

# 母語話者との対話における中上級日本語学習者の 使用語彙調査

石田 麻実

## 要 旨

中上級学習者は、中級以降多くの語を学ぶものの、話す際に使える語が増えないと感じている。しかし、実態は明らかになっていない。本稿は、中級以降の語彙の拡充という観点から、中上級学習者を対象に話し言葉の使用語彙について、その実態を一側面ではあるが、明らかにすることを目的として語彙調査を行った。調査資料は、ある程度話題が統一された学習者と母語話者（調査者）との4組の対話の中での学習者の発話とする。その中で、どのような語が産出され、中級以降に学んだ語でどのような語がどの程度含まれていたかについて分析した。その結果、4名とも中級以降の語の割合は、異なり語数全体の約30%であった。同じ話題で個人の使用語彙を比較したところ、動詞は差がなく、名詞は差が顕著に出た。

【キーワード】 語彙調査、中上級、使用語彙、話し言葉、対話、話題

## An Investigation of the Productive Vocabulary in the Conversations between Upper-intermediate Japanese Learners and a Native Speaker

ISHIDA Mami

【Abstract】 This study investigates the conversational usage of productive vocabulary by upper-intermediate Japanese learners. In order to achieve this objective, four pairs of semi-structured conversational interviews were performed in which the researcher observed and noted the usage of productive vocabulary. The results of an analysis performed on the vocabulary used indicate that their vocabulary corresponds with the vocabulary of the first level of the Japanese Language Proficiency Test (JLPT), indicating that 30% of the vocabulary used falls under the intermediate-level.

【Keywords】 vocabulary investigation, upper-intermediate Japanese learners, productive vocabulary, spoken language, conversation, topic

## 1. はじめに

中級以降、語彙学習が本格化し、学習者は多くの語を学ぶが、実際に話したり書いたりする際に使える語彙（以下、使用語彙とする）が増えないと感じている学習者は少なくない。山崎誠（2005, 『日本語教育辞典』:284）は、「使用語彙は、一般に理解語彙よりも少ないとされる」としている。

時間の制約のある話し言葉では、初級で学んだ語ばかり使ってしまう、中級以降に学んだ語があまり使えず、中上級段階になっても、話す際、使っている語が初級と変わらないとジレンマを感じている学習者の声も聞く。また、特に成人の学習者は、現在の日本語の話し言葉が自身の知的レベルと見合わないと感じていることも少なくない。日本語の学習が進むにつれて、より場に適した語を使いたい、自分の言いたいことや伝えたいことを忠実に表したいという気持ちが強くなり、話し言葉で語彙を増やしたいと感じるのだろう。しかし、これまで学習者の話し言葉の使用語彙を調査した研究自体が少なく、中上級学習者の使用語彙の実態が明らかになっていないのが現状である。実際に中級以降に学んだ語がどのくらい使われているのかについてはわからない。

そこで、本稿は、中上級学習者を対象に、対話での使用語彙について、中級以降の語彙の拡充という観点からその実態を一側面ではあるが、明らかにすることを目的とする。本稿では、母語話者（調査者）との対話で中上級日本語学習者の使用語彙として、どのような語が出現したのかを調査し、その中に中級以降に学んだ語でどのような語がどの程度含まれているかを明らかにする。対話の相手は、母語話者の教師（調査者）で統一した。その理由は、話題から語彙の分析もするため、ある程度話題をコントロールする必要があったためである。以上のことより、中上級学習者の話し言葉における使用語彙の特徴を探る。

## 2. 先行研究

### 2.1 話し言葉と書き言葉

日本語は話し言葉と書き言葉の違いが大きく、その区分は単純ではない。

水谷信子（2005, 『日本語教育事典』:348）は、話し言葉を「相手があって場面があるコミュニケーションの目的を持って用いられるものと限定する。また音声言語であっても書かれた原稿をそのまま読む場合の講演やテレビニュースは話し言葉として扱わない。」と定義している。

話すといっても、何らかの準備をして話すのと、即興で話すのとでは、全く違う。本稿は、話し言葉を自発性が高い話し言葉としたい。例えば、話す内容を予め考えて、メモしていたとする。話す際に、メモを見ながら話し、紙にある語を見てその語を産出しただけでは、話す際に自らその語を想起していないので、使用語彙とは言えない。よって、本稿は、水谷の話し言葉の定義にならう。

## 2.2 日本語の話し言葉の語彙調査の単位について

野元 (1980, 『日本人の知識階層における話しことばの実態』:9) は、「従来、言語能力の一方の面である書き言葉を主とした言語構造の面については研究が重ねられてきたが、話し言葉、言語行動などの実際の言語運用場面についての研究は極めて不十分と言わざるを得ない」と述べている。これまで、書き言葉に比べて、話し言葉の語彙調査は少ない。日本語の話し言葉を扱った語彙調査自体が少ないが、そのほとんどは対象が母語話者である。学習者の話し言葉から語彙調査をしたものは、見当たらない。

これまでに行われた代表的な話し言葉の語彙調査は、国立国語研究所『日本人の知識階層における話しことばの実態』と、国立国語研究所編 (1995) 『テレビ放送の語彙調査Ⅰ—方法・標本一覧・分析編—』、国立国語研究所編 (1997) 『テレビ放送の語彙調査Ⅱ—語彙表—』である。この2つの調査は、目的、調査対象、調査単位ともに異なる。

本稿では、『テレビ放送の語彙調査』で用いられた調査単位である長い単位を使用する。

『テレビ放送の語彙調査』は、1989年4月～6月放送の6放送局7チャンネルのテレビ放送の語彙調査であり、次の4点を目的として行われた。①音声言語と文字言語とが複合して用いられているテレビ放送語彙の多様性をとらえる、②現代のテレビ放送において、どのような語が、どのような場合に、どのくらい使われるのか、その実態を調べる、③テレビ放送の語彙をとおして、現代日本語の語彙のありさまを追究する、④テレビ放送が、その音声と画面とを通して伝える語を、どのように調査するか、その方法論を確立するための基礎を築く、という目的で行われた。調査単位として文節 (単語) 相当の「長い単位」を用いている。この調査での「長い単位」の定義は、「文を構成する上で、かかり・うけなどの構文的な機能を担う、連続した最小の構文的成分 (=文節) を1単位とする」としている。下に長い単位で区切った例を示す。

