

日本人の語彙量(理解語彙、使用語彙)調査を行うにあたっての基礎的研究

荻原 廣

1. はじめに
2. 理解語彙、使用語彙とは
3. 先行研究
 - 3-1 東京の成城小学校の児童の語彙量調査
 - 3-2 国立国語研究所の「24時間調査」
 - 3-2-1 白河市および附近の農村における実態調査
 - 3-2-2 鶴岡における実態調査
 - 3-3 阪本一郎の語彙量調査
 - 3-4 森岡健二の語彙量調査
 - 3-5 語彙数推定テスト
 - 3-5-1 語彙数推定テストの概要
 - 3-5-2 語彙数推定テストを使った語彙量調査
 - 3-5-2-1 大妻女子大学の語彙量調査
 - 3-5-2-2 佛教大学の語彙量調査
 - 3-6 オンライン・コミュニケーションを使った語彙量調査
4. 今後、日本人の語彙量(理解語彙・使用語彙)調査を行うにあたって
5. おわりに

個人の語彙量(理解語彙、使用語彙)がどのくらいあるのかについての調査は、現在に至るまで決して多く行われてきたとは言えず、中でも使用語彙についての調査は、幼児対象の調査を除くと、ほとんど行われていない。

そこで本稿では、過去の理解語彙、使用語彙の調査がどのように行われ、そこにどういった問題点があるのかを明らかにし、そのうえで、個人の理解語彙、使用語彙を調査するにはどうすればいいかについて述べ、最後に、現在、筆者が行っている理解語彙、使用語彙の調査について触れる。

1. はじめに

語彙調査は、今までいろいろと行われてきた。そして、それは日本語コーパスの整備に伴い、今後もさまざまな形で進められていくであろう。しかし、いくらコーパスが整備されても、個人の語彙量（理解語彙、使用語彙）がどのくらいあるのかについての解明にはつながらない。

個人の語彙量（理解語彙、使用語彙）に関する調査は、現在に至るまで決して多く行われてきたとは言えない。中でも使用語彙についての調査は、幼児対象の調査を除くと、ほとんど行われていないのが現状だ。⁽¹⁾

そこで本稿では、過去の理解語彙、使用語彙の調査がどのように行われ、そこにどういった問題点があるのかを明らかにし、そのうえで、個人の理解語彙、使用語彙を調査するにはどうすればいいかについて述べる。また最後に、現在、筆者が大学生を被調査者として行っている理解語彙、使用語彙の調査がどういったものかについても述べる。

2. 理解語彙、使用語彙とは

理解語彙、使用語彙とは何だろうか。過去の定義を見てみると、玉村(1984)は、「理解語彙というのは、ある個人（またはある学年の児童など）が聞いたり読んだりするときに理解することができる見出し語の集合であり、使用語彙（表現語彙）というのは、ある個人（ある学年の児童など）が話したり書いたりするときに用いることができる語の集合である。」⁽²⁾と述べている。

また、日本語教育辞典（執筆者・山崎誠）は、「理解語彙は、個人が見たり読んだりする際に理解しうる語の総体のことをいう。使用語彙には①個人が話したり、書いたりするときに使うことのできる語彙、②個人が実際に使用した語彙（「漱石の語彙」など）の2つの意味がある」⁽³⁾としている。ただし、②は夏目漱石の語彙、源氏物語の語彙など、文学作品の研究の際に用いられるものであり、一般の個人を対象とするなら①である。

そこで、ここでは今までの定義に倣いながらも、やや平易に「理解語彙というのは個人が聞いたり読んだりするときに理解できる語の集まりである。使用語彙というのは個人が話したり書いたりするときに使うことのできる語の集まりである。」と定義する。

一方、理解語彙、使用語彙の調査の少なさについては、今までいろいろと行われてきた。例えば、玉村(1984)は、「理解語彙と言ひ、使用語彙と言つても、

実際にはどうしてそれらの量を測定するのか。〈理解することができるかどうか〉をどういう方法で検査するのか、また〈使うことができるかどうか〉をどうして調べるのか、実は大きな問題なのである。ただ、〈使った〉語を調べるといふなら、確実に測定する方法はある。⁽⁴⁾とした上で、「個々人の使用語彙の調査は、まだ極端に少ない。潜在的である理解語彙と比べると、使用語彙における個人差は、質的にも量的にもたいへん大きいことはたしかである。」⁽⁵⁾と述べている。

また、林四郎(1971)は、「日本人の成人の理解語彙量は大体4万語程度であろう」⁽⁶⁾と推察したあと、「使用語彙は調査法がむずかしく、まだ調査例に接していないので、日本人の平均使用語彙量がどのくらいのものか、いっこうにわからないが、理解語彙量よりずっと少ないことは確かだろう。」⁽⁷⁾と述べている。

3. 先行研究

今までに行われた個人の理解語彙、使用語彙の調査のうち、幼児を対象とした調査を除いた主なものについて、その概要を見ていく。

3-1 東京の成城小学校の児童の語彙量調査

澤柳・田中・永田(1919)『児童語彙の研究』に東京の成城小学校の児童に対して行われた理解語彙の調査の内容が載っている。

この調査は、質問を用いて児童の理解語彙を調査したものである。なお、質問を用いた調査にしたのは、比較的多数の児童についての調査だったため、外国で多く行われてきた一個人について長い期間観察するという方法では調査できなかったためである。

1) 被調査者

1918年、東京の成城小学校に入学した30名のうち、種々の障害のあるものを省いた⁽⁸⁾25名。平均年齢は6歳5か月。

2) 調査時期

1918年4月14日から6月20日まで(日曜祭日等の休業日を除く)

3) 調査方法

調査者(藤本房次郎、田中末廣の2名)が、子供に主として直接質問をして知っているかどうか調べた。1人の児童を相手に、語彙表(金沢庄三郎の『辞林』を基本とし、その他の書籍も参考としながら6867語を選定した)について1語1語理解しているかどうかの質問をし、理解している語

に印をつけていった。

4) 調査結果

最も多くの語を知っていた児童は5162語、最も知っている語が少なかった児童は3500語、平均は4089語であった。

5) 問題点

調査が単調であったため、児童はすぐ飽き、飽きてくると知らない語でも知っていると言えたり、その逆もあつたりしたらしく、調査の大変さがかうかがえる。ただ、対面方式で調査者が知っているかどうか1語1語判断するため、正確な結果が得られたものと考えられる。しかし、この方法は、児童相手だから可能だったとも言え、これが児童より語彙量の多い中学生以上を対象とすると、無理ではないかと思われる。

なお、千葉の鳴浜小学校と岡山の師範附属小学校も同様の調査を行っている。⁽⁹⁾

3-2 国立国語研究所の「24時間調査」

国立国語研究所は「24時間調査」について、「特定の個人について一日じゅうつききりでその言語と言語生活を全部記録する調査をわれわれはかりに『24時間調査』(詳しくは『言語生活の24時間調査』)と呼ぶことにする。」と述べており、いわゆる使用語彙の調査ではない。

では、過去に行われた2つの「24時間調査」について見ていく。⁽¹¹⁾

3-2-1 白河市および附近の農村における実態調査

1) 被調査者

51歳の男性、農民。49歳の女性、商家の主婦。45歳の男性、美容院主。以上3名。

2) 調査時期

男性2人は1949年11月8日、女性は1949年12月7日。

3) 調査方法

最初は、1人の被調査者に2人の調査者がついて手書きで記録した。しかし、調査員の負担があまりにも大きいため、のちに、1人の被調査者に述べ6人の調査者をつけることとし、2人1組が午前、午後、夜と3交替で1人の被調査者につき手書きで記録した。

