

語彙文法シラバス主催 公開シンポジウム

シラバス作成を科学にする

—日本語教育に役立つ多面的な語彙シラバスの作成—

2015年2月22日(日) 10:00~16:30

国立国語研究所 2階 講堂

【基調講演】

10:00~12:00 「語彙シラバスと文章理解」……………3
石黒 圭 (一橋大学 教授)

(昼休憩)

【パネル発表】

13:00~13:40 「語彙密度から見た語彙シラバス」……………21
佐野 大樹

13:40~14:20 「外国につながる子どもたちのための語彙シラバス」……………31
中石 ゆうこ (広島市立大学 非常勤講師)
建石 始 (神戸女学院大学 准教授)

(休憩)

14:40~15:20 「子どもを持つ外国人のための語彙シラバス」……………49
森 篤嗣 (帝塚山大学 准教授)

15:20~16:00 「日本語学習者から見た語彙シラバス」……………61
劉 志偉 (首都大学東京 助教)

16:00~16:30 質疑応答

語彙文法シラバス公開シンポジウム基調講演

語彙シラバスと文章理解

石黒 圭（一橋大学国際教育センター教授）

k.ishiguro@r.hit-u.ac.jp

1. はじめに

1.1 語彙シラバスの意味

文章理解の目的は、内容理解である。内容理解のためにもっとも重要なのはトピックである。トピックがわかれば、内容の見当はつく。

内容理解を支えるのは、語彙理解である。語彙が正確にわかれば、トピックもわかるし、トピックを支える内容も正確に理解できる。

語彙は文法とは異なり、均衡コーパスのようにジャンル横断的に調べると、かえってその性格がつかみにくくなる。ジャンルごとに、また分野ごとに細かく見るほうが理解の精度は高まる。

分野にはそれぞれ固有の偏りが見られる一方、近い分野には類似性が見られ、分野が重なると共通語彙が用いられる。そのため、分野の偏りや重なりを反映した語彙シラバスがあれば、文章理解教育がより効率的になる。

学習者はすべての分野に精通する必要はない。自分と関連の強い分野にのみ詳しくければよい。分野ごとに語彙シラバスを構築する意味はここにある。

1.2 社会科学系基礎文献を例に

語彙シラバスを考えるにあたり、社会科学系基礎文献を例に取り、今村・石黒(2012)、今村(2014)、石黒(2014)を参考に、共通語彙についてまずは考えていくことにしたい。

以下の【表】は、大学の教科書として使われる商学の文献、経済学の基礎文献に出てくる二字漢語の高頻度語ベスト 15 である（今村(2014:31)の表 3「社会科学各分野トップ 30」を参考に作成）。ご覧になって、何か気づいた点はあるだろうか。

【表】商学・経済学の基礎文献に出てくる高頻度二字漢語

	1	2	3	4	5	6	7
商学	企業	利益	会計	製品	事業	市場	資産
経済学	経済	生産	消費	労働	成長	市場	価格

8	9	10	11	12	13	14	15
場合	投資	顧客	基準	資本	会社	経営	価値
需要	場合	所得	企業	投資	資本	増加	上昇

まず気づく点は、商学と経済学で出てくる言葉が似ている点である。「市場」が6位で並んでいることがまずは目を惹く。また、「企業」「場合」「投資」「資本」も両方の分野にまたがって出現している。

第二に気づく点は、その本質を表す言葉、すなわちお金に関わる言葉が多い点である。商学の「利益」「会計」「資産」「投資」「資本」、経済の「価格」「所得」「投資」「資本」はすべてお金に関係がある。言うまでもなく、商学、経済学とも、お金と、切っても切り離せない関係にある。漁業関係者が魚の名前を細かく呼び分けるように、経済学者・経営学者もお金の名前を細かく呼び分ける。ほかにも、「通貨」「貨幣」「金融」「利潤」「収入」「費用」「物価」など、リストを広く見ていくと、枚挙にいとまがない。

第三に気づく点は、第二の点ともつながるが、意味の似た類義語が多いことである。たとえば、「資産」「資本」「投資」、それに「資金」を加えた四つはとても似ている。私自身はこの表をもとに社会科学の基礎語彙を学ぶ授業を行っているが、この表を見せると、留学生はその類似にすぐに気づき、その違いを質問してくる。

質問された場合、同じ「資」という漢字が使われていることを指摘し、漢字をよく見てもらい、「買収」の「買」、「財政」の「財」、「貯蓄」の「貯」など、「貝」のつく漢字はお金を表すということにまず気づいてもらう。その「貝」のうえに「次」という漢字がついているので、「次に使うお金」、すなわち元手の意味になると伝える。字源の説明は多少正確さを欠いても、憶えるきっかけになればよいと考える。

そのうえで、「資産」は「産」だから利益を産みだすもの、「資本」は「本」だから事業活動に必要な基金、「資金」は「金」だから目的実現のためのお金だと説明する。「資産」は企業も個人も持つが、「資本」はおもに企業が、「資金」はおもに個人が持つ。「みなさんも留学資金を貯めて日本に来たのでしょうか」と言うと、その意味が明確に理解できるし、「みなさんは結婚資金を貯めていますか」と聞くと、留学生のあいだに笑いが生まれる。

なお、投資は、消費するためではなく、将来の利益を得るために金融商品などを購入することである。これだけが動詞なので、区別は容易であると思われる。

第四に気づく点は、類義語だけでなく、対義語も想定できることである。経済学分野の第二位と第三位を占める「生産—消費」の対は目に見える対義語で、この二つがセットになっていることはすぐにわかる。

また、「利益」にたいする「損失」、「需要」にたいする「供給」、「所得」にたいする「支出」、「増加」にたいする「減少」、「上昇」にたいする「下落」など、一方が表に現れていないものの、対になる言葉を想定することが可能である。このように、想像力を働かせて、見えないところを可視化すると、記憶に関連性が生まれる。

第五に気づく点は、専門語は日常語とは違う言葉が用いられやすいという点である。「企業」は専門語で、「会社」は一般語である。専門語は抽象的かつ集合的であり、一般語は具体的かつ個別的である。その証拠に、「企業に行く」とは言えないが、「会社に行く」と言うことはできる。「企業」は抽象的であるが、会社はオフィス、すなわち建物を持つ実態だからである。反対に、「IT会社」「中小会社」とはあまり言わないが、「IT企業」「中小企業」とは言える。会社は個別的な存在、企業の場合は集合的な存在なので、「IT企業」「中小企業」という結びつきが可能なのだろう。

ただし、商学の第十三位に「会社」が入っている点には注意したいところである。「会社」は、「会社法」という法律があるという意味では専門的な言葉であり、「株式会社」「持ち株会社」「合同会社」などとなった場合は専門語と考えることが可能になるからである。

ほかにも、専門語と一般語の対を、想像力によって補うことができる言葉が並んでいる。「顧客」と「お客」、「労働」と「仕事」、「価格」と「値段」などは専門語と日常語の対になる。「市場」を「しじょう」と読めば専門語、「いちば」と読めば日常語になり、「事業」も「事業者」となれば専門語、「業者」となれば日常語である。「製品」という専門語も、「品物」となれば日常語であるが、「品物」はむしろ「商品」との対とも考えられる。言うまでもなく、「製品」も「商品」も同じ「品物」であるが、「製品」は工場で作られた「品物」で、「商品」は店頭に並んでいる「品物」になる。

商学・経済学の論文・レポートを書くときに、「会社」「お客」「仕事」「値段」「業者」「品物」などを使うと、どこことなく子どもっぽい印象を与えてしまう。日常語は厳密な定義なしに使われる具体的・個別的な言葉であり、論文・レポートのような論理的な文章には不向きである。

以上、見てくるとわかるように、ある分野における高頻度語は、形態的、意味的に関連性のある語彙のネットワークを形成することがわかる。したがって、1語1語別々に憶えるのではなく、その分野の内容や背景に思いをいたしながら、体系的に憶えることが語彙学習のうえで重要であろう。

1.3 語彙シラバスの限界

このように分野共通語彙をまとめた語彙シラバスが有用なことは言うまでもない。しかし、こうした語彙シラバスを整備すれば、文章は読めるようになるのだ

ろうかと考えると、それだけでは不十分であると言わざるをえない。

私は受験生のころ、英単語帳や古文単語帳を片手に語彙を必死に憶えたことがあるが、それで英文や古文が読めるようになった記憶がない。むしろ、「単語は文脈のなかで憶えなさい」と言われ、実際の英文や古文にたくさん触れているうちに、多少なりとも文章が読めるようになったような気がする。

読解というのは言語技能の一つであり、読解という処理過程のなかでは、単語帳のなかで対応する母語の意味を想起するのとはまったく別の処理が行われていると考えられる。

本講演では、その処理の特徴について、「語形」「意味」「機能」という三つの観点から検討する。

「語形」というのは、読んで字のごとく語の形である。単語帳ではすでに語がリストになって掲載されているが、実際の日本語の文章では分かち書きがないため、文章のなかから語を正確に取りだしてくるのも一苦勞である。また、平仮名、片仮名、漢字という字種の別もあるため、その使い分けについても意識を向けなければならない。とくに、日本語学習者の場合、そうした日本語の表記システムになじめず、単語帳的な語彙力があっても、それが読解力に結びつかないことが多い。

「意味」というのも、学習者にとっても教師にとってもなじみのある言葉である。単語帳では、意味がいくつか列挙されていることが多いが、そのなかから文脈に合う意味を瞬時に選ぶのはけっして易しい作業ではない。また、語の周辺的な意味、すなわち語のニュアンスや文化的背景などを汲み取るのも、とくに第二言語の場合難しくなる。たとえば、この時期「義理チョコ」が飛び交うが、その「義理チョコ」の「義理」のニュアンスや、その背後にある社会文化的状況を理解するのは単語帳に盛られた知識だけでは難しい面があろう。

「機能」というのは、いろいろな意味で使われる言葉であるが、ここではテキスト的機能を指す。読解の対象となる文章は、多数の文が連なって構成されている。読み手は、今読んで理解している文だけでなく、前後の文の関係や文章全体の構成などを意識し、理解していなければならない。それは語の理解においてもそうである。そのような語のテキスト的機能を理解して初めて、読解のなかで語を理解したことになる。

本講演では、「語形」「意味」「機能」という三つの観点から文章理解の処理過程における語彙の性格をあぶりだし、日本語教育の教場で語彙シラバスを用いて読解指導をするときの留意点について考える。

2. 語形の把握

2.1 類似語形の把握

以下の表現を見て、一瞬おやっと思わないだろうか。

- (1a) ネガティブ・スピーカーの話し方
- (2a) トイレに載っているポテト
- (3a) 古い小屋に入ったことがある
- (4a) 東海大相撲の野球部
- (5a) まずは働け
- (6a) 今の苦い人達は余り変愛といふものを高調し過ぎる。

似た字形を見ると、誤解をしてしまうことがある。たとえば、「ろ」と「る」のように字形が似ているものは、混同の可能性はある。しかし、「ろく」を「るく」と読み間違える可能性は低い。「ろく」は「六」や「碌」などとして頭のなかの辞書に登録されているのにたいし、「るく」は頭のなかの辞書に適切な語がないからである。一方、「しろ」を「しる」と読み間違える可能性はある。両方とも頭のなかに辞書に登録されている語があるからである。インターネットの掲示板の「カキコしろ」が「カキコしる」になり、「カキコ汁」になるという言葉遊びは、そうしたことが背景にあると考えられる。

上記の(1a)～(6a)は、じつは(1b)～(6b)のように書かれているものを、パッと見た瞬間、字形の類似から(1a)～(6a)の語形として誤解してしまう可能性があるという例である。

- (1b) ネイティブ・スピーカーの話し方
- (2b) トレイに載っているポテト
- (3b) 占い小屋に入ったことがある
- (4b) 東海大相模の野球部
- (5b) まずは動け (2014年4月22日『日本経済新聞』「パナソニック 復活は本物か」見出し)
- (6b) 今の若い人達は余り恋愛といふものを高調し過ぎる。(芥川龍之介「恋愛と夫婦愛とを混同しては不可ぬ」)

2.2 表記の把握

日本語の場合、類似語形の誤解は、漢字仮名交じり文という表記体系と密接に関連する。ここではとくに、片仮名だからこそ起きる誤解を例に挙げ、その原因を考えてみたい。

カタカナ語表記によって起きる誤解の第一のタイプは、外来語を片仮名表記した場合、原音とズレがあるため誤解してしまうものである。以下の例は、桑原・山口(2014)で紹介されている興味深い例である。

- (7) レストラン情報の中の「バイキング」を「ハイキング」と発音し、「外で食べる」と誤解した。
- (8) 「ベッド」を「ペット」と発音し、動物のペットのことだと考えた。

ただし、桑原・山口(2014)によれば、このような誤解は多くなく、片仮名どおり発音できても、それが何を指しているかわからないものが大半だったという。

片仮名表記は、英語母語話者にとっては鬼門である。表記自体が日本語母語話者の発音しやすいようにアレンジされているからである。

カタカナ語表記によって起きる誤解の第二のタイプは、非外来語の片仮名表記である。たとえば、ポクロフスカ(2015 予定)によれば、ウクライナ人中級日本語学習者が文章理解をするさい、「モノ」を「メモ」とする誤りが散見されたという。原因は、中山(1998)などで指摘されているとおり、非外来語の片仮名表記というものにたいし、日本語学習者が、なじみがないということに由来すると思われる。

「モノ」と「メモ」という語形はさほど似ているとは思えないが、片仮名語は外来語という強い意識を日本語学習者が持っているのと、頭のなかの外来語の語彙辞書のなかから「モノ」に近いものを選んでくると、「メモ」になるのだろう。

柏野・中村(2013)によれば、「モノ」は、片仮名で表記される率は1.3%と低いが、絶対数で言えば、非外来語のなかでもっとも多く片仮名表記されている語である。しかし、日本語の文章に触れる頻度がさほど高くない中級日本語学習者の場合、「モノ」という表記をほとんど見かけたことがなかったということが背景にあると考えられる。

最近、朝日新聞について「アサヒ」という表記をよく見かけるようになった。朝日新聞自体、「朝日新聞デジタル」「ジュニア朝日海外電子版」といった情報配信サービスを、「アサヒ・コム」「こどもアサヒデジタル」という名称で行っていたことがあるが、「アサヒ」という表記が、「アサヒビール」「アサヒ芸能」と混同されやすい点、ネトウヨが朝日新聞を揶揄して用いる「アサヒる」のような表記を想起させるため、用いないようになりつつあるようである。このように、表記の選択自体が意味のアクセスに微妙な影響を与えることがある。

カタカナ語表記によって起きる誤解の第三のタイプは、片仮名と漢字の語形の類似性である。丸みを帯びた平仮名とは異なり、片仮名の字形は角ばっており、これが漢字との混同を引き起こす。また、長音記号が漢数字と誤解されるという問題もある。

- (9) 「アニメ版もヒット『妖怪ウォッチ』 あなどれない小学生のロコミカ」
一連のブームの火付け役ともなったアニメ版『妖怪ウォッチ』(テレビ東京系)。アニメ放送を観た子どもたちを夢中にさせる、綿密に練られた仕掛けによって、ゲームや玩具は爆発的な売れ行きを見せ、今や社会現象

にもなっている。子供たちのツボを刺激するポイントを探った。

(<http://www.oricon.co.jp/news/2038059/full/>)

ORICON STYLE 2014.5.31 配信)

非外来語の「ツボ」も誤解を招くおそれがあるが、ここで問題にしたいのは「ロコミカ」である。一瞬「ろこみか」と読んでしまう。最近では、そうした誤解を避けるために「クチコミ」という表記が選ばれることが多いようである。

2.3 分節の把握

最近、インターネットを検索していると、「スポンサードリンク」という語をよく目にする。私はこれを最初「スポンサー・ドリンク」と読んでしまっていたが、そんな飲み物があるわけではなく、「スポンサード・リンク」と読まなければならないことであとで気がついた。

文節の区切りの誤解もまた、表記と深い関係があることが予想される。例は、ポクロフスカ(2015 予定)のものを用いる。

漢字は実質語を担うことが多いため、読解過程において意味を取る場合、漢字を優先的に拾っていく。また、文節の区切りは漢字で始まることが多いため、平仮名から漢字に変わる部分を意味の切れ目として認識することが多い (Kajii et. al. 2001)。そのため、そうした標準的なパターンから逸脱するものについては、誤解を生む可能性が高いことになる。

(10) ひと昔前なら「物持ちがいい」ことにも価値があったのですが、

ポクロフスカ(2015 予定)によれば、この(10)の例を半数以上が「ひと、昔前なら」と読みあげ、「昔の人々は必要なモノとそうでないモノを峻別していた」のように理解していたという。

また、漢字から平仮名に変わる部分にも意味の切れ目を見出そうとしやすい。そのため、そこでも誤解が起きやすい。

(11) これは非常に 面倒くさいことです。

この(11)にたいしては、「非常」という名詞+「に」、「面倒」という名詞+「くさい」と捉えた学習者もいたという。

一方、補助動詞を本動詞として解釈する例も多く、

(12) はい、ステーキは、先程ミスで作りすぎたものがありますので、それを温め直して出します。

を「ステーキを作るときにミスを犯したので、温めることによって、それを直した」と解釈する学習者もいたという。

一方、平仮名や片仮名が連続する場合は、同じ字面の文字が続くことになるため、どこで切ってよいかわからなくなることが多くなる。先ほど例に挙げた「スポンサードリンク」がその例である。

やはり、ポクロフスカ（2015 予定）によれば、

(13) いいことなど、ひとつもないのです。

を読んで、『つもない』ってどういう意味なんだろう」と考えこむ学習者、

(14) あきこが店に入ると、シェフがみずから迎えてくれた。

を読んで、「ここでの『みず』というのは玄関でしょうか」と質問する学習者がそれぞれ観察されたという。

3. 意味の把握

3.1 多義の把握

語の意味は一つにかぎらない。多義語の場合、複数ある意味のうち、いずれを選択するかが問題になる。

(15) 「箱根駅伝“山の神”たちを支える彼女の存在」

それぞれの思いをたすきに背負い、苦悶の表情で走る選手たち。普通の人間には考えられないスピードで長い距離を走るわけだが、そんな選手たちも一歩競技を離れば普通の大学生。意外と言っては大変失礼だが、多くの選手には彼女がいる。

「『あなたの走りには勝ちたいという気持ちが伝わらない』といわれたこともあります」。年上の彼女にバツサリ切り捨てられたことを明かす選手がいれば、「お守りを毎年作ってくれて『応援してるから』といわれて…。頑張らなかつたら男じゃない」なんて選手も。辛い練習を乗り越えるには、温かく時には厳しい？ 彼女の存在が必要なのもかもしれない。

格好いい姿を見せたい。怒られたくない。彼女に対する思いは選手それぞれ。それに加え両親や監督、チームメートの思いを背負って走る若者たち。だからこそ、箱根駅伝にはドラマがあるのもかもしれない。

(<http://www.sanspo.com/sports/news/20141215/ath14121511390009-n1.html>)

サンケイスポーツ 2014.12.15 配信)

「山の神」は、今井正人（順天堂大学）、柏原竜二（東洋大学）、神野大地（青山学院大学）らに冠せられる比喻であり、その比喻の解釈も多義の把握に含まれるが、ここで注目したいのは見出しにある「彼女」である。

彼女は、she の翻訳としての女性代名詞的な用法、そして、付き合っている特定の女性という用法がある。見出しだけ見ると、女性代名詞的な用法と思ってしまうが、記事を読むと、選手と恋仲にある特定の女性であることがわかる。このような試行錯誤の過程に柔軟に対応できるかどうか読解力のカギである。

次のインタビューは、「適当」という語の多義性をうまく活用している。

(16) 「村上春樹は英語的な味わい 主要作品を英訳、ジェイ・ルービンさん講演」

翻訳中に村上本人ともやりとりしたが、原作者としての意見を聞くと、村上は決まって「適当にやってください」と答えたという。「しかし結局、この『適当』こそが翻訳家の仕事を全て言い表しています」とルービン名誉教授。

「翻訳家の仕事は、自分の言葉で、『適当』と思われる手法を使って、できるかぎり原文に近い文学的経験を読者に味わってもらうこと。その過程は翻訳家の主観的なものですが、それによって何千何万という読者に、その小説以外では手の届かない世界をのぞかせることができるのではないかと思います」

(<http://digital.asahi.com/articles/DA3S11473392.html>)

朝日新聞デジタル 2014.11.25 配信)

多義で気をつけたいのが、固有名詞である。海外で学ぶ学習者は概して固有名詞の区別に弱い。固有名詞はその文化のなかで生活しないとなかなか身につかないからである。

(17) 毎日の、毎日が、変わる。(仲畑貴志氏のコピー)

(17)は『毎日新聞』の広告だということで、初めて意味を持つ。最初の「毎日の」を『毎日新聞』という固有名詞だと解釈すると、「毎日新聞の毎日の紙面が変わる」「毎日新聞の紙面が毎日変わる」という意味になり、二度目の「毎日が」を『毎日新聞』という固有名詞だと解釈すると、「私が毎日読んでいる毎日新聞の紙面が変わる」「うちに毎日届く毎日新聞の紙面が変わる」という意味になりそうです。つまり、前者は「日々変わる」という点が、後者は「日々触れている」という点が、それぞれ強調されているのである。

私がふだん授業で使っている読解教科書（石黒編 2011）には以下のような文章が載っている。

(18) 「今春スタートの新ドラマ『いもうと』」

愛と真理は2歳違いの姉妹。両親は、長女の愛は厳しく育てたが、妹の真理は甘やかすことが多かった。やりたいことを何でも自由にさせた結果、ほんぼうな性格に育った真理。一方、愛は真面目で気が弱く、子どものころから真理に困らされてばかり。{中略} 対照的な性格の姉妹の青春を描いた、ほのぼののコメディ。

ところが、この文章を使って留学生に教えたところ、誤解が続出した。「姉妹」とははっきり書いてあるので、「愛」「真理」は固有名詞だということはわかるはずだと考えたのだが、現実には「愛」「真理」がそれぞれ明確な意味を持つため、固有名詞として認識されにくいようなのである。もし日本社会のなかで生活してきていたら、それが人名であることはすぐに気づけたと思われる。

3.2 含意の把握

語には含意がある。その含意を十分に汲み取って初めて筆者の意図が探れることも少なくない。そうした含意についての知識も必要である。

(19) 不正な処理を行ったので強制終了します。

かつての Windows の画面によく現れ、評判の悪かった表現である。「不正な処理」という場合、間違っただけという以上の「倫理的に問題な」という意味を含んでいる。ここに多くのユーザーが不快感を覚えた理由がある。

一時期野球界で話題になったニュースに以下のようなものがある。

(20) 上原浩治 日本と違うメジャーでの中継ぎの役割

レッドソックスの上原浩治投手（38）が中継ぎ投手のプライドを見せたのは3月だった。WBC（ワールド・ベースボール・クラシック）の1次ラウンドで不調だった日本代表のエース、楽天の田中将大投手（24）が中継ぎに回る可能性が浮上し、一部メディアが“中継ぎ降格”と報じたことに上原がツイッターで「中継ぎをバカにするな」と反論したのだ。中継ぎを軽く見るかのような報道に黙っていられなかったのだろうが、田中を思いやる気持ちもあった。

(<http://www.tokyo-sports.co.jp/sports/baseball/151371/>
東京スポーツ 2013.6.11 配信)

ここでは「降格」という言葉がネガティブなニュアンスを帯びており、それが中継ぎを専門にする投手にたいする侮辱につながると、当時クローザーに移るまえで、もっぱらセットアッパーを務めていた上原が考えたわけである。

