『中央公論』１０１年の語彙

入江（いりえ）　さやか

同志社大学文学部

キーワード

語彙調査，語種，のべ，異なり，類似度

１．はじめに

本調査は，国立国語研究所の調査である『雑誌用語の変遷』のあとを継ぐものである。『中央公論』1906 年から 2006 年までの 101 年間で，日本語がどのような変化をたどったのか，語種別の異なり語数，のべ語数，年度間の類似度などを中心に述べたい。

２．標本のぬきだしと調査単位

標本ぬきだしの単位は，1/10 ページである。以下これを「ブロック」とよぶ。ブロックの決定は行数によっている。１ブロックの言語量は，年によって異なるため，のべ 10,000 語の標本をとるために必要なブロックの数も年によって違う。１年を通じての合計ページ数も異なるので，抽出比も一定していない。以下，各年の抽出比である。

〈表１〉

	総ページ	抽出ブロック数	抽出比
1906	1,890	466	1/41
1916	4,170	462	1/90
1926	6,344	452	1/140
1936	7,016	421	1/167
1946	1,460	310	1/47
1956	4,360	337	1/129
1966	5,268	359	1/147
1976	5,090	379	1/134
1986	4,812	403	1/119
1996	4,960	418	1/119
2006	4,402	454	1/97
計	49,772	4,461	1/112

単位語の認定についても,『雑誌用語の変遷』にしたがい,〈ながい単位〉を採用した。

３．『中央公論』１０１年の語彙

3.1　各年度の度数分布

のべ語数は，各年度とも 10,000 語で，同じである。101 年間で，11 年度の調査を行い，合計

110, 000 語となる。異なり語数は, 年によって多少の差は見られるが, それほど大きな差はない。各年度の異なり語数を〈表２〉に示す。表の中段が異なり語数，下段は，平均使用度数である。

〈表２〉

年度	1906	1916	1926	1936	1946	1956	1966	1976	1986	1996	2006	計
異り	4483	4413	4536	4750	4439	4341	4649	4936	4795	4826	4959	30241
平均	2.23	2.27	2.20	2.11	2.25	2.30	2.15	2.03	2.09	2.07	2.02	3.64

『雑誌用語の変遷』では，1976 年までの調査で，異なり語数が 23161，平均使用度数が 3.45 だった。今回の継続調査で, 101 年間における異なり語数は 30241, 平均使用度数は 3.64 となり, 異なり語数は 7080 増えたが，平均使用度数はやや高くなっていはいるが，あまり変わらないことがわかった。

次に，異なり語数 30241 語のうち，複数年度に出現する語，ある特定の年度にしか出現しない語の数についてみてみる。上段が何年分に出てきたかを，中段が語数，下段は割合を示す。小数第二位を四捨五入している。

〈表３〉

年数	1	2	3	4	5	6	7	8	9	10	11	計
語数	22929	3327	1360	764	497	354	237	213	189	169	292	30241
割合	75.8	11.0	4.5	2.5	1.6	1.2	0.8	0.7	0.6	0.6	1.0	100.0

異なり語数 30241 語の 75.8%が, ある特定の１年の資料にしか出現していない。理由としては, 調査単位として〈ながい単位〉をとったことと，年ごとに 10, 000 語しかとっていないという調査資料の少なさが挙げられる。

『雑誌用語の変遷』では, 特定の１年の資料にしか出現しない語の率は, 76.5%であったから, ごくわずかではあるが，３年分を追加したことで，特定の年にしか出現しない語の割合は下がった。

3.2　高頻出語

『中央公論』における度数の高い語は次のとおりである。合計が上位 12 の語をあげたが，各年度の上位 10 の語はすべて入っている。『雑誌用語の変遷』の 1976 年までの上位語の合計をした調査と異なったのは，「いる（居る)」が４位から２位へ上がったこと，「なる（成る)」と「もの」の順位が入れ替わり，「なる」が８位に上がったことである。

