

## 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—

佐藤 尚子<sup>1)</sup>・田島 ますみ<sup>2)</sup>・橋本 美香<sup>3)</sup>・松下 達彦<sup>4)</sup>・笹尾 洋介<sup>5)</sup>

<sup>1)</sup>千葉大学国際教養学部    <sup>2)</sup>中央学院大学法学部    <sup>3)</sup>川崎医科大学医学部

<sup>4)</sup>東京大学大学院総合文化研究科    <sup>5)</sup>豊橋技術科学大学総合教育院

### Development of a 50,000-word-level Japanese vocabulary size test based on written-corpus frequency data

Naoko Sato, Masumi Tajima, Mika Hashimoto, Tatsuhiko Matsushita, Yosuke Sasao

#### 要旨

本稿では、日本語母語話者・非母語話者を問わず、多様な日本語使用者の語彙サイズを測定するテストの開発と、そのテストを用いて測定した日本人大学1年次生の語彙サイズの測定結果について述べる。先行して開発された30,000語レベルの語彙サイズテストでは、弁別力に問題があった（田島ほか、2016）ため、50,000語レベルの語彙サイズテストを開発し、実施した。

50,000語レベルのテストの結果、30,000語レベルと同様に高得点者が多かった。分析対象になった400名のうち、96.8%の学生は推計既知語数が30,000語レベルより上であり、74.5%が40,001語以上だった。今回の語彙サイズテストの結果からみれば、多くの日本人大学生の大学入学時の語彙サイズは40,000語前後と推計される。一方、一部に極端に語彙サイズが小さい学生がいることがわかった。また、頻度の低い語であっても、大学1年次生にとって日常生活等で接する機会が多い語、あるいは、漢字から意味が類推できる語の正答率が高い傾向が見られた。

#### キーワード

日本語使用者、使用頻度、語彙サイズの測定、語彙サイズテスト、推計既知語数

## 1. はじめに

近年、日本語母語話者である日本人大学生の日本語能力の低下は多くの大学で問題にされるようになってきている（日本リメディアル教育学会、2012）。日本語能力の中でも、特に、語彙力の低下が問題視されることが多い。語彙力は論理的な思考力を根底で支えるもので、語彙力の低下は大学での学びの質の低下を招く原因となる。語彙力は日本語の重要な要素であり（松下、2016）、その語彙力を測定することによって、日本語能力が、ある程度、予測できることから、「日本語IRTテスト」（NHKエデュケーショナル制作・監修）などが開発されてきた。

しかしながら、日本人大学生の語彙力を数量の面から明らかにすることは、従来、あまり行われてこなかった。大学生として入学時にどのぐらいの語彙サイズがあればいいのか、大学の学びにどの程度の語彙サイズが必要なのか、ということは明らかになっていない。

本稿の研究課題は、（1）全ての日本語使用者を対象とした語彙サイズを測定するテストはどのようにすれば開発できるか、（2）日本語母語話者である日本人大学生の語彙サイズ（推計既知語数）はどの程度か、（3）日本語母語話者である日本人大学生にとってどのような語の意味が既知であり、どのような語の意味が未知であるか、の3点である。

## 2. 先行研究

### 2.1 日本語母語話者の語彙サイズの測定

日本人の語彙サイズについて、過去にどのような調査が行われてきたかについては、萩原（2014）に詳しい。その中に取り上げられている阪本（1955）の調査では、18歳で理解語彙が約50,000語という結果が示されている。また、日本人大学生を対象とした語彙サイズの調査として、大妻女子大学で行われた調査をあげている（中尾・柴田・中谷・平林、2012）。大妻女子大学では、2010年度と2011年度に4年制文学部、短大国文科、家政科、英文科等の1、2年生の学生を対象に、NTTがウェブ上で公開している語彙数推定テストを用いて調査を行った。その結果、全員の平均語彙サイズはおおよそ34,900語であった。しかし、これらの先行研究はいずれも内省法（自己評価）に基づくものであり、客観テストではない。

