

機関番号：62618

研究種目：特定領域研究

研究期間：2006～2010

課題番号：18061008

研究課題名（和文） 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

研究課題名（英文） Application of Large-scale Balanced Corpora to Language Planning and Education: Through the Compilation of Word Lists and Kanji Character Lists of Contemporary Japanese

研究代表者 田中 牧郎 (TANAKA Makiro)

大学共同利用機関法人人間文化研究機構国立国語研究所・言語資源研究系・准教授

研究者番号：90217076

研究成果の概要（和文）：(200字程度)

「現代日本語書き言葉均衡コーパス」をもとに、語彙表と漢字表を作成し、社会でよく使われている語彙と漢字の範囲を明確にした。小学校・中学校・高等学校で用いられている全教科の教科書を対象としたコーパスを構築し、その語彙表と漢字表も作成し、社会で使われる語彙・漢字と、学校で教えられている語彙・漢字を対照できるようにした。これらのコーパス、語彙表、漢字表を用いて、国語政策と国語教育の分野における諸課題に取り組む方法を研究した。

研究成果の概要（英文）：

We clarified the scope of vocabulary and kanji character commonly used in society through compilation of word lists and kanji character lists based on "Balanced Corpus of Contemporary Written Japanese". We also built a corpus of elementary and middle school and high school textbooks, and made comprehensive lists of vocabulary and kanji character of the general public and school education. Then we studied the method to address challenges in the field of language policy and language education.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	6,200,000	0	6,200,000
2007年度	11,300,000	0	11,300,000
2008年度	20,384,000	0	20,384,000
2009年度	7,600,000	0	7,600,000
2010年度	7,600,000	0	7,600,000
総計	53,084,000	0	53,084,000

研究分野：日本語学

科研費の分科・細目：言語学 日本語学

キーワード：語彙 漢字 言語生活 国語教育 語彙教育 漢字教育 国語政策 コーパス

1. 研究開始当初の背景

国語政策と国語教育の分野で研究の論点になる重要課題の中には、現在の日本社会における言語の実態を正しく把握することが前提となるものが多いが、その把握は必ずしも十分に行われていなかった。その部分の調査研究に、現代日本語の縮図として構築される「現代日本語書き言葉均衡コーパス」

(BCCWJ) などを用いることで、国語政策と国語教育の研究を大きく進展させることができると見込まれた。

2. 研究の目的

国語政策と国語教育の分野の諸課題で共通して問題になることが多い語彙と漢字を取り上げて、次の二つのことに目標を設定し

て研究を進めた。

- (1) 小学校・中学校・高等学校で使われている教科書のコーパスを作成し、本特定領域全体で構築する BCCWJ と合わせて、現代日本語の公共的場面で使われる語彙・漢字を網羅した語彙表・漢字表を作成する。
- (2) 作成した語彙表・漢字表を通してコーパスを使うことで、国語政策・国語教育の諸課題にどのように取り組んでいくことができるかを示す。

3. 研究の方法

「2. 研究の目的」に記した二つについて、それぞれ次のような方法で取り組んだ。

- (1) 教科書コーパスと語彙表・漢字表諸種を、国語政策・国語教育の諸課題に役立てることができる基礎資料となるように設計し、次の六つを作成した。

①教科書コーパス

小学校・中学校・高等学校の全学年・全教科の教科書一種ずつの全文をコーパス化し、学校で教えられている語彙や漢字の全貌が把握できるようにした。

②教科書コーパス語彙表

上記の「教科書コーパス」に出現する語彙を、様々な観点から一覧できるように語彙表を作成した。校種別・教科別の頻度のほか、BCCWJ と比較した場合の教科別の特徴度なども掲載した。

③BCCWJ 主要コーパス語彙表

BCCWJ のうち、国語政策・国語教育にも特に有用だと思われる、「図書館書籍」「出版書籍」「新聞」「雑誌」「Yahoo!知恵袋」「ブログ」の6種のサブコーパスについての頻度のほか、頻度に基づく語彙レベルを設定して、サブコーパス間で比較対照できるように一覧にした。

