

# 大学生の日本語の使用語彙、理解語彙

——使用語彙は平均約33000語——

荻原 廣

1. はじめに
2. 使用語彙、理解語彙とは
3. 先行研究
4. 大学2年生の日本語の使用語彙、理解語彙の調査
  - 4-1. 使用語彙の調査方法とは
  - 4-2. 今回の調査を行うにあたって
  - 4-3. 今回の日本語の使用語彙、理解語彙の調査
    - 4-3-1. 被調査者
    - 4-3-2. 調査時期
    - 4-3-3. 調査方法
    - 4-3-4. 調査結果
    - 4-3-5. 考察
  - 4-4. 今回の調査における問題点
5. 大学2年生と大学4年生の日本語の使用語彙、理解語彙の比較
  - 5-1. 大学4年生の日本語の使用語彙、理解語彙
  - 5-2. 大学2年生と大学4年生の日本語の理解語彙の比較
  - 5-3. 大学2年生と大学4年生の日本語の使用語彙の比較
  - 5-4. 結論
6. 使用語彙と理解語彙の相関関係
7. おわりに

まず、大学2年生の日本語の使用語彙、理解語彙について行った調査について述べた。次に、荻原（2016）で行われた大学4年生の日本語の使用語彙、理解語彙の調査と比較をした。その結果、大学生の日本語の使用語彙は、2年生、4年生とも平均約33000語、一方、大学生の日本語の理解語彙は、2年生が平均約43000語、4年生が平均約45000語であることがわかった。また、使用語彙と理解語彙の間に相関関係は見られなかった。

## 1. はじめに

荻原（2016）で、2014年度に行われた大学4年生12名の語彙量（使用語彙、理解語彙）に関する調査（以下、「大学4年生の語彙量調査」とする）の結果、使用語彙は平均30899語、理解語彙は平均45354語ということが明らかになった<sup>①</sup>。しかし、この調査は、2014年度、2015年度の2年にわたり、大学2年生と大学4年生を対象に行った調査の半分であり、荻原（2016）の調査結果は、そのうち大学4年生に対して行った2014年度の調査結果のみであった。そこで、今回は、2015年度に行った大学2年生13名の語彙量（使用語彙、理解語彙）に関する調査（以下、「大学2年生の語彙量調査」とする）の結果について明らかにし、更に、大学2年生と大学4年生ではどういった違いがあるのかについて比較検討も行う。

## 2. 使用語彙、理解語彙とは

使用語彙、理解語彙が過去にどのように定義されてきたかは、荻原（2016）に詳しいため、ここでは触れないが、「大学4年生の語彙量調査」同様、今回の「大学2年生の語彙量調査」も、荻原（2014）の「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、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」との定義<sup>②</sup>で調査を行った。

## 3. 先行研究

荻原（2014）に、今までに行われた個人の使用語彙、理解語彙の調査のうち、幼児を対象とした調査を除いた主なものについて、その概要が書かれている。また、荻原（2016）にも、阪本の理解語彙の調査、森岡の理解語彙の調査、語彙数推定テスト<sup>③</sup>を利用した理解語彙の調査、荒牧・増川・森田・保田の使用語彙の調査について述べてあるため、ここでは触れない。

## 4. 大学2年生の日本語の使用語彙・理解語彙の調査

### 4-1. 使用語彙の調査方法とは

理解語彙に関する調査は、今まで何度も行われてきたが、使用語彙に関する調査は、ほとんど行われていない。それは、使用語彙の調査に対するアプローチは、従来、外から捕捉するしかないと考えられていて、更にその捕捉が容易

ではないからであろう。しかし、萩原（2016）では、

アウトプット、つまり書いたり話したりする際に使用できる語というのは、自分以外の誰かの使用語彙を外側から調査するなら、確かに、「書かれたり話されたりしたもの」以外に知る方法はないであろうが、被調査者自身が自分について調査するなら、当然、表出する前は、内側にあるのだから、内側で測定する（内省法で測る）ことは可能なはずであり、その個々の結果を集計すれば、使用語彙の調査となる。つまり、従来、理解語彙にのみ利用されていた調査法の内省法は、使用語彙についても利用できるというわけである<sup>④</sup>。

