

『窓ぎわのトットちゃん』語彙12

——中間の語彙に関して（試案）——

田 島 毓 堂

0. はじめに

高頻度語、上位の語、いわゆる基本語彙、あるいは、基幹語彙というものについては、多くの議論があり、それぞれの語彙はそれぞれ一定の意義を有している。そして、そういう語彙は、たとえば、日本語の語彙に限って言うならば¹⁾、内容的には同じようなものであることが多い。勿論、当然それぞれの語彙に、その語彙に特有なものがある。それは、まさに、その語彙を特徴付ける「基調語彙」（キーワード群）である。このことは、比較によって知られる。多くの語彙に共通して現れるのは、その言語を使って表現する以上は、どうしても使わなくてはならぬ語群である。まさに、林四郎氏がいう「基幹語彙」²⁾に当たるであろう。

上記の基調語彙以外の高頻度語は、決まった言葉達であり、今までも何度も示してきたように、それを見ても、特にそういう事柄に特別な関心を持つ人以外には、何も特別な感慨も催させないような、ごくありふれた言葉達であった。それらは、ほぼ決まったものであった。

一方、逆に最低頻度語、これは、実際に用いられているものを対象にするから、具体的には「度数1の語」である。その語彙に姿を現さない、度数0の語は一体どういうものか、ということについても、たとえばこういうものだということ、単にある物以外ということではなくて具体的に示すことは出来る。未だ、実際にそれを使って試したことはないが、それが、意味分野別構造分析法の一つの重要な射程でもある。

最低頻度語、つまり、度数1の語は異なり語数

にして、ほぼ半数を占めることが多く、その出現する意味分野も広範にわたる。しかし、おのずから、それがいかなる意味分野によく出現するか、度数1の語として出現する事の多い意味分野と、そういうことの少ない意味分野というものがある程度指摘できる。そして、付け加えて言えば、高頻度語が、どの語彙にも共通する語であることが多いのに対して、頻度1である語というものを前もって特定することはできない。結果的に度数1の語として出現していることが分かるだけである。そして、この度数1の語について論じた研究もいくつか、貴重な成果を上げている³⁾。

それに対して、本稿で取り上げようとしている上位の語彙でもなく、最低頻度語でもない語彙——それを「中間の語彙」と称しておくのだが——については、その扱い方そのものから未知数である。第一、何を「中間の語彙」とするかにしても、定義されているものではない。

本稿では、上位語彙でもなく、最低頻度語でもない語彙、について、それをどう定義するかという事から初めて、全く、素朴な形であるが、試案を示し、大方のご意見・ご批判をいただきたいと思う。

1. 中間の語彙

前稿『窓ぎわのトットちゃん』語彙データ（確定版2007）（「語彙論の基盤となる語素コード・単語コードの研究及びその電子辞書開発」科学研究補助金報告書 課題番号15202010 基盤研究A 2003～2006）において、語彙分析の今後の課題の一つとして、その中に、「4.3 上位語と度数1の間」として、次のように記した。すなわち、

自立語だけに限ってみると、累積使用率50%、所謂C50は124語（度数45まで）、60%では228語（度数22まで）、70%では426語（度数12まで）、80%では809語（度数6まで）このあたりまでは、上位語といえる（語彙によって若干の違いはあるし、正確なことではないが、累積使用率が10%増す毎に、所要の異なり語数はほぼ倍になっている、このことは多くの語彙に共通した数量的性質である）。一方度数1の語はほぼ5割に達する。この中間にある語、今対象にしている『窓ぎわのトットちゃん』語彙では具体的に度数6から2までの語に相当するが（度数6はC80と一部重複）、これは、異なり語でほぼ3分の1に達する（34.53%）。使用率では15%程度である。度数1の語についてはそれ相応の扱いがあるが、この中間の語彙についてはこれを放置しておいていいはずはないが、その分析方法が確立していない。今後、語彙の分析をしていく上でこの扱いについて考えていかなければならない。

と述べ、その数量を「表12 中間の語」として示した。それに依れば、

度数	語数	延べ語数	使用率	異語率
6	104	624	1.936	2.360
5	153	765	2.374	3.473
4	224	896	2.780	5.085
3	354	1062	3.295	8.036
2	686	1372	4.258	15.573
計	1521	4719	14.645	34.528

（「度数統計表全体」による）

である。

但し、この数値は、記号等を含めた全体の統計表から、度数6以下のものの数値である。記号を除いた場合及び自立語だけの場合若干異なる。その上、全体で考えると、累積使用率80%を超えるところは、以下に示すように、度数35である。従って、累積使用率80%までの語を上位語、以下、度数2以上のものまでを中間の語彙としようとする立場とは食い違うので、もう一度、そのデータを示す。

