

絵本テキストにおける高頻度語彙の分析

Analysis of high-frequency words in picture books

平博順† 藤田早苗† 小林哲生†
Hirotoishi Taira Sanae Fujita Tessei Kobayashi

1. はじめに

子どもがどのように言語能力を発達させていくかという問題は、人間を対象とした自然科学研究の重要な問題のひとつである。この問題について、これまでは主に発達心理学や心理言語学などの分野で研究が進められてきた。しかし、幼い子どもを研究対象とするため、データ取得には多大な困難が伴い、言語発達過程を十分に明らかにできるほどデータが得られているとは言い難い状況である(小林・永田, 2012a)。

こうした状況を打破するために、幼児が覚えた語彙を養育者に Web サービスを介して投稿してもらうことで、各月齢における子どもが覚える語彙の傾向を調査するなど、新しい手法での研究も最近活発化してきている(小林・永田, 2009; 小林 他, 2012b)。

子どもは 1 歳頃に初語を発し始め、1 歳半過ぎから「語彙爆発(vocabulary spurt)」と呼ばれる語彙急増期を迎える。その後も子どもは着実に語彙を伸ばしていき、英語話者を対象とした研究では、6 歳の時点で 1 万語を獲得しているとの報告もある(Anglin, 1993)。こうした語彙の成長を促す要因として、養育者から幼児への話しかけを重視する研究者は多い。しかし、それだけでは幼児が接する語彙数は少なく、子どもの急速な言語発達は説明できないと考える研究者もいる(Senechal et al., 1996)。

幼児が言葉に触れる機会は、養育者からの話しかけ以外にも、大人同士の会話を傍らで聞くこと(overhearing)や、テレビ・ラジオなどのメディアからの入力、幼児への絵本の読み聞かせなどが考えられる。中でも、絵本の読み聞かせは、言語/認知発達へのポジティブな影響を与えるという知見が得られており、幼児に対する言語の入力機会として注目されている(Scarborough and Dobrich, 1994; Bus et al., 1995; Collins, 2010)。これらの研究では、読み聞かせ場面における親の行動傾向から子どもの言語発達への影響を探る試みが主に行われてきた。しかし、読み聞かせ場面でのような語彙が子どもに輸入されているかといったことや具体的に絵本にどのような語彙が含まれているのかを研究した例はなく、幼児の各発達段階において理解・発話する語彙との相関なども分かっていない。

本稿では、そのような研究を進める上での基礎的なデータとして、実際の絵本に含まれる単語の出現頻度を調査した。調査対象は売り上げ数の多い約 500 冊の絵本とした。人手ですべての単語を調査することはコストが大きいいため、形態素解析器を使って自動的に単語区切り、品詞付与を行った。

特に日本語の絵本テキストにおいては、通常漢字で表記される単語でもひらがな・カタカナで表記されることが多いため、新聞記事を初めとする一般向けテキストを主な対象として開発された形態素解析器を使用すると、適切な単語区切りが期待できない。そのため、本研究では、ひらがな・カタカナ表記のデータを大量に用意し、解析モデルが再学習可能な形態素解析器にモデルを再学習させることで、ひらがな・カタカナ表記の語でも高精度に解析できるようにした。このような形態素解析器を利用して、絵本テキストに高頻度に出現する基本的な語彙を特定し、品詞ごとに分析を行った。

以下、2 節において、絵本テキストの収集方法および形態素解析方法等について述べ、3 節において、得られた結果および考察について述べる。最後に 4 節で結論を述べる。

2. 方法

2.1 絵本テキストの収集

分析対象となる絵本は、日本語で書かれた絵本とした。その中で、多くの子どもに読まれている絵本を分析対象としたかったが、実際にそれを調査するのは困難であったため、大規模書店における売上冊数上位の絵本を対象とすることで、それに代えることにした。具体的には 2011 年の紀伊國屋書店グループ全体において、絵本カテゴリに入る書籍を対象とした。表 1 に絵本カテゴリの売上げランキング上位 20 位までを示す。