含意の把握の場合、とくに注意したいのが副詞である。「あいにく」「さすがに」「さっぱり」「せいぜい」「せっかく」「せめて」「たいして」「つい」「どうせ」「むしろ」など、副詞には筆者の気持ちが忍びこむものである。

(21) いちいち胸に刺さる！

(『ぐでたま哲学1 & 2』大和書房の南武線車内広告)

(22) このようにして、今や「地方の時代」は、先進工業社会における新しい政治的、文化的潮流になってきているわけですが、これらに共通して見られることは、人間復興、福祉優先の立場は、地域や自治の尊重につながらざるをえないという点であります。すなわち「福祉の時代」、「人間の時代」は、まさに「地方の時代」でなければならないのであります。

(長洲一二『地方の時代』を求めて)

(21)では、副詞「いちいち」が利いている。「いちいち」は、「いちいちうるさい」のように、一つ一つのことがかんに障るといったネガティブなニュアンスで用いられることが多いが、ここではそれを逆用している。

(22)では、筆者が主張する「地方の時代」を際立たせるために「今や」「まさに」といった副詞が添えられている。

3.3 背景知識の把握

語彙は文章と記憶(知識・体験)とを結ぶインターフェイスである。そのため、語彙の意味がわかっても、記憶のなかに背景知識がないと理解できないことがある。とくに、日本での生活経験がない、あるいは短い学習者の場合、社会・文化的な背景知識が学習者の場合に不足しがちである。

冒頭で示した「義理チョコ」もそうであるし、たとえば「甲子園」という言葉もそうである。もともとは兵庫県の地名であるが、「甲子園球場」の略称として知られ、そこでは春と夏の全国高等学校野球選手権大会が開催される。全国の高校球児がそこで戦うことを目指す、いわば高校野球の聖地であり、新聞社・NHKがこぞって報道するので知名度が高い。そこから、高校生が集まる全国大会は、野球でなくても、甲子園で行われなくても、〇〇甲子園と呼ばれるようになる。たとえば、「クイズ甲子園」「ものづくり甲子園」「かるた甲子園」「ダンス甲子園」「ディベート甲子園」「俳句甲子園」などがある。

(23)では、「ソロバンをはじく」を例に取る。ソロバンはかつては習いごとの定番であったが、パソコンが発達した今ではソロバン教室に通う子どもは多くない。

しかし、慣用句としては今でも残っており、「ソロバンをはじく」は「ソロバンで計算する」、すなわち「損得を考えて行動する」という意味になる。しかし、これもまた、学習者にとっては難しい慣用句だろう。

(23) 成瀬は今年5月に国内FA権を取得も、海外FA権はない。しかし、本人が「メジャーに移籍したい」と言えば、ロッテはOKを出す可能性が高い。

「メジャー移籍なら本人の年俸が現在の1億4400万円から3億円程度にアップするばかりか、ロッテも入札上限の20億円とまではいかないまでも、8億～10億円程度の入札金が得られる。国内FA権を行使されて出ていかれるよりずっといい。メジャー側もマエケンのように大金をかけず、計算できる先発左腕が数億円で獲得できるメリットがある」とは前出の球団幹部。

ロッテは10年に西岡（現阪神）をポスティングでツインズに移籍させた実績もある。当時の入札金は約532万ドル（現在のレートで約5億6000万円）だった。

成瀬とロッテはこれから毎日ソロバンをはじくか。

(<http://www.nikkan-gendai.com/articles/view/sports/154068/2>)

日刊ゲンダイ 2014.10.15)

野球ネタをもう一つ挙げておく。ロッテ・マリーンズのマスコット・キャラクター「マーくん」は元気なカモメの男の子。毒舌で知られる。パシフィック・リーグのロッテに所属する「マーくん」は、セントラル・リーグとの交流戦の時期になると、「マーくんセ界をひとつ飛び！」というポスターをつうじて挑発をすることで知られる。以下はその2014年度版。

(24) 対ヤクルト：今年のツバメはやけに低く飛ぶね。ああ、お天気が心配だなあ…。

(25) 対巨人：最近人気の…なんだっけ…ほら、「惨劇の巨人」？ 面白いよねえ！

(26) 対横浜：あ…流れ星！ 交流戦の優勝をお願いしておかなくちゃ！

(27) 対広島：煮てよし、焼いてよし、お刺身でもよし。鯉ってホントにおいしくて大好き！

(24)は、「ツバメ」がヤクルトを指し、その「ツバメ」が低空飛行をしているとなると、ヤクルトの成績がふるわないこと。また、「ツバメ」は天気が悪くなると低く飛ぶことなどがわからないと、理解できない。

(25)は、マンガ『進撃の巨人』が流行っていること、それを「マーくん」が間違っていて憶えていることを下敷きに、理解が成立する。

(26)は、「横浜」が「ベイスターズ」、すなわち星であること、流れ星に願いをかけると叶うという文化的な背景知識が求められる。

(27)は、「広島」が「カープ」、すなわち鯉であること。また、「マーくん」がカモメであり、魚が好物であること。さらには、ロッテが広島をお得意さまにしていることが理解の前提になっている。

野球の例ばかりでは心許ないので、一般的な例も挙げておきたい。朝日新聞の声の欄からの引用である。

(28) 「(声) 少子化時代の部活選びに悩む主婦」 町田亜紀 (神奈川県 39)

娘が中学に入り、部活選びで悩んだ。人口の多い都市なのに娘の学校は生徒数が少なく、女子は学年で 20 人ほど。バスケ、テニス、吹奏楽、美術の四つしか部がない。

選択肢が多ければ「興味を持ったものやってみたら」と言えるが、球技が苦手な娘に運動部は勧めづらい。「いっそ帰宅部で好きな習い事したら」とも提案したが、「友達はみんな部活をする」「受験にも関わる」と言う。

内申点に響くというのは納得いかないが、本当ならば「やってみたら楽しいかもしれないから」と言うべきか。それとも、「習い事の方が充実するから、1人でも帰宅部にきなさい」と言うべきか。親も子もさんざん悩み、結局、娘は美術部を選んだ。

ああ、少子化の影響がここまできているのか。もっと人数がいれば部活も選べるし、帰宅部の友だちもいるかもしれないのに。夫婦とも経験のないテーマで、久々に感じた子育ての壁だった。

(<http://digital.asahi.com/articles/DA3S11152386.html>)

朝日新聞デジタル 2014.5.24)

次の例は、含意の把握とも関わる例である。認知症の母を持ち、10年を超える介護の日々を漫画に書いてきた作者へのインタビューである。

(29) (インタビュー) 認知症の親の贈り物 漫画家・岡野雄一さん

《10年を超える介護の日々を漫画に描いてきた。おかしくも切ないエピソードが反響を呼び、映画やテレビドラマにもなった。》

これはもう想定外としか言いようがありません。きっかけは編集の仕事をしていたタウン誌の片隅で描いていた8コマ漫画で、もともと飲み屋での失敗談とか、身の回りの面白い話とかをネタにしていた。その漫画も社

長の好意で描かせてもらっていたんです。20歳で上京して漫画家を志したけど、時代は劇画ブーム。僕の丸っこい絵はお呼びじゃなかった。小さな出版社で漫画誌の編集者になり、あげく離婚して息子と2人で長崎へ戻り両親宅に転がり込み、ようやく得た仕事です。日々の業務をこなす褒美に1ページ使っていていいと。

そんな日々のなかで、父が亡くなり、それをきっかけに母も少しずつボケ始めた。みそ汁の味がおかしくなったり、ふらりと出て戻らなくなったり。困ったり心配したり腹が立ったりもするんですが、漫画のネタとしては面白い。そもそも介護という意識もなくてね。認知症という言葉も知らず、「母ちゃんも年取ってボケたなあ」と。時々そんな話も描くようになったんです。

そうしたら、雑誌を持って行った小料理屋のママが、僕の漫画を読んで「うちも母がね」と泣きだしたり、客の男性が「そう、そうなんだよ」ってしみじみ言ってくれたり。反響があるとうれしいから、またボケた母の話を描く。するとまた反響がある。そんなふうにして、母の話が多くなってきた。そうこうするうち、3年近く前に自費出版した本が地元の書店で2カ月連続売り上げ1位になり、評判が広がったんです。

時代に遭遇したんでしょう。ハゲた息子が車いすの親に会いに行く、それだけの漫画なんだから。僕はサボっていただけですが、結果的に良い距離感で母と接することにつながり、読んで癒やされると感じる人がたくさんいた。僕と同じ団塊の世代が今、親を介護している。認知症の介護がいかに切実か、切実な中に救いを求める人がいかに多いか。そういうことだと思います。

(http://digital.asahi.com/articles/DA3S11438543.html?iref=com_rnavi_arank_nr01
朝日新聞デジタル 2014.11.5 配信)

インタビューの上記の部分のなかでは「ボケ」が3回、「認知症」が2回使われている。「ボケ」というと、俗語的であり、馬鹿にしているようなニュアンスを感じ取る人もあるかもしれない。新聞という媒体からすると、「認知症」のような知的で、中立的な言葉のほうがふさわしいと感じる人もいるかもしれない。

しかし、「認知症」という正式な言葉は、厳格で血の通わない印象があるのにたいし、「ボケ」には何とも言えない暖かみがあることがこのインタビューをとおして感じられる。そうした印象を裏づけるように、このインタビューを編集委員の稲垣えみ子は以下のように閉じている。

(30) 「『認知症』って味がないよネ」と岡野さん。確かにそうだ。「ボケた」なら笑い話なのに、認知症となると本人も周囲も「問題」にしてしまい、

老いを受け入れて日々を味わうことを忘れる。岡野さんの描く豊かな世界を読むと、そのもったいなさを痛感する。老後はかけがえのない時間だ。いい加減であることの大切さよ。

4. 機能の把握

4.1 結束性の把握

文章理解に必要な語彙量は膨大であり、かつ、文章ごとの特徴も多様である。語彙不足をカバーするには語彙を憶えるだけでなく、テキスト的ストラテジーも合わせて身につける必要がある。

テキスト的ストラテジーの代表的なものとして、文章のトピックに関わる語、具体的には、指示詞、反復語句、言い換え語句などを追っていくというものがある(砂川(2005)、馬場(2006)など)。

次の(31)は、よく国語の教科書に取り上げられる大岡信氏の「言葉の力」である。この文章のトピックである桜の皮から取れる美しいピンク色が、さまざまな表現に言い換えられている。

(31) 京都の嵯峨に住む染色家志村ふくみさんの仕事場で話していた折、志村さんがなんとも美しい桜色に染まった糸で織った着物を見せてくれた。そのピンクは、淡いようであり、しかも燃えるような強さを内に秘め、はなやかでしかも深く落ち着いている色だった。その美しさは目と心を吸いこむように感じられた。「この色は何から取り出したんですか」。「桜からです」と志村さんは答えた。素人の気安さで、私はすぐに桜の花びらを煮詰めて色を取出したものだろうと思った。実際はこれは桜の皮から取出した色なのだった。あの黒っぽいゴツゴツした桂の皮からこの美しいピンクの色がとれるのだという。志村さんは続けてこう教えてくれた。この桜色は、一年中どの季節でもとれるわけではない。桜の花が咲く直前のころ、山の桜の皮をもらってきて染めると、こんな、上気したような、えもいわれぬ色が取出せるのだ、と。(大岡信「言葉の力」)

4.2 文脈の把握

語の意味は文脈によってさまざまな色を帯びる。文脈によって帯びる臨時的な意味(外延的意味など)を認識できなければ、正確な文章理解は難しい。

次の(32)は、椎名誠氏のエッセイである。「オトナ」という語がこの文章のトピックであって、連鎖している。しかし、「オトナ」には2種類の意味があって、それを文脈におうじて読み分ける力が問われている。

(32) 今年の新年の集英社の広告に、田辺聖子さんの「本を読まへんオトナが

増えた。だからコドモみたいな国になってしもた」というメッセージがあり、それに激しく心を動かされた、という話を何人かに聞いた。ぼくもその一人だった。本当にそうなんだろうなあ、と実感を持って思ったからだ。

ぼくは今モノカキなんていう仕事をしているが、本を読んでいなかったらとてもこんな仕事はできなかつたらと思う。

この数年「荒れる成人式」が話題になる。あのニュースなど見ると 20 歳のオトナ（成人）なんてまるでコドモなんだなあ、とこれも実感を持って思う。ぼくも自分が 20 歳になったときのことを思い出すからだ。

20 歳がオトナならなんとオレはどうしようもないオトナになってしまったんだろう、とそのとき自分に対して思ったものだ。学校の成績は常に最下位だったし、恋人のカケラもいなかったし、個人的に大変荒れてもいたので毎日ヒトを殴ることしか考えていないコドモだった。無免許暴走運転をして事故をおこし、脳内出血して一月間絶対安静の入院生活をした。退院してリハビリの時に暇のあまり本を読むようになった。それがきっかけで読書好きになり、それからは徹底的にいろんな本を読んできた。

文を書く能力を得たのはその頃からの読書の蓄積によるものだろうと思う。たぶん潜在的に文を書く要素が少しあって、それが読書によって無自覚のうちに鍛えられていったのだろうと思う。だからあの頃あのまま本を読まずにいたら、と考えると実にまったく「ゾッ」とする。何かろくでもない罪を犯していたか、ひょっとすると死んでいたかもしれないと思う。その意味ではコドモみたいなオトナのまま無謀な事故をおこしたのはホメられないけれど、そのリハビリ期間に読書の習慣がついたのはよかったのかなあなどとも思う。ヘンに一所懸命勉強をしてこなくてよかつたなあ、とも思う。

話は全然変わるけれど、子どもの眼鏡についてである。これまで世界で日本人が断然多いなあと何時も強烈に感じていた眼鏡をかけた（たぶん近視の）子どもが、近頃は韓国人や中国人にやたらに増えているのを旅の折々確認している。この両国はいま日本以上に親が子どもの教育に熱心で、激的な勉強競争の時代になっているらしい。子どもらは勉強が忙しくて本などとても読んでいられないそうだ。

しかし勉強のしすぎは、人間としてなにかとても大事なものが「摩耗」していくような気がする。その「磨耗」していくものが「オトナ」であるためのいくつかの大事な基本的要素というようなものではないのかなとひそかに思っている。ぼくはまったく勉強しなかつたのでまだ脳細胞に隙間がいっぱいあって、そこにオトナが書いた本の知識や見識がどすどす入ってきて、まあなんとかオトナのようなものになれたような気がするのだ。

（椎名誠「コドモのような」『週刊金曜日』2004.3.12 風速計）

4.3 文章構造の把握

テキスト全体の構造（対比など）や位置（冒頭と結末）などとも関連づけて語を把握できると読み取りが正確になる。

ある語に出会ったとき、文章を読みはじめるときと読みおえるときとで印象がずいぶん違うように感じられるときがある。それは、文章を読むという体験をとおして内容理解が深まり、冒頭での理解とはまったく異なる背景で結末を読むことが可能になるからだと考えられる。

(33)は、斎藤隆介氏の童話『モチモチの木』の冒頭と結末である。豆太は、この物語をとおして大きな成長を遂げる。臆病な弱虫で、夜中には一人でせっちゃんにも行けなかった豆太は、腹痛で苦しむじさまの姿を見て、勇気を振り絞って暗い夜道を一人で走り、医者に呼びに行くのである。そうした描写が以下の{中略}に入る。

(33) 全く、豆太ほどおくびょうなやつはない。もう五つにもなったんだから、夜中に、一人でせっちゃんぐらい行けたっていい。

ところが、豆太は、せっちゃんは表にあるし、表には大きなモチモチの木がつつ立っていて、空いっぱいのかみの毛をバサバサとふるって、両手を「わあっ。」とあげるからって、夜中には、じさまについてってもらわないと、一人じゃしょうべんもできないのだ。

じさまは、ぐっすりねむっている真夜中に、豆太が「じさまあ。」って、どんなに小さい声で言っても、「しょんべんか。」と、すぐ目をさましてくれる。いっしょにねている一まいしかないふとんを、ぬらされちまうよりっからなあ

{中略}

——それでも、豆太は、じさまが元気になると、そのばんから、「じさまあ。」

と、しょんべんにじさまを起こしたとき。

したがって、結末の豆太は、もう臆病な弱虫ではない。大好きなじさまが元気になってくれたことに安堵し、甘えているだけなのである。その意味で、冒頭に出てくる「じさまあ。」と結末に出てくる「じさまあ。」は、同じ語形であっても、読みこまれる意味はまったく異なる。

そうした読み取りは文章理解のなかで初めて生まれてくるものであり、語彙リストをいくら眺めていても出てくるものではないのである。

5. おわりに

文章に出てくる語彙を効率的かつ効果的に学ぼうと、語彙シラバスは間違いなく有用である。しかし、現実の文章を理解するときは、語彙シラバスでは扱われていない運用上の知識や技術が必要である。

本講演では、そうした語彙の運用上の知識や技術について「語形の把握」「意味の把握」「機能の把握」という三つの観点から考え、「語形の把握」については「類似語形の把握」「表記の把握」「分節の把握」、「意味の把握」については「多義の把握」「含意の把握」「背景の把握」、「機能の把握」については「結束性の把握」「文脈の把握」「文章構造」に分け、それぞれ具体的な例を挙げながら論じた。

整理の仕方には再検討が必要であろうが、語彙シラバスを考える場合、読解にかぎらず人間が言語をオンラインの過程で実際に処理するメカニズムも合わせて考えないと、教育現場のなかで語彙シラバスを生かすことは難しいと考える。本講演で示した上記の整理がその一助となればさいわいである。

参考文献

- 石黒圭(2014)「感度を高める言葉の教育(2) 学術専門語の特徴」『指導と評価』60-5、pp.37-39、図書文化社
- 石黒圭編著、熊田道子、筒井千絵、Olga Pokrovska、山田裕美子著(2011)『留学生のための 読解トレーニング —読む力がアップする 15 のポイント—』凡人社
- 今村和宏(2014)「社会科学系基礎文献における分野別語彙、共通語彙、学術共通語彙の特定一定量の基準と教育現場の視点の統合」『専門日本語教育研究』16、pp.29-36
- 今村和宏・石黒圭(2012)「社会科学系基礎文献における語彙と表現—頻度と共起の様相—」『第14回専門日本語教育学会研究討論会誌』pp.15-16
- 柏野和佳子・中村壮範(2013)「現代日本語書き言葉における非外来語のカタカナ表記事情」『第4回コーパス日本語学ワークショップ予稿集』pp.285-290
- 桑原陽子・山口美佳(2014)「中国語系初級日本語学習者がホテル検索サイトを読むときの困難点」『国立国語研究所論集』8、pp.109-127
- 砂川有里子(2005)『文法と談話の接点—日本語の談話における主題展開機能の研究—』くろしお出版
- 中山恵利子(1998)「非外来語の片仮名表記」『日本語教育』96、pp.61-72
- 馬場俊臣(2006)『日本語の文接続表現—指示・接続・反復—』おうふう
- ポクロフスカ、オーリガ(2015 予定)「ウクライナ人中級日本語学習者の読解における文字列分節の課題」『一橋日本語教育研究』3、ココ出版

語彙密度から見た語彙シラバス

佐野 大樹

1. はじめに

語彙シラバス作成で重要なのは、シラバスに含める語彙を選択する際にどのような選定基準・指標を用いるかということであろう。例えば、「弱々しい」「脆弱」という同義関係にある表現をどのようにシラバスに組み込むか考慮する際に、前者は初級学習者向け、後者は上級者向けシラバスに含めることが経験的に妥当だと考えられる。しかし、なぜこのように語を選定すべきと考えるのか、そう判定する基準が常に明確でない場合もあり、裏付けとなる論拠やデータが存在しないような場合もある。このような状況を改善するため、コーパスを用いて語彙の使用傾向を計量的に把握し、より実証的なデータに基づいて語彙シラバスを作成しようという試みが近年増えてきている。

本発表ではこのような試みのひとつとして、語彙密度 (lexical density) を用いる方法を提案する。語彙密度は従来語彙の使用傾向を把握するためでなくテキストの「書き言葉らしさ」・「話し言葉らしさ」あらゆる特徴量として提案されたものである。それを本研究では、話し言葉らしいコンテキストで利用される傾向が強い語彙と書き言葉らしいコンテキストで利用される傾向が強い語彙とを判別するための指標を作成するために利用する。具体的には、代表性のあるコーパスに含まれるサンプルの語彙密度を計測し、これを用いて、語が使用されるテキスト群の語彙密度の平均値を計測したデータを作成した。例えば、上述した「弱々しい」の語彙密度平均値は 3.4、「脆弱」は 6.8 となる。このデータを用いることで、図 1 のように、語彙密度平均値が高い語から低い語をクラス分け・順位づけすることができるようになる。

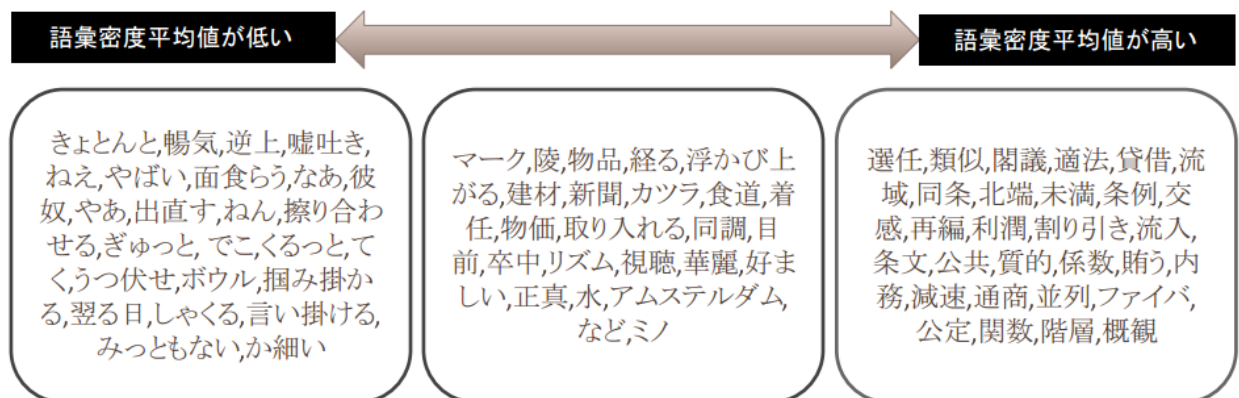


図 1: 語の語彙密度平均値

以下まず、語彙密度とは何かを概説し、次に、語彙の使用傾向をあらわす指標となるデータを作成するために語彙密度をどのように利用できるか、また、代表性のあるコーパスを用いてどのように語彙密度平均値を求められるか説明する。最後に、語彙シラバスを作成する上で、語彙密度平均値をどのように利用できるか、ひとつ事例を示す。なお詳細は後述するが、本発表では「書き言葉らしさ」とは情報の埋め込みの程度が高い (lexically dense) テキストのことを、「話し言葉らしさ」とは情報の埋め込みの程度が低いテキストのことをさすものとする。

2. 語彙密度とは？

語彙密度は、内容語 (lexical item) が節中にどれだけ埋め込まれているかによって、テキストにおける情報の密度をあらわす指標である (Halliday 1990)。たとえば、以下のふたつの文の語彙密度について考えてみる。下線が引かれているものが内容語、引かれていないものが機能語、「|」は語の境界、「<<>>」が節境界を示す¹。

- 1) <<アルコール₁を|たくさん₂|飲ん₃|で|>>c₁、<<|運転₄|する|と|>>c₂、<<|事故₅|を|起こす₆|>>c₃
- 2) <<|過度₁|の|アルコール₂|摂取|は₃、|自動₄|車|事故₅|の|原因₆|だ|>>c₁