と述べ、また国立国語研究所が行った『外来語定着度調査』（平成14年11月～平成16年8月随時実施）<sup>⑤</sup>や文化庁文化庁国語課が行っている国語に関する世論調査<sup>⑥</sup>を例に挙げ、内省法での使用語彙の調査が可能であるとした。

#### 4-2. 今回の調査を行うにあたって

今回の調査も、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」ではなく、「大学4年生の語彙量調査」同様、辞書をそのまま使用し、辞書の見出し語が理解語なら○をつけ、使用語なら符号vをつけることとした「全数調査」で行った。これは、同じく全数調査（ただし、理解語彙についてのみ）を行った森岡（1951）の「義務教育終了者に対する語彙調査の試み」<sup>⑦</sup>を参考にし、考案したものである。

なお、辞書は、「大学4年生の語彙量調査」同様、今回も収録項目数が約82000の『三省堂国語辞典第七版』を使用することにした。ただし、接頭語、接尾語、造語成分、連語、慣用句（イディオム）などを除いたいわゆる学校文法の10品詞についてのみ扱い、それ以外は、採用しないこととしたため、調査対象となる語の数は82000語より少ない。

#### 4-3. 今回の日本語の使用語彙・理解語彙の調査

##### 4-3-1. 被調査者

当初、佛教大学（通学生）の2年生20名（女性17名、男性3名）に被調査者の依頼をした。しかし、7名（女性5名、男性2名）は、提出の締め切り日である2015年度の年度末までに調査の終わった辞書の提出が間に合わず、最終的

に締め切り前に提出があったのは、13名（女性12名、男性1名）であった。なお、この13名は、全員が日本文学科所属の学生であった。

#### 4-3-2. 調査時期

「大学4年生の語彙量調査」は2014年4月～2015年3月の1年間にわたって行われたが、今回は、語彙量調査の被調査者になるにあたっての説明会を行い、その後に調査用の辞書を配布したのが7月29日と遅く、「大学4年生の語彙量調査」より約4か月も短い最長でも約8か月間の調査となった。その結果、リタイアを7名と多く出した可能性が否定できず、その点では、残念な結果となった。

なお、調査が完了した被調査者13名のうち、一番早く調査の終わった被調査者は、2015年12月22日に調査で使用した辞書を提出し、一番調査に時間のかかった被調査者は、2016年3月29日に調査で使用した辞書を提出した。

#### 4-3-3. 調査方法

調査を行うにあたって、被調査者に、「大学4年生の語彙量調査」同様、注意を書いた文書を配布した上で、口頭でそれをもとに説明した<sup>®</sup>。これは、被調査者全員、「使用語彙」「理解語彙」に対する認識を共通にしておくためである。

#### 4-3-4. 調査結果

被調査者13名が使用語彙と理解語彙をチェックした辞書をデータとして整理し、まとめたものが、以下の表1である。

**表1** 大学2年生の日本語の理解語彙と使用語彙及びその割合（使用語彙／理解語彙） 単位＝語

被調査者	理解語彙	使用語彙	割合 (使用語彙／理解語彙)
M	28018	12783	45.6%
T	33763	14499	42.9%
V	35491	24003	67.6%
P	37486	34285	91.5%
R	42858	26132	61.0%

X	43849	34693	79.1%
O	43861	40046	91.3%
U	46162	34411	74.5%
W	46853	31821	67.9%
S	47565	45567	95.8%
N	48863	42688	87.4%
Y	51240	35707	69.7%
Q	56805	54945	96.7%
平均	43293	33198	76.7%
中央値	43861	34411	

