

令和元年5月28日現在

機関番号：11101

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K02653

研究課題名(和文) 地理情報システムによる世界諸言語の比較表現の言語類型地理論的研究

研究課題名(英文) Areal-Typological Study of Comparative Constructions of the World's Languages Based on Geographical Information System

研究代表者

山本 秀樹 (YAMAMOTO, Hideki)

弘前大学・人文社会科学部・教授

研究者番号：20200834

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：研究代表者は、2017年までに、地理情報システムによってデジタル化した世界言語地図を作製し、研究代表者による、それぞれ約3,000言語、約1,800言語のデータに基づく種々の語順特徴及び格標示体系の地理的分布を表した。本研究では、比較表現について約1,300言語からデータを収集し、世界諸言語の比較表現の地理的分布を同地図上に表現、分析し、比較表現についても、語順や格標示と同様、主要なパターンが地理的に大きなまとまりを成して分布する様態などを明らかにした。

研究成果の学術的意義や社会的意義

本研究で収集した比較表現データは約1,300言語に上り、世界でも群を抜いて最大規模のデータベースが構築されたことになり、今後の比較表現研究にとって、きわめて貴重な資料となる。また、比較表現の地理的分布についても、本研究のGISデジタル世界言語地図は、地理的に連続した状態で種々の表現を可能にするもので、語順、格標示に加えて、比較表現に関しても、世界全域にわたって微視的にも巨視的にも、より正確かつ詳細に地理的分布を分析することが可能となり、利用価値の高いものである。

研究成果の概要(英文)：By 2017, I made a digitalized version of maps on the world's languages by Geographical Information System, showing areal distribution of various word order features based on my data from about 3,000 and 1,800 languages of the world respectively. In the present study, I have collected data on comparative constructions from some 1,300 languages, and have shown and analyzed the areal distribution of comparative constructions of the world's languages, demonstrating that their major patterns occupy vast continuous areas on the globe as are the cases with word order and case marking.

研究分野：言語学

キーワード：言語学 言語類型論 言語類型地理論 地理情報システム 世界言語地図 世界諸言語 比較表現

1. 研究開始当初の背景

(1) 本研究課題を行うに至った背景

研究代表者は、平成 18~21 年度の科学研究費によって、GIS (地理情報システム) によるデジタル世界言語地図を活用する研究を開始した。同科学研究費および続く平成 22~24 年度の科学研究費による研究では、言語の諸特徴の地理的分布状況を分析し、従来の言語類型論や言語系統論を統合する新しい言語類型地理論的研究を可能にするための GIS デジタル世界言語地図の作製、開発、改良を行った。さらに、研究代表者が収集した約 3,000 言語の語順データを GIS デジタル世界言語地図のデータベースに組み入れて、種々の地図表現を行い、世界諸言語における様々な語順特徴の地理的分布を分析、考察してきた。さらに平成 25~28 年度の科学研究費による研究では、語順とともに人類言語の基本的文法構造における 2 つの核を成すとも言える重要な文法特徴である格標示体系について約 1,800 言語からのデータを収集し、それらに対して GIS デジタル世界言語地図上で種々の地図表現を行い、世界諸言語における様々な格標示パターの地理的分布を分析、考察してきた。

本研究課題では、これら平成 28 年度までの科学研究費による研究をさらに推進すべく、しばしば語順や格標示と相関を成す比較表現に関して、世界の多くの言語から比較表現データを収集し、それらのデータを GIS デジタル世界言語地図のデータベースに組み入れて、様々な地図表現を行い、世界全体にわたる比較表現パターの地理的分布を考察することにした。

