

[論文]

# 日本語学術共通語彙テストの開発

田島ますみ  
佐藤尚子  
橋本美香  
松下達彦  
笹尾洋介

- 〈目次〉
- 1 はじめに
  - 2 学術共通語彙知識を測るテスト
    - 2.1 日本語学術共通語彙
    - 2.2 テストの開発
  - 3 テスト実施概要
  - 4 結果と考察
    - 4.1 基礎統計量と信頼度
    - 4.2 大学生の学術共通語彙知識
      - 4.2.1 学術共通語彙と一般語彙の結果比較
      - 4.2.2 正答率の低かった語
      - 4.2.3 テキストジャンル別頻度と正答率の相関
  - 5 まとめと今後の課題

## 1 はじめに

日本人大学生の日本語語彙力の問題は以前から指摘され、特に理解語彙量の測定に関してはいくつかの調査がある。中尾・柴田・中谷・平林(2012)は、女子大での調査で平均34,900語、松浦(2015)は大学1年生のテスト結果として平均33,611語、荻原(2016)は大学4年生での調査結果として平均45,354語という数値を報告している。

この他に、田島・佐藤・橋本・松下・笹尾(2016a)は、書き言葉コーパスの頻度順位上位30,000語から抽出した語を対象語として語彙テスト(Vocabulary Size Test for Reading Japanese 30K, 以下VSTRJ-30K)を作成し、大学1年生を中心に実施したが、高得点者が多く、天井効果が認められた。この結果を受けて対象語の範囲を頻度順位上位50,000語に広げたテスト(Vocabulary Size Test for Reading Japanese 50K, 以下VSTRJ-50K)では、40,000語前後(平均点からの推計では41,840語)という数値が報告された(佐藤・田島・橋本・松下・笹尾, 2017)。

日本人大学生の理解語彙量について40,000語前後という結果が出たが、頻度順位50,000語までの語彙テストであっても高得点者が多いので、さらに弁別力の高いテストが必要である。一つの方向性として、テストの対象語を学術系の語彙に絞る案が考えられる。学術系の語彙は、大学での学びを支える基礎となる語彙であり、これを測るテストがあれば大学生の語彙力を議論する際にも有効な数値・資料を提供できるであろう。本稿では、書き言葉コーパスの中から学術系の語彙に限定して語彙テストを開発し、大学生に実施した結果を報告する。

## 2 学術共通語彙知識を測るテスト

### 2.1 日本語学術共通語彙

学術系の語彙を明確に限定するため、今回は松下（2011）が日本語の大規模書き言葉コーパスから計量的な手法によって抽出した語彙リスト「日本語学術共通語彙リスト」に基づいてテスト開発を行った。

学術共通語彙とは、一般的な文章に比べて学術的な文章でより高い使用率を占める語彙を指す。さらに「共通」の語が示すように、特定の専門領域でのみ使用率の高い専門語彙ではなく、学術的な文章一般で使用率の高い語彙である。人文、社会、理工、生物・医学の4ジャンルのうち3ジャンル以上で特徴語となっている語彙なので（Matsushita, 2012）、文系、理系を問わず学術的・論理的な思考・活動に多用される基礎的な語彙であり、一般語彙と専門語彙の中間に位置する語彙だと言える。抽象性が高く、論理操作に不可欠である語が多い。たとえば「推察」「概観」「狭まる」「ファクター」など、大学教育においては、学部学科、専門専攻を問わず、論理的思考、あるいは講義の聴講、レポートの作成で必要不可欠な語彙である。

### 2.2 テストの開発

この「日本語学術共通語彙リスト」から、テストの対象語を、一般語彙も含めた総合的なデータベースにおける使用頻度順位を基準に選定した。頻度順位上位20,000語の中から250語に1語を抽出することとし、可能な限り抽出した語の順位から次の順位までが等間隔となるよう抽出した。そのようにして得られた80語から、外国語としての日本語教育で初級語彙とされる頻度順位、1,285位より上位の初級語彙5語を除いた75語を対象語とした。初級語彙には著しい領域特徴性がないことに加え、大学生の語彙量を測る問題としては易しすぎ、十分な弁別力を得られないと予想されるためである。

