

使用頻度と意味分野から見た 日本語と韓国語の語彙*

申 玟 澈**

(e-mail:mcshin68@hanmail.net)

目 次

0. はじめに
 1. 比較対象語彙
 2. いくつかの数値的指標からの比較
 3. C50の意味分野別構成
 4. 最低頻度語（使用頻度1）の意味分野別構成
 5. おわりに
-

0. はじめに

個々の語はそれぞれの個性をもって存在するが、一方では、韓国語、日本語、英語、万葉語、京ことば、児童語などのように、何らかの基準によっていろいろなまとまりを形成する。このように、「あるまとまりをもった語の群れ」のことを語彙といい、それを対象とする学問分野が語彙論であるが、語彙についての研究は、言語学の他の分野に比べてかなり遅れている。

語彙研究の歴史を見ると、20世紀に入ってから、アメリカを始めとする色々な国で、教育的な目的から、読書教育に必要な語彙や基本語彙などの選定のために、聖書、文学作品、新聞、雑誌、教科書などの言語資料を対象とした大規模の語彙調査が行われるようになる。

* 本論文は2009学年度韓南大学校学術研究助成費の支援により研究されたものである。

** 韓南大学校 副教授 日本語学

なお、辞書の編集や各種の言語情報処理（機械翻訳・自動抄録・情報検索など）に必要な語彙テーブル作成のためにも語彙調査が行われる（田中章夫1978）。

語彙調査を行うと、語ごとに何回使われているかという頻度が得られ、頻度の高い語から低い語への順に並べると、ある言語において基本度の高い語が上位に位置するようになる。このように、使用頻度のような語彙の数量的側面は語の基本度を測る尺度として用いられるものであるが、その他にも、語彙の数量的側面からは「Zipfの法則」¹⁾に代表される語彙の一般的特徴が明らかになっている。

また、「樺島の法則」（樺島忠夫1954・1955）と「大野の法則」（大野晋1956）によって、語彙の品詞別構成と文体の間における関係性が明らかにされた。なお、語彙の語種別構成（語の出自による語彙の分類）からは、ある言語の語彙が外来要素の影響をどれくらい受けているか、という度合いが分かる。

しかし、上記の色々な観点をを用いても、言語間の語彙の比較において、各々の語彙を区別し特徴づけるには限界がある。その原因は、従来の語彙研究²⁾では、語彙の意味的性質が捨象されてきたからである。

語彙の構成要素である個々の語は必ずある意味を示すことから、語彙にはその基本的性格として数量的側面と意味的側面が備わっている。したがって、語彙を詳細に記述するためには、この両方の側面を生かさなければならない。

そこで、本稿では、同一内容の言語資料を対象として語彙調査を行い、そこから求められる使用頻度によって数量的側面を、個々の語に国立国語研究所の『分類語彙表』における分類番号を与えることによって、意味的側面をそれぞれ生かし、この両者による日本語と韓国語の語彙の比較を試みる。今回は、高頻度語と低頻度語において、両言語はどのような類似点と相違点を呈するかを明らかにしたいと思うのであるが、それは各々の語彙の記述に繋がるであろうと思われる。そのうち、低頻度語については、今まではほとんど注目されてこなかった使用頻度1の最低頻度語に焦点を絞ってみたい。

1. 比較対象語彙

本稿では、黒柳徹子の『窓ぎわのトットちゃん』（講談社、1981）と、それを召喚子が韓国語に翻訳した『창가의 토토』（프로메테우스출판사、2000）をテキストとして

1) 「頻度×頻度順位＝不変数（ $P \times r = \text{constant}$ ）」で、どんな語彙も、使用頻度の極めて高い極少数の語と多数の低頻度語、それから、たった一回しか使われない相当数の最低頻度語から構成されることをいう。

2) ここでいう語彙は、文字通り「語の集合としての語彙」を指す。

語彙調査を行う。『窓ぎわのトットちゃん』は61の話とあとがきから構成されているが、そのうち、本稿では、巻1～10の話における語彙を分析の対象とする³⁾。

これまで、国立国語研究所によって種々の語彙調査が行われ、調査単位としては、 α 単位・ β 単位、M単位・W単位、長単位・短単位など、色々な単位が採用されてきた。しかし、その規定がかなり細かいため、それに慣れるまで相当の時間を要するばかりでなく、記憶の負担も大きい。そこで、本稿では、もっと簡単明瞭な単位として、日本語は文節、韓国語は語節を基準にして語彙調査を行うことにする。

文節は橋本進吉による命名で、「文を実際の言語としてできるだけ多く区切った最短の一区切り」であり、次のように規定される（『日本語百科大事典』大修館書店1988参照）。