順位	書籍名	著者名	出版社	初版年
1	ぐりとぐら	中川 李枝子 大村 百合子	福音館書店	1963
2	バムとケロのむりごや	島田 ゆか	文溪堂	2011
3	だるまさんが	かがくい ひろし	プロズ新社	2008
4	うんこ!	サトシ 西村 敏雄	文溪堂	2010
5	はらぺこあむし	エリック カール (もりひさし訳)	偕成社	1976
6	だるまさんと	かがくい ひろし	プロズ新社	2009
7	だるまさんの	かがくい ひろし	プロズ新社	2008
8	きんぎょが にげた	五味 太郎	福音館書店	1977
9	ちか100かいだてのいえ	いわい としお	偕成社	2009
10	100かいだてのいえ	いわい としお	偕成社	2008
11	もぐらバス	佐藤 雅彦 うちの ますみ	偕成社	2010
12	新幹線のたび へばやぶさのぞみ・さくらで	コマヤス カン	講談社	2011
13	こびとづかん	なばた としたか	長崎出版	2006
14	わたしと小鳥とすずと	金子 みすず	JULA出版局	1984
15	ぞうくんのあめふりさんぽ	なかの ひろたか	福音館書店	2004
16	うずらちゃんのかくれんぼ	きもと ももこ	福音館書店	1994
17	どうぞのいす	香山 美子 柿本 幸造	ひさかた チャイルド	1981
18	そらまめくんのベッド	なかや みわ	福音館書店	1997
19	よみかかせおはなし絵本	千葉 幹夫	成美堂出版	2011
20	ちいさなあなたへ	アリスン・マギー ピーター・レイノルズ (なががわ ちひろ訳)	主婦の友社	2008

表 1. 絵本売上ランキング (紀伊國屋書店、2011 年)

†日本電信電話(株) NTT コミュニケーション科学
基礎研究所, NTT Communication Science Laboratories,
E-mail: taira.hirotoishi@lab.ntt.co.jp

この絵本カテゴリ売上げランキングの上位 547 位までの本のうち、売り切れ等で入手ができなかったもの、および子どもを読者対象としていない絵本・詩集・お話集や、図鑑・ことわざ集といった絵本とは言い難いもの合計 42 冊を除いて、505 冊の絵本を収集した。そして、収集した絵本について、テキスト部分を電子化した。テキストの電子化は人手入力で行った。

2.2 ひらがな・カタカナに頑健な形態素解析器の作成

絵本のテキストでは、ひらがな・カタカナ表記の語の出現頻度が、通常のテキストに比べて非常に高い。そこで、ひらがな・カタカナ表記も高精度で解析が可能な形態素解析モデルを作成した。解析モデルの作成には、京都テキスト解析ツールキット kyTea (Neubig et al. 2011) を用いた。なお、品詞体系は IPA 辞書の体系を用いた。

学習コーパスとして、京大コーパス(黒橋・長尾, 1997)を IPA 品詞体系に変換した約 4 万文、Lexeed の例文約 4 万 5 千文と定義文約 7 万 5 千文を利用した。これらのコーパス中の漢字表記の各語について、読みの情報を用いてひらがな表記の語として展開し、追加データとした。そのため、実際の学習コーパスの規模は上記の約 2 倍となった。

また、辞書としては、NTT 基本語意味データベース Lexeed (笠原 他, 2004) の 2 万 7 千語、NTT 日英機械翻訳システム ALT-J/E の日本語辞書 (池原 他, 1997) 中の固有名詞約 20 万語、動植物約 5 千語を利用した。辞書についても読み情報を用いて、漢字表記をひらがな・カタカナ表記に展開したものも追加した。そのため、実際の辞書エントリ数は上記の約 3 倍である。

これらの学習コーパス・辞書を用いて、解析モデルの構築を行った。その結果、絵本からランダムサンプリングした 900 文に人手で正解データを付与したデータによる評価で、単語分割精度 97.7%、品詞付与精度 93.7% の高精度の形態素解析器が得られた。この解析器を使って、絵本テキストの形態素解析を行った。

2.3 高頻度に出現する基本単語の特定

出現頻度が高くても、特定の絵本に集中して出現する単語は基本単語とはいえないので、文書頻度 (Document Frequency; DF) が高いものを基本単語とみなした。本研究の場合、文書頻度は「ある単語が解析対象の 505 冊のうち何冊の本に出現したかのカウント」となる。文書頻度に基づき、絵本に高頻度で出現する基本単語の上位 2000 語を特定した。なお、文書頻度が同じだった場合は、延べの出現頻度の高い方を上位とした。

2.4 売れ筋絵本による基本単語 2000 語のカバー数の分析

また、絵本 505 冊について、売れ行き順に本を読んでいった場合と、ランダムな順に本を読んでいった場合で上位基本 2000 語に対して、どのくらいカバーできるかについてシミュレーション計算を行った。