75語の対象語について、短い例文を示し、語義を問う選択問題とするという点は、VSTRJ-30K および VSTRJ-50K と同様である。これらの一般語彙テストでは、四肢択一式であったが、本研究では三肢択一式を採用した。その理由は、四肢択一を三肢択一にしても項目難度への影響が小さいこと、選択肢数よりも項目数の方がテスト結果の信頼度に与える影響が大きいこと、選択肢数よりも錯乱肢の質が当て推量の抑制に効果があることといった先行研究の結果を考慮に入れたことである (Costin, 1970, 1972; Kolstad, Kolstad, & Wagner, 1986; Rodriguez, 2005)。

また、選択肢の質を高めるために、これまでのテストのものと比べて、より正確な意味知識を必要とするものにした。これまでの錯乱肢は、基本義がわかっているならば、深い意味知識はなくてもよいという方針のもと、対象語と同じような頻度順位レベルにある他の語の意味・言い換えで作ってきた。そのため、正解の選択肢とは明らかに異なるものとなり、そのことで類推や消去法などのテスト受験方略で正解が得られてしまうのではないかと考えたためである。今回の作問にあたっては、他の語の意味・言い換えで錯乱肢を作らず、対象語のより正確な理解を必要とする問題となるようにした。

図1に、問題例を示す。75の対象語に関し、例のような作問をし、問題の順番を無作為に変えた二つのバージョン (A, B) の問題冊子を作成した。テスト問題に順序効果のないことを確認するためである。

帰結：一つの帰結である。

- 1) 行ったり来たりする関係
- 2) 最後にまとまった考えや状態
- 3) 初めに出てきた問題

図1 語彙サイズテスト問題例

### 3 テスト実施概要

以上のように作成した三肢択一75問の日本語学術共通語彙テスト (Japanese Common Academic Word Test, 以下JCAWT Ver. 1) を, 2016年4月7日から18日の間に, 3大学の1年次配当日本語科目の授業内で実施した. 実施時間はVSTRJ-30K および VSTRJ-50K と同様に40分とした. 各大学で A, B, 二つのバージョンの実施結果がなるべく同数となるように留意した. 合計で A バージョン259名, B バージョン251名, 計510名のデータを分析対象とした. テスト受験者は大半が大学1年生であるが, 当該科目の再履修クラスでも実施したので, 2年生12名, 3年生2名, 4年生3名が含まれている.

## 4 結果と考察

### 4.1 基礎統計量と信頼度

テスト結果の基礎統計量を表1に示す. 1問あたりの配点を1点とした75点満点のテストで, 平均点は62.4, 標準偏差は8.0であった. 分布は高得点に偏っているものの, 天井効果による影響は限定的であると考えられる. また, クロンバック  $\alpha = .88$  であり, テスト得点の信頼度 (内的一貫性) は高いといえる. A と B の二つのバージョンによる差は検出されなかった ( $t(508) = 0.833, p = .405, \text{Cohen's } d = 0.07$ ) ため, 問題の順番によるテスト得点への影響は小さいと言える.

表1 素点の基礎統計量

平均	標準偏差	最大値	最小値
62.4	8.0	74	24

さらに、テスト実施時間についても VSTRJ (30K および50K) より短縮が可能であることが確認された。今回、VSTRJ (30K および50K) と同様40分で実施したが、受験状況を見る限り40分は不要であった。日本語母語話者の大学生で実施する場合、30分以下の時間で十分なことが確認された。

以上の結果から、大学生の学術語彙力を簡易に測定するテストとして JCAWT が有効であることが示唆された。

## 4.2 大学生の学術共通語彙知識

### 4.2.1 学術共通語彙と一般語彙の結果比較

テストの有効性を確認した上で、大学生の学術共通語彙知識の観点からテスト結果を分析する。図2はテストの素点を5点ごとに区切って示した人数分布である。高得点のほうへ傾いた分布となり、この分布傾向は VSTRJ と類似し、特に VSTRJ-50K と似たものとなった。

平均正答率は83.2%であり、これも VSTRJ-50K の平均正答率83.7%（佐藤ほか, 2017）と近似している。松下（2011）による日本語学術共通語彙リストは書き言葉コーパスの頻度順位で上位20,000語の中からの抽出であり、50,000語レベルの一般語彙テスト（VSTRJ-50K）の正答率と同じような正答