- ①一定の音節が一定の順序に並んで、それだけはいつも続けて発音される。
- ②文節を構成する各音節の音の高低の関係（アクセント）が定まっている。
- ③実際の言語においては、その前と後に音の切れ目を置くことができるのであって、必ず置かれるのではない。

上記のように、文節は音声言語に基づいた言語単位であるため、文節に分けることは割合容易であり、人による揺れも少なく、安定した単位である。これは、田島毓堂氏の学生を対象とした実験からも証明されている（田島毓堂1986）。一方、日本語の文節にほぼ相当する単位として韓国語には「語節」がある。語節は、パラダイム関係（paradigmatic relation；「範列関係」「系列関係」）とシンタグム関係（syntagmatic relation；「連辞関係」「統合関係」）によって分けられる一まとまりの語で、分かち書きの基準にもなる。なお、発音の時は、語節単位に区切って発音することが出来る（南基心・高永根1985）。よって、文節や語節を基準にして語彙調査を行うと、語彙調査の容易さと、日韓両言語間における単位の統一、という二つの条件を同時に満たすことが出来る。実際の調査においては、文節と語節に区切った後、その中を自立語と付属語に分け、それぞれを一単位とするのである。

語彙調査は全数調査を行った。その結果、日本語は異なり1237語・延べ8593語、韓国語は異なり1610語・延べ9772語の語彙がそれぞれ得られた。

2. いくつかの数値的指標からの比較

上記の語彙調査の結果を見ると、異なり・延べとも韓国語が日本語を大きく上回っているが、これは両言語の付属語の違いによるもので、韓国語が日本語より付属語の種類

3) 巻10までと範囲を限定したのは、そこまでの語彙で、ある程度の傾向が見られると考えたからである。

(特に、語尾)が多く、しかも、その頻度が高いからである。本稿の語彙調査では、<言語のすべての要素(当面は文字として固定できるもの)が語彙論の観点から見れば語彙の要素である>という比較語彙論の考えに従い、助詞、助動詞、語尾といった付属語も語彙調査の対象としているが、今回は、使用頻度と意味分野による比較を試みるので、意味に重点を置くために付属語は外すことにする。そうすると、日本語は異なり1159語・延べ4492語、韓国語は異なり1420語・延べ5171語になる。これに対して、田島毓堂(1992)に示されている、語彙を特徴づける色々な数値的指標のうち、異なり語数と延べ語数、一語平均使用回数、C50、頻度1の語などから、日本語と韓国語の語彙を比較してみたいと思う。

先ず、異なり語数と延べ語数であるが、この二つの指標は個別語彙⁴⁾の最も基本的な数量的側面を表わす数値である。対象の語彙全体を、単位語に区切り、個々に見出し語を与え、互いに他と異なる見出し語の総体が異なり語数であり、単位語の総数が延べ語数である。一般に、延べ語数は語彙の大きさ(作品の大きさ)を示す。一方、異なり語数は、語彙の大きさと深い関係があるものの、単純な比例関係ではなく、延べ語数が多くても、異なり語数は逆に少ない場合もある、といわれる。本稿の対象語彙を見ると、付属語を含めた時より差は縮まっているが、依然として、異なり語数・延べ語数ともに韓国語が日本語をかなり上回っている。そのうち、異なり語数はテキストに使われる語の種類なので、それによって語彙が豊富か貧弱かを指摘することが出来る。ここでは、韓国語の語彙が日本語より豊富である。

次に、一語平均使用回数は日本語が3.87、韓国語が3.64である。一般に語彙が大きければ、同じ語の使われる度合いが大きくなるので、当然ながら平均使用回数も増す。これは、どんな言語も、表現するのに必ず必要な語彙(基幹語彙⁵⁾)はある程度決まっているし、その使用回数は、作品が大きくなればなるほど、増すからである。しかし、本稿においては、異なり語数と延べ語数の差があるにも関わらず、両言語の平均使用回数の差はわずかしがなく、ほぼないに等しい。このことから、両言語は効率の面では似通っているといえよう。

一方、田島毓堂(1992)では、累積使用率が50%に達した時の異なり語数をC50といい、各語彙において注意すべき語、各語彙の特有語などを探る際、一つの目安として用いられると述べている。C50を見るためには度数別統計表が必要である。《表1》と《表2》は

4) 個別的に実現された言語作品・言語資料の語彙をいう。

5) 林四郎(1971)では、基本語彙を次のように五つに分けて考えている。

- ・基礎語彙-意味の論理的分析によって求められた半人工的な語彙
- ・基本語彙-特定目的のための「〇〇基本語彙」
- ・基準語彙-標準的社会人としての生活に必要な語彙
- ・基調語彙-特定作品の基調を作るのに働く語彙

