

教科書等，教材に使用される語彙の研究動向について

似 内 寛

要旨：本稿は教科書などの教材の理解度が低い場合の，児童・生徒・学生の社会的背景について研究するための，予備的考察を行う研究ノートである。主に〔教科書を分析対象として語彙の分析を行っている先行研究〕の研究結果から，教材の難易度の指標として，語彙の難易度や文章の長さのほか，文法的な要素，児童・生徒・学生が日常で使用する語彙であるか，教科書で説明している「概念」を理解するために必要な語彙など多様な要素が教科書の難易度に関係していることがわかった。また分析手法として (1)「つまずきことば」の抽出 (茂木 2013)，(2)一文あたりの平均単語数と単語の難易度を変数とした重回帰分析 (川村 2013)，(3)多項ナイーブベイズ分類など難易度付きコーパスを学習用データとして使用する難易度測定 (近藤ら 2008，佐藤 2011)，(4)語彙数をデータとした特徴語の抽出 (中條 2008，河内 2017)，(5)文法的な特徴や語彙の多義性を指標とするもの (川口 1996，宮部 2008，宮部 2015，西谷 2000)，(6)文体の難易度に注目したもの (クリスティーナ・フリヤメク・寒川 2012)，(7)教科書で説明されている「概念」とその周辺知識との関係性に注目したもの (平松ら 2015，松原 2008，内田ら 2009) など非常に多様な分析方法があることがわかった。

キーワード：教科書，語彙，難易度

研究の目的

本稿は教科書などの教材に対する理解度が低い場合の，児童・生徒・学生の社会的背景について研究するための，予備的考察を行う研究ノートである。本稿では語彙に注目した先行研究の，分析方法と分析結果について整理する。

教材に使用されている文章について，理解できるかできないかを分ける要素を探ることが目的である。それは言語学的な関心からではなく，理解ができるかできないかを分ける要素を身につけるか身に付けないかに関わる社会的背景を探ることが最終的な目的であり，そのための予備的な考察として，教材を理解することに関係した指標や，指標を抽出するための分析手法について，先行研究から知見を得ることが目的である。

対象としている研究には，日本語を母語としない日本語学習者の支援が目的の研究も含まれるが，学校での教科学習支援を目的としている，あるいは分析手法が参考になる，という理由から取り上げた。同様に英語学習についての研究も，分析手法が参考になるものを取り上げた。

教材が理解できるかできないかは，どのような語彙がどのように使用されているかという問題だけでなく，文法的に正しく文章を解釈できるか，文章で説明されている内容と関わる背景的な知識を有しているか，論理的な推論に関する能力がどの程度あるか，などいろいろな要素が考

えられる。本稿では教材の難易度に関わる語彙に関係した研究のみを取り上げる。

1. 教材の難易度の要素を抽出している研究

教材の難易度に影響している語彙を直接抽出している研究には茂木（2013）がある。茂木の研究では「つまずきことば」に注目している。茂木は「つまずきことば」について、「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）語または語彙を指す」（茂木 2013, p 343）と説明している。茂木は教科書の語彙を独自の視点で分類し、「子供の日常に近い語」/「子供の日常に遠い語」×「教科に固有の語」/「教科に固有でない語」に分けている。子供の日常に近い語であれば、普段の生活の中で習得されるが、遠い語は意識的な学習を行う必要がある。また「教科に固有の語」は科学的な概念として理解する必要がある、「教科に固有ではない語」は一般的な概念として理解する必要があるとしている。

教科書に使用されている言葉をこのように分類することにより、「つまずきことば」が、「各教科に典型的な『難しい用語』のみに限定して存在するものではない」という前提にたつて「つまずきことば」の分析を行っている。茂木の考えによれば、たとえば生活の中で馴染み深い語であっても教科の文脈では異なる使い方をされている場合があり、そのような場合に「つまずきことば」となる。茂木は「ある語の日常的・辞書的な意味を理解していたとしても、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、知っているはずなのに実は理解できていない」（茂木 2013, p 344）可能性があるとして説明している。

茂木は言及していないが、この視点は「子供の日常に近い語/遠い語」が、個人の生活体験の違いに影響される可能性があることについて考える手懸りとなる。

この研究の「つまずきことば」は小学校国語科教科書から収集されている。収集には「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」（茂木 2013, p 345）する方法を採用している。調査の結果、小学校国語科の教科書の中から1,795語の「つまずきことば」を抽出している。この研究の成果は小学校国語教科書に限定されるが、調査手法は他教科や他の学年レベルでの調査にも応用できると考える。

