

語彙調査に基づくタスクの分類 (2)

一タスク別・レベル別・品詞別使用語彙の傾向から一

橋本 直幸

1. はじめに

金澤 (編) (2014) に収録されている「YNU 書き言葉コーパス」は、日本人大学生 30 名、日本語学習者 60 名に対し、それぞれ 12 種類の作文タスクを課すことによって得られた合計 1080 サンプルを収録する約 25 万語のコーパスである。コーパスの大規模化が進む現在にあっては、決して規模が大きいといえるものではない。しかし、YNU 書き言葉コーパスの最大の特長は、一人の被調査者に対し、性格の異なる 12 のタスクを課している点にある。話題、機能、相手が異なる様々なタスクを課すことによって、学習者のもつ言語能力のさまざまな側面を捉えることが可能となる。YNU 書き言葉コーパスを研究に利用する場合、この特徴を最大限に生かすためにも、それぞれのタスクの特性をしっかりと把握しておくことが、非常に大切なことである。

また、金澤 (編) (2014) が書名に「日本語教育のための」と冠していることからわかるように、実際に学習者に課すタスク集としても使用できるものであることが望ましい。そうであるならなおさら、それぞれのタスクの性格をきちんと把握し、「このタスクを行うことによって、どのような能力が身につくか」、また「タスクを達成するためにどのような能力が必要か」を把握しておくことが必要である。

実際に、金澤 (編) (2014) では、「第Ⅱ部 評価・分析編」として、一つ一つのタスクについて、母語話者の例とレベル別¹の学習者の例を示しながら、評価の観点やレ

¹ YNU 書き言葉コーパスでは、あらかじめ決めた評価の観点に基づきタスクごとに作文を評価し、12 タスクの総合結果に基づいて被調査者を上位群、中位群、下位群に分けている。YNU 書き言葉コーパスでは、これを「グループ」と呼んでいるが、本稿では便宜上「レベル」と呼ぶ。ただし、言うまでもなく、日本語教育で一般に呼ばれる「初級」「中級」「上級」といった区分や、OPI などで判定される「Novice (初級)」、「Intermediate (中級)」、「Advanced (上級)」といった「レベル」とは異なる。

ベルによる違い、母語による違いなど多方面からタスクの特性を明らかにしている。また、「第Ⅲ部 論文編」に掲載されている橋本 (2014) では、各タスクで使用されている語彙に焦点を当て、「語彙多様性」「個人差」という観点から、それぞれのタスクを分析し、その特性を明らかにしている。

本稿は、この橋本 (2014) に続くものとして、引き続き語彙に焦点を当てたタスクの分析を行うものである。具体的には各タスクで使用されている品詞に注目し、それぞれのタスクでどの品詞がタスク達成のための重要な要素になっているかを明らかにしていきたい。

2. 研究の背景

本研究は、YNU 書き言葉コーパスを使って、それぞれのレベルの学習者における語彙の実態を調査し、タスクごとの特徴を明らかにしようとするものである。本稿での「語彙」とは、いわゆる「実質語 (content word)」と呼ばれるものを指す。

これまでの第二言語習得研究の関心の中心がいわゆる「文法」を対象としたものであり、学習者の「語彙」の使用実態や習得過程を明らかにしようとした研究が立ち遅れていたことは、長友 (1999)、佐々木 (2010) などでも指摘されている通りである。語彙習得研究と呼ばれるものであっても、山内 (2004) が指摘したように、語彙学習ストラテジーの研究や特定の語群に絞った習得研究が多く、「語彙習得研究のプロトタイプとでも言うべき「語の集まり」の習得研究がほとんど行われていない」(p.141) という状況であった。

山内 (2004) はこのような当時の現状を踏まえ、「形態素解析」と「N グラム統計」というコンピュータを利用した新たな語彙習得研究の方法を提案している。文法に比べてその構成要素が膨大である語彙を対象とした研究において、やはりコンピュータの使用が欠かせないものであるということを改めて認識させることとなった研究である。そして、この指摘から 10 年以上が経過し、大規模な学習者コーパスもさまざま整備されてきた。とは言え、文法と違いその内容に大きく左右される語彙を研究対象とする場合、やはりサンプル自体の内容も無秩序な自由作文ではなく、統制がとれたものでなければならない。もちろん習得研究を視野に入れるなら母語別・レベル別も欠かせない要素である。このような点において、「母語」「レベル」「タスク」の統制がとれた YNU 書き言葉コーパスは、語彙習得研究においてもより積極的に活用されるべきであると考えられる。

