

大統領就任演説のことば学

——人称代名詞の有標的使用——

稲 木 昭 子

1. はじめに

バラク・オバマ(Barak Obama)大統領就任演説は、全世界にリアルタイムで発信された後も、メディアで様々な観点から取り上げられてきている。そのほとんどで、就任演説がそれまでのオバマの演説とは一線を画するものであったと述べられている。これはある意味、大統領就任演説がそれまでの演説とは異なった儀式演説であったという制約からきていると言えなくもない。たとえばこれまでの演説なら、‘Yes We Can’ 等のようなオバマ流のサウンドバイト(soundbite: テレビ・ラジオのニュース番組などで短く引用される、政治家などの発信・所見など)は容易に指摘可能であったが、就任演説の中で各種のメディアが選択したサウンドバイトは実に様々であった。この就任演説を取り上げた日本のテレビ番組の一つにおいても、複数の番組登場者が別々のサウンドバイトを選択していたのは印象的であった。決め手に躊躇するということは、それだけ演説の中に盛り込まれた内容が豊富で多岐にわたっていたとも考えられる。ところが報道によっては分かりやすい英語であると書かれているものもあった。

この儀式演説が、分かりやすい英語あるいは読みやすい英語という印象をどういう風に与えることができたのであろうか。演説の一部分を引用してみる。

- (1) We remain a young nation. But in the words of Scripture, the time has come to set aside childish things. The time has come to reaffirm our enduring spirit ; to choose our better history ; to carry forward that precious gift, that noble idea passed on from generation to generation : the God-given promise that all are equal, all are free, and all deserve a chance to pursue their full measure of happiness.

一読しただけでも、語彙的にも構文的にも、必ずしも分かりやすいとは言にくいパラグラフである。たとえば3番目の文の構成をみると、この1文の語数は51語にもおよび、さらにいくつかの点で特徴が見られる。2番目の文中の‘the time has come’ と同じパターンが文頭で反復(repetition)され、これに、‘to reaffirm . . . ; to choose . . . ; to carry . . .’ がセミコロで並列(parallelism)される。コロンの後続する‘the God-given promise’ は、先行

する ‘that precious gift’ つまり ‘that noble idea’ の言い換えとなる。さらにこの同格表現 ‘the God-given promise’ の後続文中でも、同じパターンの繰り返しが見られる (all are equal, all are free, all deserve a chance)。文末の重みを活かして、‘the God-given promise’ と言い換えることにより、聴衆にしっかりとこの「神から与えられた約束」を刻みつけることができる談話構成となっている。

分かりやすい英語であるという評価は、一度この語彙的にも、構文的にも特徴のある儀式演説を読み始めると、必ずしも納得のいくものではない。むしろ複雑で難しいパラグラフという印象を受ける。この英文そのものの難易度判断に開きが出てくるのは、単なる印象からくるものであるのか、あるいは別の理由があるのであろうか。

ここでは、質的分析からくる特に易しさの評価の理由を、大統領就任演説の量的分析を行うことにより、抽出されるキーワードから、考えることにする。この分析は演説の特徴把握へとつながっていくのではないと思われる。

2. PC プログラム

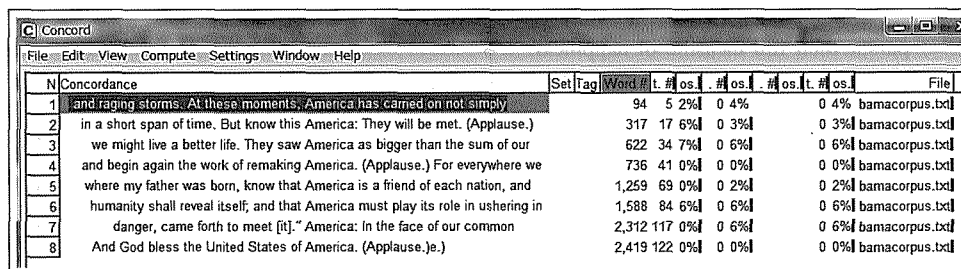
量的分析ということで、オバマ大統領就任演説をコーパス化し、作成したコーパスの言語処理を行うことにする。参考としてケネディおよびリンカーンの1回目の大統領就任演説も併せてコーパス化する。

大統領就任演説のテキストは、オフィシャル・サイト (<http://www.whitehouse.gov/blog/inaugural-address/>) からダウンロードし、WZ Editor (Version 4.00) で整形をしてコーパス化する。¹

作成したコーパスを処理するプログラムとして、Mike Scott の WordSmith Tools (Version 5.0) を利用するが、これには、主として次の3つのツールが含まれる。

