Investigating macroscopic transitions in Japanese foreign policy using quantitative text analysis[†]

Takafumi Suzuki[*]

Faculty of Sociology, Tokyo University, Tokyo, Japan
*E-mail: t_suzuki@nii.ac.jp

Abstract

This study introduces new quantitative text analysis methods into foreign policy analysis. Quantitative text analysis in the social sciences is currently aimed in two directions, namely (a) more systematic analysis using larger amounts of data sets and (b) more detailed analysis using linguistic knowledge. Our methods, by using recent techniques in natural language processing, integrate these two different trends, and achieve more systematic but detailed analysis. We apply our methods to 147 Diet speeches of Japanese prime ministers, and shed new light on the character of Japanese foreign policy. This study makes a methodological contribution to foreign policy analysis and a substantial contribution to the study of Japanese foreign policy.

[†]   Earlier versions of this study were presented at the 2008th Annual Meeting of the Japan Association of International Relations, and the ISA's 50th Annual Convention.

1 Introduction

Text is arguably the most pervasive – and certainly the most persistent – artifact of political behavior (Monroe and Schrodt, 2008). With the recent advances in natural language processing and the production of large corpora, textual data have become one of the most important resources for political researches. In the field of international relations, including foreign policy analysis, it has attracted more attention among scholars, as it sheds new light on the perceptions and interests of governments and political leaders (Hudson, 2005).

Methodologically, quantitative text analysis in the social sciences is currently aimed in two directions: (a) more systematic analysis using larger amounts of data sets and (b) more detailed analysis using linguistic knowledge. Regarding (a), some studies try to develop systematic and automatic analysis of political speeches. For example, Laver et al. (2003) try to develop methods for automatically detecting policy positions from political speeches. They succeeded in carrying out one of the most systematic analyses of political texts, but their method was basically based on counting the frequencies of content words in the texts, and did not take into account the linguistic information that the texts have. These kinds of analyses cannot shed light on the detailed changes in governments' or political leaders' perceptions and interests.

Regarding (b), some studies try to use linguistic knowledge for analyzing political speeches in order to compensate for the deficiency in (a). For example, Walker and Schafer (2007) examined the changes in Carter's perceptions using operational code analysis, and Duffy et al. (1998) investigated the process of INF negotiations using a theory of pragmatics. These studies have successfully revealed more detailed characteristics of governments' and political leaders' perceptions and interests by introducing linguistic knowledge, but for that reason, their analyses of texts had to be carried out manually, and thus were applied to very small amounts of data. As these studies have shown, there was a practical trade-off between (a) and (b) (Monroe and Schrodt, 2008). If we can develop methods integrating (a) and (b), we will be able to obtain more fruitful findings from political texts.

Against this background, this study, by using recent techniques in natural language processing, introduces new quantitative text analysis methods into foreign policy analysis. By using multivariate analysis and

machine learning with larger amounts of textual data, and introducing linguistic knowledge for analysis, our methods integrate two different trends in quantitative text analysis in the social sciences, and achieve more systematic but detailed analysis. Applying them to Japanese prime ministers' Diet addresses, we investigate macroscopic transitions in Japanese foreign policy after World War II, and thus provide new empirical findings on the character of Japanese foreign policy.[1] This study makes a methodological contribution to foreign policy analysis as well as a substantive contribution to the study of Japanese foreign policy.

2 Character of Japanese foreign policy

According to Green (2001), conventional views of Japanese foreign policy have emphasized its coping (Blaker, 1993), reactive (Calder, 1988) and immobile (Stockwin *et al.*, 1988) character. Japan has been said to be a state unwilling to take risks or assert its interests in the world (Lincoln, 1992). These views have also been embedded into descriptions of Japanese diplomatic history which have focused on security issues and the US–Japan relations and have thus necessarily highlighted how Japan responds to the United States' actions (Inoue, 2005).

Since the end of the Cold War, however, many scholars have proposed alternative or more nuanced views on Japanese foreign policy (Potter and Sudo, 2003) by investigating its history, from a long-term perspective (LaFeber, 1998; Pyle, 2007), throwing light on cases focusing on its more strategic or proactive aspects (Green, 2001; Midford, 2002; Hughes, 2004), investigating economic issues (Amyx, 2002; Ono, 2002; Pekkanen *et al.*, 2007) and aid issues (Miyashita, 1999; Fukushima, 2000) in addition to security issues, and taking into account ideational or symbolic factors (Katzenstein, 1996; Tanaka, 2000; Lind, 2004; Oros, 2008). Some historians have also started to describe the diplomatic history of Japan free from the bipolar structure of the Cold War, with the aim of shedding light on the multidimensional character of Japanese foreign policy (Irie, 1997; Inoue, 2005). These studies have clarified the importance of examining various issues from a broader perspective when

---

1 Most previous studies focused only on English texts, but a wide variety of texts in other languages are available now. Thus, it is important to apply our methods for the texts in other languages, though our methods can be used for English texts as well.

investigating a government's foreign policy, and have prompted us to re-examine the character of Japanese foreign policy.

From a methodological viewpoint, most of these studies still rely on the investigation of cases or documentary sources, while some studies have attempted to investigate the Japanese government's perception of the world by analyzing prime ministers' speeches (Watanabe, 1974; Tanaka, 1994; Suzuki, 2009). These studies emphasized systematic analysis, but in fact counted the frequencies of specific content words or concepts, thus there is room for improvement in their analyses.

3 Research agendas

This study investigates (a) macroscopic transitions in the Japanese government's regional interests after World War II, (b-1) changes in its level of interest in particular aspects of the US–Japan relations, and (b-2) changes in its rhetoric regarding the US–Japan relations. In our analysis of the US–Japan relations, we focus especially on whether or how it changed before/after the end of the Cold War, because heated discussions have been conducted on this issue (Sudo, 2007).

Before explaining each research agenda, we explain our uses of the terms in this paper. First, Japanese government's 'perception' of the exterior world represents how it sees other countries and regions. It does not equal to the political decisions themselves, but it is one of the important bases for them. Perception is regarded as consisting of two elements; interest and context of interest (Watanabe, 1974; Suzuki, 2009); 'interest' simply represents whether and how much the Japanese governments see it as important, and the context of interest represents how the government sees it. In addition to these concepts, we introduce the concept of rhetoric; it also characterizes the governments' interest, and represents how the governments see it, as well as the context. The differences between the context and rhetoric is that the context represents a rather direct characterization of the interest,[2] where rhetoric include a more nuanced and indirect characterization of the interest. On the linguistic level, the context is measured by using the simple co-occurrences of a

---

2   For example, Suzuki (2009) measured the context of interest by creating the indicators of conflict or dialogue based on words directly representing these meanings such as 'Tairitsu' (conflict) and 'Taiwa' (dialogue).

word, but our concept rhetoric is measured by using a modification structure. We will explain the details of our method in the Method section.

3.1 Transitions in regional interests

First, we investigate macroscopic transitions in the Japanese government's regional interests after World War II. A government's regional interests, i.e. the particular regions that a government is interested in, are a reflection of its perception of the exterior world, and give us important information on the character of its foreign policy (Watanabe, 1974; Suzuki, 2009). In addition, regional interest is one of the points of contention in the debates between supporters of the conventional 'reactive' Japanese foreign policy model that was based mainly on the US–Japan relations, and supporters of newer models that emphasize broader regional interests. In previous studies, the Japanese government's level of interest in particular regions has been thoroughly investigated by examining documentary sources, but the overall structure of its interests, taking into account the correlation between respective regional interests, has never been investigated. Our analysis, by using larger amounts of textual data, enables us to capture the overall picture of Japanese foreign policy after World War II. It can thus contribute to resolving conflicts between scholars as well as to providing empirical findings on the character of Japanese foreign policy.