本稿はこのような問題意識の下、タスクの分析に加え、YNU 書き言葉コーパスを用いた語彙習得研究の可能性も探っていきたいと考えている。

3. 調査の概要と手順

3.1 調査の概要

この調査では、それぞれのタスクにおいて、タスク達成のためにどのような語が重要な役割を果たしているかを調べることを目的とする。ただし、具体的な語について見る前に、まずは品詞に注目し、名詞、動詞、形容詞のうち、どの品詞がタスク達成のための鍵となっているかを明らかにする。そのため、ここではタスク別、レベル別、品詞別の語彙の使用実態を観察する。例えば、タスクの中には日本語学習者の下位群、中位群、上位群いずれも同じような語を使用しておりレベル間に差が見られないものがある。この場合、語彙の使用状況が変わっていないにもかかわらず、最終的なレベル判定が異なっているということから、語彙の使用がレベルの決定に有効に働いていないと考えることができる。一方で、レベル間に使用語彙の顕著な違いが見られるものもある。この場合、語彙がレベル決定の重要な要素となっている可能性が高い。

この調査では、タスクごとに、名詞、動詞、形容詞の使用実態を母語話者のそれと比べることで、この点を明らかにしていきたい。

3.2 調査の手順

調査では、タスクごとに日本語母語話者 30 名分の使用語彙をモデルとして、上位群、中位群、下位群の学習者の作文がそれぞれどれくらい母語話者の使用傾向に近いかを数値により確認する。日本語母語話者の日本語を学習者の目指すべきモデルとするかどうかということについては多くの議論があるところだが、下位群から中位群、中位群から上位群へとレベルが上がっていくにつれ、使用語彙の傾向が似て来ることは多くのタスクで確認できることなので、とりあえず本稿では日本語母語話者の日本語をモデルとし、分析を行うこととする。具体的にはまず、レベルごと、品詞ごとに、累積使用語数の上位 70% までを占める語にしばったうえで、仮に学習者に日本人と同様のパフォーマンスを行うことを求めた場合、学習者の使用している語でそれがどの程度可能かを計算する。これを本稿では「カバー率」と呼ぶことにする。ここで上位 70% 語にしばったのは、それ以下の語になると、どのタスクも使用頻度が 1 や 2 となり、必ずしもタスク達成に必要な語であるとは言えないと考えられるからである（なお、タスクによっては 70% で既に使用頻度が 1 や 2 になる場合もあり、そのようなタスクについては対象とする語の範囲を 50% 程度まで引き上げている）。

例えば、タスク 4（奨学金増額の依頼メール）を例にとると、日本語母語話者、日本語学習者の上位 70% までの使用語彙は次頁の表 1 のようになる。

日本語母語話者がタスク 4 において使用した動詞の上位 70% が、表 1 に挙げた 263

語である。一方、日本語学習者が使用した語で同様のパフォーマンスを行おうとすると、カバーできる語は、日本語母語話者が使用した 260 語のうち、上位群の学習者で 220 語、中位群の学習者で 223 語、下位群の学習者で 188 語である。それぞれ、日本語母語話者の使用語彙の 83.7%、77.6%、71.5% にあたる。以上の方法により、全 12 タスクをレベル別、品詞別に調査を行い、考察を加えた。

表 1 タスク 4 における上位 70% の使用語彙 ※括弧内の数字は延べ語数

日本語母語話者		言う (45)、増やす (27)、就く (24)、出る (23)、思う (21)、申す (20)、 関する (9)、行う (8)、於く (以下、7)、励む、申し上げる、差し上げる (6)、 挙がる (以下、5)、上がる、送る、聞く、存ずる、務める、抛る、受ける (以下、4)、考える、費やす、開く、増える、向ける (延べ 263 語)
日本語学習者	上位群	思う (27)、言う (24)、就く (20)、申す (14)、増やす (13)、行う (11)、関する (10)、申し上げる (10)、務める (9)、出る (8)、増す (8)、送る (以下、7)、考える、増える、抛る、存ずる (6)、伝える (6)、対する (5)、分かる (5)、減る (4)、答える (以下、3)、費やす、使う、取り上げる、開く、纏める、読む (延べ 229 語)
	中位群	思う (30)、言う (19)、就く (18)、増やす (17)、申す (14)、考える (12)、出る (10)、増える (9)、与える (以下、8)、関する、減る、抛る (6)、行う (以下、5)、従う、待つ、対する (以下、4)、伝える、務める、纏める、集まる (以下、3)、集める、受け入れる、送る、頑張る、立つ、取る、増す、向ける、持つ、分かる (延べ 223 語)
	下位群	言う (23)、就く (19)、思う (18)、申す (16)、出る (14)、増える (10)、考える (7)、増やす (7)、頑張る (6)、足りる (5)、待つ (5)、関する (以下、4)、伝える、申し上げる、住む (以下、3)、取る、増す (延べ 151 語)