なお、調査の時間は実際は「24時間」ではなく、朝起きてから夜床へ入るまでの十数時間であった。ただ、この時間は、外出の時もついて行くな

ど被験者に徹底的についてまわり、全て手書きで記録した。(相手の言葉は記録しなかった。) また、3人の被調査者が話した言葉は記録したが、書いたものは調査結果に語数としては反映されていない。(ただし、書くために何分費やしたかは報告されている。)

4) 調査結果

	話題の数	延べの文の数	文節の延べ語数	異なり語 (自立語) 数
農民	706	2638	10068	2324
商家の主婦	731	3206	9290	2138
美容院主	661	2894	8558	—

(美容院主の異なり語(自立語)数はデータが示されていない)

5) 問題点

この調査は、あくまで個人の1日の言語生活を調べたものであって、個人の使用語彙を調べるのが目的の調査ではない。しかし、調査報告の中に「この単語数が個人の総使用語数の何%を占めるかは明らかでないが、農民にしても商家の主婦にしても一日に2000語以上使っていることは興味深い⁽¹²⁾」と述べてあるように、成人の1日の会話における使用語彙がどの程度かを示したのは貴重である。

これを長期間続け、更に書いたものについても集めれば、個人の使用語彙がどの程度かわかる可能性がある。しかし、1日でも大変なこの調査を長期間継続して行うのは無理であろう。

3-2-2 鶴岡における実態調査

1) 被調査者

高級地方公務員、手工業者、商店主で、全員男性。合計3名。

2) 調査時期

高級地方公務員は1950年11月20日、手工業者は1951年10月1日、商店主は1950年11月21日。

3) 調査方法

最初から1人の被調査者に6人の調査者がつくこととした。そして、この6人が、午前、午後、夜と3交替で2人ずつ1人の被調査者につき記録した。記録は手書き以外にも、機械速記、録音機も併用した。

4) 調査結果

	話題の数	延べの文の数	文節の延べ語数	異なり語 (自立語) 数
高級地方公務員	437	1983	5528	1497
手工業者	318	1573	4752	1282
商店主	391	1347	2891	919

5) 問題点

この調査は、白河市および附近の農村における調査とほぼ同じ方法で行っているため、こちらも個人の1日の言語生活を調べたものであって、いわゆる個人の使用語彙を調べるための調査ではない。ただ、こちらも、成人の1日の使用語彙（ただし、書くのに使った語は含まれない）がどの程度かを示したのは、貴重である。

なお、白河市および附近の農村における調査と調査結果に大きな隔たりがあるのは、職業の違いの可能性が考えられるが、しかし同じ被調査者でも、日によって話す量が一定とは限らないので、正確なことはわからない。

3-3 阪本一郎の語彙量調査

阪本(1955)は、内省法を使い、次のような理解語彙の調査を試みた。これは、児童への面接調査や国立国語研究所の「24時間調査」と違って、調査者は被調査者に対して直接何かをするわけではなく、被調査者自身が内省しながら自ら調査用紙に記入する方法である。

1) 被調査者

小学校、中等学校については以下の表の通りである。

	男	女	計
小学校	9544	11045	20589
中等学校	11282	7495	18777
計	20826	18540	39366

2) 調査時期

1937年6月から9月

3) 調査方法

全国の小学校、中等学校、女学校、師範学校等に調査を依頼した。調査語彙は、最初、金沢庄三郎の『広辞林』より機械的に1000語選んだ。しかし、予備調査で1000語に対する調査時間が1時間半程度必要だったため、結局、総語数を半分の500にした。(奇数番目と偶数番目の2種類作った)更に予備調査を経て正解者数の多い語(易しい語)から順に並べ替え、100語ずつ5段階に区切り、幼少者には始めの方の段階のみを与え、年長者には全段階を与えることとした。

また、阪本は、附印法では間違いが生じるからと定義法を併用した。しかし、500語全てに定義を記入させると時間がかかるので、まず附印法でさせ、全部終わってから各段階で指定した5ないし10語ずつのみに定義を記入させた。そして、その定義の正答率を附印した語の総数に掛け、修正を行った。こうして500語中、正しく知っているのは何語か測定し、理解語彙量を推定した。

なお、阪本は、日本語の語彙の総量は、普通の程度において約10万であるから、このテストの1語は200語を代表しているとの考えから、語彙量を推定した。すなわち、50語知っていれば、理解語彙は1万語、250語知っていれば理解語彙は5万語となる。しかし、『広辞林』は76315語だったため、のちに500語から24%分を引き、その不足分を『辞苑』から補った。

4) 調査結果

標準語彙量として、以下の表を⁽¹³⁾発表した。

性別 満年月		男			女		
		測定人員	語彙量	S.D.	測定人員	語彙量	S.D.
6	0—5	690	5606	1654	645	5158	2200
	6—11	630	6142	1877	525	5739	2447
7	0—5	675	6655	2100	587	6250	2694
	6—11	642	7186	2282	580	6708	2931
8	0—5	570	7822	2464	680	7283	3167
	6—11	595	8689	2892	605	8088	3629

9	0—5	545	9812	3319	718	9256	4091
	6—11	512	11248	3968	713	10787	5179
10	0—5	579	12863	4617	690	12635	6267
	6—11	509	14940	5824	758	15075	7517
11	0—5	502	17359	7030	820	18057	8769
	6—11	541	20212	8290	975	21674	9159
12	0—5	1021	23085	9550	1300	25254	9550
	6—11	1080	25928	10067	1129	28406	9567
13	0—5	1035	28809	10267	1067	31035	9583
	6—11	1139	31636	10433	892	33478	9304
14	0—5	1069	34379	10583	957	35726	9024
	6—11	629	37026	10733	950	37783	8962
15	0—5	589	39475	10867	855	39722	8900
	6—11	588	41458	11000	682	41191	8887
16	0—5	772	43886	11133	742	42447	8874
	6—11	750	45962	11233	469	43382	8795
17	0—5	892	47721	11317	310	44161	8716
	6—11	684	49107	11400	188	44770	8693
18	0—5	606	50069	11500	102	45190	8669
	6—11	598	50656	11600	114	45400	8673
19	0—5	443	50985	11683	123	45467	8676
	6—11	371	51128	11750	88	45489	8643
20	0—5	490	51176	11800	57	45496	8609

なお、林(1971)は、上記の表とは別の表（1年ごとに男女を合算したもの）⁽¹⁴⁾を元に、今日の教育段階に当てて、以下のように大まかにまとめた。

小学校入学時 6歳 6000語

小学校卒業時	11歳	20000語
中学校卒業時	14歳	36000語
高等学校卒業時	17歳	46000語
成人期	20歳	48000語

5) 問題点

全数調査ではなく標本調査なので、当然、誤差が生じる。そこは阪本も気にしていたようで、この調査の信頼度を成城小学校、鳴浜小学校、師範附属小学校の調査を引き合いに出し、以下のように述べている⁽¹⁵⁾。

いずれもきわめて周到な用意をもって調査せられたものであるから、これらの数字は信頼するに足るものである。これと比較するとわたくしの調査では満六歳六ヶ月の男子が六一四二、女子が五七三九であって、やや多くなっている。しかし大体においてその数が近似していることは、わたくしの概算的方法の信頼性を多少なりとも裏書するものである。

ただ、標本調査とはいえ、この調査は被調査者が4万人と非常に多いうえ、6歳から20歳と幅広い年代の理解語彙の量を明らかにしており、このような大規模の個人の語彙量調査は他に類を見ない。

3-4 森岡健二の語彙量調査

森岡(1951)は、義務教育終了者がどれくらい理解語彙を持っているかを知るため、内省法を使い、次のような調査を試みた。

1) 被調査者

東京精華学園高等部に新たに入学した15名(実施前に準備調査を行った時は17名だったが、希望により1名を加えて18名となった。しかし、事故1名、病気2名を生じ、最終的には15名となったとある。)

