

『窓ぎわのトットちゃん』語彙13

——中間の語彙の品詞別構成——

田 島 毓 堂

0. はじめに

前稿『『窓ぎわのトットちゃん』語彙12——中間の語彙について（試案）——』において、いかなる語彙を中間の語彙と指定するかについて試案を示して、その数量的分布という視点から、若干の考察を行った。

1. 前稿で分かったこと

極めて素朴な試論であった。しかし、従来、上位語については早くから、又、度数1の語についても検討されてきたが、「中間の語彙」は、その概念自体問題になったことがないので、何を「中間の語彙」とするかについて先ず試案を提出した。

それに抛れば、語彙のどういう部分を対象にすべきか、つまり、記号を含めた全部なのか、自立語だけなのか、自立語付属語を含めたもの（記号を除いた自立語と付属語）なのかを決める必要があった。検討の結果としては、記号を除いた自立語と付属語を対象にすることが適切と判断され、それから上位語と、最低頻度語を除いた部分を中間の語彙と考えて論を進めるのが、現時点では適切だと判断した。

具体的に言えば、『窓ぎわのトットちゃん』語彙では次の通りである。記号を除くと、異なり語数4528語、61120度数である。上位語は度数統計表において、累積使用率80%に最も近い度数階級までの343語49045度数（累積使用率80.24%、異なり語率7.58%）とする。また、最低頻度語、即ち、度数1の語は2183語2183度数（使用率3.57%、異なり語率は48.21%）である。この両

者を除くと、異なり語数2002語9892度数（使用率16.18%、異なり語率44.21%）が中間の語彙となる。度数、2から18までの語である。

このように対象を定めた上でそれを幾つかの観点から、観察した。

すなわち、上記の通り、それが異なり語・延べ語においてどの程度の割合を占めるのか、『窓ぎわのトットちゃん』語彙では、度数2から18までの語、異なり語率、44.21%、使用率16.18%であった。また、これを、同様の基準で『南の島のティオ』語彙で選定すると、それは度数2から15までの語、異なり語率42.04%、使用率は15.72%とほぼ同様な姿を見せる。さらに、多くの語彙について検証していく必要があるが、異なり語で、4割強、使用率は1割5分前後ということになる。上位語は使用率80%と決めてあるのだが、異なり語はともに7%前後、度数1の語は異なり語率5割弱、使用率は3.5%前後になる。

また、中間の語彙が含む意味分野はそのコード数で『窓ぎわのトットちゃん』語彙では627コード（全体の72.74%）であり、『南の島のティオ』語彙では645コード、68.91%がある。つまり、7割前後のコードが含まれている。これは上位語の3倍強、度数1の語よりは1割強少ない。

さらに、その分布について見た。『窓ぎわのトットちゃん』語彙についてのみであるが、その章別分布を見ると、理論的には、度数2から18の語であるから、分布は1から18までになるが、実際には1から17まで、語数は多くても分布の小さいものと広く分布するものの両者が見られた。ただ、分布1から5までで8割以上になる。分布13から17までの語も29語有る。『南の島の

ティオ』語彙は章が10しかないので同じようには行かないが、念のために調べておくと、次のようになる。前稿に載せなかったので、次に掲げる。

『南の島のティオ』語彙の中間語彙は度数2から15までの1918語であるが、度数ごとの分布は次の通りである。

度数	2	3	4	5	6	7	8
語数	713	320	223	138	123	90	75
度数	9	10	11	12	13	14	15
語数	61	40	37	29	27	19	23

分布別の語数を示す。

分布	1	2	3	4	5	6	7
語数	442	691	315	212	119	73	44
分布	8	9	10				
語数	15	7	0				

上の表のように、分布は1から9までである。分布1から5までがやはり8割以上になる。分布10というのがないことは、『窓ぎわのトットちゃん』語彙で18がないのとよく似ている。

この中間語彙は意味分野における分布の上からも、章別の分布という事からも、かなり広く分布していることになる。広く分布して、基幹語彙的なものもあるが、上位語よりは基幹語彙的性格は当然薄いことも見て取れる。これこそが数量的観察から中間の語彙であるという性質のものがあるわけではなかった。

