

時代による総理大臣演説の文体的変化

鈴木崇史

東京大学大学院学際情報学府

影浦峽

東京大学大学院教育学研究科

概要

本研究では、1945年から2006年にかけての総理大臣国会演説の、時代による文体的変化を検討した。一文の長さ、延べ語に対する助詞の比率、助動詞の比率を時系列に観察すると、それぞれにトレンドが観察された。全ての助詞・助動詞の相対頻度を用いた主成分分析・クラスター分析によって、三木以外の総理演説が分類されることから、総理演説の助詞・助動詞使用に、もっとも大きな影響を与えている要因は時代であることが示された。分類に寄与の大きい助詞・助動詞の増加・減少傾向は、現代雑誌とほぼ対応していることから、総理演説の助詞・助動詞使用の変化は、主に日本語の変化に対応したものであると考えられる。

The Stylistic Changes of Japanese Prime Ministers' Addresses over Time

SUZUKI Takafumi
University of Tokyo

KAGEURA Kyo
University of Tokyo

Abstract

In this study, we examined the stylistic changes of Japanese prime ministers' addresses before the Diet (1945-2006). Macroscopic trends were observed in the sentence length, and the ratio of the particles and auxiliary verbs to tokens. Classification of these addresses by principal component analysis and cluster analysis with the relative frequencies of all particles and auxiliary verbs demonstrated that the most predominant factor affecting the use of particles and auxiliary verbs was the era. As the increase or decrease of the relative frequencies of all particles and auxiliary verbs which strongly contributed to classification corresponded to that in the magazines, the changes of the use of the particles and auxiliary verbs were attributed to general changes in Japanese usage.

1 はじめに

本研究では、1945年から2006年にかけての総理大臣国会演説における文体的変化を検討する。

文体論の領域では、一文の長さ、単語の長さ、品詞の相対頻度などの特徴量がしばしば観察される [2]。

これらの特徴量は、一人の著者においても時代によって変化することが知られており (e.g., Can and Patton[1])、また、大規模な語彙調査を比較することでも観察される [6, 7]。

総理演説の文体的特徴については、東 [8] が、文末表現や一文の長さや総理大臣の性格やリーダーシップとの関係を中心に検討している。また、文体的特徴に影響すると考えられる演説の作成過程については、篠田 [9] が検討しており、それによれば、通常、

総理演説は官僚が草稿を書き、それを総理が改訂するという作成過程を経るとされている¹。しかし、前者は就任直後の演説のみを分析対象としており、後者は関係者のインタビューを中心とした質的研究である。総理大臣の全国会演説を用いて、また、時代による特徴量の変化を中心として総理演説を体系的に検討した研究は存在しない。

このような背景をもとに、本研究では、以下の三つの観点を中心に、1945年から2006年にかけて行われた総理大臣の全国会演説 (施政方針・所信表明) をコーパスとして、総理演説の文体的特徴の変化を検討する。(1) 特徴量の変化になんらかのトレンドが観察されるか。(2) 時代は、他の要因、たとえば

¹秘書官によれば、三木武夫はこのような通常の演説作成過程に従わず自ら演説を起草したとされる [9]。

総理のリーダーシップや作成過程と比べ、特微量に大きな影響を与えるか。(3) 特微量の変化が観察されるとすれば、それは、総理演説特有の変化か、日本語の変化に対応したものか。

観察する特微量としては、一文の長さ、延べ語に対する助詞の比率、助動詞の比率、全ての助詞・助動詞の相対頻度を選択した。これらを選択した理由は、官僚文章の特徴が反映されると考えられ²、また、東[8]によって総理の性格やリーダーシップとの関係が指摘されているためである。

他の特微量のうち、文章に関する特徴は議会の種類(常会・臨時会)に、内容語はトピックに、文字の特徴は書記の特徴に依存する³。上記の特微量は、これらに依存しないと考えられるため(c.f., 金[5])、その点でも本研究の目的に適合的である。