3.3 調査データ・分析ツールについて

調査データは YNU 書き言葉コーパスの「修正データ」を用いた。ただし、タスク 12 の母語話者データについては、「小学校新聞で昔話を紹介する」というタスクの性質上、小学生に配慮し平仮名を多用しているものが多く、この場合、形態素解析が正確に行えないデータが多かったことから、平仮名で書かれている部分を漢字に書き直したものを使用している。形態素解析には、UniDic1.3.12 (mecab 版) を用いた。

調査対象は実質語を対象とすることから、品詞は、名詞、動詞、形容詞とした。UniDic のタグは以下の通りである。

表2 本稿での品詞の表示と UniDic での品詞タグ

本稿での品詞の表示	UniDic1.3.12 での品詞タグ
名詞 (一般)	名詞 - 普通 - 一般
名詞 (サ変可能)	名詞 - 普通 - サ変可能
動詞	動詞 - 一般
形容詞	形状詞 ² - 一般 / 形容詞 - 一般

名詞の「サ変可能」というのは、漢語サ変動詞の語幹部分にもなりうる名詞であり、名詞として使用されているものもあるが、漢語サ変動詞として使用されているものも相当数あることから、名詞 (一般) とは分けて集計した。

なお、集計にあたって、タスク 1、2 などのメールタスクのデータに含まれる「件名」「本文」などの語や、名乗り部分における学部名 (「教育」「人間」「科学」) などは、母語話者と学習者で差が出るようなものではないため、削除して集計した。また、タスク 4 は、「学長に奨学金の増額を依頼する」というタスクであるが、母語話者は「学生自治会の代表」として、学習者は「留学生会の代表」として依頼を行うようタスクを課しており、「自治」「留学」などの語も、どちらかにしか出て来ない語で、もともと一致が望めない語であることから、これらについても集計の段階で削除している。

4. 調査結果 - 特徴的な 2 タスクについて

この節では、全 12 タスクについて見る前に、特徴的な 2 つのタスク「タスク 5」と「タスク 12」について説明する。

4.1 タスク 5

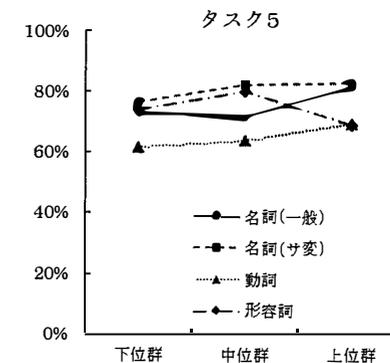
まず、タスク 5 の結果について述べる。タスク 5 の課題は「交通事故に遭い、就職活動や卒業論文に不安を感じている後輩に対し、励ましの手紙を送る」というものである³。以下の表 3 が、使用頻度の高い上位 70% を対象に、下位群、中位群、上位群の使用語彙で日本語母語話者の使用語彙をカバーできる割合である。

² 「形状詞一般」は、「静か」「穏やか」など、いわゆる形容動詞の語幹部分である。

³ 以下、タスクの詳細については、金澤(編) (2014) を参照されたい。

表3 タスク 5 の品詞別・レベル別カバー率

タスク 5	下位群	中位群	上位群
名詞 (一般)	73.1%	71.7%	81.7%
名詞 (サ変可能)	76.7%	82.2%	82.9%
動詞	61.4%	63.8%	69.4%
形容詞	74.1%	80.0%	68.9%



