

# 「天声人語」の比較語彙研究（その1）

—1946年と2000年との語彙比較を通して—

宋 正植

キーワード 語彙、意味分野別構造分析法、コード付け、カイニ乗検定、有意差

## 1. はじめに

本稿の目的は、第2次世界大戦直後（1946年）の日本語の語彙と2000年のそれとを比較し、現代日本語における語彙の特徴を明らかにすることである。

語の集合である語彙<sup>1</sup>には、基本的性格として数量性と意味性がある。語彙を分析するには、この両方の性格を考慮しなければならない。語彙を、この両面から扱うのが比較語彙研究という意味分野別構造分析法である。これは、語彙を意味分野別に分けて、その構造を分析する方法である。具体的には、なんらかの基準によって分割された意味分野<sup>2</sup>のコードを調査対象の語彙にも与え、コードごとに、あるいはいくつかのコードにまとめて集計し、いかなる分野にどれだけの語があるかを分析する方法、言い換えれば、意味分野別の語彙構造を観察する分析法である。なお、本稿においては、コード付けの基準として国立国語研究所編『分類語彙表』（1964）を用いる。

## 2. 意味分野別構造分析のためのコード付け

意味分野別構造分析を行うためには、語彙の構成要素である個々の語にコードを付けなければならない。比較語彙研究で用いるコードには、単語コードと語素コードとがある。単語コードとは、単純語・複合語を問わず1語として付けるコードのことであり、語素コードとは、語の構成要素ごとにつけるコードのことである。これら2種類のコードは基本的に『分類語彙表』に従う。ただし、『分類語彙表』に収録されていない複合語のコードや付属語のコードなどについては、田島・広瀬（1997）を基にして新設された田島（2000）に従うことにする。

『分類語彙表』と現段階までの比較語彙研究で提案されているコードは以下の通りである。なお、『分類語彙表』にはコードの番号が4桁までであるが、ここでは省略する。

●『分類語彙表』コード

1. 体の類	2. 用の類	3. 相の類	4. その他 (接続詞や感動詞など)
1. 1	2. 1	3. 1	抽象的關係(人間や自然のあり方のわく組み)
1. 2	—	—	人間活動の主体
1. 3	2. 3	3. 3	人間活動—精神および行為
1. 4	—	—	人間活動の生産物—結果および用具
1. 5	2. 5	3. 5	自然—自然物および自然現象

●新設コード (田島 2000)

5. 接頭辞	8. 助詞・助辞	12. 語尾
5.001名詞化接頭辞	8.001格助詞	12.001各語尾
5.002動詞化接頭辞	8.002並列助詞	12.002人称語尾
5.003形容詞化接頭辞	8.003係助詞	12.003性数語尾
5.004形容動詞化接頭辞	8.004接続助詞	12.004時制語尾
5.005副詞化接頭辞	8.005終助詞・間投助詞	12.005先語末語尾
(5.006 その他)	8.006準体助詞	12.006終結語尾
6. 接中辞	8.007副助詞	12.007連結語尾
6.001名詞化接中辞	(8.008 その他)	12.008転生語尾
6.002動詞化接中辞	9. 助動詞	12.009用言語尾
6.003形容詞化接中辞	9.001第四類	(12.010 その他)
6.004形容動詞化接中辞	9.002第三類	
6.005副詞化接中辞	9.003第二類	13. 前置詞・介詞
(6.006 その他)	9.004第一類	14. 意味不明
7. 接尾辞	9.005別類	15. 固有名詞
7.001名詞化接尾辞	(9.006 その他)	16. 記号
7.002動詞化接尾辞	10. 補助動詞・補助形容詞・形	17. 漢字語素
7.003形容詞化接尾辞	式動詞・アスペクト動詞	18. 外来語語素
7.004形容動詞化接尾辞	11. 関係詞	19. (予備)
7.005副詞化接尾辞		20. 連語
(7.006 その他)		

### 3. 資料の選定および語彙調査の単位

#### 3.1 資料の選定

語彙調査の対象は、『朝日新聞』の「天声人語」である。標本は、全数調査と

抽出調査とがあるが、本稿では抽出調査の方式をとる。標本の抽出はランダム抽出である。抽出された語彙の分析方法は、比較語彙研究で言う意味分野別構造分析法である。

語彙調査の標本になるのは、1946年の記事66回分と2000年の記事40回分である。記事数の違いは、両者の記事の長さ（2000年の記事の方が平均200字程度長い）に差があることを勘案した結果である。

### 3.2 語彙調査の単位

本稿では、文節を第一次の調査単位とし、そこから1つの自立語と、いくつかの付属語を析出する。田島（1997、1999）は、文節を第一次の調査単位として採用することをすでに表明している。文節を採用する理由はいくつかあるが、語彙調査に際して比較的容易に見出される単位であることが最も主要な理由である。田島が言う文節とは、発話においてそれ以上短く切って発音すればその意味が喚起できないか、あるいは無意味になってしまう限界の最初の単位である。

語彙調査の手順としては、天声人語の記事の文を一旦文節に分割した後、その中から、1つの自立語を見出し、その語に残った部分があれば、付属語を析出する。複合サ変動詞、アスペクトを示す動詞を伴った複合動詞、形容動詞、いわゆる造語成分・接尾辞を伴った語などは、文法論的には様々な議論があるが、それぞれが1語としてまとまりを持つので、これらには1つの単語コードを与える。