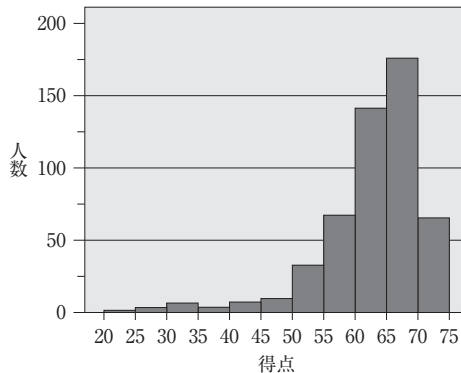


図2 素点の得点分布

率が出たことから、学術共通語彙の理解度は一般語彙に比べ低いことが示唆される。なお、2016年度の VSTRJ-50K と今回の JCAWT の受験者および選択肢作成法が異なるので正確な比較ではない。だが、どちらも同じ3大学で同じ1年次配当日本語科目で実施しているので、ある程度、同質な受験者が受けたということは言える。VSTRJ-50K の正答率が83.7%、30,000語レベルの一般語彙テスト (VSTRJ-30K) の正答率が90.1% (田島ほか, 2016a) であったことを考えれば、20,000語レベルの一般語彙テストの正答率は90.1%より高くなると考えられるが、結果は83.2%であった。よって、学術共通語彙は、頻度が高く比較的容易だろうと考えられる語であっても、大学生の理解度が一般語彙に比べて低いように思われる。

また、一般語彙と学術共通語彙で、正答率と対象語の頻度順位の相関係数 (スピアマンの順位相関) が顕著に異なったことも、今回のテスト実施結果の中で注目される点である。VSTRJ-50K では  $\rho = -.486$  であったのに対し、JCAWT では  $\rho = -.277$  と、相関が弱いという結果であった。語彙習得において頻度は重要な要素で、通常、使用頻度の高い語彙は習得者へのインプットの量が多いことを意味し、習得されやすい語彙と見なされる。また、語彙の習得がおおむね頻度順に進むことは、第二言語としての習得過程では既に確かめられている (Read, 1988)。今回の結果では、正答率と頻度の相関が弱いことが示され、学術共通語彙の習得は頻度順に進むという単純な過程ではないことが示唆される。頻度との相関が弱く、繰り返し接するだけでは習得が進んでいかないとすると、単にインプットを増やすだけの指導法では効果が限定的となるおそれがある。指導側からのより効果的な教育的介入が必要であり、また、工夫により伸ばせる部分でもあると言えるだろう。

さらに述べておきたい点としては、一部の学生が著しく低得点であったことである。これは、VSTRJ-30K および50K でも同様の結果が出ている (田島ほか, 2016a; 佐藤ほか, 2017)。今回の結果では、510名中35名が50点以下であった。彼らは、テストの対象語の3分の1以上が正確に理解できていないおそれがある。学術共通語彙が正確に理解できていなければ、その基礎の上

に積み上げられていく専門語彙の習得はさらに困難と考えられる。専門教育課程だけでなく、それ以前の基礎教育や教養科目の学習においても支障があるだろう。初年次教育の段階での何らかの対策が求められる。

#### 4.2.2 正答率の低かった語

表2に正答率が80%未満であった対象語を示す。なお、統計分析でテストの作問等に問題があると思われるものは除いた。

正答率が特に低かった語は、「踏襲」、「編著」、「テーゼ」、「所与」、「鑑みる」などの日常の会話表現ではあまり使用されない語が多い。これらの語は学術的、またはあらたまった日本語の文章を読んでいなければ接する機会が極めて少ないと考えられ、読書量の少ない現在の大学生が理解していない状況は納得のいくところである。

頻度順位が上位であるにもかかわらず低正答率であった「順序」は、興味深い例である。この対象語に関しては、選択肢を「話の続け方」「進んでい

表2 正答率の低かった語

正答率降順順位	頻度順位	対象語	正答率 (%)
1	44	踏襲	33.5
2	55	編著	46.1
3	15	順序	52.2
4	58	テーゼ	56.9
5	41	重ね合わせる	61.6
6	70	作製	61.6
7	75	所与	62.7
8	43	鑑みる	67.6
9	35	能動	69.8
10	34	指向	70.8
11	27	台頭	71.2
12	32	-考	72.4
13	50	国内外	74.1
14	68	小数	77.6
15	52	暗黙	77.8
16	25	集大成	77.8
17	62	省令	79.6