それぞれ日本語と韓国語の度数別統計表である。

それを見ると、日本語と韓国語のC50は86語と121語である。なお、最低頻度語（使用頻度1の語）は日本語が629語、韓国語が797語であり、それが異なり語で占める割合は両言語とも5割を超える。C50も最低頻度語も韓国語が日本語を上回っている。

以上、いくつかの語彙指標から日本語と韓国語の語彙の様相を探ってみたが、一語平均使用回数以外は全て韓国語が日本語をかなり上回っている。同一内容のテキストを対象とした語彙調査の結果であるのに、このような傾向が見られたのは興味深い現象であるが、それが日韓両言語間において一般的に見られる傾向であるかどうかは、さらに色々な資料を対象とした語彙調査を行い、その結果と照らし合わせてみないと分からない。それは今後の課題とする。なお、上記のような差が生じた原因はどこにあるかを究明するのも大切であるが、数値的指標のみではこれ以上の指摘は不可能である。そこで、以下においては、本稿の主眼である、使用頻度と意味分野からの比較を行ってみることにする。

《表1》 度数別統計表（日本語）

度数	語数	延語数	延語累計	使用率	累積使用率	異語累計	累積異語率
156	1	156	156	3.47	3.47	1	0.09
84	1	84	240	1.87	5.34	2	0.17
82	1	82	322	1.83	7.17	3	0.26
78	1	78	400	1.74	8.90	4	0.35
74	1	74	474	1.65	10.55	5	0.43
72	1	72	546	1.60	12.15	6	0.52
56	1	56	602	1.25	13.40	7	0.60
54	1	54	656	1.20	14.60	8	0.69
47	2	94	750	1.05	16.70	10	0.86
46	1	46	796	1.02	17.72	11	0.95
43	1	43	839	0.96	18.68	12	1.04
42	1	42	881	0.93	19.61	13	1.12
40	1	40	921	0.89	20.50	14	1.21
39	1	39	960	0.87	21.37	15	1.29
38	1	38	998	0.85	22.22	16	1.38
37	2	74	1072	0.82	23.86	18	1.55
35	1	35	1107	0.78	24.64	19	1.64
32	1	32	1139	0.71	25.36	20	1.73
30	1	30	1169	0.67	26.02	21	1.81
29	3	87	1256	0.65	27.96	24	2.07
28	1	28	1284	0.62	28.58	25	2.16
27	2	54	1338	0.60	29.79	27	2.33
26	1	26	1364	0.58	30.37	28	2.42
25	1	25	1389	0.56	30.92	29	2.50
24	2	48	1437	0.53	31.99	31	2.67
23	1	23	1460	0.51	32.50	32	2.76

22	1	22	1482	0.49	32.99	33	2.85
21	4	84	1566	0.47	34.86	37	3.19
20	4	80	1646	0.45	36.64	41	3.54
19	1	19	1665	0.42	37.07	42	3.62
18	2	36	1701	0.40	37.87	44	3.80
17	1	17	1718	0.38	38.25	45	3.88
16	2	32	1750	0.36	38.96	47	4.06
15	8	120	1870	0.33	41.63	55	4.75
14	4	56	1926	0.31	42.88	59	5.09
13	10	130	2056	0.29	45.77	69	5.95
12	9	108	2164	0.27	48.17	78	6.73
11	8	88	2252	0.24	50.13	86	7.42
10	7	70	2322	0.22	51.69	93	8.02
9	9	81	2403	0.20	53.50	102	8.80
8	15	120	2523	0.18	56.17	117	10.09
7	29	203	2726	0.16	60.69	146	12.60
6	20	120	2846	0.13	63.36	166	14.32
5	28	140	2986	0.11	66.47	194	16.74
4	56	224	3210	0.09	71.46	250	21.57
3	93	279	3489	0.07	77.67	343	29.59
2	187	374	3863	0.04	86.00	530	45.73
1	629	629	4492	0.02	100	1159	100

《表2》 度数別統計表 (韓国語)

