

原著論文

Web の計量言語学的分析からみた政治的感性の特徴

村井 源*, 松本 斉子**, 山本 竜大*, 住住 彰文*

*東京工業大学大学院社会理工学研究科, **九州大学大学院システム情報科学研究院

SENSIBILITY ANALYSIS OF POLITICIANS' WEB LANGUAGE BY QUANTITATIVE LINGUISTICS

Hajime MURAI*, Naoko MATSUMOTO**, Tatsuhiro YAMAMOTO*, and Akifumi TOKOSUMI*

* Graduate School of Decision Science and Technology, Tokyo Institute of Technology, 2-12-1 Ohokayama, Meguro-ku, Tokyo 152-8552, Japan

** Graduate School of Information Science and Electorical Engineering, Kushu University, 744 Motooka, Nishi-ku, Fukuoka-shi, Fukuoka 819-0395, Japan

Abstract: In politics, the sensibility words play an important role to move citizen's feelings with a brief slogan. The purpose of this paper is to examine the characteristics of the political affection by the feature of the sensibility words. The sensibility words have been extracted from the Web pages of Diet member of the Liberal-Democratic Party, the Democratic Party, the Communist Party, and the Social Democratic Party in Japan. The extracted sensibility words were classified and examined. It was clarified that Japanese politicians have some common characteristics about sensibility expressions irrelevant to their ideologies. Moreover, it is possible to analyze the nuance of a political slogan by utilizing the database of the words that has dependent relationships with obtained sensibility word.

Keywords: Politics, Sensibility Word, WWW

1. はじめに

人の感性のメカニズムを解明する上で、感性語に着目した研究は多くの分野でおこなわれてきた。個々人の感性は、特に芸術的な対象の評価・価値判断とも関係が深い。そのため、感性語（印象語）を用いた絵画や画像、写真の検索システム [1]、楽曲の検索システム [2] などが種々実現されてきている。検索ではないが、伝統工芸と感性語との関係をデータベース化して利用する研究も行われている [3]。また、コンピュータ上で用いるフォントをユーザーの感性語から自動作成するシステムも構築されている [4]。これらの研究は視覚や聴覚などの身体的な感覚と感性の関係性について感性語を手がかりとして解明するものである。

一方で、感性語を用いたテキストの分類なども行われ始めている [5]。これは身体的な感覚とは別に、意味的で言語的な概念のまとまりもまた、感性語という切り口から解明できる可能性があることを示唆している。

意味的で言語的な概念を伝達するために、感性語が効果的に使われてきた分野としては、読者に感情的に訴えかけるような詩（歌詞なども含めて）や種々の文学が挙げられる。

言語の感性的表現の特徴を挙げると、論理ではなく感情や直感的な価値判断に訴えかける、どちらかという主観的な表現である、ということであろう。これは、複雑で、理解に思考的操作を必要とする、どちらかという客観的で論理的な言説とは対照的な性質である。

また感性的表現は、マスメディア上の広告などで、読者の

感情を喚起する効果を狙って用いられることが多く見られる。感性的表現によって感情が喚起された場合、その感情は人を行動へと向かわせる効果が期待されるからである [6]。

近年一方で、感性、感性的表現との関係がクローズアップされてきた分野として政治が挙げられよう。小泉元首相（首相在任 2001～2006）の言語表現は感性的・感覚的で短く、小泉語と評されマスコミにも大きく取り上げられた。これを野党民主党的の小沢党首（党首在任 2006～現在）と比較し、「感性政治と理性政治の対決」と評する政治学者もあった [7]。小泉首相の用いた、短く簡潔で感性に訴える表現には、理解しやすいという好意的な受け止め方も多く、その任期中高い支持率を維持することが出来たのかもしれない。しかし、感性的な政治の表現を、論理的思考の欠如、思考停止、空疎なスローガンの蔓延であるとして、問題視した見解も少なくない [8]。

感性的な表現は、簡潔で平易であると感じさせる一方で、意味の全体を構成する論理的構造や、文脈が見えにくい側面も持っている。しかしながら、政治の領域では、非常に短い選挙期間に、政治的知識や関心の薄い層も含め、多くの人々に対して自己の優位性をアピールする必要があり、スローガン・キャッチフレーズの感性的表現が多用される傾向にある。

政治家側としては、論理の把握困難な感性的表現によって、市民が政治家の意図を推測せざるを得ない状況を築けるため、明言を避けたい状況などでは好都合であろう。しかし、逆に市民の側からすれば、政治家の真意が不明瞭な状態で投票を行うことになりかねない。このため、感性的表現が多用される政治状況中では、各候補者・政党の論理構造と将来的

目標を、市民側が的確に読み取ることが必要となっている。政治領域の感性的表現に込められた意図を、感性工学を用いてより正確に解釈し、再表現することが可能になれば、種々のミスコミュニケーションを防ぎ、より効率的で豊かな意図の伝達が実現するであろう。また、政治的領域で有効な手法が確立できれば、将来的には他の詩・文学・音楽・美術や広告などの分野にも適用が可能であると考えられる。これらの領域では、現在でも人手によるコンコルダンス解析などでニュアンス解析が行われているが、感性工学的手法を用いることで、より科学的・計量的な解析が可能になるであろう。

本論文では、感性的表現の中でも、語彙レベルで表される感性、感性語を対象として取り上げ、政治領域で用いられる頻出感性語とその用法の特徴を解析する。感性語の政治領域独特の利用法を計量的に明らかにし、そのメカニズムを解明することで、将来的には小泉語と言われるような、簡潔だが難解な政治的表現の解明に貢献できると考えている。

2. 本論文で扱う対象

2.1 政治とWWW

情報化の進んだ現代社会において、さまざまな情報がWWWなどの形態で流通しており、情報伝達に占めるインターネットの役割は日増しに重要になってきている。政治の領域でもインターネットの重要性は増している。たとえば、前首相の発行するメールマガジンが人気となったり、巨大掲示板での議論から某国会議員のblogが炎上（blogに批判的コメントが集中し機能マヒを起こすこと）したりするなど、政治のさまざまな局面でインターネットの影響を見ることができる。