3.2 Changes in the US–Japan relations

Secondly, we investigate changes in the Japanese government's level of interest in particular aspects of the US–Japan relations, and changes in its rhetoric.

3.2.1 Changes in the level of interest in particular aspects of the US–Japan relations. Previously, the US–Japan relations were discussed mainly in the context of security issues, and it was frequently noted that relations changed before/after the end of the Cold War (e.g. Green and Cronin, 1999). Recently, however, Ono (2002) has emphasized the importance of taking into account economic issues as well as security issues to shed light on more multidimentional character of Japanese

foreign policy and, by analyzing both of these issues, has argued that the US–Japan relations were understood to be stable before/after the end of the Cold War (Sudo, 2007). This study is one of the most comprehensive analyses of the US–Japan relations after World War II, and is based on two major theoretical viewpoints in the study of international relations, namely realist and liberalist, in the sense that he analyzed both security issues and economic issues, the issues these two theories, respectively, emphasized as an important factor in determining international relations. However, even this study does not investigate the cultural or symbolic issues, which are currently attracting more attention among scholars. As we mention in detail in Method section, this study includes three categories, namely the security issues, economic issues, and cultural/symbolic value issues, when analyzing the texts referring to the US–Japan relations, so we can take into account various viewpoints into account for a empirical analysis. In this way, we try to analyze the Japanese government's interest in particular aspects the US–Japan relations in a more comprehensive way.

3.2.2 Changes in rhetoric. Previous studies have examined the substantial changes in the Japanese government's interest in the US–Japan relations by examining military expenditure or levels of trade, but our methods using textual data have the advantage of allowing us to analyze the symbolic changes in the US–Japan relations, namely how the Japanese governments perceive the US–Japan relations, that enables us to analyze, for example, whether the US–Japan relations were exclusive foreign relations from others or paralleled with others, and whether it showed obedience or equivalence to the United States. In order to derive knowledge on these points, we examine the Japanese government's changes in rhetoric as well as its interest in the US–Japan relations by using linguistic knowledge in our analysis. This enables us to better understand the symbolic changes in the Japanese government's perception of the US–Japan relations.

By investigating these points, we show that our methods provide useful empirical findings for understanding the character of Japanese foreign policy after World War II, and in this way, this study makes a methodological contribution to foreign policy analysis as well as a substantial contribution to the study of Japanese foreign policy.

4 Data

4.1 Overview

The data[3] we used in this study consists of 147 Diet addresses[4] given by 25 Japanese prime ministers[5] from 1948 to 2008.[6] Diet addresses are based on Article 72 of the Constitution: 'The prime minister, representing the Cabinet, submits bills, reports on general national affairs and foreign relations to the National Assembly and exercises control and supervision over various administrative branches'. The address creation process is usually as follows: first, the bureaucrats from each ministry prepare drafts of parts of the speeches in their charge that they would like to include in the addresses, and second, the bureaucrats in the Cabinet Office merge and revise them to make drafts of the addresses, and finally, the prime ministers themselves revise and make a final versions of them (Shinoda, 1994). Because of this creation process, the addresses are the results of policy competitions between the ministries (Shinoda, 1994), and thus, the content of the addresses can be interpreted as reflecting the government's interest, although they can also reflect the prime ministers' individual interest (Suzuki, 2009).[7] Also, they have been delivered by prime ministers under similar conditions and at similar intervals over a long period of time (Watanabe, 1974), and

---

3   Replication datasets are available on our website (http://researchmap.jp/stratovarius/).

4   Japanese prime ministers deliver two types of Diet addresses, namely Shisei Hoshin Enzetsu (Speech to Express Policy: SEP) and Syoshin Hyomei Enzetsu (Speech to Express Belief: SEB). The SEP is delivered at the ordinary Diet session, while the SEB is delivered at extraordinary Diet sessions. We included both of them for our analysis, because both include foreign interest, and it was pointed out that there are few differences between them in the level of basic textual characteristics (Suzuki and Kageura, 2007).

5   The addresses of Prime Minister Yoshida that were used in this study come from his second tenure, thus we used the notation YS2.

6   Based on Suzuki (2009), we classified the prime ministers into six standard divisions of eras in Japanese diplomatic history after World War II (T1: Yoshida (second tenure) – Kishi, T2: Ikeda and Sato, T3: Tanaka – Ohira, T4: Suzuki – Uno, T5: Kaifu – Koizumi, T6: Abe and Fukuda). The table on the names, the initials, the divisions of eras, the date of assumption into office, the number of addresses, and the total and mean number of tokens for each prime minister is available on our website (http://researchmap.jp/stratovarius/). See also Iokibe (2006).

7   It may include the strategic uses of references in speeches such as rip services to other countries, but in this paper we did not distinguish them from others, because the fact that it had to use rip services is also a sign of showing its interests, and it is thus valuable to analyze. In the future, we will investigate these points in more detail.

provide a clear indication of the Japanese government's long-term interest of the world.[8] We downloaded the addresses[9] from the on-line database *Sekai to Nihon (The World and Japan)*.[10] We applied morphological analysis to the addresses using ChaSen, a Japanese morphological analysis system (Matsumoto *et al.*, 2003). As Japanese sentences do not leave a space between morphemes (words), we need to apply it to them for extracting every morpheme. As well, we applied dependency analysis to the addresses using CaboCha (Kudo and Matsumoto, 2002), a Japanese dependency analysis system. It enables us to extract syntactic relations between phrases, and examine what phrases modify what other phrases.

We extracted nouns according to part-of-speech tags assigned by ChaSen, and calculated the relative frequencies of (compound) nouns to the total number of nouns, which should provide us better interpretation of results than calculating the relative frequencies to the total number of tokens, i.e. the number of running morphemes appearing in the text (Suzuki and Kageura, 2007).[11]

4.2 Regional/country names

In order to examine regional interest, we extracted all the regional/country names[12] from the addresses. We divided the extracted regional/country names into larger categories, i.e. the United States, Soviet Union + Eastern Europe + Central Asia, Asia (general), East Asia, South-East Asia, South Asia, Middle East, Oceania + Pacific, Africa, Europe + Canada, and Latin America,[13] and calculated the value of

---

8 There are a lot of data for examining the policy implementation, such as budget data, but there is very little data available for investigating the policy 'intention' of governments, that can be analyzed from public speeches.

9 The texts of the speeches are written down by secretaries in shorthand.

10 www.ioc.u-tokyo.ac.jp/worldjpn/index.html

11 We refer to this as 'relative frequency' below.

12 We excluded city names and the names of parts of the countries because our purpose was to investigate macroscopic transitions in Japanese foreign policy, and these names tended to reflect rather occasional events (e.g. where the summit conferences were hold etc.). The list of the regional/country names is available on our website (http://researchmap.jp/stratovarius/).

13 Categories are basically based on the geographical unit. We regarded the United States as an independent region that appeared especially frequently in the addresses, and divided Asia into 'the whole of Asia' and 'each part of Asia', each of which represents different

each category by summing up the relative frequencies of all the categorized regional/country names.

4.3 Compound nouns of 'Nichi-bei' (the United States–Japan)

To examine the changes in interest in the US–Japan relations, we extracted the compound nouns that include the word 'Nichi-bei' (the United States–Japan).[14]

We classified these compound nouns into three categories, namely security issues, economic issues,[15] and cultural/symbolic value issues.[16] We classified 'Nichi-bei-Kyodo-Sengen' (the US–Japan joint communique) and 'Nichi-bei- Kyodo-Seimei' (teh US–Japan joint declaration) according to the issues mentioned in the actual communique or declaration.[17] We calculated the value of each category by summing up the relative frequencies of all the categorized compound nouns.