① Concord Tool

これはいわゆるコンコーダンス (concordance) 作成のためのツールで、検索語が前後の文脈とともに画面に表示される。



N	Concordance	Set	Tag	Word #	t. #	os. %	精. os. %	精. os. %	File
1	and raging storms. At these moments, America has carried on not simply	94	5	2%	0	4%	0	4%	bamacorpus.txt
2	in a short span of time. But know this America: They will be met. (Applause.)	317	17	6%	0	3%	0	3%	bamacorpus.txt
3	we might live a better life. They saw America as bigger than the sum of our	622	34	7%	0	6%	0	6%	bamacorpus.txt
4	and begin again the work of remaking America. (Applause.) For everywhere we	736	41	0%	0	0%	0	0%	bamacorpus.txt
5	where my father was born, know that America is a friend of each nation, and	1,259	69	0%	0	2%	0	2%	bamacorpus.txt
6	humanity shall reveal itself, and that America must play its role in ushering in	1,588	84	6%	0	6%	0	6%	bamacorpus.txt
7	danger, came forth to meet [it]. "America: In the face of our common	2,312	117	0%	0	6%	0	6%	bamacorpus.txt
8	And God bless the United States of America. (Applause.)e.)	2,419	122	0%	0	0%	0	0%	bamacorpus.txt

図1 オバマ演説の検索語 america のコンコーダンス

② WordList Tool

このツールを使用すると、コーパスに含まれる語彙の頻度数順のリスト(F), アルファベット順のリスト(A), および統計情報(S)が得られる。統計情報の中には、タイプとトークン(異なり語数), 語彙の文字数および頻度数等の情報が含まれる。

N	Word	Freq.	%	Texts	%lemmas	Set
1	THE	136	5.63	1	100.00	
2	AND	111	4.60	1	100.00	
3	OF	82	3.40	1	100.00	
4	TO	71	2.94	1	100.00	
5	OUR	68	2.82	1	100.00	
6	WE	60	2.48	1	100.00	
7	THAT	50	2.07	1	100.00	
8	A	46	1.90	1	100.00	
9	IS	36	1.49	1	100.00	
10	IN	26	1.08	1	100.00	
11	THIS	24	0.99	1	100.00	
12	FOR	23	0.95	1	100.00	
13	US	23	0.95	1	100.00	
14	ARE	22	0.91	1	100.00	
15	BUT	20	0.83	1	100.00	
16	IT	19	0.79	1	100.00	
17	ON	17	0.70	1	100.00	
18	THEY	17	0.70	1	100.00	
19	WILL	17	0.70	1	100.00	
20	NOT	16	0.66	1	100.00	
21	HAVE	15	0.62	1	100.00	
22	HAS	14	0.58	1	100.00	
23	APPLAUSE	13	0.54	1	100.00	
24	CAN	13	0.54	1	100.00	
25	WHO	13	0.54	1	100.00	
26	WITH	13	0.54	1	100.00	
27	YOU	13	0.54	1	100.00	
28	BE	12	0.50	1	100.00	
29	NATION	12	0.50	1	100.00	
30	AS	11	0.46	1	100.00	

図2 オバマ演説ワードリスト(F)の上位30位

	N	Overall	
text file	Overall	rpus.bdl	
file size	13,692	13,692	
tokens (running words) in text	2,415	2,415	
tokens used for word list	2,415	2,415	
sum of entries			
types (distinct words)	895	895	
typetoken ratio (TTR)	37.06	37.06	
standardised TTR	46.30	46.30	
standardised TTR std.dev.	37.97	37.97	
standardised TTR basis	1,000	1,000	
mean word length (in characters)	4.46	4.46	
word length std.dev.	2.31	2.31	
sentences	124	124	
mean (in words)	19.48	19.48	
std.dev.	15.46	15.46	
paragraphs	1	1	
mean (in words)	415.00	415.00	
std dev			
headings	0	0	
mean (in words)			
std.dev.			
sections	1	1	
mean (in words)	415.00	415.00	
std.dev.			
numbers removed	1	1	
stoplist tokens removed	0	0	
stoplist types removed	0	0	
1-letter words	49	49	
2-letter words	418	418	
3-letter words	595	595	
4-letter words	420	420	
5-letter words	238	238	
6-letter words	225	225	
7-letter words	197	197	
8-letter words	111	111	
9-letter words	66	66	
10-letter words	58	58	
11-letter words	25	25	
12-letter words	7	7	

図3 ワードリスト(S)

ワードリストの上位には、一般的に冠詞・接続詞・前置詞等の機能語、さらに代名詞等が抽出される。

③ KeyWords Tool

このツールは、当該のコーパスとこれよりもずっと大きなコーパスを比較して、当該のコーパスに有標的に使用される語を抽出するものである。ここでは比較対象として、British National Corpus のワードリストを利用することにする。

([http://www.lexically.net/downloads/version 4/downloading%20 BNC.htm](http://www.lexically.net/downloads/version%204/downloading%20BNC.htm))

N	Key word	Freq.	%	. Freq.	RC. %	eyeness	P	emmas	Set
1	OUR	68	2.82	93,455	0.09	332.72	000000		
2	APPLAUSE	13	0.54	541		153.12	000000		
3	WE	60	2.48	00,833	0.30	148.47	000000		
4	NATION	12	0.50	3,567		94.54	000000		
5	US	23	0.95	80,226	0.08	71.64	000000		
6	AMERICA	8	0.33	8,934		42.18	000000		
7	AND	111	4.60	24,341	2.64	29.62	000498		
8	GENERATION	5	0.21	4,843		27.74	001358		
9	OATH	3	0.12	534		26.68	002375		
10	SPIRIT	5	0.21	6,314		25.16	005246		
11	FOUNDING	3	0.12	689		25.16	005248		
12	IDEALS	3	0.12	749		24.66	006800		

