

# 地理情報システム (GIS) による エチオピアのデジタル言語地図

乾 秀行

(山口大学)

inui@yamaguchi-u.ac.jp

## 0 はじめに

本稿は、我々プロジェクトがフィールド調査で集めたエチオピアのオモ・クシ系少数言語に関する様々な言語データを、GIS(Geographical Information Systems)を用いてデジタル地図とどのように関連づけるかについて説明すると共に、現状報告するものである。

## 1 言語学の分野における言語地図作成

言語研究において地理情報は、言語地理学、言語接触、言語類型論、社会言語学、方言研究など、いろいろな言語学の分野の研究をする上で極めて重要なファクターであるが、今までは研究者が説明の必要に応じて自ら手書きで作成した言語地図を論文に添えるのが常であった。

言語地図として日本では、たとえば国立国語研究所所収の「日本言語地図」が有名である。これは日本全国各地の方言でどのような語形や発音がどこに現れるかを項目ごとに地図で表示したものであり、日本全国の方言の地理的分布を一望できる基礎資料と説明されている。国立国語研究所のホームページからは『日本言語地図』第1～5集がPDFで公開されている (<http://www5.kokken.go.jp/dash4/frameset.jsp>)。今後は地図化された情報を文字データにし、手書きの原表記や注記を含む原カードを画像データにするなどして、データベース化が図られる計画のようである。

一方、世界言語に目を向けると、世界規模で言語データを集めている Ethnologue のホームページからは、現在国別の言語地図を参照することができる (<http://www.ethnologue.com/>)。また、出版物としては包括的に世界言語の分布を地図上に表した Moseley & Asher (1994) がある<sup>1</sup>。ただし、これらはいずれ

---

<sup>1</sup>初版が出る前からすでにいくつかの誤りが指摘され、初版が出て10年以上経過し改訂版が

も地図上に言語名（系統分類含む）として分けられているだけで、直接的に音声や語彙や文法に関する言語特徴を参照することはできない。また出版物はアナログ地図であるため、実際に個々の研究者が言語研究に利用しようとしても、全体を視覚的に把握するための、あくまで参考資料としての使用に限られるのが現状である。そのため、今のところ言語地図は言語研究を進める上で副次的に参照するレベルにとどまっている。

ところでできるだけ多くの言語を対象に言語の普遍性を探究する言語類型論的研究は、統計処理中心の研究から言語地図と関連づける形へと発展してきている。たとえば松本(2000)では、世界言語を対象に言語類型地理論的観点から研究が進められており、従来の比較言語学では手に負えないような、言語の遠い親族関係を明らかにするための言語特徴が設定され、その言語特徴による世界言語地図の作成が考察する上で重要なカギとなっている。また、Haspelmath et al. (2005)では、50人を超える言語学者がそれぞれの専門領域（音韻、形態、文法、語彙など）に関して142枚の言語地図を作成し、世界で最初の包括的な言語構造特徴 (structural features) の地理的分布に関する研究を発表している<sup>2</sup>。ただし、142枚の言語地図はその特徴を持つ言語の位置に丸い点が示され色分けされて記述されているだけで、それぞれの言語特徴間の相関関係をパソコン画面上で地図として重ね合わせることはできないようである。そういう意味ではこれらの言語地図はあくまで一枚一枚独立したものであり、利用は制限されざるを得ない。

このような状況下、世界言語地図をデジタル化し、多くの言語学者がそれを共有できる環境を作ることは急務の課題であるといえる。本研究プロジェクトでは、エチオピアの地域研究に限定しているが、その手法は世界言語を対象にした言語類型論的研究にも今後役立つものと確信している。

ところでこの報告書作成中に気づいたことであるが、本プロジェクトと同じく Ethnologue を出している夏期言語研究所 (Summer Institute of Linguistics、略称 SIL) が Global Mapping International (GMI) 社と共同開発で Ethnologue の GIS 対応デジタル世界地図を作成し、World Language Mapping System (WLMS) として最近販売を始めたようである<sup>3</sup>。それは Ethnologue の第15版を基本に作成していて、アカデミック価格も存在するようである。どのようなことができるのか、今後そのシステムについて詳細に吟味してみる必要があるが、本プロジェクトの方針が世界的に見ても時代に合っていることを証明する1つの証拠といえよう。

---

待たれていたが、ようやく2006年12月に改訂第2版が出版される予定である。

<sup>2</sup>付属のCD-ROMで言語地図をパソコン画面上で見ることができる。

<sup>3</sup><http://www.gmi.org/wlms/index.htm#features>

## 2 GISの機能と言語研究での利用における可能性

GISでは画像処理の違いをベクトル形式とラスタ形式と呼んで区別している。この違いはたとえばコンピュータ上で点XとYを直線で引く場合、ベクトル形式では、平面上におかれた始点Xと終点Yの2つの座標値で位置情報を保持するのに対して、ラスタ形式では、二次元平面上に敷き詰められた画素を白黒で塗りつぶすことによって直線XYを表現することになる。データとしてはベクトル形式の方が座標値を正確に入力するので手間がかかり、地図を表示させるためのプログラムが複雑になるけれども、地図の精度は高くなりデータ量は比較的少なくてすむことになる。一方、ラスタ形式ではデータの構造は単純であるけれども、精度の高い地図を描くには画素を増やさなくてはならず、データ量が膨大となるという欠点がある。本研究ではインターネット上に情報公開し、遠隔地から検索プログラムを稼働させることを念頭に置いているので、できるだけ軽いデータにする必要がある。そのためベクトル形式で言語地図を作成している。