〈表４〉

		1906	1916	1926	1936	1946	1956	1966	1976	1986	1996	2006	計
1	ある	312	353	303	374	372	317	289	277	291	248	243	3379
2	いる	56	159	201	188	207	244	241	249	245	298	265	2353
3	する	245	241	213	210	206	167	177	174	177	187	195	2192
4	いう	146	170	172	179	132	310	249	207	195	202	180	2142
5	こと	134	166	176	181	201	193	183	150	178	174	156	1892
6	ない	137	134	123	138	129	150	137	109	97	102	113	1369
7	その	141	135	125	126	144	91	92	93	110	89	88	1234
8	なる	72	95	103	103	118	139	125	113	127	114	120	1229
9	もの	156	123	125	124	107	93	90	85	93	53	64	1113
10	この	71	73	64	68	99	73	83	75	77	82	48	813
11	それ	33	61	94	66	79	77	76	82	71	75	67	781
12	これ	114	71	61	60	76	56	44	42	57	30	32	643

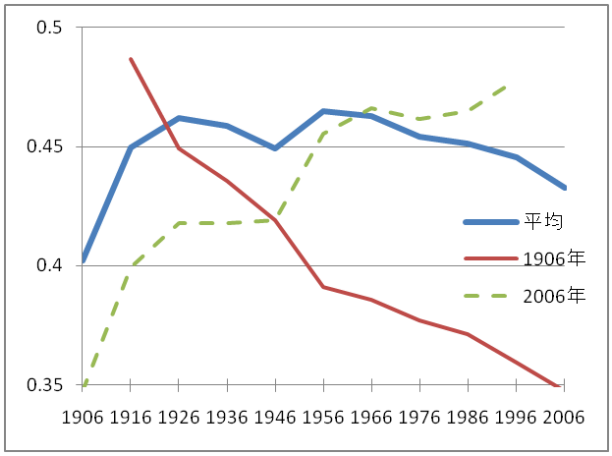
3.3　各年度の類似度

ここでは，それぞれの年の用語は，ほかの年とどの程度似ているかを数字で示す。類似度の示し方は，大きくない方（より小さい方，同じ場合にはその使用率）を合計する，という方法である。このようにして，年ごとの類似度を計算したのが，次の表である。

〈表５〉

1906	1916	1926	1936	1946	1956	1966	1976	1986	1996	2006	
	0.4868	0.4492	0.4357	0.4192	0.3910	0.3857	0.3770	0.3712	0.3593	0.3472	1906
0.4868		0.5095	0.4864	0.4614	0.4483	0.4379	0.4311	0.4230	0.4140	0.3995	1916
0.4492	0.5095		0.5002	0.4670	0.4718	0.4595	0.4589	0.4407	0.4441	0.4178	1926
0.4357	0.4864	0.5002		0.4723	0.4742	0.4612	0.4543	0.4451	0.4404	0.4177	1936
0.4192	0.4614	0.4670	0.4723		0.4695	0.4616	0.4439	0.4483	0.4315	0.4190	1946
0.3910	0.4483	0.4718	0.4742	0.4695		0.5043	0.4845	0.4800	0.4678	0.4553	1956
0.3857	0.4379	0.4595	0.4612	0.4616	0.5043		0.4843	0.4909	0.4750	0.4659	1966
0.3770	0.4311	0.4589	0.4543	0.4439	0.4845	0.4843		0.4756	0.4715	0.4614	1976
0.3712	0.4230	0.4407	0.4451	0.4483	0.4800	0.4909	0.4756		0.4738	0.4647	1986
0.3593	0.4140	0.4441	0.4404	0.4315	0.4678	0.4750	0.4715	0.4738		0.4781	1996
0.3472	0.3995	0.4178	0.4177	0.4190	0.4553	0.4659	0.4614	0.4647	0.4781		2006
（平	均）										
0.4022	0.4498	0.4619	0.4588	0.4494	0.4647	0.4626	0.4543	0.4513	0.4456	0.4327	

〈グラフ１〉



類似度は，近くの年とは高く，離れた年とは,低くなると予想される。1906年は，まさにその通りで，1916年が一番高く，そのあとは， 完全に右肩下がりである。

2006年は，特異な年である1946年のところで，類似度は低くなるが，その後，急激に高くなり，1976年でやや類似度は低くなるが，1986年から再び少しずつ高くなる。