く方向」「決まった並べ方」として提示した。正解は「決まった並べ方」で、「順序」の語義は「続け方」でも「方向」でもないのだが、「並べ方」という語義の中心部の理解が不確実なのではないかという理由が考えられる。高頻度で難度の低いと思われる基本語彙であっても理解が正確ではない語があることが示唆される。

また、「能動」「指向」「台頭」「-考」「集大成」も、頻度順位は10,000位以内で、全問題項目の中では中央より上位であるが、同レベルの他項目に比べて正答率が低かった。これらは、頻度は比較的高いが、抽象度が高く、言語活動や思考の苦手な学生にとっては正確な意味把握が難しい語彙と考えられる。

#### 4.2.3 テキストジャンル別頻度と正答率の相関

JCAWT は頻度順位を基準に対象語を選定しているので、頻度と正答率の関係をより詳細に検討してみた。書き言葉コーパスの頻度順位と正答率の相関は前述したように、 $\rho = -.277$ と弱かった。この相関を、コーパスにおけるテキストのジャンル別に見た結果を表3に示す。頻度順位はテキストのジ

表3 テキストジャンル別頻度順位と  
正答率の相関 (Spearman's  $\rho$ )

テキストジャンル	頻度順位との相関係数
全体	-.277*
文芸	-.347**
言語哲学	-.114
歴史民俗	-.145
芸術	-.219
政治法律	-.257*
経済商業	-.248*
社会教育	-.202
科学技術	-.182
生物医学	-.145
ネットフォーラム	-.393**

\*  $p < .05$ , \*\*  $p < .01$

ジャンルによって異なってくる。たとえば、「省令」は政治法律系のテキストでは頻度順位が上で、文芸系テキストでは頻度順位が低くなる。そのようにジャンルによって異なる頻度順位と正答率の相関を見た。

頻度と比較的相関が高かったジャンルは、文芸とネットフォーラムであった。これは、大学に入学してきた学生たちがどのようなテキストに接してきたかを間接的に示すものである。「頻度が高ければ習得する」という前提に立てば、頻度と正答率の相関が強いものは、それだけ接してきている、すなわち読んできているテキストジャンルだと考えられる。文芸作品は高校までの国語教育の教科書で多く取り上げられるものであるし、読書量が少ないと言われる大学生であってもライトノベルやキャラクター小説などは読むという者もいる。ネットフォーラムでの頻度順位と正答率の相関が最も強いというのは大学生の読んでいるものの現状をよく示していると思われる。

ネットフォーラムや文芸のジャンルの次に正答率と相関の高かったのは「政治法律」と「経済商業」のテキストであった。ネットフォーラム・文芸とこれら政治・経済の計四つのジャンルにおいてのみ頻度順位と正答率との間に有意な相関が検出された。政治・経済のテキストで頻度の高い語の正答率が高くなることから、ニュース・新聞などからのインプットは接触頻度が高いことがうかがえる。新聞を読まないと言われている大学生ではあるが、彼らはニュース・新聞などをある程度読んでいるのではないだろうか。相対的な比率として、彼らが読んでいるものの中でニュース・新聞は高いのではないかということが推察される。もう一つの可能性としては、学術共通語彙の中でも政治・経済系の語は、ネットフォーラムや文芸系の語と合わせて、他のテキストジャンルで多用される語よりも相対的に、頻度の高いインプットにより、すなわち繰り返し接することで習得されやすいということも考えられる。

一方、正答率と頻度順位との間に有意な相関が検出されなかったのは「言語哲学」「歴史民俗」「生物医学」などのジャンルである。その原因として考えられるのは、上述したニュース・新聞などからのインプットについての二

つの可能性と対になることである。すなわち、大学生達がこれらのジャンルのテキストを読んでいないということ、また、これらのテキストジャンルで頻繁に使われる語彙はニュース・新聞などとの単純な繰り返しの接触だけでは習得されにくいということである。