2 データ

コーパスとして、第二次世界大戦後(1945-2006)の総理大臣 27 人 28 期(吉田茂は一期と二期にわけられる)、150 本の演説を取り上げる。2006 年 4 月 18 日に、オンラインデータベース「世界と日本」⁴から、2006 年 5 月 18 日に、「総理官邸」の Web ページ⁵から、それぞれ演説をダウンロードした。

表 1 は、各総理大臣の就任年月日、任期(単位:日)、国会演説本数、総理あたり延べ語数、一演説あたりの平均延べ語数である。

形態素解析には ChaSen[3] を使い、ChaSen の品詞タグを用いて、助詞・助動詞を抽出した。抽出した助詞の異なり語数は 111、助動詞の異なり語数は 60 である⁶。このうち、漢字とかなという点でのみ区別される 8 組(16 異なり語の助詞・助動詞)の表記ゆれを統一した。最終的に分析に用いた助詞の異なり語数は 105、助動詞の異なり語数は 58 である。

²官僚は、将来の政策運営において、以前の演説が制約となることを避けるために、長く、曖昧な文章を起草すると考えられる。

³演説のコーパスは、速記者の記録に基づいている。総理大臣が速記後のテキストを加筆・修正をすることはない。

⁴<http://www.ioc.u-tokyo.ac.jp/~worldjpn/index.html>

⁵<http://www.kantei.go.jp>

⁶助詞・助動詞両方の用法をもつ語が 7 種類(あり・なり・や・な・ね・で・し)観察された。

3 結果と考察

3.1 特微量の観察

特微量になんらかのトレンドが存在するか否かを検討するために、一文の長さおよび延べ語に対する助詞・助動詞の比率を時系列に観察した。

3.1.1 一文の長さ

図 1 は、一文の長さ⁷を時系列にプロットしたものである⁸。以下の上昇・下降傾向が観察される。

- 1953-1975 YS2-14⁹-MiT2: 下降
- 1975-1993 MiT2-HM2: 上昇
- 1993-2003 HM2-KJ6: 下降

図 1 から、一文の長さの変化には大きな波があることが観察される。官僚の影響がないことで知られている三木(MiT)の演説において、一文の長さが短いことから演説作成過程の違いもまた、一文の長さに影響を与えている可能性が示唆される。