(2) 先行研究

これまで比較表現について、世界諸言語にわたるデータを利用した研究は、国内外を問わず、せいぜい百数十言語のデータを基にしたものにすぎず、しかもそれらを地理的分布の観点から研究したものはほとんどなかった。ほとんど唯一のものとしては、Haspelmath et al.(2005)(いわゆる WALIS) に所収の Leon Stassen によるものがある。これは、比較表現のパターンを 4 パターンに分類し、世界の言語から 167 言語をとり上げて、WALIS の世界地図上に分布を表現している。しかし、そもそも WALIS は、世界諸言語の中から比較的少数の言語を取り出して、当該言語の使用地域の中心地を小さな点として表現させるものにすぎず、世界諸言語にわたって、言語特徴の正確な地理的分布を詳細に分析するには不十分なものであった。

本研究の GIS デジタル世界言語地図は、諸言語の言語特徴を地理的に連続した状態で表現し、それらの地理的分布を世界全域にわたって微視的および巨視的に観察することを可能にするものである。本研究において、世界の言語から可能な限り多数の比較表現データを抽出、収集し、それらを GIS デジタル世界言語地図のデータベースに組み込めば、世界全域にわたって、実際の連続的な分布を、正確かつ詳細に分析、考察することが可能になる。

2. 研究の目的

本研究の目的は、主として以下の 3 点にまとめられる。

- (1) これまであまり多数の言語について調査、データ収集がされたことのなかった比較表現に関して、可能な限り世界の多くの言語からデータを抽出、収集し、世界最大規模の比較表現のデータベースを構築する。
- (2) 収集した比較表現のデータを、以前の科学研究費で構築した地理情報システムによるデジタル世界言語地図用データベースに組み入れて、GIS デジタル世界言語地図上で、様々な地図表現、分析を可能にする。
- (3) GIS デジタル世界言語地図上に表現された世界諸言語における比較表現の地理的分布を、言語類型地理論的な観点から分析、考察する。

3. 研究の方法

(1) 比較表現データの種類とパターン認定

本研究では、比較表現のパターンに関しては Stassen(1985) (後の WALIS 所収のものも基本的には同様) に従い、比較基準の表現方法によって、主要なパターンとして 4 種類、下位パターンを含めて 7 種類に分類することにした。すなわち、主要なパターンとしては、比較基準を場所表現で表す「場所的比較(L)」、比較基準を他動詞の直接目的語として表す「上回る型比較(E)」、比較対象と比較基準を別々の節に分け、比較の意味を 2 つの等位節で表現する「等位接続型比較(C)」、比較基準を小辞で比較対象と同じ格関係で表す「小辞型比較(P)」である。さらに、これらのうち、場所的比較は、比較基準に用いる格関係に基づいて「離格型比較(LF)」、「向格型比較(LT)」、「所格型比較(LA)」、等位接続型比較は、「反意的表現を用いる等位接続型比較(CA)」と「否定的表現を用いる等位接続型比較(CN)」に分けられる。ただし、1 つの言語がこれらのうち複数のパターンを利用できる場合も少なくない。本研究では、可能な限り、各言語において最も基本的と考えられるパターンのみを認定することにした。

さらに、本研究では、義務的な比較特性標識(たとえば英語の-er や more)を用いる言語のデータ、および比較基準と比較詞(比較する特性を表す形容詞や副詞)との基本語順データも同時に収集した。この場合も各言語で最も基本的と考えられるものを認定した。

(2) 比較表現のデータ収集

本研究のように、GIS デジタル世界言語地図用データベースに諸言語の比較表現のデータを組み込み、世界全域にわたる地理的分布を分析するためには、まず、可能な限り多くの世界諸言語からデータを収集しておく必要がある。この種の研究においては、実質的にこのデータ収集作業が、最も時間と工夫と労力を要する部分となる。そこで、各個別言語の記述文法書を中心に数多くの文献から、可能な限り多くの言語についてデータを収集していった。