＜長い単位の例＞ 型紙どおりに | 裁断して | 外出着を | 作りました |

上記の例のように、日本語教育では、「作りました」で一語として提示することがほとんどである。本稿では『テレビ放送の語彙調査』の調査単位である長い単位 (文節=単語相当) にならうこととする。

『日本人の知識階層における話し言葉の実態』は、日本語学習者の日本語の言語能力の測定基準設定を目指している。調査単位の区切り方は、原則的にβ単位を採用している。β単位は、短い単位の系列で、主として言語の形態的な側面に着目して考えた単位である。下にβ単位で区切った例を示す。

＜β単位の例＞ 型紙 | どおり | に | 裁断 | し | て | 外出 | 着 | を | 作り | まし | た |

日本語教育の場で | 作り | まし | た | と区切って提示することはほとんどないと考え、本調査ではβ単位を調査対象としなかった。

日本語教育における話し言葉の語彙調査の研究では、池田伸子 (1996) の「日本人ビジネスマンのビジネスの場における話し言葉の語彙調査」がある。池田は、日本人ビジネスマンのビジネスの場における話し言葉の語彙を調査、分析し、ビジネスマンの話し言葉の語彙の特徴を明らかにした。調査対象は、ビジネスマンの会社における発話を集めたものである。調査単位は『日本人の知識階層における話しことばの実態』と同じ単位β単位である。ビジネス場面での高頻度語彙を明らかにしたが、学習者が利用する資料であるので、長い単位の調査のほうが適していたのではないだろうか。例えば、この調査で派生語である「～的」の使用頻度が高かった (26位) ということ指摘しているが、「的」の前についた語にはどんな語が多かったのか、複合語で資料として提供したほうが、学習者にとって、役に立つのではないかと考える。

ここで、<長い単位>と<β単位>を比べる。この2つ以外にも調査単位はあるが、話し言葉を扱っている語彙調査の単位ということで、この2つの単位から考える。

#### <長い単位の例>

型紙どおりに | 裁断して | 外出着を | 作りました |

#### <β単位の例>

型紙 | どり | に | 裁断 | し | て | 外出 | 着 | を | 作り | まし | た |

本研究では、話し言葉の実態を扱うが、β単位は、言語の形態的な側面に着目して考えた単位であり、上の例では「外出着」が「外出 | 着」と区切られてしまい、派生語を一語として扱えないという問題が出てくる。また「裁断して」の漢語サ変動詞も「裁断 | し | て」と区切られてしまう。また、学習者が産出した語から、合成語が作れるかを見ることも日本語学習者の使用語彙の側面を捉える上で重要である。合成語を一語とするためには、短い単位の系列であるβ単位で区切るのは適切ではない。よって、本稿では『テレビ放送の語彙調査』の調査単位である長い単位 (文節=単語相当) にならうことにした。しかし、学習者の話し言葉には、母語話者のそれよりもかなり誤用も多く含まれると推測できるため、新たな基準も設ける必要がある。

### 3. 調査概要

#### 3.1 調査協力者

2009年春学期に私立大学の日本語中上級クラスに在籍していた中上級日本語学習者4名に協力してもらった。調査協力者は、ハンガリー語母語話者 (以下、Hとする)、イタリア語母語話者 (以下、Iとする)、韓国語母語話者 (以下、Kとする)、中国語母語話者 (以下、Tとする)。対話調査では、できるだけ話しやすい状況を作りたかったため、初対面でない学習者にお願いする予定であったが、Tのみ初対面であった。協力者には、予め調査概要を伝え、録画の許可を得た。話す話題については、学習者には事前に知らせない。次の

表1は調査協力者のプロフィールである。

表1 調査協力者のプロフィール（母語、レベル、学習歴）

No.	協力者	性別	母語	出身	日本語学習歴	日本語能力試験
1	H	男性	ハンガリー語	ハンガリー	5年	2級受験(2009)
2	I	男性	イタリア語	イタリア	7年	2級受験(2009)
3	K	女性	韓国語	韓国	2年半	1級受験(2009)
4	T	女性	中国語	台湾	2年3カ月	2級受験(2009)

### 3.2 調査方法

この節では、4つの調査方針を述べる。

1) 独話ではなく、日本語母語話者との対話を資料とする。

話し言葉は「相手があって場面があるコミュニケーションの目的を持って用いられるものと限定する」(水谷)。そのため、相手がいない独話を調査資料とするのは適切でない。また、相手が一人の対話のほうが学習者に自発的に話したいという気持ちが生まれやすいと考えた。また予備調査から学習者同士や母語話者の友人を対話の相手にした場合、対話の話題に共通性を持たせることは、今回は難しいと判断した。話題が拡散してしまうと、学習者間での使用語彙を比較することも難しくなる。ある程度、対話を構造化することは、もちろん自然な対話とは言えないが、今回は対話の自然さではなく、対話の話題を重視するために、調査者が全ての協力者の対話の相手となった。

2) 日本語学習者と日本語母語話者との4組のペアの対話を資料とする。

日本語学習者は、中上級レベルと設定した。本稿では、中上級レベルを、2009年度までの「日本語能力試験」2級程度の学習者とする。今回は、漢字圏、非漢字圏と限定しなかった。

3) 対話の時間は、複数のペアでなるべく同じくらいに揃える。

同じ調査時間でも発話量の問題はあるが、学習者4名の発話でどのような中級以降の語がどの程度産出されているか実態を見ることを目的としているため、なるべく対話の時間は揃えたほうが良いと考えた。対話の時間は40分前後を想定する。しかし、対話の流れを無視することはせず、話が続いている場合は40分で切ることはいらない。時間より次の4)で述べる5つの話題で会話することを優先する。

4) 対話の話題は複数のペアでなるべく同じ話題に揃える。

対話の話題は特別な知識を必要としない一般的なものとするようにした。例えば、経済に関心がなかったら、経済の話題は母語でも話せない可能性がある。そのような場合、話

せないのは、その話題自体に知識があるかどうかの問題になり、日本語の問題ではなくなってしまう。そのため、『ACTFL-OPI入門』(2001:40) の話題のレベルを参考にし、一般的な出来事の話から「大学」「仕事(就職)」を選んだ。話題の順番は、易しい話題から難しい話題へ移行するようにした。

