

学部入学前日本語予備教育における
学術共通語彙知識の獲得
—国立大学に入学する韓国人学習者を例として—

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

¹⁾千葉大学国際教養学部 ²⁾東京大学大学院総合文化研究科 ³⁾京都大学国際高等教育院

⁴⁾中央学院大学法学部 ⁵⁾川崎医科大学医学部

Acquisition of Japanese Common Academic Words in the Japanese Language
Preparatory Course before the Entry of Japanese National University
Undergraduate Programs: A Case of Korean Students

SATO Naoko, MATSUSHITA Tatsuhiko, SASAO Yosuke,
TAJIMA Masumi, HASHIMOTO Mika

要旨

大学の講義や教科書には、専門用語のほかに、分野に関わらず学術的な文脈で用いられる学術共通語彙が高頻度で現れる。これらの学術共通語彙の理解は、日本の大学で学ぶ留学生には必須である。本稿では、日本語学術共通語彙テストを使用して、韓国で6か月の日本語予備教育を受けた韓国人学習者（日韓共同理工系学部留学生）（以下、日韓生）がどの程度学術共通語彙の知識を獲得し、理解しているかを、国立大学で学ぶ日本人学部生と比較し、報告する。調査の結果、日本人学部生は20,000語レベルまでの学術共通語彙はほぼ問題なく理解している。一方、日韓生は、韓国での予備教育終了時に、上位層は日本人学部生とほぼ同等に理解しているのに対して、下位層は日本人学部生の最低点との差が大きかった。学部入学後に日本人学部生と同様に授業を理解していくためには、日本での6か月の予備教育期間中に学術共通語彙を知り、理解できるように教育を行う必要がある。

キーワード：

学術共通語彙、日本語学術共通語彙テスト、語彙頻度、高等教育、日本語予備教育

1. はじめに

1.1 背景

日本の大学の学部で学ぶ留学生は、入学と同時に日本語母語話者である学部生（以下、日本人学部生）と同様に日本語による講義を聴き、日本語で書かれた教科書を読むことになる。その講義や教科書には、専門語彙のほか、分野にかかわらず学術的な文脈で高い頻度で用いられる学術共通語彙が現れる。これらの学術共通語彙の理解は、日本の大学で学ぶ留学生には必須である。本稿では、韓国で6か月の日本語予備教育を受けた韓国人日本語学習者（日韓共同理工系学部留学生）（以下、日韓生）が、読解のための学術共通語彙をどの程度知っていて、理解しているか、その語彙知識の獲得状況について国立大学の学部で学ぶ日本人学生と比較し、報告する。

今回、調査を実施した日韓生は、日本と韓国の政府間で実施されている日韓共同理工系学部留学生事業によって、日本の国立大学の理工系学部留学する学生である。この事業は2000年度から開始され、毎年100名の新規学生が日本へ留学している。今回、調査を行ったのは、2016年10月に渡日した第2次第7期生である。

この事業に参加する韓国人学生は学部入学前に1年間の日本語予備教育を受ける。3月から8月まで韓国で6か月間の日本語予備教育を受け、9月下旬から10月上旬に渡日し、10月から3月まで、翌年4月から入学する国立大学で、さらに6か月の日本語予備教育を受ける。その後、学部に入學し、日本人学部生とともに日本語による授業を受ける学生である。2015年10月に渡日した第2次第6期生100名は全国の20の国立大学に配置されている¹⁾。

1.2 学術共通語彙とは

調査の対象とした学術共通語彙とは、Coxhead（2000）によって提唱された英語の Academic Word List²⁾と同様の目的で作成された日本語の語彙リストで、一般的なテキストと比べて学術的なテキストで、より高い使用率を占める語彙のことである。具体的には、対数尤度比（Log-likelihood）³⁾を用いた特徴語抽出により、人文、社会、理工、生物・医学の4ジャンル中3ジャンル以上で特徴語として選ばれた語を指す^{4,5)}。必然的に文理双方にまたがって使用されることの多い語彙であり、大学での学びにおける語彙の中で一般語彙と専門語彙の中間に位置するものである。