2.5 ウェブ日誌法で取得された早期獲得語彙との対応

ウェブ日誌法(小林・永田, 2009)によって、ウェブサービスから得られた幼児の各月齢における獲得語彙のデータと

の比較を行った。このデータから、より早期に幼児が獲得する語彙上位 30 語が、上記の絵本における高頻度出現語彙としては、何番目に当たるものかを調べ、幼児が実際に獲得する語彙と絵本に頻出する語彙との関係について考察を行った。

3. 結果と考察

3.1 高頻度語彙の分析

505 冊の絵本に対して、ひらがな・カタカナ表記に頑健な形態素解析を行った結果、延べ 47 万 6370 語、異なり語数で 2 万 4615 語の単語を得た。文書頻度に基づきソートし、絵本テキストに高い頻度で現れる基本語彙を特定した。上位 2000 語の品詞別の割合は、名詞 48%、動詞 27%、副詞 8%、形容詞 5%、助詞 4%、感動詞 3%、接続詞 2%、記号 1%、連体詞 1% であった。

表 2 に全品詞に対する上位 10 単語、表 3~5 に、名詞・動詞・形容詞それぞれについての高頻度出現の上位 10 単語を示す。表 2 より、上位は助詞または助動詞で占められていることが分かる。また、表 3 より、名詞に関しては、

順位	単語	品詞	文書頻度	出現数
1	て	助詞	471	15826
2	の	助詞	471	14506
3	に	助詞	471	10966
4	た	助動詞	469	18990
5	は	助詞	466	14639
6	が	助詞	466	10480
7	だ	助動詞	454	6680
8	も	助詞	445	11206
9	も	助詞	432	4068
10	する	動詞	430	4545

順位	単語	品詞	文書頻度	出現数
1	の	名詞	381	2349
2	こと	名詞	295	1184
3	みんな	名詞	286	1049
4	ん	名詞	280	1173
5	それ	名詞	275	861
6	たち	名詞	244	1414
7	よう	名詞	244	670
8	僕	名詞	242	1342
9	さん	名詞	226	1394
10	何	名詞	225	576

表 2. 頻出語彙 (全品詞)

表 3. 頻出語彙 (名詞)

順位	単語	品詞	文書頻度	出現数
1	する	動詞	430	4545
2	なる	動詞	353	1664
3	言う	動詞	343	2629
4	行く	動詞	267	933
5	見る	動詞	267	816
6	ある	動詞	236	627
7	来る	動詞	222	586
8	食べる	動詞	203	795
9	出る	動詞	192	514
10	思う	動詞	186	443

順位	単語	品詞	文書頻度	出現数
1	いい	形容詞	272	730
2	ない	形容詞	177	365
3	良い	形容詞	165	304
4	大きい	形容詞	154	330
5	おいしい	形容詞	128	280
6	楽しい	形容詞	114	173
7	小さい	形容詞	95	182
8	長い	形容詞	80	145
9	うれしい	形容詞	79	116
10	優しい	形容詞	78	145

表 4. 頻出語彙 (動詞)

表 5. 頻出語彙 (形容詞)

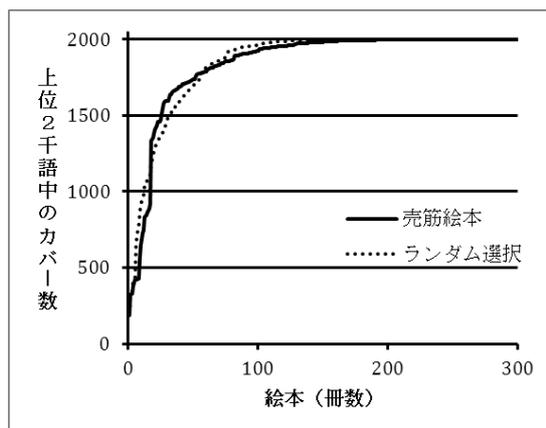


図 1. 絵本冊数に応じた上位 2 千単語中のカバー数

上位語については代名詞や非自立語が多いことが分かる。動詞・形容詞に関しては、生活する上で、基本的な単語が出現していた(表 4, 5)。なお参考のため、本稿末尾に 400 位までの語を示した。

3.2 売れ筋絵本による基本単語 2000 語のカバー数

505 冊の絵本を、販売数順に本を読んだ場合とランダムな順に本を読んだ場合で、上位 2000 語がどのくらいカバーできるかの推移についてシミュレーションした結果を、図 1 にグラフで示す。

図 1 より、売れ筋の絵本を順に読んでも、ランダムに選択しても 20 冊程度読むだけで上位 1000 単語、150 冊程度で上位 2000 単語がカバーできることが分かった。例えば、1 日 1 冊異なる本を選択して読み聞かせをすれば、約半年で上位 2000 単語に子どもが触れられることになる。