1)も 2)も内容語を 6 件含む文であるが、1)は節数 3 件であるのに対し、2)は節数 1 件である。1)と 2) の語彙密度は図 2 のようにイメージできる。1)は節という 3 つの容器に 6 つの内容語が分散して入っているが、2)は 1 つの容器に 6 つの内容語が詰め込まれている。したがって、2)のほうが 1)に比べて語彙密度が高いことになる。語彙密度は、このような意味で情報の埋め込みの程度をあらわすものと位置づけられている。

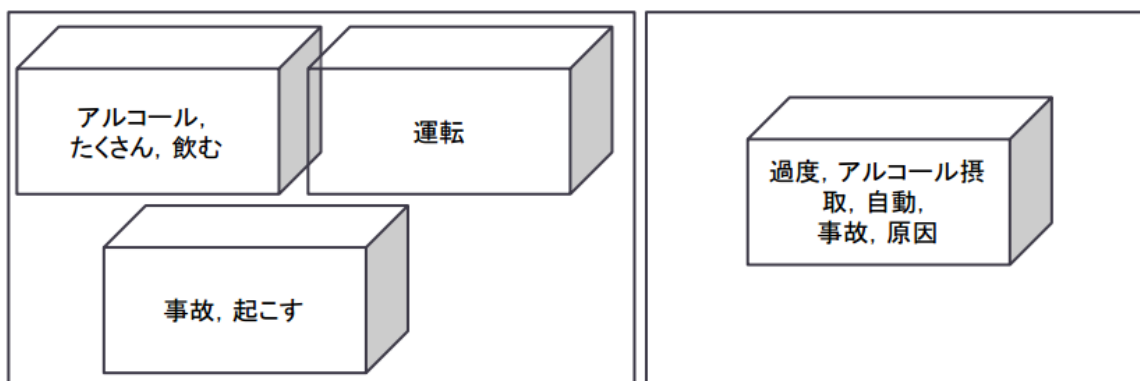


図 2: 内容語・節と情報の埋め込み

¹ なお、接頭辞、接尾辞は内容語の数にカウントしていない。

テキストの語彙密度を計算する際には、以下の式を用いて、1節に含まれる節の平均内容語数をその語彙密度とする。文の語彙密度であれば、1)の語彙密度は内容語が6件、節数が3件であるから2.0、2)の語彙密度は内容語数が6件、節数が1件であるから6.0となる。

$$\text{語彙密度} = \frac{\text{テキストに含まれる内容語の総数}}{\text{テキストに含まれる節 (ranking clause) の総数}}$$

なお、一般的に、話し言葉に比べて書き言葉のほうが動詞の名詞化などの文法的比喻 (grammatical metaphor) が使用される傾向が強いため、語彙密度が高くなると考えられている (Halliday, 1990)。たとえば、先ほど上述した例 1) と 2) は、ほぼ同じ内容をあらかわす文であるが、1) では述部として用いられている「飲む」が 2) では「摂取」、「起こす」が「原因」と置き換えられており、これらの動詞の名詞化によって節数を3件から1件に減らすことで、語彙密度が大幅に変わっている。

このような性質から、語彙密度の計測は患者に配布する医療パンフレットに情報が埋め込まれ過ぎていないかチェックするのに活用されている (Clerehai 2005)。また、アカデミックエッセイで、評価が高かったテキスト群と、評価が低かった群の語彙密度の平均値には有意差が認められることなども統計的に確認されている (佐野 2010)。日本語に関する研究では、『現代日本語書き言葉コーパス領域内公開データ』 (2008年度版) を用いて、NDC 別、出版年別、販売対称別に、語彙密度にどのような違いが認められるか調査した研究もある (佐野ほか, 2009)。例えば、図 3 は佐野ほか (2009) が C コード (図書コード) 販売対象別に語彙密度の分布を箱髭図であらわしたものである。カテゴリ別に語彙密度の中央値には違いがあり、一般的には語彙密度が高くなる傾向があるカテゴリと低くなるカテゴリがあることを示すが、同じカテゴリ内であってもサンプルごとに語彙密度が大きく異なることもあらわしている。特に、「専門」に属するサンプルは同じカテゴリに属するサンプル間での差が大きいことがわかる。

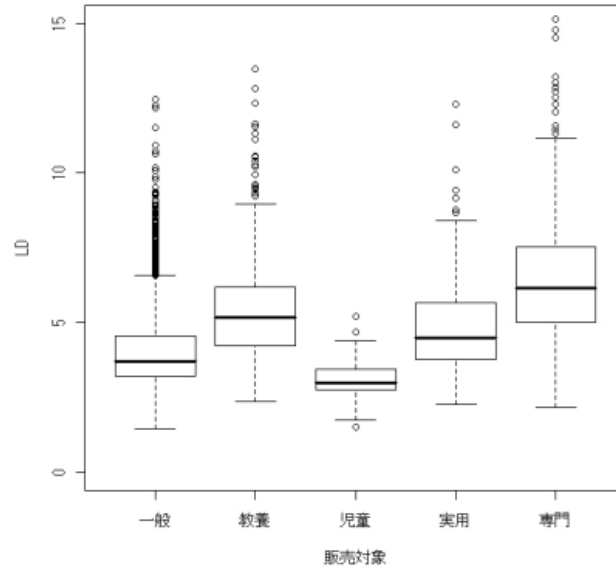


図 3: 販売対象別の語彙密度からみたサンプルの分布 (佐野ほか, 2009 より)

3. 語の語彙密度平均値の計測

このように語彙密度はテキストにおける情報の埋め込みの程度をあらわすもので、本研究ではこれを語の使用傾向をあらわす指標のひとつとして用いることを提案する。具体的には、『現代日本語書き言葉均衡コーパス』(BCCWJ)の出版サブコーパス(図書)に含まれるサンプルの語彙密度を計測し、特定の語が含まれているサンプル群の語彙密度の平均値を求めることで、当該の語がどのような語彙密度を持つテキストで使用される傾向があるかをあらわすデータを作成した。以下、詳細を示す。

語の語彙密度平均値を求めることで語の使用傾向(の一側面)を捉えられると考えたのは、ある語(W_a)が他の語(W_b)に比べて語彙密度が高いテキストで利用される傾向があるのであれば、 W_a は W_b に比べて情報の埋め込みの程度が強い(書き言葉らしい)テキストで使用される語であり、 W_b は W_a に比べて情報の埋め込みの程度が低い(話し言葉らしい)テキストで使用される語になると考えたためである。

次に、コーパスの選択についてだが、BCCWJは、出版サブコーパス(約3,500万語)、図書館サブコーパス(約3,000万語)、特定目的サブコーパス(約3,500万語)以上3つのサブコーパスを含む(前川・山崎, 2009)。出版サブコーパスは、2001年から2005年に国内で出版されたすべての書籍・雑誌・新聞に含まれる文字の総体を母集団として構築されたコーパスである。一方、図書館サブコーパスは、書き言葉の流通に着目したもので、東京都の公立図書館13自治体以上で所蔵している書籍を母集団として構築されている。特定目的サブコーパスには、白書、教科書、ベストセラー、国会会議録などから抽出されたサンプルが収録されている。本研究では、一定以上、流通していることが保証されているテキストにあらわれる語を語彙密度平均値計測のターゲットにできるように、図書館サブコーパスを用いることにした。

また、BCCWJには「固定長サンプル」と「可変長サンプル」という2種類のサンプルが用意されており、固定長サンプルは長さが約1,000文字に固定されており、統計的処理に適したサンプルとして設計されている。一方、可変長サンプルは、長さでなく文章のまとまりを重視したサンプルで、テキスト構造や文脈、文体調査などに適したサンプルとして設計されている。本研究では、統計処理を行うので、固定長サンプルを用いている。なお、サンプル中の節数が5件以下のもの（e.g. 名詞が羅列されているもの）35/10,551サンプル（0.3%）は、語彙密度の計測に利用しなかった。また、10件以下のサンプルにしかあらわれない語は、意味のある統計量を得られないと考え、語彙密度平均値の計測対象から外した。

内容語数の特定には、『現代日本語書き言葉均衡コーパス』（DVD版）に付与されている品詞情報を用いた。厳密には、語が内容語か機能語かを区別するには品詞情報だけでは十分でなく語が使用される文脈を踏まえて判断する必要があるが、本手法では多量のテキストを処理するために、以下に該当するものを内容語とみなした。

表 1: 品詞と内容語の特定

内容語として計測するもの	内容語として計測しないもの
名詞-普通名詞-副詞可能, 形容詞-一般, 副, 動詞-一般, 名詞-固有名詞, 名詞-普通名詞-サ変可能, 名詞-普通名詞-形状詞可能, 形状詞-一般, 名詞-普通名詞-一般, 感動詞-一般, 形状詞-タリ	助詞-格助詞, 助詞-係助詞, 助詞-準体助詞, 助詞-接続助詞, 助動詞, 形容詞-非自立可能, 形状詞-助動詞語幹, 動詞-非自立可能, 助詞-副助詞, 名詞-数詞, 助詞-終助詞, 連体詞, 代名詞, 接続詞, 名詞-助動詞語幹, 感動詞-フィラー, 名詞-普通名詞-サ変形状詞可能

この判定は統計的情報に基づきなされたもので佐野ほか (2009)に詳細が記載されているが、簡単に説明すると、あるコーパスにおいて当該の品詞に該当する語の出現頻度が多くなるのに比例してその異なり語数が多くなるようなものは内容語とみなし、反対に、出現頻度が多くなっても異なり語数が増えていかないような品詞は対象外とみなしている。

節数の計測には、節境界認定プログラム CBAP(Clause Boundary Annotation Program) (丸山ほか, 2004)を用いた。CBAPを用いることで、以下のように節境界を認定し、かつ、節の種類を分類できる。

<<アルコールをたくさん飲んで>> (テ節) <<運転すると>> (条件説ト) 事故を起こす。
(文末)

ただし、CBAP で認定される節全てを節数には含めていない。基本的には、以下のような名詞句の一部として機能できるもの、節に該当しないものは節数の計測対象から除外している。

連体節-形式名詞，形容詞連体節，談話標識，主題ハ

一方で、以下のような節は、節数の計測対象に含めている。

連用節，並列節ケレドモ，条件節カギリ，譲歩節テモ，理由節ノデ

以上の方法で、図書館サブコーパスに含まれるサンプルの語彙密度を計測し、当該コーパスに含まれる語の語彙密度平均値を求めた。また、平均値に加えて、語彙密度の最小値、最大値、上側ヒンジ、下側ヒンジ、中央値、分散も求めた。統計量の計算には R (<http://www.R-project.org/>) を用いた。

4. 語彙密度平均値データとその活用

4.1 語彙密度平均値データ

3 節にて説明した方法を用いた結果、16,082 語の語彙密度平均値を計測することができた。結果の一部を図 4 に示す。なお、図に記載した例は特に品詞等を限定していないが、今回利用した語には既に品詞情報が付与されているので、特定の品詞に限定してリストを作成することも可能である。

語彙密度平均6以上

語	最小値	下側ヒンジ	中央値	上側ヒンジ	最大値	分散	平均値
導体	3.51	5.22	6.04	10.1	13.4	12.6	7.54
輸出入	4.29	5.15	6.37	7.21	13.4	5.87	6.62
フロー	3.58	4.28	5.55	7.47	15	8.92	6.45
調達	2.68	4.27	5.56	7.02	25.6	13.3	6.36
赤外	2.96	3.3	4.45	4.84	19.8	27.7	6.21

語彙密度平均4～5

語	最小値	下側ヒンジ	中央値	上側ヒンジ	最大値	分散	平均値
衛星	2.91	4	4.67	5.5	10	2.29	4.95
池	1.79	3.29	3.89	5.12	25	9.87	4.69
発売	2.42	3.76	4.25	5.08	17.9	4.08	4.65
人情	1.67	3.44	3.95	4.46	8.26	1.91	4.17
町角	2.45	3.39	3.82	4.97	6.02	1.04	4.12

語彙密度平均3以下

語	最小値	下側ヒンジ	中央値	上側ヒンジ	最大値	分散	平均値
暢気	1.72	2.32	2.75	3.48	4.7	0.66	2.97
しゃくる	2.3	2.56	2.89	3.18	3.87	0.22	2.94
なあ	1.69	2.44	2.8	3.3	4.79	0.37	2.88
ねえ	1.26	2.45	2.78	3.1	4.68	0.37	2.82
挽き肉	2.32	2.49	2.7	2.87	3.2	0.08	2.71

図 4:語と 語彙密度平均値の例

4.2 語彙密度平均値データの利用—特徴語リストとの併用

語彙密度平均値データの利用方法のひとつとして、特徴語リストとの併用が考えられる。特徴語リストとは、調査対象とするテキスト群を収録したコーパスと参照コーパスにおける語の出現頻度の違いを用いて調査対象データに特徴的な語を抽出したもので、BCCWJ など、コーパスを活用できる環境が整ったこともあり、近年、様々な特徴語リストが作成されている。

ここでは一例として、日本語学習者、特に学校に通う児童・生徒向けの語彙シラバスを作成する場合を想定して、国立国語研究所が提供している『教科特徴語リスト²』と語彙密度平均値データを併用することでどのようなことが可能になるか議論する。

学校に通う児童・生徒向けの語彙シラバスを作成する上で重要な事項のひとつは、授業と日常生活で異なる意味をもつ語彙を扱うことであると考えられる。ただし、そのような語彙のリストを人手で作成するには、授業科目と日本語教育の両方に精通した知識が要求され、また、候補となる語彙を挙げることも自体も難しい作業となる。

このような「特殊」な語彙リストも特徴語リストと語彙密度平均値データを組み合わせることで作成できる可能性がある。特徴語リストだけでは、授業科目と日常的な場面で語義が異なる語彙を抽出することは難しいが、特徴語リストに含まれ、かつ、語彙密

² http://www.ninjal.ac.jp/corpus_center/bccwj/data-files/frequency-list/Domain-Specific-Vocabulary.zip

度平均値が低い語は「話し言葉らしいテキストで利用される傾向が強いという使用傾向、つまり、日常生活で利用される傾向がある一方で、科目に特徴的な語彙」でもあることから、同じ語であっても異なる二つの文脈で利用される可能性が高い語彙であると考えられるためである。例として「教科特徴語_中_理科.xls」データから語彙密度平均値下位 10 件の動詞を表 2 に示す。「熱する」や「溶ける」など理科でも日常生活でも同じ語義をもつものもあるが、「働く」（仕事をする vs.機能する）や「伝わる」（話が広がる vs.熱や音が移動する）、「保つ」（受状態が変わらない vs.エネルギーが逃げない）、「持つ」（手にする vs.性質を有する）など、理科の授業と日常生活では異なる語義を表す可能性がある語が含まれる。語彙密度平均値だけでは、科目において特殊な語義をもつ語のみを含む語彙リストを作成できなかったが、語彙シラバスを作成する際に収録語彙を選定する際には有用なデータとなると考えられる。

語彙素	特徴度	語彙密度平均値
熱する	146.21	2.9
溶ける	467.21	3.8
流れる	301.94	4.0
<u>働く</u>	325.09	4.1
<u>伝わる</u>	120.62	4.1
作る	306.76	4.2
考える	121.53	4.2
起こる	480.78	4.3
<u>保つ</u>	129	4.3
<u>持つ</u>	234	4.3

表 2: 語彙密度平均値下位 10 件、理科の特徴語の動詞

特徴度が高く語彙密度平均値が低い語彙のどの程度が当該科目にて「特殊」な意味を持ち得る語かを明らかにするためには、さらにそれぞれ語が使用されている文脈も含めて細かく調査していく必要があるものの、このことは語彙密度平均値という特徴量を既存手法と補完的に用いることで、語彙シラバス作成に新しい観点を提供できる可能性があることを示唆するものであると考える。

5. まとめと今後の展望

本発表では、コーパスを用いて語彙の使用傾向を計量的に把握し、これをシラバスに含める語彙を選択する際に利用する試みのひとつとして、語彙密度平均値データを作成した。語彙密度平均値という特徴量を既存手法と補完的に用いることで、語彙シラバス

作成に新しい観点を提供できる可能性がある。今後は、語彙密度平均値データの公開方法、語彙密度平均値を用いて難解な日本語をやさしい日本語に言い換える方法について検討していければと考えている。

参考文献

- Clerehugh, R., R. Buchbinder and J. Moodie. (2005) "A Linguistic Framework for Assessing the Quality of Written Patient Information: Its Use in Assessing Methotrexate Information for Rheumatoid Arthritis." *Health Education Research*, 20.3(2005):334-344.
- Halliday, M. A. K. (1990) *Some Grammatical Problems in Scientific English*. *Annual Review of Applied Linguistics*. 6: 13-37.
- 佐野大樹 (2010). 「学年によるエッセイの評価と語彙密度の関係の違い —British Academic Written English コーパスを用いた一考察—」 『英米文化』 40, 215-227.
- 佐野ほか (2009). 『語彙密度を利用した『現代日本語書き言葉均衡コーパス』テキスト分類の試み』 特定領域研究「日本語コーパス」平成 20 年度研究成果報告書.
http://www.ninjal.ac.jp/corpus_center/bccwj/doc/report/JC-D-08-02.pdf
- 前川喜久雄・山崎誠 (2009) 「現代日本語書き言葉均衡コーパス」 『国文学解釈と鑑賞』 74 号, 15–25.
- 丸山岳彦・柏岡秀紀・熊野正・田中英輝 (2004) 「日本語節境界検出プログラム CBAP の開発と評価」 『自然言語処理』 11(3), 39-68.
- 丸山岳彦・山崎誠・柏野和佳子・佐野大樹・秋元祐哉・稲益佐知子・田中弥生・大矢内夢子 (2011a). 『『現代日本語書き言葉均衡コーパス』におけるサンプリングの原理と運用』, 特定領域研究「日本語コーパス」平成 22 年度研究成果報告書 (JC-D-10-01), 特定領域研究「日本語コーパス」データ班.

外国につながる子どもたちのための語彙シラバス

中石 ゆうこ (広島市立大学)

建石 始 (神戸女学院大学)

1. はじめに

グローバル化の現在、日本に就労、結婚などの事情で来日する人々の増加によって、学校現場でも外国につながる児童、生徒数が増加している¹。彼らにとって日本での学校生活は、教科学習の場であると同時に第二言語としての日本語の習得の場でもある。日本語を母語とするモノリンガルの子どもたちがそれまでの日常生活において当たり前習得してきた語彙や表現もまた、彼らにとっては新しく学ぶ事柄となる。学校教育の現場において二重の学習が必要なことによって、外国につながる子どもたちは、教科の内容以外にも授業でのつまずきが起こりやすく、そのつまずきへの対策を行うことは教師にとって喫緊の課題である。

学習言語の習得において、バトラー(2011:68)の言葉を借りれば、語彙は要であると言える。そこで、現場の教師が指導の際に配慮すべき語彙の指針として、本研究では、「外国につながる子どもたちのための語彙シラバス」を作成する。その際の語彙選択の根拠となる客観的なデータとして、実際の教育現場で指導された語の記録を用いる。その際、小学生は発達の差が大きいので、学年を区切る必要があると考え、対象を3年生に限定する。この学年は、発達的に具体的思考と抽象的思考の境目の時期となり、学習についていくことが難しくなると考えられている時期である。

本研究の目指すところは、外国人児童・生徒にとって指導が必要となる語彙リストを作成することで、現場の教師が事前に丁寧に説明すべき語彙として理解しておく内容の明示化をすることである。そのために本研究は、実際に指導が必要であった語を抜き出し、指導が必要な語彙リストの試案を作成した上で、さらに教科書コーパス語彙表および先行研究の語彙リストと比較し、外国につながる子どもたちにとって重要な語を絞り込む。また、リストの語に共通する特徴を明らかにすることで、現場の教師にとって指導の難しい語の傾向を知り、それに対して丁寧に手当をすることを可能にしたいと考えている。

2. 先行研究

2-1 外国につながる子どもたち

本研究では川上(2008)に従い、「外国につながる子どもたち」を「外国籍の児童・生徒、あるいは日本国籍であるが、家庭において多言語環境で育つ児童・生徒」と定義する。その特徴として、移動が頻繁であること、学習状況が十人十色であることの二つが挙げられる。前者の特徴に関しては、空間的移動だけではなく、言語間の移動を繰り返す(川上2008:90, 100)と指摘される。物理的な移動の多さを裏付けるように、外国につながる子どもたちは、学期中も五月雨式に日本語クラスに編入(バトラー2011:18)してくる。

¹ 2012年(平成24年)5月1日現在、公立の小学校、中学校、高等学校、中等教育学校及び特別支援学校に在籍する日本語指導が必要な外国人児童生徒は27,013人に上る。(文部科学省2013)

これに加えて、後者の特徴の表わす通り、言語使用実態、学習状況が個々で異なる（バトラー2011:23）ことから、現場の教師が直面する指導の困難さを窺い知ることができる。

2-2 語彙リスト

2-2-1 教科書コーパス語彙表

本研究で用いる教科書コーパス語彙表は、国立国語研究所の特定領域研究「日本語コーパス」言語政策班が構築した「教科書コーパス」に出現する語彙に関するデータを収録したものであり、2011年に作成されている²。具体的には、2005年度に使用された、小学校・中学校・高等学校の全学年・全教科の教科書1種ずつ計144冊を対象とした「教科書コーパス」の語彙の一覧表である。教科書コーパス語彙表には、全校種・全学年・全教科を総合した頻度を知ることができれば、校種別・学年別・教科別の頻度を知ることができるというメリットがある。

2-2-2 バトラー(2011)

バトラー(2011)では、教科学習を行うに当たって非常に重要だと誰もが認める「専門語」の大切さを認めつつも、リストの対象語を「縁の下の力持ち」的な役割（バトラー2011:66）である「学習語」に限定している。学習語とは、教科書などの学習場面で出現する語であるが、「専門語」と違って日常的にも用いられることがある語であり、教室での作業の手順、教師の指示、思考のプロセスを理解したり伝達したりする時に用いられる語と定義される。「学習語」は、「専門語」に比べると指導においては盲点になりがちで、教師から十分に説明されないため習得が難しいと予想され、それを習得しておかないと教科学習に支障を来す重要な語群と捉えられている。

この語彙リストでは、国立国語研究所の「教科書コーパス」から2005年度に使用された小学校1年生から中学3年生までの9学年全科目、計87冊分のデータが使用されている。

語の選出基準は、①一般語でないもの、②多くの教科で使用されているもの、③頻度が高いものの3つである。一般語としては、旧日本語能力検定試験4級、3級の語彙に含まれる語をほぼ日常一般語に相当すると捉え、それらの語をリストから消去している。多くの教科で使用されているものとして、4科目以上の教科書で使用されているものという基準で選出している。頻度に関しては、頻度数が最低20までの語彙を選んで、頻度の高い順に語彙を並べている。

それらの一次リストについて、日本語学習児童生徒の指導に携わっている公立小・中学校の教員や、大学で学部留学生の日本語指導を行っている指導者など合計7人による絞り込み、ならびに重要度の判定が行われ、1,230語が選定されている。それを頻度順に4つのブロックに分け、それらのブロックごとに重要度（高、中、低）によって分類している。

² 教科書コーパス語彙表の詳細は田中・近藤(2011)、教科書コーパスの詳細は田中・近藤・平山(2011)を参照されたい。

2-2-3 茂木(2013)

茂木(2013)は「つまずきことば」という観点から小学校国語科教科書の語彙の分析を行っている。「つまずきことば」とは、「各教科において、学習者（児童）にとってスムーズな意味の理解が難しく、指導者による何らかの手当てを必要とする（と予想される）語または語彙」（p.343）を指す。茂木(2013)は、横軸に「子どもの日常生活からの距離」という観点、縦軸に語彙の「教科固有性」という観点を取り、教科書の語彙を4つに分類している（表1）。