2. テキストの難易度を判定するシステムに関する研究

ここで取り上げる研究の目的は、文章の難易度を判定することである。そのために教科書の学年が難易度の基準データとして利用されている。

テキストの難易度判定を行うシステムに関する研究の目的はいくつかあり、近藤・松・佐藤（2008）や佐藤（2011）においては、情報伝達手段としてのテキストのわかりやすさを客観評価

することが目的とされており, また川村・北村(2013)や安井・井出・土居(2014)では日本語を母語としない日本語学習者の支援を目的としている。

文章の難易度判定には基準が必要である。近藤・松・佐藤(2008)や佐藤(2011)では教科書を難易度判定の資料として使用している。教科書は学年により難易度が明確である, という前提のもと, 多項ナイーブベイズ分類などの学習データとして教科書を利用している。近藤らの研究は分析に unigram モデル(例えば漢字一文字)を単位としたテキストデータを利用し, 学年を難易度のレベルの基準にしている。そのため教科書の難易度に影響している要因を抽出するような分析に応用することを考えた場合に, テキストに含まれるどの1文字がどの程度難易度に影響しているかに関しては確認することができるかもしれないが, なぜその文字が難しいのか, その理由に関しては「高い学年の教科書に出現する頻度が高い」という理由以外に解釈のしようがない。しかし多項ナイーブベイズ分類は, 学習テキストを工夫することで, 同じ学年の教科書の語彙について, それを正しく解釈できるかできないかを分ける要因を探す目的で応用できる可能性があると考えられる。

川村・北村(2013)の研究では語彙の難易度と構文の複雑さに着目している。語彙の難易度判定の学習には日本語学習用の教科書(初級・中級・上級の各レベルの教科書)および専門書と白書を使用し, 重回帰分析により難易度を判定する式を決定している。指標として一文あたりの平均単語数と, 日本語能力試験のN1からN5にあたる5段階+超級(N0)の6段階に, 日本語学習用教科書をレベル別に対応させ, 「出題基準〇級の単語数/総単語数」というデータを作り, 重回帰分析に投入している。

安井ら(2014)も日本語学習者向けの支援を目的としている。具体的には「日本語中級教科書の難易度が適切かを判定する」ということを目的としている。

これらの研究は難易度の基準を教科書以外のデータにするなどで, 応用することが考えられる。たとえば学習者が難しいと感じるテキストとそうでないテキストを難易度付きの学習用データとすることで, 社会的背景の違いによる文章理解困難の要因となっている要素を見つける手がかりとなる可能性がある。

3. 教材の難易度の多様な指標について

この章では, 語彙の難易度以外に注目した研究について, どのよう視点で分析を行っているのか概観する。

川口(1986)は「外国学生用日本語教科書中級」の改訂版作成の参考とするため, 既存の日本語教科書では何が教えられているのかについて実態を調査することを目的としている。この研究では日本語を母語としない日本語学習者用のテキストの調査により, 多様な使われ方をする「こそ」「さえ」「もちろん」「たしかに」の用法と, 「待遇表現」, 「〈問題提起の意見→その根拠の提

示→問題解決への提案」といった一定の論理パターンを示す文章型」などを修得するために、十分な内容となっているかについて分析している。

また西谷(2000)は外国人児童が日本の教育を受ける際の、社会科教科教育の内容理解を促進することを目的とし、小学校2, 4, 5年生の教科書の語彙や表現の出現頻度と、日本語能力試験及び児童用日本語教科書との相関について調査を行っている。その背景として西谷は外国人児童の教室観察より「小学校4年生の中国人児童は、算数については問題がなかったが、社会科にはかなり苦戦していた。社会科の教科書の内容は日常生活と密接に結びついているものも多く、それだけに、日本人児童にとっては当たり前のことでも、外国人児童には馴染みが薄いものである。しかし、逆に言えば、日常生活に密着したものであるからこそ、社会科において扱われる語彙・表現の習得は外国人児童にとって欠くことのできないものだと考えることができる」(西谷2000, p 80)と述べている。このような背景からこの研究の目的として「(1)教科教育に結びつきやすい日本語教育を考える際の基礎資料を得る。(2)文法項目が学年別でどのように現れているか、その分布・様相を知る。(3)この基礎資料から得られる文法項目と語彙項目を日本語カリキュラム・ガイドラインおよび日本語力評価方法試案上に反映させる。」(西谷2000, p 80)の3つが設定されている。調査の対象は平成11年の3～5年生までの社会科教科書14冊のなかに使用されている全ての文字である。分析方法はこれらの教科書に使用されている語彙の出現頻度を集計し、出現頻度の高い語彙の『日本語能力試験出題基準』の級レベルを判定することで、社会科教科書の、語彙レベルの特徴を抽出している。