このような状況の中で、政治家が開設する政治家自身のホームページは一般的なものとなってきており、たとえば自由民主党の国会議員では419人中394人（94%）が自身のWebページをWWW上で公開している。これらの政治家のWebページ上で公開された言語表現は、政治的専門用語や政治的表現を駆使した国会などでの答弁とは異なり、一般の有権者に向けたものであると考えられる。このためWebページで閲覧可能な文書中には、政治家が有権者に向け直接に語りかける、先に述べたような平易でより感性的な表現が多数含まれていると期待される。またWWWというメディアの特性上、現代の政治的状況をリアルタイムに、かつ双方向的に反映するデータとしての価値もある。

政治家Webサイトは、支持者への報告や、有権者一般への政治的なアピールなどの非専門家の有権者に語りかける政治的言説が中心ではあるが、一部には私的な日記的内容も掲載されている。このため、内容面から政治家Webサイトを全て同質のデータとみなすことは困難である。しかし、本研究の対象は、有権者に政治を伝達する場合に用いられる、平易で簡潔な感性的表現である。全国の有権者が閲覧する可能性をもつWebメディアの性質は、ページの内容に関係なく共通である。本研究ではWebサイトのこのメディアの共通

表1 収集したテキストデータ

	自民	民主	共産	社民
議員数(人)	419	196	18	13
HP所有者(人)	394	190	18	13
HP所有率(%)	94%	97%	100%	100%
ファイル数	74740	37671	18831	3374
テキスト量(MB)	209	191	99	22
平均テキスト量(MB)	0.5	1.0	5.5	1.7

的性質は基本的に政治家側から認識されていると仮定し、Webサイトの全てのテキストデータを感性的表現の解析対象とする。

政治家が開設した、Webページを収集・解析の対象とすることで、政治家が有権者に直接語る言葉の研究の一基盤とすることができるであろう。また、これらのWeb上での言説と、従来の街頭演説・講演会などでの言説のあり方を比較することでWebメディア特有の言語表現的・感性的特徴が抽出できるであろう。言語的な解析と、Webメディアに特徴的なハイパーリンクなどの解析と複合させることで、Webメディアに特有の、現代的な思想伝達のダイナミズムを解明することが可能になると期待される。

2.2 日本語政治WWWテキストコーパスの構築

現在多くの領域のテキストが電子化され、コーパスやデジタルライブラリーの形で提供されるようになってきている[9]。言葉のありようを解明する上で、実際に用いられた言葉を蓄積したテキストコーパスは欠かすことができない。本研究の目的である政治領域における感性的な表現の解明のためにも、何らかのコーパスが必要となる。そこで、日本の政治家が開設するWebサイト（Webサイトを公開している、自由民主党・民主党・日本共産党・社会民主党の国会議員615人）における文書（主にhtml形式のもの）を複数の自動巡回プログラム（一部サイトは手動で収集）で収集し、タグや非文部分を除去して、形式を統一することで、WWW政治領域コーパスを構築した[10]（表1参照）。コーパスに収められたデータは、2006年12月から2007年5月にかけて収集されたものである。

3. 日本語政治WWWテキストコーパスの語彙的特徴

構築したテキストコーパスにおける用語法の特徴を、語彙レベルでいくつか示す。まず、品詞ごとの出現数とその党毎の比率を表したのが表2である。表中の▲と▽は、 χ^2 二乗検定の残差分析の結果、有意に多い項目と、有意に少ない項目をそれぞれ表している。この表中では全ての項目で $p < 0.01$ の水準で有意であった。品詞への分割と品詞の依存関係の解析にはChasen[11, 12]とCabocha[13, 14]を用いた。感性語の表記は分類語彙表[15]での表記に統一し、同訓の語彙でも、たとえば「厳しい」と「きびしい」は別の語彙として扱った。

表 2 政治コーパスで用いられる品詞数

	名詞	動詞	形容詞	副詞
自民	11,421,047	3,896,170	276,309	618,585
	▲ 70.4%	▽ 24.0%	▲ 1.7%	▽ 3.8%
民主	11,979,935	4,899,290	270,370	917,877
	▽ 66.3%	27.1%	▽ 1.5%	▲ 5.1%
共産	5,818,275	2,214,468	103,440	330,832
	▲ 68.7%	▲ 26.2%	▽ 1.2%	▽ 3.9%
社民	1,280,379	540,897	27,712	102,021
	▽ 65.6%	▲ 27.7%	▽ 1.4%	▲ 5.2%

表 3 各政党の形容詞使用の相違 (上位 50 位)