2) 調査時期

1950年7月20日から11月末日まで

3) 調査方法

竹原スタンダード和英辞典所載の見出し語をすべて印刷(384頁)して検査語彙とし、附印法と関連語の記入とを併用する形で調査した。記入法は、以下の(1)(2)の通りである。⁽¹⁶⁾

(1) 知っているか、いないかについて次の符号をつける。

- よく知っていていつも使っていると思う語
- ν 聞け・読め ば意味がわかると思う語
- △ 聞いた・読んだ ことはあるが意味のはっきりしない語
- × ぜんぜん分らない語

(2) 問題の語に、関係のある語を書き入れる。すなわち意味・言い換え・反対語・同類語その他連想などによって思いついた語を書き入れる。

- 例1) ○カワ (川・河) 山 水 小川
- 例2) νケンビ (兼備) 名 才色
- 例3) △カワウソ (川獺)
- 例4) ×ケンカ (鹼化)

なお、附印法だけでなく関連語を記入することも要求したのは、以下の4つの理由による。第1に理解が正しいか否かを検するため。第2に誤って印をつけることを防ぐため。第3に検査の単調を破り興味をつなぐため。第4に竹原の辞書にのっていない語を導く手掛りを得るため。

4) 調査結果

符号のうち、○とνとを理解語とし、△と×とを理解されぬ語として取り扱うこととした。その結果、理解語(○ν)の最高は36330語、最低は23381語、平均は30664語であった。なお、15名から回答を得た検査語彙の正確な数字は37970語である⁽¹⁷⁾。しかし、この数は、ミスプリントの校正漏れの語があつて削除した後の数のため、厳密にはスタンダード辞典の見出し語の数とは一致しない。

5) 問題点

森岡は、「竹原スタンダード和英辞典が検査語彙として適当か、それで現代日本語の生きた語彙を一応網羅しているかどうかということは、和英辞典の性質として一応現代日本語の生きた語を採録してあるものと見て用いただけであつて、この吟味は今後に残された問題である」と述べている⁽¹⁸⁾。

また、検査語彙の数が37970語という点も問題が残る。これは、義務教育終了者が対象なので、この程度でもよかつたとも推測されるが、一方で、理解度が最高(36330語)であつた被調査者は、検査語彙の95.7%も理解

していたこととなり、これがもっと検査語彙が多ければ、更に理解語彙が増えていた可能性も否定できない。とはいえ、この調査は竹原スタンダード和英辞典所載の見出し語を全て検査語彙とする全数調査であり、検査語彙の多さは、他に例がなく、被調査者数は少ないながらも、精度の高い調査と言えよう。

3-5 語彙数推定テスト⁽¹⁹⁾

テストの解説によると、語彙数推定テストは、NTT データベースシリーズ「日本語の語彙特性」第1巻の単語親密度⁽²⁰⁾データベースを用いた簡単でかつ精度の高い推定テストであるとのことである。

そこで、ここでは、まず、このテストの概要について述べ、次にこのテストを用いた理解語彙の調査について述べる。

3-5-1 語彙数推定テストの概要

1) テストの作成

このテストで使用されている単語は、機械的に選ばれたものではなく、上記の単語親密度データベースから、単語親密度を基準に選択されたものである。そして、その選ばれた単語を親密度順にならべ、目標とするテスト項目数になるように、ある一定間隔で単語を取り出し作成されている。

なお、テストは3種類あるが、それぞれ50語からなっており、ネット上で、知っている語にチェックを入れていき、全てチェックした後で、推定開始ボタンで結果を送信すると、「あなたの語彙数は××語です」という形で推定語彙量が計算される仕組みになっている。

2) 単語親密度境界の決定

高親密度単語は知っている場合がほとんどであり、低親密度単語は知らない場合がほとんどとなるはずなので、単語親密度の高い順に単語を並べたとき、だんだん単語親密度が下がるにつれて知らない単語が出現する可能性が高くなる。しかし、知っている単語と知らない単語の親密度境界を定めることは非常に困難を極める。

ここでは、単語親密度順に判断をしたときに知らない単語が2つ以上連続する単語の親密度と、知っている単語が2つ以上連続する単語の親密度との中間点を単語親密度の境界としている。

3) 語彙数の推定

2) で求めた単語親密度境界以上の単語の数をデータベース中で数えあ

げて推定語彙数としている。

4) 問題点

単語親密度境界の決定方法はいろいろあり、どの方法が実際の語彙数によく合うかはまだ判明していないとのことである。また、過去に行われた辞書を使った調査も同様であるが、この調査も推定される語彙数は常に基となる辞書またはデータベースに依存するため、推定される語彙の最大数は、辞書またはデータベースの語数以上にはならない。よって、このテストで推定された語彙量は実際に知っている単語の数よりも少ない可能性がある。

3-5-2 語彙数推定テストを使った語彙量調査

3-5-2-1 大妻女子大学の語彙量調査

柴田(2012)は、大妻女子大学でこの語彙数推定テストを使った調査をしている。

1) 被調査者

平成21年度、22年度の短大、4年制文系、家政系の一部の1、2年生の学生。

2) 調査時期

2010年6月、2011年5月

3) 調査方法

「日本語(口頭表現)」の受講生に協力を求め、無記名(専攻学科のみ記入)でテストに参加してもらった。なお、授業中に行ったため、プリントアウトした3種類のテストを配布し、知っている単語にチェックを点けて提出してもらおうという方式により調査した。

4) 調査結果

3つのテスト、それぞれの推定結果の平均を各学生の語彙量とした。その結果、推定語彙量は、4年制学部生の平均が3万6千余語、短大国文の学生が3万3千余語、短大他専攻生は、2万7千余語あり、語彙量では、4年制学部の文学系の1、2年生の平均語彙量が最も高かった。全員の平均語彙数はおよそ34900語であった。

5) 問題点

語彙力推定テストは、附印法を使ったテストであるが、ここでは、そのまま附印法のみでの調査をしている。しかし、附印法は、過去の語彙量調査でも誤差が生じるからと、単独では使われていない。先に述べたように、

坂本は附印法と定義法を併用。また、森岡は附印法と伴に関連語を記入させる方法をとっている。

また、どうして3種類のテストを合算したのだろうか。本来、この推定テストが正確なら、3種類のテストを全てしてもどれも同じ結果になるはずである。しかし、それが違う結果になるのなら、それぞれの結果を示す必要があるだろう。

ただ、大学生の読書量と語彙量の関係は、あちこちで言われながらも、データとして示したものはなかった。そこを明らかにしようとしたのは、意義のある調査と言える。

3-5-2-2 佛教大学の語彙量調査

筆者は、佛教大学でこの語彙数推定テストを使った理解語彙の調査を行った。

1) 被調査者

「日本語学概論2」の初回の受講生99名（2回生から4回生まで）に調査を行った。しかし、回答者99名のうち、全てのテストをしていない者が3名おり、その3名を除く96名の調査結果を有効回答とした。

なお、被調査者の所属学科は日本文学科、中国学科、仏教学科、人文学科であった。

2) 調査時期

2013年9月23日

3) 調査方法

筆者が講義を担当している「日本語学概論2」の受講生に協力を求めた。また、これは自由な意見を聞くアンケートなどではなく、単語を知っているかどうかを聞くだけなので、それなら責任ある回答を得るためにと記名方式にした。

授業中に行ったため、印刷した用紙を配り、それに書かせる形にした。しかし、知っている単語にチェックをさせて提出してもらう附印法だけなら誤差が生じる。だからといって、定義法や関連語を書かせるといった方法は、調査者にも被調査者にも負担がかかる。よって、より負担の少ない方法として、各語の横に意味をつけたものを用意した。（資料1）