本研究において比較表現に関するデータを抽出、収集した言語約 1,300 言語(正確には 1,272 言語)に上る。この数は、これまで最も包括的なデータを扱った WALS 所収の Stassen による 167 言語と比べても、群を抜いて最大規模ではある。しかし、これまで、研究代表者が構築した語順に関する約 3,000 言語、格標示パターンに関する約 1,800 言語と比べると、かなり下回ることも事実である。これには、特に次のような要因が関係している。伝統的に、比較表現は、各個別言語において独立した特定の言語現象として扱われていないことが多く、本研究でデータを収集する過程においても、比較表現にはまったく言及がない記述書、文法書、研究文献に遭遇することが少なくなかった。また、上述したように、比較表現の場合、1 つの言語が複数のパターンで比較を表現できる場合も少なくなく、比較は表現しようと思えば、いずれのパターンでも表現できてしまうという側面がある。そのため、言語によっては、語順や格標示以上に最も基本的なパターンを見いだすことが困難で、結局、当該言語における最も基本的なパターンをデータとして抽出することができなかつた言語も少なくなかった。

(3) GIS デジタル世界言語地図用データベースへの組み込み

上述のように収集した比較表現データを、種々の情報を含む GIS デジタル世界言語地図用データベース上に組み込んでいった。その過程では、収集した言語データにおける言語名を、同データベース上の言語名と照合させなければならない。世界の言語は、しばしば異なった分類や様々な命名がなされているため、この作業も容易ではなく、多くの時間と労力を要する作業となったが、各個別言語の地理的位置や系統等を考慮しながら、慎重に照合を進めていった。

(4) GIS デジタル世界言語地図上の表現

本研究によって作成された比較表現のデータベースを基に、GIS デジタル世界言語地図上で分布を表現していった。この GIS デジタル世界言語地図では、各言語の使用地域がそれぞれポリゴンと呼ばれる様々な形の領域として区分されており、それらを塗りつぶす方式をとっている。

まず、C、E、L、P の主要 4 パターンについて、GIS デジタル世界言語地図上に、パターンごとに色分けしたポリゴンでそれらの分布を表現した。次に、C と L に関して、それぞれの下位パターン、すなわち CA と CN、LA と LF と LT の分布を同様に表現した。さらに C、E、L、P について、Spatial Analyst というプログラムを使って分布を表現した。このプログラムは、地図上でデータの入っていない部分のデータを予測して補完するもので、正確な分布を表現することには限界があるが、空白地域の言語特徴が推測された連続的な地図表現を可能にする。また、Spatial Analyst による地図を本来の分布地図と重ね合わせた地図も表現した。さらに、GIS によるデジタル世界言語地図の一つの特性を活かして、これらの分布を地勢と合わせて表現した。これらに加えて、義務的な比較特性標識を用いる言語、比較基準と比較詞との基本語順についても、GIS デジタル世界言語地図上に分布を表現した。これらの地図表現は、いずれも GIS デジタル世界言語地図上であれば、詳細な地域を自由に拡大して表示、観察することが可能である。