	全体		自民		民主		共産		社民	
1	よい	6.6%	よい	6.6%	よい	5.9%	強い	5.9%	多い	6.3%
2	多い	6.2%	多い	6.1%	いい	5.7%	大きい	5.6%	いい	5.6%
3	強い	4.8%	大きい	4.6%	多い	5.5%	よい	5.0%	強い	5.5%
4	大きい	4.5%	強い	4.4%	新しい	4.8%	新しい	4.9%	よい	5.2%
5	高い	4.4%	高い	4.4%	強い	4.6%	多い	4.9%	大きい	5.1%
6	新しい	4.3%	新しい	4.3%	大きい	4.4%	高い	4.6%	高い	4.4%
7	いい	3.4%	厳しい	3.2%	高い	4.0%	厳しい	4.1%	厳しい	3.2%
8	厳しい	3.3%	いい	3.1%	厳しい	3.6%	はやい	4.0%	新しい	2.7%
9	はやい	2.7%	はやい	2.6%	難しい	2.8%	いい	3.9%	難しい	2.7%
10	難しい	2.3%	難しい	2.1%	はやい	2.8%	少ない	2.5%	少ない	2.7%
11	少ない	1.9%	悪い	2.0%	よろしい	2.6%	悪い	2.1%	よろしい	2.5%
12	悪い	1.9%	少ない	1.8%	悪い	2.2%	難しい	2.0%	長い	2.2%
13	長い	1.8%	長い	1.7%	少ない	2.1%	低い	1.9%	おかしい	2.2%
14	若い	1.8%	若い	1.7%	長い	1.8%	よろしい	1.7%	はやい	2.1%
15	よろしい	1.5%	よろしい	1.5%	若い	1.7%	長い	1.7%	悪い	2.0%
16	すばらしい	1.5%	すばらしい	1.4%	広い	1.4%	広い	1.6%	若い	2.0%
17	ない	1.4%	ない	1.4%	おかしい	1.4%	重い	1.6%	ひどい	1.5%
18	深い	1.3%	うまい	1.4%	低い	1.2%	ひどい	1.5%	低い	1.5%
19	近い	1.2%	深い	1.3%	近い	1.2%	深い	1.4%	深い	1.4%
20	うまい	1.2%	楽しい	1.3%	重い	1.2%	若い	1.4%	広い	1.3%
21	広い	1.2%	忙しい	1.2%	うまい	1.2%	おかしい	1.2%	重い	1.3%
22	正しい	1.0%	近い	1.2%	正しい	1.2%	近い	1.1%	正しい	1.2%
23	楽しい	1.0%	広い	1.1%	深い	1.2%	詳しい	0.9%	うまい	1.2%
24	低い	1.0%	おいしい	1.1%	すばらしい	0.9%	正しい	0.9%	近い	1.1%
25	うれしい	1.0%	低い	1.1%	ない	0.9%	きびしい	0.9%	美しい	0.8%
26	重い	0.9%	正しい	1.0%	詳しい	0.9%	ふさわしい	0.8%	小さい	0.8%
27	ありがたい	0.8%	うれしい	0.9%	幅広い	0.8%	激しい	0.8%	すごい	0.7%
28	すごい	0.8%	寒い	0.8%	望ましい	0.7%	うまい	0.8%	楽しい	0.7%
29	幅広い	0.8%	重い	0.8%	小さい	0.7%	うれしい	0.8%	詳しい	0.7%
30	おいしい	0.7%	おもしろい	0.8%	ありがたい	0.7%	すばらしい	0.7%	ものすごい	0.7%
31	美しい	0.7%	ありがたい	0.8%	うれしい	0.7%	幅広い	0.7%	弱い	0.7%
32	忙しい	0.7%	すごい	0.7%	楽しい	0.7%	小さい	0.7%	うれしい	0.6%
33	おかしい	0.7%	幅広い	0.7%	古い	0.7%	安い	0.6%	短い	0.6%
34	遅い	0.7%	遅い	0.7%	細かい	0.6%	望ましい	0.6%	安い	0.6%
35	詳しい	0.7%	詳しい	0.7%	短い	0.6%	弱い	0.6%	幅広い	0.6%
36	安い	0.7%	美しい	0.7%	ふさわしい	0.6%	短い	0.6%	暗い	0.6%
37	おもしろい	0.7%	小さい	0.7%	弱い	0.6%	すごい	0.6%	望ましい	0.6%
38	明るい	0.6%	おかしい	0.6%	遅い	0.6%	熱い	0.6%	激しい	0.6%
39	小さい	0.6%	安い	0.6%	すごい	0.6%	著しい	0.5%	遅い	0.5%
40	寒い	0.6%	弱い	0.6%	安い	0.6%	古い	0.5%	著しい	0.5%
41	古い	0.6%	明るい	0.6%	おもしろい	0.5%	力強い	0.5%	苦しい	0.5%
42	あつい	0.6%	あつい	0.6%	激しい	0.5%	楽しい	0.4%	ふさわしい	0.5%
43	激しい	0.6%	古い	0.6%	寒い	0.5%	美しい	0.4%	おもしろい	0.5%
44	ふさわしい	0.6%	暖かい	0.6%	甘い	0.4%	おもしろい	0.4%	遠い	0.5%
45	熱い	0.5%	ふさわしい	0.5%	ひどい	0.4%	苦しい	0.4%	ありがたい	0.5%
46	遠い	0.5%	激しい	0.5%	あつい	0.4%	遅い	0.4%	古い	0.5%
47	短い	0.5%	珍しい	0.5%	著しい	0.4%	とんでもない	0.4%	すばらしい	0.4%
48	弱い	0.5%	熱い	0.5%	熱い	0.4%	ものすごい	0.4%	細かい	0.4%
49	優しい	0.5%	短い	0.5%	忙しい	0.4%	忙しい	0.4%	明るい	0.4%
50	望ましい	0.5%	遠い	0.5%	苦しい	0.4%	狭い	0.4%	危ない	0.4%

また、感性的表現と関連の深い感性語が多数含まれる品詞である形容詞での頻出語が各政党でどのように相違するか、全体と各党の上位50位までを表3に示す。上位50位までで、各党の形容詞の約8割を占めている。

用いられている形容詞の語彙であるが、上位50位中4政党に共通のものが38ある。いずれかの党で上位50位以内に入る語彙で、4党共通に上位150位以内までに入らないのは「きびしい」のみである(自民党で253位)。用いられている語彙には多少の差異はあるものの、党の思想的背景を超え、政治における感性語(この場合形容詞)の用い方は非常に似通っていることが分かる。現在の日本の政治状況においては、有権者の側からは、各政党がどれも同じに見えて困る、というような意見が頻繁に聞かれる。この一つの原因として、党に関わらず類似した語彙を利用することが考えうるのである。「政治的文体」とでも呼びうる、いかにも政治家が有権者に向けて語っているような言語表現の利用によって、内容はともかく、印象としては類似したイメージが残ると推測される。

4. 感性語の分類

4.1 感性語全体の大分類

党に関わらず類似した感性語の利用法によって表現されているのは、どのような感性なのか。各党に共通の感性語利用の特徴を解明することで、政治領域での感性的特徴を明らかにすることが出来る。また同時に党における相違を解析することで、思想や政治論理とは異なる、各党の感性的特徴をも抽出できることが期待される。