調査単位については、以下のような原則と細則に従って文を分割することにする。

原則：記事を文節に分け、文節を自立語と付属語に分割する。

細則

- 1) 派生語と複合語は原則として1単位とする。  
例) 見送る (2.351)、話しつづける (2.313)、走り回る (2.1523)
- 2) 「～する」形のサ変動詞は原則として1単位とする。  
例) 説明する (2.3163) ←説明 (1.3163)、出現する (2.121) ←出現 (1.121)
- 3) 形容動詞は「形容動詞語幹(名詞) + 助動詞(だ)」を1単位とする。  
例) 安心だ (安心だ: 3.3012) 静かだ (静かだ: 3.503)
- 4) 造語成分・接尾辞は分割せず1単位とする。  
例) 授業中 (授業中: 1.364) 学校中 (学校中: 1.263)
- 5) 補助動詞・補助形容詞の場合は分割し、それぞれ1単位とする。

例「～てしまう」「～てよい」「～てほしい」

(～て：8.004)(しまう：10.1502)(よい：3.133)(ほしい：3.3012)

6) いわゆる助動詞のうち、「(ら)れる」「(さ)せる」「たい」「らしい」については、それぞれ1単位とする。その他、「た」「だ」「です」「ます」「ず」「ぬ」「べき」「まい」などについても1単位とし、一括して助動詞の新設コード「9」を与える。

7) 助数詞が付いているものは、「数字+助数詞」を1単位とする。

例「3人」「2ヶ月」「5メートル」(すべて1.196コードを与える)

8) 固有名詞

例) 人名に関しては「姓+名」を1単位とする。

また、地名・国名・書名・建造物名なども1単位とする。

9) 外国語は、日本語の調査単位に合わせてコード付けをする。

例) アイ(一人称)→(私：1.20)

10) 記号・区切り符号などは全て「16」のコードを与える。

分割された単語のコード付けは、『分類語彙表』を基準に、新設コードを併用する。

#### 4. 「天声人語」の語彙調査

「天声人語」の語彙調査の結果、1946年の66回分では、延べ語数19817個、異なり語数4578個が得られ、2000年の40回分では、延べ語数20304個、異なり語数4162個が得られた。これらは、両年の記事に含まれている全ての単語及び記号にコードを付けて得られた数値である。これらの数値を類別に分けてみると、<1.>(体の類)～<16.>(記号)の16類として分類されるが、その中で、<5.>(接頭辞)、<7.>(接尾辞)、<11.>(関係詞)～<14.>(意味不明)の類は析出されていない。これら析出されていない6類以外の9類を全て考察することもできるが、本稿では<5.>(接頭辞)～<16.>(記号)までの付属語については除外し、<1.>(体の類)～<4.>(その他)までを考察の対象とする。

ところで、語彙分析の結果は細かな数字の表になるので、表を見てもどの部分に有意差があるかは容易には分からない。その有意差を分かりやすくするためには、客観的に指摘するため統計的手法に頼らなければならない。すなわち、カイ二乗( $\chi^2$ )検定<sup>3</sup>によって有意差のある部分を検出し、両語彙に関して、有意差の生じた項目を指摘しなければならない。

申（2001:44～46）は、『統計学入門』と『教育と心理のための推計学』を参考にして、意味分野別構造分析のためにカイ二乗検定をどのように行ったのかを詳しく述べている。本稿でのカイ二乗値は、それを参考としている。その分布表を示すと以下の通りである。

●カイ二乗（ $\chi^2$ ）値

$\chi^2 \backslash P$	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.02	0.01	0.005	0.001
1	0.0000	0.0001	0.001	0.004	0.016	2.706	3.841	5.024	5.412	6.635	7.9	10.8

今回は、10%以下の危険率までを有意の差と認めるが、これは、90%以上の確率で有意差があると認めることを意味する。言語研究では一般に危険率10%以下（90%の確率）までも有意の差があると認められるが、心理学や医学などの分野では、危険率5%以下（95%の確率）までが有意の差として認められている。今回の調査においては、危険率10%以下の有意差を対象とし、その対象の項目については網掛けによって表示する。■は1946年が有意に大で、■は2000年が有意に大であることを表す。

今回、「天声人語」の語彙を調査し、カイ二乗検定を行った結果、危険率10%以下で、異なり単位と延べ単位で出現に有意差が生じていると判断される項目を網掛けによって示したのが巻末の【資料1】と【資料2】である。

## 5. 意味分野別構造分析法

意味分野別構造分析法は、先行研究によってコードの利用法が異なる。『分類語彙表』にはコードの番号が4桁までであるので、どこまで細かな分析を行うかによって小数点以下のコードの用い方が異なってくる。

本稿では、単語コード小数点以下1桁と2桁までを用いて意味分野別構造分析を行うことにする。小数点以下3桁と4桁までの分析はあまりにも分類が細かいため、全体的傾向を見るには不適切であると考ええる。

分析の手順は、まず、単語コードの異なり単位と延べ単位の小数点以下1桁までを使った分析により、両年の語彙の全体的傾向を捉えることにする。次に、単語コードの小数点以下2桁までの分析では、どの項目において出現に有意差が生じているかを指摘し、有意差のある項目を取り上げ、分析を試みることにする。

### 5. 1 小数点以下1桁までの分析

巻末の【資料1】からも分かるように、単語コードの小数点以下1桁までの分析は、語彙を視覚的に捉えやすい。これは語彙の全体像を掴むのに適している分析法である。

小数点以下1桁までの分析の結果、カイ二乗検定により10%以下の危険率で出現に有意差が生じていると判断される項目を示すと表1の通りである。

表1

異なり単位		延べ単位	
1946年が大	2000年が大	1946年が大	2000年が大
1.4、1.5、3.3、3.5	1.1	1.2、1.3、1.4、1.5、 2.1、2.3、3.1、3.3、 3.5、4.3	

小数点以下1桁までを使った分析で、異なり単位でも延べ単位でも共通して差が見られるのは、1946年の<1.4> (物品)、<1.5> (刺激)、<3.3> (意識・感覚)、<3.5> (光・音・色) の4項目である。ところが、2000年の場合は、異なり単位でも延べ単位でも、共通して差が見られる項目がなく、異なり単位で<1.1> (こそあど) の1項目のみ有意に大である。