これを一旦「中間の語彙」と仮定し、種々の観

点から「中間の語彙」が如何なるものなのかという点について考えてみたい。

2. 『窓ぎわのトットちゃん』語彙によるデータ

上位の語彙、中間の語彙、度数1の語彙それぞれについて、先ず客観的データを示す（表1）。記号や付属語等はあるいは最終的に余り関わりは持たないかも知れないが、全体、記号を除いた語彙（除記号）、自立語だけに分けて示す。上位語はC50、C60、C70、C80（C80までを上位の語彙としたので、この欄に示す数値が、上位の語彙のデータである。C50、C60、C70の項目はその内数である）に分けて示す。

このような数値の羅列だけではなかなかその語彙構造に対してイメージを喚起できないかと思う。若干の説明を加える。

先ず、全体においては、上位語は異なり語数217語、及び異なり語数率4.77%、コード数124／全コードに対する率14.37%のいずれも、その累積使用率80%に対して極めて小さいことに注意される。上位語に限られた語によって構成されていることに依る。この傾向は、除記号、自立語においても、多少それぞれ、その数、率ともに大きくなるが、なお、異なり語に対する率も、出現コードにおける率も、最大は、自立語だけの場合の、19.34%、44.27%止まりである。

これに対して、度数1の語は、どのように集計しても、異なり語においては、ほぼ半数に及び、コード率も80%を超える。この度数1の語の重要性については、前稿までに繰り返し述べたことであるので、その指摘に留める。

一方、中間の語彙と仮定した語彙は、全体、除記号、自立語とその分け方によりその所属範囲が相当異なる。即ち、累積使用率80%までの上位語と、度数1の語を除いたものということになるので、全体では、使用度数34から2までの語2147語となり、除記号では2から18までの2002語、自立語では2から5までの1390語になる。それぞれの異なり語率は、47.16%・44.21%・31.55%と逓減する。使用率も、17.07%、16.18%、12.75%、含まれるコード率も74.16%、72.74%、

『窓ぎわのトットちゃん』語彙12 (田島)

表1 『窓ぎわのトットちゃん』語彙データ

全体

	実率	最小度数	異語数	内自立語数	含まれるコード数	延べ語
上位語						
C80	80.13%	35	217	151 (付54、記12)	124	62710語
C70	69.95%	105	84	45 (付30、記9)	53	54742語
C60	60.24%	231	37	13 (付18、記6)	22	47140語
C50	50.03%	729	16	2 (付10、記4)	8	39151語
異なり語率 C80	4.77%				コード率14.37%	
中間語彙						
度数2から度数34まで						
2147語	47.16%	(自2091、付49、記7)	640コード	74.14%	(自619、付20、記1)	
延べ	13360語	17.07%				
度数1の語						
2189語	48.08%	(自2163、付20、記6)	721コード	83.54%	(自707、付13、記1)	
延べ	2189語	2.80%				

除記号

	実率	最小度数	異語数	内自立語数	含まれるコード数	延べ語
上位語						
C80	80.24%	19	343	281 (付62)	192 (自171、付21)	49045語
C70	70.18%	55	143	97 (付46)	88 (自68、付20)	42897語
C60	60.28%	123	66	38 (付28)	47 (自32、付15)	36841語
C50	50.00%	236	30	12 (付18)	20 (自11、付9)	30562語
異なり語率 C80	7.58%				コード率22.27%	
中間語彙						
度数2から度数18まで						
2002語	44.21%	(自1961、付41)	627コード	72.74%	(自611、付16)	
延べ	9892語	16.18%				
度数1の語						
2183語	48.21%	/3.57% (自2163、付20)	720コード	83.53%	(自707、付13)	

自立語

	実率	最小度数	異語数	含まれるコード数	延べ語
上位語					
C80	80.82%	6	852	367	26040語
C70	69.75%	13	419	232	22474語
C60	60.90%	22	241	149	19621語
C50	50.00%	45	124	86	16153語
異なり語率 C80	19.34%			コード率44.27%	
中間語彙					
度数2から度数5まで					
1390語	31.55%	555コード	66.95%	延べ4018語	12.75%
度数1の語					
2163語	49.10%	707コード	85.28%	延べ2163語	6.71%

注：上位語は前述の通り C80 語彙である。C70・C60・C50はその内数。「実率」とは C80 と言っても、累積使用率が丁度80%になるとは限らないので、直近の度数階級における、累積使用率を示したものの。

66.95%と同様の傾向を示し、度数1の語が全体・除記号・自立語と通増するのと逆の様相を示す。特に、自立語だけの場合の落ち込みが激しい。自立語だけの場合、度数1の語は他と大差ないが、上位語の比重がその分、増している事を示す。

