

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 9 日現在

機関番号：12102

研究種目：基盤研究(A)

研究期間：2011～2014

課題番号：23242026

研究課題名(和文)汎用的日本語学習辞書開発データベース構築とその基盤形成のための研究

研究課題名(英文) Research for the formulation of basic grounds for the construction of a general database for the development of Japanese language learners' dictionaries

研究代表者

砂川 有里子 (SUNAKAWA, Yuriko)

筑波大学・人文社会系・教授

研究者番号：40179289

交付決定額(研究期間全体)：(直接経費) 31,200,000円

研究成果の概要(和文)：日本語学習辞書編集を支援する目的で「汎用的日本語学習辞書開発データベース」を作成した。このデータベースは、「現代日本語書き言葉均衡コーパス」を利用した頻度や分布の語彙調査により選定された1万7千920項目の語彙からなる。各語彙には、経験豊富な日本語教師の主観判定に基づく6段階の難易度の他、旧日本語能力試験の等級、品詞、語種などの情報や、語義と用例(作例とコーパスからの実例)が組み込まれている。このデータベースの他、コロケーションリスト、類義語リスト、文化的発見記述文リストを作成し、これら全てを統合した情報の検索システムを(<http://jreadability.net/jev>)に公開した。

研究成果の概要(英文)：In order to support the compilation of Japanese language learners' dictionaries, we have completed a database called 'a general database for the development of Japanese language learners' dictionaries'. This database consists of 17,920 words selected via researching word frequency and dispersion using the Balanced Corpus of Contemporary Written Japanese (BCCWJ). Each word has readability scale of 6 levels determined by Japanese veteran teachers, the level of the old Japanese Language Proficiency Test, part of speech, word type, and sample sentences (both composed sentences and authentic sentences collected from BCCWJ.)

Apart from this database, we have completed a collocation list, thesaurus, and a list of short descriptions of cultural discoveries. We have developed a search system for both the database and those lists, and presented the system via internet (<http://jreadability.net/jev>.)

研究分野：日本語学・日本語教育

 キーワード：日本語教育 日本語学習辞書 二言語辞書 コーパス データベース 日本語学習辞書編集支援 Web版
日本語学習辞書

1. 研究開始当初の背景

海外での日本語学習者数は年々増加しているが、世界各地には満足な日本語学習用辞書がない地域が数多く存在する。その最大の理由は、辞書開発にかかる莫大な労力と費用と時間である。これらの負担は辞書の記述内容を搭載したデータベースを提供することにより大いに軽減され、日本語学習辞書を必要とする地域における現地語と日本語の二言語辞書の開発の促進につながると考えられる。

本プロジェクトの研究分担者と連携研究者は辞書の編纂・語法研究やコーパスを活用した言語分析の豊富な経験を有している。また、最近では大規模な日本語コーパスの構築が進み、コーパスを活用した辞書開発の環境が整いつつある。このような状況で、日本語コーパスを活用することにより、中・上級レベルの学習者向け日本語学習辞書に必要な語彙項目の選定を行い、日本語学習辞書のコンテンツに必要な情報を搭載したデータベースを構築し、公開することとした。

2. 研究の目的

これまでの日本語教育において、一般の教師や日本語教育研究者が利用可能な語彙教育用のリソースは、旧日本語能力試験の出題基準が唯一であった。しかし、この出題基準は日本語能力試験を作るために構築されたリストであること、および、1980年代に策定された基本項目に基づくものであることから、近年のコミュニケーション能力やタスク中心の日本語教育にとっての必要要件をみたしていない。この点を踏まえ、本プロジェクトでのデータベース構築に際しては以下の三つの課題を立てた。

- (1) オーセンティックな語彙項目を取り入れた教育のための語彙表を作成すること。
- (2) 語彙項目に対して様々な指標を付け、辞書開発はもちろんのこと、教育現場でのニーズに答える語彙表を作成すること。
- (3) ウェブを通じて国内外のユーザーと共有できる語彙表を作成すること。