表1. 教科書の語彙の分類

	子どもの日常に近い語	子どもの日常から遠い語
教科に固有の語	(1)	(2)
教科に固有でない語	(3)	(4)

(茂木 2013:343)

茂木(2013)によると、(1)の具体例は「動物」、(2)の具体例は「主語」、「塩酸」、「円周」、(3)の具体例は「(文章の)組み立て」や「(日本との)つながり」、(4)の具体例は「いろり」、「煎じる」などである。日常生活から遠い(2)や(4)は、そもそも学習者にとって見慣れないものであり、理解自体が学習活動に含まれる。日常生活に近い(1)は素朴な概念の理解が科学的概念に適切に置き替わらない可能性がある。(3)の場合、ある語の日常的・辞書的な意味を理解したとしても、教科書の特定の文脈に合わせた適切な解釈ができなければ、つまずきが起こる可能性がある」と指摘している。

また、茂木(2013)は小学校国語科教科書における「つまずきことば」のデータ収集を行っている。この調査が対象とした教科書は、光村図書出版の平成23年度版（平成22年検定）小学校国語科用教科書「こくご」（1年～2年）および「国語」（3年～6年）である。調査手順としては、現職小学校教員2名（教職経験25年）が、自らの経験に基づいて、これらの教科書に現れる語彙や表現から抽出したものである。このような手順で作成された「つまずきことば」のリストの元データ（2,548項目）に対して、学年間で重複している項目の削除、固有名詞や付属語（助詞・助動詞）の削除などを行った結果、1,795語を収集し、以下のような特徴を見出している³。

結果：

（語種と品詞）

- ・ 品詞は1年から6年まで比率が一定で、名詞7割、動詞2割
- ・ 動詞連用形が名詞化した語。（例）「作り」「動き」「集まり」など。
- ・ 動詞は和語が74.5%と多く、その三分の一がV+Vの複合動詞。
（例）食いちがう、見立てる、眠りこむ
- ・ 語種は1年で和語が8割、4年で和語と漢語が逆転する。

（傾向）

- ・ 名詞の5.9%は生物名。体験的実感がない場合、辞書からの知識に頼るしかない。
（例）いそぎんちゃく、コウジカビなど。
- ・ 副詞では、オノマトペが半数を占める。（例）いそいそ、ひやり
- ・ 抽象的概念の語。（例）治める、関わる、出来事

³ ただし、1,795語の全てがリストとして公開されているわけではない。

- ・ 比喩的な表現 (例) クレーンの腕
- ・ 非日常的な用法 (例) 長くはもたない, 災害に見舞われた
- ・ 文脈がないもの (例) 見出しの位置にあるもの

この研究は, 日本人児童のつまずきに関する研究であり, 外国につながる子どもたちのつまずきとの違いとは異なると予想されるが, 上記の「つまずきことば」の様々な特徴は興味深い指摘である。

茂木(2013)では, 国語科専門用語との関連として, 以下の2つの語彙表を参照している。

(A) 甲斐睦朗監修・松川利広編 (2005) 『語彙指導の方法—語彙表編』光村図書出版

(B) 中央教育研究所 (成田雅樹) (編) (2007) 『21世紀における学校教育・国語科の在り方—小学校国語科専門用語候補の選定と理解度調査及び考察—』中央教育研究所
両方のリストと重なる「つまずきことば」として以下の100語を提示している。

後書き, 後づけ (後付け), あらすじ, 意図, イメージ, 引用, うごき (動き), おくりがな (送り仮名), (※のばす) おん (音), (※音訓の) 音, 音読, 会意 (一文字), 外来語, 書き表す, 書きうつす, 画, 画数, かじょう書き (簡条書き), 聞き手, 記述, 脚本, 狂言, きろく (記録), 組み立て, 訓, 敬意, 形声 (一文字), けんじょう語 (謙譲語), 構成, 構想, 故事成語, 古典, ことわざ, 小見出し, 作者, 詩, 指事 (一文字), 詩集, 詩人, 修飾語, 修正, 熟語, 主語, 取材, 主張, 述語, 象形 (一文字), 書名, 心情, 人物, 推敲, 随筆, 清音, 設定, 説得力, せつめい (説明), 尊敬語, 対比, 濁音, たとえ, 短歌, 段落, ちょうし (調子), つくり (旁), ていねい語 (丁寧語), できごと, 伝記, 登場人物, 討論, 討論会, 童話, ト書き, 読本, (※漢字の) 成り立ち, 俳句, はっぴょう (発表), 場面, 半濁音, 反論, 筆者, 表記, 表現, 複合語, 符号, 部首, ぶん (文), 文語, ぶんしょう (文章), へん (偏), ほうこく (報告), 間, まとまり, 見出し, 要旨, ようす (様子), 要点, 要約, 読みかえす, わけ (訳), 話題

また, 「光村専門」と重なる「つまずきことば」として以下の85語を提示している。

相づち, 当てはまる, 案, アンケート, 言い切り, 言いつたえ (言い伝え), 言い回し, 意識, いろはかるた, 印象, インターネット, うなずく, えがく (描く), 絵巻物, 絵文字, 演説, 解説, 回答, 解読, 書き記す, 課題, 語りつぐ (語り継ぐ), 語る, (※辞書の) 活用, かんさつ (観察), 漢詩, かんじ (感じ), 聞き分ける, 記号, 記事, 強調, くさび形文字, 経験, けんそん (謙遜), 交流, 事がら (事柄), こめる (込める), こよみ (暦), 指ししめす (指し示す), さす (指す), 参考, しめい (指名), しめす (示す), 下手, 出典, 出版, じゅん (順), しょうかい (紹介), 状態, しょうち (承知), 証明, 助言, しりょう (資料), 記す, 神話, 宣言, 全体, せんもん (専門), 想像力, 対象, 単行本, 注, つたわる (伝わる), ていあん (提案), 手がかり, 点字, 問い, 投書, とり入れる (取り入れる), 能, 発行, 発信, 批評, 分析, 分類, 編集, 報告書, ホームページ, 翻訳, 明記, やく (訳), 読み広げる (味わい, つながり, つぶやく)

さらに, 「中央専門」と重なる「つまずきことば」として以下の27語を提示している。

歌人, 語り手, かんそう (感想), 漢文, 慣用句, キーワード, 議題, くぎり (区切り), 結末, 視点, じょうほう (情報), 説得, 絵画, 題材, 動作, ないよう (内容), ナレーター, (※語や文の) ひびき (響き), ファンタジー, 文語調, ポスターセッション, 見出し語, 民話, むすび (結び), 流行語, わらべ歌, わり付け (割り付け)

茂木(2013)の研究は興味深く, 外国につながる子どもたちの研究にも応用できる部分が多いが, 茂木自身が指摘するようにこの研究では「つまずきことば」かどうかの判断が各語に対する教師の主観的な評定に基づくので, 実際につまずきの有無は実証できないという限界がある。そこで, 本研究では実際の現場のつまずきを, 担当教師による授業記録をデータとして扱うことで明らかにしたい。

2-2-4 旧日本語能力試験出題基準語彙表

2009年までの日本語能力試験(JLPT)⁴では, 出題基準として語彙表が公開されていた。本研究では, いわゆる初級の語彙表として旧4級, 旧3級の語彙を取り上げる。国際交流基金(2002)によると, 旧4級の語彙は「原則として語彙表の728語を含む800語および「あいさつ語等表現」のなかの当該項目とする」(p.241), 旧3級の語彙は「原則として語彙表の1409語を含む1500語および「あいさつ語等表現」のなかの当該項目とする」(p.241)となっている。

2-3 どんな語が難しいか

西川・青木・細野・樋口(in printing)⁵では, 小学校入学時点ですでに表面的には日本語を母語とするモノリンガルの児童と変わらない日本語力を持っているように見える, 日本生まれの第二言語児童に対して, 「モノリンガルとは日本語力において何かが違う」, 「意外と“簡単な”日本語を知らない」という教育現場での気づきをスタート地点とした研究を行っている。この研究では, 「簡単な日本語」として, 和語動詞に材料を絞った産出調査を行い, 日本語を第二言語とする児童とモノリンガル児童の語彙力の比較をしている。その結果, 日本語を第二言語とする児童にとっては, 「(ごはんを) たく」, 「(目薬を) さす」, 「(ハンガーに洋服を) かける」など学校生活で接する機会がない語は難しく, モノリンガルの習得と大きな差が見られることを明らかにしている。この研究が明らかにした両者の言語生活の違い, および頻度による習得の差という指摘は, 従来の研究で言われてきた「学習語」か「一般語」か, 抽象的か具体的かという語彙の特性だけでは習得の差を語ることはできないことを明らかにしており, 興味深い。

本研究でも, 外国につながる子どもたちにとって難しい語を明らかにする際に, 従来の研究の枠組みを一旦横に置いて, 現場の気づきを出発点とすることにしたい。

3. データ

3-1 日本語クラスと在籍する児童

日本語クラスは, 2011年度, 2012年では指導の方針が次のように異なる。

⁴ 以下では, 旧日本語能力試験と呼ぶ。

⁵ 印刷中の原稿をいち早く参照させていただいた西川朋美さんに心から感謝申し上げます。(私信2015年2月5日)

【2011年度】

- ・一斉指導⁶
- ・取り出し授業では、国語、算数の教科書中心の授業⁷

【2012年度】

- ・複数のクラスで指導
- ・取り出し授業では、国語、算数の教科書と日本語の指導⁸

母語による指導の違いを考慮して、本研究では分析対象を「中国につながるルーツを持つ児童」8名のデータに限定する⁹。これらの児童のうち5人は中国帰国児童、すなわち、中国残留邦人の引き揚げによる定着後に日本で生まれた子弟（4世）である。残りの3人は、保護者の中国から日本への移動に伴って、2005年から2008年にかけて来日した児童である。対象児童は2011年度では5人、2012年度では3人であった。この8名のうち、日本語クラスに途中から参加した児童が2名いた¹⁰。

3-2 分析材料：授業記録

実際に指導があった語を示す資料として、本研究では広島県内のある小学校3年生の日本語クラスの授業記録（2011年度、2012年度）を用いる。なお、授業記録の書式は2年を通して一貫しており、以下のような内容が記載されている。

表 2. 授業記録の内容

・授業形態 （取り出し・入り込み）	・児童名	・指導した語
・月日	・指導者名	・指導した文字
・時限	・日本語学習クラスでの教科	・指導した文型や文法
・在籍学級の教科	・日本語学習クラスでの単元	・特記事項
	・宿題	

このうち特記事項については、指導の際に担当教師が気づいた点が児童ごとに分かれて日本語クラスの各回ごとに記述されている。その記述の中で、学習の困難点としては①児童の誤用や誤解（～と～を混同している、～と～とを勘違いしていた…など）、②苦手とする学習項目（～を知らなかった、～が難しかった…など）、③学習項目の定着の具合（まだ～を理解できていない、～を忘れてしまっている…など）の記述がされている。

⁶ 他の児童との能力差の大きい1名の児童（2月、3月のみ在籍）に関しては個別指導が行われた。

⁷ この小学校での使用教科書は2011年、2012年とも共通しており、国語、算数はともに東京書籍。

⁸ 日本語の教材は日本語教科書、市販の国語やことば関係の本から抜粋して、作成。

⁹ それ以外の児童は、2012年度のフィリピンにつながる児童1名のみである。

¹⁰ 児童の一人は、日本語指導が必要であると学期途中で判断されたことによる中途からの参加で、2011年9月～2012年3月に参加した。もう一人の児童は中国への一時帰国から戻ったのち、2012年2月～3月の参加であった。

4. 分析

日本語授業での指導記録に基づき、語彙リストに盛り込むべき語を選定する。そのために日本語クラスで複数回指導があった語をリスト化する。語の「重みづけ」で言えば、これらが重要な語であると言える。またこれらの語から、外国につながる子どもたちにとってどんな語が難しいのかを明らかにする。さらに、先行研究の語彙リストに挙げられた語との比較を通して、対象児童、日本語クラスの個別性を越えて、共通して重要な語を明らかにする。

4-1 指導記録に基づく語彙リスト

4-1-1 方法

- ① 授業記録のうち、「指導した語」、「特記事項」に出現する語を抽出する。
 - ・語の認定に当たっては、指導記録の記載を尊重した¹¹。
 - ・一つのコマ（同一日時の同一時限）で扱った語を「1回出現」と換算。
 - ・あるコマが能力別学級（複数の先生がそれぞれ別の学級で教えている）である場合、同一日時、同一時限に行われたクラスであれば、何人の先生が扱った語でも合算して「1回出現」として換算。
- ② 抽出した語を指導回数が多いものから順に並べ、指導回数が2以上のものを抽出してリスト化する。
- ③ 教科ごとの出現回数を記載する。
- ④ 品詞や意味カテゴリーなどの分析を、②のリストに挙げられた語に限って行う。

4-1-2 結果

4-1-2-1 複数回の指導があった語の数

・授業記録のうち、「指導した語」、「特記事項」に出現する語は異なり語数で1998語であった¹²。そのうち複数回指導された語は408語であった。指導記録に表れる語のうち約8割の語（79.6%、1590語）は指導が一回のみに限られることが分かった。

次に、授業記録に出現する語について教科ごとの出現回数を数えた結果を次に挙げる。

表3. 教科ごとの述べ語数・異なり語数

	述べ語数	異なり語数	異なり語数／述べ語数
国語	1569	1344	85.7%
算数	725	459	63.3%
社会	38	33	86.8%
理科	2	2	100.0%
その他 ¹³	317	291	91.8%

¹¹ 単語ではなく句になっているもの（例：読みましょう、書きましょう、足がすくむなど）や複合語になっているもの（例：数直線、つなぎ言葉、負けず嫌いなど）も含まれる。

¹² 延べ語数で2651語であった。

¹³ その他には、教科の指導をせずに日本語のみを指導したクラス、および教科名が空欄だったクラスの指導語が含まれる。

社会，理科は，月に数回の入り込み授業でのみ指導されている教科なので指導回数も限られる。そこで，ここでは取り出し授業と入り込み授業で指導される国語と算数で扱われる語の数を見たい。すると，国語は算数に比べて指導される語彙の異なり語数が多く，それに比べると算数は，一語が複数回指導されているような傾向にあることが分かった。このことから，算数は国語に比べて記憶すべき語が限られ，その意味では学習の労力がいらぬことが分かる。しかし，記憶すべき語は限られるのに，それが複数回指導されていることから見ると，それぞれの語の定着が難しいという別の指摘ができる。

4-1-2-2 どんな語が難しいのか

ここで，複数回指導された語に焦点化して，それらに共通する特徴が何かを明らかにする。そのために，指導記録の「特記事項」の項に見られた語の意味に関する困難点に関する記述に注目した。指導が重点的に必要な語彙の意味的な特徴をまとめると，以下のような語に関しては意味を捉えることが難しいことが示されていた。

- (1) 教科に特有の抽象的な概念の語
- (2) 日常生活の語
- (3) 時を表わす語
- (4) 方角を表わす語
- (5) オノマトペ
- (6) 日本の年中行事の語

これらの分類のうち，(1)教科に特有の抽象的な概念の語，(2)日常生活の語が難しい語に関しては，1970年代の終盤から発展してきたバイリンガル教育，バイリンガル研究の分野における BICS と CALP の議論と重なり，これまでも数多く議論されている。

それ以外の(3)から(6)の4項目については，教科に特有の抽象的な概念の語，および日常生活の語の細分類であるとも考えられるが，その中でもどんな語が難しいかを具体的に示すことになった。

日本語クラスで複数回指導があった語を上記の6つの分類に振り分けたところ，語によって「教科に特有の抽象的な概念の語」と「日常生活の語」のいずれのカテゴリーにも含まれることがある，「数詞・助数詞・単位を表わす語」が多いことが分かったので，一つのカテゴリーとして独立させた。

- (7) 数詞・助数詞・単位を表わす語

よって以下では，(1)から(7)の7つの分類に，日本語クラスで複数回指導があった408語を振り分けたリストを挙げる。なお，各リストに示されている「回数」は，2年間の日本語クラス（取り出しと入り込みの両方）での総指導回数である。

(1) 教科に特有の抽象的な概念の語

語	回数	語	回数	語	回数	語	回数
余り	8	数直線	4	行	3	辞書形	2
等しい	8	数量詞	4	球	3	繰り下げ	2

整数	8	繰り上がり	4	100倍	3	繰り上げ	2
分母	7	円	4	割る(割る数)	3	位	2
分子	7	横軸	4	かける(かけられる)	2	位をそろえる	2
倍	6	割る(割られる)	4	命令形	2	漢数字	2
半径	6	つなぎ言葉	4	付け加え言葉	2	横書き	2
直径	6	かける(かける数)	4	地図記号	2	繰り下げる(繰り下げて)	2
等号	6	割り切る(割り切れない)	4	目盛り	2	合計	2
動詞	5	割る	3	辺	2	繰り上がる	2
等分	5	かける	3	動詞の活用	2	形式段落	2
部首	5	筆算	3	棒グラフ	2	五十音順	2
小数	5	直線	3	分数	2	区切り	2
単位	5	小数点	3	中世	2	差	2
距離	5	主語	3	表	2	擬態語	2
かける(かけられる数)	5	送り仮名	3	熟語	2	記号	2
割り切る(割り切れる)	4	商	3	述語	2	あらすじ	2
名詞	4	曲線	3	積	2	ア行	2
中心	4	割り算	3	速度	2	にんべん	2
不等号	4	もとにする大きさ	3	商(を立てる)	2	グラフ	2
和	4	形容詞	3	段落	2		
縦軸	4	暗算	3	人口	2		

(2) 日常生活

語	回数	語	回数	語	回数	語	回数
代金	6	目薬をさす	2	使用	2	交番	2
道のり	5	都合	2	重さ	2	右	2
束にする	4	注文する	2	太平洋	2	軽い	2
空港	4	湯船	2	川原	2	厚い	2
九州	4	灯油	2	枝	2	近く	2
苦手	4	注ぐ	2	待合室	2	延期	2
洋服	3	町	2	水族館	2	横	2
借りる	3	葉	2	実	2	リボン	2
大きい	3	沸く	2	細い	2	やぶる	2
図書館	3	長い	2	少年	2	左	2
住所	3	練る	2	残り	2	昆虫	2
少ない	3	童話	2	心	2	うすめる	2
短い	3	平ら	2	多い	2	お膳	2
駅	3	道路	2	真っ暗	2	お湯を沸かす	2
銀行	3	歩道橋	2	神戸	2	フェリー	2
角	3	預ける	2	貸す	2	たたみ	2
港	3	同じ	2	趣味	2	ボート	2

おつり	3	和服	2	棚	2	くし	2
バス停	3	市役所	2	消火器	2	デパート	2
両親	2	始発	2	段	2		
予定	2	重い	2	県	2		

(3) 時の表現

語	回数	語	回数	語	回数	語	回数
今年	5	期間	3	来週	2	正午	2
来年	3	今夜	3	日間	2	午後	2
時刻	3	翌朝	2	昼	2	今週	2
時間	3	冬	2	先週	2	季節	2
去年	3	秒	2	秋	2		

(4) 方角

語	回数	語	回数	語	回数	語	回数
東西南北	5	方向	2	北東	2	南東	2
方角	3	北西	2	南西	2		

(5) オノマトペ

語	回数	語	回数	語	回数	語	回数
ヨタヨタ	2	テクテク	2	ぎゅうぎゅう	2	ガラガラ	2
トボトボ	2	スタスタ	2	ブラブラ	2		

(6) 年中行事

語	回数	語	回数	語	回数	語	回数
福袋	2	七福神	2	大晦日	2	お月見	2

(7) 数詞・助数詞・単位

語	回数	語	回数	語	回数
いくつ分	19	1両	2	mm	2
1あたり	9	十万	2	7(しち)	2
1あたり(1あたりの数)	6	dl	2	1万(いちまん)	2
1箱	5	1ダース	2	300(さんびゃく)	2
本(助数詞)	4	10等分(じゅうぶん)	2	800(はっぴゃく)	2
個分(助数詞)	3	10倍(じゅうばい)	2	5人掛け	2
m	3	1めもり	2	ml	2
cm	3	g	2	600(ろっぴゃく)	2
1人分	2	kg	2	1000万(いっせんまん)	2
本数	2	10(じゅう)	2		2

複数回の指導があったものの、上記の7つのカテゴリーに含まれない語は次の174語であった。

その他

語	回数	語	回数	語	回数
復習	5	長所	2	まで(～から～まで)	2
予習	5	予報	2	下流	2
幹	5	読みましょう	2	見やすい	2
助手	4	要求	2	緩やかな	2
題名	4	来て(きってーきて)	2	家路	2
また	4	入れ替える	2	顔から火が出る	2
ずつ	4	本箱	2	一斉に	2
でも	4	例えば	2	攻撃	2
それぞれ	4	毎に	2	期日	2
こわごわ	4	同じ数ずつ	2	角笛	2
しかし	4	路	2	高原	2
かさ	4	着陸	2	雨宿り	2
理由	3	手相	2	ゆだねる	2
役立つ	3	世界	2	化石	2
木の葉	3	足がすくむ	2	駅長	2
風速	3	指名	2	見失う	2
負けず嫌い	3	惜しい	2	教科	2
得意	3	出港	2	言いつける	2
列島	3	出版	2	橋を架ける	2
戻る	3	書きましょう	2	曲がる	2
落葉	3	書庫	2	交代	2
藤づる	3	守備	2	ろく木	2
筆者	3	助走	2	開会式	2
嵐	3	勝負	2	行進	2
放牧	3	歯科	2	左右	2
分ける	3	谷川	2	ぴったり	2
全部の数	3	持ち主	2	スピードが速い	2
石油	3	実際	2	じゃれる	2
前(まえ)	3	時間が早い	2	いばる	2
水深	3	生死	2	ウェイター	2
従う	3	炭	2	うず	2
進級	3	切手(きってーきて)	2	さわやか	2
科目	3	短所	2	その他	2
見にくい	3	川下	2	だから	2
幸運	3	省く	2	うらめしい	2
休息	3	川上	2	えこひいき	2
はした	3	上流	2	さむざむ	2
たにし	3	主人公	2	すくむ	2
あべこべ	3	手を焼く	2	そして	2
はやす	3	足が棒になる	2	それから	2
うづくまる	3	心構え	2	ふと	2
うずまく	3	打撲	2	それで	2
かすり	3	親しい	2	へそを曲げる	2

発音	2	親しげ	2	ちょうど	2
木の枝	2	第六感	2	つり橋	2
野外	2	作物	2	くちびるをかみしめる	2
投手	2	醜い	2	どうどう	2
洋	2	次に	2	なだらか	2
落ち葉	2	沼	2	くろぐろ	2
目を丸くする	2	やまびこ	2	はかり	2
問屋	2	暗記	2	ゴーカート	2
峠	2	まず	2	こずえ	2
天使	2	見通しを持つ	2	パンチ	2
木の幹	2	幸い	2	コンパス	2
長椅子	2	寒波	2	から(～から～まで)	2
平等	2	空き家	2	べっかんこ	2
門柱	2	巻尺	2	きしむ	2
予想	2	黄金	2	さえぎる	2

これらの中には、「予習」、「復習」、「開会式」、「暗記」、「コンパス」、「進級」など、学校生活の場で用いられる語が散見され、それについては一つのカテゴリーとして独立させることができるかもしれない。なお、どの語にも含まれない語として残した「その他」の174語と日常生活の語として振り分けた語との違いは、厳密には判別が困難であった。例えば、「切手」、「落ち葉」、「得意」、「持ち主」、「分ける」、「投手」、「はかり」などの語は日常語かどうか判定に迷う語であった。詳しくは、次の4.2節で述べる。