中間の語彙は、数量と、意味分野の広がりとの観点からはこのような語彙である。

3. 『南の島のティオ』の語彙の場合

さて、前稿までの考察において、対象が『窓ぎわのトットちゃん』語彙に限られていた。ために、指摘できた事柄が果たして単に『窓ぎわのトットちゃん』語彙に限られる現象なのか、もう少し一般的に言うことが出来ることなのかは、うすうす感じることは出来ても、証拠を添えて言うことは出来なかった。

『南の島のティオ』は池澤夏樹氏の10編の短編から成る美しい物語である。1992年に単行本として刊行された後、96年に文春文庫の一冊として刊行された。その語彙の概要については別稿に述べるが、表2に、前項に記した形で、『南の島のティオ』語彙のそれを示しておく。

なお、全体は異なり語数4574語、延べ62042語、記号は12種類7330回、付属語は98種類26026回の使用である。自立語は若干の不確定要素はあるが、異なり4464語、延べ28686語である。全体の規模は『窓ぎわのトットちゃん』語彙に比べて僅かに小さいというところである。

表2を見て、『窓ぎわのトットちゃん』語彙のそれと比べてみると、勿論、個々の数値に違いはあるが、全体的に見れば数量的には、驚くほどよく似た構造を示していると言える。但し、特に、『窓ぎわのトットちゃん』語彙で目立つ記号類の多さが、「全体」の表では違いとして目立つ。そのことが「全体」の表に現れているように思えるが、「除記号」或いは「自立語」においてはその影響は除去されている。そして、規模が、『南の島のティオ』の方が若干小さいにもかかわらず、異なりではほぼ同数、結局、語彙にその題材からよりバラエティが有ることが、自立語・除記号の表には現れる。

4. 中間の語彙の概観

前稿にごく大まかに「中間の語彙」として示したとき、それがほぼ、異なり語数の3分の1ぐらいに当たると述べた。ただ、上述したように、そのデータそのものが「全体」の表なら、累積使用率80%までの語と、度数1を除いたものとするべきところ、度数6から2までとした点、不正確であった。度数6の語を含めるかどうかは議論のあるところだが、C80を80%直近の度数階級とすれば、本稿で上述したようになる。

『窓ぎわのトットちゃん』語彙と『南の島のティオ』語彙の両者から、中間の語彙と規定した語彙を見ると、まさにいろんな点で中間的な語であると言える。

異なり語に占める割合は、どの表でも、上位語よりは大きく、度数1の語よりは小さい。その中に含まれるコードの率を見ても同様である。尤も、使用率は度数1の語彙よりは何れの場合も大きくなる。順位は、異語率、コード率とは逆転するけれども。上位語を累積使用率80%としているから、中間の語彙の使用率と、度数1の語の使用率をたせばほぼ20%に成るのは当然である。この、使用率は、上位語彙、中間語彙、最低頻度語彙を区分するためにだけ役立ち、以後、この分析には、主として異なり語数とコード率、そのコードが関わってくる。

自立語だけを対象にすれば簡単ではあるが、この二つの語彙を見ただけでも、付属語の様子は少し差が見られる。そうとすれば、この、中間の語彙を観察するには、ひとまず記号は差し措くとしても、自立語だけでは正確な様子を知られない可能性が予測される。それ故、本稿では、「除記号」と称した表を中心に考えていく。それは、分布という観点からも度数2から5（『南の島のティオ』語彙では2から4）では、非常に限られることが前もって知られるからである。もちろん、度数2から18の語であっても、分布ということからみれば、『窓ぎわのトットちゃん』は62章に分割されているので、これまた、最大でも全体の29%にしか行き渡らない。まさにこのことこそが中間

『窓ぎわのトットちゃん』語彙12 (田島)

表2 『南の島のティオ』語彙データ

全体・南

	実率	最小度数	異語数	内自立語数	含まれるコード数	延べ語
上位語						
C80	80.19%	20	310	246 (付59、記5)	163	49753語
C70	70.02%	62	118	73 (付40、記5)	68	43439語
C60	60.20%	148	5	23 (付22、記5)	32	37347語
C50	50.54%	389	22	5 (付13、記4)	13	31355語
異なり語率	C80 6.78%				コード率17.40%	
中間語彙						
度数2から度数19まで						
2003語	43.79% (自2250、付11)			640コード68.30% (自)		延べ10028語16.16%
度数1の語						
2261語	49.43% (自2250、付11)			795コード84.85% (自787、付8)		延べ2261語3.64%