### 2.2 使用頻度を用いた語彙サイズテストの開発

前述のNTTが開発した語彙数推定テストに使用されている語彙は、NTTデータベースシリーズ「日本語の語彙特性」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ている。当該データベースの語彙親密度は、20代前半を中心とした32名が出題された語に対して7段階評価（1：なじみがない、7：なじみがある）の自己評定を行い、その平均をとった値である。被験者32名には年齢的な偏りがあり、専門領域、語彙サイズなども

統制されておらず、かつごく少数のデータをもとに作成されている。そのため、単語親密度を主基準に抽出された語を用いたテストでは、テスト項目の代表性という観点からは妥当ではないと考えられる。それに対し、第二言語習得においてはおよそ使用頻度順に語彙の習得が進むことが確かめられている（例えばRead, 1988）。このような状況のもと、使用頻度をもとに作成された、第二言語としての日本語教育で開発されてきた語彙サイズを測るテストを母語話者にも応用できないかということを考え、30,000語レベルの語彙サイズテストが開発された（田島ほか、2016）。

### 2.3 30,000語レベルの語彙サイズテストについて

30,000語レベルの語彙サイズテストの作成と実施結果については、田島ほか（2016）で詳しく述べている。ここではその概要を述べる。

テストの形式は前述の「日本語を読むための語彙量テスト」（松下、2012）を踏襲している。このテストが参考にしたのは、英語の語彙サイズテスト（Nation & Beglar, 2007）であり、正解の意味の手がかりを最低限にする非定義文に、テスト項目である目標語を埋め込み、その定義を4つの選択肢から選ぶ形式である（図1）。目標語以外はできるだけ理解を助けるように、目標語を除いて、漢字には全て振り仮名をつけた。150問を40分で解答する形式で、配点は1問1点の150点満点である。出題順による正答率への影響を考慮し、出題順が異なる2つのバージョンを作成して実施している。

目標語は「日本語を読むための語彙データベース」の使用頻度順位30,000語までのリストから、200語につき1語を選んでテスト項目としている。

2014年4月から5月に、3大学の日本人学生を対象とする1年次配当日本語科目の授業で、合計485名の学生を対象に実施し、図2のような得点分布になった。高得点者が多く、150点満点中、平均135.2、標準偏差10.5、最高148、最低74だったが、極端な低得点の受験者が一定程度存在した。

30,000語レベルの語彙サイズテストの得点の分布の形はかなり右側に偏っており、30,000語レベルでは母語話者の語彙サイズを測るのには不十分であり、より難度の高い項

[6000 語レベル]

**礼儀:** あの人<sup>ひと</sup>は**礼儀**<sup>し</sup>を知らない。

- 1) おおよその<sup>ないよう</sup>内容をまとめたもの
- 2) 人間<sup>にんげん</sup>関係<sup>かんけい</sup>を守るための<sup>まも</sup>行動<sup>こうどう</sup>様式<sup>ようしき</sup>
- 3) ものを大切に<sup>たいせつ</sup>思う<sup>おも</sup>気持ち<sup>きもち</sup>
- 4) 理論<sup>りろん</sup>とは異なる<sup>こと</sup>現場<sup>げんば</sup>の仕事<sup>しごと</sup>