範囲（例：「占める」「特殊」）、関係（例：「優れる」「属する」）、段階（例：「当初」「現状」）、量的変化（例：「減少」「強化」）、論述の展開（例：「取り上げる」「まとめる」）などの、抽象性の高い、論理操作に不可欠な語が多い。例えば、学術共通語彙レベルI（559語）のテキストカバー率（延べ語数に占める割合）は、一般のテキストでは3%前後、会話では1%未満であるのに対して、学術的なテキストでは約11%である⁵⁾。レベルIは初級語彙（旧日本語能力試験の3級および4級の語彙約1,300語）を含まず、上位5,000位以内の中級

レベルの語彙であり、人文、社会、理工、生物・医学の4ジャンルすべてで特徴語とされる語彙である。学生が初級を終えた後に学ぶべき語彙であり、性質（目的）、レベル、語数のいずれにおいても英語のAcademic Word Listと比較可能であると思われる。Academic Word Listのテキストカバー率が学術的なテキストで約10.0%とされる²⁾のに対し、学術共通語彙レベルIは約11%なので、匹敵する語彙群であると言える（本調査のテストではレベルIだけでなく、より低頻度の、超上級レベル（2万語レベル）の学術共通語彙まで対象としている）。

特定の分野で多く使われる専門語彙は、その専門性から、授業の際に教員は、板書する、その意味を説明する、など明示的に扱うが、学術共通語彙は習得されていることを前提として授業が進められるのが一般的である。このため、学術共通語彙が理解できないと、学術的な話題が理解できず、大学での学びに支障をきたしてしまう可能性が高い。英語のAcademic Word Listの語彙の能力が一般的な学術能力をある程度予測するという研究もある⁶⁾。

本調査では、国立大学の理工系学部に入学者、日本人学部生と同じ授業を受ける日韓生が、韓国での6か月の日本語予備教育を修了した段階で、どのような学術共通語彙を、どの程度理解しているのかを明らかにすることを目的とする。調査には、学術共通語彙知識測定のために開発された「日本語学術共通語彙テスト Version 1」⁷⁾を使用した。

2. 方法

2.1 日本語学術共通語彙テスト

「日本語学術共通語彙テストVersion 1」を用いた。

ターゲット語は「日本語学術共通語彙リスト」⁸⁾から選んだ75語である。「日本語学術共通語彙リスト」の中の語を「日本語を読むためのデータベース」⁹⁾での頻度順位に従い、上位20,000語の中から250語に1語の割合で、かつ、可能な限り等間隔で選定した。また、「粗」のような接辞も一般的な単語と同様にターゲット語とした。これは、日本語の接辞には「会議室」の「室」のように実質的な意味を持つものが多く、辞書の見出しにもなっており、日本語の語彙研究では接辞は語と同等の扱いを受けることが多いためである。ただし、初級語彙（旧日本語能力試験の3級と4級の語彙）は除外した。旧4級の語彙は学術共通語彙からそもそも除外されており、3級語彙も中級以上の学習者には弁別力がないと判断したためである。

問題の形式は、ターゲット語とそれを含む文（語義のヒントを与えないような文）を示したうえで、3選択肢の中からターゲット語の言い換えとなる意味記述を正答として選ぶものである。すなわち、書かれた語の理解を試すテストである。ターゲット語が多義である場合は、NINJAL-LWP for TWC¹⁰⁾などで調べて、学術共通語彙として最も使用されている語義を用いた。

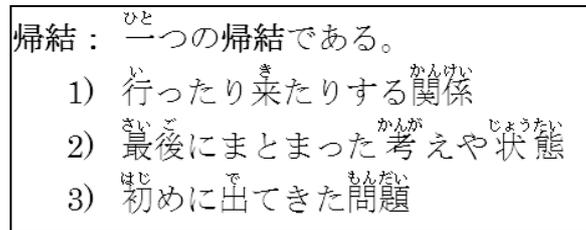


図1 日本語学術共通語彙テストの問題の形式例

問題の形式例を図1に示した。それぞれの問題の最初にターゲット語を、その横にターゲット語の品詞が分かる程度の例文を記した。これは例文から意味が判断できないようにするためである。そして、その下に選択肢を配置した。

2.2 対象

2016年8月に韓国での予備教育を修了した日韓生100名のうち、日本語を母語と回答した1名を除いた99名（すべて理系生、推定平均年齢^{注1}：19歳4か月、標準偏差8.8か月）を分析対象とした。