2. 中間の語彙の品詞分類

2.1 『窓ぎわのトットちゃん』語彙の品詞構成

以上の考察をふまえて、前稿で指定した中間の語彙の意味分野別構造を観察してみようと思うが、その前に、上記の上位語、中間語彙、最低頻度語（度数1の語）の品詞別の構造について概観しておこう（以下、「上位（語）」「中間」「下位（語）」と称することがある）。品詞別に見て何か特徴があるのか無いのかを確かめておこうと思う。

表1に、全体（記号を除いた全語彙＝除記号）

表1 『窓ぎわのトットちゃん』語彙品詞別構成（全）

	異語数	異語率 (%)	延語数	延語率 (%)
名詞	2346	51.81	13669	22.36
固有名詞	231	5.10	2013	3.29
他	3	0.06	4	0.00
動詞	1034	22.83	8039	13.15
補助動詞	37	0.81	1872	3.06
連語	15	0.33	455	0.74
形容詞	140	3.09	1481	2.42
形容動詞	191	4.21	611	0.99
副詞	280	6.18	2228	3.64
連体詞	17	0.37	890	1.45
接続詞	35	0.77	726	1.18
感動詞	76	1.67	233	0.38
体の類	2580	56.97	15686	25.66
用の類	1086	23.98	10366	16.96
相の類	628	13.86	5210	8.52
その他	111	2.45	959	1.56
助詞	87	1.92	22177	36.28
助動詞	23	0.50	6682	10.93
接辞他	13	0.28	40	0.06
自立語	4405	97.28	32221	52.71
付属語	123	2.71	28899	47.28
全体	4528	100.00	61120	100.00

の品詞別構成を示す。異なり語数・その率・延べ語数・その率を示す。そのほかに、1体の類、2用の類、3相の類、4その他類、および助詞・助動詞・接辞についても示しておく。助詞・助動詞・接辞は1から4の類に含まれないので、示したのである。なお、接辞については、語基について一語と認めるのを原則とするが、それを一語とするのが不自然でそうしなかったものを別に接辞としてカウントした。このことについては、別に論じなければならない。

表2は、上位語の品詞別の構成を異なり語数・その上位語内での異語率・全体（記号を除いた全語彙）での異語率・延べ語数・上位語内での延語率・全体での延語率を示す。

表3は、中間の語彙について、品詞別の構成を異なり語数・その中間語彙内での異語率・全体での異語率・延べ語数・中間語彙内での延語率・全体での延語率を示す。

表4には、下位語（度数1の語）について、品詞別の構成を異なり語数・下位語内での異語率・全体での異語率・延べ語数・中間語彙内での延語

『窓ぎわのトットちゃん』語彙13 (田島)

表2 『窓ぎわのトットちゃん』上位語品詞別構成

	異語数	上位語内 異語率 (%)	全体 異語率 (%)	延語数	上位語内 延語率 (%)	全体 延語率 (%)
名詞	122	35.56	2.69	7823	15.95	12.79
固有名詞	12	3.49	0.26	1438	2.93	2.35
他	0	0.00	0.00	0	0.00	0.00
動詞	67	19.53	1.47	5159	10.51	8.44
補助動詞	12	3.49	0.26	1744	3.55	2.85
連語	4	1.16	0.08	395	0.80	0.64
形容詞	17	4.95	0.37	975	1.98	1.59
形容動詞	3	0.87	0.06	91	0.18	0.14
副詞	26	7.58	0.57	1307	2.66	2.13
連体詞	9	2.62	0.19	862	1.75	1.41
接続詞	8	2.33	0.17	603	1.22	0.98
感動詞	1	0.29	0.02	21	0.04	0.03
体の類	134	39.06	2.95	9261	18.88	15.15
用の類	83	24.19	1.83	7298	14.88	11.94
相の類	55	16.03	1.21	3235	6.59	5.29
その他	9	2.62	0.19	624	1.27	1.02
助詞	46	13.41	1.01	21974	44.80	35.95
助動詞	16	4.66	0.35	6653	13.56	10.88
接辞他	0	0.00	0.00	0	0.00	0.00
自立語	281	81.92	6.20	20418	41.63	33.40
付属語	62	18.07	1.36	28627	58.36	46.83
全体	343	100.00	7.57	49045	100.00	80.24