## 5 まとめと今後の課題

以上、日本語の学術共通語彙知識を測定する目的で、JCAWTを作成し、実際に日本語母語の大学生に実施して結果を検討してみた。JCAWTがテストとして一定の弁別力と信頼性を有することが確認されたほか、大学生は、語の頻度を基準にした場合、一般語彙に比べ学術共通語彙の理解が進んでいないと思われること、学術共通語彙の理解が不十分で大学での学びに支障が出ると考えられる学生が一定数存在すること、書き言葉を読んでいないと接する機会が少なくなる語の正答率が特に低いこと、高頻度語であっても理解の正確でない語が存在すること、抽象度の高い語は理解度が低くなることなどが明らかになった。テキストジャンル別の考察に関しては、頻度と正答率の相関から大学生の読んでいるテキストジャンルと語彙の種類による習得のしやすさ・しにくさについて言及したが、仮説の域を出ていない。

今後は、JCAWTのテストとしての質を高めていくほかに、テキストジャンルとの関連も含めて、日本語母語話者の大学生が習得しにくい語彙はどういうものなのかをより詳細に検討し、大学の初年次教育やリメディアル教育に資することを目指す予定である。もともと学術共通語彙は、日本語非母語話者が日本の高等教育機関で学ぶ場合に、学術系のテキストで使用頻度の高い語彙を集中して学べば読解において効果的であるということで選定された語彙である。第二言語としての日本語学習者だけではなく、日本語母語話者である日本人大学生にも効果的な学習方法につながる方向性も模索したい。さらに、学術共通語彙知識はどのように発達していくのかについての概要をとらえるために、大学入学前の、初等・中等教育段階での習得状況の調査に

も着手している。学術共通語彙知識の発達が順調に進まない大学生はどの段階で停滞してしまうのかなどが解明できるのではないかと考えている。

本研究は、科学研究費補助金基盤研究 (C) 「グローバル化に向けた日本語の語彙テスト開発」(課題番号15K02631, 平成27年度～29年度, 研究代表者: 佐藤尚子) の助成を受けた。

### 引用・参考文献

- Costin, F. (1970). The optimal number of alternatives in multiple choice achievement tests: some empirical evidence for a mathematical proof. *Educational and Psychological Measurement*, 30 (2), 353-358.
- Costin, F. (1972). Three-choice versus four-choice items: Implications for reliability and validity of objective achievement tests. *Educational and Psychological Measurement*, 32 (4), 1035-1038.
- Kolstad, R. K., Kolstad, R. A., & Wagner, M. J. (1986). Performance on 3-choice versus 5-choice MC items that measure different skills. *Educational Research Quarterly*, 10, 4-8.
- Matsushita, T. (2012). In what order should learners learn Japanese vocabulary? A corpus-based approach. PhD Thesis. Victoria University of Wellington. <http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/4476/thesis.pdf> (2017年9月7日参照)
- Read, J. (1998). Measuring the vocabulary knowledge of second language learners, *RELC Journal*, 19 (2), 12-25.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24 (2), 3-13.
- 荻原廣 (2014). 「日本人の語彙サイズ (理解語彙, 使用語彙) 調査を行うにあたっての基礎的研究」『京都語文』21, 1-30.
- 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, 276-298.
- 佐藤尚子, 田島ますみ, 橋本美香, 松下達彦, 笹尾洋介 (2017). 「使用頻度に基づく日本語語彙サイズテストの開発—50,000語レベルまでの測定の試み—」『千葉大学国際教養学研究』1, 15-26.

- 田島ますみ, 佐藤尚子, 橋本美香, 松下達彦, 笹尾洋介 (2016a). 「日本人大学生の日本語語彙量測定の試み」『中央学院大学人間・自然論叢』41, 3-20.
- 田島ますみ, 佐藤尚子, 橋本美香, 松下達彦, 笹尾洋介 (2016b). 「日本人大学生は学術共通語彙をどの程度理解しているのか」. 『リメディアル教育学会第12回全国大会発表予稿集』, 118-119
- 田島ますみ, 佐藤尚子, 松下達彦, 笹尾洋介, 橋本美香 (2017). 「日本語学術共通語彙知識の発達—義務教育課程と高等教育課程での習得状況の比較—」『リメディアル教育学会第13回全国大会発表予稿集』, 140-141.
- 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—』44, 108-92.
- 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), 53-61.
- 松下達彦 (2011). 日本語学術共通語彙リスト Ver. 1.01 <http://www17408ui.sakura.ne.jp/tatsum/list.html#jcaw> (2017年9月7日参照)