1946年の場合は、表1からも分かるように、延べ単位だけで、<1.2> (われ・なれ・かれ・だれ)、<1.3> (心)、<2.1> (関係)、<2.3> (感覚・疲労・睡眠)、<3.1> (こそあど)、<4.3> (表現態度) などの6項目で有意に大である。2000年に比べて、1946年の方が有意に大の項目が異なり単位でも延べ単位でも多いのは注目に値する。

今回の語彙調査の結果は、【資料1】が示すように、延べ語では2000年の方が1117語多いが、異なり語では1946年の方が416語多い。これらの数字からも分かるように、延べ語数というのは、必ずしも異なり語数に比例して多くはならない。

小数点以下1桁までを使った分析において差が生じている項目の全体的な傾向を捉えてみると、1946年の場合は<1.> (体の類) から<4.> (その他) までの全ての範囲にわたって有意に大の項目が現れているが、2000年の場合は<1.> (体の類) の中で異なり単位の<1.1> (こそあど) のみが有意に大であった。

このように、1946年が有意に大の項目が多いのは、同じ意味分野を表す語が異なり単位においても延べ単位においても、2000年に比べ、1946年の方が語の

使用頻度が高いということによる。

小数点以下1桁までのコードを使った分析結果は【資料1】の通りである。小数点以下1桁までのコードを使った分析は、意味分野を表す意味範疇が広く、それによって集合されている語も多いので、全体的傾向を捉えることに過ぎない面がある。

## 5.2 小数点以下2桁までの分析

小数点以下2桁までの分析の結果、カイ二乗検定により10%以下の危険率で出現に有意差が生じていると判断される項目を示すと表2の通りになる。

表2

異なり単位		延べ単位	
1946年が大	2000年が大	1946年が大	2000年が大
1.25、1.30、1.36、 1.43、 <b>1.51</b> 、1.55、 <b>1.57</b> 、2.12、2.33、 3.12、3.15、3.30、	1.16、1.17、1.19、 1.26、1.28、1.31	1.10、1.11、1.18、 1.23、1.25、1.30、 1.33、1.35、1.36、 1.37、1.38、1.40、 1.42、1.43、1.50、 1.55、2.11、2.12、 2.33、2.34、2.36、 3.12、3.13、3.15、 3.16、3.19、3.30、 3.33、3.34、3.35、 3.50、4.32	1.16、1.17、1.19、 1.22、1.26、1.28、 1.31、3.11

単語コードの小数点以下2桁までを使った分析で、異なり単位でも延べ単位でも共通して差が見られるのは、1946年では<1.25>（公社）、<1.30>（心）、<1.36>（支配・政治・革命）、<1.43>（食料）、<1.55>（生物）、<2.12>（存在）、<2.33>（文化・風俗）、<3.12>（在不在）、<3.15>（変化・動き）、<3.30>（意識・感覚）の10項目である。

2000年の場合は、<1.16>（地位・位置・場合）、<1.17>（空間・場所）、<1.19>（量）、<1.26>（社会）、<1.28>（同盟・団体）、<1.31>（言動）の6項目である。

1946年の<**1.51**>（自然・物体）と<**1.57**>（体）の2項目は、異なり単位では有意差が生じているが、延べ単位では有意差が生じていない。ほとんどの場合、異なり単位での有意差はそのまま延べ単位での有意差に繋がる場合が多いが、そうならない場合もある。その原因が何にあるか注目する必要がある。

ところで、今回の分析は、1946年が異なり単位においても延べ単位においても有意に大である10項目のみを対象としている。2000年が大である6項目と、1946年の<1.51>と<1.57>の2項目については、稿を改めて論じることにする。

以下、1946年が異なり単位においても延べ単位においても有意に大である10項目を取り上げ、有意差の指摘および語彙の比較を試みる。

(1)<1.25> (公社) の項目の場合、【資料2】からも分かるように、異なり語では1946年149語、2000年107語で、延べ語では1946年377語、2000年218語となっており、1946年が有意に大である。<1.25> (公社) の項目に含まれている語を挙げてみると表3の通りである。

表3

1946年	2000年
日本 (92)、国 (21)、国家 (11)、連合国・アメリカ・米国 (9)、欧米・農村・日 (7)、ソ連 (6)、中国・隣邦・内地・都市 (5)、国際・都会・東京・米 (4)、ドイツ・英・英国・京都・首都・満州・外国・支那 (3)、世界国家・戦勝国・本土・露、印度・都・信州・青幫・長崎・奈良・松本・広島・東西 (2)、後進国・大帝國・帝國・敵国・列国・安慶・伊那・長野など (1)、以下省略	日本 (22)、国 (19)、個人 (8)、韓国・東京 (7)、平壤・米国・町 (4)、フランス・米、アメリカ・英国・欧米・地元・小千谷・新潟県・京都 (3)、故郷・北朝鮮・土建国家・三重県・東京都・芦浜・有珠山・欧州・大原・沖縄・関東地方・サラエボ・滋賀県・西欧・ソウル・中国・朝鮮・ドイツ・東海村・西アフリカ・バリ・フィンランド・レニングラード・ロシア・ロンドン (2)、アジア・アフリカ・イギリス・エチオピア・愛知県・岩手など(1)、以下省略

( ) の中の数字は該当する語の使用回数である。〔以下同様〕

<1.25> (公社) の項目のカイ二乗値を比較してみよう。

異なり単位のカイ二乗値は3.583である。これは危険率10%以下、つまり90%の確率で1946年が有意に大であることを表す。一方、延べ単位のカイ二乗値は52.796である。これは危険率0.1%以下、つまり、ほぼ99.9%の確率で1946年が有意に大であることを表している。

