様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　２ 年　 ７ 月 １３ 日現在

機関番号：３２６３４
研究種目：基盤研究(B)（一般）
研究期間：2016～2019
課題番号：１６Ｈ０３４２６

研究課題名（和文）「昭和話し言葉コーパス」の構築による話し言葉の経年変化に関する実証的研究

研究課題名（英文）An empirical study on the diachronic change of spoken Japanese through the construction of Showa Speech Corpus

研究代表者
丸山　岳彦（MARUYAMA, Takehiko）

専修大学・国際コミュニケーション学部・教授

研究者番号：９０３９２５３９

交付決定額（研究期間全体）：（直接経費）　　9,160,000円

研究成果の概要（和文）：本研究課題では、『昭和話し言葉コーパス』を構築し、一般に公開することにより、過去の日本語音声を分析するための基盤を整備した。過去の録音資料をコーパス化することにより、話し言葉に生じた経年変化を明らかにするための「通時音声コーパス」を実現することができる。本研究では、1950～1970年代の録音資料を収録した『昭和話し言葉コーパス』（約50時間）を構築・公開するとともに、国内外の学会における研究発表や国際シンポジウムの開催など通じて、その研究成果を発信した。

研究成果の学術的意義や社会的意義
過去の録音資料を現代の技術でコーパス化して、話し言葉の経年変化の実態を捉えようとする研究は、世界的に見てもイギリスのDCPSE、フランスのESLOなどがあるのみで、その実践例は極めて少ない。本研究では、『昭和話し言葉コーパス』の構築と公開により、日本語の話し言葉における経年変化をコーパスに基づいて研究する基盤を整備した。また、国際シンポジウムInternational Symposium on Diachronic Speech Corporaを開催し（2017年）、「通時音声コーパス」が持つ可能性を世界各国の研究者らと議論した。

研究成果の概要（英文）：In this project, we constructed the Showa Speech Corpus and released it to the public. A "diachronic speech corpus" can be achieved by collecting old recordings and connecting it to a contemporary speech corpus. We collected audio recordings created from the 1950s and 1970s, and compiled them into the Showa Speech Corpus. We also presented our research results at domestic and international conferences and international workshops.

研究分野：コーパス言語学

キーワード：コーパス言語学　昭和話し言葉コーパス　経年変化　コーパス　話し言葉　言語学　コーパス日本語学　話し言葉コーパス

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

１．研究開始当初の背景

1960 年代に端を発するコーパス言語学は、1990 年代のイギリスで活発化し、さらに世界各地で各国語のコーパスが作られるようになった。日本でも 2000 年代に入ってから国立国語研究所を中心として日本語コーパスの開発・公開が進み、現在では日本語の書き言葉コーパス・話し言葉コーパス・学習者コーパス・歴史コーパス・方言コーパス・マルチモーダルコーパスなど、多様なコーパスが利用できるようになっている。

本研究課題の申請時（2015 年）における日本語コーパスの整備状況を見た時、その時点では達成できていない種類のコーパスとして、「通時音声コーパス」があった。通常、ある言語の通時的な変化を探るために利用される通時コーパス（歴史コーパス）の収録対象は書き言葉であり、話し言葉は対象外となる。これは、日本語の場合、一定量を備えた書き言葉の資料は 7 世紀ごろまで遡ることができる一方、話し言葉の録音資料はせいぜい 100 年程度しか遡ることができず、かつそれらは散逸した状態にあるためである。もし古い時代に録音された音声資料を収集し、それらを現代の技術でコーパス化した上で、現代の話し言葉コーパスと比較することができれば、日本語の話し言葉に生じた言語変化の実態をコーパスに基づいて実証的に明らかにすることができる。

「通時音声コーパス」を構築したという研究事例は、イギリスの DCPSE（Diachronic Corpus of Present-day Spoken English）やフランスの ESLO（ENQUÊTE SOCIOLINGUISTIQUE D'ORLÉANS）など少数の事例を除き、世界的に見てもほとんど存在しない。申請者が当時所属していた国立国語研究所には 1950 年代に開始した録音作業による録音資料が多く残されていたことから、これらをコーパス化することにより、話し言葉の経年変化を実証的に分析するためのコーパス『昭和話し言葉コーパス』を構築することを計画した。