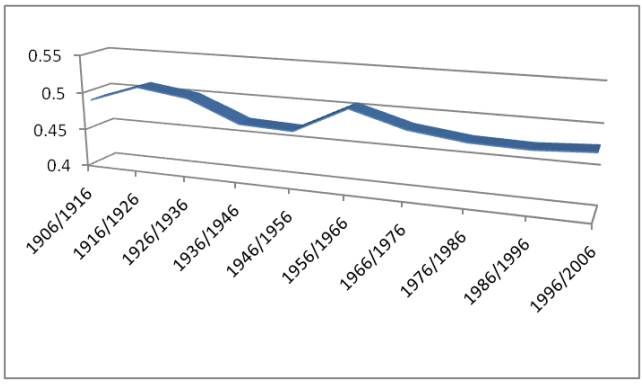
3.4　語彙の共通度

類似度のほかに共通度の調査も行う。共通度の示し方は，３つないし４つの年度の使用率をくらべて，より小さい方，同じ場合にはその使用率を合計して求める。連続した３つ，ないし４つの年度の共通度には，意味があると考える。その値が高いということは，この３つの年度の間，つまり２０年間に，あるいは，４つの年度の間，つまり３０年間に大きな変化がなかったことを示すからである。結果は次の表の通りである。

〈表６〉

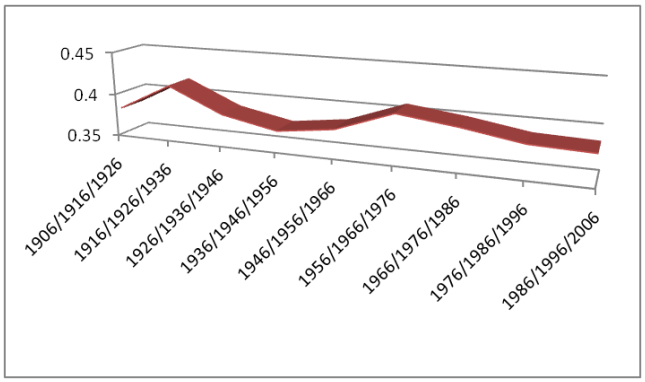
1906/1916/1926	0.3814	1906/1916/1926/1936	0.3342
1916/1926/1936	0.4136	1916/1926/1936/1946	0.3481
1926/1936/1946	0.3853	1926/1936/1946/1956	0.3338
1936/1946/1956	0.3725	1936/1946/1956/1966	0.3326
1946/1956/1966	0.3810	1946/1956/1966/1976	0.3379
1956/1966/1976	0.4050	1956/1966/1976/1986	0.3802
1966/1976/1986	0.3966	1966/1976/1986/1996	0.3510
1976/1986/1996	0.3856	1976/1986/1996/2006	0.3419
1986/1996/2006	0.3833		

〈グラフ２〉



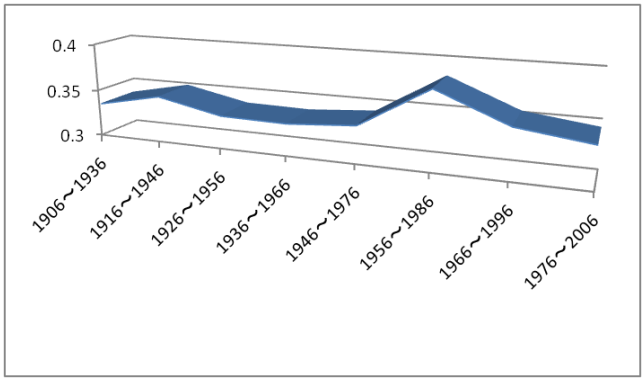
これらの表からまず読み取れることは，４年度続きのほうが３年度続きに比べて，共通度が低くなっていることである。そして，ともに２年ごとの類似度に比べれば，値は小さい。

〈グラフ３〉



３年度続きの共通度〈グラフ３〉を見ると，最も高いのは，戦前の 1916-1936 年の間で，次が 1956-1976 年である。３年度続きで見ると，大正のころが最も共通度が高く，非常に安定していた時期といえる。戦争直後の 1946 年の用語が，特異な様相を見せるので，1946 年をはさむと共通度は低くなる。戦後 10 年ほど経った 1956 年以後から，語は安定すると考えられる。しかし，安定を見せた後，徐々に共通度は下がっていく。