やり方は、まず単語だけを書いたテストをさせ、その後もう一度、今度は各語の横に意味をつけたテストをさせた。なお、多義語の場合は、1つでも知っていれば、知っているとみなした。

4) 調査結果

資料2の表を見ると、3種類のテストには、難易度に差があることがわかる。テスト1と3は平均で7000語以上も差がある。よって、易しいテストを受けると語彙量は多く、難しいテストを受けると語彙量は少なくなる。つまり、単語親密度を用いたこの語彙数推定テストは、かなりの誤差が生じるということがわかった。(資料2)

なお、単語だけを書いたテストと、各語の横に意味をつけたテストの結果を比べると、やはり誤差が生じている。1-3のテスト、全てにおいて単語だけを書いたテストと、各語の横に意味をつけたテストの結果が全く同じだった者は、96名中、わずかに3名しかいなかった。(語彙数推定の結果が全て同じだったものは6名いたが、資料2のNO91-93は、結果がたまたま同じになっただけで、実際チェックした語は、単語だけを書いたテストと、各語の横に意味をつけたテストの間で違いがあった。つまり、附印法を単独で使うと、このように間違いが生じるのである。

3-6 オンライン・コミュニケーションを使った語彙量調査

荒牧・増川・森田・保田(2012)は、Twitterに注目した。Twitterはユーザ数が1400万人(2011年1月現在)と多く、また、文字数制限のため大規模な引用がないことが理由である。

1) 被調査者(ユーザ数)

ユーザ数	約10万人(99964人)
ユーザ抽出条件	毎月5ツイート以上投稿している(継続的な発言) 総発言語数が5000以上 最初の100ツイートの中に「の」が含まれている(日本語使用者に限定)これは非日本語使用者を除くため行った。
全ツイート数	約2.5億ツイート(253,482,784ツイート)
全形態素数	約43億語(4,258,707,255語)

2) 調査時期

2009年11月3日から2010年3月25日までの143日間(約5カ月間)

3) 調査方法

一定期間にユーザが発言した語数から、潜在的な語彙数Nを推測。その推測はユーザがジップ則にしたがって語を発生していると仮定することで可能としている。

4) 調査結果

オンライン・コミュニケーション上での平均使用語彙数は8000語と推定した。そして、先行研究で推定されてきた理解語彙40000語と比較すると、多くは実際には使用されていないと結論づけている。

5) 問題点

使用語彙を明らかにしようとした試みは、画期的ではある。しかし、国立国語研究所の「24時間調査」もそうだったが、荒牧・増川・森田・保田(2012)のTwitterを使った調査も、使用語彙の対象は会話のみである。つまり、過去の発言のみが対象で、過去に書かれたものは調査対象とはなっていない。しかし、筆者が冒頭で定義したように、使用語彙とは個人が話したり、書いたりするときに使うことのできる語彙であるならば、この調査は、書く行為を調査対象としていないため、厳密に言えば、個人の使用語彙の全体像を明らかにしようとした調査とは言えない。これは、定義の段階で使用語彙を「使用された語彙」としかとらえていない、つまりは認識不足があったからではないだろうか。

また、この調査では、どの年齢層の語彙量がどのくらいかといったことが全くわからない。タイトルにあるように、日本人のオンライン・コミュニケーション上での平均使用語彙数は8000語であることはわかるが、その「日本人」の中には、中学生も高校生も大学生も社会人も混在している。なぜなら、Twitterは中学生から利用できるからである。過去の理解語彙の調査において、中学生、高校生、大学生の理解語彙の量は、明らかに違うことがわかっている。それなら、当然、中学生、高校生、大学生の使用語彙の量も違うと推定される。本来なら、そこも明らかにすべきところを全く考慮に入れていない。なお、調査結果のところでは、先行研究で推定されてきた理解語彙40000語と比較すると、多くは実際には使用されていないとあるが、その40000語は成人の理解語彙である。成人の理解語彙と中学生も高校生も大学生も社会人も混在した「日本人」の使用語彙を比べても、比較したことにはならない。

更に、個人の発話量の少なさも問題である。ユーザ抽出条件が、毎月5ツイート以上投稿していることとしているが、1ツイートは、最高で140字である。これが5ツイートなら、1か月、最低700字と原稿用紙2枚弱程度の発話にしかならないにもかかわらず、データに入れている。NHKのアナウンスルームによると、人前で話すときは、目安として、1分間に

400字詰め原稿用紙 1 枚弱がいいとのことであるが、そうすると2分程度⁽²¹⁾の発話量になる。1か月に2分（5か月なら10分）程度の発話量で、その人の使用語彙の量（ただし会話のみ）を推測するのは無理があるのではないか。

なお、この調査が短期間であることについては、自ら欠点として挙げているが、俳句の季語ではないが、語の中にはある一定の期間にのみよく使われるものがある。たとえば、クリスマスに関わる語は12月に、入学に関わる語は4月に、桜は春に、紅葉は秋になど、他の季節に現れにくい語がある。よって、5か月ではなく、最低でも1年は調べる必要があるだろう。

4. 今後、日本人の語彙量（理解語彙、使用語彙）調査を行うにあたって

以上、ここまで先行研究を見てきたが、個人の理解語彙、使用語彙を調査するには、さまざまな問題点が横たわる。また、個人の語彙量のうち、いわゆる使用語彙（期間の限定や対象が会話だけなどの制限がない使用語彙）については、調査例が無いに等しいことがわかった。そこで、ここでは、理解語彙、使用語彙の調査方法について、どのように行うべきかを考えたい。

過去に行われた調査方法はいろいろある。しかし、観察法は使用語彙の調査法としては最も正確だと考えられるが、調査コストを考えると現実的ではない。また、理解語彙の調査として調査者が直接質問する方法も、幼児期を除けば、調査に時間がかかりすぎて無理であろう。そうなると、やはり内省法ということになる。ただ、この内省法、今までは、理解語彙の調査としてのみ用いられてきた。しかし、使用語彙の調査にも使えないだろうか。

もし使用語彙の定義を「過去に個人が話したり、書いたりした語彙」とするなら、正確に調査するには、生まれてから調査時に至るまでの、毎日、24時間、話すのも書くのも記録し続けなければならないことになる。しかし、最初に定義したように「個人が話したり、書いたりするときに使うことのできる語彙」とするなら、被調査者が内省した結果「この語は私が話したり、書いたりするときに使うことのできる語だ」と判断すれば、それは使用語彙ということになり、つまりは内省法でも使用語彙の調査が可能になることになる。

実は、理解語彙と使用語彙について、平日頃から内省法で判断している職業がある。それは、日本語非母語話者に対して日本語教育を行っている日本語教員である。日本語教員は、授業の準備段階で、新出の単語1つ1つについて「書き言葉で使う」か、「話し言葉で使う」か、「どちらでも使う」か考え、そ

して更にその単語が「理解語彙」か「使用語彙」かについても考え、判断している。つまり、内省法によって、その語が理解語彙か使用語彙かを判断しているのである。これについては、実はいろいろと問題もある（つまりは、ある単語が理解語彙か使用語彙かという判断を全て個人任せにしていいのかということ）が、その点については、また別の機会に触れたい。ただ、このことは、内省法によっても、使用語彙の調査は可能だということを裏付けることになるであろう。

また、先に書いた森岡の調査だが、理解語彙の符号をつける際の基準として、以下のようにあった。

(1) 知っているか、いないかについて次の符号をつける。

- よく知っていていつも使っていると思う語
- ν 聞け・読め ば意味がわかると思う語
- △ 聞いた・読んだ ことはあるが意味のはっきりしない語
- × ぜんぜん分らない語