除記号・南

	実率	最小度数	異語数	内自立語数	含まれるコード数	延べ語
上位語						
C80	80.15%	16	383	321 (付62)	197 (自173、付24)	43849語
C70	69.96%	43	155	108 (付47)	89 (自70、付19)	38279語
C60	60.15%	97	69	37 (付32)	42 (自27、付15)	32910語
C50	49.92%	225	29	14 (付15)	20 (自12、付8)	27312語
異なり語率	C80 8.40%				コード率21.05%	
中間語彙						
度数2から度数15まで						
1918語	(自1893、付25) 42.04%			645コード68.91% (自636、付9)		延べ8602語15.72%
度数1の語						
2261語	(自2250、付11) 49.56% / 3.64%			795コード84.94% (自787、付8)		

自立語・南

	実率	最小度数	異語数	含まれるコード数	延べ語
上位語					
C80	80.85%	5	969	405	23184語
C70	69.87%	10	490	251	20031語
C60	60.66%	18	285	157	17392語
C50	49.95%	31	150	95	14319語
異なり語率	C80 21.71%			コード率44.41%	
中間語彙					
度数2から度数4まで					
1245語	27.89%		559コード61.29%		延べ3252語11.34%
度数1の語					
2250語	50.40%		787コード86.29%		延べ2250語7.84%

の語彙の一つの特徴なのだろうが。それでも、自立語だけを対象にする場合よりはかなり分布が広がる。もちろん、異なり語数としては、付属語は僅かであるが、使用率は相当になる。そのデータ

を示しておこう。

『窓ぎわのトットちゃん』の記号を除いた語彙

異なり語数4528語 延べ語数61120語

自立語4405語 延べ32221語 52.718%

付属語123語 延べ28899語 47.282%
『南の島のティオ』の記号を除いた語彙
異なり語数4562語 延べ語数54712語
自立語4464語 延べ28686語 52.431%
付属語98語 延べ26026語 47.569%

この語彙の数量的構造は、恐らく一般的なものであろう。

以上、全て、数量的観点から眺めてきた。それによって、中間の語彙を初め、その他の数量的構造について、かなりこの両語彙には共通することが見て取れた。以下、中間の語彙に焦点を当てて、数量面と共に、意味的側面について観察しよう。

5. 分布について

「中間の語彙」の意味分野別構造について分析する前に、もうひとつ、この中間の語彙がどのように分布しているのか、という点について確かめておく。

度数2から18の語2002語についてどのように分布しているかを見る。ここで言う分布とは、『窓ぎわのトットちゃん』本章61章と、あとがき1章を含め、62章にどのように使われているかということである。その実際を見る前にまず考えられることは、それらは1章だけに使われるものから、18章全体に散らばっているものまであり得るということである。一方で、あるところに集中して用いられる語もあろう。一般に度数が小さければ分布は小さく、度数が大きければ、分布範囲も広いということが言えよう。実際に、2002語の使用度数ごとの語数を示しておく。

度数	2	3	4	5	6	7	8	9	10
語数	684	353	223	152	104	89	78	59	52
度数	11	12	13	14	15	16	17	18	
語数	29	36	42	27	19	21	17	17	

次に、分布数ごとの語数を示す。

分布	1	2	3	4	5	6	7	8	9
語数	431	624	304	176	134	89	59	44	33
分布	10	11	12	13	14	15	16	17	18
語数	28	29	22	12	8	7	1	1	0

度数2の語は、分布は1か2しかない。一方度数18の語は分布1から18まであり得る。少し煩わしいかも知れないが、一々表示してみよう。それより先に、次の表からは、明示的には分からないので、前もって一言すれば、度数と分布の数が等しい語、つまり、5語あれば5章に、6語あれば6章に分布しているというように、ある章に1回ずつ出現しているという語である。これが、2002語中798語ほぼ4割ある。また、何語であろうと、1章にのみ分布しているのが431語ある。この中には、度数の高い語もあり、注目しておく必要がある。これは後に表示する（別表1分布1の語）。

表3に語彙の度数ごとの分布を示した。中間の語彙の度数ごとの分布が見られる。

『窓ぎわのトットちゃん』は幾つかに章分けできるので、このようにして、中間の語彙を観察することが出来る。これによって特定の章におけるキーワードがこの中にも含まれる事も分かる。例えば、分布1という語で、度数が、11とか、13、14、16、18というような語がある。これらは、別表に示すように、18回の「宮崎君」、16回の「マサオちゃん」、14回の「やぶける」、13回の「苦い・山伏・病気・田麩」、11回の「二十銭・温泉」、10回の「白墨・義経」などはそれぞれの章でのキーワードである。

こういう特徴のない語は、このような数量的観点からは特別な特徴を指摘できないものが多い。

ただ、この方法は、章分けされていない語彙には直接適用できないが、何らかの分割法を見いだすことにより、部分部分の特徴語を指摘することは可能であろう。

以上の分布と度数の関係を別表に示しておく。なお、別表中の「金」「申」欄に○印があるのは、それぞれ金直洙「日韓「基幹語彙」の比較研究——その選定及び意味分野別構造分析を中心に——」（『比較語彙研究の試み11』2004）、申玫瑰「日韓語彙の比較研究——「小学生基本語彙」を対象にして——」（『比較語彙研究の試み7』2001）、において、基幹語彙的な語として認定されているものである。