度数	語数	延語数	延語累計	使用率	累積使用率	異語累計	累積異語率
159	1	159	159	3.07	3.07	1	0.07
108	1	108	267	2.09	5.16	2	0.14
87	1	87	354	1.68	6.85	3	0.21
86	1	86	440	1.66	8.51	4	0.28
57	1	57	497	1.10	9.61	5	0.35
55	1	55	552	1.06	10.67	6	0.42
53	2	106	658	1.02	12.72	8	0.56
52	1	52	710	1.01	13.73	9	0.63
44	1	44	754	0.85	14.58	10	0.70
38	1	38	792	0.73	15.32	11	0.77
37	1	37	829	0.72	16.03	12	0.85
36	1	36	865	0.70	16.73	13	0.92
35	1	35	900	0.68	17.40	14	0.99
34	2	68	968	0.66	18.72	16	1.13
33	1	33	1001	0.64	19.36	17	1.20
32	2	64	1065	0.62	20.60	19	1.34
31	2	62	1127	0.60	21.79	21	1.48
29	1	29	1156	0.56	22.36	22	1.55
28	2	56	1212	0.54	23.44	24	1.69
27	2	54	1266	0.52	24.48	26	1.83

26	4	104	1370	0.50	26.49	30	2.11
25	2	50	1420	0.48	27.46	32	2.25
24	4	96	1516	0.46	29.32	36	2.54
23	4	92	1608	0.44	31.10	40	2.82
22	1	22	1630	0.43	31.52	41	2.89
21	2	42	1672	0.41	32.33	43	3.03
20	2	40	1712	0.39	33.11	45	3.17
19	4	76	1788	0.37	34.58	49	3.45
18	4	72	1860	0.35	35.97	53	3.73
17	4	68	1928	0.33	37.28	57	4.01
16	3	48	1976	0.31	38.21	60	4.23
15	5	75	2051	0.29	39.66	65	4.58
14	4	56	2107	0.27	40.75	69	4.86
13	8	104	2211	0.25	42.76	77	5.42
12	8	96	2307	0.23	44.61	85	5.99
11	7	77	2384	0.21	46.10	92	6.48
10	13	130	2514	0.19	48.62	105	7.39
9	16	144	2658	0.17	51.40	121	8.52
8	20	160	2818	0.15	54.50	141	9.93
7	22	154	2972	0.14	57.47	163	11.48
6	35	210	3182	0.12	61.54	198	13.94
5	42	210	3392	0.10	65.60	240	16.90
4	51	204	3596	0.08	69.54	291	20.49
3	114	342	3938	0.06	76.16	405	28.52
2	218	436	4374	0.04	84.59	623	43.87
1	797	797	5171	0.02	100	1420	100

3. C50の意味分野別構成

先にも述べたように、本稿では、語彙の意味的側面を見るために、『分類語彙表』の分類番号を用いるが、先ず、それがどういうものであるかを述べる。

『分類語彙表』は、単語が表わし得る意味の世界をカテゴリー化した上、その各カテゴリー（「意味分野」といってもいい）にそれぞれの単語を配当したものである。なお、各々のカテゴリーには固有の番号（「コード」ともいう）が付けられている。そのコードは整数部分（1～4）と小数部分（2～4桁）から成っており、整数部分は大分類として品詞論的な分類（<1. 体の類>、<2. 用の類>、<3. 相の類>、<4. その他>）を表わし、小数点以下の第1位は意味範囲の5部門（<.1>抽象的關係（人間や自然のあり方のわく組み）、<.2>人間活動の主体、<.3>人間活動－精神および行為－、<.4>人間活動の生産物－結果および用具－、<.5>自然－自然物および自然現象－）、第2位以下はその細分である。なお、小数点以下は<1.1110>（関係）、

<2.111> (関係)、<3.110> (関係) のように、品詞論的な4類の間で少なくとも小数点以下第2位まで相互に関係付けられている(『分類語彙表』の「まえがき」参照)。

それでは、日本語と韓国語のC50はそれぞれどんな意味分野の語から構成されているかを見てみることにする。但し、全体的傾向を捉えるために、コードは小数点以下第1位まで使って分類する。