3. 研究の方法

研究の目的にあげた三つの課題のそれぞれについて本プロジェクトが取った方法は以下の通りである。

(1)の課題を解決するために、国立国語研究所が2011年に公開した「現代日本語書き言葉均衡コーパス(BCCWJ)」を積極的に利用し、語彙の頻度や分布の調査を行った上で収録する項目を選定した。

(2)の課題を解決するために、経験豊富な日本語教師5名に依頼して、主観判定により語彙の難易度情報を付ける作業を行った。難易度は初級前半・初級後半・中級前半・中級後半・上級前半・上級後半の6段階である。この難易度の他、旧能力試験出題基準の等級、品詞、語種などの情報を付した。

(3)の課題に関しては、データの運用から配布のすべてのプロセスでCSVファイルによる管理を行った。

4. 研究成果

1万7千920項目の実質語を搭載した「汎用的日本語学習辞書開発データベース(通称「日本語教育語彙表」)」を作成した。この語彙表は、川村よし子氏が開発した『リーディングチュウ太ウェブ辞書(<http://chuta.jp>)』の約8000の見出し語に関わる情報を含む。各見出し語には、6段階の難易度、旧日本語能力試験出題基準の等級、重要度、品詞、語種、意味分類、異表記などの情報の他に、各見出し語の語義の一つ一つに複数の用例(作例とコーパスから集めた実例)を搭載している。

表1 日本語教育語彙表

項目ID	品詞	語種	難易度	等級	重要度	意味分類	異表記	用例	備考
1000000001	名詞	カタカナ	初級前半	初級前半	1	動物	クマ	クマは山に生息する。クマは肉食動物である。	
1000000002	動詞	カタカナ	初級前半	初級前半	1	動作	走る	彼は早く走る。彼は毎日走る。	
1000000003	名詞	カタカナ	初級前半	初級前半	1	動物	クマ	クマは山に生息する。クマは肉食動物である。	
1000000004	動詞	カタカナ	初級前半	初級前半	1	動作	走る	彼は早く走る。彼は毎日走る。	
1000000005	名詞	カタカナ	初級前半	初級前半	1	動物	クマ	クマは山に生息する。クマは肉食動物である。	
1000000006	動詞	カタカナ	初級前半	初級前半	1	動作	走る	彼は早く走る。彼は毎日走る。	
1000000007	名詞	カタカナ	初級前半	初級前半	1	動物	クマ	クマは山に生息する。クマは肉食動物である。	
1000000008	動詞	カタカナ	初級前半	初級前半	1	動作	走る	彼は早く走る。彼は毎日走る。	
1000000009	名詞	カタカナ	初級前半	初級前半	1	動物	クマ	クマは山に生息する。クマは肉食動物である。	
1000000010	動詞	カタカナ	初級前半	初級前半	1	動作	走る	彼は早く走る。彼は毎日走る。	

今後はこの語彙表を活用して、世界各地の現地語と日本語との二言語辞書の開発支援を進める予定である。

「日本語教育語彙表」の他に作成したデータベースとしては、(1)動詞の類義語リスト、(2)動詞のコロケーションリスト、(3)文化的発見記述文リスト、(4)機能語用例データベース「はごろも」があげられる。それぞれの内容は以下の通りである。

(1)動詞の類義語リスト

日本語教育語彙表の意味分類を手掛かりとして類義語を選別したもの。

(2)動詞のコロケーションリスト

NINJAL-LWP for BCCWJ (<http://nlb.ninjal.ac.jp>) や NINJAL-LWP for TWC (<http://nlt.tsukuba.lagoinst.info>) のコロケーション情報をもとに、日本語学習者に必要なコロケーションに関わる名詞のリストを動詞ごとに掲載し、それぞれの意味を解説したもの。