『窓ぎわのトットちゃん』語彙12 (田島)

表3 語彙の度数分布

度数	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
分布1	260	70	37	8	14	9	9	5	2	2		4	1		1		1
分布2	424	105	42	21	7	9	6	4	1		2		1		2		
分布3		170	62	34	8	7	3		6	3	3	5	1		1		1
分布4			82	38	22	12	8	2	5	2	2	1	1	1			1
分布5				51	31	16	13	10	4	3	1	1		2		1	
分布6					22	22	17	10	6	2	2	7	1				
分布7						14	13	9	5	3	6	2	4	2		1	
分布8							9	12	9	3	6	2		1	2		
分布9								7	7		2	5	4	2	2	2	2
分布10									7		3	6	5	3	3	1	
分布11										8	4	9	3	2	1	1	1
分布12											2		5	6	5	2	2
分布13												1	2	2	2	2	3
分布14													1	1	1	4	1
分布15															1	3	3
分布16																	1
分布17																	1
分布18																	1

次に、この、意味分野別構造についても続稿を用意する。

注

- 1) 他言語の語彙について具体的に観察したことがないからこのように日本語に限って述べておくが、語彙の数量的一般性として他言語に於いても同様であると考えられる。
- 2) 林四郎「語彙調査と基本語彙」『電子計算機による国語研究Ⅲ』(国立国語研究所研究報告39) 1971.
- 3) ・石井正彦「使用頻度“1”の語と文章——高校『物理』教科書を例に——」『国立国語研究所研究報告集』17、1996.
・広瀬英史「使用頻度“1”の語彙の意味構造分析——『源氏物語』(桐壺巻)の原文と谷崎潤一郎

訳を用いて——」『日本語研究センター報告』5、1998.

・拙稿「『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、2005)

・拙稿「同 9——度数1の語(その2)——」(『愛知学院大学文学部紀要』35、2005)

・拙稿「同 10——度数1の語(その3)・度数率50%以上——」(『愛知学院大学人間文化研究所紀要人間文化』21、2006)

・拙稿「同 11——度数1の語(その4)——」(『愛知学院大学人間文化研究所紀要人間文化』22、2007)

・拙稿「度数1の語について——『窓ぎわのトットちゃん』語彙から——」(『国際シンポジウム比較語彙研究』X、2006)

別表の見方

1. 別表1から12までである。
2. 項目は全て同じである。
3. 「見出語」は仮名見出し、「漢字表記」はその漢字表記できるものの漢字見出し。
4. 「Wc」その見出し語の単語コード。
5. 「度数」はその見出し語の頻度数。
6. 「分布」はその見出し語が、幾つの「章」に出現

しているかを示す。

7. 「金」「申」欄の○印は、前記の通り、それぞれ金直洙「日韓「基幹語彙」の比較研究——その選定及び意味分野別構造分析を中心に——」(『比較語彙研究の試み』11、2004)、申玫澈「日韓語彙の比較研究——「小学生基本語彙」を対象にして——」(『比較語彙研究の試み』7、2001)に有ることを示す。

別表1 『窓際のトットちゃん』分布1度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	みやざきくん	宮崎君	15.2340	18	1		
2	まさおちゃん	マサオちゃん	15.2340	16	1		
3	やぶける	破ける	2.1571	14	1		
4	でんぶ	田麩	1.4300	13	1		
5	にが	苦い	3.5050	13	1		○
6	びょうき	病気	1.5850	13	1	○	○
7	やまぶし	山伏	1.2410	13	1		
8	おんせん	温泉	1.5250	11	1		○
9	にじっせん	二十銭	1.1961	11	1		
10	はくぼく	白墨	1.4530	10	1		
11	よしつね	義経	15.2340	10	1		
12	きやたつ	脚立	1.4450	9	1		
13	こうか	校歌	1.3230	9	1		
14	さわかい	茶話会	1.3510	9	1		
15	しばい	芝居	1.3240	9	1		
16	ろー	ロー	2.3830	9	1		
17	おおさか	大阪	15.2590	8	1		
18	さあー	さあー (終助)	8.0050	8	1		
19	さのさん	佐野さん	15.2340	8	1		
20	すきー	スキー	1.3374	8	1		
21	ちひろさん	ちひろさん	15.2340	8	1		
22	ちょうせんじん	朝鮮人	15.2300	8	1		
23	とがし	富樫	15.2340	8	1		
24	ぶつ	打つ	2.1563	8	1		○
25	もどす	戻す	2.1527	8	1	○	○