ところでGISを利用する最大のメリットは、「オーバーレイ」と呼ばれる地図の重ね合わせ機能にある。つまりGIS上では複数のレイヤーを自由に重ねることでデジタル地図を作ることが可能なため、いったんある言語特徴に関するデータを地図上に関連づけてしまえば、その後必要に応じて自由に地図を重ね合わせることができる。これは言語特徴の相関性を研究する言語類型論的研究や言語接触による様々なレベルでの類似性を研究する上で、強力なツールになりえるであろう。

それ以外にもGISにはいくつか言語研究に使える機能が常備されている。たとえば「バッファリング」と呼ばれる機能である。これは、ある点（あるいは面）が周りにどの程度影響を与えるかを示すバッファを設定できることである。たとえば大言語と弱小言語との社会言語学的な要因を調べる際、大言語がどの程度影響力を与えているかを地図上に示すことが可能となる。

また「ボロノイ分割」という機能もある。これは、地理学の中の「中心地理論」と呼ばれる考え方に基づくもので、人は一番近い施設を利用するという行動パターンを基礎にして商圈分析を行うことができる。この機能を利用してたとえば言語共同体から最も近いマーケットがどこにあるのかを地図上に表示することで、数量的に人の移動を捉えることが可能になるかもしれない。今までのフィールドでの社会言語学的調査は、たいてい適切な数の被験者に同じ質問をぶつけ、それを統計的に処理して結論を導き出すという手法を採ってきた。もしこうした社会言語学的調査も言語地図と関連づけることができるならば、新たな可能性が広がるように思われる。

### 3 エチオピアの言語

エチオピアの言語は、Grimes (2000) によれば、合計 85 の言語が数えられている。ただし、言語と方言の違いなどは明確ではなく、たとえば Grimes (2000) ではエチオピアのオモ系言語の中でカチャマーガンジュレ語 (Kachama-Ganjule) は 1 つの言語として記載されているけれども、私の調査によれば両者は別の言語と考えた方がいい<sup>4</sup>。また、その後の調査でその方言とされるガンタ語も別の言語と考えられる。このように Ethnologue に記載されている言語の数はあくまで概数であり、それほどはっきりとしたものではない。

ところで、言語には一応それぞれの系統関係というものが想定されており、日本語のように系統不明の言語もなかにはあるけれども、アフリカ大陸の場合アフロ・アジア、ナイロ・サハラ、ニジェル・コンゴ、コイサンという 4 つの大語族に還元されるというのが通説である。そのうちエチオピアにはアフロ・アジア語族のセム系、クシ系、オモ系、およびナイロ・サハラ語族の東スーダン系があるとされる。本研究プロジェクトでは、主にオモ系、クシ系の少数言語に焦点を当てて、研究者が記述研究を行っている。これらの少数言語はまだ十分に記述研究がされておらず、必ずしも系統関係が証明されているわけではない。特にオモ系の言語は基本的な文法項目に関して他のアフロ・アジア語族と著しく異なる特徴を有することも注目に値する。

参考までにメンバーが現在調査研究を行っている言語を列記しておく。

#### 研究代表者および研究分担者

乾 (山口大学) : バスケト (Basketo) 語、マロ語 (Melo)、ガンジュレ語 (Ganjule)、ガンタ語 (Ganta)

柘植 (金沢大学) : アリ語 (Aari) ガリラ方言

小脇 (熊本大学) : コンソ語 (Konso)

中野 (帝京平成大学) : ハムタンガ語 (Khamtanga)、アガウ語 (Agau)

#### 研究協力者

若狭 (東京外国語大学非常勤講師) : ウォライタ語 (Wolaytta)

高橋 (筑波大学大学院) : ハマル語 (Hamer)

---

<sup>4</sup>詳細は乾 (2004) 参照。

## 4 データベース構築

エチオピアの言語地図作成には、前述した世界言語地図などを参考に、こちらで用意したエチオピアの言語地図を山口菱洋システム (<http://www.yrs.co.jp/index.htm>) に依頼し、ベクトル形式で言語の範囲を1つ1つポリゴンという面で区切って地図上に作図した。言語を点ではなく面で表す方式にしたのは、言語特徴により色分けするとき、地理的分布が連続的に綺麗に表せるからである<sup>5</sup>。そしてメンバーがフィールドから持ち帰ったデータをその言語の地図上の面とリンクする形でデータベースを構築している。