\* 実際の問題とは異なる

図1 語彙サイズテスト問題例（田島ほか、2016、p.9）

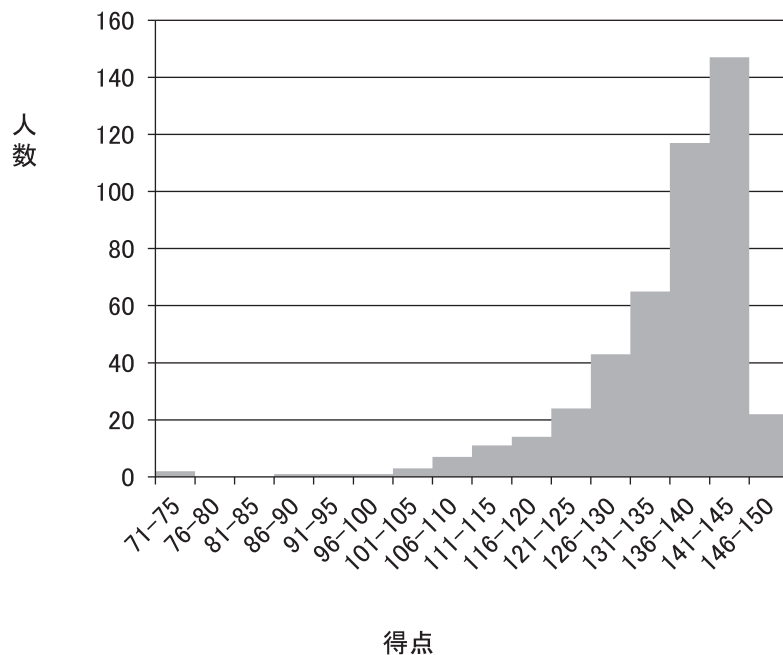


図2 30,000語レベルの語彙サイズテストの得点別人数分布  
(田島ほか、2016、p.11)

目を追加する必要があることが示唆された。

### 3. 50,000語レベルの語彙サイズテストの開発

30,000語レベルの語彙サイズテストの結果を受け、50,000語レベルまでを範囲とするテストを開発、実施することとした。

テスト形式は、松下（2012）や田島ほか（2016）と同じである。125問を40分で解答する。配点は1問1点の125点満点である。出題順による正答率の影響を考慮し、出題順が異なる2つのバージョンを作成した。語彙は「日本語を読むための語彙データベース」（松下、2011b）の使用頻度順位50,000語までのリストから、400語につき1語を選んでテスト項目とした。目標語の選定にあたっては、各語の、データベース（松下、2011b）上の頻度順位（頻度と分散度を掛け合わせた修正頻度）が、50,000を400ずつに割った各ゾーンの真ん中に近づくように無作為抽出した。15,000語レベルまでは松下（2012）の既存の項目の中から、15,000～30,000語のレベルでは、一部を除き、田島ほか（2016）から選び、それ以外は新たに選んだ。ただし、選ばれた項目のうち、データベース作成時の形態素解析の誤りによる成分や、他の項目と同じ字を使用して意味にも重なりがあり、問題項目の独立性を保てない場合など、不適切な項目は抽出しなおして次善の語に変えた。

#### 4. 50,000語レベルの語彙サイズテストの実施

2015年4月から2016年6月にかけて、3大学（国立大学1、私立大学2）で、日本人学生を対象とする1年次配当日本語科目の授業時に、合計407名に実施した。これらの受験者の解答パターンを調べるために、ラッシュモデルに基づく適合度分析（fit analysis）を行った。ラッシュモデル（Rasch, 1960）とは、客観的な測定結果が従うべき数理モデルであり、そのモデルから大きく逸脱した解答パターン（outfit  $t > 4.0$ ）を示す受験者7名を分析から除外した。この7名の受験者の平均は46.0点（正答率36.8%）とかなり低く、また「終える」や「偉い」といった多くの易しい語に不正解であった。本研究では、400名（男性299名、女性100名、無回答1名）を分析の対象とした。なお、受験者の母語はすべて日本語であり、3大学の受験者数の内訳は、109名（国立大学文系3学部・理系3学部）、282名（私立大学文系1学部）、および9名（私立大学文系1学部）である。多様な学力の3大学で実施することで、概ね大学生の平均的なレベルを推定できると判断した。