《表3》 C50の意味分野別構成

意味分野	日本語		韓国語	
	語数	語例	語数	語例
1.1	13	それ これ なに(何) こと(事) よう(様) 時間 時 次 所 中 前 みんな 全部	16	것 거 뭐 때문 때 시간 오늘 전(前) 다음 후(後) 자리 쪽 앞 안 장기 정도
1.2	11	私 誰 自分 人 ママ 子 校 長先生 先生 学校 駅 教室	13	누구 자기 사람 아이 아이들 아저씨 엄마 친동아 교장선 생님 선생님 학교 역 교실
1.3	3	気 声 話	8	마음 기분 생각 수(의존명사) 말 얘기 일 수업
1.4	9	物 お弁当 田麩 門 窓 机 蓋 定期 電車	9	도시락 텐부 반찬 문 책상 의자 뚜껑 징액권 진철
1.5	5	山 海 顔 目 手	5	소리 산 들 바다 손
2.1	7	ある いる 成る 行く 来る 開ける 閉める	14	싸다 있다 되다 시작하다 서다 다 다니다 가다 오다 나다 들다 넣다 열다 닫다 향하다
2.3	12	思う 考える 分かる 見る 聞く 言う 話す 書く 歩く する 持つ やる	19	좋아하다 알다 모르다 생각하다 보다 쳐다보다 들여다보다 듣다 외치다 말하다 묻다 얘기하다 대답하다 그리다 앉다 걷다 하다 그러다 기다리다
3.1	14	その この そう こう どう こんな 無い いい よい また 未だ 少 し ちよっと もう	23	그 이 이렇게 그렇게 그런 같다 다르다 없다 아니다 안 (‘아니’의 준말) 싶다 좋다 잘 또 다시 아직 크다 갑자기 한 조금 다 더 아주
3.3	1	嬉しい	0	
4.1	2	そしてでも	4	그리고 그래서 하지만 그런데
10	5	いる ない みる いう くださる	8	있다(보조동사) 앓다 있다(보조형용사) 지다 버리다 보다 하다 주다

15	2	トットちゃん ロッキー	2	토토 로키
20	2	じゃてる	0	
合計	86		121	

《表3》には、整数部分が<10>、<15>、<20>の意味分野もあるが、これらは『分類語彙表』にはなく、比較語彙研究で新設したものである。それぞれ補助用言、固有名詞、連語の意味分野である。新設コードについてさらに詳しいことは田島毓堂・広瀬英史（1997）を参照されたい。

C50の語彙は、ほとんどが馴染み深く何の変哲もないもの、いわば水や空気のような存在であるが、日本語の‘ママ、校長先生、田麩、定期、トットちゃん、ロッキー’と韓国語の‘창가, 친동아, 교장선생님, 덴부, 정액권, 전철, 토토, 로키’などのような語はやや特殊である。これらの語は、申玟澈（2001）において、日韓両言語の最も基本的な語彙として選定した「小学生基本語彙」にも入っておらず、それほど基本度が高いとはいえない。このように、ある言語作品・言語資料の語彙調査を行うと、C50のような高頻度語には、その言語において基本的であると考えられる語彙だけでなく、その作品や資料においてのみ頻度の高い語彙も含まれる。そのような語彙を基調語彙（または、「テーマ語彙」という。基調語彙は作品や資料の主題および素材と関係の深いもので、上記の日本語と韓国語もそうである。使用頻度から語の基本度が測れるので、使用頻度は基本語彙選定の目安となるが、ある語がいくつの作品、または、資料にわたって出現するかという「使用範囲の広さ」も考慮に入れると、基調語彙のような語彙をふるい落とすことが出来る。

一方、両言語とも全体的に小数点以下第1位が<.1>である「抽象的關係」を表わす語の割合が高い。なお、<2>（用の類）では、「人間活動-精神および行為-」（<.3>）を表わす語も多いが、これは動詞の特徴であると考えられる。

日韓両言語間においては、<1.2>（人間活動の主体）、<1.4>（生産物および用具）、<1.5>（自然物および自然現象）、<15>（固有名詞）では差がほとんどなく、<2.1>（抽象的關係）、<2.3>（精神および行為）、<3.1>（抽象的關係）ではかなりの開きがある。<1.2>、<1.4>、<1.5>、<15>で差がないのは、個別語彙のある言語から他の言語に翻訳する際、主人公と登場人物のような主体や素材、背景などはそのまま移行される性格のものであるので、ある意味では、当然の結果であろう。一方、<2.1>、<2.3>、<3.1>の意味分野に対して、互いに関係のある語を対応させてみると、‘見る：보다·쳐다보다·들여다보다, 聞く：묻다·듣다, また：또·다시’のようになり、韓国語の方に色々な表現形式が用意されている。なお、韓国語には「否定」に用いられる‘아니다, 안’と「希望」を表わす‘싶다’があるが、これらには日本語の助動詞が対応する。これを見ると、両言語の表現形式と品詞分類の違いが差の生じた一つの原因であることが分かる。

4. 最低頻度語（使用頻度1）の意味分野別構成

次に、最低頻度語（使用頻度1）について考えてみたいと思う。C50の場合と同様に、小数点以下第1位まで使って分類する。まず、最低頻度語の全体における割合を以下の《表4》に示す。

《表4》

	異語数	延語数	頻度1の語数	頻度1の異なりでの割合	頻度1の延べでの割合
日本語	1159	4492	629	54.2	14.0
韓国語	1420	5171	797	56.1	15.4