この4つの符号のうち、森岡は○とνとを理解語とし、△と×とを理解されぬ語として取り扱っているが、よく見ると、○は「よく知っていていつも使っていると思う語」とある。この「いつも使っている」というのは、それが使用語であるというようにも受け取れる。つまり、森岡は内省法でも使用語彙が調査できるとは考えなかったため○とνとを理解語としたが、生徒たちへの説明の仕方次第では、○を使用語、νを理解語として調査できたのではないかと思われる。

一方、語彙量調査では調査語彙をどうするかも問題である。簡単に行うなら、辞書からランダムに選んだ少数の単語から全体を推し量る標本調査だが、全数調査のほうがコストはかかるが、やはりより正確である。そのうえ、全数調査は、のちに単語ごとの利用率もわかるため、日本人の使用語彙の選定にもつながる。よって、可能なら、全数調査が望ましい。

また、附印法だけでは、間違いが生じることは、ここまで見てきた過去の調査でわかっている。しかし、定義法や関連語を書かせるのは調査者にも被調査者にも負担がかかる。よって、より負担の少ない方法として、筆者が佛教大学の語彙量調査の際に用いた、各単語の横に意味をつけたもので調査する方法を用いるのが望ましいと考える。

以上のことから考えると、個人の理解語彙、使用語彙の調査は、

- ①内省法で調査する。
- ②全数調査で行う。
- ③各語の横に意味をつけたもので調査する。

というのが最も良い方法だと思われる。

5. おわりに

大野晋は、『日本語練習帳』の中で、大学生の語彙量は15000か20000くらいに落ちているのではないかと書いた。しかし、大学生といっても、語彙量は、大学によっても違うし、同じ大学でも学部によって違うであろうし、さらに学生個々人によって違うはずである。おそらく、大野が言いたかったのは、細かな数字のことではなく、昔に比べて今の若者は、語彙量が減っていると思われるが、これでいいのだろうかということなのだろう。

確かに、いろいろなところで、最近の若者は本を読まなくなったので、語彙量が減っているといった話は聞く。しかし、実際に大学生の正確な語彙量がわかっていないのだから、これでは、説得力に欠ける。

そこで、対象を今回は大学生に絞って、理解語彙、使用語彙の調査しようと考えた。そして、本年の4月から、内省法を使った理解語彙、使用語彙の調査を15名の大学生を対象に行っている。これは、上記の①～③の条件で調査している。簡単に言えば、辞書に理解語なら○を付け、使用語ならチェックを入れるという内省法を使った方法で理解語彙、使用語彙の調査を行っているのである。

辞書は、年内に全て回収する予定だが、回収し次第、データ分析をし、結果を公表したい。

注

- (1) 大久保(1967)は、自己診断をしてみると、調査できたのは6～7割で、全語彙の採集は到底困難と述べている。このように、幼児でも、使用語彙の調査は難しい。
- (2) 玉村文郎(1984)『日本語教育指導参考書12 語彙の研究と教育(上)』国立国語研究所 p97
- (3) 日本語教育学会編(2005)『日本語教育事典』大修館書店 p284(執筆・山崎誠)
- (4) (2)に同じ
- (5) 玉村文郎(1984)『日本語教育指導参考書12 語彙の研究と教育(上)』国立国語研究所 p100
- (6) 林四郎(1971)「語彙調査と基本語彙」『国立国語研究所報告39電子計算機による国語研究Ⅲ』国立国語研究所 p8

- (7) (6)と同じ
- (8) 吃音で調査が困難なもの、教師に馴れず泣き出すもの、退学者、帰国子女などは除いたとある。
- (9) 平均語数は、成城小学校4089語、鳴浜小学校5019語、師範附属小学校5230語であるが、ただ、調査に使った語の数は、成城小学校6867語、鳴浜小学校11908語、師範附属小学校28661語とそれぞれ違う。成城小学校と鳴浜小学校の平均語数の差は、一見、調査に使った語の数の違いによるように見えるが、鳴浜小学校と師範附属小学校は調査に使った語の数が倍以上も違うのに、平均語数は変わらない。その理由をはっきりとは分らないが、鳴浜小学校が調査に使った語の数、約12000語あたりが調査するのにちょうどいい語数の可能性がある。
- (10) 国立国語研究所 (1951) 国研報告 2 『言語生活の実態 一白河市および附近の農村における一』 秀英出版 p280
- (11) 国立国語研究所 (1971) 国研報告 41 『待遇表現の実態 一松江24 時間調査資料から一』 にも「24時間調査」の調査報告が載っている。この調査は落合春雄氏宅に家族や訪問者に見えないように録音機を置き録音した落合春雄氏と同居家族 4 名、訪問者 24 名の計 29 名の調査だったが、先の 2 つの 24 時間調査とは違い、待遇表現の使い分けが調査の主な目的のため、また、資料処理の関係で、16 時間分 (午前 6 時から午後 10 時まで) の資料全部の分析はできず、扱ったのは半分の 8 時間分にとどまったため、文節の延べ語数や異なり語 (自立語) 数は公表されていない。しかし、16 時間分の資料全部の分析ができたとしても、当日、落合氏宅で行われた会話のほとんどは記録されたが、家族が外出した場合の発話は記録されておらず、また訪問者も、訪ねてきたその時間帯しか録音されていないため、最初から文節の延べ語数や異なり語 (自立語) 数は調査の対象外だったのであろう。
- (12) 国立国語研究所 (1951) 国研報告 2 『言語生活の実態 一白河市および附近の農村における一』 秀英出版 p285
- (13) 阪本一郎 (1955) 『読みと作文の心理』 牧書店 p129
- (14) (6)と同じ
- (15) 阪本一郎 (1955) 『読みと作文の心理』 牧書店 p130-131
- (16) 森岡健二 (1951) 「義務教育終了者に対する語彙調査の試み」『国立国語研究所年報 2』 国立国語研究所 p96
- (17) 森岡健二 (1951) 「義務教育終了者に対する語彙調査の試み」『国立国語研究所年報 2』 国立国語研究所 p100
- (18) 森岡健二 (1951) 「義務教育終了者に対する語彙調査の試み」『国立国語研究所年報 2』 国立国語研究所 p106-107
- (19) 「語彙数推定テスト」
<http://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/goi-test.html> (2014.9.30)
- (20) 単語親密度とは、その単語がどの程度「なじみ」があると感じられるかの主観的評定値である。NTT データベースシリーズ「日本語の語彙特性」第 1 巻の単語親密度データベースには、新明解国語辞典第 4 版の見出し語約 7 万語に対する親密度評定値が収録されている。この親密度は、20 代前半を中心とした 32 名が 7 段階評定 (1:なじみがない、7:なじみがある) を行った結果の平均値となっている。

<http://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/intro.html> (2014.9.30)

(21) NHK 「NHKアナウンスルーム」

<http://www9.nhk.or.jp/a-room/qa> (2014.9.30)

資料1 佛教大学の語彙量調査で使用した語の横に意味を付けたテスト。意味は『新明解国語辞典第七版』、『大辞林第三版』などから採った。

* 以下の語と意味を見て、意味を知っている語なら左の数字に○をつけなさい。(複数の意味がある場合は1つ知っていればいい)

