations.

The keyword used to search the database was the single character "*sen(選)*" (first character in the word "*senkyo*," or "election"), which identified all coverage of the House of Councillors election. Obviously, by using a single character as our keyword rather than longer phrases such as "ordinary House of Councillors election," "House of Councillors election," or "*senkyo*" (full form of the word "election"), we also identified coverage that had nothing to do with the House of Councillors election, for example coverage of the South Africa World Cup soccer qualifying matches(予選) held during the same period, which also use the character "*sen*." These false positives were removed from the data set during the post-extraction confirmation process because they were irrelevant to the House of Councillors election.

V.   Findings

1.   Quantitative Analysis

The analysis period covered is from June 8 to July 10, 2010, and Table 2 contains the broadcast hours and numbers for coverage of the 22nd House of Councillors election by individual television stations. Obviously, the House of Councillors election may not have been the main story for the entire time or all of the broadcasts. Some of the stories were about other topics and merely mentioned the House of Councillors election. However, if the focus of the analysis is to compare different stations, stories merely commenting on election are presumably proportional and there would be little difference in the results were they removed, so this analysis did not account for such differences.

Table 2: Times and Numbers

		NHK	NTV	TBS	Fuji	Asahi	Tokyo	Total
Prior to official campaign (16th)	Time	30046	30566	29758	23505	13023	3698	130596
	Number	187	184	181	128	76	45	801
During official campaign (17th)	Time	70011	28668	37212	28048	42102	7462	213503
	Number	306	187	261	162	225	66	1207
Total (33 days)	Time	100057	59234	66970	51553	55125	11160	344099
	Number	493	371	442	290	301	111	2008

∗   All of the figures in the time column are in second units.

As can be seen from Table 2, NHK devoted the most time to coverage of the House of Councillors election, with a total of 100,057 seconds (approximately 27.8 hours). It was followed by TBS at 66,970 seconds (18.6 hours), NTV at 59,234 seconds (16.5 hours) TV Asahi at 55,125 seconds (15.3 hours), Fuji at 51,553 seconds (14.3 hours) and providing the least amount of coverage, TV Tokyo at 11,160 seconds (3.1 hours). This confirms that NHK was fully cognizant of its role as a public broadcaster in its coverage of the election.

During the campaign period, average coverage per day was 69, 36, 28, 41, 27 and 11 minutes respectively (Table 3). Also, TV Asahi is notable for devoting considerably more time during the campaign period to coverage of the House of Councillors election than prior to it.

Average broadcast hours per day during the campaign period can be compared against the coverage of the 45th House of Representatives election, which was held on August 30 of the previous year.[9] For the House of Representatives election, the results were 54 minutes for NHK, 36 for TBS, 35 for NTV, 41 for TV Asahi, 23 for Fuji and 11 for TV Tokyo. TBS, Fuji, TV Asahi and TV Tokyo showed very little difference, while broadcast hours increased for NHK and decreased for NTV. We used PASW Statistics to test for equivalence ($\chi 2$ test) and found exact significance probability (both sides) to be 0.768, which indicates absolutely no statistically significant difference. While many today argue that the House of Councillors is unnecessary because the ruling party does not have a majority in it, our analysis indicates that the election was treated the same regardless, at least with respect to the quantity of television coverage.

We next show the result of analyzing the JCC coverage summaries with a text-mining software.

Table 3: Average Broadcast Hours Per Day during the Campaign Period

	NHK	NTV	TBS	Fuji	Asahi	Tokyo
House of Councillors election	69	28	36	27	41	11
House of Representatives election	54	35	36	23	41	11

∗   All figures in the table are in minute units.

2.   Qualitative Analysis

The software that we used for text-mining in our analysis was the commercially available "IBM SPSS Text Analytics for Surveys." While there is excellent free software available like KHCoder, we chose to use IBM SPSS Text Analytics for Surveys (referred to as "ISTAS" for the remainder of this study) because it provided better coordination with the Excel data set. The software was developed to analyze free answer data from questionnaire surveys, but is able to analyze text of approximately 5,000 characters length, which made it suitable for analysis of the JCC news summaries, which are a maximum of 400 characters.