《表4》を見ると、最低頻度語の異なりでの割合は5割を超えているのに対して、延べでは割合が大分下がってその存在感が薄れている。そのため、従来、基本語彙などを選定する目的で行われた一連の語彙調査では、ほとんど注目されてこなかったのである。高頻度語が、多くの作品で共通し安定した性質を持つものに対して、頻度1の語は、必ず頻度1の語になるという性質を持つものはない。そこで、頻度1の語の解釈は、「たまたま用いられた」とする 경우가多く、その他の位置づけはほとんどなかったのである。一方、石井正彦（1997）、広瀬英史（2002）、田島毓堂（2005）などは使用頻度1の語に注目し、使用頻度1の語の性格とそれが担う意味分野などを探り、その位置づけを試みている。その結果は、使用頻度1の語からも語彙について色々な指摘が可能であるとのことである。上記の研究はいずれも日本語の語彙のみを対象としているが、本稿では、韓国語との比較も試みる。

《表5》 最低頻度語（使用頻度1）の意味分野別構成

意味分野	主たる意義	日本語	韓国語
1.1	抽象的關係	85	97
1.2	人間活動の主体	18	33
1.3	人間活動－精神および行為	64	78
1.4	生産物および用具	67	68
1.5	自然物および自然現象	41	38
2.1	抽象的關係	80	114
2.3	精神および行為	91	104
2.5	自然現象	8	12
3.1	抽象的關係	72	143

3.3	精神および行為	43	45
3.5	自然現象	18	26
4.1	接続	8	7
4.3	間投および表現態度	16	22
10	補助用言	6	4
15	固有名詞	7	5
20	連語	5	1
合計		629	797

《表5》を見ると、日韓両言語とも「抽象的關係」 (<.1>) と「精神および行為」 (<.3>) の意味分野の語は多く、「人間活動の主体」 (<.2>)、「生産物および用具」 (<.4>)、「自然物および自然現象」 (<.5>) の意味分野の語は少ない、という構成である。なお、<2.5>、<3.5>、<4.3>はC50ではなかった意味分野で、これらは頻度1の語の現われやすい意味分野であるといえよう。<3.3>もC50では日本語の方に「嬉しい」1語しかないので、この意味分野も同様に考えていいと思われる。

一方、日韓両言語間の差を見ると、<1.4>、<1.5>、<15>では僅かな差しかないが、これに対しては、C50のところで述べたのと同じ理由が考えられる。最も大きな差が生じた意味分野は<3.1>で、日本語と韓国語の間で倍の差が見られる。その所属語を示すと、次の《表6》のようになる。

《表6》 <3.1>の意味分野の所属語

日 本 語	あっち あんな こっち 何で(副詞) 何んで(副詞) 一齐 ロ々 お互い 同じだ 等しい 一向 (副) まるっきり 仕方ない 必要だ 駄目 面倒だ 一応 駄目だ よく 良さそうだ 結構だ 上等だ 素敵だ 大丈夫 びん 立派だ ぎっちり 辛抱強い 強い 盛んだ ばちばち びたり ずるずる 大急行 ばらばら びったり べったり 暫く 段々 たまに またもや いやいよ 結局 既に やっと 若い 向い合わせ ぞろぞろ つるつるだ ぶくぶく 丸い よれよれ もしゃもしゃ 近い 広い 短い 薄い 遅い そろそろ 目まぐるしい ゆっくり 薄々 多少 たっぶり ほんの(副詞) なるべく 一段 大概 殆ど ほぼ 益々 散々
韓 国 語	그러하다 그리 마냥 별 어찌 여느 요('이'의 작은말) 이리저리 일일이 하나씩 당연하다 당연히 바르다(正) 사실(부사) 참되다 상관없다 어찌 어째야('어찌하여야'의 준말) 덩달아 직접 똑같다 서로 나란히 뒤죽박죽(부사) 형편없다 다름없이 영(부사) 절대로 공교롭다 별수없다 쓸데없이 필요하다 곤란하다 벅차다 쉬(쉽게) 하나같이 보통(부사) 심각하다 어처구니없다 유난히 유별나다 나쁘다 잘못(부사) 멋있었다 조출하다 지독하다 괜히 달달(의태어) 저절로 절로 짹 짹 활달하다 꿈쩍 종긋 휘청 빙 핑핑 쑥쑥 펄쩍 획 획 덩그러니 바짝 찰싹 내내 때로 설세없이 한참 항상 곧바로 당장 이내 이윽고 점점 즉시 매일 차례차례 막상 그런대로 드

