

令和 2 年 6 月 6 日現在

機関番号：32710

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K02749

研究課題名(和文) 記述・フィールド言語学向け言語資料の理論的・実践的言語ドキュメンテーション研究

研究課題名(英文) A study of language documentation for descriptive or field linguistics to establish a theory for recording of emerging data

研究代表者

大矢 一志 (Ohya, Kazushi)

鶴見大学・文学部・教授

研究者番号：80386911

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：言語ドキュメンテーションの理論的活動として、言語研究者からの要求分析と動向調査の結果を元に、言語研究者が主体的にデジタル環境上で言語資料をデータ化するために必須の知識であるデータ構造を学習するモデルカリキュラムとその実践教材を独自開発のwebシステム上で作成・公開した。また実践的活動としては、ロシア・トムスク州でのフィールド調査を実施、採録したデータを元にセリクープ語の子供向けイラスト付き語彙集の作成に向けたデータ整理にあたった。さらにロシア・カムチャッカ地方のイテリメン語のデータ整理にあたり、発語資料集の発行の他、データ整理にあたった。

研究成果の学術的意義や社会的意義

言語ドキュメンテーション活動には、個人活動としてある言語資料の収集、整理、利用、管理を計算機を使いながら研究成果を高める活動と、より多くの言語をフィールドから記述・記録し、学術研究で共有すると共に後世に言語資料を伝える活動がある。ところが、現在の多くの言語ドキュメンテーション活動は、後者の活動に資金が投じられ、結果としての成果も、共有可能な状態にある言語資料が学術的資料として個人の研究活動に容易に利用できる状況にはない。本研究では、言語ドキュメンテーションのうち前者の活動に焦点をあてた公金プロジェクトとして学術的意義は高い。

研究成果の概要(英文)：This study had targeted establishment of a documentation theory that reflects a state of data emersion in a field or data management on computational environment. Based on observation of data emersion on field notes or computational environment, and requirement analysis from language researchers and also investigation of academic or project movements on especially archive/repository projects of language resources, we confirmed that international-scale archives and/or projects based on authentic standards of data model seem to rapidly fail in attracting interest from language researchers due to too-short lifetime. Thus, this study made an online lecture on data structure that is required knowledge for language researcher who have to set digital environment for their research activities in order to make their language resources being in their lifetime. In addition, the online lecture is issued on a new web system that is easy to make, manage, and preserve the content.

研究分野：言語学

キーワード：language documentation オンライン教材 web学習システム データ構造 lectures for radio フィールド調査 セリクープ語 イテリメン語

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。

## 様式 C-19、F-19-1、Z-19(共通)

### 1. 研究開始当初の背景

言語ドキュメンテーション活動の重要性は広く理解され、共同研究の成果として大規模なアーカイブやレポジトリの作成を目指すプロジェクトが世界規模で生まれていた。ところが、それらの殆どが最終成果を言語資料の共有・公開を目的としたシステムとし、これらは言語研究者よりむしろ計算機科学者の活動実績として評価される対象物であった。その結果、アーカイブやレポジトリへの提供が求められていた言語資料のデータ構造等の国際規格は、データベース寄りの要求に応える仕様で、言語研究者が自らの研究活動で使用するデータ管理環境を全く想定していないものであった。ある意味ではデータ変換と規格の学習が負担となり、言語研究者にとっては本来の個人研究活動を阻害する提案といえるものであった。そこで研究代表者である大矢は、大規模アーカイブ・レポジトリで求められる規格化されたデータ形式の評価と、個人が管理する言語資料を規格データ形式へと変換するソフトウェアの開発に取り組んできた。また、システム開発のそもそも論として、言語資料を作成・利用するのは言語研究者であり、その言語研究者が自らの研究活動で求める要求を分析(要求分析)し、それを元にした仕様(仕様書)から、言語ドキュメンテーション活動向けのシステムを開発していないという現状に対して疑問を抱くようになった。そこで、言語研究者が自らの研究活動で求めているデジタル環境を分析するための要求分析をすることにした。これらの研究活動は基礎研究として位置づけられる。従って、研究活動の評価が得られるのかという不安はあったものの、科研費でこそ実施される研究に相応しいと考え、研究計画を立案・申請し、採択された。