カイ二乗値を比較した結果、1946年が有意に大であり、このことから、異なり単位に比べ、延べ単位の方が大きいことが分かる。特に、延べ単位におけるカイ二乗値が大きいのは、表3の<1.25> (公社) の項目に含まれている語からも分かるように、個別語の使用回数が多いためである。要するに、「日本(92)、国(21)、国家(11)、連合国・アメリカ・米国(9)など」のように、個別語



の使用頻度が非常に高いからである。

(2)<1.30> (心) の項目の場合、異なり語では1946年340語、2000年257語で、延べ語では1946年600語、2000年543語となっており、1946年が有意に大である。<1.30> (心) の項目に含まれている語を挙げてみると表4の通りである。

表4

1946年	2000年
大問題 (18)、意義 (12)、感動・お疲れ・注意・原則 (11)、志望・侵略主義 (9)、建国精神 (8)、心理・気分・想像 (7)、実用新案 (6)、悪・善悪・一隻眼・社会主義運動・初耳 (5)、大和心・感激・愛国心・精義・疑い・検閲・はず、筆法・方法・時風 (4)、驚嘆・ご不満・同胞愛・人間愛・敬意・期待など (3)、以下省略	問題 (23)、責任 (20)、そう (15)、民主主義 (13)、意味 (11)、はず (10)、気・意見 (9)、心・思い・調査・入試 (8)、大和魂・気持ち・例・計画 (7)、感じ・声・義理・経験・試験・結論 (6)、予定 (5)、好み・自由・理解・選考・科学・法 (4)、感情・気分・敬意・信頼・期待など (3)、以下省略

2000年の項目にある「そう (15)」は、伝聞を表す語である。

<1.30> (心) の項目の場合、異なり単位のカイ二乗値は5.367で、延べ単位のカイ二乗値は7.195である。異なり単位においても延べ単位においても、危険率5%以下 (95%の確率) で1946年が有意に大である。カイ二乗値を比較してみると、異なり単位に比べて延べ単位の方が少し大きいことが分かる。

<1.30> (心) の項目の場合、表4からも分かるように、1946年では「侵略主義・建国精神・社会主義運動・愛国心・同胞愛」などの語が多く見られ、2000年では「民主主義・意見・入試・試験・自由・科学」などの語が多く見られる。これらの単語から、時代の異なる両年の語彙は、それぞれ時代的背景を反映している語で構成されていることが窺がえる。

(3)<1.36> (支配・政治・革命) の項目の場合、異なり語では1946年87語、2000年57語で、延べ語では1946年145語、2000年102語となっており、1946年が有意に大である。この項目に含まれている語を挙げてみると表5の通りである。

表 5

1946年	2000年
諸政 (10)、悪性 (6)、救援・満州事変 (5)、北清事変・教育・募集 (4)、管轄・新任・立候補・一殺・遵守・指令・措置令・受賞 (3)、支那事変・政治・無血・弾劾・奨励・育児・援護・順法う・奨学・命令・強制 (2)、議会政治・建国・食糧デモ・ファッション政治・治安維持・節約令・追放令・援助・救済・助けなど (1)、以下省略	育児・陳謝 (8)、投票 (5)、政治・落選 (4)、判決・人事・選挙・しつけ・教育 (3)、管理・帝政・実刑・監禁・経営・総選挙・子育て・申請・要請・規制・注文 (2)、いじめ・称賛・激励・声援・救護・介護・看護・アドバイザー・お手伝い・指導・運営・地方選挙・授業・教育上・戒めなど (1)、以下省略

<1.36> (支配・政治・革命) の項目のカイ二乗値を比較してみよう。

異なり単位のカイ二乗値は3.788である。これは危険率10%以下(90%の確率)で1946年が有意に大であることを表す。一方、延べ単位のカイ二乗値は10.187である。これは危険率1%以下(99%の確率)で1946年が有意に大であることを表している。

<1.36> (支配・政治・革命) の項目の場合、1946年では「満州事変・北清事変・支那事変・建国・食糧デモ・ファッション政治・治安維持・節約令・追放令・援助・救済・助け」などの語が多く見られ、2000年では「育児・しつけ・教育・子育て・授業・教育上・いじめ・称賛・激励・声援・救護・介護・看護・アドバイザー・お手伝い」などの語が多く見られる。この項目の場合も、<1.30> (心) の項目と同様に、両年において時代的背景を反映している語が多く使われていることが分かる。

(4)<1.43> (食料) の項目の場合、異なり語では1946年33語、2000年10語で、延べ語では1946年79語、2000年12語となっており、1946年が有意に大である。<1.43> (食料) の項目に含まれている語を挙げてみると表6の通りである。

表 6

1946年	2000年
食糧 (28)、米 (14)、主食 (4)、パン・醤油・味噌・お茶・豆腐 (2)、豆腐・漬物・塩麴・清香・粥・握り飯・肉・ハム・バター・飼料・穀物・出し・チューインガム・御馳走・食べ物・配給米・新米・塩・煙草・DDT・応急薬・ペニシリン・味噌汁 (1)	チョコレート (5)、食物・キャンデー・クッキー・ワイン、特効薬・麻薬・ミルク (1)

<1.43>（食料）の項目の場合、異なり単位のカイ二乗値は10.275で、延べ単位のカイ二乗値は53.355である。異なり単位においても延べ単位においても、危険率1%以下（99%の確率）で1946年が有意に大である。1946年のカイ二乗値を比較してみると、異なり単位に比べ、延べ単位の方が非常に大きいことが分かる。これは、「食糧・米・主食・パン・醤油・味噌・お茶・豆腐」などの個別語の使用頻度が高いためである。表6からも分かるように、1946年では「食べ物」に関する語が非常に多い。これは、1946年が大戦直後の「食糧難」の只中の年であったことを如実に反映している。