3.3 ウェブ日誌法で取得された早期獲得語彙との対応

表 6 に、ウェブ日誌法(小林・永田, 2009)による早期に幼児が獲得する語彙上位 30 語が、絵本における出現語彙としては何番目の出現単語かについての関係を示す。

この表を見ると、子どもが早期に獲得する傾向のある幼児語(まんま, ねんね, ぶーぶー等)は、絵本での頻度はあまり高くないように思われる。一方で、幼児語でない、子どもが使うような単語(ある, 手, お母さん等)は絵本でも頻度が高い傾向にあった。

順位	項目	平均獲得月齢	投稿数	絵本基本単語での順位
1	まんま	13.56	118	7584
2	あーあ(失敗)	15.55	37	19377
3	いないいないばあ	15.58	136	-
4	パパ	15.90	154	1182
5	ママ	15.96	198	667
6	はい	16.13	117	1069
7	ワンワン(犬)	16.34	188	2165
8	ないない(片づける)	16.53	61	-
9	バイバイ	16.54	143	18886
10	おっばい	16.81	83	2060
11	よいしょ	16.83	54	988
12	ねんね(寝る)	17.26	104	4364
13	くっく(靴)	17.68	60	-
14	ニャンニャン(猫)	17.82	107	9887
15	アンパンマン	18.07	185	-
16	お父さん	18.18	72	300
17	お母さん	18.18	87	122
18	ペットの名前	18.19	27	-
19	たっち(立つ)	18.32	43	-
20	ある	18.38	80	52
21	パン	18.43	67	420
22	どうぞ	18.43	121	632
23	あれっ	18.57	32	3570
24	うまい	18.65	36	414
25	バナナ	18.66	71	1210
26	お茶	18.69	75	1263
27	葉っぱ	18.81	64	467
28	手	18.82	38	128
29	いや	18.83	121	411
30	ぶーぶー(車)	19.00	79	6808

表 6. 幼児が早期に獲得する語彙(小林・永田, 2009)の絵本基本単語における順位

4. 結論

本研究では、子どもに対する言語入力の特徴を探るために、絵本を題材とした分析を行った。ひらがな・カタカナ表記に頑健な形態素解析器を作成、使用することによって、日本語の絵本に頻出する語彙のリストを自動的に作成することができた。こうした試みにより得られた絵本の頻出語彙リストは、これまで一度も作成されたことがなく、本研究分野における貴重な資料となる可能性が高い。ただし、単語の区切り方、品詞体系に何を使用するか等については議論の余地があり、今後、幼児の早期獲得語彙リストにおける単語区切りとの整合性についても検討を加えながら、絵本テキストにおける語彙と幼児が実際に獲得している語彙についての相関について分析を進めていきたい。

参考文献

- Anglin, J. M. (1993). Vocabulary development: a morphological analysis. *Monograph of the Society for Research in Child Development* **58** (10).
- Bus, A. G., van IJzendoorn, M. H. and Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: a meta-analysis on intergenerational transmission of literacy. *Review of Educational Research* **65**, 1-21.
- Scarborough, H.S. and Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review* **14**, 245-302.
- Senéchal, LeFevre, Hudson and Lawson (1996). Knowledge of picture-books as a predictor of young children's vocabulary. *Journal of Educational Psychology* **88**, 520-536.
- Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly*, **25**, 84-97.
- Neubig, G., Nakata, Y. and Mori, S. (2011). Pointwise Prediction for Robust, Adaptable Japanese Morphological Analysis, Proceedings of ACL 2011, Short Papers, 529-533.
- 池原悟, 宮崎正弘, 白井論, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦 (1997). 日本語語彙大系 CDROM 版. 岩波書店
- 笠原 要, 佐藤 浩史, Francis Bond, 田中 貴秋, 藤田 早苗, 金杉 友子, 天野 昭成 (2004). 「基本語意味データベース:Lexeed」の構築, 2004-NLC-159, 75-82
- 黒橋禎夫, 長尾眞 (1997). 京都大学テキストコーパス・プロジェクト, 言語処理学会第 3 回年次大会発表論文集, 115-118.
- 小林哲生, 永田昌明 (2009). ウェブを用いた幼児言語発達研究: 大規模縦断データの収集. 言語処理学会第 15 回年次大会発表論文集, 534-537
- 小林哲生, 永田昌明 (2012a). 日本語学習児の初期語彙発達. 情報処理, **53**(3), 229-235.
- 小林哲生, 南泰浩, 永田昌明 (2012b). 縦断および横断データを用いた幼児早期出現語の獲得月齢の特定. 言語処理学会第 18 回年次大会発表論文集, 743-746