4-1-3 まとめ

ここでまとめとして、表3に各カテゴリーに含まれる語数と語の総計408項目に対する各カテゴリーの語の割合を示す。

表4. 指導が難しい語のカテゴリーと語数

	項目数	総項目中の割合
教科に特有の抽象的な概念	86	21.1%
日常生活	82	20.1%
数詞・助数詞・単位	29	7.1%
時	19	4.7%
方角	7	1.7%
オノマトペ	7	1.7%
日本の年中行事	4	1.0%
その他	174	42.6%
計		408項目

この結果から、日本語クラスでは、「一般語」が82項目、教科に特有の抽象的な概念である「専門語」が86項目と同じ程度指導されていることが分かった。続いて細分類を見ても、日本語クラス担当教師による特記事項の記述にもあった通り、「時」、「方角」

という意味的特徴のある語、「オノマトペ」、「年中行事」は、比較的多く指導されていることが分かった。それに加えて、「数詞・助数詞・単位」の指導も多かった。以上の結果から、従来の研究で用いられる二分類である、特有の抽象的な概念の語およびは日常生活の語という観点だけでなく、それらの大分類の中で、特にどのような語が困難であるのかを、より焦点化して捉えることが、具体的な困難点を把握するためには必要であることが指摘できる。

次節では、外国につながる子どもたちに共通して重要な語の一般化を目指して、先行研究の語彙リストに挙げられた語との比較を行い、語の重みづけ、絞り込みを試みたい。

4-2 教科書コーパス語彙表・他の語彙リストとの比較

4-2-1 教科書コーパス語彙表との比較

教科書コーパス語彙表との比較の方法、結果は以下の通りである。

【方法】

- ①教科書コーパス語彙表のうち、まず、教科書全体（頻度2以上）、小学校全体（頻度2以上）、小学校国語（頻度2以上）、小学校数学（頻度2以上）の語彙と指導した語の重なり具合を Excel の VLOOKUP 関数を用いて分析した。
- ②小学校前半（頻度2以上）、小学校前半国語（頻度2以上）、小学校前半数学（頻度2以上）の語彙と指導した語の重なり具合を Excel の VLOOKUP 関数を用いて分析した。

【結果】

- ①教科書全体と重なる語が 293 語（71.8%）、小学校全体と重なる語が 265 語（65.0%）、小学校国語と重なる語が 165 語（40.4%）、小学校数学と重なる語が 161 語（39.5%）となった。
- ②小学校前半と重なる語が 156 語（38.2%）、小学校前半国語と重なる語が 93 語（22.8%）、小学校前半数学と重なる語が 86 語（21.1%）となった。
- ③全てと重なる語が以下の 41 語（10.0%）であった。

（例）余り・距離・かける（かけられる数）・本（助数詞）・東にする・円・それぞれ・戻る・短い・時間・前・大きい・少ない・駅・m・町・同じ・長い・毎に・表・秒・分ける・例えば・次に・重い・多い・世界・左・交代・横・右・県・まず・教科・午後・10・ちょうど・1万・300・ぴったり・7

上記の 41 語を 4-1-2-1 で分類したカテゴリーから見ると、日常生活の語が約3分の1、その他が約3分の1という結果になった。教科書コーパス語彙表との比較ではあるものの、日常生活で使われる語の重なりが大きく、3年生の段階では、日常語はそれだけ重要であることが伺える。

4-2-2 バトラー(2011)との比較

バトラー(2011)との比較の方法、結果は以下の通りである。

【方法】

- ①バトラー(2011)の語彙リストと指導した語の重なり具合を Excel の VLOOKUP 関数を用いて分析した。

【結果】

①重なっていたのは以下の 26 語 (6.4%) である。

(例) 等しい・直径・単位・距離・中心・和・円・直線・役立つ・石油・従う・期間・曲線・割る・入れ替える・表・段・平ら・予想・要求・注ぐ・左右・作物・使用・実・グラフ

バトラー(2011)の学習語彙リストは、「一般語」でも、「専門語」でもない、「学習語」から成る。今回の指導記録の分析では、バトラーのリストの中では、26 語しか複数回の指導がなかったことは注目すべきである。しかし、「学習語」がそもそも指導されないのか、それとも指導はされるが複数回の指導が必要ないのかについては、今回の分析では分からなかった。これに関しては、今後、授業記録に表れた一回のみ指導された語に分析対象を広げて明らかにする必要がある。なお、これらの中には、本研究では「教科に特有の抽象的な概念の語」とされる「直径」、「グラフ」、「和」、「割る」などの語や、「日常生活の語」に分類された「使用」、「実」が含まれる。これに関しては、バトラー(2011)においても、「学習語」「専門語」「一般語」の区別は難しく、とくに「一般語」に関しては、バトラー(2011)では、旧日本語能力試験 4 級、3 級の語彙表と定義されるが、「日本語能力検定試験が子どもを対象を絞ったものではないこと、小中学生の日常一般語として適切であるかに関しては、疑念の余地がないわけではない」(p.128) という記述がある。そこで、次節では、本研究のリストと旧日本語能力試験の語彙とを比較してみる。

4-2-3 旧日本語能力試験との比較

旧日本語能力試験の語彙表との比較の方法、結果は以下の通りである。

【方法】

①旧日本語能力試験 4 級、3 級の語彙表と指導した語の重なり具合を VLOOKUP 関数を用いて分析した。

【結果】

①重なっていたのは以下の 74 語 (18.1%) である。

(例) 倍・復習・予習・今年・空港・円・また・でも・しかし・ずつ・戻る・短い・洋服・来年・理由・借りる・時間・図書館・住所・今夜・大きい・角・去年・銀行・駅・港・おつり・かける・来週・町・同じ・発音・葉・長い・表・予定・都合・両親・冬・沸く・昼・辺・例えば・枝・趣味・秋・細い・重い・多い・貸す・人口・今週・世界・先週・心・左・棚・横・軽い・厚い・右・季節・県・まず・午後・交番・曲がる・近く・それから・そして・それで・ちょうど・デパート・だから

旧日本語能力試験の語彙表には、本研究の分類における「日常生活の語」37 語、「時を表わす語」13 語見られた。それに加えて、「その他」の語が 18 語見られたが、この中には「また」、「でも」、「しかし」といった接続詞や「ずつ」、「ちょうど」「例えば」など、文の構成に関わる語が多いことも特徴の一つと言える。

5. 考察

5-1 リストの比較を通じた考察

教科書全体、小学校全体、小学校国語、小学校数学、小学校前半、小学校前半国語、小学校前半数学の全てと重なる41語は、教科書の違い、教科の違い、学年の違い、地域の違い、児童の属性などを超えた、普遍的に重要な語を表すと考えられる。

これに対して、その41語とは重ならないものの、今回の調査の対象に最も近い小学校前半国語や小学校前半数学¹⁴と重なる語として以下のものが挙げられる。

【小学校前半国語と重なる語】

今年・幹・中心・題名・しかし・洋服・役立つ・得意・方角・理由・借りる・主語・住所・今夜・送り仮名・行・来週・発音・葉・峠・落ち葉・都合・冬・段落・昼・方向・道路・枝・左右・秋・細い・出版・述語・持ち主・今週・川原・少年・実・心・真っ暗・棚・横書き・一斉に・軽い・区切り・曲がる・近く・いばる・そして・にんべん・ふと・ガラガラ

【小学校前半数学と重なる語】

等号・倍・代金・復習・単位・道のり・かさ・直線・筆算・時刻・角・割り算・暗算・割る・割る(割る数)・cm・おつり・入れ替える・段・目盛り・予想・辺・切手(きってーきて)・実際・人口・残り・正午・リボン・繰り下げる(繰り下げて)・位・巻尺・厚い・合計・800・はかり・1ダース・1メモリ・mm・600・dl・g・kg・10倍・グラフ・ml

これらは国語、数学において重要な語であり、教科の指導の際に理解ができるように丁寧に扱う必要がある。

一方、どの語彙リストとも重ならない語彙も107語(26.2%)見つかった。

(例) いくつ分・1あたり・1あたり(1あたりの数)・東西南北・不等号・数量詞・数直線・割り切る(割り切れる)・割り切る(割り切れない)・こわごわ・かける(かける数)・つなぎ言葉・藤づる・負けず嫌い・進級・全部の数・小数点・一箱・もとにする大きさ・形容詞・個分(助数詞)・見にくい・割る(割られる数)・バス停・100倍・はやす・命令形・同じ数ずつ・読みましょう・目を丸くする・長椅子・木の枝・目薬をさす・付け加え言葉・福袋・木の幹・門柱・湯船・動詞の活用・歩道橋・地図記号・来て・棒グラフ・本箱・足が棒になる・手を焼く・手相・水族館・七福神・十万・時間が早い・重さ・川上・第六感・書きましょう・親しげ・書庫・始発・市役所・辞書形・消火器・足がすくむ・谷川・太平洋・待合室・大晦日・主人公・ヨタヨタ・形式段落・ろく木・まで(～から～まで)・位をそろえる・一人分・繰り上がる・五十音順・開会式・空き家・繰り下げ・繰り上げ・擬態語・見やすい・駅長・見通しを持つ・橋を架ける・漢数字・顔から火が出る・どうどう・ウェイター・つり橋・すくむ・えこひいき・お湯を沸かす・かける(かけられる)・じゃれる・スタスタ・5人掛け・スピードが速い・その他・へそを曲げる・テクテク・くちびるをかみしめる・トボトボ・くろぐろ・ゴーカート・1000万・べっかんこ・から(～から～まで)

¹⁴小学校なので「算数」だが、教科書コーパス語彙表の表現に従って、「数学」とした。

これらの107語のうち18語は、本研究で設定した「難しい語」のカテゴリーの細分類(3)から(7)に分類される語であった。従来の語彙リストには重ならない語ではあるが、現場の教師から難しさの指摘がある語であることから、最終的なリストから削除せずにおきたい。それらは具体的には下の表5の語である。

表5. 先行研究のリストと重ならないが「指導が難しい」と考えられる語

番号 ¹⁵	語	番号	語	番号	語	番号	語
4	東西南北	6	福袋	7	1あたり (1あたりの数)	7	一人分
5	ヨタヨタ	6	七福神	7	一箱	7	5人掛け
5	スタスタ	6	大晦日	7	個分(助数詞)	7	1000万
5	テクテク	7	いくつ分	7	100倍		
5	トボトボ	7	1あたり	7	十万		

107語から、上記の18語を引いた残りのうちの約半分(51語)の語は、「その他」に分類される語であった。これらの語には、例えば「藤づる」、「はやす」、「くちびるをかみしめる」、「つり橋」、「足がすくむ」、「手相」、「第六感」など、教科書の題材に左右されるような個別的な特徴を表している語が多く含まれる。これらの語に関しては、今回の対象校と同じ教科書を用いる際には、説明が必要になる語と捉えておけばよいだろう。

5-2 指導が必要になる語に共通する特徴

指導記録の分析を通して、日常生活の語、教科に特有の抽象的な概念の語のいずれもが難しいことが分かった。このうち、日常生活の語は、西川他(in printing)の指摘するように、生活体験、特に日本語を母語とするモノリンガル児童であれば小学校に入学するまでの日常生活で触れることが多い語である。日本語を母語とするモノリンガル児童であれば当たり前前の生活に密接な語もまた、3年生の日本語クラスでは、比較的多く指導されていることが分かる。このことは、外国から日本の小学校へ編入してくる児童に対する日本語指導における語彙の選定が、日本語母語話者の視点では困難であることを示唆すると同時に、小学校レベルの日本語指導よりも学齢の低い段階からの日本語指導が果たす役割が小学校レベルの日本語力の土台として重要であることを示唆している。

さらに、従来は日常生活の語、教科に特有の抽象的な概念の語と分類される語の中でも、次のようなカテゴリーの語は、指導が困難な語であることが示唆された。

- ・数詞、助数詞、単位を表わす語
- ・時を表わす語
- ・方角を表わす語
- ・オノマトペ
- ・日本の年中行事の語

これらの中には、日本語を母語とするモノリンガル児童にも共通して困難な語と外国につながる子どもたちに特有の困難な語があると考えられる。例えば教科に特有の抽象的

¹⁵ ここでの番号とはカテゴリーの番号のことである。

な概念である「分母」「直径」などの語、「南西」などの方角の語は、両者に共通した困難な語であろう。一方、「今年」などの時の表現、「ヨタヨタ」などのオノマトペ、「お月見」などの日本の年中行事の語、「銀行」「リボン」などの日常生活の語は、外国につながる子どもたちに特有の困難点であることが予想される。数詞、助数詞、単位に含まれる語に関しては、dl (デシリットル)、mm (ミリメートル) などの単位は両者に共通して困難な語であると考えられるが、「10万」などの数詞、および「～本」などの助数詞は外国につながる子どもたちに困難な語であると考えて良いだろう。これらの困難点に関する実際の様相に関しては、現場の教師へのアンケート、日本語を母語とするモノリンガル児童と外国につながる子どもたちを対象にした語彙調査を通して、今後明らかにしなければならない。

6. 今後の方向性

今後の方向性としては、工藤(1999)、樋口・古屋・頼田編(2011)といった他の語彙リストとの比較を通して、語の重みづけ、絞り込みを行うことでリストをより一般化していかなければならない。工藤(1999)は外国人児童生徒を対象とした語彙リストの先駆的存在であり、樋口・古屋・頼田編(2011)は扱っている語彙数が最も多い。そのような語彙リストと比較することによって、本研究が提示する語彙リストの位置づけがより明確になるだろう。

また、さらなる今後の展開として、日本語モノリンガル児童と外国につながる子どもたちとの「難しい語」の比較および、日本語クラスでの指導経験が豊富な教師と、未経験の教師の「難しい語」の判断の比較を行いたい。これらの展開を通して、本研究は今後、複眼的な視点から語彙リストの客観性を検証すると同時に、現場での気づきの契機となる研究になることを目指している。

引用文献

- 川上郁雄(2008)『『移動する子どもたち』のプロフィシェンシーを考える—JSLバンドスケールから見える『ことばの力』とは何か—』鎌田修, 嶋田和子, 迫田久美子(編)『プロフィシェンシーを育てる——真の日本語能力をめざして』, 92-109, 凡人社
- 工藤真由美(1999)『児童生徒に対する日本語教育のための基本語彙調査』ひつじ書房
- 国際交流基金(2002)『日本語能力試験 出題基準 [改訂版]』凡人社
- 田中牧郎・近藤明日子(2011)「教科書コーパス語彙表」田中牧郎他『特定領域研究「日本語コーパス」言語政策班報告書 言語政策に役立つ, コーパスを用いた語彙表・漢字表等の作成と活用』, 55-63, 文部科学省科学研究費特定領域研究「日本語コーパス」言語政策班
- 田中牧郎・近藤明日子・平山允子(2011)「教科書コーパス」田中牧郎他『特定領域研究「日本語コーパス」言語政策班報告書 言語政策に役立つ, コーパスを用いた語彙表・漢字表等の作成と活用』, 7-54, 文部科学省科学研究費特定領域研究「日本語コーパス」言語政策班
- 西川朋美・青木由香・細野尚子・樋口万喜子(in printing)「日本生まれ・育ちのJSLの子どもの日本語力—和語動詞の産出におけるモノリンガルとの差異—」
- バトラー後藤裕子(2011)『学習言語とは 何か—教科学習に必要な言語能力』三省堂

樋口万喜子・古屋恵子・頼田敦子編(2011)『進学を目指す人のための教科につなげる学習語彙 6000 語』ココ出版

茂木俊伸(2013)「小学校国語科教科書における『つまずきことば』の分析」鳴門教育大学研究紀要 28, 343-355

文部科学省(2013)「日本語指導が必要な児童生徒の受入れ状況等に関する調査(平成 24 年度)」の結果について」

(http://www.mext.go.jp/b_menu/houdou/25/04/_icsFiles/afieldfile/2013/04/03/1332660_1.pdfよりダウンロード)

子どもを持つ外国人のための語彙シラバス

森篤嗣 (帝塚山大学)

1. はじめに

言うまでもないことであるが、対象者別の日本語教材を作成するとき、必要なことは対象者のニーズを把握することである。ニーズは多岐に渡るが、言語的な側面に目を向けたとき、もっとも大きな領域を占めるのが語彙であろう。ある分野を知るということは、そこで使われる言葉を知ることとほぼ等しい。例えば、子どもを持つ外国人が子育て・教育ということに関わっていくとき、保育所や小中高等学校で使われる言葉を知っている必要がある。

本発表は、上記のような考えに基づき、対象者のニーズに即した語彙シラバスを模索するものである。対象者のニーズは、対象者や関係者へのインタビューやアンケートなどで把握するのが一次的である。本発表では二次的ではあるが、学校配布物と連絡帳という媒体を調査し、そこで使われている言葉から探索を試みる。

2. 対数尤度比による特徴度の抽出

本発表では、学校配布物と連絡帳を仮に一つのコーパスとみなし、これらに特徴的に現れる語の抽出を試みる。学校配布物ないし連絡帳を「対象コーパス」、現代日本語書き言葉均衡コーパス (以下、BCCWJ とする) を「参照コーパス」として設定し、対数尤度比を計算する。本発表では、参照コーパスに基づいて計算した対象コーパスの対数尤度比を「特徴度」と呼ぶこととする。

石川(2006:8)によれば、対数尤度比 (log-likelihood score) は、コロケーションを検出する上でもっともバランスのとれた指標であるとされる。Leech, Rayson & Wilson(2001:16-17), 大学英語教育学会基本語改訂委員会 (編) (2002) なども、語彙表構築の基礎指標として対数尤度比を利用している。

対数尤度比を導く式は、下記のようになる (Hunston 2002:70)。 e はネイピア数 (約 2.718 となる無理数) と呼ばれるもので、自然対数の底である。

$$G^2 = 2 \times \sum \text{実測値} (\log e (\text{実測値}) - \log e (\text{期待値}))$$

そして、対数尤度比はコロケーション検出だけではなく、特徴度の抽出にも用いることができる (寺嶋 2009, 石川 2012)。まず、「見出し語 X」について見出し語 X の頻度と見出し語 X 以外の頻度を 2×2 の分割表を用意する。

石川(2012:129)を参考に、仮に見出し語 X の頻度が対象コーパス、参照コーパス共に 50 で、対象コーパスの総語数が 200、参照コーパスの総語数が 99,800 とした場合、下記のような分割表となる。これが粗頻度（頻度そのもの）の分割表である。

表 1 粗頻度の 2×2 分割表

	対象コーパス	参照コーパス	計
見出し語 X	50	50	100
見出し語 X 以外	150	99,750	99,900
計	200	99,800	100,000

そして、表 1 内の全てについて、 $(\text{頻度}) \times \log e (\text{頻度})$ 、すなわち、頻度と、頻度の自然対数をかけ合わせて、頻度×頻度の自然対数表を作成する。

表 2 頻度×頻度の自然対数表

	対象コーパス	参照コーパス	計
見出し語 X	*195.6	*195.6	460.5
見出し語 X 以外	*751.6	*1,148,164.6	1,150,041.3
計	1,059.7	1,148,790.2	*1,151,292.6

表 2 の*が実測値、それ以外が期待値である。公式の Σ は総和を示すので、(実測頻度×実測頻度の自然対数)の総和から、(期待頻度×期待頻度の自然対数)の総和を引き、その差を 2 倍したものが対数尤度比の値となる。表 1、表 2 の場合、対数尤度比は 496.65 となる¹。

寺嶋(2009:82)では、特徴度の指標として、カイ 2 乗検定及びイエーツの補正公式と対数尤度比を比較し、対数尤度比はイエーツの補正公式と比べ、コーパスサイズの影響を受けにくく、安定的に統計量を抑える傾向があるとしている。特徴語抽出において、統計量が高くなりすぎるのは、当該語を過大に評価することにつながる。したがって、現時点においては、特徴語抽出において対数尤度比を用いることが適切であると判断する。

¹ 対数尤度比を手軽に計算するために、石川慎一郎氏の研究室 HP では、「言語コロケーション統計指標自動計算 Excel シート (β版)」を公開している。ただし、「※検証中です。自己責任で結果を確認下さい」という注意書きもあることに留意して欲しい。ちなみに、「t 検定自動計算 Excel シート (β版)」もあわせて公開されている。

<http://language.sakura.ne.jp/s/stat.html>

3. 学校配布物に現れる語の特徴度

ここで対象コーパスとするのは、兵庫県神戸市と石川県金沢市の小学校 2 校で配布された配布物である。MeCab0.98 及び UniDic1.3.12 を用いて形態素解析をした結果、延べ語数は 421,560 語、異なり語数は 9,299 語である（いずれも短単位）。

3. 1 学校配布物の名詞

上記の 9,299 語のうち、品詞が「名詞 - 普通名詞」だけを取り出し、5,682 語を得た。さらに、この 5,682 語を参照コーパスである BCCWJ の語彙表と突き合わせ、BCCWJ の頻度を付与した²。そして、語彙シラバスに取り上げる語を探索するという観点から、ある程度、日本社会で使われる語であることを考慮に入れ、BCCWJ で頻度 100 以上という条件を付けた。その結果、5,862 語のうち 4,898 語が条件に合致し、これを本発表の対象とした。

この 4,898 語について、学校配布物（対象コーパス）と BCCWJ（参照コーパス）のそれぞれの粗頻度の相関係数は 0.570 であった。後に示す動詞や形容詞と比較すると、この値はかなり低い。すなわち、学校配布物（対象コーパス）と BCCWJ（参照コーパス）の名詞については、一致度は低く、学校配布物の語彙が特殊であることを示唆している。以下、特徴度順に上位 50 語を示す。

表 3 学校配布物の名詞（特徴度上位 50 語）

	見出し語	配布物	BCCWJ	特徴度		見出し語	配布物	BCCWJ	特徴度
1	学校	2584	38116	9590.02	26	油	294	3509	1203.12
2	給食	684	1388	4932.61	27	火	370	7245	1182.17
3	学年	544	2041	3360.56	28	豚肉	204	976	1172.72
4	学習	725	8983	2918.06	29	子供	942	59470	1167.12
5	会	1506	61475	2898.03	30	砂糖	266	3158	1091.09
6	牛乳	487	2283	2817.50	31	担任	196	1081	1076.95
7	参観	319	214	2804.42	32	登校	203	1508	1006.22
8	本校	271	172	2400.54	33	行事	230	2486	982.81
9	児童	589	7936	2280.52	34	教室	315	6654	963.67
10	授業	510	5077	2256.22	35	家庭	409	13306	942.18
11	下校	250	157	2217.67	36	運動	463	18080	924.13
12	献立	294	506	2196.72	37	片栗粉	144	483	917.39

² BCCWJ の語彙表は下記のサイトで公開されている。

http://www.ninjal.ac.jp/corpus_center/bccwj/freq-list.html

13	保護	709	15156	2154.53	38	集金	121	196	915.44
14	人参	381	2034	2115.86	39	協力	429	15437	914.91
15	懇談	299	842	1994.20	40	指導	439	17095	878.24
16	夏休み	334	1892	1819.30	41	挨拶	270	5334	858.31
17	渦	256	630	1765.15	42	鶏肉	144	702	822.97
18	御飯	423	5570	1655.33	43	大会	294	7243	819.20
19	学級	275	1476	1524.49	44	サラダ	206	2629	817.80
20	冷	178	187	1460.24	45	ボランティア	233	3913	808.81
21	玉葱	279	2087	1379.46	46	全校	111	230	796.58
22	活動	723	31440	1315.43	47	予定	362	13518	749.46
23	体育	308	3444	1297.08	48	キャベツ	158	1348	743.34
24	年	117	246629	1251.79	49	市	706	50766	741.85
25	皆	828	44925	1215.47	50	算数	122	501	734.25

食べ物や料理に関する語が多いのは、学校配布物に「給食だより」が含まれており、比較的にその割合が大きいことが影響している。バランスの取れた語彙表を作成するためには、学校配布物のジャンルもできるだけ偏らないように考慮する必要がある。