中條・西垣・長谷川(2008)は英語教育をテーマとして、ゆとり教育時代の英語教科書に使用される英単語の語彙数の変化傾向について分析している。分析手法として注目した点は教科書語彙の語彙レベルの基準として、イギリス英語の成人コーパスBNCと、「Dale and O'Rourke(1981)のThe Living Word Vocabulary(米国の1～16年生の半数以上が当該単語の意味を理解できる学年が示されている)」と、TOEIC, TOEFLなどの言語資料を用いていることである。これらの基準資料の語彙数と比較することで、教科書に使用されている語彙のレベルや、ネイティブスピーカーが使用する語彙のカバー率を算出し明らかにしている。

宮部(2008)は小学校社会科教科書を対象に教科書に出現する「他動詞とそれと組み合わせる『を格』の名詞」とに注目し、連語論的な観点(連語のなかで他動詞がどのような語彙的な意味で用いられており、『を格』の名詞とどのようなむすびつき方をしているか)(宮部2008, p 69)から分析している。その背景には日本語を母語としない子どもにとって、社会科教科書においては、多義語である他動詞が派生的な意味で使用されることが多く難しい、という認識がある。この研究における分析手法の特徴的な点は、(1)社会科教科書と理科教科書で語彙の比較を行っている点と、(2)難易度の指標として「を格+他動詞」が「対象へのはたらきかけ」「対象の所有ややりとり」「対象への心理的な関わり」の3つ語彙的な意味となることに注目している点である。(1)は社会科教科書の語彙的な特徴を、理科教科書を基準とすることで示していて、(2)の難易

度の指標の出現頻度を社会科と理科で比較し, 社会科教科書の難易度に関する語彙的な特徴を抽出している。

内田ら (2009) は「現代日本語書き言葉均衡コーパス」における語彙頻度に基づき「ジャンル別の語彙使用の特徴を明らかにすることを目的」(内田ら 2009, p 442) としている。この研究でいう「ジャンル」とは日本十進分類法 (NDC) による総記, 哲学, 歴史, 社会科学, 自然科学, 技術・工学, 産業, 芸術・美術, 言語, 文学である。これらのジャンルのコーパスに加えて国語, 数学, 理科, 社会, 外国語, 技術家庭, 芸術, 保健体育の 8 科目の教科書コーパスを分析対象とし, 「現代日本語書き言葉均衡コーパス」のサブコーパスデータを用いて, 語彙の共起関係から算出した修正ファイ係数を用いたクラスター分析から, 各ジャンルにおける特徴語を抽出している。この研究の分析手法で特徴的な点は (1) 教科書コーパスと, それ以外のジャンル別コーパスを分析対象としている点と, (2) クラスター分析を用いてジャンル間の近さを明らかにしている点, (3) 「ジャンルの上位各 100 語のリストからそれぞれのジャンルに独自に出現する語 (以下, 特徴語とする) を抽出し, それらをフレームの観点から分析」(内田ら 2009, p 444) している点である。(1) と (2) から教科ごとにどのような分野の語彙と類似性が高いのかについての知見が得られ, 教科を理解するための背景知識をどのような分野の書籍から得る必要があるのかを考察する資料となっている。また (3) の「フレーム」とは言葉を理解するための背景知識であり「FrameNet という現在構築中であるオンライン上のフレーム辞書に記述されている」(内田ら 2009, p 444) と説明されている。「特徴語」と「フレーム」という概念が示されていることで, 教科がどのような背景知識を必要としているかについて分析することができる手法である。