付録

以下に頻出語彙上位 400 語までの語を示す。

順位	単語	品詞	文書 頻度	延べ 頻度
1	て	助詞-接続助詞	471	15826
2	の	助詞-連体化	471	14506
3	に	助詞-格助詞-一般	471	10966
4	た	助動詞	469	18990
5	は	助詞-係助詞	466	14639
6	が	助詞-格助詞-一般	466	10480
7	だ	助動詞	454	6680
8	を	助詞-格助詞-一般	445	11206
9	も	助詞-係助詞	432	4068
10	する	動詞-自立	430	4545
11	で	助詞-格助詞-一般	423	3224
12	。	記号-句点	402	26550
13	、	記号-読点	391	28018
14	!	記号-一般	391	4784
15	「	記号-括弧開	388	11248
16	」	記号-括弧閉	387	11232
17	いる	動詞-非自立	387	3142
18	ない	助動詞	382	2369
19	の	名詞-非自立-一般	381	2349
20	か	助詞-副助詞／並立助詞／終助詞	376	1662
21	よ	助詞-終助詞	373	2406
22	から	助詞-格助詞-一般	363	1670
23	ます	助動詞	362	12595
24	と	助詞-接続助詞	360	2725
25	お	接頭詞-名詞接続	360	2014
26	う	助動詞	360	1765
27	なる	動詞-自立	353	1664
28	で	助詞-接続助詞	347	1433
29	言う	動詞-自立	343	2629
30	?	記号-一般	341	2105
31	と	助詞-格助詞-引用	339	2287
32	と	助詞-格助詞-一般	336	1477
33	に	助詞-副詞化	335	1556
34	です	助動詞	327	2986
35	ね	助詞-終助詞	327	1067
36	と	助詞-並立助詞	317	1524
37	くる	動詞-非自立	297	1039
38	こと	名詞-非自立-一般	295	1184
39	みんな	名詞-一般	286	1049
40	ん	名詞-非自立-一般	280	1173
41	それ	名詞-代名詞-一般	275	861
42	いい	形容詞-自立	272	730
43	もう	副詞-一般	270	685

44	その	連体詞	268	982
45	行く	動詞-自立	267	933
46	見る	動詞-自立	267	816
47	へ	助詞-格助詞-一般	259	1175
48	ん	助動詞	249	1001
49	たち	名詞-接尾-一般	244	1414
50	よう	名詞-非自立-助動詞語幹	244	670
51	僕	名詞-代名詞-一般	242	1342
52	ある	動詞-自立	236	627
53	この	連体詞	235	707
54	から	助詞-接続助詞	229	661
55	しまう	動詞-非自立	227	827
56	でも	接続詞	227	638
57	さん	名詞-接尾-人名	226	1394
58	大きな	連体詞	225	588
59	何	名詞-代名詞-一般	225	576
60	や	助詞-並立助詞	224	803
61	来る	動詞-自立	222	586
62	れる	動詞-接尾	221	704
63	なる	動詞-非自立	221	509
64	まで	助詞-副助詞	219	544
65	ある	連体詞	217	624
66	また	接続詞	217	439
67	そう	副詞-助詞類接続	216	601
68	時	名詞-非自立-副詞可能	214	524
69	な	助詞-終助詞	209	416
70	そう	名詞-接尾-助動詞語幹	208	447
71	くれる	動詞-非自立	206	500
72	食べる	動詞-自立	203	795
73	どこ	名詞-代名詞-一般	203	455
74	目	名詞-一般	195	488
75	ながら	助詞-接続助詞	194	524
76	これ	名詞-代名詞-一般	194	494
77	いく	動詞-非自立	194	485
78	出る	動詞-自立	192	514
79	たい	助動詞	192	467
80	思う	動詞-自立	186	443
81	声	名詞-一般	183	554
82	家	名詞-一般	182	834
83	・・	記号-一般	179	1221
84	もの	名詞-非自立-一般	179	454
85	帰る	動詞-自立	178	392
86	中	名詞-非自立-副詞可能	178	380
87	の	助詞-終助詞	177	451
88	と	助詞-副詞化	177	412