3. 2 学校配布物の動詞

前節と同じ学校配布物（対象コーパス）を用い、「動詞 - 一般」と「動詞 - 非自立可能」を取り出し、1,032語を得た。そして前節と同じく、BCCWJで頻度100以上に限定したところ、981語を得た。

この981語についても前節と同じく、学校配布物（対象コーパス）とBCCWJ（参照コーパス）のそれぞれの粗頻度の相関係数を求めたところ、0.953と極めて高い値となった。これは学校配布物（対象コーパス）とBCCWJ（参照コーパス）での見出し語の頻度傾向がほとんど変わらないことを示唆している。

表4 学校配布物の動詞（特徴度上位50語）

	見出し語	配布物	BCCWJ	特徴度		見出し語	配布物	BCCWJ	特徴度
1	言う	712	803148	2900.01	26	整える	75	3169	139.90
2	願う	917	20349	2725.95	27	呼ぶ	15	28626	139.18
3	下さる	1413	72328	2203.68	28	守る	152	12235	136.04
4	有る	1593	956900	1716.45	29	迎える	123	8541	135.25
5	頂く	706	35439	1122.02	30	目指す	124	8895	130.63

6	取り組む	276	5835	843.93	31	深める	59	2046	129.47
7	致す	481	23823	774.47	32	渡す	90	5041	127.62
8	居る	3163	1121183	457.26	33	分かる	165	88041	126.61
9	為る	8336	2563860	421.10	34	知れる	41	39167	122.82
10	頑張る	246	11856	406.65	35	煮る	58	2156	120.49
11	噛む	112	2391	340.48	36	見る	604	223222	109.73
12	過ごす	176	7211	337.22	37	誇る	21	136	109.31
13	仕舞う	118	98178	269.15	38	与える	9	20769	108.99
14	対する	67	69928	236.55	39	努める	80	4775	105.72
15	遣る	118	90348	225.73	40	学ぶ	104	7685	105.24
16	申し上げる	167	9918	221.96	41	申し込む	50	1836	104.9
17	食べる	332	32739	211.52	42	死ぬ	4	16081	99.10
18	於く	110	84355	211.14	43	得る	55	40470	96.34
19	漉す	43	484	180.54	44	繋げる	35	904	94.60
20	買う	6	27664	175.52	45	越える	2	13692	92.88
21	始まる	186	14746	170.28	46	会う	10	18940	91.78
22	付ける	379	44215	169.65	47	育む	34	908	89.91
23	成る	1718	564043	148.29	48	求める	14	20798	89.29
24	知る	477	64503	144.42	49	変わる	32	29331	88.53
25	干す	53	1393	141.63	50	通す	106	8978	87.72

上位 50 語の特徴度が名詞に比べて相対的に低いことがわかる。また、表 4 を見ても、学校配布物に使われやすいと思われる動詞であるとは言えるが、それが学校配布物の特徴を示していて、語彙シラバスに盛り込む必要があるかと言われれば、そうでもないだろう。

基本的に動詞については、対象者のニーズを考慮しない一般的に頻度が高い語彙シラバスを共用してもよいだろう。

3. 3 学校配布物の形容詞／形状詞

前節と同じ学校配布物（対象コーパス）を用い、「形容詞」と「形状詞」を取り出し、343 語を得た³。そして前節と同じく、BCCWJ で頻度 100 以上に限定したところ、328 語を得た。

この 328 語についても前節と同じく、学校配布物（対象コーパス）と BCCWJ

³ UniDic における「形容詞」は日本語教育での「イ形容詞」、「形状詞」は「ナ形容詞」に相当する。

(参照コーパス)のそれぞれの粗頻度の相関係数を求めたところ、0.838と名詞よりも動詞寄りの値となった。

表5 学校配布物の形容詞／形状詞(特徴度上位50語)

	見出し語	配布物	BCCWJ	特徴度		見出し語	配布物	BCCWJ	特徴度
1	無い	499	458395	1390.57	26	安い	2	8366	52.01
2	大切	342	12962	699.56	27	大丈夫	4	9295	48.90
3	楽しい	253	12606	404.87	28	嫌	11	12746	46.70
4	健やか	50	506	219.65	29	蒸し暑い	14	277	44.46
5	暑い	107	4249	210.53	30	色々	91	10321	43.45
6	沢山	109	4884	192.91	31	良い	625	198994	42.17
7	燦々	26	114	153.47	32	小さい	14	13114	40.39
8	みたい	8	24220	138.87	33	可愛い	6	8977	38.69
9	ドライ	38	799	116.55	34	そんな	2	6523	38.18
10	遅しい	34	884	91.50	35	そう	86	38826	37.87
11	可能	49	36362	87.47	36	辛い	2	6340	36.82
12	非常	10	17652	82.82	37	確か	22	15282	33.78
13	欲しい	178	20740	79.89	38	同じ	15	12402	33.75
14	素晴らしい	81	6323	75.85	39	少ない	49	25175	33.48
15	重要	22	22629	75.61	40	近い	13	11435	33.17
16	爽やか	36	1485	68.57	41	ださい	11	251	32.14
17	豊か	76	6132	67.77	42	若い	25	16013	31.55
18	好き	30	24502	65.82	43	強い	60	28368	31.18
19	明らか	6	12887	65.81	44	多様	2	5540	30.93
20	早め	41	2101	63.79	45	様	1297	373477	30.16
21	寒い	75	6247	63.71	46	有名	6	7581	29.49
22	高い	73	40495	62.74	47	小まめ	14	516	29.28
23	正しい	88	8436	59.05	48	高度	2	5299	29.17
24	古い	1	8107	56.24	49	木目細か	10	236	28.62
25	めでたい	39	2169	55.72	50	クリーン	16	735	27.67

形容詞は、語数そのものについては少ないため、語彙シラバスの中で大きな割合を占めることはない。しかし、動詞に比べると、やや特徴的な語も含まれるため、語彙シラバスを考えるにあたっては考慮する必要があると言える。

4. 連絡帳に現れる語の特徴度

次に対象コーパスとするのは、山形県山形市などで収集した外国出身の母親と保育園・幼稚園との連絡帳である。MeCab0.98 及び UniDic1.3.12 を用いて形態素解析をした結果、延べ語数は 8,456 語、異なり語数は 1,016 語である（いずれも短単位）。データの総量が少ないことと、前節での検討結果から、連絡帳については名詞のみを対象とする。「名詞 - 普通名詞」が 473 語あり、BCCWJ で頻度 100 以上に限定して 420 語を得た。相関係数は 0.487 である

表 6 連絡帳の名詞（特徴度上位 50 語）

	見出し語	配布物	BCCWJ	特徴度		見出し語	配布物	BCCWJ	特徴度
1	先生	61	26992	288.99	26	子供	21	59470	29.57
2	世話	33	4249	235.87	27	学期	4	662	26.60
3	願い	32	4408	224.46	28	最近	12	21299	26.05
4	園	30	8739	166.34	29	帰り	6	3364	25.67
5	有り	22	5177	131.14	30	一緒	12	22245	25.16
6	糊	13	747	113.59	31	親	10	14637	25.06
7	墨	13	990	106.36	32	表	12	22851	24.61
8	日	45	87635	90.72	33	ママ	6	3711	24.55
9	友達	20	12796	80.56	34	払い	4	885	24.32
10	教え	11	1537	76.82	35	遊	3	265	23.66
11	欠け	7	149	74.82	36	セント	4	1048	23.00
12	思い	16	12685	57.98	37	金曜	5	2648	21.93
13	様	2	373477	45.61	38	出し	4	1402	20.73
14	知恵	10	5567	42.92	39	毎日	8	11882	19.84
15	風邪	8	2712	41.98	40	寝	3	527	19.60
16	本当	19	35074	39.98	41	心配	8	12210	19.46
17	今日	19	35121	39.94	42	食後	3	578	19.05
18	母	14	18720	37.33	43	保育	6	6110	18.95
19	頼り	6	1308	36.65	44	便り	3	644	18.42
20	可	7	2478	36.14	45	人	1	156890	18.30
21	話	21	48837	36.12	46	年中	3	678	18.11
22	玩具	6	1385	35.98	47	此の頃	4	2021	17.91
23	一杯	9	7136	32.61	48	蛭	2	113	17.54
24	バス	7	3537	31.34	49	真	5	4263	17.44
25	パパ	6	2284	30.13	50	大根	4	2151	17.43

データ量が少ないため、偏りはあるが、学校配布物と比較してみると、より身近な語が多くなっている。とりわけ、表 6 については、連絡帳という性格上、読むだけではなく、書く必要がある語としてリストアップされている点に注目したい。理解だけでよいのか、産出を見据えるのかは、語彙シラバスを検討する上で重要な課題である。

5. 形態素解析で「見えなくなる語」

表 3 と表 6 を見てみると、「会」や「市」、「園」、「人」など、一文字でどちらかといえば単語というより接尾語に近い語があることに気づく。UniDic の解析単位である短単位は、単純語など短い単位を基本としているため、合成語が多い専門用語の分析には適していない面がある。語彙シラバスを検討するには、短単位よりも長い単位が必要である。

UniDic には短単位のほかに、長単位がある。長単位は概ね文節を実質語部分と機能語部分に切り分けた単位である。小澤ほか(2014:582)の例を挙げると以下のようになる。

- (1) 【文】 固有名詞に関する論文を執筆した。
- (2) 【長単位】 固有名詞／に関する／論文／を／執筆し／た／。
- (3) 【短単位】 固有／名詞／に／関する／論文／を／執筆／し／た／。

名詞部分が複合名詞化し、「に関する」のような複合格助詞が取り出され、サ変動詞も結合することがわかる。ところが、長単位での形態素解析は、現段階では容易に使える状態にはなっていない⁴。参照コーパスである BCCWJ は既に長単位で形態論情報が付与されているものが公開されているが、対象コーパスについては、自分で形態素解析をする必要があるため、長単位の形態素解析辞書がなければ、長単位で形態素解析することができない⁵。

そこで、学校配布物のうちの一部を使って、短単位での形態素解析結果のうち、「名詞」が連続した語を強制的に結合することにより、短単位で「見えなくなる語」について考察してみる。使用したデータは、学校配布物のうち約 1/20、

⁴ 小澤ほか(2014)では、機械学習に基づく長単位解析ツール Comainu を発表している。しかし、Comainu はあくまで機械学習に基づいた長単位自動付与ツールであり、長単位の形態素解析辞書ではない。

⁵ 田中・近藤(2011:91)では、「短単位で切り出されたものを長い単位にまとめる規則（例：熱中/症→熱中症、起立/性/低/血圧→起立性低血圧）を独自に立て、擬似的な長単位で集計する作業を行った」とあるが、残念ながら報告書には「擬似的な長単位」の詳細が書かれていない。

短単位で 23,600 語分である。これを名詞結合した結果、16,729 語になった。3 節の「学校配布物」と区別するため、このデータは「学校配布物の一部」と呼ぶ。そして、前節までと同じように学校配布物の一部(対象コーパス)と BCCWJ 長単位(参照コーパス)を突き合わせ、特徴度(対数尤度比)を計算した。ただし、ここでは BCCWJ 長単位で 100 語という制限はかけていない。

表 7 学校配布物の一部の複合名詞(特徴度上位 50 語)

	見出し語	配布物の一部	BCCWJ 長単位	特徴度		見出し語	配布物の一部	BCCWJ 長単位	特徴度
1	連絡帳	27	59	191.43	26	自由帳	3	2	26.39
2	五限	14	8	125.73	27	体操服	5	33	25.85
3	プール開き	10	6	89.24	28	提出物	4	14	25.19
4	時間割	13	59	75.94	29	スクール水着	4	14	25.19
5	学習教室	7	3	65.06	30	元気一杯	7	111	25.02
6	運動会	19	519	49.49	31	御知らせ	15	718	24.95
7	四月	8	10590	42.45	32	御用意	4	18	23.43
8	個人懇談	5	5	41.36	33	配布物	3	5	22.56
9	自転車安全教室	5	7	38.94	34	柿山伏	3	5	22.56
10	ミニトマト	10	146	37.25	35	生活目標	3	6	21.70
11	夏休み前	6	33	33.00	36	サマースクール	4	25	21.08
12	実行委員	7	63	32.25	37	自分達	4	5246	20.91
13	水泳帽子	4	5	31.82	38	一冊目	4	26	20.79
14	汗吹きタオル	4	5	31.82	39	色鉛筆	5	68	19.27
15	授業参観	6	40	30.92	40	五年生	9	384	16.64
16	夏休み中	7	70	30.90	41	四年生	10	489	16.30
17	登校日	4	6	30.73	42	学校生活	6	166	15.50
18	御家庭	9	165	29.81	43	火曜日	8	348	14.54
19	鍵盤ハーモニカ	4	7	29.78	44	各教室	3	26	14.03
20	六年生	14	512	29.47	45	予備日	3	27	13.82
21	挨拶運動	5	23	29.09	46	御協力	14	1054	13.80

22	一年生	20	1109	28.67	47	生活科	4	69	13.69
23	学級懇談会	4	9	28.17	48	施設見学	3	29	13.43
24	二冊目	6	57	27.05	49	御都合	4	78	12.81
25	三年生	18	956	27.00	50	学習面	3	33	12.72

表 7 は短単位が少なくとも 2 以上結合した語であるため、短単位による形態素解析で「見えなくなる語」である。さらにいえば、BCCWJ 長単位にも存在しない複合名詞も 14 語あった。

表 8 学校配布物の一部の複合名詞のうち長単位にも存在しない語

見出し語	配布物の一部	見出し語	配布物の一部
赤白帽子	7	標準服	3
学年懇談会	5	水泳学習	3
連絡袋	5	体育服	3
児童個人票	4	校区探検	3
体育袋	4	音読カード	3
飼育活動	4	緊急連絡カード	3
計算カード	4	学級委員さん	3

6. まとめ

本発表では、学校配布物と連絡帳を対象コーパスとし、BCCWJ を参照コーパスとして、子どもを持つ外国人を対象者としたときの語彙シラバスのニーズを探索した。これにより、子どもを持つ外国人の語彙シラバスについて、データに基づく議論の下地は見えてきた。

しかしながら、実際の語彙シラバスに落とし込むにあたっては、データになくとも必要な語はないか、逆にデータにあっても不必要な語はないか検討する必要がある。語の補充については、シソーラス（類語辞典）などの検討が必要であろうし、語の削除については、子どもを持つ外国人やその周囲の関係者の主観判断を仰ぐ必要もある。語彙シラバスの作成は簡単なことではないが、データに基づいた語彙シラバス作成に向けて努力を続けていきたい。

謝辞

学校配布物データを提供して下さった松田真希子先生（金沢大学）、本田弘之先生（北陸科学技術先端大学院大学）、連絡帳データを提供して下さった内海由美子先生に深く御礼申し上げます。連絡帳データについては、プライバシー保護のため、文字化データそのものの提供ではなく、形態素解析済みデータ

での提供を受けたことを申し添えておきます。

参考文献

- 石川慎一郎 2006 「言語コーパスからのコロケーション検出の手法－基礎的統計値について」『統計数理研究所共同研究レポート』190, 1-28
(http://language.sakura.ne.jp/s/ilaa/ishikawa_20060328.pdf)
- 石川慎一郎 2012 『ベーシックコーパス言語学』ひつじ書房
- 小澤俊介・内元清貴・伝康晴 (2014) 「BCCWJに基づく長単位解析ツール Comainu」『言語処理学会第20回年次大会発表論文集』pp.582-585
- 大学英語教育学会基本語改訂委員会(編) 2003 『大学英語教育学会基本語リスト：JACET List of 8000 Basic Words』大学英語教育学会.
- 田中牧郎・近藤明日子(2011) 「難解用語の抽出と序列化におけるコーパスの利用－医療用語を例に－」『言語政策に役立つ、コーパスを用いた語彙表・漢字表等の作成と活用』特定領域研究「日本語コーパス」言語政策班
- 寺嶋弘道 2009 「日本語教育語彙を選定するための統計的指標--尤度比検定、カイ2乗検定、イエーツの補正公式の特徴」『ポリグロシア』17, 立命館アジア太平洋研究センター, 71-83
(<http://r-cube.ritsumei.ac.jp/handle/10367/895>)
- Hunston, S. 2002. *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Leech, G., Rayson, P., & Wilson, A. 2001. *Word frequencies in written and spoken English: based on the British National Corpus*. London: Pearson Education.

日本語学習者から見た語彙シラバス

劉 志偉 (首都大学東京)

liuzw@tmu.ac.jp

1. はじめに

カタカナ語が学習者にとっての語彙習得の難点の1つであることは日本語教育従事者の共通認識であろう。本発表は、日本語学習者(学習経験者)〔以下、学習者と称する〕の立場から見た語彙シラバスの一環としてのカタカナ語について検討を行う。

2. 用語について

2.1 カタカナ語

日本語学の見地からすれば、カタカナ語は俗称とされているが、本発表は(1)近代以降に日本に入ってきた中国語も外来語に含まれること(沖森編 2010:88、木村義之執筆)(2)欧米語からの借用語のみならず、大和言葉に区分されるオノマトペや混種語についても言及すること、の2点から記述の便宜上「カタカナ語」という名称を用いることとする。

3. 問題の提起

3.1 日本語母語話者にとってのカタカナ語

カタカナ語は、時折社会問題にもなりかねない。2014年6月12日岐阜県の男性(72)がNHKを相手取り、公共放送にカタカナ語(「外来語」)が多用されていることにより精神的苦痛を受けたとして裁判を起こしたことが記憶に新しい。

年配の方のみならず、若い世代にとってもカタカナ語について問題がないわけではない。2015年2月1日午前11時頃、NHKで放送された菅義偉官房長官による人質事件に関する会見の画面に「一方的なプロパガンダ」¹という字幕が映っていた。20代と30代の友人がそれを理解しきれなかったようであった。

カタカナ語について言語学の専門家の見地も二分しているように思われる。分かりにくい「外来語」について言葉遣いを工夫して、わかりやすく伝える外来語言い換え案の活動が2006年8月～2010年6月の間に行われていた。一方、中国文化が浸透していた時代に本来外来語であった漢語の日本語化は、時代の需要に合った方法であることと同様、アメリカ等の欧米文化が浸透している目下においてはカタカナの外来語の定着を待つしかないこと、また言い換えにより漢語の同音異義語の問題を大きくしてしまう恐れがあること、と主張する山口(2006)がある。

¹ propaganda : 情報戦、心理戦、宣伝戦、世論戦。

3.1 日本語学習者にとってのカタカナ語

個人的には、カタカナを含む現代日本語の表記体系が日本人の器用さの現れの1つと見ることができると考えている。しかし、一学習者としてはやはりカタカナ語の習得が難しい。

上級以上に限って言えば、(旧)日本語能力試験出題基準改訂版(国際交流基金 2004)や樋口ほか(2011)に挙げられているカタカナ語は、日本での実生活または進学のためのものとしては十分なものであるとは到底言えない。

4. 本発表の概要について

4.1 母語の違いに応じて考える必要があることから

中国語を母語とする日本語学習者の立場から私見を述べる。

4.2 学習レベルに応じて考える必要があることから

本発表は上級以上でニア・ネイティブレベルを目指す学習者のカタカナ語教育について考える。ここで「ニア・ネイティブ」と称しているのは、学習者が日本生まれ日本育ちや幼少時から日本に滞在し続けているといった特殊な条件に当たらない限り完全なる日本語母語話者にはなり得ないことによる。

4.3 本発表の目的と目標

本発表は、学習者にとってカタカナ語のうちどのようなものが難しいか、カタカナ語の学習過程においてどこが難しいかを示すと同時に、学習者の要望を込めた解決案を提示する。

5. 本発表で扱う資料について

本発表は、発表者自らの日本語学習メモを研究資料として用いる。これは、発表者が来日後、日頃日本語母語話者と接しながら日本語を体得する中で、分からなかった表現や気になった表現を常に手帳に記録してきたものである。記録期間は2003年～2014年である²。ただし、2014年は本プロジェクトの関係で意図性をもってメモをとっている面がある³。またその記録媒体も手帳ではなく、携帯のメモ帳を使用している。

学習メモを精査すると、その内訳(延べ語数)は概ね語彙(1427)と文法(340、問題意識を含む)とに分けられる。その詳細は表1の通りである。

² メモを取り始めた2003年は、発表者が旧日本語能力試験一級を取得してから約4年後であったことから、本発表で言及する学習レベルに合致するものと考えられる。

³ 従来と同様、その場で気になる表現(メディア等)は勿論、さらに過去に経験してできなかった語彙を思い出したものや、当時は難しかったであろう表現も取り入れている。

表 1 学習メモ (2003 年～2014 年)

	語彙	詳細	文法	詳細	問題意識	小計	バックグラウンド
2003	15	・キャラクター・関西弁	4		0	19	来日したばかりで、手書きの図で日本のキャラクターを教えてもらったことが印象的であった。
2004	227	・カタカナ語が多い・品位・暮らし/ファッション/レジャー/若者	50	・関西弁・ンの体系	12	289	高い学習意欲があった。中国国内の教科書では学習できなかった語彙や表現に強い関心を示していた。2005 年は就職活動を行っていたためか、やや数が少ない。
2005	86	・アクセント・歌手名(カタカナ表記)・若者言葉	12		12	110	
2006	404	・アクセント・歌手名(カタカナ表記)・カタカナ語・慣用句・教育/教養・暮らし/レジャー/若者	17		0	421	
2007	0		2		3	5	後期課程に進学したものの、将来に対する不安や、研究と生活を両立することに苦悩する。
2008	135	・書き言葉・カタカナ語	17		2	154	研究に専念することを決意し、再び従来の姿勢に落ち着く。論文を書くために書き言葉に注目。
2009	17		0		14	31	学習の安定期に入る(?) 習慣的に気になる表現などをわずかに記録。また、自分自身が上手く説明できないような内容を今後の研究のネタ(問題意識)として記していた。
2010	15		24		7	46	
2011	8		26		1	35	
2012	0		4		9	13	論文のネタ(問題意識)としてメモを取っていた。
2013	0		14		2	16	
2014	520	・アクセント・うる覚え・誤推測・カタカナ語・教育/教養/暮らし/ネット/ファッション/レジャー/若者	95	・接続形式・通時論の視点	13	628	プロジェクトの関係で意図的にメモを取るようになっている。様々な媒体を通して気になった表現、記憶(過去の学習経験)、主観(中国人留学生にとっては難しいであろう表現)、友人の指摘・協力
小計	1427		265		75	1767	

表 1 の一番右側にある「バックグラウンド」は、発表者の日本語学習ライフヒストリーに基づくものである。語彙の「詳細」の列を確認すると、カタカナ語が常に語彙学習の難点となっていることが分かる。

6. どのようなカタカナ語が難しいか

6.1 学習者にとって難しいと考えられるカタカナ語—学習メモを手がかりに— 学習メモを手がかりに、以下学習者にとって難しいと考えられるカタカナ語

の一覧を挙げる。カテゴリーは発表者の主観に基づくタグによる⁴。なお、()内は記録した年を指す。波状の下線を引いた単語は、とくに学習者にとって難しいものを表している。

6.1.1 教育関係

- ・化学 (物質等)
 - ミネラル (2006)
- ・作文記号
 - セミコロン (2004)
- ・自然
 - エルニーニョ (2004)、ストーム (2006)、スコール (2006)、マーズ (2006、2014)、ジュピター (2014)、ネプチューン (2014)、ビーナス (2014)、マーキュリー (2014)
- ・植物
 - チューリップ (2014)、ポプラの枝 (2014)
- ・数学 (単位等)⁵
 - ヘクトパスカル (2006)
- ・生物
 - アルマジロ (2006)、マングローブ (2006)、コブラ (2008)、ピラニア (2014)、オランウータン (2014)