表2 動詞のコロケーションリスト

動詞	コロケーション	品詞	難易度	等級	重要度	意味分類	異表記	用例	備考
走る	早く走る	副詞	初級前半	初級前半	1	動作	早く	彼は早く走る。	「早く」は動作の速度を示す。
走る	毎日走る	副詞	初級前半	初級前半	1	動作	毎日	彼は毎日走る。	「毎日」は動作の頻度を示す。
走る	早く走る	副詞	初級前半	初級前半	1	動作	早く	彼は早く走る。	「早く」は動作の速度を示す。
走る	毎日走る	副詞	初級前半	初級前半	1	動作	毎日	彼は毎日走る。	「毎日」は動作の頻度を示す。
走る	早く走る	副詞	初級前半	初級前半	1	動作	早く	彼は早く走る。	「早く」は動作の速度を示す。
走る	毎日走る	副詞	初級前半	初級前半	1	動作	毎日	彼は毎日走る。	「毎日」は動作の頻度を示す。
走る	早く走る	副詞	初級前半	初級前半	1	動作	早く	彼は早く走る。	「早く」は動作の速度を示す。
走る	毎日走る	副詞	初級前半	初級前半	1	動作	毎日	彼は毎日走る。	「毎日」は動作の頻度を示す。
走る	早く走る	副詞	初級前半	初級前半	1	動作	早く	彼は早く走る。	「早く」は動作の速度を示す。
走る	毎日走る	副詞	初級前半	初級前半	1	動作	毎日	彼は毎日走る。	「毎日」は動作の頻度を示す。

これらの活動を推進するにあたり、毎年2回の全体集会と2回の分担者会議、および各班による不定期の会議を開催し、研究成果の報告と研究方針および研究計画の討議を行った。さらに、期間中に4回の国際研究集会「日本語学習辞書の支援を考える」を開催し、本プロジェクトの活動を世界の日本語教育関係者に周知するとともに、現地の日本語学習辞書の開発や使用の状況についての情報を集め、世界各地の日本語教育関係者と日本語・現地語の二言語辞書開発についての討議を行った。

- (1) 2012年3月18日 オーストリア国際研究集会 ウィーン大学
- (2) 2013年3月2日・3日 マレーシア国際研究集会 Crystal Crown Hotel, Johor Bahru
- (3) 2013年11月16日・17日 中央アジア国際研究集会 カザフスタン国立大学
- (4) 2014年12月7日 最終年度国際研究集会 筑波大学

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)

[雑誌論文](計22件)

砂川 有里子、「コーパスを活用した類義語分析—「連動的な変化の進展」を表す用法—」日本語教育連絡会議論文集、査読有、26、2014、49-60

Yuriko SUNAKAWA, Jae-ho LEE, Mari TAKAHARA, The Construction of a Database to Support the Compilation of Japanese Learners' Dictionaries, Acta Linguistica Asiatica, 査読有, 2, 2012, 97-115

杉本 武、コーパスに見る類義表現: 「気がつく」と「気づく」、文藝言語研究 言語編、査読有、62、2012、37-51

砂川 有里子、日本語教育へのコーパスの活用に向けて、日本語教育、査読有、140、2011、4-18

[学会発表](計99件)

砂川 有里子・李 在鎬・高原 真理、日本語学習辞書支援のためのデータベース構築、

EJAS, AJE ヨーロッパ日本語教育シンポジウムパネルセッション、2014.8.30、リュブリャーナ大学(リュブリャーナ、スロベニア)

山内 博之、語彙から始める日本語教育、第88回第二言語習得研究会(関東地区)招待講演、2014.6.14、お茶の水女子大学(東京都文京区)

野山 広、カール フォルスグラフ、田中 牧郎、小木 曾、迫田 久美子、言語に関するデータベース、コーパスの構築とその活用について、第5回「日本語教育とコンピュータ」国際会議(Castel/J)パネルセッション、2012.8.22、名古屋外国語大学(愛知県日進市)

砂川 有里子、コーパスを活用した日本語語彙リストの構築、日本語教育国際研究大会・パネルセッション、2012.8.19、名古屋大学東山キャンパス(愛知県名古屋市)