④学校・社会対照語彙表

上記の「教科書コーパス語彙表」と「BCCWJ 主要コーパス語彙表」に収録した語彙情報のうち、中等教育において特に活用されることが多いと思われるものを抜き出し、学校の語彙と社会の語彙とが比較対照できるような一覧表を作成した。

⑤教科特徴語リスト

「教科書コーパス」をもとに、教科に特徴的な語彙のリストを作成した。②の「教科書コーパス語彙表」は短い単位による語彙表であったが、本リストは長い単位に基づき、人手による修正も加え、教科専門用語の扱いについて検討できるリストを作成

した。

⑥NDC ジャンル別漢字出現頻度表

BCCWJ の「出版書籍」データを用いて、日本十進分類 (NDC) のジャンル別の漢字頻度表を作成し、ジャンルの特徴漢字の抽出を行った。

- (2) 上記の語彙表・漢字表をどのように活用すれば、国語政策や国語教育の諸課題に役立つのかについて、具体的な課題に即して研究した。

国語政策の課題に対しては、医療分野を例に、コーパスにおける語彙の出現状況を分析して難解な専門用語を抽出して序列化する研究と、漢字のジャンル別頻度やどのような語の表記に用いられるのかを分析して、漢字政策に新視点を取り入れる研究を行った。

国語教育の課題に対しては、学校教育に期待される語彙力や漢字力を養成するための諸課題を幅広く取り上げ、コーパスや語彙表・漢字表をどのように使っていけばよいかを研究した。語彙教育・漢字教育のあり方の検討に活用するアプローチと、教師や生徒の授業活動で直接活用するアプローチの両面から取り組んだ。

4. 研究成果

- (1) 「3. 研究の方法」の(1)に記した六つの基礎資料は、それぞれ次のような意義を持つ。

①教科書コーパス

全学年・全教科を対象とした教科書コーパスは従来になかったものである。また、その一部を著作権処理を施して BCCWJ の一部に組み入れて公開したことにより、今後多くの研究者や教師達によって、研究の基礎資料として広く活用されると期待できる。

②教科書コーパス語彙表

全学年・全教科の教科書で使用されている語彙が一覧できるようになり、BCCWJ から明らかになる社会で使われる語彙と比較した場合の教科ごとの特徴も明示されたことにより、今後の学校教育や教材研究などの基礎資料として高い価値を有すると考えられる。

③BCCWJ 主要コーパス語彙表

BCCWJ の特徴は、多種多様なサブコーパスから構成され、その相互比較が行えるところにある。語彙に関してこれを行いやすくした本語彙表は、語彙の難易度や重要度を推定する基礎資料として、政策や教育の様々な

局面で活用されると見込まれる。

④学校・社会対照語彙表

中等教育は、社会で役立つ教育内容をもたせることが求められているが、本語彙表から、社会での必要度に比べて学校教育での扱いが不足している語彙や概念などを抽出することが可能であり、教育課程や教科書編集のあり方を考える重要な知見を導き出すことが期待される。

⑤教科特徴語リスト

各教科の教科専門用語を抽出したものであるが、従来あった経験的にリスト化された教科用語集と比較して網羅的であることや、また、全教科が同じ基準で抽出されていることなどの利点がある。教科教育のあり方について、全体を見わたしながら検討するのに格好の資料であると言える。

⑥NDC ジャンル別漢字出現頻度表

2010年に内閣告示となった、新しい「常用漢字表」では、その検討の過程でジャンル別の漢字頻度についての情報を必要とされながら、そのデータが必ずしも十分に用意されなかった。本漢字表は、様々なジャンルをバランスよく含むBCCWJの特徴を生かしたそれであり、今後の漢字政策の基礎データとなると考えられる。

(2)国語政策に対しては、本研究の成果は次のように位置づけられる。

①医療用語を例に、コーパスの語彙頻度を用いて、難解な専門用語の抽出や分類を実施し、医療専門家・非専門家に対するアンケート調査の結果と対照することで、この方法の評価を実施したところ、コーパスの有効性が確かめられた部分と、他の方法と連携する必要性が求められた部分があった。実際に、専門家と非専門家のコミュニケーションを円滑化するための政策課題に役立てて行くには、さらに何段階かの研究が求められるが、本研究は、そこに進むための足がかりとして有効だと考えられる。