本節では、感性的特徴が最もよく現れていると考えられる形容詞を対象にする。そして、用いられた感性語の全体的な特徴を明らかにするため、主要な感性語がどのような意味分類に属しているかを区分する。意味区分は、自然言語処理や認知科学分野で頻繁に用いられる分類語彙表[15]にしたがう。

一つの語彙が、分類語彙表では複数の意味分類に属する場合、その語彙の政治コーパスにおける係り受け関係の頻度より、意味分類を判定する。ある語彙の、過半数を超える係り受け用法が一つの意味分類に属する場合、その分類を政治領域における主要な意味分類と認定し、その語彙の出現頻度すべてを主要な意味分類の頻度に加える。過半数を超える係り受け用法がない場合、上位の用法の出現比率の概算で各意味分類に頻度を割り当てる。以降の解析では、対象とする語彙は、各党の上位50位以内に入る形容詞とする。50位以内の形容詞で、各党の全出現形容詞数の約8割をカバーしている。

分類語彙表に基づく形容詞の意味分類を表4に示す。意味分類は分類語彙表における中項目に基づく。表中の▲と▽は、 χ^2 乗検定の残差分析の結果、有意に多い項目と、有意に少ない項目を、*と**はそれぞれ $p < 0.01$ の有意水準と、 $p < 0.05$ の有意水準を示している。

表4 使用頻度上位形容詞の意味分類

	自民	民主	共産	社民
真偽	2975 ** ▽	3129 ** ▲	925 ** ▽	329 * ▲
類	4066 ** ▲	3610 ** ▲	309 ** ▽	345
存在	4182 ** ▲	2451 ** ▽	1185	54 ** ▽
様相	56808 ** ▽	58250 ** ▲	17746 ** ▽	5628
力	18857 ** ▽	16841 ** ▽	7822 ** ▲	2188 ** ▲
時間	30383 ** ▽	28689 ** ▲	11661 ** ▲	2187 ** ▽
量	81771 ** ▲	69968 ** ▽	29214 ** ▲	7994 ** ▲
心	28054 ** ▽	26179	10535 ** ▲	2949 ** ▲
生活	3629 ** ▲	1039 ** ▽	402 ** ▽	95 ** ▽
行為	1649	1320 ** ▽	801 ** ▲	153
待遇	9801 ** ▽	9902	5096 ** ▲	913 ** ▽
経済	1962 ** ▲	1558 ** ▽	670	166
自然	8661 ** ▲	4842 ** ▽	1293 ** ▽	558 * ▽

表4によれば、政治敵領域において用いられる形容詞のうちで最も多いのは「量」のカテゴリーに属するものであり、大小関係や長短・遠近などに関係している。次に多いのが、「様相」に関係する語彙であり良し悪しなどの評価に関係する。全体的に多い順に、「時間」「心」「力」と続くが、「力」までの上位5カテゴリーで全体の約七割を占めている。

表4に示された分類のうちで、占める割合の大きい量、様相、時間、心、力に関して、中項目に含まれる実際の単語を分析し、政治における感性の特徴を抽出する。

4.2 量に関する感性

分類語彙表中の「量」のカテゴリーに属する形容詞で政治コーパスにおいて頻出したものを表5に示す。以下の表における単語の表示順は、分類語彙表における分類順とする。

表5に示された形容詞はそれぞれ、長さや大小、長短、高低などの程度や量に関係している。しかし、それぞれの程度の尺度において、無標形容詞で程度のはなはだしいもの(大小なら大、遠近なら遠、長短なら短…)の方がそうでないものよりも出現頻度が圧倒的に多い。例外は唯一、共産党において「長い」よりも「短い」の頻度が高い点のみである。なお係り受け関係を調べると、短にかかる名詞は、「時間」「期

表5 「量」に関する形容詞の出現頻度

	自民	民主	共産	社民
多い	17128 ** ▲	14889	5018 ** ▽	1756
少ない	5285 ** ▽	5645 * ▲	2607 ** ▲	735 ** ▲
長い	5034 ** ▲	4800 ** ▲	620 ** ▽	608 ** ▲
短い	1420 ** ▽	1705 ** ▽	1736 ** ▲	168 ** ▽
高い	12038 ** ▲	10910 ** ▽	4763	1216 * ▽
低い	2721 ** ▽	3359 * ▲	1942 ** ▲	421 * ▲
深い	3562 * ▲	3121 ** ▽	1443	397
遠い	478 ** ▲	317.67 ** ▽	79 ** ▽	44
広い	3205 ** ▽	3870 ** ▲	1703 ** ▲	372
幅広い	2096	2174 ** ▲	712 ** ▽	164 ** ▽
狭い	891	932	392	87
大きい	12544 ** ▽	11847 ** ▽	5823 ** ▲	1409
小さい	1782 ** ▽	1920 ** ▲	709	211
あつい	1694 ** ▲	1105 ** ▽	314 ** ▽	102 ** ▽
寒い	1758 ** ▲	1260	270 ** ▽	75 ** ▽
暑い	877 ** ▽	1100 * ▲	565 ** ▲	148 ** ▲
とんでもない	322 ** ▽	523	419 ** ▲	63

表 6 「量」に関する形容詞の程度による差

	合計	自民	民主	共産	社民
程度大	19.8%	20.3%	19.2%	19.5%	21.5%
程度小	5.1%	4.4%	5.0%	7.1%	5.9%

間」が圧倒的多数で、共産党は他党よりも「短い」「時間」(の労働)を強調していることがうかがえる。

程度の大きいことを示す感性語(「多い」,「長い」,「高い」,「深い」,「遠い」,「広い」,「幅広い」,「大きい」,)とそれに対応して程度が小さいことを示す感性語(「少ない」,「短い」,「低い」,「狭い」,「小さい」)を比較した場合の出現率を表 6 に示す。

4.3 様相に関する感性

分類語彙表中の「様相」のカテゴリーに属する形容詞で政治コーパスにおいて頻出したものを表 7 に示す。

表 7 から読み取れることとして、肯定的な評価(よい・いい・よろしい・美しいなど)が否定的な評価(悪い・難しい・危ない)に比べて、圧倒的に多いということがある。肯定的な感性表現と否定的な感性表現を比較した場合の比率を表 8 に示す。