#### 5. 結果と考察

##### 5.1 バージョンによる違い

表1は2つのバージョンの記述統計、対応のないt検定結果、および効果量を示している。両バージョンの平均値に統計的に有意な差は検出されず、かつわずかな効果量であった（ $r=0.08$ ）ため、練習効果や疲労効果などの出題順序による影響は無視できる程度であったと考えられる。また、テスト結果の信頼性については、クロンバック  $\alpha = .90$ であり、内的一貫性の高い結果が得られた。

##### 5.2 語彙レベルに関する結果と考察

基礎統計量を表2に示す。素点の平均が104.6（満点125）であった。標準偏差は12.4であった。

表1 50,000語レベルの語彙サイズテスト各バージョンの統計量

	人数 (名)	平均 (点)	標準偏差 (点)	対応のない $t$ 検定結果	効果量
バージョンA	204	103.5	13.0	$t = 1.70, df = 398$ $p = 0.90$	$r = 0.08$
バージョンB	196	105.6	11.8		

表2 50,000語レベルの語彙サイズテストの基礎統計量

	平均	標準偏差	最高	最低
得点（素点）	104.6	12.4	124	30

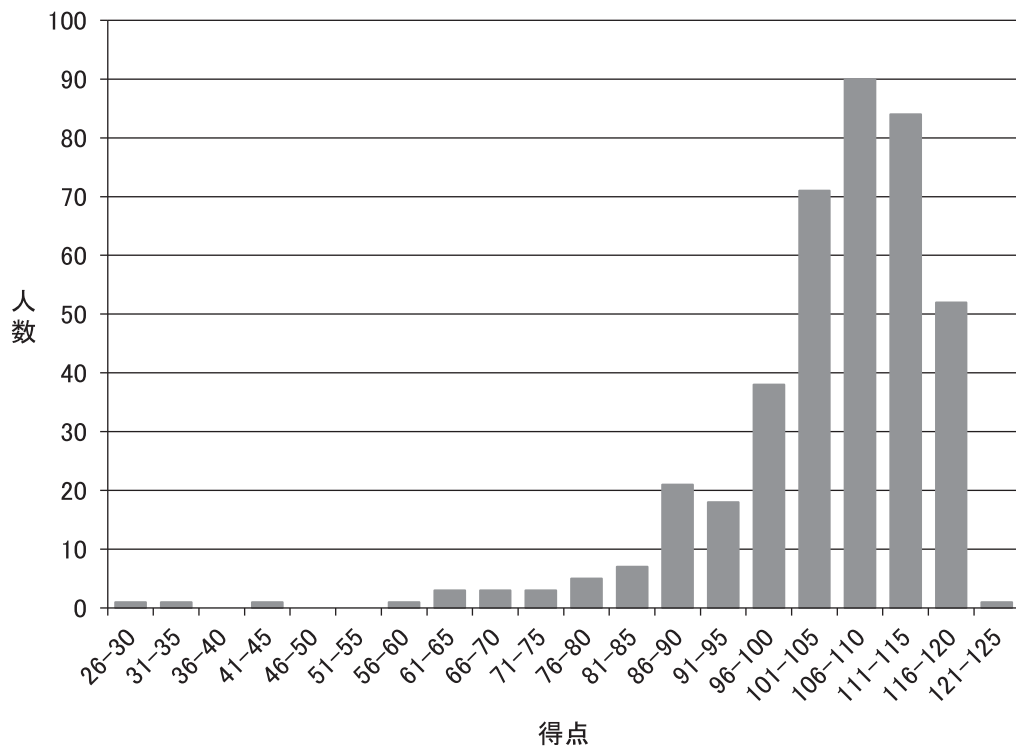


図3 50,000語レベルの語彙サイズテストの得点別人数分布

図3は得点を5点ごとに区切った人数の分布を表したグラフである。図2の30,000語レベルとほぼ同様に、分布は右側に偏り、高得点者が多かった。

母語話者の場合、このような多肢選択式テストの形式では30,000語レベル以上の低頻度の語彙であっても正答を選択することができる語が多く存在することがわかった。目標語を知っているかどうか測ることを重視したため、錯乱肢（不正解選択肢）に正解とは明らかに異なる意味記述を並べたことが一因であろう。語義の正確な理解を問う作成法ではなかったため、語義の一部を知っていれば正解できる項目があったということである。今後は、消去法などのテスト方略がどの程度影響を与えているかを検討する必要があるであろう。