### 2. 研究の目的

言語研究者が自らの個人研究活動において、言語資料を計算機上で入力・整理・管理・保存・再利用する、データの全ライフサイクルがどのように行われているのかを観察・分析し、そこから要求分析をすることで、言語ドキュメンテーション活動に必要な仕様をまとめることを、本研究の大きな目的とした。これらを計画書では、実践的フィールド上ドキュメンテーション、実践的フィールド外ドキュメンテーション、理論的フィールド上ドキュメンテーション、そして理論的フィールド外ドキュメンテーションの4つの視点で整理している。但し、この中でも手つかずの領域で、かつ情報処理の一番の勘所でもある、データ化される情報そのものの分析を、理論的フィールド上研究の優先目標とした。フィールド上でどのようなデータが記録・記号化され、それがどのように計算機上で入力されているのかという手順の観察・分析である。その主たる観察対象はフィールド・ノートとした。但し、フィールド・ノートの分析は、分析方法すら存在していない状況であることから、目に見えた研究成果としては理論的フィールド外研究として高汎用性のデータ変換ソフトの開発を目指した。また実践的研究としては、フィールド上の活動としては言語調査を、フィールド外の活動としては、言語資料の整理をすすめる言語ドキュメンテーション活動そのものを目標とした。

### 3. 研究の方法

理論研究の具体的な活動として、(1)変換ソフトの作成、(2)規格動向の調査、(3)フィールド・ノートの分析、(4)言語研究者が自らデータ変換をする際またデータそのものを作成する際に有用または必要となるデータ構造の知識を学習する教材の開発、(5)そのモデルシラバスの作成などを設定した。実践研究の具体的な活動には、(a)西シベリア地方、カムチャツカ地方でのフィールド調査、(b)採取した言語資料の整理とアノテーション付加等によるデータ編集を設定した。

#### 4. 研究成果

上記の研究計画は実施にあたり変更されることになった。活動1年目(2017年度)においてインフォーマントの死去により計画されていたフィールド調査が実施できず、フィールド・ノートの観察も中止されることになった。また、フィールド・ノート上の記述現象の観察については、本科研費採択にあたり、研究者同士の情報共有と公開についての倫理的課題の検討が採択条件として付加されたことから、学内倫理委員会でこれを検討、結果として研究遂行が認められたものの実践的フィールド調査が次年度以降になった事態も勘案し、本活動計画からフィールド・ノートからの要求分析を除外することにした。本研究事業以前からフィールド・ノート分析には取り組み、文献調査にも着手していたこれまでの実績もあることから、基礎研究として計画当初から成果物の期待は薄く、またデータ変換等の他の理論研究に注力することで本事業の目標到達に大きな支障はないと判断し、フィールド・ノート調査を活動項目から外すこととした。

(1)変換ソフトの作成は、言語研究者が自ら作成する言語資料と、(2)規格動向調査から得られた変換ターゲットとなるデータ形式の情報から、実装されることになる。ところが、前回の科研費で調査・確認されていた研究のトレンドが、本科研費期間中に大きく様変わりした。具体的には統一データ形式によるアーカイブへの要求が薄れてきたことである。大規模プロジェクトとして期待されたアーカイブ・レポジトリが短期プロジェクトとして終焉し、結果、データ提供の協力を求められた言語研究者は、国際規格に準拠したデータ形式による大規模プロジェクトに対して自らの研究活動として活用できるシステムであるとの評価を与えることが困難な状況になっていた。すなわち、変換ソフトを使い提供する予定であったアーカイブ・レポジトリの存在意義が、個人研究活動からすると薄れてしまったといえる。本研究は、個人活動の要求分析を目的としたことから、当初の研究計画とに関わらず、この要求の変化には応えるべきと判断した。よって、変換ソフトそのものを研究目標から外すこととした。これに代わり本研究で重要度が増したのが(4)データ変換を実践する際に言語研究者にとっても重要となるデータ構造の知識を学ぶ教材の作成と、(5)そのカリキュラム・モデルプランの策定である。言語研究者がデータ管理を主導するための知識習得の機会を与えることが本事業の主たる項目となった。結果として、提案していたカリキュラム・モデルプランに沿った、言語研究者向けの学習教材を、独自に開発をしたweb教材システム上で作成、公開までの成果を出すことができた。図らずも、このweb学習システムは、本科研費事業終了後の、新型コロナウイルス禍の中で求められたオンライン教材システムとして絶大な効力を発揮し、これは本科研費事業の副次的な活用として留めておく価値がある。

実践的言語ドキュメンテーション活動としては、(a)フィールド調査として、ロシア・トムスク州においてセリクープ語の聞き取り調査を実践した。但し、当初の計画としてあったカムチャツカ地方での調査は、インフォーマントの死去等の状況下で、実施できなかった。

(b)フィールド調査で収集した言語データの整理は順調に進み、カムチャツカ地方で収録していた言語資料等を元に対話資料を出版した。また、トムスク州で収録された言語資料についても子供向けのイラスト付き語彙集の作成に向けたデータ整理をすすめた。