(5)<1.55>（生物）の項目の場合、異なり語では1946年53語、2000年19語で、延べ語では1946年97語、2000年28語となっており、1946年が有意に大である。<1.55>（生物）の項目に含まれている語を挙げてみると次の表7の通りである。

表7

1946年	2000年
花（20）、桜（6）、朝顔（5）、南瓜（4）、梅（3）、茄子・麦・ツツジ・満州大豆・蘭・新芽・種蒔・根（2）、盆栽・胡瓜・木・木々・キャベツ・夏草・栗・櫟・サヤエンドウ・麦・ジャガイモなど（1）、以下省略	生物（4）、花（3）、男女・男性・お花・種（2）、雌雄・植生・草・若草・栗・桜・ツタ・バラ・豆・桃・実・草花（1）

<1.55>（生物）の項目の場合、異なり単位のカイ二乗値は13.112で、延べ単位のカイ二乗値は42.256である。異なり単位においても延べ単位においても、危険率0.1%以下（99.9%の確率）で1946年が有意に大である。カイ二乗値は異なり単位に比べ、延べ単位の方が非常に大きいことが分かる。

<1.55>（生物）の項目において、特に延べ単位におけるカイ二乗値が大きいのは、表7からも分かるように1946年のこの項目では「花・木・草など」の個別語の使用頻度が高いためである。これらの単語が多く用いられているのは、大戦直後の日本における食糧事情と関係がある。食べ物がなかったこの時代は、花や草などが観賞用ではなく「食糧」だったのである。

(6)<2.12>（存在）の項目の場合、異なり語では1946年63語、2000年40語で、延べ語では1946年373語、2000年345語となっており、1946年が有意に大である。<2.12>（存在）の項目に含まれている語を挙げてみると表8の通りである。

表 8

1946年	2000年
出来る (122)、有る (109)、生じる (18)、呈する (15)、残す (9)、絶える・無くなる (8)、在る (7)、依存する (6)、隠れる (4)、暴く・無くする・投げ捨てる・排する (3)、居る・体现する・登場する・起こす・対立する・抹殺する・擲つ (2)、淘汰する・駆除する・一掃する・要する・除く・取る・捨てる・消滅する・遺失する・亡びる・亡び尽くす・失う・尽す・消すなど (1)、以下省略	成る (119)、有る (97)、出来る (37)、居る (17)、出る (12)、起こる (10)、残る (9)、登場する・消える (3)、ござる・現れる・提出する・起きる・残す・成立する・保つ (2)、控える・出現する・表す・実現する・起こす・生じる・出来上がる・廃止する・廃業する・失う・失せる・消えかける・混在する・絶える・無くなる・滞在するなど (1)、以下省略

<2.12> (存在) の項目の場合、異なり単位のカイ二乗値は3.222で、延べ単位のカイ二乗値は3.313である。異なり単位においても延べ単位においても、危険率10%以下 (90%の確率) で1946年が有意に大である。

この項目は、1946年と2000年との間に語の数量的違いはあるものの、「存在・出現・消滅」などほぼ同様の語で構成されており、時代を反映する特徴的な語はそこに見られなかった。

(7)<2.33> (文化・風俗) の項目の場合、異なり語では1946年57語、2000年34語で、延べ語では1946年74語、2000年52語となっており、1946年が有意に大である。この項目に含まれている語を挙げると表9の通りである。

表 9

1946年	2000年
食いつぶす (9)、執る・拾う (3)、はやりだす・勤める・売り食いする・同居する・味わう (2)、崇る・開ける・処する・禍する・転職する・歴任する・怠る・在官する・従事する・携わる・勤まる・働く・居着く・着る・食う・食べる・暮らす・食べる・着ける・脱ぐ・安住する・亡命する・嫁する・蒙る・埋蔵する・遊ぶ・欣喜雀躍する・つかむ・握る・飲むなど (1)、以下省略	働く (5)、住む (4)、祈る・歩く・摘む (3)、引退する・応募する・暮らす・歩む・拾う (2)、被爆する・就職する・退職する・転職する・怠る・出張する・携わる・移り住む・生活する・食べる・召す・棲息する・留学する・合格する・祝う・遊ぶ・踊る・抱く・担ぐ・しゃがむ・這う・繰る・握り合う・持つなど (1)

<2.33> (文化・風俗) の項目のカイ二乗値を比較してみると、異なり単位のカイ二乗値は3.385である。異なり単位では危険率10%以下、つまり90%の確

率で1946年が有意に大である。一方、延べ単位のカイ二乗値は5.207である。これは危険率5%以下（95%の確率）で1946年が有意に大であることを表している。

<2.33>（文化・風俗）の項目の場合、1946年では「食いつぶす・食う・食べる・売り食いする・飲むなど」の語が多く、2000では、「働く・住む・就職する・出張する・留学するなど」の語が多い。表9に挙げられている動詞の例から、漠然とではあるが、両語彙間に時代的な違いは窺い知ることができる。

(8)<3.12>（在不在）の項目の場合、異なり語では1946年37語、2000年21語で、延べ語では1946年97語、2000年61語となっており、1946年が有意に大である。この項目に含まれている語を挙げてみると表10の通りである。

表10

1946年	2000年
とんと（37）、空虚（6）、無い（5）、なかる（4）、虚ろ（3）、ありそうだ・決して・断じて・珍しい・必ずしも・絶対に・必然的・無用に・捨て難い・脱げ難い・無理だ（2）、一向・がち（曇り～）・ちっとも・無き・亡き・稀に・必ず・仮に・緊急・偶然・特高・必要だ・及び難い・期し難い・着き易い・易しい・難しい・空しい・無理・やむをえないなど（1）	無い（23）、難しい（5）、決して（4）、がち（曇り～）・珍しい・必ず・仮に（3）、どうしても・必ずしも・絶対・（2）、空しい・空（あき）、一向・緊急・不要だ・無理に・余儀ない・解決しにくい・可能だ・至難・容易に（1）