While the software enables many different forms of analysis, this study focused on the most basic, a comparison of stations based on extracted vocabulary. In the earlier House of Councillors election, the tide had turned against the DPJ and it was expected to be roundly defeated. The main factors for this were comments on raising the consumption tax, the Ozawa problem, and the question of relocating Futenma Base. To illustrate the differences among stations, we focused on the Hatoyama/Ozawa problem, in other words, the extent to which stations covered the question of "politics and money." Comparing the differences among stations during that election will indicate what stations focused on in their coverage of the most recent House of Councillors election, and what their basic position is with regards to the DPJ. However, it should be emphasized that this is not an analysis of the actual coverage, but of summaries of the news created by professionals at a private company. The conclusions that we reach should therefore be viewed only as probabilities and potentials.

The first morphological analysis using ISTAS extracted keywords in a variety of morphemes (the minimum unit of meaning in a language: nouns, verbs, adjectives, adverbs, particles) from a total of 2,008 news summaries. Within this, we extracted a number of morphemes containing the word "Ozawa," for example "Ozawa," "Ichiro Ozawa," "Ichiro Ozawa Office," "dump Ozawa" and "DPJ Ozawa Group" (39 in total). We therefore edited the extraction dictionary and redid the extraction so that everything was identified in the

form of "Ozawa," including the morphemes "money" and "DPJ Iwate prefectural federation." The results identified 301 news summaries containing the morpheme "Ozawa."

Table 4 illustrates the differences among stations in the number of news stories commenting on "Ozawa." The equivalence test ($\chi 2$ test) found an exact significance probability (both sides) of 0.1% for the data in Table 4, indicating that it satisfied the significance level. We therefore compared this against the total number of news stories to determine from the adjusted residual the degree to which news mentioned "Ozawa." The adjusted residual was positive for NTV, TBS and Fuji at 3.0, 0.9 and 1.0 respectively. It was negative for NHK, TV Asahi and TV Tokyo at -2.9, -0.3 and -2.6 respectively. The first three stations, led by NTV, had an excess of coverage commenting on "Ozawa" compared to their total coverage, while the second three, led by NHK, under-covered the story. This can be assumed to be the result of differences in attitudes of each station towards the DPJ.

Table 4: Comments on "Ozawa"

	NHK	NTV	TBS	Fuji	TV Asahi	TV Tokyo
Ozawa	51	78	73	50	43	6
Adjusted residual	-2.9	3.0	0.9	1.0	-0.3	-2.6
Total	493	371	442	290	301	111
Adjusted residual	2.9	-3.0	-0.9	-1.0	0.3	2.6

VI. Conclusions

This paper summarizes the findings of a content analysis of television news coverage of the 22nd House of Councillors election, held in July 2010. It compares stations on broadcast hours to examine the volume of information on the election provided to viewers and finds that NHK was, unsurprisingly, the highest, followed by TBS, NTV, TV Asahi, Fuji and TV Tokyo in that order. The paper also compares the average volume of information provided per day against the House of Representatives election of August 2009 and found no statistically significant difference between the two, an indication that television stations quantitatively treat all national elections the same.

In addition, the paper makes use of a commercial database of television news provided by a private company rather than the traditional analysis relying on human coders, and it attempts a new method of content analysis using text-mining software to examine summaries of news stories. This method indicates the possibility of reducing the labor

required for content analysis of television news, studies that researchers are often hesitant to undertake because of the significant amounts of labor involved.

The analysis contained in this paper is tentative at best. The directions employed could be made more fruitful by text-mining full transcriptions of the audio portions of news broadcasts, rather than just summaries. Currently, this would need to rely on human input, but voice recognition software[10] has already achieved extremely high levels of precision and processing capacity, and in the future it may be possible to use it to eliminate the need for human input. At some point, there may be databases of the complete text of television news audio, rather than just summaries. In addition, the development of software capable of analyzing the video portions would add new dimensions to computer-based content analysis of television news alongside analysis of the audio portion.