対話には、全ての協力者との5つの大きな話題を入れる。対話は、話題ごとに質問項目もいくつか準備し、ある程度コントロールする。しかし、それにこだわりすぎることはせず、対話の流れを壊さないように、臨機応変の対応を心がけた。よって、大きな話題の中身は、対話の流れに任せた部分もあり、ペアごとに小話題、さらにその詳細な話題は異なることもあった。話す話題については、事前に知らせない。

### 3.3 本調査

#### 3.3.1 概要

2009年8月から9月に中上級日本語学習者と母語話者との対話調査を行った。母語話者は調査者である。調査時間は40分前後とし、対話は全て録画した。対話後、対話に関する簡単なアンケートとインタビューを実施した。

#### 3.3.2 対話調査

対話の相手となる調査者の態度・心構えとして同じくインタビュー式会話であるOPI<sup>1</sup>のテスターがすべきことを参考にした。しかし、この調査はテストではないため、加筆修正を行った。協力者のレベル判定を行うわけではないので、言語的挫折をさせるようなことはしない。協力者が言葉につまったり、困ったりしているようなときは、助け舟を出す。その理由は、母語話者の支援によって使用語彙になる場合もあるからである。また、学習者から質問された場合も教えないようなことはしない。また、協力者の沈黙を尊重するようにする。沈黙したからといって、すぐに質問を変えたり、話題を変えたりしない。その理由は、返答に窮しているのではなく、協力者が単語を思い出したり、内容を模索したりしていることもあり、その際、会話を遮ってしまうと協力者の語の産出を妨げてしまう可能性があるからである。

調査時は、夏だったため、話しやすい日本の夏の話から対話を始めた。次の5つの大きな話題の流れで、各協力者と対話を行った。

#### <調査当日の対話の話題の流れ>

**【夏】 → 【出身地】 → 【余暇の過ごし方】 → 【大学】 → 【仕事(就職)】**

5つ目の話題「仕事」まで終わったら、対話を終了とした。

### 3. 3. 3 本調査の調査時間と対話の話題

#### 1) 調査時間

調査時間は約40分の予定であった。協力者3名は40分前後であったが、Kは約20分長くなってしまった。この理由は、対話中は相手の話に集中してしまい、なかなか時計を見ながら対話をするということができなかったためである。これは反省点である。Kとの対話は40分で切ることもできたが、その時点で5つ目の話題「仕事」について全く話していなかった。この調査では、時間の統制より話題を揃えることが重要であるので、協力者の了解を得て時間を延ばした。調査協力者と4名と各協力者の対話時間については、次の表2の通りである。

表2 調査協力者及び対話調査時間

No.	協力者	母語	対話時間
1	H	ハンガリー語	38分
2	I	イタリア語	42分
3	K	韓国語	62分
4	T	中国語	42分

#### 2) 本調査で話された話題

次のページの表3は対話で話された話題を表にまとめたものである。話題は上から下へ時系列順に並んでいる。大きな5つの話題【夏】→【出身地】→【余暇の過ごし方】→【大学】→【仕事】の流れで話されたが、各話題において、その内容、小話題は異なった。

ここでの大きな話題の区切りは、調査者が意図的に話題を変えようとした発話時から区切っている。表3から、大きな話題を構成している小話題が、共通しているものもあるが、そうでないものがあることがわかる。大きな話題ごとに、準備しておいた小話題もあったが、対話の中でそれを制限するようなことをせず、対話の流れを大事にした。そのため、協力者間で小話題の違いが出るのはやむをえない。しかし、5つの大きな話題の流れ【夏】→【出身地】→【余暇の過ごし方】→【大学】→【仕事】で対話調査を行うことはできた。

表3 会話調査で話された話題

大話題	小話題			
	協力者H	協力者I	協力者K	協力者T
①夏	夏	夏	夏	夏、台風
②出身地	出身地	出身地 自国の料理	出身地 幼少時の海外滞在	出身地 自国の料理
③余暇の過ごし方	自国での過ごし方 日本で：旅行	日本で：旅行 自国での過ごし方	日本で：東京案内 自国で：サークル活動	自国で：インターンシップ 日本で：旅行
④大学	日本の大学生 大学生活	日本の大学生 大学生活	日本の大学生 サークル	日本の大学生 飲み会、サークル 所属学部
			就職活動	
	大学入試	大学入試	受験勉強 (自国の高校生) 大学入試 浪人	自国の大学 受験勉強 大学入試
	自国の大学の制度 自国の高校生 浪人	自国と日本の教師との比較 教育制度 学部の選択	自国の大学の評価制度 自国の大学への不満 学部の選択	授業内容
⑤仕事	希望の仕事 就職活動 就職先の希望条件	希望の仕事 就職活動 就職先の希望条件 アルバイト	希望の仕事 就職活動 就職先の希望条件 進路(大学院)	希望の仕事 就職活動 就職先の希望条件 自国の大学生の進路

#### 4. 中上級日本語学習者の使用語彙リストの作成

調査単位は、「語」とした。今回は、実質的意味を持つ名詞、動詞、イ形容詞、ナ形容詞、副詞、連体詞と接続詞を扱うこととし、感動詞、固有名詞は対象としなかった。また、語形に誤りがあるものは使用語彙として認めなかったため、それらを除いて、学習者別に使用語彙リストを作成した。

##### 4.1 調査対象

調査対象は、会話開始から終了までの中上級日本語学習者の発話のみとし、母語話者の発話は含まない。学習者のみの発話を対象とした理由は、今回の調査の目的は、母語話者とのやり取りがうまくできているかについて見る分析ではなく、学習者が母語話者との対話でどのような語を使ったのかを見るためだからである。そのため、今回の調査では、母語話者の語彙調査は必要ないと考え、対話全体を対象にした語彙調査ではなく、学習者の発話のみを対象とした語彙調査とした。3.3.3で述べたように、調査時間には差が出たが、この語彙調査は、学習者が産出した語彙量を調査することを目的としているのではなく、共通した5つの話題でどのような語彙を産出したか、学習者の使用語彙の構成、異なり、豊かさを調査することを目的としている。そのため、Kは対話開始から終了まで60分かかっ