また、日韓生と一緒に学ぶ日本人学部生がどの程度学術共通語彙知識を獲得しているかを知るために、日韓生が配置された実績がある国立大学A大学、B大学の2校の文系および理系学部^{注2}に所属する日本人学部生1年生79名、2年生2名の計81名（文系32名、理系49名、推定平均年齢：18歳11か月、標準偏差8.6か月）も対象とした。A大学は10学部、学部生10,000人、B大学は4学部、学部生4,000人規模の大学である。

2.3 実施方法

2段組みとして1ページに左右7問ずつ、各ページ14問をA4サイズの紙に両面印刷した。計75問について、表紙を含めた7ページの冊子形式にまとめ配付した。解答の番号をマークシートに記入させ、マークシートの読み取り機で解答を処理した。

問題冊子にはテスト時間が40分と記されているが、実際には受験者全員が30分以内で終了した。

日本人学部生に対しては、2016年4月から10月にかけてそれぞれの所属大学においてテストが実施された。日韓生に対しては、2016年8月の来日直前の時期に韓国の大学で実施された。

3. 結果

3.1 テストの信頼性

当該テストの信頼性係数（クロンバックの α ）は $\alpha = .93$ であり、内的一貫性の高い結果が得られた。

3.2 日本人学部生の結果

1問1点とした75点満点のテストで、国立大学の日本人学部生は、平均点68.7、最高点74、最低点60、標準偏差3.1だった。国立大学の日本人学部生は高得点者が多く、頻度順位20,000語レベルまでの学術共通語彙であれば、ほとんど問題なく理解していると言える(表1)。

また、文系の学生と理系の学生に統計的有意差は検出されなかった ($t(79)=1.15$, $p=0.25$, $n.s.$, $d=0.260$)。専門分野に限らず学術的テキストに頻出する語彙という性質上、文系・理系の違いに関わらず、国立大学の母語話者である日本人学生は学部1年生段階で、学術共通語彙はほぼ理解できていることが示唆された。

3.3 日韓生の結果

75点満点のテストで、日韓生は、平均点53.9、最高点72、最低点23、標準偏差10.9だった(表1)。日本人学部生との間には有意な差があり、効果量(Cohen's d) も大きい ($t(117.0)=13.0$, $p<.001$, $d=1.84$)。

韓国の予備教育では、日韓生は日本語能力別に6つの班に分けられる。第2次第7期生(2016年度)も韓国での予備教育開始前に行われた試験の結果で、学習期間と日本語能力によって班分けが行われた。第1班～第3班が日本語学習期間6か月以下、第4班、第5班が学習期間6か月以上、第6班が学習期間1年以上、または、日本語能力試験N1、N2に合格していることを基準に分けられている¹¹⁾。ただし、予備教育期間中には定期試験などの結果に基づいて、所属する班が変更になる場合がある。

日韓生の得点を、韓国における予備教育の1班から6班の班別に、表2に示した。また、日本人学部生と得点を比較するため、図2に分布を示した。

日本人学部生の最低点(60点)に届かない日韓生は66名(66.7%)であった。日韓生の

表1 日本人学部生と日韓生の学術共通語彙テストの結果：基礎統計量(75点満点)

	平均	最高	最低	標準偏差
日本人学部生(81名)	68.7	74	60	3.1
文系生(32名)	69.2	74	60	3.5
理系生(49名)	68.4	73	60	2.8
日韓生(99名)	53.9	72	23	10.9

表2 学術共通語彙テストにおける日韓生の班別得点別人数分布および平均・標準偏差 (99名、75点満点)

得点	0-30	31-40	41-50	51-60	61-70	71-75	計	平均	標準偏差
1班	4	4	6	2			16	38.9	8.8
2班		4	5	7			16	47.8	7.6
3班		3	2	10	2		17	51.9	7.6
4班				10	7		17	59.7	4.1
5班			1	9	7		17	58.9	6.1
6班				3	12	1	16	65.6	4.5
全体	4	11	14	41	28	1	99	53.9	10.9