図4 オバマ演説キーワードの上位12位

このキーワードは、当該のコーパスとこれよりも大きなサイズのレファレンス・コーパス（reference corpus）における語の頻度数を比較して計算された結果でてくるものである。この時点で行われる検定は、カイ二乗検定とログ尤度検定である。より大きなコーパスにおける頻度数と比較して、‘unusually frequent’ となる語がキーワードとして抽出される。

このツールの結果として抽出されるキーワードには、一般的に3種類のものがある。一つ目は固有名詞。二つ目はテキストの‘aboutness’を示すキーワードで、これはテキストを読む人が、PC分析に頼らずとも普通にキーワードと判断するような種類のものである。三つ目は頻度の高い語で、これは読み手によって通常キーワードとは判断されにくい種類のものである。したがってこのキーワードは、特に抽出された語が含まれる文脈の精査が必要となり、当該の語が何故抽出されたのかを考察する必要がでてくる。

3. PC 分析

3.1 ワードリスト(S)

ワードリスト(S)の抽出結果から3つの点を取り上げる。

① 図3、資料の図12、資料の図15のワードリスト(S)には、各テキストに含まれる一文字語の語数、二文字語の語数などの情報が含まれているので、これをグラフで表示してみると、次頁のような割合になる。

3つのテキストの割合が非常に似通っていることが表示される。この文字数と語数には際だった開きは見られない。

② 1文の語数に関しては、オバマ演説は総語数が2415語で、1文は平均19.48語となる(図3)。これはケネディの25.30(図12)、リンカーンの26.96(図15)と比べると、三つの中で1番低い数値であり、1文の語数が少ないことになる。

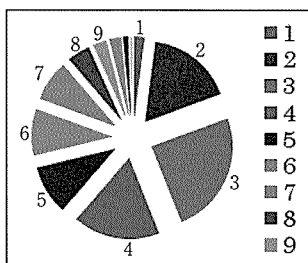


図5 オバマ演説の語数の割合

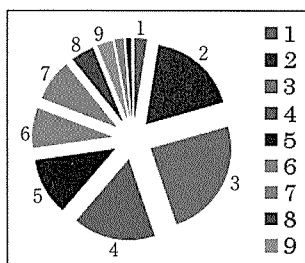


図6 ケネディの割合

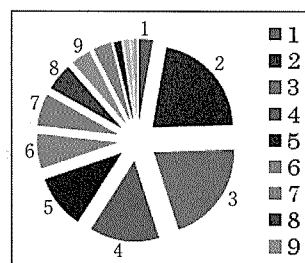


図7 リンカーンの割合

③ 演説のタイプ／トークンの値(type/token ratio)は、オバマ演説の場合で 37.06 と表示される(図 3)。ケネディとリンカーン一回目の大統領就任演説の値はそれぞれ 39.02, 27.64 となる(図 12, 図 15)。この三つの中では、リンカーンの割合が一番低い、つまり異なり語数の割合が低いので、テキストはより読み易いと判断される。²

3.2 キーワードとワードリスト(F)

オバマの演説のキーワードは図 4 であるが、上位にあるのが、2 位の 'applause' を除外すると、'our, we, us' などの一人称複数の代名詞群である。さらに、名詞 'nation, America', および接続詞 'and' が続く。

これをケネディの演説のキーワード及びリンカーンの演説のキーワードと比較してみると、ケネディでは、'let, pledge, our, we, sides, citizens' (資料の図 13) と続き、一部オバマ演説中のキーワードと重複するが、リンカーンの、'constitution, union, states, shall, be upon' (資料の図 16) とは明瞭に異なる。

オバマ演説のキーワードの一人称複数代名詞及び接続詞の 'and' は一つの特徴となる。これらはレファレンス・コーパスに比べて 'unusually' に高頻度で出てくる語ということになるが、興味あることに、これらの語は通常、ワードリストにおいても上位に位置する語である。

ワードリスト上位 15 位に含まれる人称代名詞は、オバマ演説では 'our, we, us' で、上位 30 位まで拡げると、'they' が入ってくる。これがケネディ演説では、上位 15 位で 'we, our, us' が含まれ、上位 30 位までには他の人称代名詞は含まれない。リンカーンの場合、上位 15 位では 'I' しか含まれず、上位 30 位まででは、'you' が含まれる。

この高頻度の語がオバマ演説の場合には、同時にキーワードとして有標化されて出てきていることになる。文やパラグラフ構成において基本的な語が、キーワードとして抽出されたわけであるから、これらは、テキストが何について書かれたものであるかという事を伝えるものというよりは、文やパラグラフ構成上の何かを示唆するキーワードとして働くと考えられる。