Notes

1. For example, the Asahi Shimbun used the interviewing method to conduct a survey of voters nationwide after the 22nd House of Councillors election (July 28 to August 9, 2010) (responses: 1,125) and found that 54% and 41.8% of voters listed newspapers and television respectively as their first and second sources of information (multiple response question).
http://adv.asahi.com/modules/box/index.php/content0093.html (February 28, 2011).
2. The following studies may be of reference regarding the structure of newspaper article databases. Ishii (1994).
3. KHCoder is free software for the "quantitative analysis of Japanese text data" developed by Koichi Higuchi of Ritsumeikan Univeristy. For details, see the KHCoder web site run by Mr. Higuchi. (http://khc.sourceforge.net/) It should be noted that there has been a great deal of computer-based content analysis performed in the United States ever since the 1960s and the development of "General Inquirer" by Philip J. Stone (Stone, Dunphy, Smith, and Ogilivie,1966). For more information on computer-based content analysis in the United States, see Neuendorf (2002) and Krippendorff and Bock (2009).
4. See for example Hosogai (2009). The Hosogai paper describes the analytical process in detail.
5. Mr. Higuchi, the developer of KHCoder, lists research using the software on the KHCoder web site. (http://khc.sourceforge.net/bib.html)
6. However, the ideals are higher than for a company like JCC that provides this service. For example, JCC describes its activity concepts thus: "We have promoted information that was traditionally lost over time to an item of true value in social activities by adding scientifically-based analysis and monitoring of the giant media of television and the Internet, thereby increasing the level of intelligence. "
http://www.jcc.co.jp/concept.html
7. These kinds of concerns are also noted in the following research from the United States, one of the most advanced countries in television news archives. Fan (1988, p.39).
8. No particular significance should be read into the author's choice to use JCC rather than other commercial services. It was merely the service that could be used from the author's research

environment and in no way constitutes a recommendation to use the services of a specific company.

9. Coverage of the House of Representatives election was extracted from the JCC database in the same manner as data on the House of Councillors election, which is the main focus of the analysis in this paper. Some of the findings were reported at the GCOE-CGCS International Symposium held in March 2010 at Keio University (Kohno 2010).

10. Software that enables computers to recognize and analyze words spoken by human beings and convert them into digitalized text data. For example, NEC has developed a voice-recognition system on a commercial basis. See the following web site for an outline of the system and its development history. http://www.nec.co.jp/vsol/index.html

References

Fan, David P. (1988). *Predictions of Public Opinion from the Mass Media: Computer Content Analysis and Mathematical Modeling*. New York: Greenwood Press.

Hosogai, Ryo (2009). "Seijigaku ni okeru tekisuto bunseki no tekiyo" (Application of Text Analysis in Political Science), *The Waseda Study of Politics and Public Law*, No. 91.

Inamasu, Kazunori and Kenichi Ikeda (2009). "Tayoka suru terebi hodo to, yukensha no senkyo e no kanshin oyobi seiji e no kanyo to no kanren: Senkyo hodo no naiyo bunseki to daikibo shakai chosa no yugo o tosite" (The Relationships between Exposure to News Shows and Interest in the 2007 Election and Involvement in Politics: Combining the Quantitative Content Analysis of Election Coverage with Social Research Data), *Social Psychology Studies*, Vol. 25, No. 1.

Inamasu, Kazunori (2010). "Tekisuto deta bunseki o mochiita 2009 nen shuinsen hodo no burui" (Categorization of the 2009 House of Representatives Election Coverage using Text Data Analysis), *Proceedings of the FY2010 Meetings of the Japan Association of Electoral Studies* (Social Psychology and Media Interest Group), May 15, 2010, Meiji University.

Ishii, Masahiro (1994). "Nihon keizai shimbun kiji detabesu no hinshitsu kanri" (Quality Assurance in the Nikkei Shimbun Article Database), *The Journal of Information Science and Technology Association*, Vol. 44 No. 4.

Iyemgar, Shanto & Simon, Adam (1993). "News Coverage of the Gulf Crisis and Public Opinion: A Study of Agenda Setting, Priming, and Framing," *Communication Research*, 20, 365-383.

Kano, Tsuneo (1961). "Anpo hantai toso kiji no naiyo bunseki" (Content Analysis of Articles on the Security Pact Dispute), *Newspaper Studies*, No. 119.

Kohno, Takeshi (2010). "A New Method of Content Analysis of TV News," Session11 Information and Governance in Civil Society, at International Symposium on Designing Governance for Civil Society (March 6, 2010, at Keio University).

Krippendorff, Klaus & Bock, Mary Angela (2009). *The Content Analysis Reader*. Thousand Oaks: Sage Publications.

Neuendorf, Kimberly A. (2002). *The Content Analysis Guidebook*. Thousand Oaks: Sage Publications.

Stone, Philip J., Dunphy, Dexter C., Smith, Marshall S., & Ogilvie, Daniel M.(eds.). (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge: MIT Press.

Uozumi, Shinji (1999). "Beikoku no terebi nyusu shiryo akaibuzu" (Television News Archives in the United States), *The Journal of Information Science and Technology Association,* Vol. 49 No. 3.