表 8 の肯定表現と否定的表現の比率は、党の思想的な差異によらずほぼ同じである。肯定/否定は心的機能としては評価に対応するものである。そのため、制度や社会的事象を評価する必要の高い政治領域において、「様相」カテゴリーに分類される語彙の使用頻度が高いことは妥当であろう。

しかし、肯定的表現を否定的表現より多く(3倍程度)用いて語るという現象が、政党によらず発生するのはいかなるメカニズムに基づくのであろうか。各政党はそれぞれ社会的な事物に対しての評価が異なっているはずであり、各党において現在の政治状況への評価が分かれば、その肯定/否定の度合いに差が出るほうが、より自然であろう。テキストに含まれたコンテンツとは関係なく、評価的語彙の使用頻度が類似しているということは、政治的な思想やコンテンツとは関係のない、普遍的な「政治的感性」の存在を示唆する。「政

表 7 「様相」に関する形容詞の出現頻度

	自民	民主	共産	社民
すばらしい	4105 ** ▲	2509 ** ▼	736 ** ▼	122 ** ▼
よい	18237 ** ▲	15944 ** ▼	3999 ** ▼	1553 ** ▼
いい	9309 ** ▼	15483 ** ▲	5217 ** ▲	1445 ** ▲
よろしい	4172 ** ▼	6981 ** ▲	1809	706 ** ▲
悪い	5154 ** ▼	5971	2203 ** ▲	561
ふさわしい	1624	1667 ** ▼	861 ** ▲	135 ** ▼
美しい	2001 ** ▲	956 ** ▼	445	229 ** ▲
難しい	6329 ** ▼	7578 ** ▲	2097 ** ▼	738
危ない	398 ** ▼	608	260 ** ▲	104 ** ▲

表 8 「様相」に関する形容詞の肯定/否定

	合計	自民	民主	共産	社民
肯定的	14.8%	14.3%	16.1%	12.6%	15.1%
否定的	4.7%	4.3%	5.2%	4.4%	5.1%
肯定/否定	3.1	3.3	3.1	2.9	3.0

表 9 「時間」に関する形容詞の出現頻度

	自民	民主	共産	社民
遠い	478 ** ▲	318 ** ▼	79 ** ▼	44 ** ▲
新しい	11757 ** ▼	12953 ** ▲	5044	743 ** ▼
古い	1752 ** ▲	1766 ▲	514 ** ▼	126
若い	4927 ** ▲	4517 * ▼	1434 ** ▼	553 ** ▲
はやい	7574 ** ▼	7517 ** ▼	4169 ** ▲	572
遅い	1954 ** ▲	1618	421 ** ▼	149 * ▲

表 10 「時間」に関する形容詞の傾向

	合計	自民	民主	共産	社民
次世代的	9.1%	8.8%	9.2%	10.3%	6.7%
旧世代的	1.4%	1.5%	1.4%	1.0%	1.2%
次世代/旧世代	6.7	5.8	6.8	10.5	5.8

治的感性」においては、有権者に何かを語るときに、肯定的な話だけでも、否定的な話だけでもなく、それらを 3:1 程度に混ぜることで、話の印象のバランスをとるという戦略が働いているものと考えられる。

4.4 時間に関する感性

分類語彙表中の「時間」のカテゴリーに属する形容詞で政治コーパスにおいて頻出したものを表 9 に示す。

時間に関する形容詞の使用頻度から分かることとしては、「新しい」「若い」「はやい」のような次世代的な感性を感じさせる語彙の頻度が多いことがある。これに対し、「古い」「老いた」「遅い」などは旧世代的な印象を喚起させる語彙であると思われる。これらの比を表 10 に示す。

従来の多くの政治的な分類では、保守と革新のように政党を分類することが広く見られた。しかし、現在日本の政治状況においては、最も代表的な保守党と見られる自民党と、それに対する対抗馬と目される民主党の間でも、時代的な感性に大きな差は見られない。唯一特徴的なのが共産党であり、他党に比べて新世代的なイメージを喚起させる語彙の使用比率が多くなっている。それに対して社民党は野党であるにもかかわらず、保守政党である自民党よりもより新世代を感じさせる語彙の利用率が少ないことが示唆される。

4.5 心に関する感性

分類語彙表中の「心」のカテゴリーに属する形容詞で政治コーパスにおいて頻出したものを表 11 に示す。

「楽しい」「うれしい」「明るい」「おもしろい」という快感情に関する語彙と、「ひどい」「苦しい」という不快感情に関する語彙の出現頻度に政党ごとの差異が見られることが表 11 より読み取れる。これらの種類の語彙使用比率を表 12 に示す。

表 12 より、まず与党と野党で大きな感性的差異を見ることが出来る。与党自民党は圧倒的に快感情に関する語彙の使用が多い。これに対して、野党は軒並み不快感情の出現比率が相対的に高く、最も高い共産党では快感情と不快感情が同率で出現している。これに対し、野党ながら第二党の地位を占める民主党は両者の中間的な値をとっている。快感情と

表 11 「心」に関する形容詞の出現頻度

	自民	民主	共産	社民
楽しい	2798 ** ▲	1813 ** ▼	776 ** ▼	207 ** ▼
うれしい	2662 ** ▲	1779 ** ▼	452 ** ▼	172 ** ▼
明るい	1796 ** ▲	909 ** ▼	372 ** ▼	105 ** ▼
おもしろい	1829 ** ▲	1389 ** ▼	435 ** ▼	134 ** ▼
ひどい	905 ** ▼	1136 ** ▼	1528 ** ▲	425 ** ▲
力強い	1108 ** ▲	928 ** ▼	497 ** ▲	43 ** ▼
苦しい	879 ** ▼	1031	428	147 ** ▲
優しい	1282 ** ▲	907 ** ▼	192 ** ▼	87 * ▼
ありがたい	2311 ** ▲	1895	290 ** ▼	126 ** ▼
熱い	1482 ** ▲	1075 ** ▼	574 ** ▲	71 ** ▼
望ましい	1254 ** ▼	1981 ** ▲	664	155 * ▼
詳しい	1921 ** ▼	2315 ** ▲	949 ** ▲	203 ** ▼
細かい	905 ** ▼	1739 ** ▲	388 ** ▼	112 ** ▼
甘い	1074	1179 ** ▲	323 ** ▼	89 ** ▼
おかしい	1963 ** ▼	3812 ** ▲	1213	605 ** ▲
重い	2421 ** ▼	3198 ** ▲	1646 ** ▲	355