ここで、オバマ演説の代名詞の ‘I, my, me, you, your, we, our, us’ の各々の頻度数をまとめると右の表になる。

一人称複数の使用頻度数が目立つ。そこでこれらの代名詞がテキストのどの辺りで使用されているのかを視覚的に表示すると、次のプロットとなる。

テキストでは、一人称単数、二人称の使用が極端に押さえられ、一人称複数が最初から最後まで通して使用されているということが明示される結果となった。さらに ‘you, your’ は集中して使用されているということが分かる。

表 1 人称代名詞頻度数

単語	頻度数
I	3
my	2
me	0
you	13
your	3
we	60
our	68
us	23

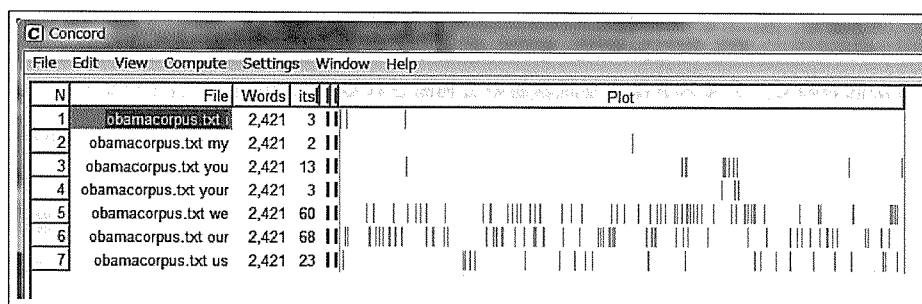


図 8 検索語 I/my/me/you/your/we/our/us のプロット

以下一人称単数、一人称複数、二人称の順で、これらの語が含まれるコンコーダンスを取り上げることにする。

‘I, my, me’ の検索結果は図 9 となり、演説ではわずかに 5 回しか使用されていない。

Concord			
File Edit View Compute Settings Window Help			
N	Concordance	Set	Tag Word #
1	My fellow citizens: I stand here today	my	1
2	My fellow citizens: I stand here today humbled by the task	i	4
3	the sacrifices borne by our ancestors. I thank President Bush for his service to	i	28
4	generation must lower its sights. Today I say to you that the challenges we face	i	284
5	capitals to the small village where my father was born, know that America	my	1,254

図 9 検索語 I/my/me のコンコーダンス

このうちの 3 例が演説の冒頭にでてくる。

- (2) My fellow citizens : I stand here today humbled by the task before us, grateful for the trust you’ve bestowed, mindful of the sacrifices borne by our ancestors.

I thank President Bush for his service to our nation – (applause) – as well as the generos-

ity and cooperation he has shown throughout this transition.

聴衆への呼びかけ、自分の立場、前大統領への謝辞の箇所である。この3回の I-stance が終わると、これに引き続くパラグラフで、‘we’ が初出される。

- (3) At these moments, America has carried on not simply because of the skill or vision of those in high office, but because we, the people, have remained faithful to the ideals of our forebears and true to our founding documents.

この ‘we, the people’ (我々合衆国人民) で、オバマは、I-stance から inclusive we-stance へと立ち位置を変化させ、聴衆を全て巻き込むというディスコースをとる。しかもこの ‘we’ は続く ‘the people’ とともに、アメリカ合衆国憲法の前文冒頭部からの引用であり、演説者が聴衆との精神的な一体感を生み出す極めて効果的な表現となる。

このパラグラフから始まる「包括的な ‘we’」は、話し手が聞き手を巻き込み、共感、仲間意識を盛り上げる効果的なことばとして機能するが、さらにこの語の使い方には特徴が見られる。いくつかの一人称複数代名詞のコンコーダンスをコンテキストを拡大して取り上げて、この点を確かめることにする。

‘us’ の例を取り上げる。

- (4) For us, they packed up their few worldly possessions and traveled across oceans in search of a new life. For us, they toiled in sweatshops, and settled the West, endured the lash of the whip, and plowed the hard earth. For us, they fought and died in places like Concord and Gettysburg, Normandy and Khe Sahn. Time and again these men and women struggled and sacrificed and worked till their hands were raw so that we might live a better life.

‘For us’ (我々のために) を3回文頭で繰り返し、アメリカの礎を作りあげた先人たちの苦難の道のを具体的に列挙し、そのお陰で今日の包括的な ‘we’ (我々) がよりよい生活を送ることができるのである、とアメリカの歴史そのものが他者への奉仕によって成り立っていることを強調する。このパラグラフは ‘us’ の例として引用しているもののもう1点、‘and’ が多用されていることに気付く。

‘our’ の例を取り上げる。

- (5) This is the journey we continue today. We remain the most prosperous, powerful nation on Earth. Our workers are no less productive than when this crisis began. Our minds are no

less inventive, our goods and services no less needed than they were last week, or last month, or last year. Our capacity remains undiminished. But our time of standing pat, of protecting narrow interests and putting off unpleasant decisions – that time has surely passed. Starting today, we must pick ourselves up, dust ourselves off, and begin again the work of remaking America.

