

## 科学研究費助成事業 研究成果報告書

平成 26 年 6 月 20 日現在

機関番号：32504

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23520697

研究課題名(和文)ドラヴィダ諸語の辞書アプリケーション作成

研究課題名(英文)Making a dictionary application for dravidian languages

研究代表者

箕原 辰夫 (Minohara, Tatsuo)

千葉商科大学・政策情報学部・教授

研究者番号：90279802

交付決定額(研究期間全体)：(直接経費) 1,500,000円、(間接経費) 450,000円

研究成果の概要(和文)：シンガポールおよび、南インドのタミルナードゥ州のチェンナイ、ケララ州のコーチン、カルナータカ州のバンガロールにおいて、タミル語、マラヤーラム語、カンナダ語の現地話者からの単語の発音について、それぞれ3000語ほどの収録を行なった。XMLで作成した共通の辞書データから、これらを学習用の辞書として、WebおよびiOS、Androidのアプリケーションとして公開していく。また、iBook、Kindle版の出版物も無償で提供して行く予定である。また、収録できなかったテルグ語については、今年度、個人研究費を用いてアーンドラ・プラデーシュ州のハイデラバードに赴いて収録を行なう予定にしている。

研究成果の概要(英文)：In this research, I recorded the pronunciations by the native speakers at Singapore and three cities in South India: Chennai in Tamil Nadu State, Kochi in Kerala State, and Bangalore in Karnataka State. The major three languages, Tamil, Malayalam, and Kannada were recorded in the recording tours. About 3000 words were recorded in each language. The recorded sound files and XML files which describe the dictionary data will be combined and published as Web pages in my public site, as free applications in iOS and Android, and as free digital publications in iBook and Kindle. The left major language, Telugu will be recorded by using the private research budget at Hyderabad in Andhra Pradesh State in this year.

研究分野：人文学

科研費の分科・細目：言語学・外国語教育

キーワード：辞書 ドラヴィダ語 タミル語 マラヤーラム語 カンナダ語 アプリケーション テルグ語

## 1. 研究開始当初の背景

近年、故大野晋氏が、日本語の祖語としての候補としてタミル語との関連を提唱するなど、南インドと日本との間に言語間の繋がりがあるという話題が出てきたのにも拘わらず、日本では、南インドで話されているドラヴィダ諸語についてのテキストが非常に少ないあるいは皆無の状況であった。出版されていても、すぐに廃版に近い形になっており、入手が困難であり、また『ドラヴィダ諸語の語源辞典』(A Dravidian Etymological Dictionary: T. Burrow と M. B. Emeneau による)の翻訳はあるが、これは一般的な語学学習には使えず、更には現地で用いられている文字では表記されておらず、翻字としての英字を用いた表記になっている状況であった。

研究開始から3年の間に、タミル語に関しては、袋井由布子氏による『タミル語入門』の本が出版された。また、山田圭子氏による『基礎テルグ語』は、2010年の出版であるが、継続的に手に入れることができる。しかしながら、2007年出版の山ノ下達氏による『マラーラム語文法』は、入手困難になっている。また、カンナダ語に関しては、日本語での出版物は存在しない。加えて、ドラヴィダ諸語を連合した形での紹介は、英語では R. Caldwell や B. Krishnamurti などの著書があるが、日本語では皆無である。南インドの言語を包括的に知るためには、これらの諸語を統一した形での学習用辞書やテキストの存在が必須であると考えていた。

## 2. 研究の目的

ドラヴィダ諸語を横断的に概括する局面と、それぞれの言語を個別に学ぶ局面の両方に使える学習用辞書を Web 教材およびアプリケーションとして作成するのが、この研究の目的である。特に、発音については、タミル語では Spoken Tamil と呼ばれるほど、表記と発音が口語では異なっている。そこで、単語について現地話者の発音の収録が必要と考えた。Web やアプリケーションとしてのデジタル教材を作成するのであれば、録音した発音を含めて出版するのは簡単にできる。この研究では、学習用辞書に録音素材を入れることを目的としている。また、この研究の目的として、南インドと日本との国際交流を支援するものと考えているので、成果物については無償での辞書・教材の提供を第一に考えている。

## 3. 研究の方法

まず、タミル語からシンガポールにおいて発音を収録することにした。シンガポールでは、タミル語が公用語の1つになっており、南インドのタミル・ナードゥ州とも連携しながら、タミル語教育が盛んに行なわれている。現地

話者の収録に関しては、シンガポール国立大学でタミル語教育の教鞭をとられている Thinnappan 教授にお世話になることとした。教授の教え子の学生(特に発音が聞き取りやすい女性の学生)3名から収録を行なった。シンガポールにおいては、袋井由布子氏の『旅の指さし会話帳:南インド』の巻末についている用語集および Web 上の素材から選定して、重複も含めて、3000 語程度のタミル語の発音を収録できた。