タスク 5 は、グラフからも分るとおり、レベルに関わりなくカバー率が近く、その結果、グラフの線が平坦なのが特徴である。これは、下位群、中位群、上位群のそれぞれで使用されている語に大きな違いがないことを示している。そしてこれは、語彙がレベルを決定するための大きな要素とはなっていないとも言える。全体的にカバー率が高いのも、このタスクが被調査者である大学生、留学生に身近な話題であり、「就職活動」「卒業論文」など馴染みのある語が多かったためでもあると考えられる。

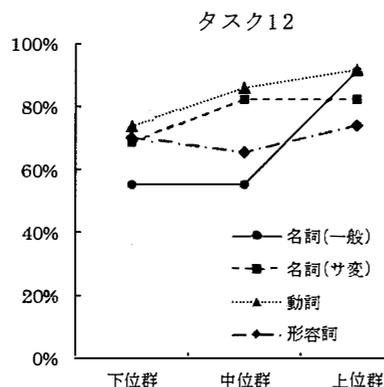
4.2 タスク 12

次に、タスク 12 の結果について述べる。タスク 12 の課題は「小学校新聞の昔話コーナーで、自分の国の昔話を紹介する」というものである。被調査者の出身国である日本、中国、韓国に共通する昔話として「七夕伝説」を書かせている。被調査者には、タスクに取り組む前に、母語による七夕伝説の物語を読ませている。

結果は以下、表 4 のとおりである。表およびグラフの見方については、前節に示した通りである。

表4 タスク12の品詞別・レベル別カバー率

タスク12	下位群	中位群	上位群
名詞(一般)	55.3%	55.3%	91.3%
名詞(サ変可能)	68.5%	82.3%	82.3%
動詞	73.7%	86.1%	91.7%
形容詞	70.0%	65.5%	74.1%



タスク12の特徴は、前節で紹介したタスク5と違い、レベル間の差が大きいことであり、その結果、グラフの傾斜も大きいものとなっている。例えば、名詞(一般)について見てみると、下位群・中位群が55.3%であるのに対し、上位群は91.3%まで上昇しており、母語話者の使用語集とほぼ同じところまで達していることがわかる。具体的には、「七夕」「織姫」「彦星」「天の川」など七夕伝説に直結した語が、下位群、中位群の上位70%には入っていない。いずれも母語話者、上位群の学習者では使用頻度の高い語であるから、これらの語を習得しておくだけでも、少なくとも語集の面ではかなり母語話者のパフォーマンスに近づくことができると言える。動詞については、「織る」という語が下位群のみ使用できていないこと、「暮らす」「許す」「困る」などは中位群まで使用できておらず、上位群のみの使用であることがわかる。