‘Our workers’（我が国の労働者は劣らず生産性があり）、‘Our minds’（我々の心は劣らず創意に富み）、‘our goods and services’（我々の商品とサービスは劣らず必要とされ）、‘Our capacity’（我々の能力は衰えてはいない）と ‘our. . .’ のパターンでたたみかけて、‘our time’（〈これまでの立場に固執し、限られた利益を守り、不愉快な決断を先送りにする〉我々の時代）は過ぎ去ったと主張する。そこで、‘we’（我々）は、アメリカ再生に向け、今日から立ち上がり、ほこりを払い、再生という仕事をもう一度始めなければならない、と訴える。

次に ‘we’ の例を取り上げる。

- (6) For everywhere we look, there is work to be done. The state of our economy calls for action, bold and swift. And we will act, not only to create new jobs, but to lay a new foundation for growth. We will build the roads and bridges, the electric grids and digital lines that feed our commerce and bind us together. We’ll restore science to its rightful place, and wield technology’s wonders to raise health care’s quality and lower its cost. We will harness the sun and the winds and the soil to fuel our cars and run our factories. And we will transform our schools and colleges and universities to meet the demands of a new age. All this we can do. All this we will do.

至る所にやらなければならぬ仕事がある、と始めて、これから先なすべき仕事を列挙していく。‘we will act’（我々は〈新しい仕事を創造するだけでなく、成長のための新しい基盤を築くために〉行動する）、‘We will build’（我々は〈道路や橋などを〉建設する）、‘We’ll restore’（我々は〈科学をその本来の場に〉戻す）、‘(we will) wield’（我々は〈驚くべき科学技術を〉使う）、‘We will harness’（我々は〈車を走らせ工場を動かすために太陽や風などを〉利用する）、‘we will transform’（我々は〈新しい時代の要請に見合うように学校などを〉変える）と、‘we will’ を6回列挙することにより、自分たちのこれからの仕事内容を確認し、‘we can do’（我々は〈これら全てを〉することができる）、‘we will do’（我々は〈これら全てを〉行う）と結論づける。

さらにこのパラグラフにおいても、‘and’ が13回も使用され、動詞の目的語が列挙される等のより具体的な例示の明示化が行われる。

このように、多くのパラグラフにおいて、一人称複数代名詞が単発的に使用されるのではなく繰り返し使用されて、その結果常に相手を巻き込んで心理的結束を盛り上げていく効果を生み出している。

次に ‘you, your’ の集中する箇所をみる。

Concord			
File Edit View Compute Settings Window Help			
N	Concordance	Set	Tag Word #
1	must lower its sights. Today I say to you that the challenges we face are real.	you	287
2	and slaughtering innocents, we say to you now that our spirit is stronger and	you	1,469
3	is stronger and cannot be broken – you cannot outlast us, and we will defeat	you	1,481
4	cannot outlast us, and we will defeat you. (Applause.) For we know that our	you	1,489
5	society's ills on the West, know that your people will judge you on what you	our	1,640
6	West, know that your people will judge you on what you can build, not what you	you	1,644
7	that your people will judge you on what you can build, not what you destroy.	you	1,647
8	you on what you can build, not what you destroy. (Applause.) To those who	you	1,652
9	and the silencing of dissent, know that you are on the wrong side of history, but	you	1,672
10	history, but that we will extend a hand if you are willing to unclench your fist.	you	1,688
11	a hand if you are willing to unclench your fist. (Applause.) To the people of	our	1,693
12	nations, we pledge to work alongside you to make your farms flourish and let	you	1,707
13	pledge to work alongside you to make your farms flourish and let clean waters	our	1,710
14	a local restaurant can now stand before you to take a most sacred oath.	you	2,182
15	it safely to future generations. Thank you. God bless you. And God bless the	you	2,409
16	generations. Thank you. God bless you. And God bless the United States of	you	2,412

図 10 検索語 you/your のコンコーダンス

- (7) And for those who seek to advance their aims by inducing terror and slaughtering innocents, we say to you now that our spirit is stronger and cannot be broken – you cannot outlast us, and we will defeat you.

このパラグラフでは、‘you’ は、テロ行為や罪のない人たちへの大量殺戮を誘導し、自らの目的を前進させようとする人たちに対して、使用されている。つまりアメリカ国民を巻き込んだ「我々」の「敵」を、‘you’ で指していることになる。

次の箇所では、この ‘you’ の 9 例が集中する。

- (8) To the Muslim world, we seek a new way forward, based on mutual interest and mutual respect. To those leaders around the globe who seek to sow conflict, or blame their society's ills on the West, know that your people will judge you on what you can build, not what you destroy. (Applause.) To those who cling to power through corruption and deceit and the silencing of dissent, know that you are on the wrong side of history, but that we

will extend a hand if you are willing to unclench your fist. (Applause.) To the people of poor nations, we pledge to work alongside you to make your farms flourish and let clean waters flow; to nourish starved bodies and feed hungry minds.