4.4 Modifier/modificand

To examine the changes in rhetoric regarding US–Japan relations, we extracted the modifier/modificand of the compound nouns including the

---

interests. We combined the Pacific with Oceania; it can also be combined with the United States or Asia, and thus we tried both of these combinations, but no major changes in the results were needed for that, and thus we show the results when we combined the Pacific with Oceania in this paper. We also combined Eastern Europe and Central Asia, for which these were few references, into the Soviet Union as part of the Communist bloc, and combined Canada with Europe as part of the capitalism bloc except US Morphemes that refer to more than one region, such as 'Bei–kan' (the United States–Korea) were counted in both categories (the United States and East Asia in this example).

14　Compound nouns make up important terms in Japanese (Nakagawa, 2000), and are suitable for analyzing the Japanese government's level of interest in particular aspects of the US–Japan relations. The list of the compound nouns is available on our website (http://researchmap.jp/stratovarius/).

15　We define these two issues as 'substantial value' issues.

16　We define terms representing 'cultural issues' as those representing scientific, academic, and cultural communication, and terms representing 'symbolic value issues' as those representing the US–Japan relations generally in an ambiguous manner. In fact, very few terms representing cultural issues appeared in the addresses, thus we combined cultural issues and symbolic value issues into one category.

17　'Nichi-bei-Kyodo-Sengen' appeared in the addresses given on 29 January 1958 (KN3), and 'Nichi-bei-Kyodo-Seimei' appeared in the addresses given on 1 February 1960 (KN9) and 3 August 1968 (SE12) were classified as both security issues and economic issues. 'Nichi-bei- Kyodo-Seimei' appeared in the address given on 17 July 1971 (SE19) and in the address given on 29 January 1972 (SE21); these were classified as security issues, economic issues, and cultural/symbolic value issues.

word 'Nichi-bei'.[18] We extracted compound nouns from the modifier/modificand, and pruned the prefixes and suffixes,[19] combined the notational differences[20] and omitted the functional nouns as stop words.[21] It is difficult to categorize rhetorical morphemes beforehand, thus we used an exploratory approach to analyze the rhetoric, which is explained in the next section.

5 Statistical methods

We applied statistical tests and principal component analysis, and we introduced new machine-learning methods in this study. In this section we explain these methods, especially the last one in detail.

5.1 Statistical test

The distribution of word frequency follows the power law distribution instead of the normal distribution, thus the Wilcoxon rank-sum test was used to test the significance between the two groups. We used it to analyze the Japanese government's level of interest in particular aspects of the US–Japan relations.

5.2 Principal component analysis

We made a text-feature matrix using the values of categories calculated on the basis of the relative frequencies of regional/country names,[22] and calculated the variance–covariance matrix. We carried out principal component analysis using this matrix. Calculating principal components from the regional categories, we examined the principal axes of Japanese foreign policy after World War II.

---

18   Strictly speaking, modifier and modificand have a different meaning, but we analyzed them simultaneously in the same experiment because both of them represent rhetoric regarding the US–Japan relations, and that gives a better performance in the experiment, a point that we will explain in the next section.

19   These are 'zenpan' (whole), 'sei' (quality), 'kankei' (relation), 'syo' (some), 'ima' (now), 'amari' (about), 'jou' (on), and 'chu' (during).

20   We united 'Yuko-Shinrai' (friendship–trust) and 'Shinrai-Yuko' (trust–friendship), and 'Patona' (partner), 'Patonasippu' (partnership), and 'Gurobaru-Patonasippu' (Global partnership).

21   These are 'igai' (except), 'tame' (for), 'mono' (thing), 'you' (-like), 'no' (of), 'kan' (inter-), 'moto' (under), 'gachi' (tend to), 'sa' (-; make nouns), 'sorezore' (every), and 'saki' (before).

22   Each row represents an address, and each column represents the value of the category.

5.3 Random forests

We made a text-feature matrix using the modifier/modificand of the compound nouns including the word 'Nichi-bei'.[23] We applied the random forests classifier (Breiman, 2001) using this matrix.

Random forests is a supervised machine-learning method. Machine learning is an approach that uses a large training data set x to tune the parameters of an adaptive model and predict the target vector t (Bishop, 2006), and has currently become the most important approach in text classification (Sebastiani, 2002), although it has rarely been applied in the field of political sciences.

Random forests is a improved method of bagging (Breiman, 1996). They are ensemble-learning methods. The basic idea of ensemble-learning methods is to improve the classification performance of previous (sometimes not so robust) statistical methods by repeatedly performing the experiments and calculating the mean or majority votes on the results. However, the results will always be the same when using exactly the same data to perform these experiments, thus ensemble-learning methods including bagging usually use bootstrap samplings from the original data to repeat the experiments. The main improvement in random forests from bagging is the extraction of a random subset from each bootstrapping sample, which enlarges the variances in bootstrapping samples, and this provided an improved performance over bagging (Breiman, 2001; Jin, 2007a,b).

We performed our method as follows.[24] We first replicated the text-feature matrix Mi,j 1,000 times with replacement, and extracted random subsets of $\sqrt{j}$ variables from each replicated data. We constructed an unpruned decision tree for each sample using the Gini index (Breiman, 2001; Bishop, 2006; Jin, 2007a,b). We constructed a new classifier by a majority vote of the set of trees. Two-thirds of the bootstrap samples were used for constructing the model and the other third were left out for testing the model.[25]

---

23 Each row represents an address, and each column represents the relative frequency of the modifier/modificand of the compound nouns including the word 'Nichi-bei'.

24 Regarding the details and other applications on random forests, see Breiman (2001), Jin (2007a,b), Suzuki (2009), and Wu and Kumar (2009).

25 It is called an out-of-bag test, and is used in random forests as a default test method.

We evaluated the classification performance using the precision (P_i), recall rates (RR_i), and F_i values. P_i measures the proportion of the texts that are relevant in the set of all texts classified, and RR_i measures the proportion of the texts that are correctly identified as such among all relevant texts. F_i indicates the summary statistics of P_i and RR_i (Tokunaga, 1999; Jin and Murakami, 2006).

We calculated the variable importance using the following formula (Breiman, 2001):

$$VI_{acu} = \frac{mean(C_{oob} - C_{per})}{s.e.},$$

where C_{oob} is number of votes cast for the correct class in the out-of-bag data, C_{per} number of votes cast for the correct class when m variables are randomly permuted in the out-of-bag data, and *s.e.* is standard error.

The mean value of subtractions for all trees formulated above represents the variable importance for a permuted variable. It represents the degree to which a class loses its specific character when a type of word, i.e. a different morpheme appearing in the texts, changes into another type of word.

A study reported that it can perform the best for text classification in Japanese (Jin and Murakami, 2006). Moreover, this method has more advantages for our task, i.e. it calculates important variables contributing to the classification in the experiment. By calculating the variable importance, and selecting the variables with high VI_{acu}, we extracted in an exploratory manner the modifier/modificand that distinguish the addresses before/after the end of the Cold War, and examined the changes in the rhetoric regarding the US–Japan relations before/after that point.

6 Results

6.1 Transitions in regional interests

We carried out principal component analysis using the variance–covariance matrix constructed on the basis of regional categories. The proportion of variance accounted for by the first principal component was 30.14%, while the cumulative proportion of variances accounted for by the first two principal components was 49.27%, and that by the first three components was 65.49%. We thus focused on the results for the first three components.

Three-dimensional plots constructed using the first three principal components are shown in Figure 1. The principal component scores of addresses are plotted according to era. The results indicate that the plots of addresses shifted from the right-hand side to the left-hand side (PC1), from foreground to background (PC2), from top to bottom (PC3), along with the changes in era. The results imply that macroscopic transitions in the Japanese government's regional interests can be seen by examining these three principal components.