3.1.2 助詞の比率

図 2 は、延べ語に対する助詞の比率を時系列にプロットしたものである。助詞の比率については、以下の上昇・下降傾向が観察される。

- 1945-1993 HN1-HM1: 上昇
- 1993-2006 HM1-KJ11: 下降

助詞の比率は、時代が経るにつれて増加したのち下降していることが観察される。リーダーシップや演説作成過程との関係は見いだせない。

⁷吉田(一期)二本目までの演説は、読点でくぎられていないため、欠損値として扱う。

⁸本研究では、総理大臣を単位とした文体的特徴の分析を中心とするが、時系列の図においては、変化をより細かく提示するために演説単位で図示する。

⁹吉田(2期)の 14 本目の演説をあらわす。以下同様。

表 1: 第二次世界大戦後の総理大臣演説

氏名	略称	就任年月日	任期(日)	演説本数	総延べ語数	平均延べ語数
東久邇宮稔彦	HN	1945年08月17日	54	1	5779	5779.0
幣原喜重郎	SK	1945年10月09日	226	1	2981	2981.0
吉田茂 1	YS1	1946年05月22日	368	3	5426	1808.7
片山哲	KTe	1947年05月24日	292	2	9072	4536.0
芦田均	AH	1948年03月10日	220	1	3506	3506.0
吉田茂 2	YS2	1948年10月15日	2248	16	26685	1667.8
鳩山一郎	HI	1954年12月10日	745	5	8318	1663.6
石橋湛山	IT	1956年12月23日	65	1	2665	2665.0
岸信介	KN	1957年02月25日	1383	9	18089	2009.9
池田勇人	IH	1960年12月08日	1433	11	35220	3201.8
佐藤榮作	SE	1964年11月09日	2798	21	56731	2701.5
田中角榮	TK	1972年06月07日	886	4	15097	3774.3
三木武夫	MiT	1974年12月09日	747	5	18222	3644.4
福田赳夫	FT	1976年12月24日	714	5	17030	3406.0
大平正芳	OM	1978年12月07日	554	4	15629	3907.3
鈴木善幸	SZ	1980年06月17日	864	4	13878	3469.5
中曽根康弘	NY	1982年11月27日	1806	10	47422	4742.2
竹下登	TN	1987年11月06日	576	4	18448	4612.0
宇野宗佑	US	1989年06月03日	69	1	3669	3669.0
海部俊樹	KTo	1989年08月10日	818	5	23541	4708.2
宮澤喜一	MK	1991年11月05日	644	4	19110	4777.5
細川護熙	HM	1993年08月09日	263	3	14609	4869.7
羽田孜	HT	1994年04月28日	64	1	4185	4185.0
村山富市	MuT	1994年06月30日	561	4	20454	5113.5
橋本龍太郎	HR	1996年01月11日	932	5	26809	5361.8
小渕恵三	OK	1998年06月30日	616	5	20702	4140.4
森喜朗	MY	2000年04月05日	387	4	20149	5037.3
小泉純一郎	KJ	2001年04月26日	1980	11	46996	4272.4

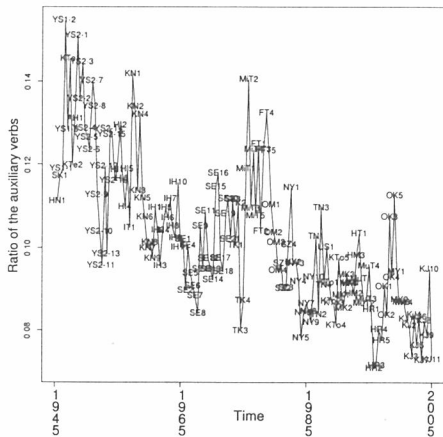


図 3: 助動詞の比率

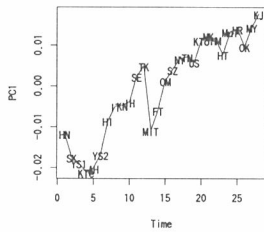


図 6: 第一主成分の推移

助詞・助動詞は矢印で示す。第一主成分の寄与率は46.03%、第二主成分までの累積寄与率は64.89%であり、これらによって文章の特徴づけをすることが可能である。

図5からは、次のようなグループが観察された。

- 1956-1978 石橋・岸・池田・福田: IT, KN, IH, FT
- 1964-1989 佐藤・田中・鈴木・竹下・中曽根: SE, TK, SZ, TN, NY
- 1989-2006 宇野・海部・宮澤・細川・村山・橋本・小淵・森・小泉: US, KTo, MK, HM, MuT, HR, OK, MY, KJ

図6は、第一主成分得点を時系列にプロットしたものである。第一主成分に注目すると、現在の総理演説は主成分得点が高く、過去の総理演説は主成分得点の低いことから、第一主成分が時代という要因と強く関係していると考えられる。

主成分分析の結果からは、橋本 (HR)・小泉 (KJ) を含め、多くの総理演説が時代ごとに分類されることが示されたが、三木は同時代の総理演説と異なった傾向を示している¹¹。

3.2.2 クラスタ分析

主成分分析の結果を別の手法で確認するために、クラスタ分析を用いて総理演説を分類した。クラスタの距離は相関距離、方法は群平均法を用いた。距離と手法は、結果に主成分分析との対応関係がもっともよくあられるものを選択した。クラスタ分析の結果の信頼性を確認するために、マルチスケールブートストラップ法 [4] を用いて、クラスタの信頼度 (p 値) を計測した。