別表2 『窓際のトットちゃん』分布2度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	ひよこ	雛こ	1.5620	16	2		○
2	やすちゃん	泰ちゃん	15.2340	16	2		
3	げんざい	現在	1.1641	14	2	○	○
4	しゃしん	写真	1.3220	12	2	○	
5	はくちょう	白鳥	1.5620	12	2		
6	ふね	船	1.4660	10	2	○	○
7	いちだん	一段	3.1992	9	2		
8	ばじゃま	バジャマ	1.4220	9	2		○
9	ばんつ	パンツ	1.4230	9	2		
10	ぶるーまー	ブルーマー	1.4230	9	2		
11	おじょうさん	お嬢さん	1.2130	8	2		
12	って	て (連語)	20.3120	8	2		
13	はしご	梯子	1.4470	8	2		
14	はやし	林	1.5270	8	2		○
15	もうふ	毛布	1.4201	8	2		○
16	やさい	野菜	1.5510	8	2	○	○

『窓ぎわのトットちゃん』語彙12 (田島)

17	いっとう	一等	3.1992	7	2		
18	かた	方	1.2020	7	2	○	○
19	しじゅうしちし	四十七士	1.2340	7	2		
20	だいじょうぶだ	大丈夫だ	3.1340	7	2	○	○
21	てつじょうもう	鉄条網	1.4161	7	2		
22	にほんご	日本語	1.3101	7	2	○	
23	はだか	裸	1.5700	7	2		○
24	やける	焼ける	2.5161	7	2		○

別表3 『窓際のトットちゃん』分布3度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	おさげ	お下げ	1.3334	18	3		
2	べんけい	弁慶	15.2340	16	3		
3	すばい	スパイ	1.2410	14	3		
4	おさいふ	お財布	1.4280	13	3		
5	おぼえる	憶える・覚える	2.3050	13	3	○	○
6	およぐ	泳ぐ	2.3370	13	3	○	○
7	しぬ	死ぬ	2.5820	13	3	○	
8	ていき	定期	1.4590	13	3		
9	かいすいぎ	海水着	1.4220	12	3		
10	しめる	閉める	2.1553	12	3		
11	ひろう	拾う	2.3393	12	3	○	○
12	おどる	踊る	2.3370	11	3	○	○
13	おなべ	お鍋	1.4520	11	3		
14	だるくろーず	ダルクローズ	15.2340	11	3		
15	おほか	お墓	1.4400	10	3		
16	きょういく	教育	1.3640	10	3	○	○
17	さいしょさん	税所さん	15.2340	10	3		
18	だの	だの(副助)	8.0070	10	3		
19	としよしつ	図書室	1.4430	10	3		
20	はた	旗	1.4580	10	3		○
21	えんぴつ	鉛筆	1.4530	8	3		○
22	きゃらめる	キャラメル	1.4340	8	3		○
23	まったく	全く	3.1990	8	3	○	○

別表4 『窓際のトットちゃん』分布4度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	まるやませんせい	丸山先生	15.2440	18	4		
2	りずむ	リズム	1.1940	15	4		○
3	のぼる	登る	2.1540	13	4	○	○
4	おかず	おかず	1.4300	12	4		○
5	かう	買う	2.3760	12	4	○	○
6	あんた	あんた	1.2000	11	4		
7	はる	張る・貼る	2.1340	11	4	○	○
8	えいご	英語	1.3101	10	4	○	
9	おしり	お尻	1.5720	10	4		
10	きっぷ	切符	1.4590	10	4	○	
11	すかーと	スカート	1.4230	10	4		○
12	ほる	掘る	2.1571	10	4	○	○
13	いし	石	1.5111	9	4	○	○
14	さげる	下げる	2.1515	9	4	○	○
15	いらっしやる	いらっしやる	2.1527	8	4		○
16	おーけすとら	オーケストラ	1.3231	8	4		
17	おさえる	押える	2.1564	8	4	○	○
18	おねがい	お願い	1.3041	8	4		
19	おはなしする	お話しする	2.3130	8	4		
20	しんぶんし	新聞紙	1.4110	8	4		
21	すぎる	過ぎる	2.1900	8	4	○	○
22	だめだ	駄目だ	3.1330	8	4	○	○