To examine them, we next focused on the eigenvectors of regional categories. Table 1 describes the eigenvectors of categories for the first, second, and third principal components. The results indicate that the first principal component represents the axis of the United States–Europe in Japanese foreign policy, as the United States has a significantly high value, and Europe + Canada follows this, while the Middle East and Africa have lower values. The results indicate that the second principal component represents the axis of Asia in Japanese foreign policy, as Asia
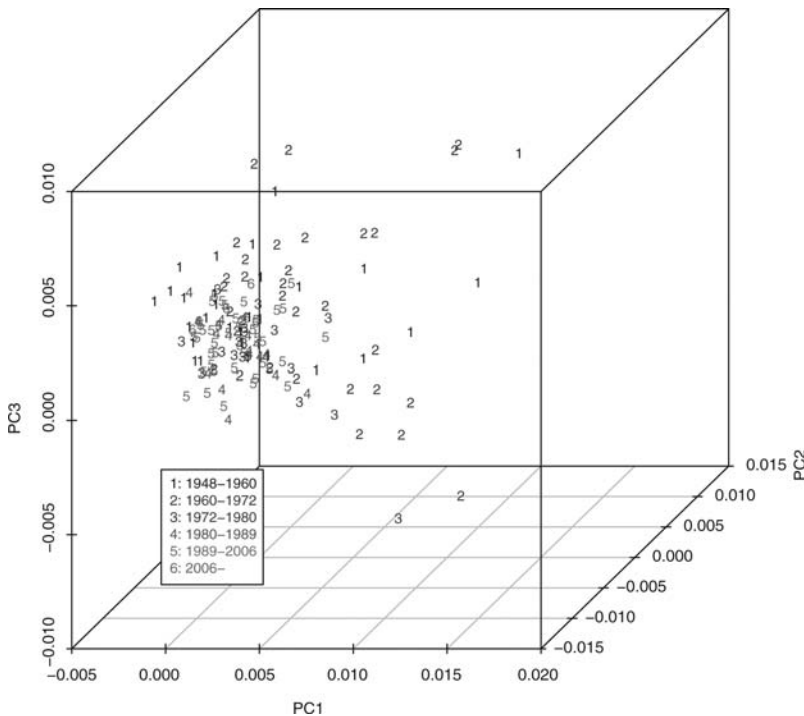
Figure 1 Three-dimensional plots constructed using the first three components (the principal component scores of addresses are plotted according to era).

Table 1 Eigenvalues of categories for the first, second and third principal components

Category	PC1	PC2	PC3
United States	0.744	0.135	−0.533
Soviet, Russia, Eastern Europe, Central Asia	0.015	0.069	−0.057
Asia (in general)	0.186	0.766	0.369
East Asia	−0.009	0.203	−0.361
South-East Asia	0.285	0.230	0.326
South Asia	0.009	0.019	−0.007
Middle East	−0.046	0.009	−0.446
Pacific	−0.002	0.075	−0.063
Europe, Canada	0.555	−0.509	0.357
Africa	−0.007	−0.017	−0.012
Latin America	0.143	−0.177	0.101

has the highest value, and South-East Asia and East Asia follow this, while Latin America and Europe + Canada have lower values.

The results indicate that the third principal component represents the axis of crisis handling including the Middle East situation, the China situation, and the North Korea situation, as the United States, Middle East and East Asia have higher absolute values.[26] We conclude that Japanese foreign policy after World War II has two important axes, namely United States–Europe and Asia, and in addition to these two important axes, the axis of crisis handling sometimes has an important role.

To examine the transitions in these important axes, we next examined the transitions in principal component scores, especially focusing on the changes before/after the end of the Cold War. Figure 2 represents the transitions in first, second, and third principal component scores. Dotted polygonal lines represent the principal component score of addresses and vertical lines represent the end of the Cold War.[27] We fitted linear model

26 It may seem that the results indicate that the third principal component represents the axis of the United States because the United States has the highest absolute value, but we named it the axis of crisis handling after we examined the addresses that had high scores qualitatively and found out that they contained many references to the crisis handling. For example, Kaifu referred to Gulf War many times in his third address (KTo3) and Koizumi referred to Iraq War many times in his addresses (KJ1-KJ10) (see also Fig. 2). We will examine this point in detail in a future study.

27 In the period of Yoshida (YS2) and Sato (SE), the variances of principal component scores were large, because they had a small number of tokens, and occasional factors like the

using these scores (PC1 and PC2) before and after the end of the Cold War (Yoshida (YS2) – Uno (US) and Kaifu (KTo) – Yasuo Fukuda (FY)), respectively. Solid lines represent the regression lines using the scores before the end of the Cold War, and the dotted lines represent those using the scores after it (upper and middle figures). Regression lines indicate that the level of interest in the United States–Europe has changed, from increasing before the Cold War to decreasing after it (upper figure), while the level of interest in Asia has consistently increased or at least constant through the end of the Cold War (middle and lower figure).[28] As well, the results indicate that Gulf War and Iraq war were the two important crises after the end of the Cold War (circled in the lower figure).

We conclude that the axis of the United States–Europe has had the most important role in Japanese foreign policy after World War II, but the axis of Asia has gradually increased in importance, and especially after the end of the Cold War, the shift from the axis of the United States–Europe to that of Asia has become more pronounced.

6.2 Changes in the US–Japan relations

6.2.1 Changes in the level of interest in particular aspects of the US– Japan relations. Next, we show the results for changes in interest in the US–Japan relations. Figure 3 represents the transitions in the values of categories, namely security issues, substantial value issues (economic issues and security issues), and all the issues (security issues, economic issues, and cultural/symbolic value issues) formed by compound nouns including the word 'Nichi-bei'. Dotted polygonal lines represent the values of addresses and a vertical line represents the end of the Cold War.[29] We fitted linear model using the values before and after the end

---

Vietnam War and reversion of Okinawa had a large impact on the latter case. We will develop more robust methods controlling for these points in a further study.

28 It seems that the second principal component (middle figure) shows a temporally lower score in Koizumi's period (KJ; 2001–2006), but that was caused by a higher absolute score for the third principal component (due to the Iraq war), and the third principal component also includes interest in Asia, thus we conclude that interest in Asia has consistently increased in that period as well.