[図書](計10件)

山内 博之、凡人社、[新版]ロールプレイで学ぶ中級から上級への日本語会話、2014、130

藤田 保幸(編著)、砂川 有里子、杉本 武他、和泉書院、形式語研究論集、2013、383(35-60)

砂川 有里子、李 在鎬、石川 慎一郎、くろしお出版、日本語教育のためのコーパス調査入門、2012、248

益岡 隆志編、砂川 有里子他、ミネルヴァ書房、はじめて学ぶ日本語学、2011、264

[その他]

ホームページ等

<http://jisho.jpn.org>

<http://jreadability.net/jev>

6. 研究組織

(1) 研究代表者

砂川 有里子(SUNAKAWA, Yuriko)
筑波大学・人文社会系・教授
研究者番号: 40179289

(2) 研究分担者

松崎 寛(MATSUZAKI, Hiroshi)
筑波大学・人文社会系・准教授
研究者番号: 10250648

小野 正樹(ONO, Masaki)
筑波大学・人文社会系・教授
研究者番号: 10302340

山内 博之 (YAMAUCHI, Hiroyuki)
実践女子大学・文学部・教授
研究者番号: 20252942

李 在鎬 (LEE, Jae-Ho)
筑波大学・人文社会系・准教授
研究者番号: 20450695

井上 優 (INOUE, Masaru)
麗澤大学・外国語学部・教授
研究者番号: 30213177

築島 史恵 (YANASHIMA, Fumie)
政策研究大学院大学・政策研究科・
客員教授
研究者番号: 40401723

今井 新悟 (IMAI, Shingo)
筑波大学・人文社会系・教授
研究者番号: 50346582

杉本 武 (SUGIMOTO, Takeshi)
筑波大学・人文社会系・教授
研究者番号: 70196749

千葉 庄寿 (CHIBA, Shoujyu)
麗澤大学・外国語学部・教授
研究者番号: 70337723

堀 恵子 (HORI, Keiko)
東洋大学・人間科学総合研究所・
客員研究員
研究者番号: 70420809

迫田 久美子 (SAKODA, Kumiko)
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・日本語教育研究・情報セ
ンター・教授
研究者番号: 80284131

(3)連携研究者

長谷川 守寿 (HASEGAWA, Morihisa)
首都大学東京・人文科学研究科・准教授
研究者番号: 50272125

石田 プリシラ・アン (ISHIDA, Priscilla
Ann)
筑波大学・人文社会系・准教授
研究者番号: 10400607

鈴木 智美 (SUZUKI, Tomomi)
東京外国語大学・留学生日本語教育センタ
ー・准教授
研究者番号: 70332632

橋本 直幸 (HASHIMOTO, Naoyuki)
福岡女子大学・文学部・講師
研究者番号: 30438113

小林 ミナ (KOBAYASHI, Mina)
早稲田大学・日本語教育研究科・教授
研究者番号: 70252286

大関 浩美 (OZEKI, Hiromi)
麗澤大学・外国語学部・准教授
研究者番号: 50401584

五十嵐 陽介 (IGARASHI, Yosuke)
広島大学・文学研究科・准教授
研究者番号: 00549008

バルデシ プラシャント
(PARDESHI, Prashant)
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・言語対照研究系・
教授
研究者番号: 00374984

福永 由佳 (FUKUNAGA, Yuka)
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・日本語教育研究・情報
センター・研究員
研究者番号: 40311146

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

近藤 明日子 (KONDO, Asuko)
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・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

川端 一光 (KAWAHASHI, Ikko)
明治学院大学・心理学部・専任講師
研究者番号: 20506159

仁科 喜久子 (NISHINA, Kikuko)
東京工業大学・名誉教授
研究者番号: 40198479

宇津呂 武仁 (UTSURO, Takehito)
筑波大学・システム情報系・教授
研究者番号: 90263433
山本 和英 (YAMAMOTO, Kazuhide)
長岡技術科学大学・工学部・准教授
研究者番号: 40359708