

令和 5 年 6 月 22 日現在

機関番号：33915

研究種目：若手研究

研究期間：2019～2022

課題番号：19K13204

研究課題名(和文)複合辞および連語形成の通時的・計量的研究

研究課題名(英文)A Historical and Quantitative Study of Compound Functional Expressions and Collocations

研究代表者

市村 由貴(渡辺由貴)(WATANABE, Yuki)

名古屋女子大学・文学部・准教授

研究者番号：10569776

交付決定額(研究期間全体):(直接経費) 3,200,000円

研究成果の概要(和文):本研究は、『日本語歴史コーパス』(CHJ)から取得した多様な資料にみられる語の連続について、通時的かつ計量的に明らかにするものである。CHJから抽出されたN-gramデータに文法および意味用法に関する情報を付与した。同時に、N-gramの使用状況の資料間の共通度も比較・検討した。口語資料においては共通した複合辞や引用表現が多用されていること、待遇表現や語りの定型的表現等においてそれぞれの資料性が強くあらわれること等を明らかにした。さらに、語の連続という観点から日本語のモダリティ表現の通時的变化を検討した。

研究成果の学術的意義や社会的意義

本研究において語の連続を一律に抽出し、その一覧表を作成することにより、複合辞・連語、あるいは出現頻度の高い語の連続形式を数値に基づいて客観的に提示することが可能となった。公開された成果は、日本語史分野における複合辞・連語研究の基礎的データとなりうるものである。また、本研究は、現代語で用いられている複合辞の成り立ちや、複合辞化・連語化の理論構築、辞書の記述・古典研究・古典教育への応用にも資するものである。

研究成果の概要(英文):This study aims to analyze the sequences of words obtained from various materials of the Corpus of Historical Japanese (CHJ) diachronically and quantitatively. Information on grammar and semantic usage was integrated with N-gram data extracted from CHJ, and the degree of commonality of each type of material was compared and examined based on the usage of the sequence of words. It was determined that common compound words and quoted expressions were frequently used in the colloquial materials, and that the characteristics of each type of material strongly appeared in the attitudinal expression and stereotyped expressions of narratives. Furthermore, diachronic changes in the modality expressions of Japanese were examined from the viewpoint of word sequences.

研究分野：日本語学

キーワード：複合辞 連語 コーパス 中世語 日本語史 N-gram

## 1. 研究開始当初の背景

複合辞・連語の史的な研究については多くの蓄積があるが、これまでは各研究者が複合辞あるいは連語と思われる形式にあたりをつけて考察していく手法が主なものであった。しかし、内省のききにくい日本語史における複合辞・連語の抽出と認定にあたっては、計量的な側面からの考察が重要であると考え、2016年度より科研費の課題として、日本語歴史コーパス（以下、CHJ）のデータに基づき、複合辞や連語が多用されるようになったといわれる中世・近世語の複合辞および連語の研究にとりくんできた。

この研究を進める中で、さらなる課題がみえてきた。まず、①同時代の資料であっても、用いられる複合辞・連語的表現の傾向に違いがある点である。例えば、『虎明本狂言集』では「である」相当の複合辞の種類がキリシタン資料と比べ豊富である。また近年、複合辞と認めうる形式は中古にもみられると指摘されており、②どのような語が一語化し、複合辞・連語の要素となりやすいのかを検討するためには、中古から近代に至るまでの多くの資料について分析する必要があると考えた。さらに、複合辞・連語は語の連続が固定化したものであるが、③一度固定化した形式が、時代とともに衰退することもある。例えば、中世に多用されていた「にて候」は、「候」という語の使用頻度が下がるとともに衰退したと考えられる。しかし、同じく中世に多用されていた接続助詞相当の「ほどに」は、「ほど」という語自体は現代でも使われているのにもかかわらず衰退し「から」「ので」等の助詞にとってかわられており、形式の衰退には様々な要因があると考えられる。

これらの課題を解決するためには、語の連続を広い視野から計量的に見渡す必要があった。

## 2. 研究の目的

本研究の目的は、中古から近代における複合辞・連語の通時的な使用状況を数値的裏付けにもとづいて時代別・資料別に示した上で、その形成・衰退過程を明らかにすることである。具体的には以下の3点の解明を目指した。

- ①日本語史上にみられる複合辞・連語にはどのようなものがあり、それぞれの時代・資料においてどのくらい使われているか（通時的な複合辞・連語一覧表の作成）。
- ②通時的観点から、どのような語が一語化し、複合辞や連語の要素となりやすいのか。
- ③一度固定化した複合辞や連語が衰退する際のメカニズムはどのようなものであるか。

## 3. 研究の方法

形態素解析がなされたCHJのデータから、連続する一定の単語数を切り出した「単語 N-gram」を抽出し、その使用頻度をもとに資料ごとに多用されている複合辞・連語的表現を整理し、あわせて複合辞・連語的表現の出現状況の観点から、各資料の位置づけを検討した。