テスト1

1	チャンピオン	①選手権保持者。②第一人者。
2	祝日	①何かの祝いの意をこめた休日。②「国民の祝日」の略。
3	爆発	①化学(物理)反応が急激に進む結果、圧力が異常に高まり、音(光・熱)を伴って破壊作用を起こすこと。②発散できずに心中にたまっていった感情がおさえきれなくなって一度に荒しい行動になって現われること。
4	ライン	①線。②航(空)路。③列。行。④系列。⑤経営のいろいろな部門ごとに、縦に分けた組織(に属する人)。
5	さつま芋	畑に作る多年草。茎はつるになって地にはう。根は塊状ででんぷんに富み甘くて食用。品種が多い。甘薯。
6	毒ガス	人畜に損傷を与えるために使われる有毒ガスの総称。主として戦争で使われる。
7	枝豆	枝が付いたままの、熟す前のダイズ。塩ゆでにして、ビールのおつまみなどにして食べる。青豆。
8	過ごす	①時間を費やす。②度を越して何かをする。
9	朝風呂	「あさゆ」(朝立てる湯。また、朝の入浴。)の口頭語的表現。
10	そもそも	①問題となる事柄を論じるのに先立って、その根源にさかのぼって事を説き起こす様子。②以下に述べる事が、話題の中心となる事柄を理解してもらう上で前提となることである意を表わす。
11	見極める	物事の動向や真偽・本質などをしっかりと見抜く。
12	あべこべ	順序・位置・関係などが、本来あるべき状態(今まで)とは逆であること(様子)。
13	本題	話題や議論などで主として取り上げる事柄。
14	エンゲル係数	生活費の中で食費の占める割合を表わす数値。この係数が高いほど生活水準が低いとされる。
15	泊まり込む	何かの事情で、帰宅せずにそこに泊まる。
16	預け入れる	銀行などの自分の口座に新たにお金を預ける。

17	言い直す	①〔前に言った事の誤りを訂正して〕もう一度言う。②ほかのもつと適当な(やさしい)言葉で言う。
18	たしなみ	①好きで親しむ芸事(などの一応の心得)。②人前でしゃべったり失敗ないように日ごろ言動に気を配ること。
19	英文学	イギリス文学(を研究する学問)。(広義では、アメリカ文学を含む)
20	はまり役	〔演劇・仕事などで〕その人に最も適した役(目)。適役。
21	ごろ合わせ	①ある文句の口調をまねて別の文句を作ること。②主に連続した数字などに意味をこじつけてよませること。
22	労力	何かをするのに要する人手。
23	忍ばせる	人に知られないようにして、何かをする。忍ばす。
24	勃発	事件などが表面化すること。起こること。
25	宿無し	泊まる家や、住む家が無いこと(人)。
26	目白押し	多人数が込み合って並ぶこと。また、物事が集中してあること。例「今年は洋画の話題作が目白押しだ」
27	請負い	請け負うこと(負ってする仕事)。
28	塗り箸	漆塗りの箸。
29	気丈さ	非常の際にも動転せず、平常心で事が処理出来る様子のこと。
30	茶番	①茶の接待をする人。②〔江戸時代、芝居の楽屋で茶番の下回りなどが始めたからという〕手近な物などを用いて行う滑稽な寸劇や話芸。③底の割れたばかばかしい行為や物事。茶番劇。
31	大腿骨	ふとももの部分にあつて、上半身を支えている骨。
32	術中	相手が仕掛けたわな。
33	泌尿器	尿の分泌・排泄に関係する臓器。腎臓・尿管・膀胱・尿道から成る。
34	血税	①明治時代、兵役義務の称。②血のするような苦勞をして納める税。
35	悶着	感情がもつれたり意見が分かれたりして互いに争うこと。
36	腰元	昔、貴人のそばに仕えた侍女。
37	裾模様	①婦人の礼服などの裾の方につけた模様。②裾の方に模様をつけた着物。
38	旗竿	旗をつける竿。
39	かんじき	雪の中に踏み込まないように、はきものの下につける輪のような形のもの。
40	百葉箱	地上の気象を観測するために屋外に置く、よろい戸のついた白い木箱。〔中に、自記温度計・湿度計などを入れる〕
41	迂曲	うねり曲が(って遠回りす)ること。
42	告諭	〔部下・目下の者に〕注意すべき事柄などを言い聞かせること。

43	辻番	町辻や一定地域の警備をすること。辻番所によってその付近を警備すること。また、その役目や人。
44	ライニング	①洋服の裏地。また、裏地を付けること。②腐食・摩耗などの防止のために、あるものの内側に他の材料をはりつけること。裏付け。裏張り。
45	輪タク	〔タクはタクシーの略〕自転車の後部（側面）に客が乗る席を設けた乗り物。〔第二次世界大戦後、一時期流行した〕
46	懸軍	（後方との連絡が無いまま）敵地に深く入り込むこと。また、その軍隊。
47	陣鐘	陣中で、軍勢の進退などの合図に打ち鳴らした銅鑼、または半鐘。
48	泥濘	道などのぬかっている所。ぬかるみ。
49	パララックス	〔写真で〕フィルムに映る像と、ファインダーに映る像とのずれ。視差。
50	頑冥不靈	頑迷で無知なこと。また、そのさま。

テスト2

1	アレルギー	注射や、特定の飲食物・薬の摂取によって、体質上、正常者とは異なる過敏な反応を示すこと。ペニシリンショックやじんましんなど。〔広義では、特定の人・物事に対する拒絶反応を指す〕
2	夫	〔妻に対して〕その人と結婚している男性。
3	食生活	日常生活のうちで、毎日の食事に関係のある面。
4	冷える	①冷たく（寒く）なる。また、そう感じる。②それまでの盛り上がった状態や感情が失われる方向に傾く。
5	無責任	①その事について責任を負うべき立場にないこと（様子）。②その事について責任を負うべき立場にありながら、責任を負おうとしない（免れようとする）こと（様子）。
6	黒板	〔学校などに備える〕白墨で字や図を書くための黒色・緑色の板。
7	実物大	実物と同じ大きさ。
8	玄米	〔白米と違って〕もみがらを取り去っただけで、まだ精白していない米。
9	本心	①本来の正しい心。②その人の本当の心。
10	混浴	男女が同じ浴場で同時に入浴することが出来ること。
11	バウンド	堅い地面（床）に当たったボールなどが、はずんではね返ること。バンド。
12	富	人間の生活を豊かにするのに役立つ物資・資源。
13	ブルマー	女性のはく、（運動用の）下ばき。すそにゴムを入れてしぼってある。
14	知恵の輪	幾つかの輪を、うまくつなぎ合わせたり抜き離したりして遊ぶもの。

15	強烈さ	堪えられないほど強い様子のこと。
16	不用心	①どろぼうなどに対する警戒が足りない様子だ。②ぶっそうな様子だ。
17	湯たんぼ	中に湯を入れ寝床などに入れて、足・腰などをあたためる道具。
18	気掛かり	好ましくない事態が予測されて、不安な気持をふっ切れない様子だ。
19	天然色	映画などで、自然の色彩に似せて表わした色。
20	書き記す	「書く」意の改まった表現。
21	ノアの方舟	〔旧約聖書の物語で〕ノアの家族と数種の動物が乗って洪水をのがれた、四角な大きい船。
22	波しぶき	波と波とがぶつかったり波が岩などにぶつかったりして立てるしぶき。
23	企てる	あることをしようと計画する（して着手する）。
24	勃発	事件などが表面化すること。起こること。
25	存じ上げる	「知る・思う」意の謙譲・丁寧語。
26	山越え	①山を越えること（所）。②昔、関所札を持たない者が、間道を抜けて山を越えたこと。
27	古びる	古くなって、往年の華やいだ美しさが失われる。
28	興ざめる	何かがきっかけとなって、せっかくのおもしろみ（愉快的気分）が無くなる。
29	針供養	二月八日〔地方によっては十二月八日〕に、裁縫を休んで、使えなくなった針を集めて、供養すること（行事）。
30	かっちり	むだなすきまやたるみがどこにも認められない様子。例「蓋をかっちり（と）閉める」
31	仰々しさ	そのものの見かけや言動などが、一般に予測される程度を越えていて、そうまでする必要があったのかという印象を与える様子のこと。
32	港湾	船の碇泊や乗客・貨物の上げおろしに便利な水域とその付属設備との総称。
33	モラリスト	道徳的な人。また、道徳について語る人や書く人。
34	あばら屋	住む人が無く、壊れかかった家。廃屋同然の意で、自宅の謙称にも用いられる。
35	耳新しさ	耳に聞いて新鮮であること。初めて聞くこと。聞いて珍しいこと。
36	土付かず	〔すもうで〕全勝。
37	切磋	①骨・角・石・玉などを刻みみがくこと。②知徳・学芸をねりみがくこと。学問、技芸などに努め励むこと。
38	丸まっつい	いかにもまるまるとしている。まるまるとして可愛らしい。