では、この表をもとに、以下に考察する。

#### 4-3-5. 考察

大学2年生の日本語の理解語彙を少ないほうから並べたのが図1である。

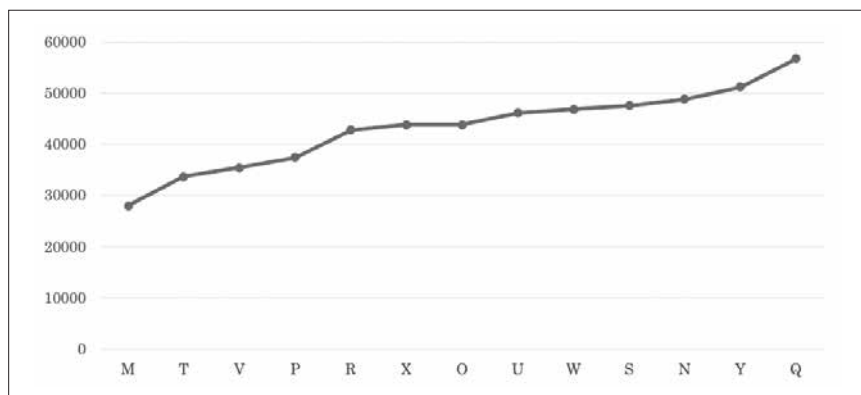


図1 大学2年生の日本語の理解語彙

縦軸は語彙量、横軸は、被調査者に与えられた記号である。右へ行くにしたがって、緩やかに上昇しており、個人差があることが、読み取れる。最小値、最大値、平均値、中央値、標準偏差は以下の通りである。

最小値 28018語

最大値 56805語  
平均値 43293語  
中央値 43861語  
標準偏差 7521語

今回の調査では、最小値と最大値の差は、約2.03倍であった。つまり、同じ大学の同じ学科で同じ学年でも、理解語彙の最も多い学生と最も少ない学生の間には、2倍以上の差があり、結果として、理解語彙というのは個人差が大きいということがわかった。

一方、平均値は43293語であった。なお、中央値は43861語で、中央値は平均値の約1.01倍と、ほぼ差がなかった。

次に、大学2年生の日本語の使用語彙を少ないほうから並べたのが図2である。

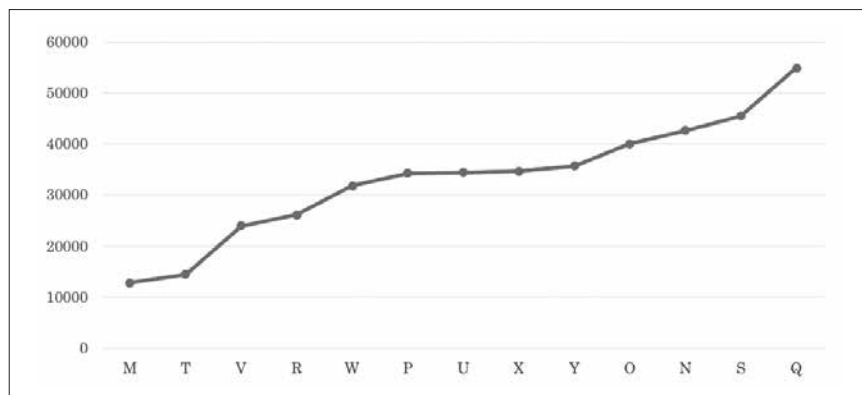


図2 大学2年生の日本語の使用語彙

縦軸は語彙量、横軸は、被調査者に与えられた番号である。右へ行くにしたがって、より大きく上昇していて、個人差が、理解語彙に比べ、よりはっきりと窺える。最小値、最大値、平均値、中央値、標準偏差は以下の通りである。

最小値 12783語  
最大値 54945語

平均値 33198語  
中央値 34411語  
標準偏差 11357語

今回の調査では、最小値と最大値の差は、約4.30倍であった。これは、理解語彙の最小値と最大値の差が約2.03倍だったのに比べると、更に個人差が大きいと言える。

一方、平均値は33198語であった。なお、中央値は34411語で、中央値は平均値の約1.04倍と、ほぼ差がなかった。

#### 4-4. 今回の調査における問題点

今回も、調査を行った際に、問題となった点がいくつかあったので、ここでその点について挙げておく。

まず、問題となったのは、荻原（2016）同様、今回も調査の際に、漏れている箇所が見つかったことである。やはり、全数調査という長期間にわたる調査ゆえ、一定の割合で記載漏れが生じる可能性は否定できない。13名中、記載漏れのあった4名（被調査者Q、W、X、Y）について、それぞれ、抜けたページ数を総ページ数で割ると、被調査者Qは、理解語彙が1.8%程度、使用語彙が1.9%程度、被調査者Wは、理解語彙、使用語彙ともに0.8%程度、被調査者Xは理解語彙、使用語彙ともに0.5%程度、被調査者Yは、理解語彙が0.1%程度、使用語彙が0.7%程度、それぞれ語彙量が増える可能性がある。