たが、対話時間全てを対象とした。

## 4.2 使用語彙調査手順

語彙調査は、会話文字化資料作成→協力者の発話部分の取り出し→形態素解析→形態素解析の結果を手作業で修正→単位切り（長い単位）→代表形をとる→品詞付与→分析の手順で行った。以下、各工程について詳しく述べる。

### 4.2.1 会話文字化資料作成

録画資料を文字化した。対話の文字化は、宇佐美まゆみ（2007）「改定版：基本的な文字化の原則（Basic Transcription System for Japanese: BTSJ）」にならって行った。文字化は学習者だけでなく、母語話者の発話も行った。文字化資料には、話題、発話文番号、発話者、発話内容、備考の項目を設け、協力者別に作成した。備考欄には、表情など気になった点を記録した。

### 4.2.2 単位切り

まず、文字化した資料から調査協力者の発話した文のみを取り出し、形態素解析にかけた。解析器として、「茶釜：ChaSen」<sup>2</sup>を使用し、解析結果をMicrosoft Excelに出力し、表形式に変換した。これが語彙リストの基となる。出力された表に、話題欄、発話分番号欄を加え、それぞれ入力した。この語はどの話題で、また、どの発話文で産出された語か検索しやすいためである。さらに、語形の誤りの欄も設けた。語形誤りがあるかどうかの情報は、使用語彙を認定する際に、必要だからである。形態素解析は、機械的な処理のため、不自然な処理が出てくる。形態素解析の結果は、全て自分の目で確かめ、誤りがある部分は手作業で修正した。

### 4.2.3 代表形

単位語から代表形にする作業を行った。以下に例を2つ示す。

例1) 単位語：「行ったことがある」→代表形：「行く」

例2) 単位語：「有名だったんです」→代表形：「有名」

学習者の場合、同音異義語が学習者の発音（アクセント）からでは判断しにくい。例えば、「キョウカイ」に対する「協会」「教会」「境界」のような場合は、文脈からどの代表形を判別する。また、ある語に接辞がついた場合は、それぞれ別の代表形（例「私」と「私たち」）になる。数詞も同様に別の代表形（例「1年間」、「2年間」）になる。

単位語から代表形にする作業で、語形から何を意味しているのかわからないものがあった。また、語形が不十分であっても、聞き手が意味を推測して意味が伝わったものもあつ

た。後者は、聞き手によるところが大きく、別の聞き手なら推測がきかず、意味が伝わらなかった場合も考えられる。そのため、語形に誤りがあった場合は、聞き手が推測して意味がわかったかどうかは考慮に入れず、リストの備考欄に「不明」と入力した。

#### <「不明」としたものの基準>

##### 1) 語形に誤りがあるもの

例) 「アチガツ」(Iの発話)

Iは「八月(はちがつ)」と言ったつもりだった。語形が不十分であるが、聞き手に意味が伝わっていた。しかし、語形が不十分のため不明として処理した。

##### 2) 日本語にその語形があるが、その語形の意味で使用しなかったもの

例) 「キョウカイ」(Kの発話)

文脈から、Kは、「協会」「教会」「境界」のどの意味で使ったわけでもないこと判断できた。フォローアップインタビューからは、Kは「郊外」を「キョウカイ」として、発音していたことが確認された。この例のように、日本語の単語としてその語形が存在するが、明らかに文脈上からその語形が持つ意味を持って使われていないと判断できる場合は、「不明」として処理した。

##### 3) 学習者が語構成の知識で作った複合語で、意味の推測が困難なもの

例) 「ペットスル」(Hの発話)

「ペットする」は、Hが「かわいがる」が思い出せなく、英語の「Pet」に「する」をつけ動詞を作った。しかし、英語のPetの動詞の意味を知らなければ、Hが発した「ペットする」の意味は相手に伝わらないだろう。

##### 4) 日本語以外の外国語の単語

例) ポース → 韓国語(英語のauraの意味)

「不明」として処理されたものは、語彙リストとは別にまとめた。

#### 4.2.4 品詞の付与

次に、代表形に品詞を付与する作業を行った。林四郎監修(2012)『例解新国語辞典(第八版)』の品詞分類の基準に従った。長い単位で区切ったため、品詞分類は、名詞、動詞、形容詞、形容動詞、連体詞、副詞、接続詞、感動詞となる。なお、品詞分類に「形容詞」、「形容動詞」と記述がある。しかし、日本語教育では一般的に「形容詞」を「イ形容詞」、「形容動詞」を「ナ形容詞」と呼ばれるため、本稿でもその名称に従う。

『例解新国語辞典(第八版)』は、日本の中学生向けの辞書のため、見出しの語句が58,000と、一般の国語辞典と比べて少ないが、本稿では、中上級日本語学習者が対話で使った語を調べるためには、58,000の語句で十分だと考えた。この辞典を選んだ理由は、一般の国語辞典よりも意味・用例が詳しく記述されていたためである。これは、学習者の使用した

語に品詞を付与する際に大変役に立った。『例解新国語辞典（第八版）』に従い、品詞を付与することとした。しかし、学習者の発話で使われた用法の記述がないものは、文化庁（1990）『外国人のための基本語辞典（第三版）』、金田一京助ほか編著（2011）『新選国語辞典（第九版）』も参照した。品詞の判断に迷うことも少なくなかった。その場合、文字化資料に戻り、文脈から判断した。次に例を示す。

例 1) 「あの」「その」 → 連体詞か感動詞（フィラー）か

例 2) 「あと」 → 文頭で接続詞として使われているのか

例 2) の「あと」の用法は、『例解新国語辞典（第八版）』には記述がなかった。しかし、『三省堂国語辞典（第六版）』（2008）には、接続詞の記述があった。明らかに接続詞として機能しているため、この場合「あと」の品詞は接続詞とした。

品詞の付与が終わってから、語彙リストの対象となっていない感動詞、固有名詞は除いた。本稿では、実質的意味を持つ名詞、動詞、イ形容詞、ナ形容詞、副詞、連体詞と接続詞を扱うこととしている。

#### 4.2.5 『日本語能力試験主題基準（改定版）』による級の付与

会話で使用された語彙の難易度を見る指標として、『日本語能力試験主題基準（改定版）』（2002）に従い、代表形に 1 級から 4 級を付与し、記載がない語は「級外」とした。本稿では、日本語能力試験改訂前である 2009 年までの受験級の語彙表に従っている。