89	やる	動詞-自立	177	396
90	ない	形容詞-自立	177	365
91	いる	動詞-自立	176	445
92	てる	動詞-非自立	176	362
93	そして	接続詞	175	524
94	一緒	名詞-一般	172	328
95	日	名詞-非自立-副詞可能	170	473
96	入る	動詞-自立	169	401
97	が	助詞-接続助詞	168	451
98	上	名詞-一般	165	345
99	良い	形容詞-自立	165	304
100	ちゃん	名詞-接尾-人名	160	1332
101	聞く	動詞-自立	160	425
102	ば	助詞-接続助詞	160	341
103	だけ	助詞-副助詞	158	309
104	今日	名詞-副詞可能	158	285
105	ぞ	助詞-終助詞	157	416
106	木	名詞-一般	156	505
107	じゃ	助詞-副助詞	156	340
108	来る	動詞-非自立	155	408
109	られる	動詞-接尾	154	335
110	大きい	形容詞-自立	154	330
111	ちゃう	動詞-非自立	151	297
112	私	名詞-代名詞-一般	147	585
113	小さな	連体詞	147	403
114	次	名詞-一般	147	344
115	下	名詞-一般	146	311
116	どう	副詞-助詞類接続	145	259
117	みる	動詞-非自立	144	258
118	さあ	感動詞	143	294
119	夜	名詞-副詞可能	142	286
120	お母さん	名詞-一般	141	702
121	人	名詞-一般	141	573
122	そこ	名詞-代名詞-一般	141	278
123	待つ	動詞-自立	141	246
124	…	記号-一般	140	1658
125	顔	名詞-一般	140	303
126	朝	名詞-副詞可能	140	275
127	できる	動詞-自立	139	297
128	作る	動詞-自立	138	392
129	手	名詞-一般	138	326
130	持つ	動詞-自立	137	335
131	だれ	名詞-代名詞-一般	137	287
132	ここ	名詞-代名詞-一般	137	274
133	な	助動詞	136	295
134	ので	助詞-接続助詞	135	305
135	今度	名詞-副詞可能	134	234

136	前	名詞-副詞可能	133	280
137	あつ	感動詞	133	226
138	居る	動詞-自立	132	261
139	おなか	名詞-一般	131	310
140	出す	動詞-自立	131	232
141	飛ぶ	動詞-自立	130	288
142	ところ	名詞-非自立-副詞可能	130	227
143	おいしい	形容詞-自立	128	280
144	せる	動詞-接尾	128	250
145	なあ	助詞-終助詞	128	250
146	行く	動詞-非自立	128	242
147	まだ	副詞-助詞類接続	126	209
148	取る	動詞-自立	125	277
149	見える	動詞-自立	125	235
150	気	名詞-一般	125	204
151	分かる	動詞-自立	124	252
152	ちょっと	副詞-助詞類接続	124	233
153	一杯	副詞-一般	124	201
154	けど	助詞-接続助詞	122	224
155	みんな	名詞-代名詞-一般	122	164
156	の	助詞-格助詞-一般	121	210
157	自分	名詞-一般	120	247
158	さ	名詞-接尾-特殊	119	251
159	いつも	副詞-一般	119	242
160	有難う	感動詞	119	224
161	花	名詞-一般	118	415
162	たり	助詞-並立助詞	118	304
163	成る	動詞-自立	118	205
164	綺麗	名詞-形容動詞語幹	116	256
165	少し	副詞-助詞類接続	116	206
166	入れる	動詞-自立	115	238
167	とても	副詞-助詞類接続	115	238
168	でも	助詞-副助詞	115	201
169	君	名詞-接尾-人名	114	907
170	元気	名詞-形容動詞語幹	114	174
171	楽しい	形容詞-自立	114	173
172	水	名詞-一般	113	293
173	体	名詞-一般	113	240
174	びっくり	名詞-サ変接続	113	205
175	大	接頭詞-名詞接続	111	264
176	あの	連体詞	111	247
177	空	名詞-一般	111	232
178	頭	名詞-一般	111	231
179	中	名詞-一般	111	196
180	誰	名詞-代名詞-一般	110	220