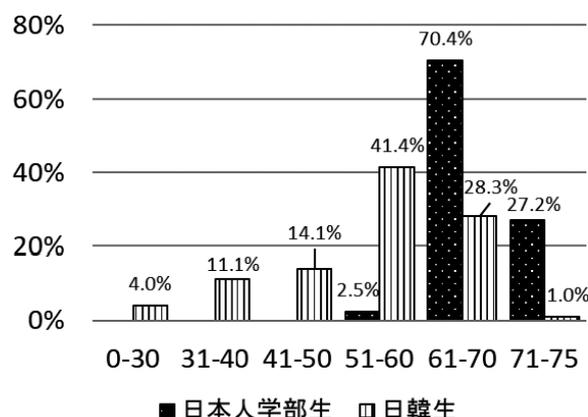


図2 韓国における予備教育修了時の日韓生と日本人学部生の学術共通語彙テストの得点分布

上位層は日本人学部生とほぼ同等に学術共通語彙を理解している一方、3分の2の日韓生は日本人学部生の最低点に達していなかった。特に40点以下の下位15名（15%）は日本人学部生の最低点との差が20点以上開いた。20点の点数の差は、ターゲット語の抽出方法、すなわち、頻度順位で250語につき1語を選ぶという方法から換算すれば、理解語彙数で一般語彙を含めて約5,000語の差に相当する。

本研究グループでは、今回の調査で使用した「日本語学術共通語彙テストVersion. 1」を改良した「日本語学術共通語彙テストVersion. 2」を使用して、日本人大学生と小中学生に実施した^{註2}。その調査結果によれば、大学Bの平均得点（70.6点）と中学生（1年生～3年生）の平均得点（50.5点）の差は20.1点であった¹²⁾。日韓生の下位15名（15%）と日本人学部生の最低点との差20点は、これと同等と考えられる。

表3 日韓生の正答率が90%以上だった語

ターゲット語	頻度順位	正答率 (%)
ビジュアル	14,372	99.0
オープン	3,880	98.0
能動	9,871	97.0
理解	1,375	97.0
エンタープライズ	17,624	96.0
吸収	3,126	96.0
視点	2,375	96.0
過程	1,875	94.9
死活	19,136	93.9
小数	18,110	91.9
溶解	12,887	91.9

3.4 項目別正答率

3.4.1 正答率の高かった語

日韓生の正答率が90%以上だった語は11語だった(表3)。日本人学部生は、この11語はすべて正答率90%以上だった。11語のうち5語は頻度順位^{注3}が5,000語より上位の、いわゆる中級レベルの語彙であった。

一方、韓国語でも使われる外来語である「ビジュアル」「エンタープライズ」は頻度順位が低いものの、正答率が高かった。また、「死活」(韓国語で사활sa hwal)、「小数」(韓国語で소수so su)、「溶解」(韓国語で용해yong hae)も頻度順位は低いが、正答率が高かった。

しかしながら、下位15名(15%)の学生の解答を分析してみると、「溶解」(平均正答率91.9%)の正答率は60.0%、「小数」の正答率(平均正答率91.9%)は73.3%しかなかった。

3.4.2 正答率の低かった語

半数以上の日韓生がわからなかった、正答率が50%未満だった12語を表4に示す。

12語のうち、正答率が40%未満だった8語は20,000語レベルで抽出したテスト対象語の中では頻度順位が11,000位以下の、比較的低いものであり、日本語教育では接する機会が少なく、語彙知識を得るのが難しい語彙だと言えよう。

しかし、正答率が40%以上50%未満の語では、頻度順位が6,000位よりも上位の、出現頻度が高い語彙が見られる。テストにおける全ターゲット語の頻度と正答率の相関を見てみると、日本人学部生が $p=.318$ であるのに対して、日韓生は $p=.144$ だった^{注4}。日韓生に

表4 日韓生の正答率が50%未満だった語

ターゲット語	頻度順位	正答率 (%)
露地	15,885	16.2
踏襲	12,120	24.2
テーゼ	15,636	26.3
粗一	13,887	31.3
間引き	13,114	32.3
重ね合わせる	11,391	33.3
所与	19,833	38.4
作製	18,600	39.4
貢献	2,624	40.4
編著	18,903	40.4
かなう	4,299	44.4
懸念	5,377	49.5

表5 正答率に50ポイント以上差があった語

ターゲット語	頻度順位	日本人学部生 正答率 (%)	日韓生 正答率 (%)	日本人学部生と 日韓生の差
間引き	13,114	98.8	32.3	66.4
粗一	13,887	92.6	31.3	61.3
貢献	2,646	98.8	40.4	58.4
かなう	10,623	97.5	44.4	53.1
テーゼ	15,636	79.0	26.3	52.7
懸念	5,377	100.0	49.5	50.5