29 Incidentally, we measured the relation between the value of category 'U.S' (sum of the relative frequencies of the regional/country names falling into that category) and the sum of the relative frequencies of the compound nouns including the word 'Nichi-bei'. The correlation coefficient of these two values was 0.81, and the portion of the latter for the former was 0.45.
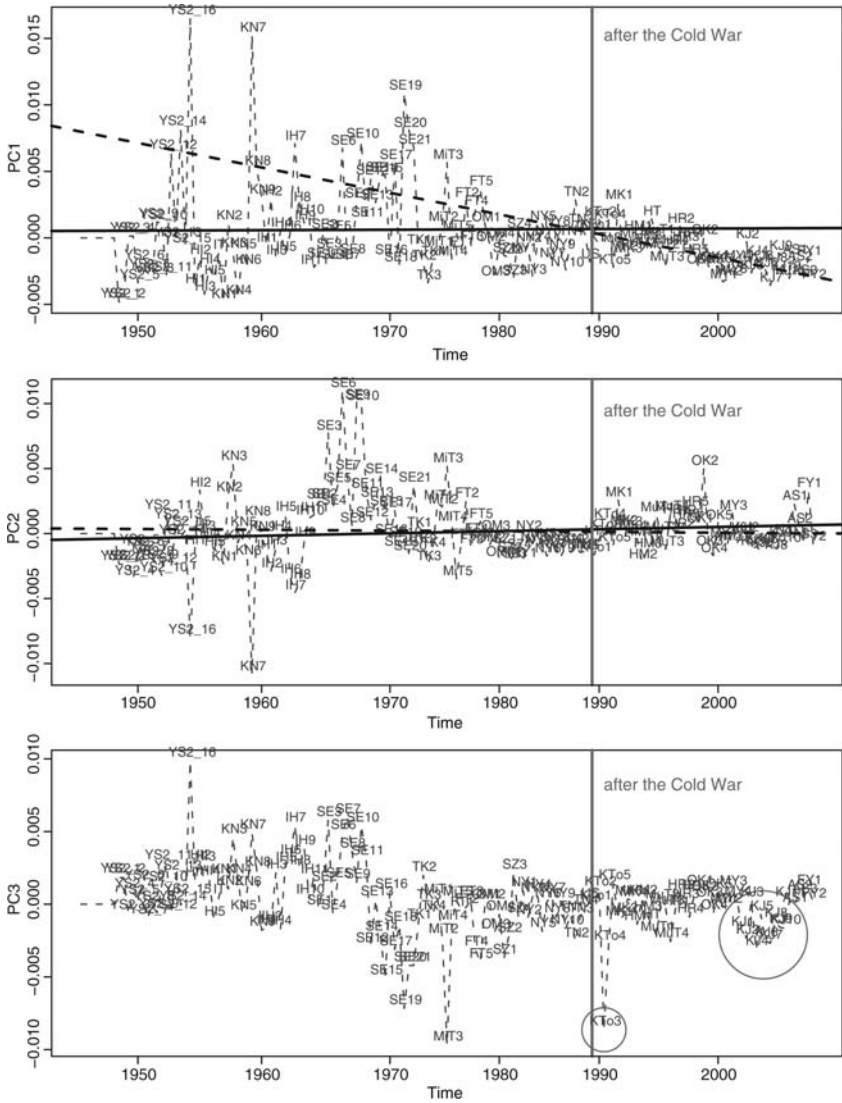
Figure 2 Transitions in first, second, and third principal component scores (dotted polygonal lines: principal component score for each address; vertical lines: the end of the Cold War; solid and dotted lines: regression lines using the scores before (solid) and after (dotted) the Cold War).

of the Cold War. Solid lines in Figure 3 represent the regression lines using the values before the end of the Cold War, and the dotted lines represent those using the values after it (security issues, all the issues respectively). We applied the Wilcoxon rank-sum test to the addresses

between Yoshida (YS2) – Uno (US) and Kaifu (KTo) – Yasuo Fukuda (FY) to test whether there was a significant difference before/after the end of the Cold War in the values of categories. Table 2 shows the mean values of categories formed by compound nouns including the word 'Nichi-bei' with the results of statistical tests.

Regression lines in Figure 3 show that the references to 'Nichi-bei' as a whole have changed, from increasing before the end of the Cold War to decreasing after it. Table 2 shows that the references to 'Nichi-bei' as a whole, and those related to cultural/symbolic value issues decreased after the end of the Cold War, though significant differences were not observed. On the other hand, regression lines in Figure 3 show that those related to security issues have consistently increased or at least constant through the end of the Cold War. Table 2 shows that those related to security issues and substantial value issues significantly increased after the end of the Cold War, and those related to economic issues increased as well, though the latter was not significant.

These results, corresponding to the regression lines in the first principal component shown in the upper figure of Figure 2, indicate that Japanese government has decreased its relative interest in the United States since the end of the Cold War. However, if we narrow our focus to security issues or substantial value issues, we can see that it has increased
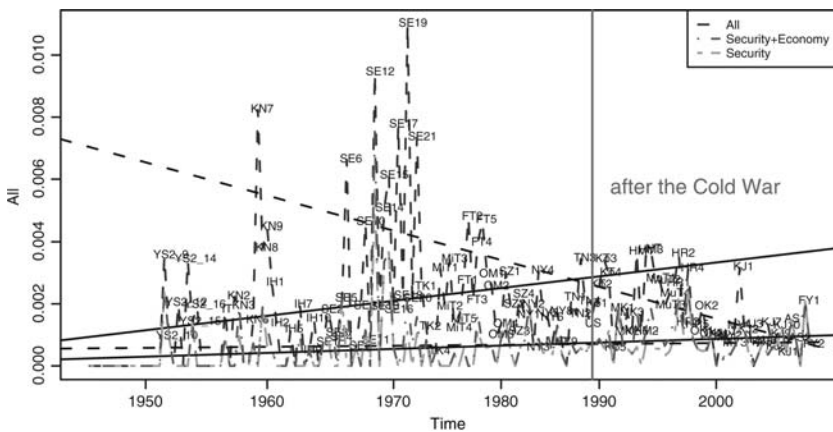
Figure 3 Transitions in compound nouns including the word 'Nichi-bei' (dotted polygonal lines: the values of addresses; a vertical line: the end of the Cold War; solid and dotted lines: regression lines using the values before (solid) and after (dotted) the Cold War, all the issues and security issues).

Table 2 Mean values of categories formed by compound nouns including the word 'Nichi-bei' before/after the end of the Cold War with the results of statistical tests

	Mean relative frequency		
	Before	After	P-value
Security issues	0.511	0.734	≪0.001*
Economic issues	0.124	0.142	0.185
Substantial value issues	0.636	0.876	≪0.001*
Cultural/symbolic value issues	1.328	0.768	0.208
All the issues	1.964	1.644	0.919

*$P < 0.01$.

its relative interest in the United States. We conclude that Japan has narrowed its interest in the US–Japan relations from a comprehensive interest including cultural/symbolic value issues to one centered on substantial value issues, and of the substantial value issues, the security issues have become more important since the end of the Cold War.

6.2.2 Changes in rhetoric. To examine the changes in rhetoric regarding the US–Japan relations before/after the end of the Cold War, we applied random forests using the text-feature matrix constructed on the basis of modifier/modificand of compound nouns including the word 'Nichi-bei' (all the issues, security issues, economic issues, cultural/symbolic value issues, respectively). We classified the addresses into two classes, namely Yoshida (YS2) – Uno (US) and Kaifu (KTo) – Yasuo Fukuda (FY), and extracted the important variables (i.e. modifier/modificand) that strongly contributed to the classification, using VI_{acu} calculated by random forests.

Table 3 shows the precisions (P_i), recall rates (RR_i), and F_1 values for the respective experiments. The results show that the modifier/modificand of 'Nichi-bei' classified the addresses before/after the end of the Cold War well for all the issues, security issues, and cultural/symbolic value issues, although not so well for economic issues as the number of tokens was small in this case.

We then examined the changes in rhetoric regarding the US–Japan relations before/after the end of the Cold War, for all the issues, security

Table 3 Classification performances using random forests with modifier/modificand of 'Nichi-bei'

	P_i	RR_i	F_1 value
All the issues	74%	30%	42%
security issues	100%	34%	51%
economic issues	$-^a$	0%	$-^a$
cultural/symbolic value issues	65%	28%	39%

[a]Denominator is 0.

issues, and cultural/symbolic value issues. Table 4 shows the 10 most important variables for classification with high VI_{acu} in the respective categories with the notation that these modifiers/modificands were frequently used ($+$) or infrequently used ($-$) after the end of the Cold War. In the following discussion, we mainly focused on the left-hand column because it used all the modifier/modificands, and thus the most important results. To deepen our discussion, we used the other two columns showing that these modifier/modificands were important in which context. We also used a qualitative examination of actual texts in addition to the quantitative results for our interpretation.