5. 全タスクの結果と考察

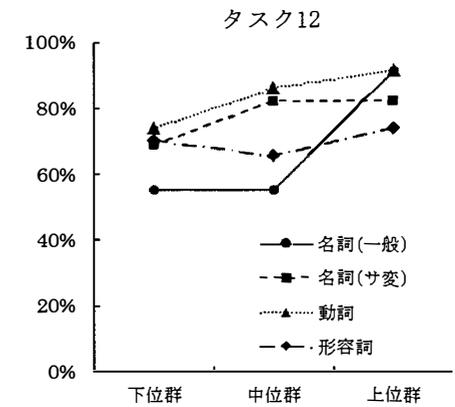
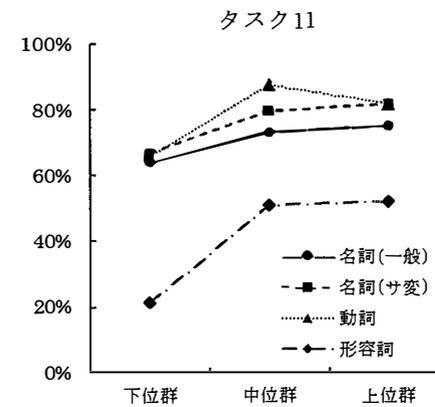
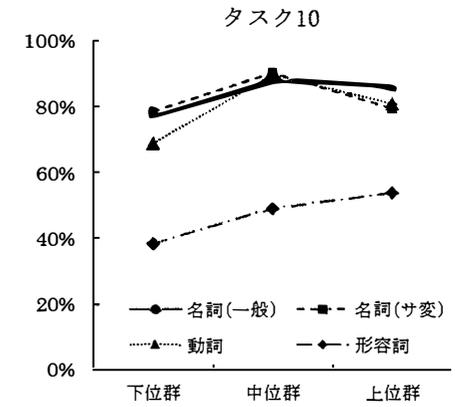
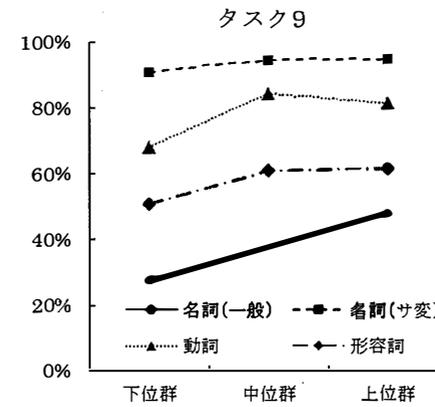
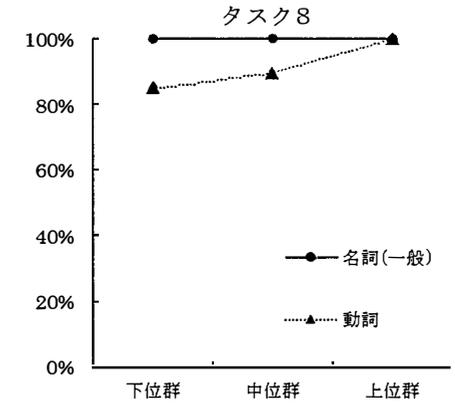
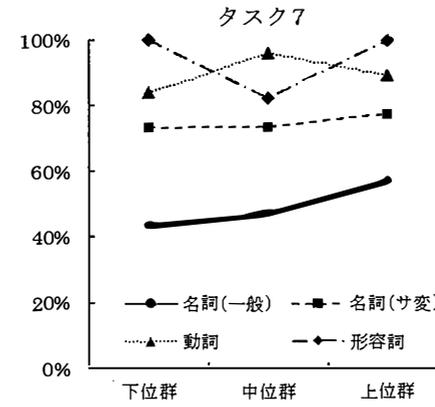
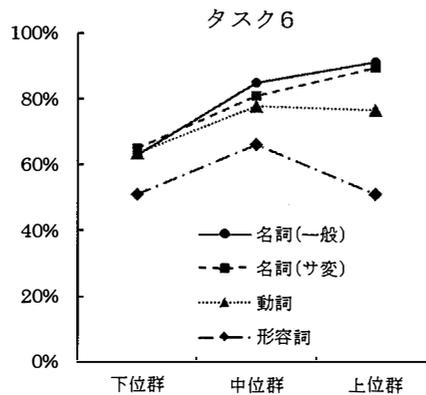
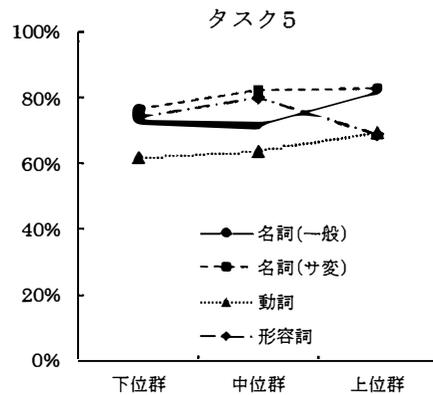
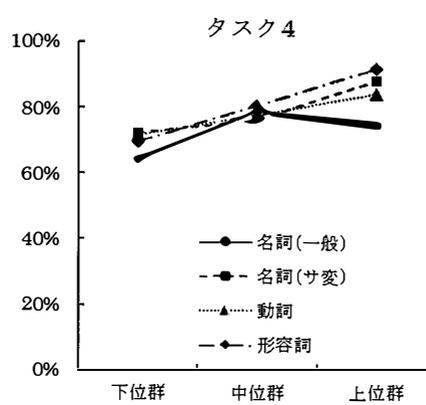
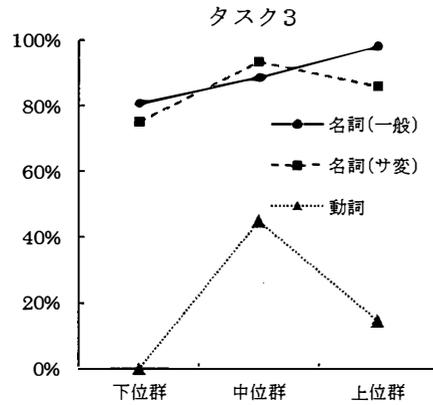
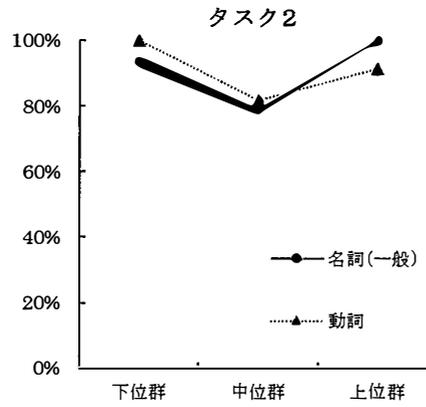
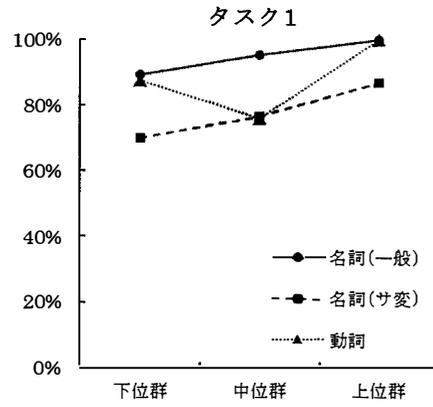
5.1 結果