<3.12>（在不在）の項目のカイ二乗値を比較してみると、異なり単位のカイ二乗値は3.045である。異なり単位では危険率10%以下、つまり、90%の確率で1946年が有意に大である。一方、延べ単位のカイ二乗値は10.414である。これは危険率1%以下、つまり、ほぼ99%の確率で1946年が有意に大であることを表している。

この項目における有意差は、表10からも分かるように、1946年では異なり語が多く、その使用頻度も高い。特に「～難い」という語が多く用いられている。

2000年では「～難い」という語がなく「解決しにくい」という語だけがある。このことから、「～し難い」という表現が「～しにくい」という表現に交替しているのではないかと考えられるけれども、今回の語彙調査では量的に語数が少ないこともあって断定は出来ない。

(9)<3.15> (変化・動き) の項目の場合、異なり語では1946年22語、2000年8語で、延べ語では1946年28語、2000年11語となっており、1946年が有意に大である。この項目に含まれている語を挙げてみると表11の通りである。

表11

1946年	2000年
おのずと・積極的だ・めっきり・ぱっと・どっしり・圧倒的(2)、進歩的・活発に・盛んだ・自然に・積極的・滔々・不安だ・どっと・どンドン・ひしひし・自発的・能動的だ・スクスク・ビツタリ・圧倒的だ・退廃的だなど	積極的(3)、ふらふら(2)、変幻自在だ・放題・ぱっと、繰り返されそうだ・ばらばら・圧倒的など(1)

<3.15> (変化・動き) の項目の場合、異なり単位のカイ二乗値は5.291である。異なり単位では危険率5%以下(95%の確率)で1946年が有意に大である。一方、延べ単位のカイ二乗値は8.416である。これは危険率1%以下(99%の確率)で1946年が有意に大であることを表している。

表11からも分かるように、1946年が有意に大であるのは、2000年に比べ、1946年の方が「変化・動き」を表す語が多く用いられているためである。

(10)<3.30> (意識・感覚) の項目の場合、異なり語では、1946年110語、2000年70語で、延べ語では、1946年156語、2000年130語となっており、1946年が有意に大である。この項目に含まれている語を挙げてみると表12の通りである。

表12

1946年	2000年
無念だ(15)、確かに(9)、学問的だ(7)、たい(4)、優生学的(3)、しみじみ・嬉しい・愉快・痛ましい・可憐だ・賢明だ・道徳的・未経験だ・確か・不思議・合理的だ・真に・非人道(2)、明らかに・精神的・痛切だ・つくづく・敏感・不潔感・無我愛・無考え・無関心・官能的・面白い・窮屈・気楽に・快い・呑気だ・不快・喜ばしい・楽だ・がっかり・悲しい・寂しい・心痛・辛い・悲惨だ・悲痛・惨めだ・遺憾だ・恥ずかしい用心深い・意味深いなど(1)、以下省略	たい(29)、明らかに・辛い・嬉しい・確かに(4)、怖い・残念・不思議・無残だ(3)、明らかだ・痛い・ふと(意識)・ほっと・恐ろしい・重苦しい・寂しい・馬鹿だ・上手い・無能・曖昧・行方不明(2)、うっかり・かつと・しみじみ・痛烈・つくづく・印象深い・味わい深い・面白い・楽しい・のんびり(楽しく)・悠悠と・うんざり・おどろおどろしい・悲しい・苦しい・悲劇的・悲痛・いらいら・残念だなど(1)、以下省略

<3.30>（意識・感覚）の項目の場合、異なり単位のカイ二乗値は5.614で、延べ単位のカイ二乗値は4.096である。異なり単位においても延べ単位においても、危険率5%以下（95%の確率）で1946年が有意に大である。

この項目の場合、表12からも分かるように、1946年では「～的だ」「～的」などの語が多いが、2000年では「悲劇的」の1語だけである。語の使い方は記事を書く人のスタイルにも関係するであろう。しかし、「～的」表現が、1946年に比べ、2000年で特に少ないのは、個別語の移り変わりに関係しているのではないだろうか。

## 6. おわりに

以上、1946年と2000年の「天声人語」を対象として、単語コード小数点以下1桁と2桁までのコードを用い、両年の語彙を分析した。今回の分析では、危険率10%以下（90%の確率）で有意差が生じた項目、その中でも特に、1946年が有意に大である項目だけを取り上げて、その有意差の指摘と項目別語の比較を試みた。その結果、2つの時期の「天声人語」の語彙に現れた大まかな傾向として、1946年の方が延べ単位においても異なり単位においても、有意に大の項目が多かった。

1946年が有意に大であるということは、1946年の語彙が比較対象とする2000年の語彙に比べてより豊富であることを意味する。要するに、小数点以下2桁までの分析の結果、有意に大であった1946年の<1.25>（公社）～<3.30>（意識・感覚）の10項目は、2000年の同じ項目に比べ、用いられている語の数が多いということである。

しかし、今回の分析は、小数点以下1桁と2桁までを取り上げたので、語彙の違いに関する細かな原因を究明するまでには至っていない。1946年が出現に有意に大である意味分野を指摘し、その傾向を明らかにする程度にとどまっている。

今後の分析では、2000年が有意に大である6項目と、1946年の<1.51>と<1.57>の2項目を取り上げなければならない。さらに、有意差が生じていない項目にも焦点を当て、有意差が生じない項目での語彙の違いも考察する必要がある。

語彙を調査し分析するには、『分類語彙表』が小数点以下4桁までであるので、4桁までの分析が可能である。今回は、小数点以下2桁までの分析を行ったが、今後、3桁と4桁までの分析法による考察も必要であろう。また、研究対象を「天