‘those leaders’ (〈紛争の種をまき散らそうとしたり、自らの社会の悪を欧米諸国にせいしようとする〉世界の指導者) に対して、あなたたちの国民が、あなたが何を壊すかではなく、あなたが何を築くことができるのかによって判断するということを知りなさい。

‘those’ (〈腐敗と偽り、異議を押さえ込むことによって権力にしがみつく〉人) に対して、あなたが歴史の間違った側にいることを知りなさい、しかしあなたが、あなたの握った拳を開こうとするのであれば、我々は手を差しのべるということも知りなさい。

さらに、‘the people’ (〈貧しい国の〉人々) に対して、あなたたちの農場を豊かにし、清らかな水が流れるよう、あなたたちとともに働くことを誓う。

‘To the Muslim world, To those leaders, To those, To those people of poor nations’ と同じパターンを文頭に置いて、イスラム世界やアメリカに敵対する国の独裁者に対して、‘you, your ...’ で呼びかけて、己の立場を認識するのであれば手を差し出す用意があることを述べる。あるいは我々を好ましく思っていない人々に対して、‘you, your ...’ で呼びかけて、援助を約束する。

これらの ‘you, your’ は、目の前の聴衆を巻き込んだ「我々」が、いわば我々にとって不都合な「敵」に向かって言うときに使用されたものであり、この点は、‘we, our, us’ の指示対象と際だった差となる。

4. おわりに

オバマ大統領の就任演説の量的分析からくるキーワードとして、特に一人称複数の代名詞と接続詞 and が抽出されたが、これらはテキストの、頻度によるワードリストでも通常上位に位置する語でもある。ということはこれらのキーワードは Scott の定義するキーワードの第三番目の種類の高頻度に使用されるキーワードということになり、この場合特に当該の語の含まれる文、およびパラグラフの中でその分析をする必要がある。つまりこれらはいわゆるテキストがどのような種類のもので、どのような事について書かれているのかといったような種類のキーワードではないということになる。オバマの演説の場合にはキーワードの含まれるコンコーダンスから、抽出されたキーワードがテキストのディスコース上重要な語として使用されていることが説明される。

オバマの就任演説が、サウンドバイト的な表現でその特徴を説明できにくいということが言われていたが、量的分析からでてきたキーワードは図らずも、その語彙自体からテキスト

の種類を特定化するものではなく、ディスコースに結びつけられるものであるということが示唆された。

演説は、包括的な ‘we’ を使用しかつ集中して使用することにより、聞き手を巻き込み、共感意識を高めていく。同じパラグラフの中で、一人称複数の代名詞を含むパターンを繰り返す。さらに多様性に富む言い換え、and の多用による並列といったディスコース的特徴を駆使して、格調の高い演説を作り出していることになる。

資料

N	Word	Freq.	%	Texts	%	emmas	Set
1	THE	86	6.30	1	100.00		
2	OF	65	4.76	1	100.00		
3	TO	43	3.15	1	100.00		
4	AND	41	3.00	1	100.00		
5	WE	30	2.20	1	100.00		
6	A	29	2.12	1	100.00		
7	IN	26	1.90	1	100.00		
8	OUR	21	1.54	1	100.00		
9	THAT	20	1.47	1	100.00		
10	NOT	19	1.39	1	100.00		
11	FOR	16	1.17	1	100.00		
12	LET	16	1.17	1	100.00		
13	BUT	13	0.95	1	100.00		
14	THIS	12	0.88	1	100.00		
15	US	12	0.88	1	100.00		
16	BOTH	10	0.73	1	100.00		
17	ALL	9	0.66	1	100.00		
18	ARE	9	0.66	1	100.00		
19	CAN	9	0.66	1	100.00		
20	ANY	8	0.59	1	100.00		
21	DO	8	0.59	1	100.00		
22	IT	8	0.59	1	100.00		
23	SIDES	8	0.59	1	100.00		
24	WORLD	8	0.59	1	100.00		
25	BE	7	0.51	1	100.00		
26	BY	7	0.51	1	100.00		
27	FROM	7	0.51	1	100.00		
28	IS	7	0.51	1	100.00		
29	NEW	7	0.51	1	100.00		
30	OR	7	0.51	1	100.00		
31	PLEDGE	7	0.51	1	100.00		
32	THOSE	7	0.51	1	100.00		

図 11 ケネディ演説のワードリスト(F)

N	Overall	1
text file	Overall	rpus.txt
file size	7,565	7,565
tokens (running words) in text	1,365	1,365
tokens used for word list	1,366	1,366
sum of entries		
types (distinct words)	533	533
type/token ratio (TTR)	39.02	39.02
standardised TTR	43.90	43.90
standardised TTR std.dev.		
standardised TTR basis	1,000	1,000
mean word length (in characters)	4.39	4.39
word length std.dev.	2.26	2.26
sentences	54	54
mean (in words)	25.30	25.30
std.dev.	17.15	17.15
paragraphs	1	1
mean (in words)	366.00	366.00
std.dev.		
headings	0	0
mean (in words)		
std.dev.		
sections	1	1
mean (in words)	366.00	366.00
std.dev.		
numbers removed	3	3
stoplist tokens removed	0	0
stoplist types removed	0	0
1-letter words	34	34
2-letter words	247	247
3-letter words	330	330
4-letter words	226	226
5-letter words	156	156
6-letter words	110	110
7-letter words	114	114
8-letter words	64	64
9-letter words	43	43
10-letter words	24	24
11-letter words	13	13
12-letter words	1	1