において、頻度と正答率の相関は非常に弱かった。

3.4.3 日韓生と日本人学部生の正答率に差がある語

日韓生と日本人学部生の間で正答率に差があった語はどのようなものだろうか。日韓生の正答率が50%未満だった語の中で、日本人学部生の正答率の差が50ポイント以上あった6語を表5に示す。これらの語は、日韓生の正答率が非常に低く、日本人学部生の半数以下の日韓生しか理解できていないと言える。

4. 考察

日韓生の一部は留学先の日本人学生と同レベルの学術共通語彙の知識を持ち、理解しているが、15%程度の学生は日本人中学生と変わらない程度の学術共通語彙知識しか持たずに日本へ留学することになる。このような学生が、学部入学後に日本人学部生と同様に授業を理解していくためには、渡日後6か月行われる予備教育期間中に数多くの学術共通語彙が理解できるようにしていく必要がある。学術共通語彙は日常会話の語彙や文芸書の語彙と重なりが少ない。例えば、学術共通語彙は約75%が漢語なのに対し、文芸語彙は和語が約71%⁵⁾、学術共通語彙の100万語当たりの平均頻度(テキストカバー効率:TCE)は、学術テキストでは144.2であるのに対し、非学術テキストでは30.5しかない⁵⁾。したがって、学術的な語彙を学ぶには学術テキストを使うのが最も効率がよい。漢字学習の負担は相当なものであるため、学術共通語彙を学習するための効率的な教材・素材が必要であろう。

頻度との関係も興味深い。語彙頻度と語彙習得レベルには相関があるのが一般的であり^{13,14)}、頻度順位の高い一部の語彙は日本語の授業で学習する機会が多いと考えられるが、学術共通語彙に限定すると、日本人大学生の正答率では頻度と0.3以上の順位相関があるのに対し、日韓生の正答率との順位相関は0.14程度で非常に弱かった。学術共通語彙は、日本語に触れる機会が相対的に少ない外国語環境では、相関が弱くなるという可能性が見られ、また、第一言語の知識や英語の知識で理解できる語彙が非学術的な語彙より多いということも、頻度との相関を弱めている原因であると推測できる。いずれにしても、渡日前の韓国の予備教育終了時では、高頻度語であっても、理解できていないことがあり、頻度と語彙知識の獲得・理解との関係において、日本人学部生と日韓生には違いがあるということに留意すべきであろう。

「ビジュアル」「エンタープライズ」は韓国語でも使われる外来語であることから正答が選びやすかったと考えられ、また、「死活」(韓国語で사활sa hwal)、「小数」(韓国語で소수so su)、「溶解」(韓国語で용해yong hae)は漢字が初級・中級レベルであることから、正確な漢字の知識がなくても、第一言語である韓国語の語彙知識を利用して意味が類推できる可能性が高い語彙であったと考えられる。

しかしながら、下位15名(15%)の学生の「溶解」「小数」の正答率が全体の平均より大きく劣っていたことから、日本語のレベルが上がってこないで第一言語の知識が十分活用できないことがあると言える。第一言語を利用することによって、語彙知識の獲得が進むと考えられる場合には、対応する第一言語の単語を積極的に教えたほうがよいであろう。

学術共通語彙には、その意味を知らなければ学術的な活動全般に支障をきたすものと相対的に影響が限定的なものがある。例えば、表5では「貢献」「懸念」など出現頻度が高いものが前者に、「間引き」「テーゼ」など出現頻度が低いものが後者に当たる。出現頻度が高く、学術的な活動全般に大きく関わる語彙が学べるよう日本語教育の中で積極的に取り上げる必要がある。

5. 終わりに

以上、日本人学部生と比較しつつ、日韓生が、どの程度、学術共通語彙知識を獲得しているかについて考察した。

国立大学に在籍する日本人学部生は20,000語レベルの学術共通語彙はほとんど問題なく理解しているのに対し、韓国における予備教育修了時の日韓生は、上位層は日本人学部生とほぼ同等に学術共通語彙を理解しているものの、下位層は日本人学部生の最低点との差が大きかった。学部入学後に日本人学部生と同様に授業を理解していくためには、渡日後6か月行われる予備教育期間中にこれらの学術共通語彙を学び、理解できるようになる必要がある。

日韓生の学術共通語彙の理解状況と語彙頻度との相関は非常に弱く、高頻度語彙でも理解できていない語彙があること、また、日韓生は韓国語でも使われる外来語や漢語など、第一言語の語彙知識が活かされると考えられる語彙は、頻度順位にかかわらず正答率が高いことがわかった。一方で、学術共通語彙が非学術的な語彙と異なった文脈で使用されることから、正答率が低い語彙は少なくとも海外においては教室内においても、教室外においても接する機会が多くない語彙だと考えられる。正答率が低い語彙をどのように教育していくか考える必要がある。