4.1節ではレベル差の小さいタスク5について、4.2節ではレベル差の大きいタスク12について見たが、この節では、すべてのタスクについて紹介する。ここで示す表とグラフの見方については、4節に示したものと同一である。なお、表中、数値を示さず「—」と記してあるものは、上位70%の語をとった際に、延べ語数が多いグループでも50語に満たなかったものであり、今回の調査からは除外したものである。

表5 全12タスクの品詞別・レベル別カバー率

タスク	下位群	中位群	上位群	タスク	下位群	中位群	上位群
タスク1				タスク2			
名詞(一般)	89.4%	95.4%	100.0%	名詞(一般)	93.5%	79.0%	100.0%
名詞(サ変可能)	70.1%	76.6%	87.0%	名詞(サ変可能)	—	—	—
動詞	87.7%	76.0%	100.0%	動詞	100.0%	81.7%	91.3%
形容詞	—	—	—	形容詞	—	—	—
タスク3				タスク4			
名詞(一般)	80.8%	88.7%	98.3%	名詞(一般)	64.4%	78.6%	74.3%
名詞(サ変可能)	75.2%	93.4%	86.1%	名詞(サ変可能)	72.0%	76.1%	87.6%
動詞	0.0%	45.0%	23.5%	動詞	71.5%	77.6%	83.7%
形容詞	—	—	—	形容詞	69.6%	80.4%	91.3%
タスク5				タスク6			
名詞(一般)	73.1%	71.7%	81.7%	名詞(一般)	62.9%	84.8%	91.2%
名詞(サ変可能)	76.7%	82.2%	82.9%	名詞(サ変可能)	64.8%	81.0%	89.4%
動詞	61.4%	63.8%	69.4%	動詞	63.6%	77.7%	76.5%
形容詞	74.1%	80.0%	68.9%	形容詞	51.0%	66.0%	51.0%
タスク7				タスク8			
名詞(一般)	43.9%	47.3%	56.9%	名詞(一般)	100.0%	100.0%	100.0%
名詞(サ変可能)	73.5%	73.5%	77.5%	名詞(サ変可能)	—	—	—
動詞	84.0%	96.0%	89.1%	動詞	84.8%	89.6%	100.0%
形容詞	100.0%	82.2%	100.0%	形容詞	—	—	—
タスク9				タスク10			
名詞(一般)	27.7%	36.8%	47.5%	名詞(一般)	77.6%	88.1%	85.6%
名詞(サ変可能)	91.0%	94.5%	94.5%	名詞(サ変可能)	78.7%	90.4%	79.4%
動詞	68.1%	84.3%	81.1%	動詞	68.8%	89.6%	80.8%
形容詞	50.8%	60.7%	61.5%	形容詞	38.4%	49.3%	53.8%
タスク11				タスク12			
名詞(一般)	63.9%	72.8%	75.0%	名詞(一般)	55.3%	55.3%	91.3%
名詞(サ変可能)	66.5%	79.5%	81.4%	名詞(サ変可能)	68.5%	82.3%	82.3%
動詞	65.6%	87.4%	81.4%	動詞	73.7%	86.1%	91.7%
形容詞	21.0%	50.6%	51.9%	形容詞	70.0%	65.5%	74.1%