図 12 ワードリスト(S)

KeyWords									
File Edit View Compute Settings Window Help									
N	Key word	Freq.	%	Freq.	RC. %	eyness	P	lemmas	Set
1	LET	16	1.17	25,382	0.03	91.34	000000		
2	PLEDGE	7	0.51	658		79.12	000000		
3	OUR	21	1.54	93,455	0.09	78.27	000000		
4	WE	30	2.20	00,833	0.30	67.75	000000		
5	SIDES	8	0.59	6,202		56.90	000000		
6	CITIZENS	5	0.37	3,260		37.26	000000		
7	US	12	0.88	80,226	0.08	35.62	000002		
8	NATIONS	5	0.37	4,115		34.96	000004		
9	BOTH	10	0.73	67,457	0.07	29.51	000528		
10	AMERICANS	4	0.29	2,745		29.40	000559		
11	GLOBE	3	0.22	650		28.92	000725		
12	FELLOW	4	0.29	4,273		25.91	003555		
13	#	3	0.22	04,421	1.61	-26.34	002827		

図 13 キーワード

WordList						
File Edit View Compute Settings Window Help						
N	Word	Freq	%	Texts	%en	
1	THE	256	7.03	1	100.00	
2	OF	146	4.01	1	100.00	
3	TO	135	3.71	1	100.00	
4	AND	105	2.89	1	100.00	
5	IN	77	2.12	1	100.00	
6	BE	76	2.09	1	100.00	
7	IT	59	1.62	1	100.00	
8	THAT	59	1.62	1	100.00	
9	A	56	1.54	1	100.00	
10	IS	49	1.35	1	100.00	
11	I	44	1.21	1	100.00	
12	BY	42	1.15	1	100.00	
13	NOT	42	1.15	1	100.00	
14	OR	35	0.96	1	100.00	
15	ALL	34	0.93	1	100.00	
16	AS	34	0.93	1	100.00	
17	NO	31	0.85	1	100.00	
18	CAN	28	0.77	1	100.00	
19	ANY	27	0.74	1	100.00	
20	WILL	27	0.74	1	100.00	
21	THIS	26	0.71	1	100.00	
22	FOR	25	0.69	1	100.00	
23	CONSTITUTION	24	0.66	1	100.00	
24	YOU	21	0.58	1	100.00	
25	ARE	20	0.55	1	100.00	
26	HAVE	20	0.55	1	100.00	
27	PEOPLE	20	0.55	1	100.00	
28	UNION	20	0.55	1	100.00	
29	WITH	20	0.55	1	100.00	
30	IF	19	0.52	1	100.00	
31	SO	19	0.52	1	100.00	
32	STATES	19	0.52	1	100.00	
33	BUT	18	0.49	1	100.00	
34	GOVERNMENT	18	0.49	1	100.00	
35	THERE	18	0.49	1	100.00	
36	WHICH	18	0.49	1	100.00	

図 14 リンカーン演説のワードリスト(F)

WordList		
File Edit View Compute Settings Window Help		
N	Overall	1
text file	Overall	opus.txt
file size	21,006	21,006
tokens (running words) in text	3,639	3,639
tokens used for word list	3,640	3,640
sum of entries		
types (distinct words)	1,006	1,006
typetoken ratio (TTR)	27.64	27.64
standardised TTR	40.33	40.33
standardised TTR std.dev.	45.66	45.66
standardised TTR basis	1,000	1,000
mean word length (in characters)	4.66	4.66
word length std.dev.	2.75	2.75
sentences	135	135
mean (in words)	26.96	26.96
std.dev.	18.62	18.62
paragraphs	1	1
mean (in words)	640.00	640.00
std.dev.		
headings	0	0
mean (in words)		
std.dev.		
sections	1	1
mean (in words)	640.00	640.00
std.dev.		
numbers removed	5	5
stoplist tokens removed	0	0
stoplist types removed	0	0
1-letter words	101	101
2-letter words	789	789
3-letter words	749	749
4-letter words	511	511
5-letter words	389	389
6-letter words	253	253
7-letter words	232	232
8-letter words	197	197
9-letter words	144	144

図 15 ワードリスト(S)