表3 『窓ぎわのトットちゃん』中間語彙品詞別構成

	異語数	中間語内 異語率 (%)	全体 異語率 (%)	延語数	中間語内 延語率 (%)	全体 延語率 (%)
名詞	1005	50.19	22.19	4627	46.77	7.57
固有名詞	85	4.24	1.87	441	4.45	0.72
他	1	0.04	0.02	2	0.02	0.00
動詞	481	24.02	10.62	2394	24.20	3.91
補助動詞	20	0.99	0.44	123	1.24	0.20
連語	9	0.44	0.19	58	0.58	0.09
形容詞	73	3.64	1.61	456	4.60	0.74
形容動詞	89	4.44	1.96	421	4.25	0.68
副詞	132	6.59	2.91	799	8.07	1.30
連体詞	4	0.19	0.08	24	0.24	0.03
接続詞	18	0.89	0.39	114	1.15	0.18
感動詞	44	2.19	0.97	181	1.82	0.29
体の類	1091	54.49	24.09	5070	51.25	8.29
用の類	510	25.47	11.26	2575	26.03	4.21
相の類	298	14.88	6.58	1700	17.18	2.78
その他	62	3.09	1.36	295	2.98	0.48
助詞	30	1.49	0.66	192	1.94	0.31
助動詞	6	0.29	0.13	28	0.28	0.04
接辞他	5	0.24	0.11	32	0.32	0.05
自立語	1961	97.95	43.30	9640	97.45	15.77
付属語	41	2.04	0.90	252	2.54	0.41
全体	2002	100.00	44.21	9892	100.00	16.18

表4 『窓ぎわのトットちゃん』度数1の語品詞別構成

	異語数	度数1内 異語率 (%)	全体 異語率 (%)	延語数	度数1内 延語率 (%)	全体 延語率 (%)
名詞	1219	55.84	26.92	1219	55.84	1.99
固有名詞	134	6.13	2.95	134	6.13	0.21
他	2	0.09	0.04	2	0.09	0.00
動詞	486	22.26	10.73	486	22.26	0.79
補助動詞	5	0.22	0.11	5	0.22	0.00
連語	2	0.09	0.04	2	0.09	0.00
形容詞	50	2.29	1.10	50	2.29	0.08
形容動詞	99	4.53	2.18	99	4.53	0.16
副詞	122	5.58	2.69	122	5.58	0.19
連体詞	4	0.18	0.08	4	0.18	0.00
接続詞	9	0.41	0.19	9	0.41	0.01
感動詞	31	1.42	0.68	31	1.42	0.05
体の類	1355	62.07	29.92	1355	62.07	2.21
用の類	493	22.58	10.88	493	22.58	0.80
相の類	275	12.59	6.07	275	12.59	0.44
その他	40	1.83	0.88	40	1.83	0.06
助詞	11	0.50	0.24	11	0.50	0.01
助動詞	1	0.04	0.02	1	0.04	0.00
接辞他	8	0.36	0.17	8	0.36	0.01
自立語	2163	99.08	47.76	2163	99.08	3.53
付属語	2	0.91	0.44	20	0.91	0.03
全体	2183	100.00	48.21	2183	100.00	3.57

率・全体での延語率を示す。

表2・表3・表4で、体・用・相・その他・助詞・助動詞・接辞について示すことは、表1同様である。

表1は全体(除記号)での比率である。表2・3・4とは明らかに異なる。

まず、体の類(名詞・固有名詞を加えたもの)の延語率(=使用率)は上位語・中間語・下位語の各層内部での比率では、下位語(62.07%)が最も高い比率で、中間51.25%・上位18.88%の順にかなりの差を持って並ぶ。しかし、それを全体における使用率から見るならば、丁度逆になり(下2.21%・中8.29%・上15.15%)、特に固有名詞に限れば、上位語に7割出現する。