次に、最終年度には当初の予定通り、南インドの各州に赴いて、現地話者の発音を収録した。話者の選定は、ツアー会社の現地ガイドに委ねることになった。ここで問題になったのは、4つの言語、タミル語、テルグ語、カンナダ語、マラーラム語で収録する単語を統一するかどうかであった。辞書的には、単語の4言語の共通の最小集合を仮定するということになる。ただし、辞書としては音声の部分はなしで、いくらでも拡張することは可能である。要は、発音が収録される単語として、語彙の最小集合を仮定するかどうかであった。これに関しては、共通の辞書というのが、収録される段階で何も用意されておらず、唯一使用可能であったドラヴィダ諸語の語源辞典では、収録単語数が非常に多く、選定に時間が掛かることが問題になった。そこで、やむを得ず、各単語を別々の辞書のソースに頼ることにした。この間に、Collins My First English シリーズが刊行されたので、今後の単語選択にはこの本に収録されている単語の、最小集合として編纂して行くことを予定している。ただし、現在も Collins My First English - Tamil Dictionary は手元に届いておらず、amazon での入荷を待っている状態である。

単語の共通集合については、南インドの共通の固有名詞、あるいは風物についての名詞を収録することにした。また、今回の研究の目的から、袋井氏の本の分類を参考にして、日本語を母語とする旅人が南インドを訪ねたときに、様々な局面で使う単語についても、タグ付けして、局面ごとにカテゴライズできるようにも考えた。

ラテン語の辞書を作るときに、行なった単語動詞の関連性(動詞形、名詞形、形容詞形、形容動詞形、対義語、類似語などの関連語)について、この辞書でもリンク構造を採ることにした。これらの関連性や、収録した音声ファイルへのリンクについては、XML 形式のファイルで作成し、辞書アプリケーションや Web アプリケーションでは、それをファイルとして読み込んで、内部でオブジェクトモデルによるリンク構造として保持する形に実装している。

収録の成果の公表としては、Web サーバを用

意し、指定したアドレスから辞書にアクセスできるように整備しつつある。また、iOSおよびAndroidタブレット用の辞書アプリケーションも無償で提供していく。加えて、iBookおよびKindleの電子書籍の形で学習用辞書を提供することを予定している。

#### 4. 研究成果

この研究で実際シンガポールや南インドに訪れた目的の1つに、現地における言語教育がどのような形で行なわれているのかの事情を探ることにあつた。タミル語に関しては、シンガポールにおいてもチェンナイにおいても、大学を中心に、若年層までかなりの教育体制が整っていることが実際にシンガポール国立大学のThinnappan教授にお会いして伺うことができた。特に、シンガポールではタミル語の学習塾的な若年層を対象にした教室も設けられている。ところが、タミル以外の言語においては、南インドではあまり現地の言葉を研究対象にしていないことがわかった。これは、1つには市販されている教科書の欠落に現れている。コーチン、バンガロールで地元の本屋を訪ね歩いて、教科書がどの程度置いてあるかを見てみたが、学校で配られるものの他には簡単なテキスト材料しかなかった。その理由は、現地の就職事情にも依存しているのではないかと思われる。バンガロールは、IT cityとして名をあげつつあるが、結局そのような情報関係の職に就く為には、英語が必修であり、教育としては英語教育に重点が置かれている。それを反映してか、コーチンにせよ、バンガロールにせよ、IT関係の英語で書かれたテキストブックは、在庫が多く、中古本も売れている状況になっている。社会的に上位の階層に行くには、現地の言葉よりもインド共通で使える英語が話せる方が有利になっているという背景が伺えた。そのため、地元だけで使える言語の育成、研究、発展までには到っていないという実情が見え隠れした。飽くまでも、当地人に取っては、地元のローカルな方言的な扱いとしか感じられなかった。

逆に英語圏の話者や、インド出身でない我々を含めた外国人の方が、現地の言葉に関心があるように思える。今回の研究の英語版の辞書を作成し、インド国内で配布することによって、それぞれの州の間での意思疎通や、連帯感が生まれるのではないかと感じた。現地ガイドの人は、北インドのヒンディー語を母語とする人であったが、現地での収録につきあってくれて、実際に音を聞いてみて、マラーラムやカンナダの方が、ヒンディー語の祖語であるサンスクリット語からの借用が多い為、ヒンディー語と似ている単語が多いということを教えてくれた。このような状況については、Wikipedia等でも、同じことが書かれているが実際に収録してみて、現地の人感想として聞いたのは収穫であると考

えている。

なお、現地の手配の関係の不備から、テルグ語については現地話者の手配ができず、収録ができなかった。これについては、今年度の大学からの個人研究費の一部をこの収録のためにハイデラバードに赴いて、補完することを予定している。