２．研究の目的

本研究課題の申請時における当初の研究目的は、以下の 3 点である。

(1) 1950～1970 年代の録音資料（講演、雑談）を体系的に整理し、種々のアノテーションを施して、50 時間分（独話 25 時間、対話 25 時間）の『昭和話し言葉コーパス』を構築すること

(2) 『昭和話し言葉コーパス』を現代語の話し言葉コーパスと比較・対照して、音韻・語彙・文法・文体・発話行為などの観点から、話し言葉の経年変化の実態を明らかにすること

(3) 上記を通して、「コーパスに基づく話し言葉の経年変化に関する実証的研究」という、新しい研究領域の方法論を開拓すること

３．研究の方法

(1) 録音資料のカタログ化、メタデータの定義

『昭和話し言葉コーパス』に収録する対象として、1950 年代から 1970 年代にかけて国立国語研究所で作成された録音資料を収集・整理し、カタログを作成する。その際、録音時期、発話の種類（独話・会話）、発話場面、発話者の属性（性、年齢、社会層）などを指標としてメタデータを定義し、録音資料ごとに付与する。これによって、コーパスに含まれる各録音資料の性格を規定し、資料全体を体系づけるとともに、さまざまな場面・発話者を含むような構成・配分方法を検討することができる。また、このようなメタデータの付与によって、現代の話し言葉コーパスと比較・対照する際に有用な指標を得ることができる。

(2) 時間情報付き転記テキストの整備

録音資料を新たに書き起こし、転記テキストを作成する。この際、発話の実質的な内容だけでなく、フィラー（「えー」や「あのー」など）や言いよどみ・言い間違いなどの「非流暢性」についても、できるだけ忠実に書き起こす。非流暢性の書き起こし基準は、『日本語話し言葉コーパス』で定義された基準に準拠する。これにより、『昭和話し言葉コーパス』と『日本語話し言葉コーパス』の間で、非流暢性の分布を精緻に比較することができる。また、「私」のように複数の読みを許す場合も、発音された形が同定できるような表記基準によって転記を行なう。

(3) 形態素解析

作成した転記テキストを、形態素解析辞書「UniDic」によって形態素解析し、人手による修正を経た上で、「短単位データベース」を構築する。これらは、国立国語研究所における既存のコーパスと同一の体系とすることにより、各コーパスとの比較が可能な状態にしておく。

(4) 『昭和話し言葉コーパス』と研究成果の一般公開

構築した「短単位データベース」およびメタデータをまとめて、『昭和話し言葉コーパス』公開用データとして準備し、国立国語研究所で管理・運用されているウェブ上のコーパス検索アプリケーション「中納言」で一般公開する。これと並行して、『昭和話し言葉コーパス』の構築および分析に関する研究成果を、国内外の学会・シンポジウム等で発表する。

４．研究成果

本研究課題の実施による成果（研究の主な成果、得られた成果の国内外における位置づけとインパクト）として、以下の 3 点を挙げる。

(1) 『昭和話し言葉コーパス』の公開

2018 年度末に『昭和話し言葉コーパス』（モニター公開版：DVD 版）を公開し、さらに 2019 年度末に『昭和話し言葉コーパス モニター公開版』をウェブ上のコーパス検索アプリケーション「中納言」で一般公開した。2020 年度末には、完成版の『昭和話し言葉コーパス』を「中納言」で一般公開する予定である。

(2) 「話し言葉の経年変化」に関する分析結果の発表

『昭和話し言葉コーパス』の構築および分析の結果は、国内外の学術学会・ワークショップ・シンポジウムなどでの研究発表、論文や書籍の刊行などを通じて、その成果を発信した。