図7が結果である。囲まれたクラスタが95%以上の p 値をもつ信頼度の高いクラスタである。

クラスタを大きく二つにわけると、三木 (MiT) を除くすべての総理演説が、1954年を境に、吉田(二期)(YS2)以前と、鳩山(HI)以後にわかれる。また、主成分分析で三グループにわかれた、鳩山以降の総理演説は、クラスタ分析によって1987年を境に、中曽根(NY)までと竹下(TN)からの二グループにわかれた¹²。

クラスタの境界となる時期と、一文の長さ、助動詞の比率、助動詞の比率の波とは明確な対応関係は観察されなかった。

主成分分析およびクラスタ分析を通じて、唯一三木の演説は同時代の総理演説とグループ化されなかった。三木は官僚の草稿を用いず、自ら演説を起草していたことで知られ、このような作成過程の大

¹¹ 散布図からは、鳩山(HI)以前の総理演説、羽田(HT)の演説も、孤立しているが、これは延べ語数が少ないためであると考えられる。

¹² 吉田(第二期)(YS2)以前の演説が細かいクラスタにわかれ、また羽田(HT)の p 値が低いことも、主成分分析の結果と同様、延べ語数が少ないためであると考えられる。

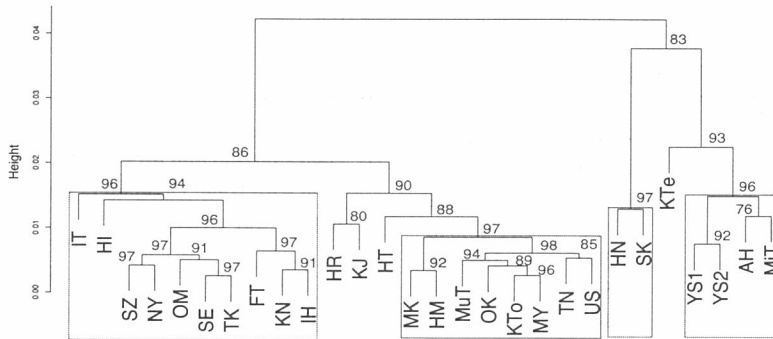


図 7: クラスタ分析の結果

きな違いが、時代の影響を超えて、助詞・助動詞の相対頻度に影響したと考えられる。

橋本 (HR)・小泉 (KJ) の演説は同時代の総理演説とクラスタ化されたが、 p 値は 95% 以下であった。総理演説のリーダーシップが、助詞・助動詞の相対頻度に影響し、それがクラスタの p 値に影響した可能性がある。しかし、田中 (TK)・中曽根 (NY) の演説にはこのような傾向が見られないことから、リーダーシップの影響は、存在したとしてもわずかなものであったと考えられる。

クラスタ分析の結果からは、総理演説が大きく時代ごとにクラスタ化されることが示された。これは、時代という要因が総理演説の助詞・助動詞使用にもっとも大きな影響をあたえていることを示している。

3.3 現代雑誌の語彙調査との比較

総理演説にみられた時代による変化が、総理演説固有のものか日本語の変化によるものか検討するために、第一主成分と強く関係する助詞・助動詞の相対頻度の増加・減少傾向を、1956 年と 1994 年に行われた国立国語研究所による現代雑誌の語彙調査 [6, 7] の結果と比較した。

第一主成分の因子負荷量に着目すると、上位 10 変

数は、{な (助動詞), を, や (助詞), で (助詞), などの, た, へ, が, とともに}であり、下位 10 変数は{あり (助動詞), で (助動詞), は, まし, に, ます, なる, も, ない, ぬ}である。

20 語のうち、現代雑誌の語彙調査と見出し語の対応づけが可能なのは 14 語であった¹³。図 2 は、現代雑誌の語彙調査におけるこれらの助詞・助動詞の相対頻度である。これらの増加・減少傾向を、1956 年から 1994 年までの総理演説と比較すると、{へ, も}をのぞく、12 語において、増加・減少傾向が一致していた¹⁴。