以下は、各タスクについて、そのカバー率の推移をグラフで示したものである。



5.2 考察

タスク 5、タスク 12 については 4 節でそれぞれ説明したので、ここではそれ以外のタスクの結果について、類似する傾向をもつものをまとめながら考察を加える。以下では、適宜、橋本（2014）の結果である各タスクの「語彙多様性」と「個人差」との関係についても考えながら考察を行う。そのため、表 6 として橋本（2014）の結果を示しておく⁴。

表 6 タスクごとの「語彙多様性」と「個人差」（橋本（2014）より）

タスクの概要		語彙多様性	個人差
タスク 1	図書貸与希望のメール（先生）	低	低
タスク 2	図書貸与希望のメール（友達）	低	低
タスク 3	折れ線グラフの説明	低	高
タスク 4	奨学金増額依頼のメール	中	中
タスク 5	悩み相談の手紙	高	中
タスク 6	病院閉鎖に対する新聞投書	中	中
タスク 7	観光地案内のメール	高	高
タスク 8	携帯メールによる事件報告	低	低
タスク 9	広報誌での料理紹介	高	高
タスク 10	早期英語教育に対する意見（先生）	中	中
タスク 11	早期英語教育に対する意見（友達）	中	高
タスク 12	小学校新聞での昔話紹介	高	低

5.2.1 タスク 1・タスク 2 について

タスク 1 とタスク 2 は、語彙多様性が低く、またメール文自体の長さもそれほど長くないため、あまり多くの語を必要としない。どちらかというとな機能語（依頼の仕方）に焦点があつたタスクであり、語彙としては、「図書」「授業」「研究」「レポート」「借りる」などの基本的な語彙を習得していれば問題なくこなせるタスクである。総じてカバー率が高くなっているのも、もともと使用している語の数が少ないことによる。70%の語以上を対象としていることは述べた通りであるが、その場合、タスク 1、タスク 2 とともにどの品詞も異なり語数は 10 語に満たない。中位群が下位群より低いものもあるが、1~2 語程度の差であり、レベル間に本質的な違いがあるとは認めにくい。

⁴ 橋本（2014）では、語彙多様性と個人差の数値を出した後に、それぞれ 3 等分して、高いものから「A」「B」「C」としてグループ別にまとめているが、本稿の表 6 では「高」「中」「低」として示す。数値の算出方法については、橋本（2014）を参照のこと。

5.2.2 タスク 3 について

タスク 3 は、「デジタルカメラの販売数についての折れ線グラフを説明する」というタスクであるが、結果からわかるとおり、動詞のカバー率が極端に低い。売上高の説明について、サ変動詞である「増加」「減少」が最もよく使用されるのは、母語話者、学習者に共通のことであるが、一方で和語動詞では、学習者が「上がる」「下がる」といった語を使用しているのに対し、母語話者がほとんど使用していないこと、そのかわりに母語話者は「売り上げる」「落ち込む」などの表現を使用していることなどが、母語話者と学習者の傾向の違いを生んでいる。複合動詞の使用が日本語学習者にとって難しいこと、日本語教育において重要であることは多くの先行研究で指摘されているところであるが、今回の調査でもやはり同様の傾向を示していると言えよう。

5.2.3 タスク 4・タスク 6 について

タスク 4 は「奨学金増額を学長に依頼するメール」、タスク 6 は「市民病院の閉鎖に対する新聞への投稿」というタスクである。どちらも比較的似た傾向を示しているため、まとめて考察する。

まず、タスク 4 についてであるが、名詞（一般）において、中位群から上位群への落ち込みがあるものの、どの品詞も最も低いグループと最も高いグループの間に 10 ポイント以上の差があり、レベル差を見るには比較的適切なタスクと言える。繰り返すことになるが、このようなタスクは次に覚えるべき語が何であるかを示しやすい。タスク 4 では、名詞（サ変）の「勉強」「負担」「考慮」、動詞の「費やす」「申し上げる」「存ずる（存じあげる）」などが、タスク 6 では「近隣」「現行」といった語が、レベル向上の鍵となる語である。また、タスク 4、タスク 6 とともに、橋本（2014）で示された語彙多様性、個人差ともに「中」であり、語彙の面でとくに注意が必要なタスクではないようである。タスク 6 のカバー率も、形容詞を除いてはバランス良く上昇しており、良いタスクだと言える。