次に、これも荻原（2016）と同様ではあるが、推定される語彙の最大数が辞書の語数以上にはならないことである。よって、過去に行われた理解語彙の調査が全てそうだったように、このテストで推定された語彙量においても、実際に知っている単語の数よりも少ない可能性があり、それは、使用語彙においても同じである。

また、今回も、全数調査であったため、調査協力者（被調査者）が20名（うち、データとして使用できたのは13名）と少なかった。ただ、グラフを見ると、同じ大学で同じ学年、しかも同じ学科の学生であるにもかかわらず、今回も、使用語彙においても理解語彙においても、かなりの個人差があることがわかった。

## 5. 大学2年生と大学4年生の日本語の使用語彙、理解語彙の比較

### 5-1. 大学4年生の日本語の使用語彙、理解語彙

では、荻原（2016）で行われた「大学4年生の語彙量調査」と今回行った「大学2年生の語彙量調査」を比較し、そこから見えてくることについて述べていきたい。そこで、まずは、「大学4年生の語彙量調査」がどうであったかについて述べる。

「大学4年生の語彙量調査」には、被調査者12名が使用語彙と理解語彙をチェックした辞書をデータとして整理し、まとめたものが掲載されているが、それが以下の表2である。

表2 大学4年生の日本語の理解語彙と使用語彙及びその割合（使用語彙／理解語彙） 単位＝語

被調査者	理解語彙	使用語彙	割合 (使用語彙／理解語彙)
G	28142	16890	60.0%
C	34840	30616	87.9%
K	38785	4096	10.6%
H	41909	30814	73.5%
L	42244	33069	78.3%
F	42473	12408	29.2%
J	44473	34820	78.3%
D	47702	25446	53.3%
I	50243	40158	79.9%
E	54029	41192	76.2%
B	58362	46088	79.0%
A	61045	55187	90.4%
平均	45354	30899	68.1%
中央値	43473	31942	77.3%

大学4年生の日本語の理解語彙の最小値、最大値、平均値、中央値、標準偏差は以下の通りである。



最小値 28142語  
 最大値 61045語  
 平均値 45354語  
 中央値 43473語  
 標準偏差 9131語

大学4年生の日本語の使用語彙の最小値、最大値、平均値、中央値、標準偏差は以下の通りである。

最小値 4096語  
 最大値 55187語  
 平均値 30899語  
 中央値 31942語  
 標準偏差 13900語

では、表1と表2より、最小値、最大値、平均値、中央値、標準偏差について、比較してみる。

## 5-2. 大学2年生と大学4年生の日本語の理解語彙の比較

**表3** 大学2年生と大学4年生の日本語の理解語彙の比較 単位＝語

	2年生 (13名)	4年生 (12名)
最小値	28018語	28142語
最大値	56805語	61045語
平均値	43293語	45354語
中央値	43861語	43473語
標準偏差	7521語	9131語

こうして比較すると、最小値は、ほぼ変わらないが、最大値は、4年生の方が4240語も多く、また、平均値も4年生の方が2061語多かった。また、標準偏差を見ると、4年生の方がばらつきが大きいことがわかる。