N	Key word	Freq	%	Freq	RC %	eymes	P/enu
1	CONSTITUTION	24	0.66	3,853		198.88	000000
2	UNION	20	0.55	17,000	0.02	100.15	000000
3	STATES	19	0.52	17,873	0.02	91.41	000000
4	SHALL	17	0.47	19,817	0.02	74.77	000000
5	BE	76	2.09	51,633	0.66	72.67	000000
6	UPON	15	0.41	22,806	0.02	58.39	000000
7	LABOR	5	0.14	164		57.12	000000
8	CONSTITUTIONAL	8	0.22	2,981		52.94	000000
9	ANY	27	0.74	22,180	0.12	52.19	000000
10	SECEDE	4	0.11	61		51.68	000000
11	LAW	14	0.38	26,483	0.03	48.76	000000
12	GOVERNMENT	18	0.49	56,343	0.06	46.20	000000
13	SURRENDERED	5	0.14	497		46.16	000000
14	LAWS	8	0.22	4,769		45.59	000000
15	OATH	5	0.14	534		45.45	000000
16	LAWFULLY	4	0.11	154		44.43	000000
17	EXPRESSLY	5	0.14	600		44.29	000000
18	PLAINLY	5	0.14	709		42.64	000000
19	SLAVE	5	0.14	605		40.44	000000
20	NO	31	0.85	29,618	0.23	35.89	000001
21	SLAVERY	4	0.11	460		35.77	000001
22	CITIZENS	6	0.16	3,260		35.26	000004
23	MINORITY	6	0.16	3,420		34.69	000009
24	ALL	34	0.93	77,566	0.28	34.64	000010
25	CAN	28	0.77	11,093	0.21	31.69	000152
26	FUGITIVE	3	0.08	156		31.54	000166
27	PEOPLE	20	0.55	16,198	0.12	30.51	000303
28	NOT	42	1.15	31,072	0.43	30.01	000401
29	ENFORCED	4	0.11	980		29.78	000456
30	ANARCHY	3	0.08	260		28.51	000903
31	ADMINISTRATION	6	0.16	6,408		27.38	001642
32	WHILE	13	0.36	54,789	0.06	26.65	002411
33	CLAUSE	5	0.14	3,632		26.54	002548
34	DISSATISFIED	3	0.08	460		25.25	005015
35	SENTIMENTS	3	0.08	460		25.12	005368
36	PERPETUAL	3	0.08	492		24.72	006613
37	WAS	5	0.14	63,917	0.87	-34.97	000004
38	#	5	0.14	04,421	1.61	-33.57	000000

図 16 キーワード

注

- 1 テキスト中にある ‘(applause)’ はあえてコーパス中に残しておいた。
- 2 この読み易さに関連して、青山学院大学開発のオンライン版「英文語彙難易度解析プログラム」(Word Level Checker) を使用してみる。WLC は教材の難易度を客観的に測定するための支援ツールとして 2006 年度からウェブサイトで公開されている。このツールに 2008 年に「リーダビリティ」(テキストの「読み易さ」の指標) の自動判定機能が追加され、その信頼性もすでに報告されている。ここでは難易度解析に使用する辞書を JACET 8000 と設定して、スコアを出した。この ARI, CLI のスコアのいずれもが、米国の学年レベルとして算出されており、日本の大学生の到達目標は 11 程度とされている。このツールで三つのコーパスの「読み易さ」を測定した結果は、表 2 となる。大統領就任演説の難易度は、Obama, Kennedy, Lincoln の順で読み易いと判断された。これはもちろん単にテキストの表面的、形態的な特徴のみによってリーダビリティを判断しているためにあくまでも参考データということになるが、それぞれの値の差がほぼ同じであることから大まかな判断は可能であり、大学生の到達レベルで考えたとき、オバマの演説は読み易いと判断できる。

表2 WLC による readability の結果

	Lincoln	Kennedy	Obama
Automated Readability Index (ARI)	14.8	12.9	9.6
Coleman-Liau Index (CLI)	10.8	10.1	9.2
Average	12.8	11.5	9.4

参考文献

- 東照二. 2009. 『オバマの言語感覚』 東京：NHK 出版.
- 荒木一雄・安井稔（編）. 1992. 『現代英文法辞典』 東京：三省堂.
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. Essex : Pearson Education Ltd.
- Huddleston, R. and G. K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge : Cambridge University Press.
- 池田拓郎. 1992. 『英語文体論』 東京：研究社出版.
- ケニイ, アンソニー（著）. 吉岡健一（訳）. 1982. 『文章の計量』 東京：南雲堂.
- 松尾弼之. 2004. 『大統領の英語』 東京：講談社.
- McEnery, T. and A. Wilson (eds.). 2001. *Corpus Linguistics : An Introduction*. Edinburgh : Edinburgh University Press.
- Meyer C. F. 2002. *English Corpus Linguistics : An Introduction*. Cambridge : Cambridge University Press.
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Harlow : Longman.
- Renouf, A. 1987. "Corpus Development." In Sinclair, J. M. (ed.), *Looking Up*. London : Harper Collins Publishers. pp.1-40.
- Scott, M. 2002. "Comparing corpora and identifying key words, collocations, frequency distributions through the WordSmith Tools suite of computer programs." In Ghadessy, M. et al. (eds.), *Small Corpus Studies and ELT : Theory and practice*. Amsterdam and Philadelphia : John Benjamins Publishing Company. pp.47-67.
- . 2005. "The Behaviour of Key Words," A Keynote Speech of Corpus Linguistics 2005. The University of Birmingham.
- . 2009. (version 5.0). *WordSmith Tools*. Oxford : Oxford University Press.