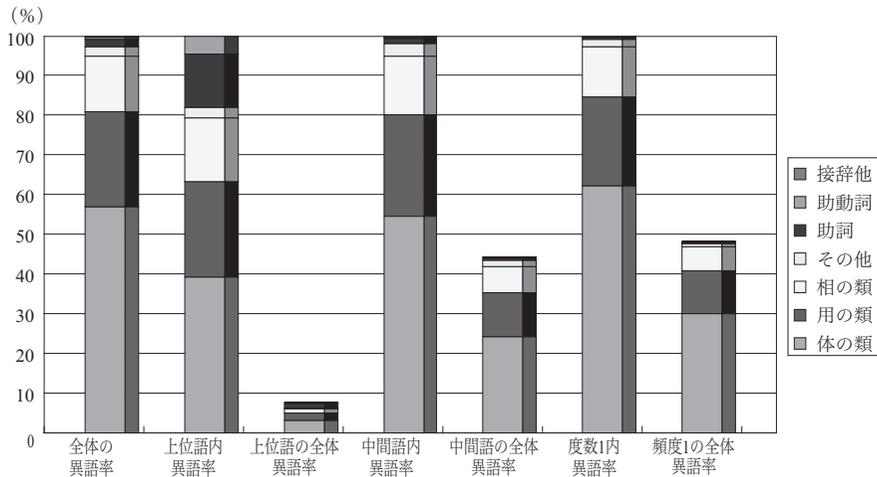
体の類の異なり語の率についてみよう。全体では5割を超える(56.97%)なのであるが、下位語29.92%・中間24.09%・上位2.95%の順に下がり、各層内部での比率も同様に下62.07%・中54.49%・上39.06%の順位である。特に、上位語はその内部でも39%と他よりかなり低い、全体で

の率は極端に低い。つまり、上位語では、体の類の語は少数しかないが、高頻度語でしめられ、そのカバーする使用率が極めて高いということである。ただし、全体においても体の類(名詞類)は25.66%である。そしてそれは自立語のほぼ半数に当たる。

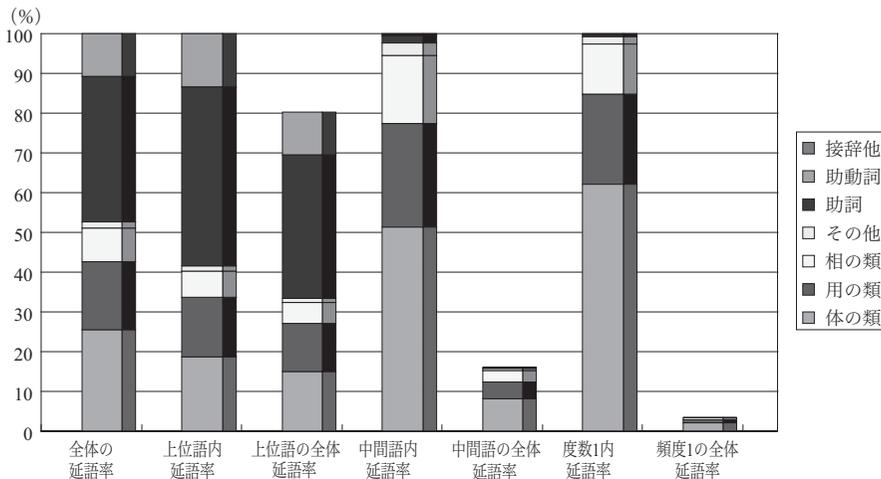
グラフ1に体の類・用の類・相の類・その他の類・助詞・助動詞・接辞の異なり語率(全体のそれと、それぞれ上位語・中間語・下位語のそれぞれについて、各層内での比率と全体における比率)、グラフ2に体の類・用の類・相の類・その他の類・助詞・助動詞・接辞の使用率(全体における比率と、それぞれの層内における比率)を示した。