The left-hand column of Table 4 shows that the most important modifier/modificand distinguishing before/after the end of the Cold War regarding all the issues was 'Kokusai-Kyocho' (international cooperation). We examined the transition in the relative frequency of 'Kokusai-Kyocho' as modifier/modificand of 'Nichi-bei' from 1948 to 2008[30] and found that 'Kokusai- Kyocho' frequently appeared in the addresses of Koizumi. In fact he frequently used the expressions 'Nichi-bei Domei to Kokusai-Kyocho wo Gaiko no Kihon Hoshin to shite (we regard both the US–Japan relations and international cooperation as basic courses in Japanese foreign policy)' in his addresses.[31] We conclude that the juxtaposition with international cooperation was the most salient characteristic of rhetoric regarding the US–Japan relations after the end of the Cold War.

30 An additional figure is available on our website (http://researchmap.jp/stratovarius/).

31 See, for example, the address given on 20 January 2006.

Table 4 Ten most important variables with high VI$_{acu}$ (modifiers/modificands of the compound nouns including the word 'Nichi-bei') distinguishing before/after the end of the Cold War

	VI$_{acu}$(all)			VI$_{acu}$(sec)			VI$_{acu}$(cul/sym)		
Rank	Variable	Value	+/−	Variable	Value	+/−	Variable	Value	+/−
1	Kokusai-Kyocho	41.065	+	Shinrai	47.512	+	Kijiku	26.798	+
2	Kijiku	16.284	+	Kokusai-Kyocho	45.726	+	Renkei	15.259	+
3	Shinrai	9.832	+	Kinou	18.124	+	Kyoryoku	9.379	+
4	Kinou	9.309	+	Wakugumi	6.288	+	Mikoku-kan	5.176	+
5	Renkei	7.557	+	Juyo	6.258	+	Kyoko	4.839	+
6	Kyoryoku	5.039	+	Kyoka	4.692	+	Kakko	3.984	+
7	Kakko	4.768	+	Mokuteki	2.830	+	Yuko	3.428	−
8	Kochiku	3.631	+	Iji	0.440	−	Kochiku	2.844	+
9	Nikoku-kan	3.483	+	Enkatsu	0.071	−	Ken-an	2.536	−
10	Yuko	2.566	−	Hokan	0.056	−	Kiso	2.126	+

all, all the issues; sec, security issues; cul/sym, culture/symbolic value issues. Translations of each modifier/modificand are as follows. Kokusai-Kyocho, international cooperation; Kijiku, axis; Shinrai, confidence; Kinou, function; Renkei, coalition; Kyoryoku, cooperation; Kakko, inflexible; Kochiku, construction; Nikoku-kan, bilateral; Yuko, friendly; Wakugumi, framework; Juyo, important; Kyoka, enhancement; Mokuteki, purpose; Iji, maintenance; Enkatsu, smooth; Hokan, supplement; Ken-an, outstanding issues; Kiso, basic.

The left-hand column of Table 4 shows that the second most important modifier/modificand distinguishing before/after the end of the Cold War was 'Kijiku' (axis). The center and right-hand columns in Table 4 shows that 'Kijiku' was important in the context of cultural/symbolic values issues, instead of security issues. We examined the transition in the relative frequency of 'Kijiku' as a modifier/modificand of 'Nichi-bei' from 1948 to 2008[32] and found that 'Kijiku' appeared not only in the era after the Cold War, but also first appeared in the Miki period (MiT; 1974–1976), and was used in the latter half of the Detente era and in the New Cold War era as well. In fact, 'Kijiku' was used in the addresses to refer to the importance of the US–Japan relations, especially among several diplomatic relations. For example, Miki, after indicating that close relationships between the United States and Japan, Japan and China, and Japan and the Soviet Union were necessary parts of Japanese foreign policy, stated: 'Nakanzuku, sono Kaname wo Nasu no wa Nichi-bei Kankei de arimasu. Yuruginaki Nichi-bei no Shinrai Kankei koso ga Azia/Taiheiyo Chiiki no Antei no Kijiku de ari masu. (Among them, the US–Japan relations have the core role. A solid confidential relationship of trust between the two countries is an axis for stability in the Asia-Pacific region.)[33] This result corresponds to the results shown in Figure 2, namely the increased importance of Asia from the latter half of the Detente era. We interpret these results as follows: the US–Japan relations were overwhelmingly predominant in Japanese foreign policy before the latter half of the Detente era, and Japan had had no need to refer to 'Nichi-bei' as 'Kijiku', while from that period on Japan has had to explicitly refer to 'Nichi-bei' as 'Kijiku' along with the increased importance of multidirectional diplomacy.

We conclude that the second important changes in rhetoric regarding the US–Japan relations was recognition of these relations as an important axis among several foreign relations ('Kijiku'), and this change stemmed from Japan's multidirectional foreign policy.

Additionally, we should note that it appeared in the context of the cultural/symbolic value issues, instead of the security issues. This is because 'Kijiku' was frequently mentioned with 'Nichibei-kankei' or

---

32 An additional figure is available on our website (http://researchmap.jp/stratovarius/).

33 Address given on 16 September 1975.

'Nichibei-kan' in general instead of any specific matter, after the US–Japan relations themselves were substantialized, especially after the Cold War.

The left-hand column of Table 4 shows that 'Shinrai' (confidence), 'Kinou' (function), 'Renkei' (coalition), and 'Kyoryoku' (cooperation) were also important modifiers/modificands for distinguishing before/after the end of the Cold War, and frequently appeared, while 'Yuko' (friendly) appeared infrequently in that era. The center column of Table 4 shows that 'Shinrai' and 'Kinou' were distinguishing modifiers/modificands regarding security issues, and the right-hand column of Table 4 shows that 'Renkei', 'Kyoryoku', and 'Yuko' were distinguishing modifiers/modificands regarding cultural/symbolic value issues.

As political terms in Japanese, 'Shinrai' and 'Kinou' refer to specific and substantial partnership between the two countries, while 'Yuko' refers to general friendship between them. In the actual addresses, Sato, for example, described both substantial partnership and general friendship between the United States and Japan, by frequently referring to 'Yuko' as well as 'Shinrai' in the US–Japan relations in the context of the reversion of Okinawa,[34] while Koizumi and Abe described only substantive partnership between the United States and Japan, the former by stating: 'Nichi-bei Anzen Hosyo Taisei no Shinraisei wo Kojo Saseru to tomoni, Ryokoku no Jizoku Kano na Keizai Seicho wo Hakaru' (Improving confidence in US–Japan Security Arrangements and attempting sustainable economic development of both countries)[35] and the latter by emphasizing the 'Kinmitsu-na Renkei' (close coalition) of the United States and Japan.[36] We interpret that these changes in rhetoric were a reflection of changes in the US–Japan relations from a comprehensive and general friendship to a specific and substantial partnership, at least at the level of the Japanese government's interest. We also believe that this change in rhetoric characterized a new phase in the US–Japan relations after the end of the Cold War, in the context of strengthening and redefining the US–Japan relations.[37]

---

34   See the address given on 25 November 1970.

35   Address given on 4 February 2002.

36   Address given on 29 September 2006.

37   We should investigate these points in a further research by conducting more detailed qualitative analyses as well as quantitative analyses.

7 Discussion

Based on these findings, we propose that Japanese foreign policy after World War II can be characterized as follows: (a) Japan's regional interests shifted from US-centered to multidirectional; (b) the US–Japan relations were 'substantialized' after the end of the Cold War.

Before we explain our propositions, we introduce a popular model explaining Japanese foreign policy since the end of the Cold War, namely the 'reluctant realism' model (Green, 2001). By comparing our work with his discussion, we can clearly show the original aspects of our discussion.