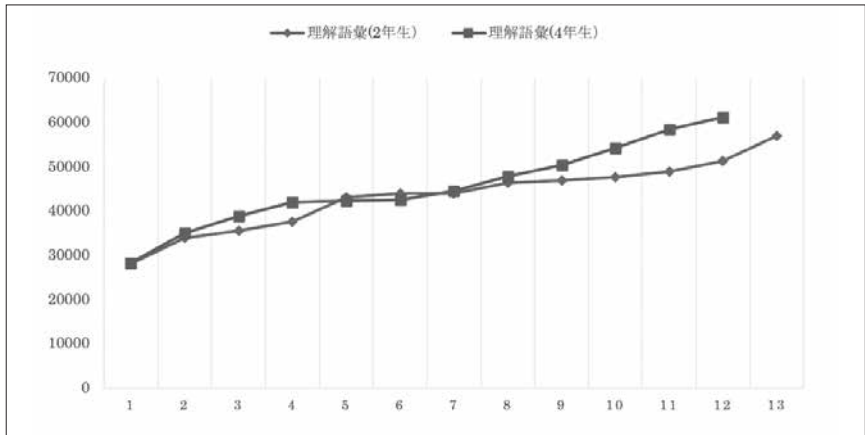


図3 大学2年生と大学4年生の日本語の理解語彙の比較

図3は、大学2年生と大学4年生の日本語の理解語彙を少ない方から並べたグラフを重ねて比較したものだが、これを見ても、大学4年生のほうが全体的に理解語彙がやや多いことが窺える。

### 5-3. 大学2年生と大学4年生の日本語の使用語彙の比較

表4 大学2年生と大学4年生の日本語の使用語彙の比較(1) 単位=語

	2年生(13名)	4年生(12名)
最小値	12783語	4096語
最大値	54945語	55187語
平均値	33198語	30899語
中央値	34411語	31942語
標準偏差	11357語	13900語

こうして比較すると、最大値は、ほぼ変わらないが、最小値は4240語と2年生の方が3倍以上も多く、また、平均値も2年生の方が2299語多かった。これは、どういうことだろうか。理解語彙の平均値は4年生の方が約2000語多かったにもかかわらず、使用語彙の平均値は逆に2年生の方が約2000語多いという一見矛盾した結果が出た。しかし、これには荻原(2016)の以下のような記述

にヒントがある。

最後に、最小値と最大値の差だが、最小値の被調査者Kは、調査終了時、本人より、自分は無口なため、使用語彙は、他の人よりずっと少ないのではないかという発言があった。使用語彙には話す際の語彙以外に書く際の語彙も含まれるが、無口なタイプは使用語彙が少ない可能性があるのではないだろうか<sup>9)</sup>。

この発言は、使用語彙が4096語と最小値だった4年生の被調査者が調査終了時に語ったことだが、ここには「使用語彙には話す際の語彙以外に書く際の語彙も含まれるが、無口なタイプは使用語彙が少ない可能性があるのではないだろうか。」とある。しかし、「自分は無口なため、使用語彙は、他の人よりずっと少ないのではないか」(下線部は筆者による)という発言から改めて考えると、被調査者がチェックした語は、「話す際に使用する語彙」だったのではないかと推察される。実際、使用語彙が4000語程度では、大学の卒業論文を書くには、いかにも少ないことからそのように推察できる。

そこで、この使用語彙が4096語と最小値だった4年生の被調査者を除いて、計算し直すと、以下のような結果となった。

表4 大学2年生と大学4年生の日本語の使用語彙の比較(2) 単位=語

	2年生(13名)	4年生(11名)
最小値	12783語	12408語
最大値	54945語	55187語
平均値	33198語	33335語
中央値	34411語	33069語
標準偏差	11357語	11813語

この結果を見ると、大学2年生の日本語の使用語彙と大学4年生の日本語の使用語彙は、最小値、最大値をはじめ、全ての項目でほとんど差がないことがわかる。

#### 5-4. 結論

大学2年生の日本語の理解語彙と大学4年生の日本語の理解語彙は、平均値において、大学4年生のほうが約2000語多かった。一方で、大学2年生の日本語の使用語彙と大学4年生の日本語の使用語彙は、平均値において、ほとんど差がなかった。ここから考えられるのは、大学2年生で使用語彙はほぼ上限に達し、それ以降増えるのは、主に理解語彙なのではないだろうかということである。つまり、大学3年生以降に増えるのは、専門教育における専門用語のような語彙であり、そしてそれらは主に理解語彙であるのではないだろうか。

以上、ここまでの考察により、この論文においては、大学生の日本語の使用語彙は、平均約33000語<sup>⑩</sup>、一方、大学生の日本語の理解語彙は、2年生が平均約43000語、4年生が平均約45000語と結論付ける。