現代雑誌と総理演説では、その文章の目的や性質が異なり、口語と文語という違いがあるにもかかわらず、これらの傾向の一致が見られたことから、総理演説の助詞・助動詞使用の変化は、主に日本語の変化に対応したものであると考えられる。

4 おわりに

本研究では、1945 年から 2006 年にかけての総理大臣演説における文体的変化を観察した。

¹³{とともに}は見出し語に含まれず、{あり, で, まし, ます, なる} は 1956 年の語彙調査の見出し語が統一されていた。

¹⁴{た}は、上位 10 変数に含まれるが、1956 年から 1994 年をとると、総理演説においても減少傾向である。

表 2: 現代雑誌の語彙調査における助詞・助動詞の相対頻度の増減

品詞	1956年 (%)	1994年 (%)
な 助動詞	18.200	21.346
を 助詞	76.791	79.199
や 助詞	4.332	7.332
で 助詞	24.780	38.991
など 助詞	3.592	21.346
の 助詞	151.109	164.488
た 助動詞	70.814	61.947
へ 助詞	5.135	3.904
が 助詞	63.936	67.783
は 助詞	84.286	79.830
に 助詞	102.217	89.267
も 助詞	26.711	35.231
ない 助動詞	21.693	13.974
ぬ 助動詞	57.040	0.064

一文の長さ、助詞の比率、助動詞の比率を時系列に観察すると、それぞれにトレンドが観察された。助詞・助動詞の相対頻度を用いた主成分分析およびクラスター分析によって、作成過程の異なることで知られる三木以外の総理演説が時代ごとに分類されることから、総理演説の助詞・助動詞使用にもっとも大きな影響を与える要因は時代であることが示された。分類に寄与の大きい助詞・助動詞の増加・減少傾向は、現代雑誌とほぼ対応していることから、総理演説の助詞・助動詞使用の変化は、主に日本語の変化に対応したものであると考えられる。

今後は、個々の助詞・助動詞の変化率をより詳しく測定することで、総理演説と現代雑誌の語彙調査との対応関係をより詳細に検討することが課題である。

謝辞

有益なコメントをいただきました、田中明彦・原田至郎・辻慶太・芳鐘冬樹・大竹洋平の各氏、匿名の査読者、東京大学教育学研究科図書館情報学研究室のメンバーに感謝致します。技術的助言をいただきました、阿辺川武氏に感謝いたします。テキスト作

成に際してご助力いただきました、東京大学東洋文化研究所田中研究室のスタッフに感謝いたします。

参考文献

- [1] Fazli Can and Jon M. Patton. Change of writing style with time. *Computers and the Humanities*, Vol. 38, pp. 61–82, 2004.
- [2] Anthony Kenny. *The Computation of Style: an introduction to statistics for students of literature and humanities*. Pergamon Press, Oxford, 1982.
- [3] Yuji Matsumoto, Akira Kitauchi, Tatsuo Yamashita, Yoshitaka Hirano, Hiroshi Matsuda, Kazuma Takaoka, and Masayuki Asahara. *Japanese Morphological Analysis System ChaSen ver.2.2.1*. (<http://chasen.naist.jp>), 2000.
- [4] Hidetoshi Shimodaira. Assessing the uncertainty of the cluster analysis using the bootstrap resampling. *Proceedings of the Institute of Statistical Mathematics*, Vol. 50, No. 1, pp. 33–44, 2002.
- [5] 金明哲. 助詞の分布における書き手の特徴に関する計量分析. *社会情報*, Vol. 11, No. 2, pp. 15–23, 2002.
- [6] 国立国語研究所. 現代雑誌九十種の用語用字, 1962-1964.
- [7] 国立国語研究所. 現代雑誌の語彙調査－1994年発行70誌－, 2005.
- [8] 東照二. 歴代首相の言語力を診断する. 研究社, 東京, 2006.
- [9] 信田智人. 総理大臣の権力と指導力—吉田茂から村山富市まで. 東洋経済新報社, 東京, 1994.