クリスティーナ・フメリヤク・寒川 (2012) は, 日本語教科書のリーダビリティ (読みやすさ) の要素を, 「日本語を母語としない日本語学習者向けの中級読解教科書に掲載された文章を読みやすくするための, 書き換えに使用されたストラテジー」を分析することにより, テキストの中の書き換えられた要素から, 読みやすさと関係する要素を明らかにしている。この研究の特徴的な点としては語彙の難易度や文法的な特徴などとは異なる指標でテキストの難易度に関係する要素を抽出している点である。分析は, 原文と書き換え分の相違をコンピュータにより抽出し, 人手により「書き換えの種類, レベル (語彙, 構文, 談話など) と書き換え内容の記述を付与」(クリスティーナ・フメリヤク・寒川 2012, p 286) するという方法がとられている。この分析により「単純化」「明示化」「標準化」という 3 つのストラテジーによりリーダビリティが向上しているという結果が得られている。

立山 (2014) は日本語を母語としない児童生徒の, 日本の学校における教科学習の支援をテーマとし, 「同じ語であっても教科によって使われ方が異なっていたり, 使用法が特殊であったりする語についての分析を行い, 教科による特徴語についての研究を精緻化させることを目的と」(立山 2014, p 317) している。教科によって使用されている意味に特徴があるといえる「伝わる」「伝える」「結ぶ」を取り上げ, 「現代日本語書き言葉均衡コーパス」の「教科書サブコーパス」

から全学年全校種を対象に検索し、『明鏡国語辞典』に従って目視で意味の分類を行っている。分析対象とした言葉は多義的であり、例えば「伝わる」について立山は、国語辞典掲載の7つの意味を示した後に、それを「1-1-1 電気・熱・音や刺激などがある道筋を通る。またそのようにしてある所に達する。1-1-2 液体が沿って流れる。1-2 情報などがある道筋を通る。またそのようにしてある所に達する。」(立山2014, p 318)の3つに分類し、集計している。この研究で注目する点は(1) テキストの難易度の指標として語彙の多義性を採用している点と、(2) 教科によって同一の言葉の意味を異なって解釈しなければならないことを示している点である。特に(2)については、多義的な言葉について学習者が日常的に使用している意味解釈と異なる解釈をしなければならない場合に、それが原因で科目の学習に支障を来す可能性があるということは大きな問題提起である。

宮部(2015)は「発達に合わせた認知的・抽象的な理解能力は、学校での教科学習を通じて獲得しなければならない、そのためには教科学習の場で話される日本語や教科書に書かれた日本語が理解できなければならない」(宮部2015, p 19)と述べ、小学校理科と社会科の教科書の分析を行っている。具体的には教科書に出現する「条件形」の解釈の多様性と、その中で特に教科書に特徴的に使用される用法の抽出を試みている。条件形は① 一般的・習慣的なことから、② 仮定、③ 予定的なことから、④ すでにあることから、⑤ そのほか、の5つの意味に分類される。そこで、理科と社会科における各条件型の意味と用法の現れかたの集計と、理科教科書・社会科教科書と成人向けの書籍(自然科学分野)のテキストにおける意味・用法の現れ方の割合と比較を行い、教科書における条件型使用の特徴を分析している。この研究において教科書の難易度を分析する上で特徴的な点は(1) 小学校の理科と社会科を取り上げている点と、(2) 成人向けの書籍と教科書を比較している点と、(3) 「条件形」を難易度の指標としている点である。特に宮部は(2)のように成人向け書籍と比較して教科書に特徴的な語彙の使用法を明らかにする必要性に関して、「日常会話はそれほど問題ないのに、教科学習についていけない」年少者が存在すること、「やぶる」という他動詞が「紙をやぶる」「武田信玄をやぶる」「ルールをやぶる」のように複数の意味を持っていることを例に「もし、教師が、子どもたちが「紙(もの名詞)をやぶる」を理解できるため、これ以外の『やぶる』の意味を教えずに進んでしまったとしたら、より意味が抽象化した『〈ひと名詞〉をやぶる』や『〈こと名詞〉をやぶる』における『やぶる』の意味を、子どもたちが理解できていない可能性があるのではないか。このように、教科書には(教科学習では)、見落とされやすい語彙や文法が存在するといえる」(宮部2015, p 21)と説明している。日常的に慣れ親しんでいる言葉の意味と異なる、教科書に特徴的な言葉の解釈が多く存在するとすれば、教科書の難易度に影響する要因と考えられる。

河内(2017)は「学力差の背景に語彙がある」という問題意識を「平成28年8月1日に公表された『次期学習指導要領に向けたこれまでの審議のまとめ(素案)』(教育課程部会教育課程企画特別部会資料3-1)」と共有し、中学校全教科の語彙の分析を行っている。この研究の意義に