各層内での、異なり語率と、使用率は比例する(上→中→下)が、全体に対する比率は逆比例する(異語率下→中→上、使用率上→中→下)。上位語・下位語には大差がある。中間語彙は体の類については、全体と層内の異語率・全体と層内の使用率のすべてで上位語と下位語の間にある。

『窓ぎわのトットちゃん』語彙13 (田島)



グラフ1 『窓ぎわのトットちゃん』の異語率の積み上げグラフ



グラフ2 『窓ぎわのトットちゃん』の延語率の積み上げグラフ

グラフ中の項目名とその意味について記しておく。「全体の異語率」「全体の延語率」は対象語彙全体における体・用・……それぞれの異なり語率・延べ語率（使用率）。「上位語内異語率」「中間語内異語率」「下位語内異語率」はそれぞれ上位語内、中間語彙内、下位語内での体の類・用の類・……の異なり語数のそれぞれの層内における比率を示す。「上位語の全体異語率」「中間語の全体異語率」「下位語の全体異語率」は上位語・中間語彙・下位語それぞれの全体に対する、体の類・用の類・……の異なり語率を示す。「……延語率」の場合は上記「異語率」の部分で「延語率」と読み替える。

用の類（動詞等）の比率は、各層内部での異なり語率は余り変わらない。語数は中間と下位語でほぼ等しく、上位語で少ない。中間は補助用言の数が多。層内の率で言えば、上位語の補助用言

は数・率とも大変高い。用の類は、全体に対する比率が、中間・下位はほぼ等しい（中11.26%・下10.88%）が、上位語では非常に小さい（1.83%）。使用率では全体で16%強である。全体でも、

各層内部でも、体の類に次ぐ。全体に対する比率では、上位語で12%弱、下位語では1%にも満たず、中間は4%強に過ぎない。名詞ほどではないが、やはり、高頻度の動詞が、上位語に含まれ、使用率を上げている。層内の使用率は、中間26.03%が一番大きく、下位22.58%・上位14.88%の順である。層内での異なり語率はそんなに差はないが、中間25.47%・上位24.19%・下位22.58%の順であるから、その順とも異なる。用の類では層内の使用率・異なり語率とも、中間の語彙においていずれも高く、全体に対する異なり語率でも高いことは注意しておいてよからう。体の類とはここが違っている。中間語彙で用の類が優位にある。

相の類（形容詞・形容動詞・副詞・連体詞）は異なり語では各層内部で言えば、上位語16.03%・中間14.88%・下位12.59%の順であるが、差は僅かである。全体に対する比率では上位で小さい1.21%ことが目につくが（中6.58%・下6.07%）、このことは体の類でも用の類でも同じである。上位語は、異なり語数が少ないから、その異なり語の全体に占める割合は非常に小さい。延べでは全体に対する比率は、上位5.29%・中間2.78%・下位0.44%の順、層内では中間17.18%・下位12.59%・上位6.59%の順で、層内使用率では中間語彙で高いのが目立つ。全体に対する異なり語率でも下位語と差は僅かだが中間語彙で一番高い。相の類でも、中間語彙は総体的に優位である。

その他の類（接続詞・感動詞等）は、全体で異なり語率2.45%と少ない。上位語と、中間内部ではそれより僅かに高い比率を示す（上2.62%・中3.09%・下1.83%）。下位語ではそれより少し低率である。中間語彙は全体に対する異なり語率、層内での異なり語率および使用率において相対的に高い。下位語では、当然であるが、全体に対する使用率は極めて小さい。一回だけ使われるような「その他の類」の語は語数（40語）にすればそんなに少ないわけではないが、使用率から見ればきわめて小さい（0.06%）。全体に対する使用率以外ではいずれも中間語彙で率が高くなっている。つまり、その他の類においても、中間語彙が