また、下位層の学生は、第一言語である韓国語に同じ漢字語があっても、類推ができない傾向が見られた。第一言語の活用を考慮に入れた教材開発を行うのも一つの方法と考えられる。

今後の課題は、ここで示された得点が、大学での学びの中で、実際にどのような影響を及ぼすかを明らかにすることである。文章作成力や読解力など、ほかの日本語能力を測定するテストの結果や教科の成績と照合するなどして、どの程度の得点であれば、大学での学びが円滑に行われるのかを明らかにしていきたい。

謝辞

本研究を行うにあたり、金蘭美先生（横浜国立大学）より韓国語の語彙についてご教示をいただきました。御礼申し上げます。

付記

本研究はJSPS科研費JP15K02631の助成を受けた。

注

注1 調査対象者が記入した満年齢の平均値に6か月（0.5年）を足したものである。

注2 Version.1とVersion.2は語彙レベルの統制の方法は同じであるが、Version.2では小中学生にも受けられるように漢字レベルも統制した。Version.1を実施した、同じ2大学の1年生に、翌年、

Version.2を実施したところ、ほぼ同じ平均点であったことから、大学生の場合は類似したレベルのテストであると考えられる。

注3 「日本語を読むための語彙データベース」⁹⁾の留学生用語彙ランキングを使用した。

注4 頻度は正規分布を仮定できないためピアソンの相関係数 r ではなく、スピアマンの順位相関 ρ (ロー)を用いた。

参考文献

- 1) 日韓共同理工系学部留学生事業協議会：日韓共同理工系学部留学生（～第2次第6期生）大学配置状況（2015）
- 2) Coxhead, A.: A new academic word list, *TESOL Quarterly*, 34(2), pp.213-238 (2000).
- 3) Dunning, T.: Accurate methods for the statistics of surprise and coincidence, *Computational Linguistics*, 19(1), pp. 61-74 (1993)
- 4) Matsushita, T.: In what order should learners learn Japanese vocabulary? A corpus-based approach. PhD thesis, Victoria University of Wellington (2012)
- 5) 松下達彦：コーパス出現頻度から見た語彙シラバス，森篤嗣編，ニーズを踏まえた語彙シラバス，くろしお出版，pp. 53-77 (2016)
- 6) Morris, L., & Cobb, T.: Vocabulary profiles as predictors of the academic performance of teaching English as a second language trainees, *System*, 32(1), pp. 75-87 (2004)
- 7) 田島ますみ・佐藤尚子・橋本美香・松下達彦・笹尾洋介：日本語学術共通語彙テストの開発，人間・自然論叢，第45号，中央学院大学，pp. 19-31 (2018)
- 8) 松下達彦：日本語学術共通語彙リストVer.1.01 (2011)，URL：<http://www17408ui.sakura.ne.jp/tatsum/list.html>（最終確認日：2019年8月6日）
- 9) 松下達彦：日本語を読むための語彙データベース（VDRJ）Ver. 1.0（教師用）(2011)，URL：<http://www17408ui.sakura.ne.jp/tatsum/database.html#vdrj>（最終確認日：2019年8月6日）
- 10) 筑波大学・国立国語研究所・Lago言語研究所：NINJAL-LWP for TWC，URL：<http://nlt.tsukuba.lagoinst.info>（最終確認日：2019年8月6日）
- 11) 趙顯龍：第2次7期生の予備教育の内容，2016年第2次7期生韓国予備教育課程の現状報告資料2,2016年度日韓共同理工系学部留学生事業協議会（2016）
- 12) 田島ますみ・佐藤尚子・松下達彦・笹尾洋介・橋本美香：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—，日本リメディアル教育学会第13回全国大会発表予稿集，日本リメディアル教育学会，pp. 140-141 (2017)
- 13) Read, J.: Measuring the Vocabulary Knowledge of Second Language Learners, *RELC Journal*, 19(2), pp. 12-25 (1988)
- 14) 篠崎大司・松下達彦・川村よし子：上級日本語文字・語彙eラーニング教材の開発，日本言語文藝研究，第16号，pp. 90-108 (2016)