②各漢字がどのようなジャンルに用いられるか、どのような語の表記に用いられるかのリストを作成したことにより、今後の漢字政策の基礎資料として役立てられることが見込まれる。漢字政策として近年改訂が行われた「人名用漢字」「常用漢字表」の検討の過程では、こうしたリストが必要とされたが、必ずしも十分なものがなかった。今後は、本研究が示したような方法で、コーパスに基づく漢字リストを作成して漢字政策課題に役立てていくことが

できると思われる。

(3)国語教育に対しては、本研究の成果は次のように位置づけられる。

①本領域全体で構築したBCCWJと、本研究で作成した「教科書コーパス」をもとに作成したさまざまな語彙表は、社会で使われている語彙の中で学校で教えられている語彙がどのような位置にあるのかを、多角的に知ることができるものになっている。こうした語彙表は、現在の学校教育が、社会で必要とされていることを教えているのかについて検証する手段として用いることができる。

また、社会的な重要語彙がどの教科でどのように教えられているかを対照することができるようになったことで、教科間で連携をはかりながら語彙や概念を教える方策を検討することができるようになった。

②上記のような語彙表によって、学校教育の場を中心に社会一般を視野に入れた語彙情報のデータベースが整備されたことになるが、このデータベースを起点に、語彙教育における新しい研究が生まれていくことが期待できる。例えば、生徒の作文の語彙を調査して、上記の語彙情報と対照すれば、生徒の発達段階と社会的な必要度に応じた語彙教育の方法を実証的に研究していくことができるようになる。

③教育漢字や常用漢字について、それがどの教科やどのジャンルによく用いられるか、またどの範囲の語に用いられるかが総覧できるようになったことは、学校教育における漢字指導のあり方を再検討する資料が整ったと見ることができる。例えば、漢字は従来は国語科だけで教えてきたが、社会科や理科など教科に特徴的な語彙に用いられる漢字の場合は、その教科での概念教育のなかで扱っていくことなどが提言できるものと考えられる。

④語彙を介して多様な文章を引くことができるコーパスは、作文教育のツールとしての発展性も有していると考えられる。本研究で作成した語彙リストや作文課題になりやすいテーマ語彙集に、コーパスから引き出して多様な文章例を添えた教材を整備することで、テーマ型作文を書く際の発想源をつくることなどが、有効だと考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線) ※協力者に点下線を引いた。

[雑誌論文] (計 19 件)

- ① 棚橋尚子、教科に特徴的な漢字に関する考察—他教科における漢字指導の可能性—、国語科教育、査読有、67号、2010、pp. 11-18

- ② 河内昭浩、作文指導におけるコーパスの活用—高等学校での小論文指導を通して—、『解釈』、査読有、56巻5・6号、2010、pp. 27-36
- ③ 田中牧郎、相澤正夫、「難解用語の言語問題への具体的対応—「外来語」と「病院の言葉」を分かりやすくする提案—」『社会言語科学』、査読有、13巻1号、pp. 95-108
- ④ 田中牧郎、言語政策に役立つコーパスを用いた語彙表・漢字表などの作成と活用、人工知能学会、査読有、24巻5号、2009、pp. 665-672
- ⑤ 鈴木一史、漢字の読み書き問題の通過率に影響を及ぼす要因について—中学三年生の調査をもとに—、解釈、査読有、638・639号、2008、pp. 9-17
- ⑥ 近藤明日子、田中牧郎、学校教科書の語彙—語種を観点として—、日本語学、27巻10号、2008、pp. 26-35
- ⑦ 小椋秀樹、相澤正夫、現代雑誌70誌における漢字の使用実態と常用漢字表—国語施策へのコーパス活用に向けた基礎調査—、日本語科学、査読有、22号、2007、pp. 125-146
- ⑧ 鈴木一史、教科書コーパスによる国語科学習語彙の選定方法の検討、東京大学教育学部附属論集、51、2007