かなり優位にある。

本稿の対象には付属語も入る。付属語と自立語の延べでの比率は47.28対52.71、つまり若干自立語が多い。その全体に対する比率で見れば、助詞36.28%（異語率1.92%）、助動詞10.93%（異語率0.50%）である。上位語に延べで46.83%（助詞35.95%、助動詞10.88%）が出現し、中間には、使用率は併せて0.35%、下位には0.01%である。圧倒的に上位語に多く出現する。語数は全部で123語しかなく、一語平均の使用回数も235.0回に達し、自立語の7.3回に比べて圧倒的なものだから、使用度数から見れば、殆ど上位語に出現するのは当然であろう。助詞・助動詞については数値は異なるが、割合をグラフに書けばほぼ同じ形になる。中間・下位にはほとんど無いと言ってよからう。一括して、付属語と扱えばいいと思う。

2.2 『南の島のティオ』語彙との対照

さて、これを『南の島のティオ』語彙と対照させてみる。どの項目も驚くほど類似した数値を持つ。

名詞だけ見ると、トット語彙では22.36%、ティオ語彙は25.85%と差があるように見えるが、上記には固有名詞を含んでいないので、それを名詞類として一括すると、それぞれ25.66%、26.09%と僅かな差である。動詞は13.15%と13.97%、補助動詞と一括すると16.96%、16.87%と0.1%の差しかない。形容詞類（相の類）は8.52%と8.46%、その他の類だけが、1.56%と0.88%と他に比べて少し目立つ。助詞は36.28%と36.77%、助動詞が10.93%と10.90%となっている。自立語は52.71%と52.30%、付属語は47.28%と47.70%となっており、殆ど等しいと言っていいくらいである。

『南の島のティオ』全体の語彙の品詞別の構成を表5に示す。

また、本稿の主題である中間の語彙についても見ておこう。中間の語彙もやはり酷似している（表6）。体の類の全体に対する比率はトット語彙8.29%、ティオ語彙8.11%である。用の類は4.21%体4.25%、相の類は2.78%対2.76%、その他の類は0.48%と0.29%、その他の類は多少差が目

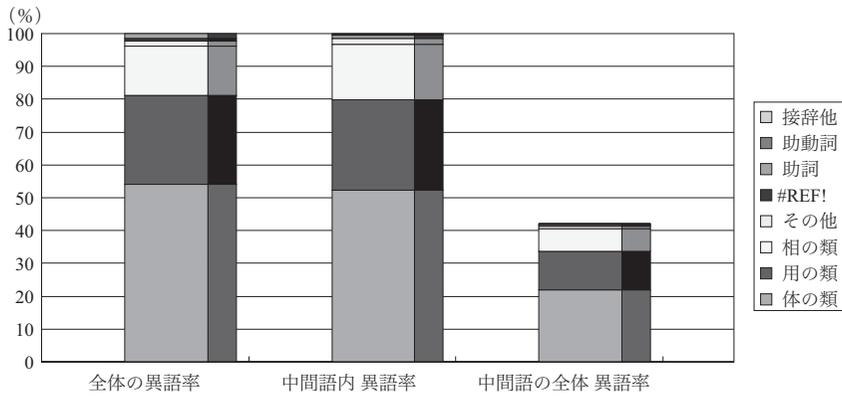
『窓ぎわのトットちゃん』語彙13 (田島)

表5 『南の島のティオ』語彙品詞別構成 (全)