ここで 1, 2 級語彙選定基準の一部を示す。

< 1, 2 級語彙選定基準 > 『日本語能力試験出題基準（改定版）』：53

1, 2 級の語彙として、次のような語は原則として採用しないこととする。

(1) 他の採録語と重複すると考えられる語

①部分からの類推が可能な複合語

②形容詞、形容動詞の語幹に「み」「さ」がついた語

③接頭語、接尾語が付いた語

④補助動詞、補助形容詞の付いた語

この基準では、合成語が採用されていないことがわかる。しかし、学習者が産出した語には、合成語が多く使われているため、合成語の級の付与に対して、新たに以下の 2 つの基準を設けることにした。

①独立して使われる 2 語が結びついた合成語つまり複合語は「級外」とする。

例) 「労働時間」：「労働」（2 級）、「時間」（4 級）である。しかし、「労働時間」は別の一語であるため級外とする

例) 「国際交流」：「国際」（3 級）、「交流」（2 級）が「国際交流」で一語であるため級外とする。

②独立して使われる語と造語成分が結びついた語、派生語は、基本的に造語成分の級とする。

例)「外国語」→「～語」は4級である→4級と認定

級の付与で修正が必要な語がいくつかあった。例えば、「何も」は2級語彙の語彙表に載っていた。学習者の発話では、「何もありません」で「何も」が使われていたが、初級で習うものであるのに、なぜ2級なのだろうか。出題基準の語彙表には記述がなかったが、おそらく「何もそんなことしなくてもいいのに」というような文脈で使われる「何も」を2級としているのだろうと判断し、「何もありません」の「何も」は2級とせず、級外扱いにした。このように、疑問に感じた点は修正した。

#### 4.2.6 集計の方法

まず、品詞ごとにExcelのシート別に作成した。そして、代表形の読みの列（カタカナ表記）を50音順に並べ替えてから、代表形の列をグループごとに集計した。この作業によって、代表形ごとの使用頻度を出した。

#### 4.3 学習者の使用語彙リスト

##### 1) 対象とする品詞

今回、使用語彙リストの対象は、実質的意味を持つ名詞、動詞、イ形容詞、ナ形容詞、副詞、連体詞と接続詞を扱うこととし、感動詞、固有名詞は除いた。

##### 2) 使用語彙の認定基準

本稿における使用語彙の認定基準を示す。

<基準1> 学習者が自発的に使おうとした語であること。

補足：①語が産出するまで少々時間がかかっていた語も認める。

②自ら産出しようとしたが、なかなか産出できず、母語話者からの何らかの助けによって、最終的に正しい語形を産出できた語も使用語彙と認める。

③元々その語自体を知らなく母語話者に直接たずねて語形を教えてもらい、文脈で使うことなく単に反復しただけの語は使用語彙とは認めない。

④自発的にその語を使って意味を表現しようと思ったのではなく、母語話者が使った語の意味が分からず、聞き返しただけの語は使用語彙とは認めない。

<基準2> 語彙リスト作成段階で備考欄に「不明」としたものは含まない。

##### 3) 作成した学習者の使用語彙リスト (Excel) の性格

個人ごとの使用語彙が、品詞、級、話題、使用頻度からみることができた。次の章では、作成した語彙リストを分析し、各協力者の使用語彙の特徴、そして、協力者全体からみる特徴を明らかにしていく。

## 5. 調査結果と考察

### 5.1 各協力者の使用語彙全体の分析

4.で作成した使用語彙リストから中級以降の語がどの程度含まれ、どのような語が現れたか調査し、個人の使用語彙を品詞別、話題別に分析した。

#### 5.1.1 各協力者の使用語彙の延べ語数・異なり語数

表4 協力者別の異なり語数と延べ語数

	異なり語数	延べ語数	一語あたり
H	343	1020	2.97
I	325	1024	3.15
K	486	2139	4.40
T	282	798	2.83

表4から、Kが異なり語数、延べ語数ともに多いことが分かる。これは、調査時間が他の協力者より長かったためである。次に「一語あたり」の欄をみると、異なり語数、延べ語数ともに最も多かったKが4.40と最も高く、他の学習者よりも同じ語をよく使っていることを意味する。今回の調査は、中上級学習者の使用語彙にどのような語が現れたかを明らかにすることが目的である。異なり語数は語彙の豊かさを見る指標になるため、各協力者の使用語彙の異なり語数から分析する。これからの分析は全て異なり語数を基に行う。

#### 5.1.2 級別構成

4.2.5で代表形に『日本語能力試験出題基準（改定版）』に従い、級を付与した。表5は、5.1.1で示した個人別の使用語彙（異なり語数）における日本語能力試験の級別構成である。なお、各協力者の左の数値は異なり語数、右の括弧の中は異なり語数全体に占める割合を示す。

表5 「異なり語数の級別構成」(左の数値は異なり語数、右は異なり語数全体に占める割合)

	4級	3級	2級	1級	級外	合計
H	169 (49.3%)	69 (20.1%)	59 (17.3%)	13 (3.8%)	33 ( 9.6%)	343 (100%)
I	151 (46.5%)	80 (24.6%)	60 (18.5%)	5 (1.5%)	29 ( 8.9%)	325 (100%)
K	238 (49.0%)	84 (17.3%)	93 (19.1%)	13 (2.7%)	58 (11.9%)	486 (100%)
T	135 (47.9%)	52 (18.4%)	55 (19.5%)	7 (2.5%)	33 (11.7%)	282 (100%)

表5から4名とも4級の割合が最も高く、4級と3級を合わせると、Hが69.4%、Iが71.1%、Kが66.3%、Tが66.3%となる。4名の間で少々差はあるものの、3、4級の語、いわゆる初級の語が約70%を占めることになる。3、4級の語は基礎語彙であり、高頻度の語彙であるので当然の結果といえる。

しかし、2級は、Hが17.3%、Iが18.5%、Kが19.1%、Tが19.5%となっており、異なり全体の20%近くを占めたという結果は意外であった。中級以降の語彙として、2級と1級を合わせると、4名とも20%に上る。よって、中級以降の語彙は、全体で20%産出されたということになる。級外まで合わせれば、その割合は30%以上となる。

KとTに限っては、3級より2級の語のほうが多かった(Kは3級84語・17.3%に2級93語・19.1%、Tは3級52語・18.4%、2級55語・19.5%)のは、興味深い結果であった。これにはK、Tの母語の関与の可能性が考えられる。