(3) 国内外における位置づけとインパクト

2017 年 9 月、イギリス・フィンランド・イタリア・フランスからの研究者を招聘して国際ワークショップ International Symposium on Diachronic Speech Corpora を国立国語研究所において開催した。さらに、フランス、イタリア、チェコ、日本の各地で、『昭和話し言葉コーパス』をはじめとする日本語コーパスに関するワークショップを開催し、研究成果を発信した。

今後の展望として、2020 年度末に『昭和話し言葉コーパス』完成版の一般公開を控えているほか、フランス国立図書館が刊行する論文集 *Mémoire sonore duJapon：le disque, la musique et la langue* への論文の掲載（現在印刷中）、さらに本研究課題での研究成果を 1 冊に収めた論文集の刊行（2021 年度末出版予定）を予定している。

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　0件／うち国際共著　0件／うちオープンアクセス　1件）

１．著者名 丸山岳彦	４．巻 282
２．論文標題 『昭和話し言葉コーパス』の計画と展望 ―1950年代の話し言葉研究小史―	５．発行年 2016年
３．雑誌名 専修大学人文科学研究所月報	６．最初と最後の頁 39-55
掲載論文のDOI（デジタルオブジェクト識別子） http://doi.org/10.34360/00007004	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計22件（うち招待講演　7件／うち国際学会　4件）

１．発表者名 Takehiko Maruyama
２．発表標題 Annotation and analysis of Japanese spontaneous speech
３．学会等名 大阪府立大学大学院人間社会システム科学研究科 講演会 Annotation of speech corpora: sharing experiences（招待講演）（国際学会）
４．発表年 2018年

１．発表者名 丸山岳彦
２．発表標題 通時音声コーパスの可能性：『昭和話し言葉コーパス』の構築と分析
３．学会等名 東京外国語大学 語学研究所 定例研究会（招待講演）
４．発表年 2018年

１．発表者名 松下晶子・丸山岳彦
２．発表標題 脚本テキストに基づくコーパス文体論の可能性 ―テレビドラマ脚本に注目して―
３．学会等名 言語資源活用ワークショップ2018
４．発表年 2018年

１．発表者名 Takehiko Maruyama and Hanae Koiso
２．発表標題 Diachronic Change of Spoken Japanese in the 20th Century: a corpora-based Study
３．学会等名 50 ans de linguistique sur corpus oraux : apports a l'etude de la variation（国際学会）
４．発表年 2018年

１．発表者名 丸山岳彦
２．発表標題 「通時音声コーパス」から見る話し言葉の経年変化
３．学会等名 日本語文法学会 第19回大会（招待講演）
４．発表年 2018年

１．発表者名 Takehiko Maruyama
２．発表標題 Diachronic change in spoken Japanese in the 20th century
３．学会等名 フランス国立図書館 国際シンポジウム Memoire sonore du Japon : le disque, la musique et la langue（招待講演）（国際学会）
４．発表年 2019年

１．発表者名 丸山岳彦
２．発表標題 『昭和話し言葉コーパス』の構築(3)：その進捗状況と問題点
３．学会等名 シンポジウム「日常会話コーパス」Ⅳ
４．発表年 2019年

１．発表者名 佐野真一郎
２．発表標題 ピッチレンジと言語外的要因 ―昭和話し言葉コーパス・日本語話し言葉コーパスを用いて―
３．学会等名 シンポジウム「日常会話コーパス」Ⅳ
４．発表年 2019年

１．発表者名 相澤正夫
２．発表標題 SP盤演説レコードがひらく音声変異研究の可能性
３．学会等名 「通時コーパス」シンポジウム2019
４．発表年 2019年

１．発表者名 丸山岳彦
２．発表標題 言語学から見た脚本アーカイブズの可能性
３．学会等名 脚本アーカイブズシンポジウム2019『脚本で振り返る「平成」という時代』（招待講演）
４．発表年 2019年

１．発表者名 佐野真一郎
２．発表標題 自然発話における単音・促音の持続時間変異 ―最小対、対立保持、情報性の観点から―
３．学会等名 シンポジウム「日常会話コーパス」III
４．発表年 2018年