表 12 快感情と不快感情の傾向

	合計	自民	民主	共産	社民
快感情	2.6%	3.3%	2.2%	2.0%	2.2%
不快感情	1.0%	0.6%	0.8%	1.9%	2.1%
快／不快	2.7	5.1	2.7	1.0	1.1

不快感情には、各政党の現状に対する受け止め方が現れているのであろうか。

また、民主党においては、「詳しい」「細かい」などの話の詳細に関する感性語、「甘い」「おかしい」など評価に関する感性語が多いという特徴が見て取れる。これは、与党自民党の政策に対して、詳細に批判的に評価する立場にある民主党の特性が現れた可能性がある。

4.6 力に関する感性

分類語彙表中の「力」の категория に属する形容詞で政治コーパスにおいて頻出したものを表 13 に示す。また、これらの語彙を、強さを表す「強い」「すごい」「ものすごい」と、弱さを表す「弱い」に分けて、その比を表 14 に示す。力、強さを強調する傾向はあまり差がないが、共産党が若干強さを強調する傾向にあることがうかがえる。逆に民主党は、他の党と比較すると弱さを強調する傾向にあることが伺われる。

表 13 「力」に関する形容詞の出現頻度

	自民	民主	共産	社民
強い	13149	12335 ** ▼	6138 ** ▲	1535 ** ▼
弱い	1310 ** ▼	1631 ** ▲	663	199
すごい	2127 ** ▲	1580 ** ▼	619 ** ▼	207
ものすごい	740 ** ▼	977 ** ▲	402	203 ** ▲

表 14 「力」に関する形容詞の比

	合計	自民	民主	共産	社民
強さ	5.9%	5.8%	5.5%	6.9%	7.0%
弱さ	0.6%	0.5%	0.6%	0.6%	0.7%
強さ／弱さ	10.5	12.2	9.1	10.8	9.8

表 15 政治感性語データベース内のデータ

	自民	民主	共産	社民
係り受けパターン	100542	97634	47627	16038
係り受け総数	327779	312530	124025	31522
名詞種類	15570	14492	9243	4500
形容詞種類	2155	2089	1455	921

5. 係り受け関係による感性的ニュアンス特定

5.1 政治感性語データベースの構築

頻出する感性語の分類によって、全体的な感性の特徴を抽出することが出来る。しかしこれのみでは感性的表現自体の解析を行うことは困難である。

感性的な表現の解析を行うためには、全体的な感性の特徴に合わせて、解析の対象となる個々の感性的表現がどのような文脈で用いられているかを知る必要がある。そこで本論文では、日本語政治 WWW テキストコーパスに対して形態素解析と構文解析を行った。さらに、感性語と、感性語と係り受け関係にある名詞を抽出し、政治感性語データベースを構築した。政治感性語データベースに収めたデータを表 15 に示す。

政治感性語データベースは、政治領域で頻出する感性語に対して、それと係り受け関係にある単語の頻度を政党ごとに計量したものである。データベース用スクリプトプログラム (Perl により実装) を使用することで、ある感性語と係り受け関係にある名詞の政党ごとの違いと、逆にある名詞と係り受け関係にある感性語の各政党の差異を出力することが出来るよう設計されている。

政治感性語データベースは、一つの感性語とその係り受け関係にある名詞のデータベースであるため、複数の感性語を含む長文の複雑な感性的表現の計量的意味解析は、これのみでは行うことが出来ない。しかし、一つしか感性語を含まないような短いフレーズや、ある対象に対する感性的な印象の解析は行うことが出来る。

将来的には、現在の感性語単体に対応したデータベースを拡張し、複数の感性語を複合した感性的表現のデータの蓄積や、複合的表現によって作られた構文の解析機能などを盛り込むことで、より一般的な政治的言説における感性的表現の計量的意味解釈が可能になると考えられる。

5.2 政党間の感性的理解

政治感性語データベースを用いて、政党間の感性的な印象を抽出することができる。自民・民主に対して自党内の議員が付与した感性語と、他党の議員が付与した感性語の上位 5 位を表 16 に示す。回数はコーパス内での頻度を表す。

表 16 によれば、自民党も民主党も自分の党を「新しい」と評価しているが、他党からも「新しい」と評価されているのは民主党の方である。自民党は逆に他党からは圧倒的に「古い」政党であると評価されている。自民党の基本的な自己評価としては「新しく」「強い」政党であるものの、自党内でも「古い」・「旧い」という意識は強い。一方で他党からの自民党の評価としては、「古く」「悪い」党であるということが

表 16 自民・民主を修飾する感性語

自民→自民		他党→自民		民主→民主		他党→民主	
新しい	192	古い	75	新しい	143	新しい	15
古い	136	多い	14	若い	21	多い	5
多い	26	新しい	8	強い	12	少ない	4
強い	12	悪い	6	多い	10	危うい	3
旧い	8	長い	6	よい	7	強い	3
		はやい	6				

言えよう。0章の頻出感性語の分析結果より、政治的感性の共通的な特徴は「新しく」「強い」ことを強調するということがであった。政権党である自民党は野党から見れば「弱く」はないため、他党からの自民党の批判は「古い」という点に集中することになる。逆に「新しい」野党の民主党は、自己評価としては「強い」ものの、他党からは「危うい」という批判を受けることとなる構図を見て取ることが出来る。