筆者の予想は、異なり全体での2級語彙はもっと割合が低く、基礎語である3級語彙のほうが多く産出されているだろうということだったため、この結果は興味深かった。中上級の学習者から、中級以上の語彙がなかなか使えないという声があったが、実態はそうでもないのではないだろうか。本稿で定義している中級以降の語と学習者が意識している中級の語が異なっている場合も考えられるが、少なくとも初級の語ばかりではなく、中級レベル以上の語も話し言葉で使用語彙になっていたことがわかった。

### 5.1.3 品詞構成

4名とも名詞、動詞、副詞、イ形容詞、ナ形容詞、連体詞、接続詞の順に多かった。4名とも名詞が最も多い。接続詞、連体詞については、語数も10以下と少なく、品詞構成の比率も2%未満のため、各協力者間で出ている語もあまり変わらない。そのため、品詞別の分析対象としないことにする。

## 5.2 品詞別にみた個人の使用語彙

各協力者の使用語彙を品詞ごとに分析していく。ここでは、1、2級の語彙を中級レベル以上の語彙とみなし、そのような語がどのぐらい使用語彙として使われていたのか、そして、それはどのような語であったか、ということを中心に分析していく。また、使用語彙のうち、中級レベル以上の語で協力者間において共通している語とそうでない語についても分析する。

### 5.2.1 名詞

次のページの表6から、4名とも4級の割合が、Hが45.5%、Iが39.6%、Kが47.2%、Tが43.2%と最も多い。しかし、4名とも、2番目に多いのが、3級でなく2級であるという

表6 名詞の級別構成（左の数値は異なり語数、右は異なり語数全体に占める割合）

	4級	3級	2級	1級	級外	合計
H	81 (45.5%)	28 (15.7%)	31 (17.4%)	10 (5.6%)	28 (15.7%)	178 (100%)
I	72 (39.6%)	36 (19.8%)	43 (23.6%)	4 (2.2%)	27 (14.8%)	182 (100%)
K	144 (47.2%)	38 (12.5%)	68 (22.3%)	10 (3.3%)	45 (14.8%)	305 (100%)
T	70 (43.2%)	18 (11.1%)	38 (23.5%)	7 (4.3%)	29 (17.9%)	162 (100%)

う結果が出た。これは、他の品詞と異なる結果である。2級の語数は、Hが31語で17.4%、Iが43語で23.6%、Kが68語で22.3%、Tが38語で23.5%となっている。KとTに至っては2級と3級の語数の差が顕著である。Kは、3級の語数38に対して、2級の語数が68と2級のほうが30多い。Tは、3級の語数18に対して、2級の語数が38と倍以上になっている。2級名詞は漢語が増えるため、KとTの結果は母語の影響が考えられる。しかし、母語の面でK、Tより不利と考えられる非漢字圏学習者H、Iも2級名詞が3級より多いという結果は、興味深い。しかし、HとIは学習歴5年と7年なので、その長さが要因として考えられる。1, 2級の名詞で4名とも使用していたのは、「感じ」「給料」「就職」「大学院」であった。4名とも「仕事」の話題で「給料」を使っていた。「就職」「大学院」は、「大学」「仕事」の2つの話題で出ている。「感じ」は、話題を選ばず、一般性が高い名詞であるためか、話題をまたいで出ている。

次に、1名しか使用しなかった名詞、つまり協力者間で異なりが見られた1, 2級の名詞を次の表7に示す。

表7 異なりが見られた1, 2級の名詞（1級の語彙は\*をつけてある）

	H	I	K	T
夏	日陰*、皮膚	温度	—	—
出身地	氷、雰囲気	お土産、市場、地方	馬、種類、体育、都市、農業	—
余暇	アクセント、入浴* 運転者、ビール	芸術、城、仏像*	協会*、照明*、新人* 通訳、量、レベル	お勧め*、温泉、作品 時期、展覧会、美術*、 老人
大学	実際、状態、全国 フィルター* ポイント*、面 法学部	エンジニア*、期待、 義務、義務制*、国立、 事実、点数、部分、 間違い、論文	学年、空間*、自習室 実力、受験、主専攻、 書類、推薦、未*、副、 ストレス、スピーチ、 センター、大学ごと、 人間、不満、未来、迎 え、面接、申し込み*	科目、言語学、広告、 ゼミ、タイトル*、 メッセージ*、 目的
仕事	インタビュー、能力、 せい(所為)、履歴*、 倒産*	支店、人気、物価	課題*、休学*、単位、 小学生、通訳者	確認、学歴*、可能性、 心理、内容

表7を見ると、各協力者で、大学の話題で中級以降の語彙がよく出ていることがわかる。大学の話題から、協力者間での異なっている語を見ると、Kが中級以降の名詞を最も多く産出していることがわかる。Kが大学の話題で産出した語彙は、大学の話題の中でも大学受験について話しているときに使った語がほとんどである。大学受験の話題はどの協力者にも振っているが、Kはこれだけ中級以降の語彙を使って説明していたということである。Tは、大学の授業の内容について話しているときに産出した語彙である。Tは、日本の学部の授業の話をしたが、他の学習者はその話題について話さなかったので、異なった語として出たのだろう。

「級外」になった名詞は複合名詞（例：就職活動、ソフト開発など）とカタカナ語（コンパ、コンビニなど）が多く、略語も見られた。複合名詞を産出するためには、構成している語の知識だけでなく、日本語の語構成の知識も必要である。そのため、級外となっている語の多くは、初級の語とは言えないだろう。中上級日本語学習者は、自分から語を作り出すという力も持っていそうである。

## 5.2.2 動詞

表8 動詞の級別構成（左の数値は異なり語数、右は異なり語数全体に占める割合）

	4級	3級	2級	1級	級外	合計
H	34 (49.3%)	21 (30.4%)	9 (13.0%)	3 (4.3%)	2 (2.9%)	69 (100%)
I	27 (44.3%)	19 (31.1%)	12 (19.7%)	1 (1.6%)	2 (3.3%)	61 (100%)
K	41 (43.6%)	29 (30.9%)	15 (16.0%)	1 (1.1%)	8 (8.5%)	94 (100%)
T	20 (41.7%)	16 (33.3%)	11 (22.9%)	0 (0.0%)	1 (2.1%)	48 (100%)