声人語」だけではなく、他のジャンルにまで広げ、現代日本語の語彙を総合的に考察する必要がある。

## 注

- 1 ごい [語彙]：ある範囲（言語体系・分野・作者・作品など）で使われる単語の全体。また、その一つ一つを集めたもの。ポキャブラリー。学習研究社（1988）『学研国語大辞典（第二版）』
- 2 辞書とは別に、語の検索に用いられる語彙表のようなものがある。「シソーラス」‘thesaurus’と呼ばれるのがその代表と言える。シソーラスには広狭2義があるが、狭義には類語・関連語辞典を指す。英語について、PeterMark Rogetの“Roget's Thesaurus”(Thesaurus of English Words and Phrases 1852年刊)が有名である。英語の単語を‘Abstract Relations’‘Space’‘Matter’‘Intellect’‘Volition’‘Affections’の6つのカテゴリーに分類したものである。各語は分類番号を付された類概念の下に分類されていて、同一範疇の語、同一レベルの語は同一番号の下にグルーピングされている。日本語について、林大氏がまとめられた『分類語彙表』（国立国語研究所資料集6 1964年刊）がある。この『分類語彙表』は、雑誌90種に用いられた高使用率の約7,000語を中心にし、他の語彙資料をも参照して得られた計32,600語を、名詞類・動詞類・形容詞類・その他の4類に分かち、分類コードを立てて分類したものである。国立国語研究所（1985）「日本語教育指導参考書13」『語彙の研究と教育（下）』（p.143~145）
- 3 カイ二乗検定は、規模の異なるA、B二つの標本において、特定性質をもつものがA標本ではa例、B標本ではb例あり、特定性質のないものがA標本ではC例、B標本ではd例あったとした場合、それぞれの標本の規模を考慮に入れても、aとbとで差が認められるかどうかは分からないので、それを検定する統計技法である。

## 参考文献

- 浅見徹 1971「古代の語彙Ⅱ」『国語史3 語彙史』大修館書店  
 金岡孝 1977「語彙研究の歴史」『岩波講座日本語9』岩波書店  
 樺島忠夫1989「語彙の計量」『講座日本語と日本語教育第6巻 日本語の語彙・



意味（上）』明治書院

- 阪倉篤義1971「語彙史の方法」『講座国語史 第3巻 語彙史』大修館書店
- 佐竹昭弘 1956「語彙の構造と思考の形態」『国語学』27 武蔵野書院
- 国立国語研究所 1962・1963・1964『現代雑誌九十種の用語用字』国立国語研究所報告21・22・23
- 国立国語研究所 1964「国立国語研究所資料集6」『分類語彙表』
- 国立国語研究所 1985「日本語教育指導参考書13」『語彙の研究と教育（下）』
- 国立国語研究所 1987『雑誌用語の変遷』国立国語研究所報告89
- 田島毓堂 1992 a「語彙論の課題——集团的規範と個別的実現——」『名古屋大学国語国文学』71
- 田島毓堂 1992 b「語彙論的語の単位試論——意味単位と分類単位——」『日本語論究2 古典日本語と辞書』和泉書店
- 田島毓堂 1995 a「比較語彙論の構想——異文化比較のために——」『国際開発研究フォーラム』2
- 田島毓堂 1995 b『意味コードの付け方——語彙比較研究のために——』開発・文化叢書14
- 田島毓堂篇 1997『比較語彙研究の試み』開発・文化叢書21
- 田島毓堂 1999『比較語彙研究序説』笠間書院
- 田島毓堂篇 2000『語彙研究の試み6』開発・文化叢書36名古屋大学大学院国際開発研究科
- 申 揖 2001「日韓語彙の比較研究—「小学生基本語彙」を対象としての試み」（『比較語彙研究の試み7』）開発・文化叢書37名古屋大学大学院国際開発研究科
- 水谷静夫 1965「語彙論の述語をめぐって」『国語学』62 武蔵野書院
- 室山敏昭 1986「語彙の社会言語学的研究」『日本語学』（Vol.5）明治書院
- 学習研究社 1988『学研国語大辞典（第二版）』

## 【資料1】 小数点以下1桁までの意味分野

意味範疇	コード	異なり単位			延べ単位		
		1946年	2000年	$\gamma$ 自乗値	1946年	2000年	$\gamma$ 自乗値
こそあど	1.1	771	906	34.126	1787	1973	1.866
われ・なれ・かれ・だれ	1.2	765	701	0.027	1520	1464	7.152
心	1.3	882	748	2.405	1503	1462	5.690
物品	1.4	193	146	2.928	315	239	15.396
刺激	1.5	229	155	8.474	374	320	7.957
関係	2.1	418	348	1.612	1236	1162	8.937
感覚・疲労・睡眠	2.3	481	443	0.043	1093	1052	5.100
光・音・色	2.5	45	46	0.315	70	92	1.882
こそあど	3.1	395	319	2.697	1088	921	26.294
意識・感覚	3.3	210	151	5.062	317	240	15.679
光・音・色	3.5	38	22	2.902	52	31	6.586
接続	4.1	27	34	1.620	581	586	0.693
表現態度	4.3	33	29	0.018	131	97	7.223
接中辞	6	0	1	1.081	0	3	2.831
助詞	8	43	44	0.306	6406	5409	214.183
助動詞	9	17	16	0.010	1137	1318	5.410
補助動詞	10	19	22	0.600	706	464	66.708
固有名詞	15	4	2	0.486	5	2	1.461
記号	16	8	29	14.082	866	3469	1,595.492
	合計	4578	4162		19187	20304	