	디어 먼저 어느새(부사) 여전히 이만(부사) 이미 이제껏 이제야 채 하마터면 젊 다 낡다 가웃하다 등그렇다 매끈매끈하다 비스듬히 우르르 줄줄이 길다랗다 나 지막히 넓다 높다 멀다 잘다 짧다 우람하다 조그맣다 촘촘하다 급하다 느닷없이 느릿하다 슬슬 찬찬히 천천히 다소 많이 잔뜩 조금씩 그만이다 꼬박 다들(부사) 다만 전(全;관형사) 그야말로 그토록 이쯤 가장 고작 단지 달랑 대개 더더욱 더 없이 맨 무려 한층 너무나 몹시 별로 상당히 심하다 웬만큼 하도
--	---

《表6》を見ると、日本語の‘一斉、一向、必要だ、面倒だ、一応、結構だ、上等だ、大丈夫、立派だ、辛抱強い、大急行、段々、結局、多少、一段、大概、散々’など、韓国語の‘일일이、당연하다、당연히、사실、상관없다、직접、형편없다、절대로、필요하다、곤란하다、보통、심각하다、유별나다、지독하다、활달하다、항상、당장、점점、즉시、매일、차례차례、여전히、다소、전(全)、단지、대개、한층、별로、상당히、심하다’などのように、漢語や漢語の成分が含まれている語が目立つ。このことから、日韓両言語とも、<3> (相の類) の語のうち、漢語や漢語の成分を含むものは使用頻度1の語に現われやすい、ということが指摘できる。特に、その傾向は副詞において著しい。一方、指示語においては、日本語が‘あっち、あんな、こっち’3語であるのに対して、韓国語は‘그러하다、그리、어찌、여느、요、이리저리、어째、어째야’8語であり、韓国語の方に縮約形があったり、語感の小さい語があったりで、バラエティーに富む。それ以外に、‘저절로、절로、잘다、조그맣다、찬찬히、천천히、많이、잔뜩’なども同様である。このように、韓国語では、<3.1>の意味分野に属する語に、似たような意味を表わす語が多く形成されており、その結果、使用頻度1の語に現われやすい語が日本語より多いのであろう。

5. おわりに

『窓ぎわのトットちゃん』という個別語彙を対象としていくつかの数値的指標とC50および使用頻度1の語の意味分野別構成から日本語と韓国語の語彙比較を行なってみた。

その結果、数値的指標による比較では、一語平均使用回数はわずかな差しかないのに対して、異なり語数、延べ語数、C50では韓国語が日本語をかなり上回っている。しかし、このような傾向が今回の調査に限ったことであるか、または、日韓両言語間において常に見られるかは、他の色々な語彙調査の結果と照らし合わせてみないと分からない。それとその原因の究明は今後の課題である。

次に、C50の意味分野別構成では、両言語とも全体的に小数点以下第1位が<1>である「抽象的關係」を表わす語の割合が高く、<2> (用の類) では、用言という特徴

から、「人間活動-精神および行為-」(<.3>)を表わす語も多く含まれている。なお、個別語彙において主人公と登場人物のような主体や素材、背景などと関係の深い<1.2>(人間活動の主体)、<1.4>(生産物および用具)、<1.5>(自然物および自然現象)、<15>(固有名詞)の意味分野ではほとんど差がないのに対して、<2.1>(抽象的關係)、<2.3>(精神および行為)、<3.1>(抽象的關係)ではかなりの開きがある。その原因は両言語の表現形式と品詞分類の違いなどにあると考えられる。

また、使用頻度1の語の意味分野別構成では、<2.5>(自然現象)、<3.3>(精神および行為)、<3.5>(自然現象)、<4.3>(間投および表現態度)は頻度1の語が現われやすい意味分野であることが確認できた。一方、<3.1>(抽象的關係)の意味分野では、日本語と韓国語の間ではほぼ倍の差が見られるが、この意味分野に属する韓国語に似たような意味を表わす語が多く含まれており、使用頻度1の語に現われやすい傾向が日本語より強いのであると考えられる。それから、日韓両言語とも、<3>(相の類)の語のうち、漢語や漢語の成分を含むものは使用頻度1の語に現われやすい、ということが指摘できる。