表8から、4級と3級を合わせた割合は、Hは79.7%、Iは75.4%、Kは74.5%、Tは75.0%を占めている。今回の調査では、漢語サ変動詞（例：「就職する」）も含んでいる。次に2級の異なり語数を見ると、Hが9、Iが12、Kが15、Tが11となっている。この中で4名に共通して出現していた動詞は「就職する」であった。仕事の話で就職の話をしたときに4名とも出現していた。

次に協力者間で異なりが見られた1, 2級の動詞を次の表9に示す。

表9 協力者間で異なりが見られた1, 2級の動詞（\*は1級動詞）

	H	I	K	T
出身地	—	接する	農業する	—
余暇	当てる、混ぜる	過ごす、日焼けする	—	のんびりする
大学	限る、提出する	あきらめる、間違う 甘やかす、解決する、 努力する	延長する、自殺する、 受験やる、整理する、 通訳やる、通訳する、 発表する	交流する、申請する、 分析する、迷う、 目指す
仕事	受け入れる*、減らす、 倒産する*、 組み合わせる*	稼ぐ、繋がる	休学する*、出会う	—

表9で、注目したいのは、H、Iは、漢語サ変動詞より和語動詞を多く産出している点である。Iは、「あきらめる」「甘やかす」「解決する」「努力する」「間違う」は大学の話題で自国の教師と日本の教師を比較しているときにのみ産出しており、その話題の中だけで5つの2級動詞が出現していた。また、Hも和語動詞のほうが多いが、その中でも1級動詞を3語使用していた。この3語はどれも仕事の話で出現したものである。「倒産する」と「受け入れる」は母語話者の助けもあり正しい語形を産出できたものであるが、「組み合わせる」は言い淀みもなくすぐ産出できていた。日本語学習者にとって、複合動詞の使用は難しいと言われているが、Hは「受け入れる」「組み合わせる」を産出することができた。HとIと反対に、KとTは漢語サ変動詞のほうが多かった。Kに至っては「出会う」のみ和語動詞で、他の中級以降の動詞は漢語サ変動詞であった。また、Kは名詞に「する」をつけるだけでなく、漢語名詞に「やる」をつけて動詞を作っていた。これは他の協力者には見られなかった。Tは「のんびりする」を余暇の話題で産出したが、それ以外は全て大学の話題で出現した語であった。

3, 4級も含めた漢語サ変動詞の異なり語数と全体に占める割合は、Iが18語で17.6%、Hが13語で18.9%、Kが20語で22.2%、Tが8語で16.0%であった。Tは漢字圏だが、漢語サ変動詞の数も割合がI、Kより低く、語数が8語と4名で最も少なかったのは意外な結果であった。Tは他の3名より動詞の異なり語数自体が少ないが、その中でも漢語サ変の割合が少ないということは、元々持っている漢字の知識をまだあまり活用できていないのかもしれない。

中級以降の動詞では、非漢字圏学習者のほうがより難易度の高い和語動詞を産出していた。非漢字圏学習者は、漢語にあまり頼っておらず、漢語サ変を多用していない可能性がある。反対に、中国語母語話者、韓国語母語話者は、中級以降の動詞は、漢語に頼り、漢語サ変は増えても、和語動詞はなかなか使えないという傾向があるようだ。

### 5.2.3 イ形容詞

4名とも4級が最も多く、3,4級で90%を占めていた。Hに至っては、2級以上は全く出ていない。Iは「詳しい」「蒸し暑い」の2語、Kは「しょうがない」の1語、Tは「詳しい」「秋っぽい」の2語であった。「詳しい」は、I、Tで共通している。Iは日本の夏を「蒸し暑い」と表現した。日本の夏を4名全員に振ってあるが、Iのみ「蒸し暑い」と表現した。Hは日本の夏に対して「湿度が高い」「ハンガリーより暑く感じられる」と表現した。イ形容詞は一語あたりの使用頻度が他の品詞に比べると最も多かった。中上級になるまでに、中級以降のイ形容詞は、勉強してきても、実際に会話で使うことは難しく、名詞、動詞に比べると、伸びなく、貧弱だということがわかった。

### 5.2.4 ナ形容詞

4名とも4級が最も多かった。その割合は、Iが55.6%、Kが55.6%、Tが56.3%と半分以上を占めているのに対して、Hは38.1%と半分以下であった。しかし、Kは、2番目に2級が27.8%を占め、Hは2級と3級が同じ割合となっており、中級以降の語の割合が低い点もイ形容詞と異なり、興味深い。

ここで、1,2級語彙の内容を協力者ごとに表すと、Hは「当たり前」「基本的」「逆」「最高」「重要」「制度的」の5語、Iは「かわいそう」の1語のみ。Kは「おしゃれ」「全体的」「日本的」「無駄」「楽」「活発(1級)」の6語で、Tは「国際的」「楽」の2語であった。接辞「～的」(2級)がついている「ナ形容詞」は、I以外に見られた。5.2.3イ形容詞では、Hは2級のイ形容詞が全く出ておらず、4級が90%を超えていたが、ナ形容詞では協力者間で中級以降の語が最も多く出ていたのは興味深い結果だった。また、Hは「クレージーな」とナ形容詞も使っていた。カタカナ語に「な」を付けると「ナ形容詞」になるという知識があった。今回、他の協力者には「カタカナ語+な」はなかった。

### 5.2.5 副詞

4名とも4級、3級の順に多かったが、Hの2級の割合が29.3%と、他の3名と比べて高いことに注目したい。2級の語数が、Iは0、Kは2、Tが1なのに対して、Hは10というのはかなり多いとはいえるのではないだろうか。下にHの2級語彙の内容を示す。

<Hの2級使用語彙(副詞)>:「意外(と)」「一応」「いつか」「かなり」「結構」「こんなに」「絶対」「ただ」「直接」「どんなに」(計10語)

Hの2級語彙で他の学習者と共通しているのは、Kの「ただ」と「どんなに」のみであった。よって、Hのみが使用中級以降の副詞は8語となる。個人によって伸びる品詞が異なるのは、非常に興味深い。