#### 6. 使用語彙と理解語彙の相関関係

図4は、使用語彙が4096語と最小値だった4年生の被験者を除いた大学2年生、大学4年生、合計24名の使用語彙、理解語彙を表わしたグラフである。

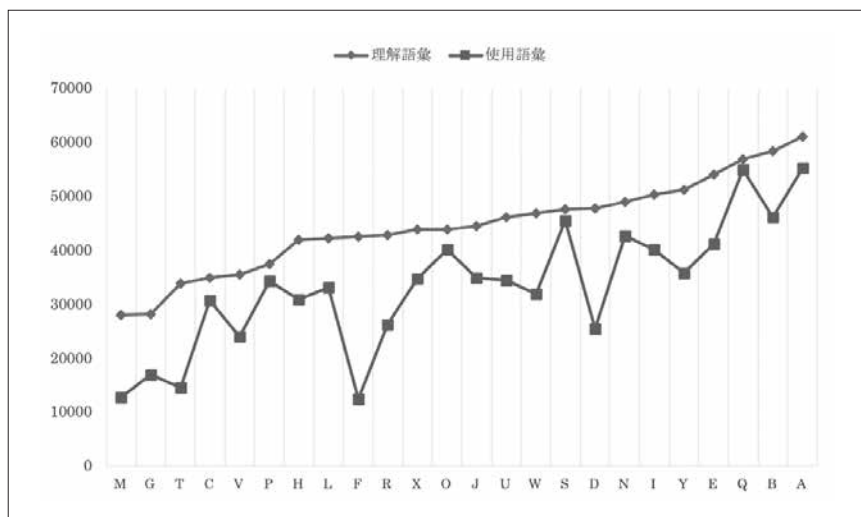


図4 大学生の日本語の使用語彙と理解語彙の相関関係

縦軸が語彙量、横軸が被調査者に与えられた番号である。中原（2015）は、「理解語彙の増加によって、場面に応じて適切に用いることのできる使用語彙

も増加する可能性がある。』<sup>①</sup>としているが、使用語彙と理解語彙の差があまりない被調査者と差が大きな被調査者が散らばっており、このグラフからは、理解語彙が多いからと言って、必ずしも使用語彙も多いとは一概には言えないことが読み取れる。

## 7. おわりに

荻原（2016）で、大学4年生の日本語の使用語彙と理解語彙のデータを分析し、それについて述べたが、今回は、大学2年生の日本語の使用語彙と理解語彙のデータを分析し、それについて述べ、更に、2つの学年のデータを比較し分析した。

この2014年～2016年の2年間にわたって行った35名の調査（うち10名がリタイアにより分析対象者は25名）に関する分析は今回で終了する。

しかし、この成人25名の日本語の使用語彙調査の結果は、今後、抜けた部分のデータの再調査をし、24名分（25名から使用語彙が最小値だった4年生の被験者1名を除いた24名）のデータを整備した後、今度は、学生ごとではなく、単語ごとの使用率のデータとして転用する予定である。そして、約82000語を使用率順に並べたデータをもとに、成人の日本語の使用語彙の選定作業を行う予定である。

なお、この日本語の使用語彙の選定は、日本語教育における使用語彙選定の資料となり、やがては、テキストに、これは使用語彙、これは理解語彙と、教師個人の判断ではなく、客観的なデータとして載せることができるようになるだろう。

最後に、大変労力のかかる調査に協力してくれた佛教大学の35名の被調査者の学生たちに、あらためて感謝したい。

### 注

①荻原廣（2016）「大学4年生の日本語の使用語彙は平均約3万語、理解語彙は平均約4万5千語」『京都語文』23 佛教大学国語国文学会 pp290-291

②荻原廣（2014）「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（日本語学特集）」『京都語文』21 佛教大学国語国文学会 p2

③語彙数推定テストは、2020年6月4日に「令和版語彙数推定テスト」として新しいテストとなり公開されている。（それまでのテストは「平成版語彙数推定テスト」となっている）

NTTコミュニケーション科学基礎研究所「単語親密度データベース」

<http://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/index.html> (2020. 9. 27閲覧)