## 【資料2】小数点以下2桁までの意味分野

意味範疇	コード	なり単位			延べ単位		
		1946年	2000年	$\gamma$ 自乗値	1946年	2000年	$\gamma$ 自乗値
こそあど	1.10	40	35	0.027	322	266	9.114
類・例	1.11	49	38	0.546	118	92	4.887
有無	1.12	31	22	0.796	46	40	0.829
様相	1.13	51	52	0.342	81	75	0.698
力	1.14	12	20	2.845	30	33	0.024
作用	1.15	86	80	0.022	120	108	1.502
位置・地位・場合	1.16	166	202	8.140	397	492	5.620
空間・場所	1.17	76	108	9.239	167	227	6.124
形・型・姿・構え	1.18	29	18	1.643	49	36	2.799
量	1.19	231	331	30.616	457	604	13.266
われ・なれ・かれ・だれ	1.20	85	71	0.282	264	295	0.419
家族	1.21	34	36	0.409	64	71	0.075
相手・仲間	1.22	12	15	0.681	16	51	16.396
人種・民族	1.23	223	176	2.064	390	263	32.977
成員・職	1.24	121	111	0.005	176	216	2.155
公私	1.25	149	107	3.583	377	218	52.796
社会	1.26	60	101	15.013	106	178	14.523
機関	1.27	48	41	0.086	85	80	0.569
同盟・団体	1.28	32	43	2.858	43	92	15.184
心	1.30	340	257	5.367	600	543	7.195
言動	1.31	115	151	9.199	210	365	33.995
創作・著述	1.32	34	28	0.150	52	48	0.468
文化・歴史・風俗	1.33	79	72	0.000	141	116	4.081
義務	1.34	48	47	0.132	68	78	0.237
交わり	1.35	66	51	0.771	98	77	3.868
支配・政治・革命	1.36	87	57	3.788	145	102	10.187
納得	1.37	65	48	1.212	108	75	8.007
仕事	1.38	48	37	0.574	81	58	5.240
物品	1.40	13	6	1.958	24	10	6.593
資材	1.41	21	25	0.837	24	31	0.540
衣料・錦・皮・糸	1.42	21	12	1.679	25	14	3.761
食料	1.43	33	10	10.275	79	12	53.355
住居	1.44	21	26	1.120	36	41	0.104
道具	1.45	39	26	1.522	52	47	0.616
燈火	1.46	26	28	0.389	46	54	0.268
地類（土地利用）	1.47	19	13	0.627	29	30	0.008
刺激	1.50	8	11	0.802	34	21	3.860
自然・物体	1.51	48	27	4.091	48	56	0.247
宇宙・空	1.52	32	27	0.081	55	53	0.237
生物	1.55	53	19	13.112	97	28	42.256
動物	1.56	18	12	0.698	38	56	2.511
体	1.57	54	33	3.304	74	71	0.349
生命	1.58	18	26	2.329	28	35	0.433
関係	2.11	40	30	0.640	368	257	26.940
存在	2.12	63	40	3.222	373	345	3.313
整備	2.13	18	10	1.592	19	17	0.253

力	2.14	1	0	0.892	2	0	2.112
変化	2.15	265	243	0.010	422	493	2.279
時間・時刻	2.16	8	10	0.452	17	22	0.390
位置・方向	2.17	4	7	1.127	6	10	0.787
形	2.18	2	0	1.801	2	0	2.112
過不足・優劣など	2.19	17	8	2.447	27	18	2.349
感覚・疲労・睡眠	2.30	190	173	0.000	366	347	2.193
呼び・名	2.31	57	69	2.612	253	293	1.121
創作	2.32	4	6	0.611	5	8	0.533
文化・風俗	2.33	57	34	3.875	74	52	5.207
行為	2.34	17	16	0.010	107	80	5.606
交わり	2.35	35	39	0.771	56	68	0.584
支配・統治	2.36	54	39	1.216	84	54	8.364
所有・取得	2.37	34	40	1.236	88	85	0.362
農工業・設備・処理	2.38	33	27	0.165	60	65	0.017
光・音・色	2.50	13	10	0.157	15	17	0.037
煙・乾湿	2.51	9	12	0.762	10	17	1.442
生・生育	2.58	23	24	0.223	45	58	0.991
こそあど	3.10	42	33	0.372	210	214	0.152
関係	3.11	27	32	1.040	59	92	5.491
在不在	3.12	37	21	3.045	97	61	10.414
簡潔	3.13	61	49	0.421	193	171	2.894
力	3.14	6	5	0.020	7	5	0.456
変化・動き	3.15	22	8	5.291	28	11	8.416
時	3.16	79	65	0.360	245	153	27.083
位置・場所	3.17	4	2	0.486	5	2	1.461
形	3.18	5	3	0.325	5	3	0.619
程度	3.19	112	101	0.004	239	209	4.114
意識・感覚	3.30	110	70	5.614	156	130	4.096
ことば	3.31	8	4	0.978	9	4	2.217
風俗	3.33	18	14	0.191	23	14	2.732
身上	3.34	43	31	0.980	66	41	7.366
交渉・交際	3.35	5	1	2.295	8	1	5.850
公式・公平	3.36	13	17	0.984	25	27	0.005
経済	3.37	12	14	0.403	29	23	1.075
光・音・色	3.50	27	18	1.050	41	24	5.472
火・水・気象	3.51	3	0	2.711	3	0	3.171
体	3.57	1	0	0.892	1	0	1.054
生育	3.58	7	4	0.555	7	7	0.011
接続	4.11	27	34	1.620	581	586	0.693
表現態度	4.31	27	22	0.145	61	51	1.553
呼びかけ応答	4.32	4	4	0.017	68	43	7.158
あいさつ	4.33	2	3	0.303	2	3	0.147
接中詞	6.00	0	1	1.081	0	3	2.831
助詞	8.00	43	44	0.306	6406	5409	214.183
助動詞	9.00	17	16	0.010	1137	1318	5.410
補助動詞	10.00	19	22	0.607	706	464	66.708
名詞一固有名詞	15.00	4	2	0.486	5	2	1.461
記号	16.00	8	29	14.082	866	3469	1,595.492
合計		4578	4162		19187	20304	