タミル語については、Spoken Tamil に関しての本なども購入し、収録の際に、口語で用いている発音でお願いしてみた。しかしながら、正書法やその発音に関しては、文法書に従ったものを尊重するとのことを、現地話者から聞き、袋井由布子氏の主張されているとは異なることがわかった。口語での発音に関しては、現地話者にしてみれば、あくまでも現地での「方言」としての扱いであった。たとえば、チェンナイでの現地話者は、「それはチェンナイ方言になってしまうから」ということで、口語での発音は拒否された。タミル語の公式な発音の仕方については、正書法のものを採用することによって、さまざまな「方言」を話すタミル語の話者での共通の意思疎通ができることである。そのため、袋井氏がカタカナで振っている「読み」とはかなり異なる、正式（表音として文字に書かれている通り）の発音で収録した。このことは、中国語における北京語と上海語などの関係に準ずるものではないかと考えている。実際に収録を行なうことによって、このような考えを現地話者から聞いたことは大きな成果であると考えている。

文字学の分類からは、音素文字（アルファベット）、アブジャッド（ヘブライ・アラビア文字）、アブギダ（インド系の文字）、音節文字までのユーラシア大陸を通しての分類の流れが、言語文法上にも現れていることが、辞書作りの過程で明らかになった。アブジャッドは、元々セム系の文字であるので、音素文字を子音中心にしただけなので、音素文字の範疇に非常に近い。ところが、アブギダは、ハングル文字や日本の仮名文字の音節文字に近い形になっている。子音を中心に、母音記号を追加するという形だが、音の捉え方から考えれば、むしろ「音節」を基本にしていると考えられる。そのようなアブギダに分類されるインド系文字のハイブリッド性が、辞書収録の過程で調べたドラヴィダ諸言語の文法的な要素にも色濃く現れてきているのが確認できた。

ドラヴィダ諸言語においても、ラテン系の言語に見られるような、名詞の格変化と、動詞の人称変化、時制による変化を持つ。実際に、変化の種類が多いことから考えると、テルグ語は、ラテン語（インド・アリア系）に近い文法構成になっている。屈折語の言語として分類できるだろう。これは、テルグ語・カ

ンナダ語が、ヒンディー語を中心とする北インドの言語に語彙も依存していることとも大きな関連性があるだろう。それに比べて、タミル語は、むしろ膠着語に近くなっている。もちろん、名詞の格変化や、動詞の変形などを持つが、タミル語ではテルグ語に比べて、名詞の格変化が、寧ろ日本語の「てにをは」の助詞を付ける方式に近い。加えて、タミル語では一般名詞において、単数形が中心で、複数形はあまり使用される機会がない。このような部分は、日本語の用法と似ているように思える。同じドラヴィダ諸言語の中でも、タミル語系列とテルグ語系列では、かなり異なる部分が存在することを確認することができた。

今回は、タミル語・マラーラム語については、主格だけを収録することにした。カンナダ語だけは、上記のことを鑑みて、一般名詞について、すべての格を録音してみたが、単語数だけが増えて、話者が疲労してしまい、効果があるとは思えなかった。そこで南インドでそのあとから録音した他の2言語については、この経験を活かすこととなった。テルグ語についても、これから収録するにあたっては、代名詞などを除いては、主格だけに限ることを考えている。

なお、授業(担当する表現メディア論)の文字の学習の回において、デーバナーガリー文字と共に、タミル文字を書かせてみた。実際に日本語の氏名の表記をするときに使う子音の選定なども考慮が必要だとわかった。加えて、学生からは様々な質問が出たので、その部分を反映させることを考えている。

## 5. 主な発表論文等

この研究は、論文投稿や学会発表、あるいは産業財産権に主眼を置くものではなく、タイトルにあるように辞書を一般の学習者が使えるようなチャンネルを通して配布することにある。そのため、Web ページからの発信と、タブレット用のアプリケーションの無料配布を中心として成果物として考えたい。なお、雑誌論文は、2011年に大学紀要に発表している辞書の基本構造を考えた「スマートフォン用ラテン語学習辞書の編纂」を基礎としているもので、2011年の論文ではラテン語を対象としているが、これをドラヴィダ諸言語に当てはめて、辞書を実装する試作品を一般に提供することが、この研究の成果の公表であると考えている。

〔雑誌論文〕(計1件)

箕原 辰夫、ドラヴィダ諸語の辞書作成、千葉商科大学紀要、査読無、第52巻第1号、2014、(9月刊行予定)

〔学会発表〕(計0件)

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ(Webサイト):

シンガポールでのタミル語の辞書:

<http://marinegreen.pi.cuc.ac.jp/draavidian/tamil>

ドラヴィダ諸語の統一辞書:

<http://marine.pi.cuc.ac.jp/dictionary/draavidian>

Android/iOS のアプリケーション:

Google Play/Apple AppStore から、ドラヴィダ語学習辞書(英語版は A Learning Dictionary for Dravidian Languages)として無償配布する予定

Kindle/iBook の電子書籍:

Kindle Store, iBook Store から、『ドラヴィダ語入門』(英語版は、A Primer to Dravidian Languages)として無償配布する予定

すべての配布物の完成まで、あと数年は掛かると思われるが、これには補助とする研究費を必要とするものではないので、大学からの個人研究費を充てながら、公表して行きたい。

## 6. 研究組織

(1) 研究代表者

箕原 辰夫 (MINOHARA, Tatsuo)

千葉商科大学・政策情報学部・教授

研究者番号: 90279802