## 5. 主な発表論文等

〔雑誌論文〕 計4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 1件）

1. 著者名 Nagasaki, Iku	4. 巻 154
2. 論文標題 The focus construction in early modern Kolyma Yukaghir	5. 発行年 2018年
3. 雑誌名 Gengo Kenkyu	6. 最初と最後の頁 123-152
掲載論文のDOI（デジタルオブジェクト識別子） 10.11435/gengo.154.0_123	査読の有無 有
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 -
1. 著者名 小野智香子	4. 巻
2. 論文標題 イテリメン語の類型的特徴 --系統か接触か--	5. 発行年 2019年
3. 雑誌名 第33回北方民族文化シンポジウム網走報告書 環北太平洋地域の伝統と文化 3 カムチャツカ半島・千島列島	6. 最初と最後の頁 37-42
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 無
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 -
1. 著者名 大矢一志	4. 巻 20
2. 論文標題 言語学の視点から整理するバフチンとオープンダイアログ	5. 発行年 2018年
3. 雑誌名 比較文化研究	6. 最初と最後の頁 1-21
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 無
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 -
1. 著者名 小野智香子	4. 巻 19
2. 論文標題 イテリメン語テキスト8	5. 発行年 2017年
3. 雑誌名 千葉大学ユーラシア言語文化論集	6. 最初と最後の頁 211-219
掲載論文のDOI（デジタルオブジェクト識別子） 10.20776/S21857148-19-P211	査読の有無 無
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9件（うち招待講演 2件 / うち国際学会 4件）

1. 発表者名 Nagasaki, Iku
2. 発表標題 A Kolyma Yukaghir Corpus with Morphological and Syntactic Annotation
3. 学会等名 4th International Symposium on Northern Languages and Cultures (国際学会)
4. 発表年 2020年

1. 発表者名 ,
2. 発表標題 -
3. 学会等名 国際会議「カムチャツカ地方の先住少数民族の言語と文化の保存と発展」(国際学会)
4. 発表年 2019年

1. 発表者名 ,
2. 発表標題 ?
3. 学会等名 国際会議「カムチャツカ地方の先住少数民族の言語と文化の保存と発展」(国際学会)
4. 発表年 2019年

1. 発表者名 長崎郁
2. 発表標題 ユカギールの言語と文化
3. 学会等名 シベリアの文化に触れてみる, 鶴見大学比較文化研究所(招待講演)
4. 発表年 2018年

1. 発表者名 小野智香子
2. 発表標題 イテリメンの言語と文化
3. 学会等名 シベリアの文化に触れてみる, 鶴見大学比較文化研究所 (招待講演)
4. 発表年 2018年

1. 発表者名 Ono, Chikako
2. 発表標題 Typological features of Itelmen and its neighboring languages
3. 学会等名 NINJAL International Symposium: Approaches to Endangered Languages in Japan and Northeast Asia: Description, Documentation and Revitalization
4. 発表年 2018年

1. 発表者名 Ono, Chikako
2. 発表標題
3. 学会等名 Conference on Uralic, Altaic and Paleo-Asiatic languages in the memory of Alexander P. Volodin
4. 発表年 2018年

1. 発表者名 Nagasaki, Iku
2. 発表標題 Ergativity in Early Modern Kolyma Yukaghir
3. 学会等名 28th Dulzon Readings (国際学会)
4. 発表年 2018年

1. 発表者名 長崎郁
2. 発表標題 グロス付きテキストから統語情報付きテキストへ --ユカギール語を例に--
3. 学会等名 科研費基盤研究(B)「シベリア少数言語の統語構造に関する類型論的研究」第1回研究会
4. 発表年 2018年

〔図書〕 計2件

1. 著者名 ,	4. 発行年 2020年
2. 出版社 北海学園大学出版会	5. 総ページ数 123
3. 書名 : , , &#1298; , &#700; &#399; &#1223; &#700; , &#700;	

1. 著者名 / . . . . .	4. 発行年 2018年
2. 出版社 科学研究費補助金基盤研究(B)『シベリア少数言語の統語構造に関する類型論的研究：従属節の構造と節連結を中心に』研究班	5. 総ページ数 226
3. 書名 (セリクーブ語ナリム方言の昔話)	

〔産業財産権〕

〔その他〕

Data Mode, Lectures for Radio <a href="https://infor.org/lr/lddm/lddm_top.html">https://infor.org/lr/lddm/lddm_top.html</a>
--

## 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	長崎 郁  (Nagasaki Iku)  (70401445)	名古屋大学・人文学研究科・特任講師    (13901)	
研究分担者	小野 智香子  (Ono Chikako)  (50466728)	北海学園大学・工学部・准教授    (30107)	