- ④荻原廣 (2016)「大学4年生の日本語の使用語彙は平均約3万語、理解語彙は平均約4万5千語」『京都語文』23 佛教大学国語国文学会 p284
- ⑤国立国語研究所 (2007)『国立国語研究所報告126 公共媒体の外来語 —「外来語」言い換え提案を支える調査研究—』国立国語研究所 p272
- なお、『外来語定着度調査』は、理解率を「使ったことがある」で調査しているが、「使ったことがある」かどうかを内省法で聞いた場合、使ったことがあるかどうか自信が持てず、使えるが、使ったことがあるとまでは自信を持って言えないので使用語彙に入れないなどと考え、結果として、正確な調査結果が出ない可能性があるため、回答の選択肢は、「使う」あるいは「使うことがある」で調査したほうがよいであろうと思われる。
- ⑥複数回行われているが、例えば、『平成13年度国語に関する世論調査』(平成14年1月調査)のQ4を見ると、「あなたは、ここに挙げた(1)から(10)の言葉を使うことがありますか。また、意味がわかりますか。」との間で、回答の選択肢は「使う」「使わないが、意味は分かる」「使わないし、意味も分からない」「分からない」となっていて、これは「使う」＝「使用率」、「使わないが、意味は分かる」＝「理解率」を調べる調査と言えるだろう。つまり、調査対象となる語を使用するかどうかについて内省法で答えてもっているのである。
- ⑦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
- ⑧荻原廣 (2016)「大学4年生の日本語の使用語彙は平均約3万語、理解語彙は平均約4万5千語」『京都語文』23 佛教大学国語国文学会 pp288-289
- ⑨荻原廣 (2016)「大学4年生の日本語の使用語彙は平均約3万語、理解語彙は平均約4万5千語」『京都語文』23 佛教大学国語国文学会 p293
- ⑩一見すると、使用語彙が多いように見えるが、それは、「高使用頻度語彙＝使用語彙」ではないにもかかわらず、「よく使う語＝使用語彙」であるというイメージが一般にあるからではないか。しかし、使用語彙の中には、当然ながら、めったに使わないが使える語彙というものが多い存在しており、それが使用語彙の語彙量を押上げていると言える。この件については、以下において詳しく考察されている。
- 荻原廣 (2016)「大学4年生の日本語の使用語彙は平均約3万語、理解語彙は平均約4万5千語」『京都語文』23 佛教大学国語国文学会 pp291-292
- ⑪中原香苗 (2015)「『文章読解』における授業外語彙学習の取り組み」『教育開発センタージャーナル』6 神戸学院大学教育開発センター p101

#### 【参考文献】

- 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 p1-8
- 荻原廣 (2014)「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 (日本語学特集)」『京都語文』21 佛教大学国語国文学会 p1-30
- 荻原廣 (2016)「大学4年生の日本語の使用語彙は平均約3万語、理解語彙は平均約4万

- 5千語』『京都語文』23 佛教大学国語国文学会  
見坊豪紀・市川孝・飛田良文・山崎誠・飯間浩明・塩田雄大編（2014）『三省堂国語辞典』  
第七版 三省堂  
国立国語研究所（1964）『国立国語研究所報告25 現代雑誌九十種の用語用字 第三分冊：  
分析』秀英出版  
国立国語研究所（2007）『国立国語研究所報告126 公共媒体の外来語 ―「外来語」言い  
換え提案を支える調査研究―』国立国語研究所  
阪本一郎（1955）『読みと作文の心理』牧書店  
田島ますみ・佐藤尚子・橋本美香・松下達彦・笹尾洋介（2016）中央学院大学人間・自然  
論叢41 p3-20  
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一』『大妻  
女子大学紀要一文系―』44 大妻女子大学 p1-17  
中原香苗（2015）『『文章読解』における授業外語彙学習の取り組み』『教育開発センター  
ジャーナル』6 神戸学院大学教育開発センター p91-101  
日本語教育学会編（2005）『日本語教育事典』大修館書店  
文化庁（2002）『平成13年度 国語に関する世論調査 日本人の言語能力を考える』財務  
省印刷局  
文化庁（2015）『平成26年度 国語に関する世論調査 漢字の形・言葉遣い・新しい言葉』  
ぎょうせい  
松浦年男（2015）「大学初年次の学生に対する日本語語彙力調査の試行」『北星学園大学文  
学部北星論集52（2）』北星学園大学 p53-61  
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 昭和  
25年度』秀英出版