### 5.3 推計既知語数

今回、実施したテストはデータベースの使用頻度順に400語に1語を抽出して、作成している。得点が1語（=1問）につき1点であることから、「テストの得点×400語」がそれぞれの学生の推計既知語数になる。表3は、テストの得点から推計する既知語数を4,000語ごと（得点では10点ごと）に区切って、それぞれの人数を示したものである。

前回のテストでは、30,000語レベルまでを測定したため、それ以上の語彙サイズについては明らかにできなかったが、50,000語レベルまでを測定した今回のテストでは、96.8%（387名）の学生の推計既知語数が30,000語レベルより上であることが確認できた。また、図3より74.5%（298名）の学生が推計既知語数40,001語以上（得点で101点以上）だった。



表3 得点による推計既知語数レベル別人数

得点 (点)	推計既知語数レベル (語)	人数 (名)
26-35	10,001-14,000	2
36-45	14,001-18,000	1
46-55	18,001-22,000	0
56-65	22,001-26,000	4
66-75	26,001-30,000	6
76-85	30,001-34,000	12
86-95	34,001-38,000	39
96-105	38,001-42,000	109
106-115	42,001-46,000	174
116-125	46,001-50,000	53
総計		400

この中で特に得点が低い3名は、推計既知語数18,000語以下である。田島ほか(2016)も、30,000語レベルの語彙サイズテストで既知語数18,000語以下とされる学生が458名中3名いたと報告されている。田島ほか(2016)でも指摘しているように、18,000語は日本語能力試験の最上位レベル(N1)で必要とされる語彙数に相当する。これらの学生は、非母語話者が論理的にやや複雑な文章や抽象度の高い文章を理解するのに必要とされる最低限度の語彙サイズしかないということになる。非母語話者は第一言語では第二言語の日本語より大きな語彙サイズを持っていると考えられる。一部の日本人学生は、第一言語が日本語である日本人学生の語彙サイズが非母語話者に必要とされるのと同等の語彙サイズしか持たないといえる。これらの学生には、大学での学びに多くの困難があると考えられる。

以上より、今回の語彙サイズテストで測った語彙の知識に限れば、概ね、日本人大学生の大学入学時の語彙サイズは40,000語前後であり、一方、一部に極端に語彙サイズが小さい学生がいるということがいえる。

#### 5.4 正答率が低かった語彙

今回のテストでは、どのような語彙の正答率が低かったのだろうか<sup>[1]</sup>。表4に正答率が低かった15語を示した。

30,001～40,000語レベルの語が2語、40,001～50,000語レベルの語が8語あり、低頻度の語彙の正答率が低い。一方で、「アナロジー」「ひがみ」「百花」「起債」は語頻度ランクではそれほど低くないものの、大学1年次生にとってはあまり触れる機会のない語であると考えられる。また、外来語が5語ある。和語・漢語において漢字で表記されていれば、その漢字の知識から意味が類推できる場合もあるが、外来語の場合は表記による手がかりがなく、触れる機会のない語については正答することは難しいためだと考えられる。