181	一番	名詞-副詞可能	110	198
182	中	名詞-固有名詞-地域-一般	109	241
183	遊ぶ	動詞-自立	109	203
184	ところが	接続詞	108	182
185	走る	動詞-自立	107	250
186	開ける	動詞-自立	107	211
187	子	名詞-一般	106	421
188	もっと	副詞-一般	106	189
189	のに	助詞-接続助詞	106	178
190	落ちる	動詞-自立	104	174
191	歩く	動詞-自立	103	215
192	すぐ	副詞-助詞類接続	103	192
193	見つける	動詞-自立	103	190
194	より	助詞-格助詞-一般	103	179
195	立つ	動詞-自立	103	174
196	付く	動詞-自立	103	149
197	なさる	動詞-非自立	102	201
198	いう	動詞-自立	102	192
199	付ける	動詞-自立	102	167
200	あ	感動詞	101	169
201	日	名詞-一般	100	174
202	山	名詞-一般	99	342
203	始める	動詞-非自立	99	174
204	沢山	副詞-一般	99	172
205	みたい	名詞-非自立-形容動詞語幹	98	151
206	二人	名詞-一般	97	307
207	足	名詞-一般	97	200
208	知る	動詞-自立	97	198
209	ああ	感動詞	97	179
210	ぬ	助動詞	97	165
211	一つ	名詞-一般	96	202
212	一人	名詞-一般	96	173
213	こんな	連体詞	96	139
214	小さい	形容詞-自立	95	182
215	今	名詞-副詞可能	95	163
216	好き	名詞-形容動詞語幹	94	172
217	どんどん	副詞-助詞類接続	93	148
218	居る	動詞-非自立	93	142
219	外	名詞-一般	92	141
220	子供	名詞-一般	90	294
221	やる	動詞-非自立	90	204
222	聞こえる	動詞-自立	90	153
223	出来る	動詞-自立	90	145
224	1	名詞-数	89	255
225	風	名詞-一般	89	153

226	やって来る	動詞-自立	89	145
227	やっぱり	副詞-一般	88	127
228	あれ	名詞-代名詞-一般	88	121
229	匹	名詞-接尾-助数詞	87	313
230	助ける	動詞-自立	87	187
231	とき	名詞-非自立-副詞可能	87	145
232	探す	動詞-自立	86	192
233	考える	動詞-自立	86	158
234	ほら	感動詞	86	119
235	で	助動詞	86	109
236	森	名詞-一般	85	264
237	お前	名詞-代名詞-一般	85	261
238	音	名詞-一般	85	197
239	乗る	動詞-自立	85	186
240	叫ぶ	動詞-自立	85	169
241	ところ	名詞-一般	85	146
242	すっかり	副詞-一般	85	110
243	君	名詞-代名詞-一般	84	233
244	みる	動詞-自立	84	124
245	何だか	副詞-一般	84	109
246	きっと	副詞-一般	83	121
247	雨	名詞-一般	82	225
248	本	名詞-一般	82	164
249	まま	名詞-非自立-副詞可能	82	135
250	急ぐ	動詞-自立	82	132
251	どうどう	副詞-一般	82	125
252	3	名詞-数	81	222
253	なんて	助詞-副助詞	81	150
254	後	名詞-一般	81	107
255	だから	接続詞	80	146
256	長い	形容詞-自立	80	145
257	大変	名詞-形容動詞語幹	80	117
258	友達	名詞-一般	79	183
259	口	名詞-一般	79	119
260	うれしい	形容詞-自立	79	116
261	ずっと	副詞-一般	78	147
262	優しい	形容詞-自立	78	145
263	笑う	動詞-自立	78	131
264	大好き	名詞-形容動詞語幹	78	127
265	ん	助詞-終助詞	78	119
266	上げる	動詞-自立	78	114
267	生まれる	動詞-自立	77	173
268	泣く	動詞-自立	77	157
269	戻る	動詞-自立	77	144

270	歌う	動詞-自立	77	142
271	眠る	動詞-自立	77	136
272	うち	名詞-非自立-副詞可能	77	128
273	明日	名詞-副詞可能	77	111
274	そこで	接続詞	76	192
275	そんな	名詞-形容動詞語幹	76	99
276	こんなに	副詞-助詞類接続	76	92
277	住む	動詞-自立	75	176
278	話	名詞-一般	75	168
279	降る	動詞-自立	75	139
280	嬉しい	形容詞-自立	75	107
281	うさぎ	名詞-一般	74	392
282	2	名詞-数	74	191
283	こっち	名詞-代名詞-一般	74	100
284	ねえ	感動詞	73	140
285	毎日	名詞-副詞可能	73	115
286	間	名詞-一般	73	111
287	だ	助詞-格助詞-一般	73	109
288	気持ち	名詞-一般	73	104
289	忘れる	動詞-自立	73	101
290	川	名詞-一般	72	196
291	やってくる	動詞-自立	72	126
292	飛び出す	動詞-自立	72	101
293	もらう	動詞-非自立	72	100
294	部屋	名詞-一般	71	159
295	書く	動詞-自立	71	140
296	ため	名詞-非自立-副詞可能	71	133
297	方	名詞-非自立-一般	71	127
298	お父さん	名詞-一般	70	316
299	早く	副詞-一般	70	93
300	初めて	副詞-一般	70	91
301	本当	名詞-形容動詞語幹	70	89
302	かわいい	形容詞-自立	69	125
303	高い	形容詞-自立	69	119
304	いつ	名詞-副詞可能	69	119
305	本当	名詞-一般	69	118
306	なんて	副詞-一般	69	114
307	困る	動詞-自立	69	100
308	魚	名詞-一般	68	200
309	町	名詞-一般	68	194
310	寝る	動詞-自立	68	125
311	力	名詞-一般	68	125