6.1.2 教養関係

- ・貨幣
 - フラン (2006)、マルク (2006)、ユーロ (2006)、ポンド (2008)、ルーブル (2008)、インドルピー (2008)
- ・国名、地理関係
 - トルコ (2004)、スパニッシュ (2006)、ワイキキ (2006)、クリミア (2014)、ジュネーブ (2014)、スコットランド独立 (2014)、ベルギー (2014)
- ・作品名
 - ブリュッセル小便小僧 (2004)、モナリザ (2006)、ダヴィデ像 (2006)、マーライオン (2014)、ムンクの叫び (2014)
- ・人物名
 - ベートーヴェン (「ベートーベン」、2006)、ゴッホ (2006)、レオナルド・

⁴ ここでは敢えて国立国語研究所 (2004) 『分類語彙表』にある詳細な分類を取らない。網羅的に分類することにより、学習者にとって特に難しいカテゴリーが逆に埋もれてしまうからである。学習者の見地から学習者の主観に基づく区分を行ったのである。

⁵ 後述の IT 用語とも関わる。森山 (2012: 74-75) に、古代ギリシア語やラテン語からの学術用語が日常語として広まってきているとし、数学用語 (アルファ・ベータ・ガンマ・デルタ・パイ) や単位 (メガ・ギガ・テラ) について述べた箇所がある (石田美代子執筆)。

ダ・ヴィンチ(2006)、ミケランジェロ(2006)、オーギュスト・ロダン(2006)、
ダンテ(2006)、ヘーゲル(2006)、ガリレオ(2014)、ペガサス(2014)、
ルノアール/ルノワール(2014)

・世界遺産

モアイ像(2006)、スフィンクス(2006)、アンコールワット(2014)、グラ
ンドキャニオン(2014)

・歴史関係

ルネッサンス(2006)

6.1.3 暮らし関係

・加工食品

フライドポテト(2006)、セリアル(シリアル、2006)、タカピオ(「タピオ
カ」、2008)、プルーン(2014)、ピスタチオ(2014)

・果物

ドリアン(2006)

・健康

ペンダコ(2004)、バイパス手術(2014)

・商業施設

ロッテリア(2004)、サイゼ(2014)

・商品名

インフォバー(2004)⁶、ソフナー(「ソフラン」、2006)、クレンザー(2006)

・食材

キャビア(2006)、トリュフ(2006、2014)、フォアグラ(2006)

・生活行動関係

MT/マニユアルトランスミッション(2005)、AT/オートマティクトラ
ンスミッション(2005)、ウエルダン(2006)、レア(2006、2008)、アル
デンテ(2014)

・生活用品

スプレー(2004)、ブリーチ(2006)、ペティナイフ(2006)、ルーズリーフ
(2006)、ワックス(2006)、メンソール(煙草、2008)、ハンモック(2014)、
ブラインド(2014)、クレンジング(2014)、コンディショナー(2014)、シ
ャンプー(2014)、トリートメント(2014)、パラソル(2014)、プリペイド
(2014)、マウスウォッシュ(2014)、リンス(2014)

・調味料

マスタード(2004、2008、2014)、タバスコ(2014)

・日常

イルミネーション(2006)、ハローワーク(2006)、ホームルーム(2006)、

⁶ ガラケーの一種である。こうした固有名詞を語彙として取り上げる必要があるかどうかについて検討する余地があるが、学習者が気になっていたことを示すためにここで敢えて挙げることにした。

バイパス (高速道路、2014)

- 宝石

エメラルド(2006)、サファイア(2006)、ダイヤモンド(2006)、ルビー(2006)

- ペット (犬、猫)

チワコ (「チワワ」、2003、2014)、コリー (2006)、ブルドッグ (2006)、
アメリカンショートヘア (2014)、ゴールデンレトリバー (2014)、チャ
ウチャウ (2014)、トイプードル (2014)、フレンチブルドッグ (2014)、ミ
ニチュアダックスフンド (2014)、ラブラドルレトリバー (2014)

- 料理

スクランブルエッグ (2006)、ゴーヤチャンプル (2006)、グラタン (2014)、
グリル (2014)、ケバブ (2014)、シュラスコ (2014)、タコライス (2014)、
ドリア (2014)、ピラフ (2014)、リゾット (2014)

6.1.4 ファッション関係

- 小物

ヘアバンド (2004)、ダテ眼鏡 (2004)、ミサンガ (2008)

- 服装 (模様)

網タイツ(2004)、スカーフ(2004)、リュックサック(2004)、学ラン(2004)、
セーラー服 (2004)、ブレザー (2004)、B系 (2004)、ガスロリ (「ゴスロ
リ」、2004)、ブリーフ (2004)、インナー (2005)、ウェア (2005)、ニッ
ト (2005、2014)、キャップ (2005、2014)、ポロシャツ (2005)、ダンパ
ーカー (「ダウンパーカー」、2006)、マント (2006)、パンスト (2006)、カ
ジュアルシューズ(2014)、グレースーツ(2014)、スキニージーンズ(2014)、
ストライプ (2014)、タキシード (2014)、タンクトップ (2014)、ディレク
ターズスーツ (2014)、デッキシューズ (2014)、トップス (2014)、ネック
ウォーマー (2014)、ハット (2014)、ボーダー (2014)、ボトムス (2014)、
レギンス (2014)、ブラックスーツ (2014)

- ブランド名

プラダ (2004)、ディオール (「ディオール」、2004)、ジバンシー (2004)、ニ
コス (2004)、シャネル (2004)、ブルガリ (2004)、グッチ (2004)、コー
チ (2004)、カルティエ (2006)、アンダーアーマー (2014)、ナイキ (2014)、
ラコステ (2014)

- ヘアスタイル

アフロ (2014)

6.1.5 レジャー関係

- 玩具

シーソー (2006)、ジャングルジム (2006)、ルービックキューブ (2014)、
クラッカー (2014)、マトリョーシカ (2014)

- 娯楽/ゲーム

カジノ (2004)、スロット (2004)、パラパラ漫画 (2004)、ホラー映画 (2004)、
イントロクイズ (2006)、リンボーダンス (2014)

- ・ イベント
 - ハロウィン (2004)、カーニバル (2004)、エイプリルフール (2008)
- ・ 楽器
 - リコーダ (「リコーダー」、2014)
- ・ 音楽 (歌手名)
 - エグザイル (2005)、エブリリトルシング (2005)、ゾーン (「ゾーン」、2005)、ミーシャ (2005)、バンプオブチキン (2005)、ミスタチルドレン (「ミスターチルドレン」、2005)、ドリームズカムトゥルー (2005)、ビーズ (2005)、サザンオールスターズ (2005)、ポルノグラフィティ (2005)、スピッツ (2005)、ドゥアズインフィニティ (2005)、ケミストリー (2005)、スピードマックス (「スピード・マックス」、2005)、ケツメイシ (2005)、ジャンヌダルク (「ジャンヌダルク」、2005)
- ・ キャラクター
 - アトム (2003)、スヌーピー (2003)、ドラえもん (2003)、ドラミちゃん (2003)、ハローキティ (2003)、ミッキー (2003)、オズ (「オズ」、2004)、テディベア (2014)、ドナルド (2014)、ピノキオ (2014)
- ・ スポーツ
 - 大リーグ (「大リーガー」、2004)、パラリンピック (2004)、ダンクシュート (2004)、バック宙 (「バク宙」、2004)、トラック (2004)、フィールド (2004)、ストレッチ (2004)、ムーンウォーク (2004)、ドラフト (2004)、スイング (2004)、スクワット (2006)、ゲレンデ (2006)、フリースロー (2006)、ロッククライミング (2008)、ラーニング (「ランニング」、2009)、スケルトン (2014)、ダブルダッチ (2014)、ネイマール (2014)、スロースタート (2014)

6.1.6 IT、ソーシャルネットワーク関係用語

ドメイン (2004)、レス数 (2004)、グーグル (2004)、カキコ (2004)、直レス/レス (2005)、アイコン (2006、2014)、ワイヤレス (「ワイヤレス」、2006)、フリーズ (2008)、スレッド/スレ (2008)、トリミング (2014)、シリ (2014)、カーソル (2014)、ギガ (2014)、コンテンツ (2014)、セキュリティ (2014)、テラ (2014)、バイン (2014)、プレビュー (2014)、メガ (2014)、マナーモードにする (2014)

6.1.7 若者言葉

頭テンパる (2004)、借りパク (2004)、スルー (2006)、ダベる (2008)、オイタする (2010)、オワコン臭 (2014)、ディスってる (2014)、ナルシスト (2014)

6.1.8 オノマトペの一部⁷

⁷ 中国語のオノマトペは日本語ほど発達しておらず、VR 構造等の構文で対応する場合が多々ある。つまり、中国語に対応しない日本語のオノマトペが存在するということである。また、同じ音に対する捉えかたが異なる点が顕著とも言

スッカラカン (2004)、ぎゃふんと言わせる (2008)
 コロコロリン (コオロギ、2008)、ジージー (アブラゼミ、2008)、リーン
 リーン (スズムシ、2008)

教育関係・教養関係については、学習者の専門といった個人差が認められよう。暮らし関係・ファッション関係・レジャー関係・IT 関係・若者言葉については、学習者の生活環境、交友関係、個人の趣味等の差を考え合わせる必要がある。オノマトペに関しては日中両言語の相違が大きな影響を与えるものと考えられる。しかし、上に挙げたカタカナ語と次に取り上げる形容詞に代表される抽象概念を表すカタカナ語のほとんどは、日本語母語話者にとってほぼ誰でも知っているような語彙である。ニア・ネイティブを目指すには避けて通れない内容である。具体的な学習上の問題点として、中国語は知っていても日本語で産出できないという問題点のほか、そもそも言葉を耳にしても物自体をイメージすることができない。これらのカタカナ語の習得が捗らないのは、次に述べる「漢字・漢語を頼りたがる心理」と「経験不足」とに起因するものと考えられる。

6.2 難しいと感じる理由

表2に示したように、カテゴリーとしての個々の語彙が母国語に存在するかどうかということと、経験の有無がカタカナ語の習得の成否に大きく関わるものと思われる。言い換えれば、カタカナ語の習得が難しいとされる主な要因に、中国語母語話者が漢語を頼りたがる心理によるものと経験不足によるものとの2点が挙げられる。

表2 カタカナ語の習得が難しい理由

共通存在するカテゴリーについて	経験済み、且つ語彙として習得	○
	経験済みだが、漢語を頼ってしまうため	△×
	経験不足のため、知る術がない ⁸	×
中国に存在しないカテゴリーについて	経験済み、且つ語彙として習得	○
	経験済みだが、難しい(対応する漢語がない)	△×
	経験不足のため、知る術がない	×
日本に存在しないカテゴリーについて	不問：基本的に使う必要がないため	

6.2.1 その一、漢字表記を頼りたがる—カタカナ語の形容詞・抽象概念を例に— 中国語にも欧米由来の借用語が存在するが、漢字表記を用いて表すことが—

える。例えば、携帯アプリ LINE の着信音の一種に関しては、発表者は「リンリンリン」としか聞こえない(再現できない)のに対し、日本語母語話者は口を揃えて「ピロリロリン」と言う。字面上からみても日本語のオノマトペが複雑であると捉えることができる。

⁸ トランプの「ハート」「ダイヤ」「クローバー(クラブ)」「スペード」「ジョーカー」が例として挙げられる(劉 2014)。

般的である⁹。往々にして、漢字そのものから事物をイメージすることができる。漢字表記を頼りたがることにより、カタカナ語学習の効率を憚るものの好例として挙げられるのがカタカナ語表記の形容詞や抽象概念である。()は漢字を伴った表記を用いた同義語である。学習者はカタカナ語よりこれらの語彙を好んで使うのである¹⁰。これらのカタカナ語に対して学習者は文脈依存で何となく理解しているつもりであるが、このことは寧ろカタカナ語の学習を疎かにしているのである。学習者は必ずしも原語からその意味を理解しているわけではない。漢字のように字形から意味の連想ができないため、カタカナ語を語彙として記憶するのは容易ではない。

・形容詞系¹¹

リーズナブル (経済的、2014)、ランダム (無差別、2008)、ヒステリック (異常な興奮状態、2014)、ハスキー (しゃがれた、かすれた、2014)、ナイーブ (素朴な、繊細な、2006)、たふ (頑丈な、2004)、スタイリッシュ (おしゃれな、2014)、カジュアル (普段の、2014)、アンティーク (骨董品、古めかしい、2004、2006)、アクロバチック / (超人的な、2006) 超カオスな (混沌な、2014)、デイリー (毎日の、日常の、2004)、タイト (きつい、2008、2014)、センシティブ (刺激的、2006)、ストイック (自分に厳しい、禁欲的、2006、2014)、シュール (奇抜な、2014)、エゴ (自己中、わがまま、2006)、エキゾチックな (つややか、なまめかしい、2006)、アンチ (反、2006)、アトランダム / アットランダム (無作為、手当たり次第、2008)、アクロバットな (超人的な、2006)、アグレッシブ (活発的、2014)

・抽象概念¹²

モノトーン (単調、2006)、ミーハー (もの好き、2014)、マンネリ (代わり映えしない、2004)、プレゼンス (提供、2004)、ブランク (空白、2014)、

⁹ 巧克力 (チョコレート)、沙发 (ソファー)、咖啡 (コーヒー) のような外来語の音を示す場合もあるが、(個人) 电脑 (パソコン) の意識や、英特网 (インターネット) のように音と意の両方を示すもののほうが圧倒的に多い。

¹⁰ 中国語母語話者はアルファベットでいう場合がある。「ジャック」「クイーン」「キング」「ショート」「トール」「グランデ」「ベンティ」がそれである。経験不足とも切り離せない関係にある。

¹¹ このほか、「スタティック (静的、2004)」「パッシブ (受動的、2006)」「エキセントリック (奇妙、奇怪、2014)」「レトリカル (修辭的、2008)」がある。

¹² このほか、「レファレンス (参照、2004)」「レイオフ (一時的な解雇、2006)」「メッカ (聖地、2006)」「マチュア (成熟、2004)」「ペイオフ (預金保険、2006)」「ブレーン・ストーミング (集団思考、2006)」「コンベンション (集会、2006)」「コンプライアンス (法令順守、2014)」「コンサイス (小型辞書の商標)」「コスモポリタニズム (世界市民主義、2008)」「レトリック (修辭、2008)」がある。

ファインプレー (すばらしい技、2014)、ニュートラル (中立、2004)、ダイジェスト (要約、2004)、スリル (恐怖、2004)、スランプ (調子の悪い時期、2014) スパルタ (厳しい、2006、2014)、ステータス (地位、2014)、ジェラシー (嫉妬、2004、2006)、シーソーゲーム (駆け引き、2006)、キックバック (見返り金、2004)、カースト (階級、2014)、オブラート (ぼやかす、2014)、オーシャンビュー (海の景色が見えること、2014)、エイジング (老化、2014)、アンチエイジング (若返り、2014) レセプション (公式な接見、2006)、レスポンス (反応、2004)、レジェンド (伝説、2014)、ユートピア (理想郷、2014)、モラトリアム人間 (先延ばし人間、2014)、メルヘン (ドイツで発生した散文による空想的な物語、2006)、マジョリティー (大衆、2004)、ポテンシャル (潜在能力、2006)、ベクトル (方向、指向性、2006、2014)、プラトニック (精神面のみのつながり、2006、2014)、バロメーター (指針、指標、2006)、パラレルワールド (並行世界、2014)、バザー (市場、2004)、パイオニア (開拓者、2014)、バーチャル (仮想、2004)、ネームバリュー (世間での知名度、2010)、ドッペルゲンガー (分身、2014)、テリトリー (縄張り、領域、2004)、デポジット (預け金、2014)、デトックス (解毒、毒素排出、2014)、デッドヒート (猛烈な接戦、2014)、デジャヴ (既視感、2006、2009、2014)、シルエット (影、輪郭、2014)、サブカル (副次文化、2006)、コントラスト (色相、2014)、クレーター (凹凸、2006)、オーバーヒート (いっぱいいっぱいという状態、手におえない、2008)、エゴイズム (自己中心主義、2006)、イントロダクション (入門、2004)、インスピレーション (直感、2014)、負のスパイラル (負の連鎖、2014)

漢字・漢語、即ちカタカナ語のような全くイメージの付かない表記以外のものを求めるという心理は¹³、カタカナ語の形容詞や抽象概念以外のものにおいても認められる。

・非抽象的概念¹⁴

レプリカ (模造品、2014)、リスナー (聞き手、2006)、ユース (青年、2006)、コレステロール (胆固醇、2004)、オーディエンス (観客、観衆、2008)、BGM/バックグラウンドミュージック (背景曲、2006)、ファインダー (枠、2006、2014)、ゴムパッキン (ゴム製品の一部、2008)、アダルトチルドレン (ピーターパン症候群の大人、2004)

・動詞

¹³ 逆の場合がないわけではない。例えば、ヒアリング上では「不正行為」より「カンニング」のほうが分かりやすい。語彙教育の成果によるものと思われる。カタカナ語を教材に取り入れる必要性の証左の一つである。

¹⁴ このほか「シネコン (複合映画館、2006)」「インジケーター (計測器、2006)」「エピテーゼ (医療用具として体の表面に取り付ける人工物のこと、2004)」が記されている。

スキップ（跳ね回る、2008）、アイコンタクト（目を合わせる、2004）

6.2.2 その二、経験不足—洋菓子の種類を例に—

生活習慣や文化の差によって、個々のカテゴリーにおいて使用する語彙に差が認められる。日本語母語話者にとって誰しもが知っているような固有名詞等であっても、学習者が生活環境の制限により日本に長く滞在していても経験していなければ知り得ないような語彙が数多く存在する。発表者もこれまでこれらのカタカナ語を聞いても全くイメージできないというような苦い体験をしばしば経験してきた。ここでは、発表者が来日後最も痛感させられたカタカナ語の語彙として、洋菓子の種類（関係の語彙）を例に取り上げる¹⁵。パン、ケーキ、クッキー、クリーム、チョコレートだけではニア・ネイティブに近づけない。

マシュマロ（2004）、ヨックモック（メーカー名、2014）、マドレーヌ（2014）、エクレア（2014）、カスタード（2014）、クランチ（2014）、コーンスターチ（2014）、コーンフレーク（2014）、タルト（2014）、パウンドケーキ（2014）、フィナンシェ（2014）、ブラウニー（2014）、マカロン（2014）、ラスク（2014）、生チョコ（2014）

6.3 解決方法に関する提案

形容詞・抽象概念や洋菓子の種類に代表されるようなカタカナ語の学習の躓きは、いずれもインターネット教材の開発によって解決できる。

前者について言えば、使用頻度等に基づいて、和語の形容詞を中心に多く取り入れている現行の教材も無論重要であるが、カタカナ語の形容詞をふんだんに取り入れた違ったタイプの教材をも、極力ネイティブレベルに近づこうとする学習者に用意すべきである。

後者に関しては、山内（2013）を発展させ、固有名詞・略語・若者言葉・流行語といった、学習者にとって経験不足に陥りがちな語彙¹⁶を図鑑式提示でインターネット教材（eラーニングの教材・学習材）に取り入れる。インターネット教材の開発により日本語教材の陳腐化という重大な問題点を解消することができるだけでなく、海外という学習環境の制限による経験不足の差を縮めることもできる。

インターネット教材の開発にあたっては、以下の注意点があろう。一つは、

¹⁵ 中国語の言い方が分からないような「マリネ」「ピクルス」等の食べ物や、「レア」「ミディアム」「ウェルダン（ウェルダム）」のような焼き加減の言い方も同様なことが言える。

¹⁶ 固有名詞までは教材に取り入れる必要がないというのはネイティブ教師側の視点である。寧ろこれらの語彙のほうが母語話者の手助けを最も要するものである。また、固有名詞等を取り入れるとキリがないという意見もあろうが、インターネット教材形式ではこの点も克服できるものと思われる。

上級以上ともなれば個々の専門教育のための語彙指導¹⁷よりは、基本意識・教養としての語彙へのシフトが重要である。その語彙選定は、日本の中学校・高校で用いられている社会や理科等の教材を視野に入れるべきである。学習者は、日本の中学校や高校の教育を受ける経験がないため、母語話者にとって誰でも知っているような基礎知識・教養の語彙の日本語の言い方を知る機会がほぼないからである。

7. カタカナ語学習のどこが難しいか（再現の難しさ）

6節では「漢字を伴う語彙を頼りたがる心理」と「経験不足」とが原因で中国語を母語とする学習者のカタカナ語の習得を難しくしていることについて述べた。同義または類義の和語や漢語語彙の学習に気を取られカタカナ語の学習が疎かになってしまう。その結果、経験不足による未習のカタカナ語はもとより、既習のカタカナ語に関しても産出するどころか、理解するのにも支障を来してしまう。これに対し、この節ではカタカナ語のいわゆる「産出」についても触れておきたい。ただし、ここでは「産出」という従来の用語を使用せず、「再現」を用いる。「再現」についてはカタカナ表記の語彙を産出する場合(7.1)と、アルファベットで示される原語表記をカタカナで日本語風に言う場合(7.2)とに分けて述べる必要があるからである。

7.1 見たことまたは聞いたことのあるカタカナ語（既習語彙を含む）をうまく再現できない場合

7.1.1 うろ覚え

・長音

パクチ（パクチー、2003）大リーグ（大リーガー、2004）、ブレザ（ブレザー、2004）、オソドックス（オーソドックス、2006）、クオリティ（クオリティー、2006）、オーバヒート（オーバーヒート、2008）、シュレッタ（シュレッダー、2014）、レパトリー（レパートリー、2014）、ボキャブラリ（ボキャブラリー、2014）

・促音

オッズ（オズ（の魔法使い）、2004）
ホチキス（ホッチキス、2014）、アクレシブ（アグレッシブ、2014）、シャールフル（シャッフル、2014）

・撥音

ラーニング（ランニング、2009）
プレミア（プレミアム、2014）、コマ（コンマ（何秒）、2014）

・濁音

ポルノサイド（ポルノサイト、2004）、マスタード（マスタード、2008）、

¹⁷ 専門日本語教育を否定するものではない。発表者自らの経験からすれば、受験、そして入学後、個人の努力と研究室の環境で専門語彙の学習は特に問題にならないものと思われる。

オブラード (オブラート、2014)、ヘッドフォン (ヘッドフォン、2014)、
ドライ ((家庭教師の) トライ、2014)、ドラック (脱法ドラッグ、2014)
アクレシブ (アグレッシブ、2014)、シュレッタ (シュレッダー、2014)

・拗音

トリフ (トリュフ、2014)

・ウの発音¹⁸

ダンパーカー (ダウンパーカー、2006)、オールランド (オールラウンド、
2014)

・交替

スチュエーション (シチュエーション、2004)、バレエーション (バリエー
ション、2014)、フアフア (フワフワ、2014)、ムヤムヤ (モヤモヤ、2014)
アクチブ (アクティブ、2006)、アクロバチック (アクロバティック、2006)、
デジャブ (デジャヴ、2006)

タカピオ (タピオカ、2008)、オードソックス (オーソドックス、2014)、
ボキャブラリ (ボキャブラリー、2014)

・リズム¹⁹

セレブティアー (セレブリティアー、2004)

・誤推測

バック宙 (バク宙、2004)、オリボイル (オリーブオイル、2014)、ヨレヨ
レ (ヨボヨボ (お爺さん)、2014)

・中国式英語の発音の影響

いつもサンシングを求めている (something、サムシング、2004)、フル (full、
フルー、2014)、ナチュラル (natural、ナチュラル、2014)