なお、フィールド調査には小型で高性能な GPS(Global Positioning System) である GARMIN 社製の eTrexVistaC (英語版) も携帯し、メンバーそれぞれの調査地点で、標高、緯度、経度情報を測定して持ち帰り、パソコンの世界地図 (GARMIN 社製 WorldMap) にその地点を記録している。これは言語特徴以外の、たとえばデジタルビデオカメラで撮影した映像情報などをデータベース化する際、面よりも緯度・経度の正確な点で表した方が適切であると判断したからである。

現時点で構築されているデータベースは次のとおりである。

### エチオピアの言語全体に関するデータ

系統関係、話者数、基本語順。

### メンバーの調査言語に関するデータ

プロジェクトにおいてこれまで公表した論文 (PDF 形式)、語彙リストとその音声ファイル (WAV 形式)、映像情報。

## 5 言語地図作成の課題

言語地図作成にあたり、4点ほど課題を挙げておく。これらはまだ解決できていない問題であり、今後研究を進めていく過程でどう処理すべきか考えなければならぬであろう。

まず最初に、言語の境界線をどうするかという問題がある。多くの先行研究でも言語地図は面で区切られているものや点で表示されているものがある。確かに話者数が100人を切るような危機に瀕した言語の場合、周辺の言語の中に完全に埋没しており、地図上に面として捉えることは殆ど不可能である。言語共同体として機能しているかも疑問である。このような場合、どうしても点で捉えざるを得ないと思われる。

---

<sup>5</sup>その問題点については次章で論じる。

次に、現時点の言語の話されている地点を地図上に正確に表すのかがいかにどうかという問題がある。たとえば南北アメリカ大陸やオーストラリア大陸のように、長年その土地に住んでいた民族<sup>6</sup>がヨーロッパ列強の国々に居住地を奪われ、現在では民族や言語そのものが完全に消滅してしまったり、違った土地に強制移住させられて言語交替を余儀なくされて、自分たちの本来の言語を使用していないような状況に置かれている場合である。それらの地域を現在話されている英語やスペイン語やポルトガル語などにしてしまったら、世界言語の地理的分布はかなり偏った地図になってしまい、言語研究をする上で限られた目的でしか利用できなくなってしまうであろう。何らかの言語資料があり、かつ元々の居住地域がわかっている場合には、できる限り復元してその言語の地域を面で捉えることが大切であると思われる。

三番目に、言語の話されている地域をどう特定するかという問題である。これはそれぞれの地域の信頼できる先行研究を頼りにせざるを得ない面がある。前述の包括的な世界言語地図なども信頼できる先行研究をもとに作図されているはずであり、当面は最も信用できる資料といえよう。また実際に現地でフィールド調査を経験すれば自明のことであるけれども、一人の調査者がその言語の話されている地域を正確に把握することはほとんど至難の業である。特に個人あるいは集団は常に同じ所に留まっているわけではなく、インフォーマントとの出会いも必ずしもその言語の話されている中心地である必要もない。つまり対象とする言語の母語話者さえ見つければ、調査はどこでもできるという面がある。この点言語学が人類学や民族学と大きく異なる点である。したがって言語学の分野の研究成果だけで民族の居住地を面として捉えることは難しく、他分野の研究成果が不可欠といえよう。

最後に同じ地域での多言語使用や二言語使用をどう表すかという問題である。エチオピアでも都市部だけでなく、農村部でも場面によって日常生活においてエチオピアの公用語アムハラ語やオロモ語が使用されることも多い。これは前述のヨーロッパの言語と土着の言語との関係とは少し異なる問題である。長い年月をかけて複数の言語が接触し、お互いに影響を受けて変化をするというのは、言語の収束的変化の例であり、単に社会言語学的研究という面だけでなく、もっと重要な示唆を与えてくれる可能性がある。この点はできる限り言語地図に反映できるような表示を考えなければならない。これはGISのオーバーレイ機能を使えばおそらく解決できると思われる。

---

<sup>6</sup>原住民や先住民などと呼ばれている民族である。

【参照文献】

- Grimes, B.F. (ed.) 2000 *Ethnologue: Languages of the World. 14th edition.* Dallas: Summer Institute Linguistics.
- Haspelmath, M. & M.Dryer & D.Gil & B.Comrie (eds.) 2005 *The World Atlas of Language Structures. (Book with interactive CD-ROM)* Oxford: Oxford University Press.
- 乾 秀行 2004 「ガンジュレ人の言語使用」『一般言語学論叢』7, 筑波一般言語学研究会, 73-93.
- 松本克己 2000 「世界諸言語の類型地理と言語の遠い親族関係」『言語類型地理論シンポジウム論文集』（遠藤光暁編＜中国における言語地理と人文・自然地理＞7）96-135.
- Moseley, Ch. & R.E.Asher (eds.) 1994 *Atlas of the World's Languages.* New York: Routledge.