【参考文献】

- ・石井正彦 (1997) 「語彙調査の新展開」 (『「語彙研究法」報告2 語彙研究の可能性』名古屋大学大学院文学研究科)
- ・大野晋 (1956) 「基本語彙に関する二三の研究-日本の古典文学作品に於ける-」 (『国語学』24)
- ・樺島忠夫 (1954) 「現代文における品詞の比率とその増減の要因について」 (『国語学』18)
- ・樺島忠夫 (1955) 「類別した品詞の比率に見られる規則性」 (『国語国文』250)
- ・申玟澈 (2001) 「日韓語彙の比較研究-「小学生基本語彙」を対象として-」 (『開発・文化叢書37 比較語彙研究の試み7』名古屋大学大学院国際開発研究科)
- ・田島毓堂 (1986) 「語の単位-語彙論的見地から-」 (『松村博司先生喜寿記念国語国文学論集』右文書院)
- ・田島毓堂 (1992) 「語彙指標-語彙の数量的側面と語彙研究への視点-」 (『日本語研究と日本語教育』名古屋大学出版会)
- ・田島毓堂 (2005) 「『窓ぎわのトットちゃん』語彙8-度数1の語- (その1)」 (『国際シンポジウム比較語彙研究IX・語彙研究セミナーVI』語彙研究会)
- ・田島毓堂・広瀬英史 (1997) 「語素コードに関する提案-比較語彙論のために (その2)-」 (『「語彙研究法」報告2 語彙研究の可能性』名古屋大学大学院文学研究科)
- ・田中章夫 (1978) 『国語語彙論』 (明治書院)
- ・林四郎 (1971) 「語彙調査と基本語彙」 (『電子計算機による国語研究III』国立国語研究所報告39)
- ・広瀬英史 (2002) 「延べから見た使用頻度“1”の語の語彙の分析」 (『名古屋大学日本語学研究室 過去・現在・未来』名古屋大学大学院文学研究科)
- ・南基心・高永根 (1985) 『표준 국어문법론 개정판 (標準国語文法論改訂版)』 (塔出版社)

要 旨

本稿では、『窓ぎわのトットちゃん』という個別語彙を対象として語彙の基本的性格としての数量的側面と意味的側面の両方を生かした日韓両言語の語彙の比較を試みた。

まず、異なり語数・延べ語数、一語平均使用回数、C50などのような数値的指標からの比較では、一語平均使用回数の差はほぼないに等しいのに対して、異なり語数、延べ語数、C50では全て韓国語が日本語をかなり上回っている。同一内容のテキストを対象とした語彙調査の結果であるのに、このような結果が見られたのは興味深い現象であるが、それが、今回の調査に限ったことであるか、日韓両言語間において常に見られる傾向であるかは、他の色々な語彙調査の結果と照らし合わせてみないと分からない。それとその原因の究明は今後の課題である。

次に、使用頻度と意味分野との関係から両言語の語彙を比較してみた。

高頻度語のC50の意味分野別構成では、両言語とも<1>（抽象的關係）を表わす語の割合が高く、<2>（用の類）では、<3>（人間活動-精神および行為-）を表わす語も多い、という傾向が見られた。なお、個別語彙のある言語から他の言語に翻訳する際、主人公と登場人物のような主体や素材、背景などはそのまま移行される性格のものであるので、それと関係の深い<1.2>（人間活動の主体）、<1.4>（生産物および用具）、<1.5>（自然物および自然現象）、<15>（固有名詞）の意味分野ではほとんど差が見られなかった。一方、<2.1>（抽象的關係）、<2.3>（精神および行為）、<3.1>（抽象的關係）ではかなりの開きがあり、両言語の表現形式と品詞分類の違いが差の生じた一つの原因であると考えられる。

また、使用頻度1の語の意味分野別構成では、日韓両言語とも「抽象的關係」（<1.1>）と「精神および行為」（<3>）の意味分野の語は多く、「人間活動の主体」（<2>）、「生産物および用具」（<4>）、「自然物および自然現象」（<5>）の意味分野の語は少ない、という傾向が見られる。なお、<2.5>（自然現象）、<3.3>（精神および行為）、<3.5>（自然現象）、<4.3>（間投および表現態度）は頻度1の語が現われやすい意味分野であることが確認できた。C50の場合と同様、<1.4>（生産物および用具）、<1.5>（自然物および自然現象）、<15>（固有名詞）の意味分野では僅かな差しか見られなかった。最も大きな差が生じた意味分野は<3.1>（抽象的關係）で、この意味分野に属する韓国語に似たような意味を表わす語が多く形成されており、その結果、使用頻度1の語に現われやすい傾向が日本語より強いのであると考えられる。それから、日韓両言語とも、<3>（相の類）の語のうち、漢語や漢語の成分を含むものは使用頻度1の語に現われやすい、ということが指摘できる。このように、従来の語彙研究ではほとんど注目されてこなかった使用頻度1の語からも、その意味分野別構成を用いると、語彙について色々な指摘が可能である。

キーワード：数量的側面 意味的側面 異なり語数 延べ語数 一語平均使用回数 C50
最低頻度語 意味分野

투 고 : 2009. 8. 31
1차 심사 : 2009. 9. 12
2차 심사 : 2009. 9. 26