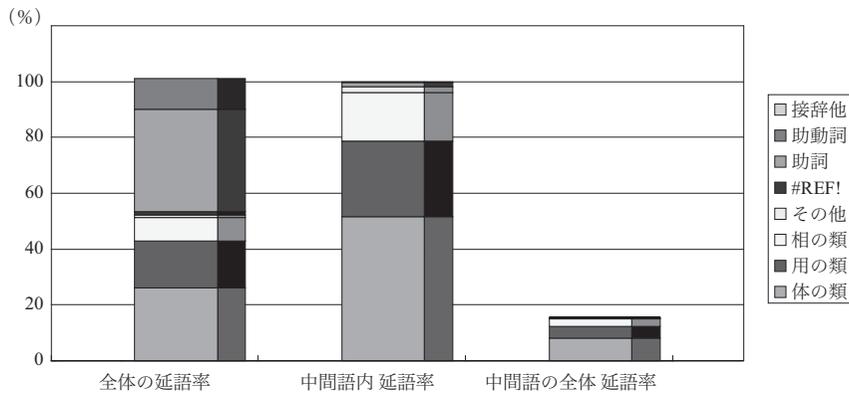
品詞	全異語	率全異 (%)	全延語	率全延 (%)
名詞	2435	53.38	14142	25.85
固有名詞	18	0.39	121	0.22
他	0	0.00	0	0.00
動詞	1209	26.50	7644	13.97
補助動詞	23	0.50	1549	2.83
連語	5	0.11	36	0.07
形容詞	146	3.20	1143	2.09
形容動詞	238	5.22	643	1.18
副詞	277	6.07	1844	3.37
連体詞	18	0.39	999	1.83
接続詞	35	0.77	351	0.64
感動詞	35	0.77	130	0.24
体の類	2466	54.06	14277	26.09
用の類	1237	27.12	9228	16.87
相の類	679	14.88	4629	8.46
その他	70	1.53	481	0.88
助詞	80	1.75	20118	36.77
助動詞	23	0.50	5961	10.90
接辞他	7	0.15	17	0.03
自立語	4452	97.59	28616	52.30
付属語	110	2.41	26096	47.70
全体	4562	100.00	54712	100.00

表6 『南の島のティオ』中間語彙の品詞別構成

品詞	異語数	中間語内 異語率 (%)	全体 異語率 (%)	延語数	中間語内 延語率 (%)	全体 延語率 (%)
名詞	997	51.98	21.85	4382	50.94	8.01
固有名詞	7	0.36	0.15	52	0.60	0.10
他	0	0.00	0.00	0	0.00	0.00
連語等	1	0.05	0.02	2	0.02	0.00
動詞	512	26.69	11.22	2213	25.73	4.04
補助動詞	12	0.63	0.26	99	1.15	0.18
連語動	2	0.10	0.04	13	0.15	0.02
形容詞	80	4.17	1.75	427	4.96	0.78
形容動詞	108	5.63	2.37	444	5.16	0.81
副詞	128	6.67	2.81	596	6.93	1.09
連体詞	6	0.31	0.13	42	0.49	0.08
接続詞	18	0.94	0.39	83	0.96	0.15
感動詞	16	0.83	0.35	76	0.88	0.14
体の類	1005	52.40	22.03	4436	51.57	8.11
用の類	526	27.42	11.53	2325	27.03	4.25
相の類	322	16.79	7.06	1509	17.54	2.76
その他	34	1.77	0.75	159	1.85	0.29
助詞	22	1.15	0.48	130	1.51	0.24
助動詞	3	0.16	0.07	24	0.28	0.04
接辞他	6	0.31	0.13	19	0.22	0.03
自立語	1887	98.38	41.36	8429	97.98	15.41
付属語	31	1.62	0.68	173	2.01	0.32
全体	1918	100.00	42.04	8602	100.00	15.72



グラフ3 『南の島のティオ』の異語率の積み上げグラフ



グラフ4 『南の島のティオ』の延語率の積み上げグラフ

立つ。助詞は、0.31%と0.24%、助動詞はともに0.04%である。自立語15.77%対15.41%、付属語は0.41%対0.31%である。酷似していると評しても良からう。

こういう枠組みで比較してみると、上位語においても同様である。と言うことは、結局下位語においても同様であることを示している。確かに、微細な差はあるが、今用いた、「体・用・相・その他・助詞・助動詞」「自立語・付属語」という品詞的枠組みでは両語彙は大変似ている。さら

に、他の語彙との比較が必要であるが、この上位語・中間語彙という対象の定め方もその外れではないということが言えると思う。

中間語彙をこのように定めることについては一応合理的だと言えようが、中間語彙としての特徴は別の問題である。単に数量の問題としてみるのには限界がある。

本稿では、意味分野別構造まで述べられなかった。次稿を期する。