〈グラフ４〉



次に４年度続きの共通度〈グラフ４〉を見ると，最も高いのは，戦後の 1956-1986 年である。30 年という期間で見ると，全体が平均化され，大正期の安定は消えてしまう。しかし，そのようにして見ても，戦後は安定しているのである。

ただし，戦後 20 年以上経った 1966 年以後の調査では，共通度は再び下がっていく。

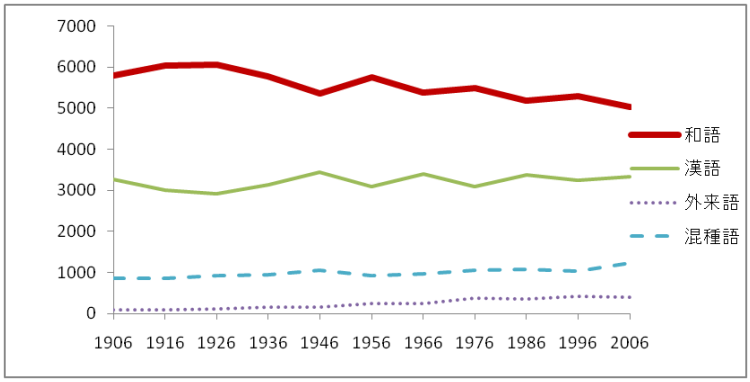
４．語種分布の変化

語種は，和語・漢語・外来語・混種語の４つにわけた。結果は次の通りである。

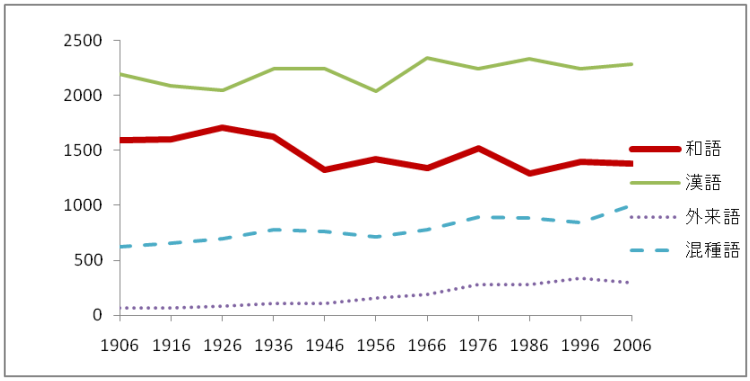
〈表７〉

	年度	1906	1916	1926	1936	1946	1956	1966	1976	1986	1996	2006
のべ	和語	5799	6049	6069	5773	5353	5751	5390	5482	5184	5296	5039
	漢語	3260	3012	2906	3135	3448	3090	3409	3087	3377	3241	3327
	外来語	87	87	112	148	154	243	238	371	352	429	404
	混種語	854	852	913	944	1045	916	963	1060	1087	1034	1230
異り	和語	1595	1600	1709	1623	1319	1424	1335	1522	1293	1396	1381
	漢語	2189	2090	2047	2242	2246	2040	2341	2241	2335	2245	2281
	外来語	72	69	86	108	113	161	190	281	282	338	301
	混種語	627	654	694	777	761	716	783	892	885	847	996
	計	4483	4413	4536	4750	4439	4341	4649	4936	4795	4826	4959

〈グラフ５〉のべ語数



〈グラフ６〉異なり語数



〈グラフ５〉〈グラフ６〉から，のべ語数，異なり語数ともに，外来語，混種語は増加傾向にあり，和語は減少傾向が見られるといえる。年によって，和語・漢語の増減が見られるが，その理由についてはまだ調査していない。

以上，簡単ではあるが，『中央公論』の 101 年の語彙について述べた。

【参考文献】

石井久雄(1990)「『中央公論』1986 年の用語」国立国語研究所報告 101，研究報告集 11　秀英出版

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　秀英出版