１．発表者名 丸山岳彦
２．発表標題 『昭和話し言葉コーパス』の構築(2)： その進捗状況と問題点
３．学会等名 シンポジウム「日常会話コーパス」III
４．発表年 2018年

１．発表者名 服部匡
２．発表標題 「～テございます」の使用傾向の推移―「～テいる」「～テある」との対応関係に注目して―
３．学会等名 シンポジウム「日常会話コーパス」III
４．発表年 2018年

１．発表者名 山口昌也
２．発表標題 国会会議録における言語表現の時間的変化の予備的分析 ―衆議院本会議を対象に―
３．学会等名 シンポジウム「日常会話コーパス」III
４．発表年 2018年

１．発表者名 シンポジウム「日常会話コーパス」III
２．発表標題 What's left for diachronic research of Japanese Speech?
３．学会等名 International Symposium on Diachronic Speech Corpora（国際学会）
４．発表年 2017年

１．発表者名 山口昌也
２．発表標題 国会会議録における言語表現の時間的変化の予備的分析
３．学会等名 言語資源活用ワークショップ2017
４．発表年 2017年

１．発表者名 丸山岳彦
２．発表標題 『昭和話し言葉コーパス』の構築
３．学会等名 シンポジウム「日常会話コーパス」I
４．発表年 2016年

１．発表者名 服部匡
２．発表標題 話し言葉のコーパスを用いた変化・変異研究の可能性
３．学会等名 シンポジウム「日常会話コーパス」II
４．発表年 2017年

１．発表者名 丸山岳彦
２．発表標題 『昭和話し言葉コーパス』におけるメタデータの設計
３．学会等名 シンポジウム「日常会話コーパス」II
４．発表年 2017年

１．発表者名 丸山岳彦
２．発表標題 『昭和話し言葉コーパス』の設計・構築と分析
３．学会等名 言語処理学会第26回年次大会
４．発表年 2020年

１．発表者名 丸山岳彦
２．発表標題 録音資料から知る、20世紀の日本語の変化
３．学会等名 第13回NINJALフォーラム（招待講演）
４．発表年 2018年

１．発表者名 丸山岳彦
２．発表標題 通時音声コーパスの可能性：『昭和話し言葉コーパス』の構築と分析
３．学会等名 東京外国語大学 語学研究所 定例研究会（招待講演）
４．発表年 2018年

〔図書〕　計2件

１．著者名 丸山岳彦	４．発行年 2019年
２．出版社 くろしお出版	５．総ページ数 178
３．書名 ことばと文字 11号 （「日本語コーパスをめぐる現状と展望」）	

１．著者名 MARUYAMA, Takehiko	４．発行年 2020年
２．出版社 Znanstvena zalozba FF.	５．総ページ数 242
３．書名 Japanese Language from Empirical Perspective: Corpus-based studies and studies on discourse. (On the Possibility of a Diachronic Speech Corpus of Japanese)	

〔産業財産権〕

〔その他〕

『昭和話し言葉コーパス』
https://pj.ninjal.ac.jp/conversation/showaCorpus/index.html

International Symposium : Diachronic Speech Corpora
http://pj.ninjal.ac.jp/conversation/event/sympo2017.html

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	山崎　誠 (YAMAZAKI Makoto) (30182489)	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・言語変化研究領域・教授 (62618)	
研究分担者	山口　昌也 (YAMAGUCHI Masaya) (30302920)	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・音声言語研究領域・准教授 (62618)	
研究分担者	佐野　真一郎 (SANO Shin-ichiro) (30609615)	慶應義塾大学・商学部（日吉）・准教授 (32612)	

６．研究組織（つづき）

	氏名 （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	相澤　正夫 (AIZAWA Masao) (80167767)	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・言語変化研究領域・教授 (62618)	
研究分担者	久能　三枝子（高田三枝子） (TAKADA Mieko) (90468398)	愛知学院大学・文学部・准教授 (33902)	
連携研究者	服部　匡 (HATTORI Tadasu) (40228490)	同志社女子大学・表象文化学部・教授 (34311)	