別表5 『窓際のトットちゃん』分布5度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	ぶーる	プール	1.4700	17	5		
2	さっこちゃん	サッコちゃん	15.2340	15	5		
3	はこ	箱	1.4513	15	5	○	○
4	ぐるーぶ	グループ	1.1981	14	5	○	○
5	いろ	色	1.5020	13	5	○	○
6	のじゅく	野宿	1.3331	12	5		
7	かぎね	垣根	1.4420	11	5		
8	ぬぐ	脱ぐ	2.1130	11	5	○	○
9	はこぶ	運ぶ	2.3830	11	5	○	○
10	しぜん	自然	1.5100	10	5	○	○
11	ちから	力	1.1400	10	5	○	○
12	のーと	ノート	1.4590	10	5	○	○
13	むすぶ	結ぶ	2.1554	10	5	○	○
14	うあいおりん	ヴァイオリン	1.4560	9	5		
15	おはなし	お話	1.3130	9	5		
16	こくばん	黒板	1.4530	9	5		○
17	しまう	仕舞う	2.1502	9	5	○	
18	しょうにまひ	小児麻痺	1.5860	9	5		
19	せなか	背中	1.5720	9	5	○	○
20	と	と (並列)	8.0020	9	5	○	○
21	ぴあの	ピアノ	1.4560	9	5	○	○
22	ようじ	用事	1.3460	9	5		○
23	わかい	若い	3.1660	9	5	○	

別表6 『窓際のトットちゃん』分布6度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	たいがく	退学	1.3340	14	6		
2	いす	椅子	1.4470	13	6		○
3	おてら	お寺	1.2630	13	6		
4	かわ	川	1.5250	13	6	○	○
5	なれる	成れる	2.1220	13	6	○	
6	みず	水	1.5130	13	6	○	○
7	やくそく	約束	1.3530	13	6	○	○
8	ゆか	床	1.4440	13	6		○
9	いけ	池	1.5270	12	6	○	○
10	なつやすみ	夏休み	1.3320	12	6	○	○
11	あな	穴	1.1840	11	6	○	○
12	いただく	いただく (補)	10.3770	11	6	○	
13	あたる	当たる	2.1120	10	6	○	○
14	かわいい	可愛い	3.3020	10	6	○	○
15	しきしゃ	指揮者	1.2452	10	6		
16	はずかしい	恥ずかしい	3.3012	10	6	○	○
17	むかし	昔	1.1642	10	6	○	○
18	よる	依る	2.1110	10	6	○	○

別表7 『窓際のトットちゃん』分布7度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	しごと	仕事	1.3800	17	7	○	
2	さんぽ	散歩	1.3371	15	7		
3	ちんどん屋さん	チンドン屋さん	1.2418	15	7		
4	なる	なる (補)	10.1220	14	7	○	○
5	ひく	弾く	2.3370	14	7		○
6	ゆめ	夢	1.3002	14	7	○	○
7	らんどセル	ランドセル	1.4514	14	7		○
8	おく	おく (補)	10.1515	13	7	○	
9	へや	部屋	1.4430	13	7	○	○

『窓ぎわのトットちゃん』語彙12 (田島)

10	ありがとう	有り難う (挨拶)	4.3310	12	7	○	○
11	おねえさん	お姉さん	1.2140	12	7		
12	じゆうだ	自由だ	3.3460	12	7	○	
13	とびこむ	飛び込む	2.1531	12	7	○	○
14	みじかい	短い	3.1920	12	7	○	○
15	むかう	向かう	2.1700	12	7	○	○
16	かく	掻く	2.3391	11	7		○
17	つく	つく (形)	10.1560	11	7	○	○
18	もぐる	潜る	2.1516	11	7	○	○

別表8 『窓際のトットちゃん』分布8度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	しっぱ	尻尾	1.5720	16	8		
2	におい	臭い・匂い	1.5040	16	8	○	○
3	きのう	昨日	1.1642	15	8		○
4	にほん	日本	15.2590	13	8		○
5	はく	履く	2.3330	13	8		○
6	せき	席	1.1720	12	8	○	○
7	とぶ	飛ぶ	2.1523	12	8	○	○
8	どっち	どっち	1.1000	12	8	○	
9	のこる	残る	2.1240	12	8	○	○
10	ふしぎだ	不思議だ	3.3060	12	8	○	○
11	よこ	横	1.1730	12	8	○	○
12	いる	入る	2.1530	11	8		
13	おかしい	可笑しい	3.3010	11	8	○	○
14	とおる	通る	2.1524	11	8	○	○

別表9 『窓際のトットちゃん』分布9度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	かいだん	階段	1.4420	18	9	○	○
2	どあ	ドア	1.4460	18	9	○	○
3	あめりか	アメリカ	1.2590	17	9		
4	よる	夜	1.1635	17	9	○	○
5	くび	首	1.5710	16	9	○	○
6	たべもの	食べ物	1.4300	16	9	○	○
7	さ	さ (終助)	8.0050	15	9	○	○
8	すばらしい	素晴らしい	3.1340	15	9	○	○
9	ちゃ	ちゃ (連語)	20	14	9		
10	でかける	出かける	2.1521	14	9	○	○
11	のぞく	覗く	2.3090	14	9	○	○
12	ひかる	光る	2.5010	14	9	○	○
13	おおい	多い	3.1950	13	9	○	○
14	こまる	困る	2.3010	13	9	○	○
15	じめん	地面	1.1750	13	9	○	
16	なんど	何度	1.1950	13	9		
17	やっと	やっと	3.1650	13	9	○	○
18	あたり	辺り	1.1780	12	9	○	○
19	は	歯	1.5760	12	9	○	○