5.2.4 タスク 7・タスク 9 について

タスク 7、タスク 9 は語彙多様性、個人差ともに「高」のタスクである。タスク 7 は料理の説明、タスク 9 は観光地の説明という、ともに「説明」のタスクであり、学習者が何（どこ）を選ぶかによって、使用される語が全く違ってくるタスクである。その影響はとくに名詞において顕著であり、今回のカバー率の調査においても、特に名詞の使用において母語話者との違いが明確になっている。ただ、それでも上位群に行くに従って母語話者と似た傾向を示すようになってきているのは、やはりそれぞれの

説明において基本的な語については共通する部分を持っていることの証拠であり、完全に個人差に委ねられるわけではないということである。また、橋本（2014）ではこのようなタスクを行う際の語彙指導の方法についても提案しているので参照されたい。

5.2.5 タスク8について

タスク8は、携帯メールによる事件報告のタスクであり、全体的に文章が短いため今回のような調査には向いていない。名詞、動詞ともに高いカバー率を示しており、語彙自体は難しくないということが指摘できる。「んだって」などを使った伝聞の方法や、短い文章での順序立てた説明の方法、「んだけど」などを用いた簡潔な前提の示し方など、文法面で多くのストラテジーを必要とするタスクと言える⁵。

5.2.6 タスク10・タスク11について

タスク10、タスク11は早期英語教育に対して意見を述べるメール文であり、タスク10が先生に、タスク11が友達に宛てて意見を述べるものである。どちらも内容は同じなので、語彙の使用についてはほぼ同じ傾向を示していることがわかる。他のタスクと異なり、形容詞については、日本語母語話者の使用している数が比較的多く、その結果、学習者が十分にカバーできておらず、低い結果となっている。使用されている形容詞自体は日本語母語話者のものであっても、さほど難しいものではないため、どのような箇所日本語母語話者が形容詞を使用しているか明らかにする必要がある。この点は今後の課題としたい。

6. まとめ

本稿では、試みとして YNU 書き言葉コーパスを使用し、レベル別、タスク別に学習者の使用語彙の実態を明らかにした。学習者の語彙の使用実態を明らかにすることは、語彙習得研究はもちろんのこと、語彙教育の発展のためにもより積極的に行われるべきであるが、一方で使用できるデータについてはやはり課題も多い。YNU 書き言葉コーパスはタスク別になっていること、レベル別になっていること、日本語母語話者との比較が可能なが大きな特長であるが、その一方で細かな分類を行うほど、各グループのサイズが小さくなっていくことが欠点でもある。今回の調査でいえば、わずかな語の多寡でもカバー率に大きく反映してしまうことが随所にあった。

タスク別、レベル別である YNU 書き言葉コーパスの特長を生かしつつ、それぞれ

のサンプル数を引き続き増やしていくことが、今後の YNU 書き言葉コーパスの課題であり、それが可能になれば語彙習得研究においても、これまで以上に積極的に活用されるものとなるだろう。

〈参考文献〉

- 金澤裕之（編）（2014）『日本語教育のためのタスク別書き言葉コーパス』ひつじ書房
 佐々木嘉則（2010）『今さら訊けない…第二言語習得再入門』凡人社
 長友和彦（1999）「第二言語としての日本語の習得研究—概観、展望、本科学研究の位置づけ—」『第2言語としての日本語の習得に関する総合研究』（平成8年度～10年度科学研究費補助金件空成果報告書／研究代表者：カッケンブッシュ寛子）pp.9-41
 橋本直幸（2014）「語彙調査に基づくタスクの分類—「語彙多様性」と「個人差」の観点から—」、金澤裕之（編）『日本語教育のためのタスク別書き言葉コーパス』pp.287-303、ひつじ書房
 橋本直幸（2015）「書き言葉コーパスから見た文法シラバス」、庵功雄・山内博之（編）『データに基づく文法シラバス』pp.67-86、くろしお出版
 山内博之（2004）「〔展望論文〕語彙習得研究の方法—茶筌と N グラム統計—」『第二言語としての日本語の習得研究』第7号、pp.141-161

⁵ この点については、YNU 書き言葉コーパスを利用して機能語の分析を行った橋本（2015）でも触れているので参照されたい。