39	無体さ	①どうしてそのような目にあわなければならないのか、どう考えても理由が見出せないようなこと。②〔法律で〕物質的な形をそなえていないこと。
40	のしあわび	アワビの肉を薄くむいて伸ばし、干したもの。〔儀式用のさかなとした〕
41	陸半球	地球の半球のうちで、陸の占める面積が最大となるものの称。その中心〔=極〕は、パリの西南西約二二〇キロの北緯四八度、東経三〇分にあり、全面積の四九パーセントが陸で、地球の全陸地の八四パーセントを含む。
42	道化方	〔歌舞伎で〕こっけいな所作を得意とする役柄（役者）。
43	自小作	自作（自分の所有地を自力で耕作する農業）を主として、小作（地主から土地を借りて農業をすること）も兼ねること。また、その人。
44	糸道	三味線を多くひいたために、左の人さし指の爪の先に出来たくぼみ。
45	皇大神宮	伊勢神宮の中の内宮。祭神は天照大神。
46	道芝	道ばたに生えた芝。
47	藻塩	昔、海藻に海水をそそぎ焼いて水に溶かし、そのうわずみをとって煮つめた塩。また、それを作るためにくんだ海水。
48	労音	「勤労者音楽協議会」の略。勤労者のための音楽鑑賞組織。昭和24年（1949）大阪で発足、以後、各地に広がり、同30年には全国労音連絡会議が結成された。
49	身体装検器	音の反射を利用して、からだに隠している金属などを着物の上から見つける器械。空港などで使われる。
50	蜀江の錦	①昔、中国の蜀で作った、精巧な錦。②京都西陣で織り出した錦の一種。蜀錦。

テスト3

1	楽しみ	楽しいと思うこと（対象）。
2	ボディー	①人間のからだ。②胴体（に相当する部分）。③人台
3	色々	〔もと、違った色が数多く有る意〕そのものの状態・性質が一通りでなかったり 関係するところが多面にわたっていたり あれこれと方法を尽くしていたりすること（様子）。
4	始発	①そこから出発すること（列車・バス）。②その日一番早く運行されること（列車・電車・バス）。
5	シーソー	長い板の中央部をささえ、両端に人が乗って互いに上下させるもの（遊び）。
6	定期預金	銀行が三か月・六か月・一年・二年などの期限を定めて預かる預金。普通預金に比べて、利率が高い。

7	黒こしょう	よく熟さない胡椒を外側の黒い皮と一緒に干したもの。かおりが高い。
8	ややこしい	事柄がこみいって、扱いがめんどうな様子だ。
9	なにしろ	事情がどうであれ、結論的にはその一点に尽きると判断する様子。
10	色男	①美男子。②情夫。間男。
11	手間	①仕事を完成するのに要する労力・時間。②〔一日幾らの日当てで仕事をする〕職人の仕事。③「手間賃」の略。
12	接点	①〔幾何学で〕接線が曲線（曲面）と共有する点。〔広義では、二つの物事が接触する点を指す。〕②〔写真で〕カメラにストロボを接続する所。
13	全速力	出せる限りの速力。フルスピード。全速。
14	模範	それを見習うべき、りっぱなやり方（手本）。
15	蓄積	〔資本・知識・エネルギーなどを〕いざという時役立てるために、たくわえておくこと。〔疲労の持ち越しなどについても言う。〕
16	運搬	物品を、目的地へ運ぶこと。
17	引き止める	①〔手・そでなどを〕引っ張るようにして止める（やめさせる）。②立ち去ろうとするものをとどめる。
18	いずこ	どこ。
19	引き離す	①引っ張って（無理に）離す。②〔競走などで〕あとに続く者との距離・間隔を大きく隔てる。
20	のたれ死に	行き倒れ（のような哀れな死に方）。
21	大それた	思い上がりもはなはだしく、その身分にある者としてあるまじき言動をする。
22	言い渡す	それに従うよう、決定したことを告げる。
23	上げ下ろし	手に持った物や荷物などを上げることと下ろすこと。
24	入り江	海・湖などが陸地に食い込んだようになっている地形。
25	プラント	①工場施設・機械装置（の一切）。②ミルク プラントの略。
26	スパンコール	ぴかぴか光る装飾用の薄い金属またはプラスチックなどの小片。舞台衣装やドレスなどに縫いつける。
27	事務次官	その省の事務処理の最高責任者として大臣を輔佐する役の人。
28	国民体育大会	各都道府県代表選手により毎年行なわれる全国的なスポーツ大会。昭和二十一年に始まった。国体。
29	置き所	①置くべき場所。②置いた場所。
30	コロニー	①植民地（の植民者が集団で居住している所）。②治りかかった患者・身体障害者・精神遅滞児などを集めて、開放的な自然環境の中で、保護したり訓練したりする施設。③集落。

31	お召し替え	「着替え・乗換」の尊敬語。
32	長らえる	長く生きのびる。
33	茶の湯	客を茶室に招き入れ、茶をたててすすめること（作法）。茶道。
34	腹帯	①〔寝冷えを防いだり、健康のために〕腹に巻く帯。腹巻。②岩田帯。妊婦が五か月目（の戌の日）から腹に巻く帯。③〔鞍をつけるために〕馬の腹にしめる帯。
35	うらぶれる	おちぶれたり不幸な目にあったりして、みじめなありさまになる。
36	坂東	関東地方の古称。
37	寺子	寺子屋（江戸時代、庶民の子供に読書・習字などの初等教育を行なった所）に入門した子供。
38	鳩派	相手の主張の採るべきところは採り、事を穏やかに解決しようとする考え方を持つ人たち。〔広義では、和平論者を指す〕
39	変光星	見かけの明るさが周期的に変化する恒星。
40	辻説法	道ばたで通行する大衆に対してする説法。
41	夜の目	夜、眠る目。夜、眠るべき目
42	戸車	戸の上（下）に取りつけて開閉をなめらかにする、小さな車輪。
43	縞物	縞の模様を織り出した織物。
44	頻々	〔望ましくないことが〕何度も続けて起こる様子だ。
45	庶流	庶子の系統。〔広義では、次男・三男などで分家した家筋を指す。〕
46	累世	「累代」（何代にもわたって代を重ねること）の意の古風な表現。
47	ズルチン	フェナセチンを原料とした、砂糖の代用品。現在では使用禁止。
48	御璽	天皇の御印。玉璽。
49	陪従	身分の高い人のお供をすること（人）。
50	さしったり	①かねてから待ち構えていたときに発する語。待っていました。②物事をしくじったりなどして、残念に感ずるときに発する語。しまった。

資料2 佛教大学の語彙量調査のデータをまとめたもの

NO	1	1の意味付	2	2の意味付	3	3の意味付	1-3平均	1-3の意味付平均
1	64800	64800	63100	63100	60300	61700	62733	63200
2	58800	64600	43800	67800	52600	55600	51733	62667
3	53100	59900	48200	56800	41500	48700	47600	55133
4	59400	59400	57400	48200	41500	40500	52767	49367
5	55200	53100	55200	55200	57100	57100	55833	55133
6	53100	53100	50800	50800	60300	54600	54733	52833