- ①複合辞・連語アノテーション用 N-gram データの作成：

CHJ から複数資料を選定して2語～7語の N-gram データを抽出し、アノテーション作業用に整備した。

- ②複合辞・連語アノテーション：

上記①の N-gram データに、複合辞・連語の分類情報のアノテーションを行った。

- ③分析対象資料の拡充：

中世以降の言語資料から追加選定した複数資料をテキストデータ化し、①②のデータの比較対象として整備した。

- ④アノテーションされたデータに基づく複合辞・連語の研究：

複合辞・連語的表現の一覧表の作成および複合辞・連語の構成要素の分析、資料間・時代間の使用状況の比較等を行った。

## 4. 研究成果

- (1) 複合辞・連語の使用状況の調査およびこれをもとにした日本語史資料の位置づけについて

『虎明本狂言集』『天草版平家物語』等を中心に、複数の日本語史資料を N-gram の側面から比較し、使用頻度の高い表現や資料間の共通度等を明らかにした。2gram・3gram においては、使用頻度上位に複合辞や助詞・助動詞の連続、接続詞相当の連語、疑問や引用に関する表現等が多くみられ、同時代の口語資料間では、使用頻度上位の表現の多くは共通していた。一方、特定の資料のみに多くみられる表現は、当該資料の文体面・内容面の特徴を反映するものであった。また、この N-gram の使用状況に基づき、各資料の性格・位置づけを検討した。成果は「短単位 N-gram からみた『虎明本狂言集』と『天草版平家物語』の表現の特徴」（青木博史・小柳智一・吉田永弘編『日本語文法史研究 5』）、「短単位 N-gram からみた『天草版平家物語』と『天草版伊曾保物語』の表現の特徴」（青木博史・岡崎友子・小木曾智信編『コーパスによる日本語史研究 中古・中世編』）等において公開した。

- (2) 複合辞と助動詞の史的変遷について

「と-思う」「と-考える」のような文末思考動詞の史的変遷について、「だろう」や「べし」等の推量の助動詞との類似性や、推量の助動詞の体系の変化とのかかわりから検討した。古代語から近代語への流れの中で、衰退した助動詞「べし」のもっていた意味範囲のうち、過去の不確かな記憶について述べる場合等、現代語の「だろう」では担えない部分を、文末思考動詞が引き継いでいる状況が示唆された。また、洋学資料における使用状況から、“I think”のような欧文脈によらず、明治初期にはすでに推量表現に準じて用いられる「と-思う」が日本語の話し言葉のなかで定着していたこと、一方で「と-考える」「と-推察される」のような文末思考動詞のバリエーションについては、欧文直訳体での使用が、その定着を促す契機となった可能性があることを指摘した。成果は「文末思考動詞と推量の助動詞—「と思う」と「べし」の類似性を中心に」（早稲田大学日本語学会編『早稲田大学日本語学会設立60周年記念論文集 第2冊 言葉のはたらき』）、「洋学資料における文末思考動詞」（『国文学研究』194）等において公開した。

### (3) 歴史コーパスの構築・利用に関する解説の作成

コーパスを利用した日本語史研究に関する解説の作成を行った。具体的には、国立国語研究所『日本語歴史コーパス』の「室町時代編Ⅰ狂言」・「室町時代編Ⅱキリシタン資料」について、解説（いずれも共著）を作成し、各コーパスの概要・特徴・背景等を述べた（青木博史・岡崎友子・小木曾智信編『コーパスによる日本語史研究 中古・中世編』）。また、青木博史・高山善行編『日本語文法史キーワード事典』において、項目「通時コーパス」の執筆を行った。『日本語歴史コーパス』の活用法を示したものとして、田中牧郎編『コーパスで学ぶ日本語学 日本語の歴史』の「室町時代」を執筆した。

5. 主な発表論文等

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

1. 著者名 渡辺由貴	4. 巻 194
2. 論文標題 洋学資料における文末思考動詞	5. 発行年 2021年
3. 雑誌名 国文学研究	6. 最初と最後の頁 104(1)-92(33)
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 渡辺由貴
2. 発表標題 文法（シンポジウム「早稲田の日本語研究のこれから」 記念論集2巻の分野）
3. 学会等名 早稲田大学日本語学会 設立60周年記念大会
4. 発表年 2021年

1. 発表者名 Kurumi KATAYAMA, Toshinobu OGISO, Yuki WATANABE
2. 発表標題 Construction of a Corpus of “Christian Materials” for the Study of Colloquial Japanese of the Muromachi Period
3. 学会等名 Digital Humanities Conference 2019（国際学会）
4. 発表年 2019年

〔図書〕 計5件

1. 著者名 青木 博史、岡崎 友子、小木曾 智信	4. 発行年 2022年
2. 出版社 ひつじ書房	5. 総ページ数 340
3. 書名 コーパスによる日本語史研究 中古・中世編	

1. 著者名 早稲田大学日本語学会	4. 発行年 2021年
2. 出版社 ひつじ書房	5. 総ページ数 432
3. 書名 早稲田大学日本語学会設立60周年記念論文集	

1. 著者名 田中 牧郎	4. 発行年 2020年
2. 出版社 朝倉書店	5. 総ページ数 192
3. 書名 日本語の歴史	

1. 著者名 青木 博史、小柳 智一、吉田 永弘	4. 発行年 2020年
2. 出版社 ひつじ書房	5. 総ページ数 312
3. 書名 日本語文法史研究 5	

1. 著者名 青木 博史、高山 善行	4. 発行年 2020年
2. 出版社 ひつじ書房	5. 総ページ数 164
3. 書名 日本語文法史キーワード事典	

〔産業財産権〕

〔その他〕

-

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
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------