### 5.3 話題からの分析

今回の調査は、複数のペアで話題を揃えるようにした。しかし、大きな話題は揃えたものの、話題の中身はペアによって異なっている。そこで、もっと話題を狭めた部分での学習者の使用語彙を比較した。その結果を下の表10に示す。ここでは、小話題「大学受験のシステム」の分析結果を取り上げる。この話題を選んだのは、共通の狭い話題の中で、学習者間での使用語彙にかなり違いが出たためである。

表10 話題「大学受験のシステム」で学習者が使用した語彙（名詞と動詞）

	初級の名詞	中級以降の名詞	動詞
H	10語：1番、2番、3番、今、学生、高校大学、どこ、勉強、何	4語：制度、卒業試験、入学試験、ポイント	5語：行く、決める、する、もらう、入る
I	2語：大学、ところ	3語：学費、入学試験、部分	5語：入学する、入る、始まる、払う、決める
K	9語：今、妹、学校、試験、それ、テスト、人、何、私	11語：自己紹介書、実力、受験、書類、大学ごと、推薦、成績、特技、早め、未来、面接	7語：ある、行く、受ける、書く、する、使う、入る
T	3語：一つ、学校、テスト	1語：成績	4語：ある、受ける、違う、決まる

表10の右をみると、動詞は初級の語彙だけで、全く中級以降の語彙が出現していないことがわかる。また、全員に重なっている動詞はないものの、2、3名で重なっている動詞も多い。しかし、名詞は、中級以降の名詞も多く出現しており、学習者で差が出ていることもわかる。特にKとTの差は顕著である。Tは、大学受験のシステムの説明の途中で「は一、難しい」と言っ、必要な語が出てこなく、説明を曖昧に終わらせてしまったため、聞き手によく意味が伝わらなかった。反対にKは、中級以降の語彙を使って、自国の受験制度について十分に説明することができていた。ある話題で、相手に意味を伝えるためには、動詞より名詞のほうがその役割が大きく、必要となってくる。この話題で出現した動詞は初級の語彙だが、名詞は中級以降の語彙が産出していた。同じ話題、同じ相手との対話であるが、特に名詞では個人差が見られた。

### 6. まとめと今後の課題

本研究は、語彙学習が本格化する中上級日本語学習者を対象に、対話における使用語彙について、その実態を明らかにするために、中上級学習者と母語話者との話題を揃えた会

話を資料とした学習者の語彙調査を行った。その結果、以下の点が明らかになった。

- ・学習者の使用語彙における異なり語数の構成は、名詞、動詞、副詞、ナ形容詞、イ形容詞の順に多い。
- ・中級以降で学ぶ語が異なり全体の20%を占め、級外の語まで合わせれば、30%になるという実態から、少なくとも初級の語ばかりで話しているわけではないことが明らかになった。品詞では、名詞が中級以降の語の割合が最も高かった。
- ・中上級日本語学習者の使用語彙に中級以降のイ形容詞はほとんど見られず、イ形容詞が貧弱であることがわかった。ナ形容詞、副詞は中級以降の語が20%ぐらい見られる学習者もいた。個人によって伸びる品詞が異なる可能性が示唆された。
- ・中級以降の動詞では、韓国語母語話者、中国語母語話者は漢語サ変の割合が多いのに対して、非漢字圏学習者は、和語動詞の割合が高い。漢字圏学習者は、漢語に頼る傾向があり、和語動詞が増えていかない可能性がある。
- ・ある話題で、相手に意味を伝えるためには、動詞より名詞のほうがその役割が大きく、中級以降の名詞が必要となってくる。同じ話題、同じ相手との対話であるが、産出する名詞では個人差が見られた。ある話題において必要な中級以降の名詞が使えるかどうか、相手に言いたいことを伝えるための鍵となっていることがわかった。

本研究の調査は、従来の研究とは異なり、同じ話題で同じ母語話者との対話を資料とし、中上級日本語学習者の使用語彙調査を行ったことに意義があるといえる。この調査により、中上級日本語学習者の話し言葉の実態を一側面ではあるが明らかにした。しかし、本研究は小規模である。これから、さらに多くのデータが必要である。そして、同じ条件で日本語母語話者を対象とした使用語彙調査を行い、今回の調査結果と比較する必要があるだろう。

## 謝辞

本稿の執筆にあたり、早稲田大学大学院日本語教育研究科小宮千鶴子教授から、大変丁寧なご指導、貴重なご助言をいただきました。ここに感謝の意を表します。

## 注

1. OPIとは、外国語学習者の会話のタスク達成能力を、一般的な能力基準を参照しながら、対面のインタビュー方式で判定するテストである。
2. 「茶釜：ChaSen」とは形態素解析器であり、入力文を単語単位に分割し品詞を付与するツールである。茶釜システムは、広く自然言語処理研究に資するため無償のソフトウェアとして開発されたものである。

## 参考文献

- 池田伸子（1996）「日本人ビジネスマンの話し言葉における語彙調査—ビジネス用日本語教育システム開発の基礎として—」『日本語教育』88号：117-127
- 宇佐美まゆみ（2005）「改定版：基本的な文字化の原則（Basic Transcription System for Japanese : BTSJ）」
- 金田一京助・佐伯梅友・大石初太郎・野村雅昭<編著>（2011）『新選国語辞典 第九版』小学館
- 見坊豪紀・金田一京助・金田一春彦・柴田武・市川考・飛田良文<編著>（2008）『三省堂国語辞典 第六版』三省堂
- 国際交流基金、日本国際教育支援協会（2002）『日本語能力試験 出題基準（改定版）』凡人社
- 国立国語研究所（1980）『日本人の知識階層における話しことばの実態 資料集』
- 国立国語研究所（1980）『日本人の知識階層における話しことばの実態 調査の概要と分析』
- 国立国語研究所編（1995）『テレビ放送の語彙調査Ⅰ—方法・標本一覧・分析編—』秀英出版
- 国立国語研究所編（1997）『テレビ放送の語彙調査Ⅱ—語彙表—』大日本図書
- 日本語教育学会編（2005）『新版 日本語教育事典』大修館書店
- 林 四郎 監修、篠崎晃一・相澤正夫・大島資生 編著（2012）『例解新国語辞典 第八版』三省堂
- 文化庁編『外国人のための基本語用例辞典 第三版』（1990）大蔵省印刷局
- 牧野成一他（2001）『ACTFL-OPI入門—日本語学習者の「話す力」を客観的に測る—』（アルク）