〔学会発表〕(計26件)

- ① 鈴木一史、使用語彙の発達過程と日本語コーパスの可能性、日本国語教育学会大学部会、2010年12月、筑波大学附属中学校
- ② 田中牧郎、近藤明日子、河内昭浩、鈴木一史、棚橋尚子、「学校・社会対照語彙表」の作成と活用、日本語学会2010年度秋季大会、2010年10月、愛知大学
- ③ 河内昭浩、語彙を豊かにする作文指導—「コーパス」を活用した実践と今後の教材開発に向けて—、全国大学国語教育学会第119回鳴門大会、2010年10月、鳴門教育大学
- ④ 佐野大樹、田中牧郎、丸山岳彦、「病院の言葉」の類型の推測とモデル化—「現代日本語書き言葉均衡コーパス」における語の使用度数を用いた—考察、日本言語学会第140回大会、2010年6月、筑波大学
- ⑤ 田中牧郎、一般語彙リストの作成について、イタリア日本語教育協会第4回日本語・日本語教育シンポジウム、2010年3月、ローマ大学
- ⑥ 田中牧郎、コーパスの語彙頻度を用いた教育語彙の検討、韓国日本語学会第20回国際学術発表会、2009年9月、韓国・建国大学

- ⑦ 斎藤達哉、公共性の高い文章における日本語表記の在り方—仮名で書くか漢字で書くか、日本言語政策学会、2009年6月、昭和女子大学
- ⑧ 相澤正夫、小椋秀樹、白書コーパスに基づく常用漢字の使用実態調査、社会言語科学会第21回大会、2008年3月、東京女子大学
- ⑨ 田中牧郎、金愛蘭、桐生りか、近藤明日子、コーパスによる難解語・重要語の抽出—医療用語を例に一、社会言語科学会第21回大会、2008年3月、東京女子大学

〔図書〕(計8件)

- ① 田中牧郎、相澤正夫、斎藤達哉、棚橋尚子、近藤明日子、河内昭浩、鈴木一史、平山允子、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」言語政策班報告書)、2011、254
- ② 田中牧郎、相澤正夫、斎藤達哉、棚橋尚子、鈴木一史、近藤明日子、平山允子、「教科書コーパス」中学校教科書語彙表(短単位)(1)、2009、540
- ③ 田中牧郎、相澤正夫、斎藤達哉、棚橋尚子、小椋秀樹、鈴木一史、近藤明日子、平山允子、金愛蘭、桐生りか、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」言語政策班報中間報告)、2008、291

〔産業財産権〕

○出願状況(計0件)

○取得状況(計0件)

〔その他：審議会への提出資料〕(計3件)

- ① 国立国語研究所、「俺」等の表記について(頻度と内訳)、文科審議会国語分科会幹事小委員会資料、2008年7月
- ② 国立国語研究所、「現代日本語書き言葉均衡コーパス」に基づく漢字音訓一覧表、文科審議会国語分科会漢字小委員会漢字ワーキンググループ用資料、2008年6月

6. 研究組織

(1) 研究代表者

田中 牧郎 (TANAKA Makiro)

国立国語研究所・言語資源研究系・准教授
研究者番号：90217076

(2) 研究分担者

相澤 正夫 (AIZAWA Masao)

国立国語研究所・時空間変異研究系・教授
研究者番号：80167767

斎藤 達哉 (SAITO Tatsuya)

専修大学・文学部・准教授

研究者番号：90321546

棚橋 尚子 (TANAHASHI Hisako)

奈良教育大学・教育学部・教授

研究者番号：20272271

(3) 連携研究者

近藤 明日子 (KONDO Asuko)

国立国語研究所・コーパス開発センター・

プロジェクト特別研究員

研究者番号：30425722

(4) 研究協力者

河内 昭浩 (KAWAUCHI Akihiro)

群馬県立館林高等学校教諭

鈴木 一史 (SUZUKI Kazufumi)

東京大学教育学部附属中等教育学校教諭