表4 正答率が低かった語

	項目	語頻度ランク	語種	正答率
1	アナロジー	19,557	外	22.5%
2	ミリタリズム	49,351	外	46.0%
3	パンタグラフ	36,550	外	46.5%
4	巻首	49,952	漢	47.5%
5	サテライト	33,351	外	49.0%
6	ひがみ	22,753	和	49.8%
7	完徳	49,151	漢	52.8%
8	架する	43,551	混	54.3%
9	業師	42,751	和	57.3%
10	多頭 (飼い)	40,350	漢	59.3%
11	講ずる	47,550	混	60.8%
12	ボランタリー	44,351	外	63.3%
13	百花 (繚乱)	29,351	漢	63.8%
14	起債	29,151	漢	64.0%
15	阿鼻 (叫喚)	26,151	外	66.5%

### 5.5 正答率が高かった語

では、どのような語の正答率が高かったのだろうか。表5に正答率97.5%以上の16語を示した。16語のうち、10,000語レベル以下のものが11語で、日頃から接する機会が多いことが考えられる。それ以外の「練乳」「怪獣」「パンケーキ」「童心」「アナゴ」の5語は、コーパスでは頻度の順位が低いが、大学1年次の学生の日常生活では接する機会の多い語、あるいは、漢字から意味が類推できる語であると考えられる。

## 6. おわりに

以上、日本人大学生に対し50,000語レベルの語彙サイズテストを実施し、その結果について考察した。

研究課題(1)のテスト開発については、英語の語彙サイズテスト(Nation & Beglar, 2007)同様の簡潔な意味記述を選ぶ型式で、部分的な語彙知識を測定することはできるが、30,000語レベルの語彙サイズテストと同様、得点の分布が右に偏り、受験者能力の弁別力の観点からは問題があると考えられる。これは、日本語の場合、漢字表記により語構成要素から語全体の意味を類推しやすい語が多く、正確な語義を知らなくても正解が選べるためだと考えられる。今後の研究では、テスト方略など、解答に至るプロセスをインタビュー等で確認し、受験者の能力にあった問題項目を増やしたり、語構成の知識や語の産出を問うテストを作成したりすることが必要となるであろう。



表5 正答率が高かった語

	項目	語頻度ランク	語種	正答率
1	養う	4,522	和	99.0%
	偉い	1,994	和	99.0%
3	練乳	33,951	漢	98.8%
4	青春	5,964	漢	98.5%
	ミス (する)	2,944	外	98.5%
6	怪獣	17,351	漢	98.3%
7	パンケーキ	37,150	外	98.0%
	(会議) 室	549	漢	98.0%
	学者	2,729	漢	98.0%
10	ムード	6,747	外	97.8%
	(家族) 連れ	5,146	和	97.8%
12	出番	9,151	混	97.8%
	ゆとり	5,306	和	97.8%
	やっと	1,354	和	97.8%
15	童心	28,352	漢	97.5%
	アナゴ	21,351	和	97.5%

研究課題（2）の日本語母語話者である日本人大学生の語彙サイズは、今回の語彙サイズテストで測った語彙の知識に限れば、概ね、日本人大学生の大学入学時の語彙サイズは40,000語前後と推計される。一方、一部に極端に語彙サイズが小さい学生が存在していることもわかった。今回は3校の学生に基づいて考察したが、平均的な学生像を描き出すには、より代表性の高いサンプルに基づいて調べる必要がある。しかし、少なくともこの語彙サイズテストを一部に存在する語彙力の低い学生を見つけ出すツールとして活用することは可能である。

研究課題（3）は、日本語母語話者である日本人大学生にとって、どのような語彙の意味が既知であり、どのような語彙の意味が未知であるか、である。正答率の点からは、一部の語彙については、データベースの使用頻度と関わらず、大学1年次生にとって、触れる機会が多いと思われる語彙は正答率が高く、触れる機会の少ないと思われる語彙は正答率が低い。また、漢字から意味が類推可能な語は正答率が高い傾向があった。このような点から、日本人大学生が高校までにどのような語彙を習得してきたか、大学での学びにはどのような語彙が必要で、それらの語彙をどのように教育していくのか検討していく必要がある。