7.1 Reluctant realism

Green insists that Japanese foreign policy since the Cold War has changed from the reactive, coping, and immobile character that were mentioned in previous literature to a more proactive 'reluctant realism' brought about by declining Japan's economic power and the rise of China. Japan's reluctant realism can be summarized as follows (Green, 2001, 6–8):

P1: a greater focus on balance of power;
P2: with growing realism, frayed idealism;
P3: a higher sensitivity to security;
P4: a more determined push for 'independent' foreign policy;
P5: a focus on Asia;
P6: a more fluid foreign policy-making process;

These points can be elaborated as follows: Japan's relations in East Asia were primarily determined by the conjunction of mercantile interests and US strategy in the past, but they now tend also to reflect a self-conscious competition with China; while the Japanese debate about a world role in the 1980s was framed in terms of the 'Kokusai-Koken' (international obligation), or 'Giri' (obligation) of the world's largest creditor nation, today foreign policies must be justified to the public in terms of 'Kokueki' (national interest); US 'Gaiatsu' (external pressure) on Japan is less salient in Japan's own defense debate, and in reaffirming alliance ties with the United States, Japan has simultaneously attempted to expand its own influence and agenda in East Asia and international organizations; Japan is pushing harder for a more independent

diplomatic identity in the world, and taking initiatives in Asia; the trends in Japanese domestic politics have created a more pluralistic and less predictable foreign policy-making process.

7.2 Shift from a US-centered to multidirectional foreign policy

Turning back to our propositions, we first argue that (a) Japan's regional interests shifted from US-centered to multidirectional through the postwar period. Our analysis of regional interests showed that the first principal component, the axis of the United States–Europe, decreased in importance after the end of the Cold War, while the second principal component, the axis of Asia, consistently increased or at least constant from the latter half of the Détente era. Previous discussion of Japanese foreign policy in the New Cold War era has mainly focused on security issues, and thus the US–Japan relations, but our analysis clarifies that Japan consistently increased its interest in Asia in this era as well.

This finding supports P1, the greater focus on balance of power, and P5, a focus on Asia, in Green's argument, but adds more. Though Green mainly focused on Japan's foreign policy character since the end of the Cold War, and discussed Japan's focus on Asia in that era, we emphasize here that Japan's interest in Asia has consistently increased since the latter half of the Detente era. Some previous studies have indicated that Japan already intended by the 1950s to 'emerge in Asia' (Oba, 2004), and increased its interest in the Asia-Pacific region in the 1960s, and created the special concepts of the 'Asia-Pacific region', and 'Pacific-rim region' in its foreign policy in 1970s (Watanabe, 1992). Our findings indicate that Japan has consistently increased its interest in Asia,[38] and provided empirical support for these studies.

Our findings on the rhetoric regarding U.S.-Japan relations corresponds to this finding; our analysis shows that U.S.-Japan relations started to be recognized as an important axis among several foreign relations ('Kijiku') in the Detente era for the first time, reflecting the fact that Japan started to develop a multidirectional foreign policy interest in that era. From these findings, we conclude that an important characteristic of postwar Japanese foreign policy was the shift from a US-centered to a multidirectional foreign policy.

---

38 As we mentioned in the previous section, the results in the Sato period were less conclusive, thus whether Japan has increased its interest since before the latter half of Detente era has not been confirmed yet. We will confirm this point in a further study

7.3 'Substantialization' of the US–Japan relations

We next argue that (b) the US–Japan relations have been 'substantialized' since the end of the Cold War. Our concept of substantialization consists of the findings that (b-1) Japan narrowed its interest in the US–Japan relations from a comprehensive interest including cultural/symbolic value issues to one focused on substantial value issues, especially security issues, and (b-2) the US–Japan relations were referred to side by side with 'Kokusai-Kyocho', and with substantial rhetoric.

Our analysis of the changes in interest regarding particular aspects of the US–Japan relations shows that Japan has decreased its interest in the US–Japan relations as a whole since the end of the Cold War, while, focusing only on substantial value issues, it increased its interest in them in that era. Green insists that the US–Japan relations changed after the end of the Cold War regarding security issues, as was shown in P2 (growing realism with frayed idealism), but Ono (2002), by extending his analysis to economic issues as well as security issues, insists that there was no changes. Our analysis, by covering cultural/symbolic value issues as well as security and economic issues, indicates that the US–Japan relations have indeed changed since the end of the Cold War, from comprehensive to substantial.

Symbolic value issues, representing the US–Japan relations generally in an ambiguous manner, are, in other words, not essential to the US–Japan relations. This surplus energy is needed only when the diplomatic relations between the two countries are unilateral. We believe that a new Japan, by reducing any redundancy and referring to the substantial value issues specifically, is not aiming at breaking down the US–Japan relations, but is showing a new intention to redefine these relations.[39]

Our analysis of the changes in rhetoric strengthens this argument. While Green referred to a change in rhetoric from 'Kokusai-Koken' as an idealistic value in the 1980s to 'Kokueki' as a realistic value in the 1990s in explaining P2, our analysis indicates that the juxtaposition of 'Kokusai-Kyocho' with 'Nichi-bei' was a more salient change in rhetoric before/after the end of the Cold War.[40] This change in rhetoric represents Japan's intention to posit the US–Japan relations in the

---

39   We will investigate this point in a further study using a more detailed qualitative analysis of the documents as well as a quantitative analysis.

40   This finding corresponds to interviews with related people (Shinoda, 2007, 37).

international context, which is a reflection of pushing for a diplomatic identity independent of the United States, but in a more strategic and complex way than simply claiming its own national interest ('Kokueki').

Further, our analysis shows that rhetoric regarding the US–Japan relations shifted to emphasizing 'Shinrai' and 'Kinou' in security issues, and from 'Yuko' to 'Renkei' and 'Kyoryoku' in cultural/symbolic value issues after the end of the Cold War, both representing a change from general friendship to substantive partnership. These changes in rhetoric, as well as the shift to substantial value issues, clearly show that the US–Japan relations have changed since the end of the Cold War.

We refer to this change in the US–Japan relations since the end of the Cold War as 'substantialization', and conclude that this is an important characteristic of postwar Japanese foreign policy. To sum up, we conclude that our analyses, using new quantitative text analysis methods, have shed new light on the character of Japanese foreign policy.

8 Conclusion

This study introduced new quantitative text analysis methods into foreign policy analysis. By using multivariate analysis and machine learning with larger amounts of textual data, and introducing linguistic knowledge for analysis, we succeeded in carrying out systematic as well as detailed analysis. We applied our methods to Japanese prime ministers' Diet addresses, and shed new light on the character of Japanese foreign policy. This study makes both a methodological contribution to foreign policy analysis and a substantial contribution to the study of Japanese foreign policy.

Along with the development of theoretical perspectives in foreign policy analysis, Katzenstein and Okawara (2001–2002) have indicated the importance of using several theories according to the nature of the respective research questions, instead of sticking to one grand theory (Sudo, 2007). In historical studies, the importance of multidimensional descriptions of diplomatic history was beginning to be recognized (Inoue, 2005).

This study, by extending the factors to be analyzed, and by analyzing the macroscopic correlation structure of regional interests, shares a common standpoint with these studies. Our methods, complement with qualitative studies, will enable us to analyze the character of foreign

policy taking into account various viewpoints. In future, we will develop our methods further, compare our results with other data, and provide more findings on the character of Japanese foreign policy.

Acknowledgements

Funding

References

Amyx, J.A. (2002) Moving beyond bilateralism? Japan and the Asian monetary fund. *Asia Pacific Economic Papers*, Australia-Japan Research Centre, 331. ideas.repec.org/p/csg/ajrcau/331.html. (11 March, 2011, date last accessed).

Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*. New York: Springer Science + Business Media, LLC.