5.3 「美しい国, 日本」

政治的なスローガンには、前述の通り短く簡潔な感性語を用いた表現が多用される。しかしながら、このような感性語を用いた表現は主観的で文脈の理解が困難である。そこで、安倍政権（2006～現在）のスローガンである「美しい国, 日本」を例として、感性語を用いた政治的スローガンの意味を、政治感性語データベースを用いて計量的に解析する。

政治感性語データベースから出力された、「美しい」という感性語と係り受け関係にある名詞の、自民党と民主党の相違を表 17 に示す（頻度上位 20 位）。また係り受け関係にある名詞を意味内容で分類した結果を表 18 に示す。

表 17 および表 18 の分類から、自民党が「美しい国」といった場合、「美しい」が表すのは主には自然環境や町の景観などに関する事柄である。ただし、「伝統」「文化」などの社会的な事柄も一部に含んでいる。一方民主党では、比較的社会的な「美しさ」は少ないものの、自然環境や言葉の美しさは多いということが読み取れる。

ただし、係り受け頻度の分布までは純粋に科学的で客観的な結果と言えるが、単語の意味分類によるニュアンス把握には複数の可能性が考えられ、分類者の主観が入り込む余地もある。また、得られた表は客観的データであるが、このままでは全体的なニュアンスを直感的に把握するのは困難である。このため特定の感性語に込められたニュアンスを、客観的で計量的に解析し、しかし解釈者の側は感性的にも理解しうる表現方法が望ましい。このような表現手段として考えられる手法の一つに、近年 Web 上で頻りに用いられるようになっているタグクラウドがある。

タグクラウドは、ある Web ページに込められた意味を表現するのに、関連する複数の単語を、ページとの関連度の強さによって大きさを変化させて列挙し、全体的なイメージを直感的に把握できるように工夫されたものである。

「美しい国, 日本」という表現に込められた、自民党と民主党のニュアンスの差をタグクラウドとして表現したのが、図 1 図 2 である。なお「美しい国, 日本」のニュアンスであるた

表 17 「美しい」が修飾する名詞上位 20

自民		民主	
国	358	国	175
日本	108	自然	40
自然	100	日本	34
国土	40	花	20
景観	35	日本語	20
星	30	風景	20
環境	28	景観	15
日本語	25	海	14
国家	22	言葉	12
地球	22	環境	10
言葉	20	国土	8
花	20	地球	8
街	19	森	6
風景	18	季節	6
人	17	町並み	5
姿	16	町	5
町	16	感動	5
豊か	16	紅葉	5
街並み	12	海岸	5
伝統	10	街	5
文化	10		

表 18 「美しい」が修飾する名詞の分類 (上位 20 位)

自民			民主		
分類	出現数	比率	分類	出現数	比率
国家	488	50.6%	国家	209	50.0%
自然環境	240	24.9%	自然環境	122	29.2%
町・景観	100	10.4%	町・景観	50	12.0%
社会・人	75	7.8%	言葉	32	7.7%
言葉	45	4.7%	社会・人	5	1.2%
その他	16	1.7%	その他		0.0%



図 1 自民党にとっての「美しい国, 日本」のニュアンス



図 2 民主党にとっての「美しい国, 日本」のニュアンス

め「国」と「日本」はタグクラウドから省略している。

本論文で取り上げた「美しい国、日本」のフレーズは感性語を一つしか含まず、そのため係り受け関係のみを蓄積したデータベースでもニュアンスの差異を把握しやすいケースである。しかし、より複合的な感性的表現であっても、原理的には、ここで例としてあげたタグクラウドのような、計量的なデータに基づくが感性的にも把握可能な表現手段は、意味の正確な伝達に有効であると考えられる。タグクラウドなどを用いることで、感性的表現の背景的な論理やニュアンスの差異が分かりにくいという欠点をカバーしつつ、しかし直感的に分かりやすいという利点を保持することが可能であると考えられる。

6. 結論：政治的感性の特徴

本研究で明らかになった、形容詞の使用頻度から見た、政治的感性の特徴を以下にまとめる。

6.1 政治に普遍的な感性

政党によらない、政治的に普遍的な感性の特徴として次の点を挙げる事ができる。

- ・大小などの「量」に関係する語彙と、よし悪しなどの「様相」に関係する語彙が多用される。
- ・「量」「様相」の次には、「時間」「心」「力」に関係する語彙が多用される。
- ・「量」に関しては、程度のはなはだしいことを示す語彙を多用する。
- ・「様相」に関しては、肯定的評価と否定的評価を、ある程度の比率で混合して用いる。ただし肯定的評価の方が圧倒的に多い。
- ・「時間」に関して、「新しく」「若く」「はやい」ことを強調する。

これらの特徴は、思想信条や、置かれた政治的な立場によらず見られるものであり、現代日本の政治に特徴的な感性的傾向といえる。逆に言えば、これらの感性的側面においては、各政党の個性が出ておらず、有権者の側からは各政党が類似して見えるといえよう。

6.2 政党に特徴的な感性

政党に特徴的な感性として今回特定されたのは、共産党における「新しさ」「若さ」などの新世代的な印象を感じさせる語彙の強調である。他の政党でも同様の傾向が見られるが、共産党での使用頻度は他よりも多かった。

また、自民>民主>共産・社民 の順で感情的な快に関する語彙の使用率が多く、逆に不快に関する語彙の使用率はその逆の順位で高かった。ただし、この傾向は、党としての感性的な傾向か、あるいは現在党が置かれた政治的地位的な側面によるのかは不明である。

6.3 感性的ニュアンスの差異

感性語と係り受け関係にある名詞を収集した、政治感性語データベースを構築した。政治感性語データベースを用いることで、政党の感性的な自己イメージと他者が持つイメージの差異を計量的に示した。

また、感性語を用いた政治的スローガンのニュアンスを、政治感性語データベースを用いて解析可能であることの例として、安部首相の「美しい国」を取り上げ、そのニュアンスを計量的に分析した。結果として、自民党的な美しさは、主には自然環境や町の景観などに関する事柄であるが、民主党に比較して「伝統」「文化」などの社会的な色彩が濃いことがわかった。