ついで河内は「各教科書の語彙の実態の提示, 国語科で指導すべき他教科の語彙の把握」(河内 2017, p 52)と説明している。具体的には2005年の学習指導要領下と, 2013年の学習指導要領下の, 中学校の全教科書のコーパスから語彙表を作成し, 2013年度版に特徴的に出現している語を対数尤度比較検定によって抽出している。そして抽出された語に「現代日本語書き言葉均衡コーパス」の各語に付与されているa~eのレベルを付与し特徴語を精選している。この研究のユニークな点は新旧の, 学習指導要領下の教科書語彙を比較するという視点である。河内は言及していないが, このような分析により, 学習指導要領の改定の影響による世代ごとの語彙力に差について考察することが可能になる。

4. 教材に出現する「概念」に注目している研究

この章では教科で学習する「概念」を理解するために, 「概念」の説明に使用される周辺的な言葉を抽出する研究について概観する。そのなかでも学習する「概念」同士の関係性や, それを理解するために必要とされる背景知識や, それら同士の近さなどに関する研究を取り上げる。

教科書等で説明される「概念」を語彙から抽出する研究として, 松原(2008), 平松ら(2015)などがある。

松原(2008)は, 「教科書の記述内容について, 自己組織化マップを用いて分析する方法を明らかにすることを目的」(松原 2008, p 1)としている。この研究の背景には「国際調査で, 理科における我が国の子供の記述内容に科学的根拠や理由などを説明する力が十分でないことが指摘されている」(松原 2008, p 1)という現状への危惧が存在する。松原は科学的思考や科学的知識は言語化・記号化された抽象度の高い知識であるとし, 問題解決などを通して「具体的な自然事象をイメージ化し, 抽象化していく」ものであり, 「教師や子どもどうしの対話, 教科書などの既に言語化されたもの」が影響して形成されると考えている。そのような背景から教科書に記載されている内容の量と質の両面からの分析が必要であるとし, 教科書のテキストデータを, 自己組織化マップを用いて分析する方法を提案している。この論文で使用されているデータは中学校理科第2分野の, 2つの出版社の教科書の, 「火山」に関する単元に出現する語彙の, 1文中の出現回数データである。この分析により特定のキーワードがどのような単語と関連づけられながら解説されているかが, 自己組織化マップの中で視覚的に明らかにされている。このような分析の視点は, 新しい知識の取得の前提として必要となる周辺知識にどのようなものがあるのか, を明らかにできる点で, テキストの難易を度考える上で重要であると考えられる。

平松ら(2015)は, 「地球規模の環境問題と日常生活をつなぐ方策を用いた教育が, 家庭科の教科書にどのように取り入れられているか」(平松ら 2015, p 2)を明らかにすることを主な目的として家庭科教科書を分析している。テキストマイニングを用いて「環境」という語と関連が深い語句を調査し, それを通して「持続可能社会」や「CLA (Life Cycle Assessment)」といった

言葉が出版社の違いと、小中高校での扱いの違いを分析している。分析には KH Coder による共起ネットワークの作成 (Jaccard 係数を使用) と、多次元尺構成法による多変量分析、対応分析、クラスター分析が使用されている。この分析により「環境」という語がどのような言葉と結びついているのかが明らかになっている。前述の松原の研究と同様に、教科書の難易度を考える上で、キーワードの背景的な知識を特定する手法として興味深い。

5. ま と め

以上の、主に語彙をデータとして分析する研究から、いくつかのサジェスションを得ることができた。難易度についての指標には (1) 語彙そのものの難易度、(2) 文章の長さ、(3) 文法的な要素、(4) 児童・生徒・学生が日常で使用する語彙か、(5) 説明されている「概念」を理解するために必要な語彙の抽出、などが有効であることがわかった。また分析の手法としても (1) 「つまずきことば」の抽出 (茂木 2013)、(2) 一文あたりの平均単語数と単語の難易度を変数とした重回帰分析 (川村 2013)、(3) 多項ナイーブベイズ分類など難易度付きコーパスを学習用データとして使用する難易度測定 (近藤ら 2008、佐藤 2011)、(4) 語彙数をデータとした特徴語の抽出 (中條 2008、河内 2017)、(5) 文法的な特徴や語彙の多義性を指標とするもの (川口 1996、宮部 2008、宮部 2015、西谷 2000)、(6) 文体の難易度に注目したもの (クリスティーナ・フリヤメク・寒川 2012)、(7) 教科書で説明されている「概念」とその周辺知識との関係性に注目したもの (平松ら 2015、松原 2008、内田ら 2009) など非常に多様な分析方法があることがわかった。