312	全部	名詞-副詞可能	68	105
313	向こう	名詞-一般	68	101
314	ゆっくり	副詞-助詞類接続	68	99
315	く	動詞-自立	67	107
316	会う	動詞-自立	67	107
317	そんな	連体詞	67	99
318)	記号-括弧閉	66	313
319	(記号-括弧開	66	307
320	様	名詞-接尾-人名	66	239
321	海	名詞-一般	66	203
322	くださる	動詞-非自立	66	141
323	連れる	動詞-自立	66	138
324	飲む	動詞-自立	66	127
325	という	助詞-格助詞-連語	66	123
326	向かう	動詞-自立	66	101
327	沢山	名詞-形容動詞語幹	66	101
328	最後	名詞-一般	66	88
329	振る	動詞-自立	66	80
330	逃げる	動詞-自立	65	168
331	涙	名詞-一般	65	108
332	静か	名詞-形容動詞語幹	65	102
333	どんな	名詞-形容動詞語幹	65	94
334	出掛ける	動詞-自立	65	92
335	心配	名詞-サ変接続	65	89
336	あなた	名詞-代名詞-一般	64	224
337	色	名詞-一般	64	122
338	くれる	動詞-自立	64	120
339	不思議	名詞-形容動詞語幹	64	92
340	起きる	動詞-自立	64	92
341	開く	動詞-自立	64	85
342	穴	名詞-一般	63	172
343	悲しい	形容詞-自立	63	113
344	ばかり	助詞-副助詞	63	111
345	歌	名詞-一般	63	110
346	吹く	動詞-自立	63	86
347	猫	名詞-一般	62	423
348	赤い	形容詞-自立	62	197
349	乗せる	動詞-自立	62	100
350	あげる	動詞-自立	62	89
351	やっと	副詞-助詞類接続	62	85
352	—	記号-一般	61	220
353	はい	感動詞	61	134
354	ご飯	名詞-一般	61	114
355	背中	名詞-一般	61	103
356	まあ	感動詞	61	98

357	匂い	名詞-一般	61	98
358	こう	副詞-助詞類接続	61	95
359	くる	動詞-自立	60	109
360	良く	副詞-一般	60	99
361	すてき	名詞-形容動詞語幹	60	75
362	熊	名詞-一般	59	270
363	道	名詞-一般	59	109
364	怖い	形容詞-自立	59	93
365	教える	動詞-自立	59	93
366	運ぶ	動詞-自立	59	90
367	間	名詞-非自立-副詞可能	59	86
368	動く	動詞-自立	59	83
369	昔	名詞-副詞可能	58	188
370	名前	名詞-一般	58	128
371	うん	感動詞	58	104
372	事	名詞-非自立-一般	58	99
373	呼ぶ	動詞-自立	58	93
374	突然	副詞-助詞類接続	58	74
375	しっかり	副詞-助詞類接続	58	73
376	これから	副詞-助詞類接続	58	72
377	じゃあ	接続詞	57	92
378	早速	副詞-助詞類接続	57	82
379	赤ちゃん	名詞-一般	56	167
380	一	名詞-数	56	161
381	白い	形容詞-自立	56	139
382	窓	名詞-一般	56	98
383	し	助詞-接続助詞	56	98
384	隠れる	動詞-自立	56	96
385	ご覧	名詞-一般	56	95
386	在る	動詞-自立	56	93
387	月	動詞-自立	56	81
388	帰り	動詞-自立	56	81
389	だめ	名詞-形容動詞語幹	56	78
390	俺	名詞-代名詞-一般	55	189
391	強い	形容詞-自立	55	122
392	恐い	形容詞-自立	55	93
393	広い	形容詞-自立	55	90
394	所	名詞-非自立-副詞可能	55	85
395	つく	動詞-自立	55	85
396	って	助詞-格助詞-連語	55	85
397	暫く	副詞-一般	55	76
398	——	記号-一般	54	223
399	心	名詞-一般	54	121
400	悪い	形容詞-自立	54	116