Blaker, M. (1993) 'Evaluating Japan's diplomatic performance', in G.L. Curtis and M.E. Sharpe (eds), *Japan's Foreign Policy after the Cold War: Coping with Change*. New York: M.E. Sharpe.

Breiman, L. (1996) 'Bagging predictors', *Machine Learning*, 24, 123–140.

Breiman, L. (2001) 'Random forests', *Machine Learning*, 45, 5–23.

Calder, K.E. (1988) 'Japanese foreign economic policy formation: explaining the reactive state', *World Politics*, 40(4), 517–541.

Duffy, G., Frederking, B.K. and Tucker, S.A. (1998) 'Language game: dialogical analysis of INF negotiations', *International Studies Quarterly*, 42, 271–294.

Fukushima, A. (2000) 'Official Development Assistance (ODA) as a Japanese foreign policy tool', in T. Inoguchi, and J. Purnendra (eds), *Japanese Foreign Policy Today*, pp. 152–176. New York: Palgrave.

Green, M.J. (2001) *'Japan's Reluctant Realism'*, New York: Palgrave.

Green, M.J. and Cronin, P.M. (eds) (1999) *'The U.S.–Japan Alliance: Past, Present, and Future'*. New York: Council on Foreign Relations, Inc.

Hudson, V.M. (2005) 'Foreign policy analysis: actor-specific theory and the ground of international relations', *Foreign Policy Analysis*, 1, 1–30.

Hughes, C.W. (2004) *'Japan's Re-Emergence as a 'Normal' Military Power?'*. Oxford: Oxford University Press.

Inoue, T. (2005) 'The design of diplomacy in postwar Japan', pp. 67–80. The Annual of Japanese Political Science Association 2004.

Iokibe, M. (2006) *Sengo Nihon Gaiko Shi*. Tokyo: Yuhikaku.

Irie, A. (1997) 'Reisen ni kodawaranai sengo-shi', Nihon Kokusai Seiji Gakkai Newsletter, p. 79.

Jin, M. (2007a) *R ni yoru Deta Saiensu*. Tokyo: Morikita Shuppan.

Jin, M. (2007b) 'Authorship identification using random forests', *Proceedings of the Institute of Statistical Mathematics*, 55(2), 255–268.

Jin, M. and Murakami, M. (2006) Authorship identification with ensemble learning, 137–144. IPSJ SIG Computers and the Humanities Symposium 2006.

Katzenstein, P.J. (1996) *Cultural Norms and National Security: Police and Military in Postwar Japan*. Ithaca, NY: Cornell University Press.

Katzenstein, P. and Okawara, N. (2001–2002) 'Japan, Asia-Pacific security, and the case for analytical eclecticism', *International Security*, 26, 153–185.

Kudo, T. and Matsumoto, Y. (2002) 'Japanese dependency analysis using cascaded chunking', CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops), Taipei, pp. 63–69.

LaFeber, W. (1998) *The Clash: U.S.–Japan Relations throughout History*. New York: W.W. Norton & Company.

Laver, M.J., Benoit, K. and Garry, J. (2003) 'Extracting policy positions from political texts using words as data', *American Political Science Review*, 97(2), 311–331.

Lincoln, E.J. (1992) *Japan's New Global Role*. Washington, DC: Brookings Institution.

Lind, J.M. (2004) 'Pacifism or passing the buck? Testing theories of Japanese security policy', *International Security*, 29(1), 92–121.

Matsumoto, Y., Kitauchi, A., Yamashita, T., Hirano, Y., Matsuda, H., Takaoka, K. and Asahara, M. (2003) Japanese Morphological Analysis System ChaSen ver.2.2.3 (chasen.naist.jp).

Midford, P. (2002) 'The logic of reassurance and Japan's grand strategy', *Security Studies*, 11(3), 1–43.

Miyashita, A. (1999) 'Gaiatsu and Japan's foreign aid: rethinking the reactive–proactive debate', *International Studies Quarterly*, 43, 695–732.

Monroe, B.L. and Schrodt, P.A. (2008) 'Introduction to the special issue: the statistical analysis of political text', *Political Analysis*, 16, 351–355.

Nakagawa, H. (2000) 'Automatic term recognition based on statistics of compound nouns', *Terminology*, 6(2), 195–210.

Oba, M. (2004) *Ajia-Taiheiyo Tiiki Keisei e no Doutei: Kyokai Kokka Nichi Go no Aidentiti Mosaku to Tiiki Syugi*. Tokyo: Minerva Publishing.

Ono, N. (2002) *Sengo Nichibei Kankei no Kokusai Seiji Keizai Bunseki*. Tokyo: University of Keio Press.

Oros, A.L. (2008) *Normalizing Japan: Politics, Identity, and the Evolution of Security Practice*. Stanford, CA: Stanford University Press.

Pekkanen, S.M., Solis, M. and Katada, S.N. (2007) 'Trading gains for control: international trade forums and Japanese economic diplomacy', *International Studies Quarterly*, 51(4), 945–970.

Potter, D. and Sudo, S. (2003) 'Japanese foreign policy: no longer reactive?', *Political Studies Review*, 1, 317–332.

Pyle, K.B. (2007) *Japan Rising: the Resurgence of Japanese Power and Purpose*. New York: Public Affairs.

Sebastiani, F. (2002) 'Machine learning in automated text categorization', *ACM Computing Surveys*, 34(1), 1–47.

Shinoda, T. (1994) *Souridaijin no Kenryoku to Shido-ryoku: Yoshida Shigeru kara Murayama Tomiichi made*. Tokyo: Toyo Keizai Shinpo-sya.

Shinoda, T. (2007) *Reisen go no Nihon Gaiko*. Tokyo: Minerva Publishing.

Stockwin, J.A.A., Rix, A. and George, A. et al. (1988) *Dynamic and Immobilist Politics in Japan*. Honolulu: University of Hawaii Press.

Sudo, S. (2007) *Kokka no Taigai Kodo*. Tokyo: University of Tokyo Press.

Suzuki, T. (2009) 'Investigating Japanese government's perceptions of the postwar world as revealed in prime ministers' Diet addresses: focussing on East-West and North-South issues', *International Relations of the Asia-Pacific*, 9(2), 317–338.

Suzuki, T. and Kageura, K. (2007) 'Exploring the microscopic textual characteristics of Japanese prime ministers' Diet addresses by measuring the quantity and diversity of nouns', PACLIC21: Proceedings of the 21st Pacific Asia Conference on Language, Information and Computation, Seoul, pp. 459–470.

Tanaka, A. (1994) 'Rhetorics and limitations of Japan's new internationalism', *Japanese Studies Bulletin*, 14(1), 3–33.

Tanaka, A. (2000) *Word Politics*. Tokyo: Chikuma Shobo.

Tokunaga, T. (1999) *Joho Kensaku to Gengo Syori*. Tokyo: University of Tokyo Press.

Walker, S.G. and Schafer, M. (2007) 'Systematic procedures for operational code analysis: measuring and modeling Jimmy Carter's code', *International Studies Quarterly*, 42, 175–190.

Watanabe, A. (1974) 'Taigai ishiki ni okeru 'senzen' to 'sengo': syusyo gaisyo no gikai enzetsu no bunseki ni motozuku jakkan no kosatsu', in S. Satou and R. Dingman (eds), *Kindai Nihon no Taigai Taido*, pp. 225–274. Tokyo: University of Tokyo Press.

Watanabe, A. (1992) *Ajia/Taiheiyo no Kokusai Kankei to Nihon*. Tokyo: University of Tokyo Press.

Wu, X. and Kumar, V. (eds) (2009) *The Top Ten Algorithms in Data Mining*. Boca Raton: CRC Press.