政治感性語データベースは、感性語と係り受け関係にある名詞のデータベースであり、一つしか感性語を含まないような短いフレーズや、ある対象に対する感性的な印象解析のみ行うことが出来る。将来的には、複数の感性語による表現のデータを蓄積し、構文の意味解析機能などを盛り込むことで、一般的な政治的言説における感性的表現の計量的意味解釈が可能になると考えられる。

また得られた結果を計量的のみならず感性的にも把握可能にするには、タグクラウドなどの、計量的データに基づくが直感に訴えるような表現手段が有効であると考えられる。

7. 政治的感性表現の理解に向けて

本研究では、WWW上の政治領域のテキストデータを収集した。そして、感性的表現に深く関係する感性語に焦点をおき、全体的傾向と党ごとの使用傾向を解析し、政治領域における感性的特徴を抽出した。

本研究では語彙レベルでの使用頻度と係り受け解析の結果を扱ったが、さらに詳細な分析のためには、各語が出現する文脈や意図の推定が必要になると考えられる。

また、本研究の結果と従来のメディアにおける言説と特徴を比較することで、Webメディアに特有の感性的傾向が明らかになると思われる。さらに、ハイパーリンクによる連携関係の解析とあわせることで、Web上での政治力学と思想的なグループ関係なども明らかにすることが可能であろう。

将来的には、本論文で得られた政治的感性語とその使用頻度、係り受け関係などの特徴を、機械が扱うことのできる形式での辞書的データ（オントロジー）とすることで、感性的で理解の困難な政治的表現の機械的な理解に向けた基礎的資料を構築したいと考えている。過去の文脈をデータベース化し、自動的な機械による意味処理を実現することによって、人間が話者の過去の文脈を逐一把握して、ありうる可能性を列挙して解釈する場合と比較して、大幅なコスト削減が実現できると期待される。

8. 謝 辞

本研究は、21 世紀 COE プログラム「大規模知識資源の体系化と活用基盤構築」の言語・文献、知識資源分野の研究の一環として行われた。

参 考 文 献

- [1] 椋木雅之, 田中大典, 池田克夫: 対義語対からなる特徴空間を用いた感性語による画像検索システム, 情報処理学会論文誌, 42, 7, pp.1914-1921, 2001.
- [2] 池添剛, 梶川嘉延, 野村康雄: 音楽感性空間を用いた感性語による音楽データベース検索システム, 情報処理学会論文誌, 42, 12, pp.3201-3212, 2001.
- [3] 宮川明大, 細川美佳子, 杉本匡光, 柴田義孝: 仮想現実空間下における伝統工芸分野への感性情報処理技術応用に関する考察, 2002, 54, pp.91-96, 2002.
- [4] 堀田創, 野澤貴, 萩原将文: 感性ルールベースを用いた日本語フォント自動作成システム, 情報処理学会論文誌, 48, 3, pp.1491-1501, 2007.
- [5] 原田隆史: 書評中の感性キーワードを用いた小説の分類, 情報知識学会誌, 15, 2, pp.57-62, 2005.
- [6] Frijda, N. H. :The Emotions, Cambridge University Press, 1986.
- [7] 凌星光: <http://weblog.lingxingguang.net/>, 2007/6/21 参照.
- [8] 小椋佳: インタビュー領空侵犯, 日本経済新聞朝刊, 2004/8/30.
- [9] 長塚 隆, 神門 典子: アジアの電子図書館研究・開発の動向, 情報管理, 48, 12, pp.785-792, 2006.
- [10] 村井源, 山本竜太, 往住 彰文: 日本語政治 WWW テキストコーパスの構築, 日本認知科学会第 24 回大会発表論文集, to appear, 2007.
- [11] ChaSen-- 形態素解析器: <http://chasen-legacy.sourceforge.jp/>
- [12] 浅原正幸, 松本裕治: 統計的日本語形態素解析に対する拡張 HMM モデル, 情報処理学会研究報告 2000-NL-139, pp.25-32, 2000.
- [13] CaboCha/ 南瓜 :Yet Another Japanese Dependency Structure Analyzer, <http://chasen.org/~taku/software/cabocha/>
- [14] Taku Kudo, Yuji Matsumoto :Fast Methods for Kernel-Based Text Analysis, ACL 2003, 2003.
- [15] 国立国語研究所: 分類語彙表 - 増補改訂版, 大日本図書, 2004.



村井 源 (非会員)

2007 年東京工業大学大学院社会理工学研究科博士後期課程修了. 博士 (工学). 現在, 東京工業大学大学院社会理工学研究科助教, 東洋英和女学院大学非常勤講師. 言語を切り口として, 思想の特徴を計量的に解析する手法の研究を中心に行っている. 人工知能学会, 情報知識学会, 認知科学会, 日本カトリック神学会などの会員.



松本 斉子 (正会員)

2006 年東京工業大学大学院社会理工学研究科博士後期課程修了. 博士 (学術). 現在, 九州大学大学院システム情報科学研究院特任准教授. 同大の次世代研究スーパースター養成プログラム学術研究員として. 『感性を切り口とした「心」の科学拠点の創成』プロジェクトに従事. 発話意図や人工物に対する愛着感情など, 高次感性に関わる研究を中心に進めている.



山本 竜太 (非会員)

東京工業大学 21 世紀 COE 特別研究員 研究分野: 政治学. 特に現代日本政治, 政治コミュニケーション, 東アジアにおける政治と ICT の関係. 主要業績: 「2003 年衆議院選挙にみるホームページの利用の変化と政策・公約に関する分析」日本選挙学会『選挙学会紀要』No.5. pp. 79-95. 慶応義塾出版会. 2005 年他.



往住 彰文 (正会員)

北海道大学大学院修士課程修了 (1977). 博士 (行動科学) (北海道大学, 2000). 北海道大学部文学部助手, 聖心女子大学文学部助教授を経て, 東京工業大学大学院社会理工学研究科教授. 専門は言語と感情の認知科学. Cognitive Science Society, 日本認知科学会, 人工知能学会など会員.