・中国語の影響

ハサフスタン (ハサ克斯坦、カザフスタン、2014)、アフニスタン (阿富汗、
アフガニスタン、2014)、クロチア (クロ地亜、クロアチア、2014)、モン
トリオ：蒙特利ル、モントリオール、2014)

7.1.2 アクセント

・中国語母語話者が頭高型のアクセントをしがちなことが原因？

ユニクロ (2006)、クレーター (2006)、ミサンガ (2008)、フリーズ (2008)、
アイフォン (2014)、ベクトル (2014) メンチ ((を切る)、2014)、レスポ
ンス (2014)

・頭高型にならないもの

イニシャル (2006)、ミネラル (2006)、シラバス (2014)、ヘンゲル (2014)

¹⁸ 中国語母語話者の場合、「バックグラウンド」「ダウンロード」「カウントダウン」のように、撥音の前にあるウの脱落が多いように思われる。また、「ブラウニー」が「ブランニー」のように「ウ」が撥音に聞こえる場合もある。

¹⁹ 「ボキャブラリー」を間違えて「ボキャブリー」と言ってしまうのもこのタイプに区分できる。

- ・中国式英語の発音の影響
 アイランド (2006)、イタリア (2006)、データ (2014)、ノーマル (2014)
- ・中国語の影響
 モナリザ (蒙娜丽莎、2006)
- ・アクセントによる意味の区別ができない場合
 ネット (網、2014)、ネット (インターネット、2014)

7.2 原語表記（主としてアルファベット）からカタカナ語表記へうまく再現できない場合

7.2.1 中国語母語話者にとって「原語表記からカタカナ語への再現」が必要なのか。

欧米語由来の借用語（以下、洋語と称する）の日本語化については、主として開音節化、促音化、日本語に存在しない子音・母音の日本語化（音素の置きかえ）等が挙げられる。現在は、英語（アメリカ英語、イギリス英語）由来の外来語が多いとされるが、通時論的に見ると、カタカナで表記されている洋語の由来は様々である。個々の語の綴りや前後の音韻環境により洋語の日本語化における一定の傾向が見られるものの、その詳細が煩雑かつ複雑である。例外が多く、和製英語や略語などの問題が加わると包括的にルール化するのは困難であるため、未だに日本語教育への応用は道のりが長いと言わざるを得ない。また、同じ洋語の日本語化に関しても通時論的なゆれ（ex. ザビエル→シャヴィエル、リーガン→レーガン）と共時論的なゆれ（ex. ワウリンカ／バブリンカ、マレー／マリー）が認められる。そして母語の違いに応じて日本語教育を考える立場からすれば、中国語母語話者は（1）ピンイン表記でアルファベットを扱うことはあるものの、発音や綴りのルール等が欧米諸語のそれとは大きく異なること（2）学習者には個々の洋語の出典の判別が無理であること（3）洋語なのか和製英語なのかの判別が容易でないこと、等の点から原語表記から日本語化のルールを中国語母語話者に徹底して教える意義はさほど大きくない。かといって全く必要がないというわけではない。発表者の実体験からすれば、デンモクで入力する歌手名や曲名、実際の会話の中に出てくる施設名等の固有名詞に限っては、原語表記からカタカナ語への再現が必要な場合があると思われる。

7.2.2 中国語母語話者による原語表記からカタカナ語への再現過程に見られる傾向

学習者にとって原語表記からカタカナ語へと再現する必要があるのは、歌手名といった固有名詞に限る。実際、発表者が2006年に書き記した歌手名やグループ名のメモが残っている²⁰。ここではこれを手がかりにその概要について触れ

²⁰ いくら日本語を勉強しても、日本人と日々接していても、アルファベット表記の歌手名、グループ名、曲名をうまく再現（デンモクで入力または検索）できなかったため、カラオケはそっちのけでデンモクを見ながら取ったメモである。

るが、その前に「経験不足」の有無を考え合わせる必要があることを断っておきたい。その好例として音楽グループ [AAA] が挙げられる。情報として知らなければ、学習者は [トリプルエー] ならぬ、[エーエーエー] または [スリーエー] と言ってしまうもおかしくないからである。実際にメモには発表者が当時 [TRF] を [チャフ] ²¹、[19] ²² を [ジューキュー] と産出してしまった例が見られる。一方、発表者が [175R] を [イナゴライター] ²³、[D-51] を [ディーゴイイチ] とほぼ正しく再現できたのは、情報として頭に残っていたためと推察される。従って、経験不足を解消することが原語表記からカタカナ語への再現にとって重要且つ有効な手段の一つであると言えよう。ただし、耳を頼りに獲得した「既知情報」の場合、特殊拍を中心とする様々な「うろ覚えのカタカナ語」が生じる場合がある²⁴。

無論、すべてのカタカナ語を経験不足の解消をもって解決することは事実上不可能である。未知の原語表記からカタカナ語を再現するにあたって、中国語母語話者に見られる傾向及び特徴を言及しておきたい。再現の手順としては、まず原語表記の語彙（ここでは英語を指す）を知っているものとそうでないものとに分けて考える必要がある。知っている英語の語彙の場合、中国語の英語の発音による負の転移が見られる一方、既習のカタカナ語や英語の日本語化のルールを積極的に応用する一面も認められる。

Every Little Thing エブリリトルスイング²⁵（中国語英語発音の影響）

Dreams Come True ドリムズカムトゥリー²⁶（中国語英語発音の影響）

X JAPAN エックスジャパン（既習のカタカナ語）

NEWS ニュース（既習のカタカナ語）

Coming Century カミングセンチュリー（既習ルール ing）

MY LITTLE LOVER マイリトルラバー（既習ルール le）

これに対し、知らない英語などの原語表記について中国語母語話者は綴りそのものに注目しつつ、開音節化や促音化を凶ろうとすると同時に、「仮名を表記する場合のアルファベット／ピンインの読み」や「第二外国語としての英語の影

²¹ 正しくは「ティーアールエフ」。

²² 正しくは「ジューク」。必ずしもカタカナ語で表記しなければならないものではないが、関連語彙として敢えてここで挙げることにする。

²³ 正しくは「イナゴライダー」。

²⁴ 例えば、「イナゴライダー (175R)」を「イナゴライター」（濁音）、「カーペンターズ (Carpenters)」を「カーペンダズ」（長音・濁音）、「カトゥーン (KAT-TUN)」を「カトゥー」（撥音・促音）、「ミーシャ (Misia)」を「ミーシャン」（濁音）と書かれている箇所が挙げられる。この箇所の原語表記とは無関係であるが、「トリュフ」を「トリフ」（拗音、2014）、「ソフラン」を「ソフナー」（R と N との子音交替、2006）のような聴覚によるうろ覚えのカタカナ語の誤用例がある。

²⁵ 正しくは「エブリリトルシング」。

²⁶ 正しくは「ドリームズカムトゥルー」。

響」等の影響を受けるものと思われる。(「 」内は正しい表記を表す)

- b.c.d.g.k.l.n.r.s.t 等の開音節化が見られる。

the brilliant green ザブリランドグリーン (「ザ・ブリリアントグリーン」)

L'Arc~en~Ciel ラルクエンクエル (「ラルクアンシエル」)

ZARD ザード

globe ゴロブ (「グローブ」)

H Jungle with t エッチジャングルウイズティー (「エイチジャングルウイズティー」)

Wink ウィンク

Crystal Kay クリスタルケイ

DEEN ディーヌ (「ディーン」)

LINDBERG リンドブラグ (「リンドバーグ」)

RIP SLYME リップスリメ (「リップスライム」)

The Beatles ザビトルズ (「ザ・ビートルズ」)

- 語中、語尾の c.cc.ch.ds.p.sh.x に対し、促音化する傾向がある。

Ricky Martin リッキーマーティン

Cocco コッコ

Def Tech デフテック

Kinki Kids キンキキッズ

RIP SLYME リップスリメ (「リップスライム」)

Dragon Ash ドラゴンアッシュ

I WiSH アイウィッシュ

MAX マックス

一方、促音化を要しない箇所に促音を入れてしまう場合もあった。

ABBA アッパー (「アバ」)

PUFFY パッフィ (「パフィー」)

KAT-TUN カットウー (「カトウーン」)

- 仮名のアルファベット/ピンインの読み (誤用例のみを挙げる)

Janne Da Arc ジャンネダアルク (「ジャンヌダルク」)

Britney Spears ブリトネィスペアズ (「ブリトニースピアーズ」)

Nina ニナ (「ニーナ」)

Le Couple レクオプル (「ルクプル」)

- 第二外国語としての英語の影響 (誤用例のみを挙げる)

Hide ハイデー (「ヒデ」)

Whiteberry ウイトベリー (「ホワイトベリー」)

MALICE MIZER マリックミザー (「マリスミゼル」)

- 中国語の影響 (誤用例のみを挙げる)

Faye Wong フィーオン

(「フェイウオン」) (王菲)

Whitney Houston ヒトニヒュストン

(「ホイットニーヒューストン」) (惠特尼休斯顿)

Mariah Carey マライヤキャリ²⁷

(「マライアキャリー」) (玛丽亚凯莉)

7.2.3 (部分) 解決に向けての提案

その一

中国語といえば、漢字・漢語というイメージが強いが、中国語の発音を表すためのピンインがアルファベットを用いており、中国語母語話者にとってアルファベットは決して馴染みのない表記ではない。ただし、その発音とスペリング上の規則が英語のそれと全く異なるため、アルファベットの綴りを目にした場合、英語表記としてというよりはピンインまたは仮名を表記する場合のアルファベットとして認識されやすい傾向があると思われる²⁸。このことが中国語母語話者による日本語の長音の再現を難しくしているのである。下記のような長音の有無が最も顕著な現れと言っても過言ではない²⁹。

- ・長音表記の箇所には長音できない場合 ar.ea.er.i.ie.iew.o.ou.ow

REIRA starring YUNA ITO レイアスタリングユナイトウ (「レイラ スターリング ユナイトウ」)

Sugar Soul feat.Kenji シュガソールフィートケンジ (「シュガソールフィーチャリング ケンジ」)

The Beatles ザビートルズ (「ザ・ビートルズ」)

SINGER SONGER シンガソング (「シンガソング」)

Nina ニナ (「ニナ」)

BENNIE K ベニケイ (「ベニケイ」)

FIELD OF VIEW フィールドオブビュー (「フィールドオブビュー」)

FIELD OF VIEW フィールドオブビュー (「フィールドオブビュー」)

m-flo エムフロ (「エムフロ」)

ECHOES エゴス (「エゴス」)

globe グロブ (「グロブ」)

ZONE ゾン (「ゾン」)

Whitney Houston ヒトニヒューストン (「ホイットニーヒューストン」)

THE YELLOW MONKEY ザイエロモンキー (「ザ・イエロモンキー」)

- ・長音不要な箇所には長音化してしまう場合 a.e.ua

²⁷ アとヤの交替は母語話者でも見られる。ずっと「イケア」を「イケヤ」と言ってきた日本人の友人がいる。発表者もそうであったが、「～屋」が影響したのも確かである。

²⁸ 日本語学習者の英語のレベルにばらつきがある。個人差にもよるが、必ずしも英語を得意としない学習者が多くいることは事実であろう。

²⁹ 清濁に関しては中国語に有声音と無声音との弁別があるため、英語等原語表記の日本語化における清濁については中国語母語話者にとって比較的問題にならない。

ABBA アッパー (「アバ」)

Wham! ワム (「ワム」)

shela シェラ (「シェラ」)

Tommy february6 トミーフェブラリ (トミーフェブラリ)

以上の誤用例を確認すれば、長音化の有無が中国語母語話者の学習者にとって難しいことは一目瞭然である³⁰。7.2.1 で述べた通り、原語表記からカタカナ語へ再現するための完全ルール化の必要はないが、中国語母語話者が間違いやすい傾向を見出し、ポイントを絞って提示することにより、学習者の誤用を減らすような指導者側の工夫は可能であろう。以下、「Y」を伴う原語表記からカタカナ語へ再現する際に長音になるかどうかを例に取り上げる。メモの精査を通して、ffv.ney.rey.ry 等「Y」を伴う原語表記を日本語化する過程において伸ばすべきところを伸ばさない傾向が強く認められる。

PUFFY パフイ (「パフイ」)

Whitney Houston ヒトニヒュストン (「ホイットニーヒュストン」)

Mariah Carey マライヤキャリ (「マライアキャリ」)

CHEMISTRY ケミストリ (「ケミストリ」)

Tommy february6 トミーフェブラリ (「トミーフェブラリ」)

JUDY AND MARY ジュディアンドマリ (「ジュディアンドマリ」)

小椋 (2013: 829) が述べるように、長音化にならない例外があることは事実である³¹。しかし、今回のメモを確認してみると、my.ray.rry.ry 等発表者が長音として認識した以下の例も存在する。

Tommy february6 トミーフェブラリ (「トミーフェブラリ」)

Whiteberry ワイトベリ (「ホワイトベリ」)

Coming Century カミングセンチュリ

そして、固有名詞としては、(melody. メロディ、Do As Infinity ドゥアズインフイニティ) のように長音記号を伴わないことになっているが、実際には長音化しても間違いとは言えない例を考え合わせると、「Y」を伴う場合は長音化になりやすいことをルール化したほうが原語表記からカタカナ語へ再現する際に役立つものとする³²。これは庵 (2011) の「100%を目指さない文法」の理念として実行可能な好例の一つである。

³⁰ 英語知識のおかげで長音が問題にならない例 (ar.ee) もある。**ZARD** ザード、**Ricky Martin** リッキーマーティン、**SPEED** スピード。

³¹ 小椋 (2013) は英語の語末が-dy.-gy.-ry-ty を対象としている。長音の有無についてはゆれが認められる。例えば、「カテゴリ」と「カテゴリー」もその1例である。

³² 発表者は Tully's を「タリズ」と発音していると指摘されることがあり、上記のルールを知っていれば、最初から「タリーズ」という発音ができていたかもしれない。ちなみに「ar」を伴う Starbucks (スターバックス) の長音箇所に関しては問題とならないようである。

その二

もう一つはアルファベットの日本語の発音を指導項目として取り上げるべき点を強く主張したい。この点にまつわる再現をうまくできたのは、発表者が意図的に習得したからである³³。

K ケー

T.M.Revolution ティーエムレボリューション (「ティーエムレボリューション」)

T-SQUARE ティースクエア

V6 ブイシックス

X JAPAN エックスジャパン

L←→R エルアール

ただし、Hに関しては、以下のような当てはまらないものがある。

H Jungle with t エッチジャングルウィズティー (「エイチジャングルウィズティー」)

HY エッチワイ (「エイチワイ」)

したがって、固有名詞においてはHを「エイチ」を用いる場合があるというルールを提示する必要があるだろう。

A	B	C	D	E	F	G
エー	ビー	シー	ディー	イー	エフ	ジー
H	I	J	K	L	M	N
エッチ	アイ	ジェー	ケー	エル	エム	エヌ
O	P	Q	R	S	T	
オー	ピー	キュー	アール	エス	ティー	
U	V ³⁴	W	X	Y	Z	
ユー	ブイ	ダブリュー	エックス	ワイ	ゼット	

その三

第二外国語としての英語学習（中国式英語）による負の影響についてはすでに前述の通りである。しかし、発表者自らのカタカナ語学習の過程を振り返ると、日本語化したカタカナ語の原語表記（英語語彙）を知らなければ、逆にカタカナ語の学習を疎かにしてしまう面もある。例えば、学習メモに次の2例がある。

いつもサムシングを求めている (2004)

リスナー (2006)

³³ 来日後、アルファベットの日本語の発音と中国語のそれとが異なることに気づき、友人に書いてもらったあと暗記したことを明確に覚えている（紙自体は紛失した）。

³⁴ 「V」の日本語風発音「ブイ」は学習者にとって難しいアルファベットの一つである。勿論、CDTV（カウントダウンティービー）のような固有名詞の場合もある。

「something」「listener」のような英語語彙に由来するものであるが、発表者はその場でこれらを理解できずにいた。発表者自らのこうした学習上の反省を踏まえた上で発表者は、カタカナ語の語彙提示について「カタカナ語→原語→和語和訳（若しくは中国語訳）」というモデルを提案する。なぜならば既知の英語であれば、日中英の三カ国語の共通情報として利用することができ、未知の英語でも、学習者は原語表記を確認することで記憶を深めることが期待され、カタカナ語の学習の手助けになる場合があるからである。

結びにかえて

本発表は発表者自身の日本語学習メモを手がかりに、カタカナ語に焦点を当て、中国語を母語とする学習者にとってどのような語彙が難しいか、学習上どこが難しいかについて私見を述べた。

難しいカタカナ語として、教育関係・教養関係・暮らし関係・ファッション関係・レジャー関係・IT（ソーシャルネットワーク）関係・若者言葉等のカテゴリーに区分することができる。欧米語を由来とする借用語のみならず、和語とされるオノマトペの一部（生き物の鳴き声）も学習者にとって難しいカタカナ語の1つである。これは中国語のそれとずれがあるためか、中国語に対応する語彙が存在しないか、のどちらかによる。また、固有名詞に限ったことではなく、カタカナ語の形容詞に代表されるように日本語化したカタカナ表記の抽象概念も中国語母語話者にとってとりわけ難しいものと思われる。さらにカタカナ語の学習を憚る理由としては、中国語を母語とする日本語学習者が漢字・漢語語彙を頼ろうとする心理と経験不足との2点が挙げられる。カタカナ語の学習を改善するためには、カタカナ語を多く取り入れた教材開発が求められるが、これまで重要視されてこなかった日本語教育界の社会問題とも言える教材の陳腐化を根本的に改善していくためには、発表者はインターネット教材の開発を学習者の要望としても強く希望する。そして、教材開発を行う際には、構成員の世代差を考慮に入れるほか、学習者（学習経験者）の視点をも取り入れて頂くことを懇願する。

一方、カタカナ語学習の難点について、カタカナ表記の語彙を正確に再現できない場合と、原語表記からカタカナ語へ再現できない場合との2つを取り上げた。前者については、特殊拍を中心とした学習者によるうる覚えとアクセントとに見られる多くの誤用例を列挙した。間違いやすい語彙を教材に意図的に取り入れることと、語彙指導の際はカタカナ語に限らずアクセントを併記することが求められる。そして、原語表記からカタカナ語への再現を要する場合は一部の固有名詞に限られることを明らかにした。英語を中心とする洋語から日本語への完全ルール化が困難であること、中国語を母語とする学習者が英語とは異なるアルファベットの使用上の規則を用いていること等複数の要因から、中国語母語話者への徹底した原語表記のルール指導に対しては否定的な立場を取る。そのかわり、部分的な解決に向けて以下の3点を提示する。1) 中国語母語話者に見られる、再現する際の誤用傾向を明らかにし、「Y」を伴う表記を

例に部分的な規則指導を提案する。2) アルファベットの日本語風発音を指導する必要性を強調する。3) 自らのカタカナ語学習の失敗談を踏まえた上で、「カタカナ語→原語→和語和訳（若しくは中国語訳）」という語彙提示のモデルを提案する。

参考文献

- 庵功雄（2011）「100%を目指さない文法の重要性」森篤嗣・庵功雄編『日本語教育文法のための多様なアプローチ』ひつじ書房
- 石綿敏雄（2001）『外来語の総合的研究』東京堂出版
- 稲垣滋子（1991）「外来語表記の基準と慣用」『日本語教育』74、日本語教育学会
- 榎垣実（1963）『日本外来語の研究』研究社
- 王伸子（2011）「中国語母語話者の日本語外来語彙習得に関する諸問題」『専修人文論集』88、専修大学
- 大曾恵美子（1991）「英語の音形の日本語化」『日本語教育』74、日本語教育学会
- 大滝靖司（2013）「日本語借用語における2種類の促音化」『国立国語研究所論集』6
- 小椋秀樹（2013）「大規模コーパスを活用した外来語表記のゆれの調査」『立命館文學』630
- 郝苗苗（2008）「中国語における外来語の受け入れの概要」『人間文化研究』6、長崎純心大学
- 川越いつえ（1995）「借用語にみる促音化とリズム衝突」『言語研究』108
- 沖森卓也編著（2010）『（日本語ライブラリー）日本語概説』朝倉書店
- 甲斐睦朗監修・松川利宏編（2005）『語彙指導の方法—語彙表編—』光村図書出版
- カッケンブッシュ寛子・大曾美恵子（1990）『日本語教育指導参考書 16 外来語の形成とその教育』大蔵省印刷局
- 河原崎幹夫（1991）「カタカナの指導—外来語の表記のしかた」『日本語教育』36
- 国際交流基金・日本国際教育協会編著（1994）『日本語能力試験出題基準』（改訂版 2004）、凡人社
- 国立国語研究所（1984）『国立国語研究所報告 78 日本語教育のための基本語彙調査』
- 国立国語研究所編（2004）『分類語彙表 増補改訂版』大日本図書
- 呉大綱（2005）「中国語の中の外来語—音訳か意識かそれとも音訳と意識のミックスか」『梅花女子大学文化表現学部紀要』2、梅花女子大学
- 小林ミナ・カッケンブッシュ寛子・深田淳（1991）「外来語にみられる日本語化規則の習得—英語話者の調査に基づいて—」『日本語教育』74、日本語教育学会

- 顧冰馨 (2012) 「外来語に見られる開音節化規則の習得－中国語母語話者への調査について」『言語教育研究』2、桜美林大学
- Giovanni Borriello・Daniele Petrella・袁華清 (2012) 『日語主題分類図解詞典』北京語言大学出版社
- 澤田田津子 (1985) 「外来語における母音添加について」『国語学』143、国語学会
- 章輝夫ほか (1990) 『中日分類辞典』外文出版社
- 陣内正敬 (2008) 「日本語学習者のカタカナ語意識とカタカナ語教育」『言語と文化』11、関西学院大学
- 陣内正敬・田中牧郎・相澤正夫 (2012) 『外来語研究の新展開』おうふう
- 鈴木義昭・王文編 (2002) 『日本語から引ける中国語の外来語辞典』東京堂
- 中川秀太 (2015) 「漢語・外来語の略語」『日本語学』34-2
- 樋口万喜子・古屋恵子・頼田敦子編 (2011) 『進学を目指す人のための教科につながる学習語彙 6000 語 (日中対訳)』ココ出版
- プレムモトワニ (1991) 「日本語教育のネック－外来語－」『日本語教育』74
- 水谷修監修 (2000) 『日英中韓カタカナ語見くらべ辞典』講談社
- 宮島達夫・高木翠 (1974) 「外来語の表記の変化とゆれ」『計量国語学』71
- 森山卓郎 (2012) 『日本語・国語の話題ネター－実は知りたかった日本語のあれこれ－』ひつじ書房
- 山内博之編 (2013) 『実践日本語教育スタンダード』ひつじ書房
- 山口仲美 (2006) 『日本語の歴史』岩波書店
- 李視岐 (1985) 『日本語外来語』山西人民出版社
- 李峰栄・板谷雄二 (2006) 「中国人向けカタカナ語 e ラーニング教材の開発と視覚的刺激の有効性の検証」『朝日大学経営学部電子計算機室年報』15、朝日大学
- 劉志偉 (2014) 『超級日本語学習者の求める語彙シラバスに関する一私案－学習者が自ら動いて習得する語彙を中心に－』第 45 回語彙・辞書研究会予稿集

参考インターネットサイト

内山和也 「実用カタカナ語 2010」

http://www7a.biglobe.ne.jp/~nifongo/data/kata_2010.html

国立国語研究所 「日本語の現在」

<http://www.ninjal.ac.jp/archives/genzai/>

茂木俊伸 「外来語の意味・用法に関する文献目録」

http://www.let.kumamoto-u.ac.jp/literature/asia/nihonbungaku/tmogi/lw_biblio/index.html