別表10 『窓際のトットちゃん』分布10度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	ともえがくえん	トモエ学園	15.2630	17	10		
2	いける	いける	2.1310	16	10	○	○
3	かなしい	悲しい	3.3011	16	10	○	○
4	つまり	つまり (接続)	4.1150	16	10	○	○
5	きこえる	聞こえる	2.3092	14	10	○	○
6	じょうずだ	上手だ	3.3050	14	10	○	
7	むね	胸	1.5720	14	10	○	○
8	かなり	かなり	3.1993	13	10	○	
9	さあ	さあ (感動)	4.3200	13	10	○	
10	せ	背	1.5720	13	10	○	○
11	はず	筈	1.3066	13	10	○	○
12	ひとたち	人達	1.2020	13	10	○	

別表11 『窓際のトットちゃん』分布11度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	なみだ	涙	1.5770	18	11	○	○
2	はな	花	1.5530	17	11	○	○
3	ちゃんと	ちゃんと	3.1130	16	11	○	○
4	ありがたい	有り難い	3.3012	15	11		
5	くらす	クラス	1.2810	15	11	○	○
6	あまり	余り (副詞)	3.1993	14	11	○	
7	うまい	巧い・旨い	3.3050	14	11	○	○
8	くつ	靴	1.4260	14	11	○	○
9	うれしそうだ	嬉しそうだ	3.3010	13	11		
10	こっち	こっち	3.1000	13	11	○	○
11	ざんねんだ	残念だ	3.3012	13	11	○	
12	すっかり	すっかり	3.1980	13	11	○	○
13	たしかだ	確かだ	3.3060	13	11	○	○
14	たちあがる	立ち上がる	2.1513	13	11		○
15	どンドン	どンドン	3.1501	13	11	○	○
16	ぼく	僕	1.2000	13	11	○	○
17	もどる	戻る	2.1527	13	11	○	○

別表12 『窓際のトットちゃん』分布12-17度数上位語

	見出語	漢字表記	Wc	度数	分布	金	申
1	さがす	探す	2.3060	18	12		○
2	そう	そう (接尾)	1.1300	18	12		
3	たいせつだ	大切だ	3.3700	17	12	○	○
4	とつても	とつても (副詞)	3.1993	17	12		
5	うつくしい	美しい	3.5020	16	12	○	○
6	せいとたち	生徒達	1.2410	16	12		
7	つれる	連れる	2.1556	16	12	○	○
8	なぜ	何故	3.1100	16	12	○	○
9	もの	もの (終助)	8.0050	16	12	○	
10	うまれる	生まれる	2.5810	15	12		○
11	おじぎ	お辞儀	1.3390	15	12		
12	おと	音	1.5030	15	12	○	○
13	さて	然て	4.1170	15	12		
14	とおい	遠い	3.1920	15	12	○	○
15	となり	隣り	1.1771	15	12	○	○
16	かんがえ	考え	1.3061	18	13	○	○
17	なんか	なんか (副助)	8.3160	18	13	○	○
18	りょうて	両手	1.5730	18	13		○
19	しずかだ	静かだ	3.5030	17	13	○	
20	つつこむ	突っ込む	2.1531	17	13		○

『窓ぎわのトットちゃん』語彙12 (田島)

21	おちる	落ちる	2.1540	16	13	○	○
22	こんど	今度	1.1671	16	13	○	○
23	かっこう	格好・恰好	1.1800	15	13		
24	すごい	凄い	3.1400	15	13	○	○
25	すがた	姿	1.1800	18	14	○	○
26	じゆうがおか	自由が丘	15.2590	17	14		
27	なあ	なあ (終助)	8.0050	17	14	○	○
28	のせる	乗せる	2.1541	17	14	○	○
29	まず	先ず	3.1650	17	14	○	○
30	あげる	上げる	2.3770	18	15	○	○
31	のる	乗る	2.1541	18	15	○	○
32	よそ	余所	1.1771	18	15	○	○
33	おもいだす	思い出す	2.3050	17	15	○	○
34	きつと	きつと	4.3110	17	15	○	○
35	しかも	然かも	4.1110	17	15	○	
36	ちょうど	丁度	3.1140	16	15		○
37	しろい	白い	1.5020	18	16	○	
38	せつめいする	説明する	2.3136	18	17	○	○