今後の課題としては、非母語話者との比較、取り上げる語彙のジャンル、発達過程の探

索がある。今回は母語話者のみのデータだが、すでに非母語話者のデータの収集も始めている。非母語話者との比較は、留学生等の受け入れに当たって必要なカリキュラムを考えるうえで基礎となるデータを提供するはずである。テストの形式としては、語彙サイズテストでは30,000語レベルも50,000語レベルも過度に厳密な精度で正確な語義を問わないように、正解と類似するような意味記述の選択肢は入れないという方針を採ってきたが、これでは日本人母語話者には易しいと考えられる。選択肢をどのようなものにするのか検討が必要である。ただし、50,000語レベルより低頻度の語彙は専門用語やいわゆる死語が増え、語彙サイズを測るのには不適切であり、50,000語レベルより語彙の範囲を拡大した一般語彙テストの開発は適切ではないと判断される。このようなことから、専門分野に関わらず、大学で学ぶためにはどのような分野でも必要な学術系の語彙に関するテストの開発を行う方向で検討を進めている。松下（2011a）は、学術共通語彙は日本の大学で学ぶ留学生にとっても必要な語彙だと主張しており、学術共通語彙のテストの開発は留学生の日本語教育にも貢献できると考える。さらに、30,000語レベルと50,000語レベルの語彙サイズテストでは日本人の大学1年次生を測定の対象としているが、母語話者がどのように語彙を習得してきたか、小学校高学年ぐらいから発達段階に沿って、語彙サイズを測定することも必要であろう。

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。

## 注

[1] 正答率が低かった語のうち、選択肢に問題があるなど、統計学的にみて、出題した問題に原因があって正答率が低いと考えられる5語については削除した。

## 引用・参考文献

- Nation, P. & Beglar, D. (2007). A vocabulary size test, *The Language Teacher*, 31(7), pp. 9-13.
- Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests*. Copenhagen: Danmarks Paedagogiske Institut.
- Read, J. (1998). Measuring the vocabulary knowledge of second language learners, *RELC Journal*, 19 (2), pp. 12-25.
- 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、pp. 1-30.
- 押尾和美・秋元美晴・武田明子・阿部洋子・高梨美穂・柳澤好昭・岩元隆一・石毛順子（2008）新しい日本語能力試験のための語彙表作成にむけて、国際交流基金紀要、4、pp. 71-86.
- 阪本一郎（1955）読みと作文の心理、牧書房
- 田島ますみ・佐藤尚子・橋本美香・松下達彦・笹尾洋介（2016）日本人大学生の日本語語彙サイズ測定の試み、人間・自然論叢、41、中央学院大学、pp. 3-20.
- 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一、44、pp. 108-92.
- 日本リメディアル教育学会（2012） 国語リメディアル教育と大学生のための日本語教育、大学における学習支援への挑戦—リメディアル教育の現状と課題—、ナカニシヤ出版、pp. 144-156.
- 松下達彦（2011a） 日本語の学術共通語彙（アカデミック・ワード）の抽出と妥当性の検証、2011年日本語教育春季大会予稿集、pp. 244-249.
- 松下達彦（2012） 「日本語を読むための語彙サイズテスト」の開発、2012年日本語教育国際研究大会予稿集第1分冊、日本語教育学会、pp. 310.
- 松下達彦（2016） コーパス出現頻度から見た語彙シラバス、ニーズを踏まえた語彙シラバス（森篤嗣編）、くろしお出版、pp. 53-77.
- NTTコミュニケーション科学基礎研究所 語彙数推定テスト <http://www.kecl.ntt.co.jp/icl/lirg/resources/goitokusei/goi-test.html>（最終確認日：2016年11月6日）
- 日本語能力試験 認定の目安 <http://www.jlpt.jp/about/levelsummary.html>（最終確認日：2016年11月6日）
- 松下達彦（2011b） 日本語を読むための語彙データベースVer.1.1、<http://tatsuma2010.web.fc2.com/>（最終確認日：2016年11月6日）