7	53100	53100	45900	50800	37100	39300	45367	47733
8	34100	53100	26100	40500	20200	27600	26800	40400
9	53100	53100	37100	39100	41500	34100	43900	42100
10	53100	53100	39300	36700	37100	37100	43167	42300
11	53100	53100	37100	34700	35200	35200	41800	41000
12	53100	51800	33500	45900	37100	41500	41233	46400
13	50500	50500	48100	64500	48700	48700	49100	54567
14	50500	50500	35400	35400	28400	29300	38100	38400
15	54600	49300	52800	52800	49300	49300	52233	50467
16	48200	48200	58100	58100	41300	29300	49200	45200
17	48200	48200	53100	53100	43400	31800	48233	44367
18	48200	48200	48200	47200	35100	35100	43833	43500
19	59900	47100	58100	54500	41300	41300	53100	47633
20	48200	47100	43400	43400	45900	45900	45833	45467
21	48200	47100	37700	41300	33300	35100	39733	41167
22	47100	47100	39300	25600	37100	29100	41167	33933
23	47100	46100	41300	41300	41300	39200	43233	42200
24	46100	46100	33300	37100	43400	43400	40933	42200
25	46100	46100	37700	35600	43800	33300	42533	38333
26	39500	46100	39300	35400	41500	29700	40100	37067
27	53100	46100	38000	28800	35100	27600	42067	34167
28	46100	46100	33400	28400	35200	32000	38233	35500
29	45900	45900	35200	35200	39300	31900	40133	37667
30	44800	44800	43400	43400	45100	41500	44433	43233
31	50500	44800	41300	41300	43400	35100	45067	40400
32	57500	44800	45900	33600	37100	37100	46833	38500
33	45800	44800	35400	33300	48500	48500	43233	42200
34	45100	44100	43400	43400	41500	33300	43333	40267
35	45800	43800	45900	45900	35100	40500	42267	43400
36	43500	43500	37100	43400	37100	37100	39233	41333
37	30500	42800	30400	33500	22200	25500	27700	33933
38	43200	42500	35100	25600	37100	37100	38467	35067
39	47100	41600	43400	43400	37100	37100	42533	40700

40	55200	41600	43400	36000	39200	32500	45934	36700
41	53100	41500	51900	27500	37100	24700	47367	31233
42	41300	41300	30400	45900	32000	35100	34567	40767
43	55900	41100	46100	50800	31900	35300	44633	42400
44	31700	41100	31700	35100	30400	27400	31267	34533
45	48200	41100	31900	31900	39200	37100	39767	36700
46	48200	39500	41300	55100	20800	29700	36767	41433
47	30500	39500	23400	30400	21200	22300	25033	30733
48	47100	39500	31900	24000	35100	29900	38033	31133
49	45900	39100	45900	39600	37100	41300	42967	40000
50	46100	39100	33500	39300	32000	35100	37200	37833
51	39100	39100	48500	38000	43400	37100	43667	38067
52	41100	39100	48500	37300	37100	32300	42233	36233
53	39100	39100	37700	27500	37600	37600	38133	38067
54	41100	37100	38400	38400	28700	28700	36067	34733
55	37100	37100	19800	27500	40500	37100	32467	33900
56	44800	37000	20400	20400	27400	27400	30867	28267
57	46100	36000	38400	18600	33300	25700	39267	26767
58	39100	35600	48200	48200	29900	29900	39067	37900
59	37100	35600	31900	31900	28800	28700	32600	32067
60	28200	35200	33300	28800	37100	27200	32867	30400
61	39100	35100	30200	33300	28700	28700	32667	32367
62	45800	35100	30200	27500	37100	30900	37700	31167
63	36800	34700	31900	33500	33500	32000	34067	33400
64	44100	34600	40500	37200	32300	32900	38967	34900
65	44100	33500	39300	39300	43400	37100	42267	36633
66	44100	33500	40500	38000	30300	23500	38300	31667
67	41100	33500	37100	37100	39300	33300	39167	34633
68	33300	33500	28800	28800	27400	21100	29833	27800
69	37100	33500	28800	23400	24900	20200	30267	25700
70	41600	32900	28800	27500	37100	30900	35833	30433
71	46100	32800	36800	33600	38000	31800	38800	32733
72	32000	32000	30400	27500	26100	32000	29500	30500

73	31900	31900	35400	28700	27600	30200	31633	30267
74	48200	30600	45900	21400	44700	16500	46267	22833
75	30300	30200	31900	31900	23500	23500	28567	28533
76	37100	30200	27500	23400	30900	30900	31833	28167
77	35100	30200	27500	22200	21100	21100	27900	24500
78	32500	28900	38000	26100	29100	28600	33200	27867
79	41100	28600	39300	45900	24700	29000	35033	34500
80	33300	26800	23400	19800	25200	20200	27300	22267
81	25600	25600	36800	28800	27400	27400	29933	27267
82	33500	24700	27500	27500	21100	24900	27367	25700
83	23400	23400	27500	22600	21400	21400	24100	22467
84	35100	23200	27500	30400	23500	21100	28700	24900
85	25500	22500	27500	20000	21100	14900	24700	19133
86	35100	21000	38400	17400	27700	12800	33733	17067
87	28800	20600	26100	26100	23500	20200	26133	22300
88	15300	15300	16100	13000	16100	16100	15833	14800
89	45800	13500	38900	26100	30300	18900	38333	19500
90	15900	13500	8700	8700	9000	10900	11200	11033
91	39100	39100	35200	35200	45900	45900	40067	40067
92	41300	41300	33300	33300	28400	28400	34333	34333
93	37100	37100	39200	39200	37100	37100	37800	37800
94	61800	61800	60100	60100	61700	61700	61200	61200
95	58800	58800	45900	45900	43400	43400	49367	49367
96	37100	37100	37700	37700	32000	32000	35600	35600
平均	43283	40119	37982	36504	35244	33035	38821	36588

引用文献

- 荒牧英治・増川佐知子・森田瑞樹・保田祥(2012)「オンライン・コミュニケーション上での平均使用語彙数は8,000語である」『研究報告自然言語処理 (NL)』2012-NL-208
 巻9号 情報処理学会
- 大久保愛(1967)『幼児言語の発達』東京堂出版
- 大野晋(1999)『日本語練習帳』岩波書店
- 語彙数推定テスト解説
<http://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/intro.html>
- 国立国語研究所(1951)国研報告2『言語生活の実態—白河市および附近の農村における—』秀英出版

- 国立国語研究所(1953)国研報告5『地域社会の言語生活—鶴岡における実態調査—』秀英出版
- 国立国語研究所(1971)国研報告41『待遇表現の実態—松江24時間調査資料から—』秀英出版
- 国立国語研究所(2000)国研報告116『日本語基本語彙—文献解題と研究—』明治書院
- 阪本一郎(1955)『読みと作文の心理』牧書店
- 澤柳政太郎・田中末廣・永田新(1919)『児童語彙の研究』同文館
- 玉村文郎(1984)『日本語教育指導参考書12 語彙の研究と教育(上)』国立国語研究所
- 中尾桂子・柴田実・中谷由郁・平林一利(2012)『「文章表現」指導内容再考のための一考察—学生の語彙量、記述上の形式的規則に見られる問題点の観察をもとに—』『大妻女子大学紀要—文系—』No44
- 日本語教育学会編(2005)『日本語教育事典』大修館書店
- 林四郎(1971)「語彙調査と基本語彙」『国立国語研究所報告39電子計算機による国語研究III』国立国語研究所
- 森岡健二(1951)「義務教育終了者に対する語彙調査の試み」『国立国語研究所年報2』国立国語研究所