これらの分析は日本語学習者を想定したもの、英語学習に特化したもの、小学校国語、理科や社会科に分析対象を限定した研究や、特定の語彙の文法的な特徴にのみ注目した研究など、それぞれの研究の目的と性格によって対象や指標が限定的である。理解度の低さの社会的背景を分析するという目的に対して、現実的に利用可能な分析方法を精査することは今後の課題である。

参 考 文 献

- 1) クリスティーナ・フリヤメク・寒川, 2012, 「日本語学習者にとって読みやすい文章について」国立国語研究所コーパス開発センター『第2回コーパス日本語学ワークショップ予稿集』, pp 285-290.
- 2) 中條 清美, 西垣 知佳子, 長谷川 修治, 内山 将夫, 2008, 「「ゆとり教育」時代の高校教科書語彙を考える—1980年代と2000年代の高校英語教科書語彙の比較分析からの考察」英語コーパス学会『英語コーパス研究』15号, pp 57-79.
- 3) 平松 あい, 山本 大輔, 栗栖 聖, 花木 啓祐, 2015, 「家庭科へのLCA的思考法導入に向けた教科書のテキスト分析」日本LCA学会『日本LCA学会誌』11巻1号, pp 2-10.
- 4) 河内 昭浩, 2017, 「中学校教科書語彙の研究」群馬大学教育学部『群馬大学教育学部紀要, 人文・

社会科学編』66巻, pp 51-64.

- 5) 川口 義一, 1986, 「中級教科書の語彙・文型」早稲田大学語学教育研究所『講座日本語教育』第22分冊, pp 14-27.
- 6) 川村 よし子, 北村 達也, 2013, 「日本語学習者のための文章の難易度判定システムの構築と運用実験」CAJLE カナダ日本語教育振興会『Journal CAJLE』vol. 14, pp 18-30.
- 7) 近藤 陽介, 松吉 俊, 佐藤 理史, 2008, 「教科書コーパスを用いた日本語テキストの難易度推定」言語処理学会『言語処理学会第14回年次大会発表論文集』, pp 1113-1116.
- 8) 松原 道男, 2008, 「自己組織化マップを用いた理科教科書内容の分析」金沢大学教育学部『金沢大学教育学部紀要 教育科学編』57巻, pp 11-16.
- 9) 宮部 真由美, 2008, 「小学校社会科教科書の他動詞の使用について・連語論の観点から: 子どもに対する教科学習の日本語支援のために」文教大学『文学部紀要』22巻1号, pp 69-90.
- 10) 宮部 真由美, 2015, 「教科学習の日本語を知る: 小学校理科教科書と社会科教科書の条件形を例として」文教大学『文学部紀要』28巻2号, pp 19-38.
- 11) 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『鳴門教育大学研究紀要 鳴門教育大学 編』28巻, pp 343-355.
- 12) 西谷 まり, 2000, 「小学校の社会科教科書の語彙と表現: 日本語能力試験レベル及び外国人児童教育のための日本語教科書との対照」一橋大学留学生センター『一橋大学留学生センター紀要』3号, pp 79-93.
- 13) 佐藤 理史, 2011, 「均衡コーパスを規範とするテキスト難易度測定」情報処理学会『情報処理学会論文誌ジャーナル』52巻4号, pp 1777-1789.
- 14) 立山 智絵, 2014, 「教科書コーパスにおける教科別特徴語」国立国語研究所コーパス開発センター『第5回コーパス日本語学ワークショップ予稿集』, pp 317-322.
- 15) 内田 諭, 藤井 聖子, 2009, 「『日本語コーパス』における語彙のジャンル別特徴—クラスター分析とフレームの観点から—」言語処理学会『言語処理学会第15回年次大会 発表論文集』, pp 442-445.
- 16) 安井 朱美, 井手 友里子, 土居 美有紀, 2015, 「中級日本語読解教材の書き換えにおける一考察: リーダビリティスコアと日本語教師の評価を基に」南山大学国際教育センター『南山大学国際教育センター紀要』15号, pp 23-37.