4. 研究成果

(1) 世界最大規模の比較表現データベースの構築

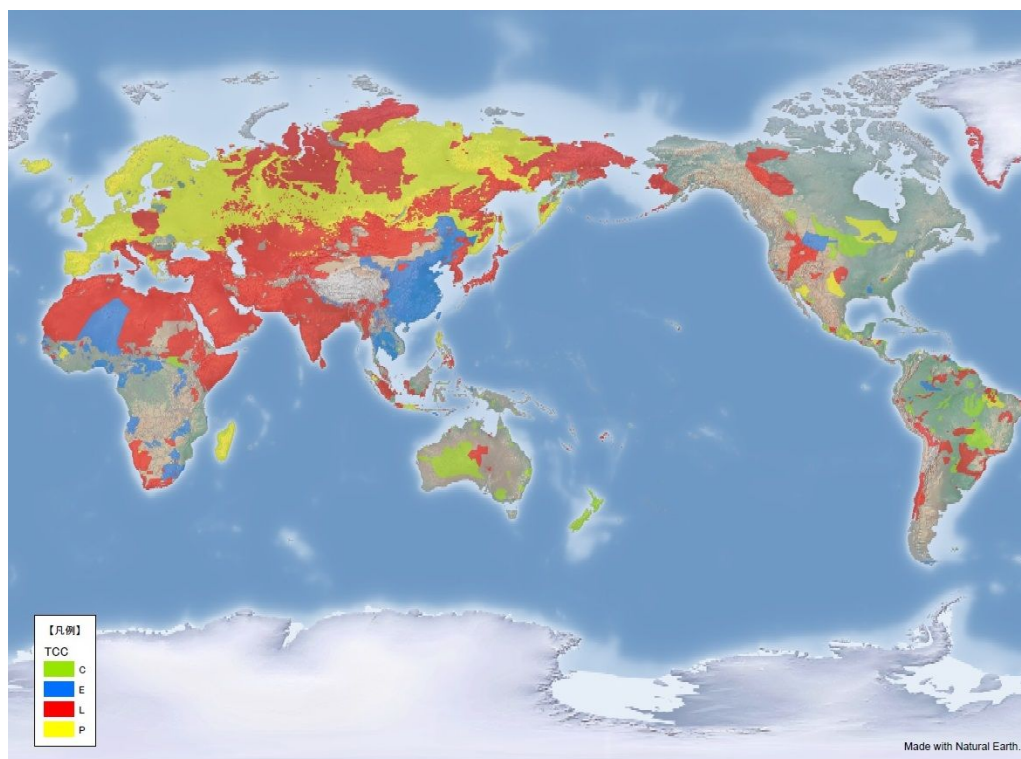
本研究において結果的に比較表現のデータを抽出し、GIS デジタル世界言語地図のデータベースに組み込むことのできた言語の数は 1,272 言語に上る。これは、上の「研究の方法」(2)において述べたような要因により、研究代表者のこれまでの語順や格標示のデータに比べると下回っているが、それまでの百数十言語に比べ、世界でも群を抜いて最大規模の比較表現データベースが構築されたことになる。この程度まで収集すれば、今後データをさらに拡大したとしても、統計的な分布についても、地理的な分布についても、おそらくそれほど大きな違いは生じず、かなり正しい分布が導き出されたと考えられる。実際、これらのデータを基にした今回の GIS デジタル世界言語地図においても、それほど重大な空白箇所を残さない形で分布が表現された。

(2) GIS デジタル世界言語地図上に現れた地理的分布の分析

本研究では、収集した比較表現データを GIS デジタル世界言語地図のデータベースに組み込み、「研究の方法」(4)で述べたような種々の地図表現を行った。そして、それぞれの GIS デジタル世界言語地図上に現れた比較表現の地理的な分布を分析、考察して、その成果を著作にまとめることができた。個々の分布に関する分析は詳細にわたるので、本報告書では、比較表現

の主要 4 パタンの大まかな地理的分布の現れに関して、その概要を簡潔に報告するとどめる。

下は、本研究で GIS デジタル世界言語地図上に表現した地図の一例で、C、E、L、P の主要 4 パタンの地理的分布を地勢と合わせて表現したものである。左下の凡例は、上から、「等位接続型比較(C)」、「上回る型比較(E)」、「場所的比較(L)」、「小辞型比較(P)」を表している。



地図上に現れたこの分布を全体的に見て最初に気づくこととしては、比較表現パタンの地理的分布は、語順や格標示と同様、しばしば同一パターンが連続して現れやすい傾向が見られるということである。例外的に比較表現パタンの複雑な分布が見られる地域として特に目を引くのは、アメリカ大陸である。アメリカ大陸についても、世界的に最も優勢な場所的比較が最も広く多く分布していることに加えて、オーストラリアやニューギニアと同様に等位接続型比較もまた比較的広く分布しているといった、ある程度の分布傾向は見えてとれる。アメリカ大陸は、多くの言語特徴に関してしばしば最もバリエーションに富んだ分布を示すが、比較表現パターンに関して、やはり同様な複雑な分布を示している様子が見えてとれる。

以下、場所的比較、上回る型比較、等位接続型比較、小辞型比較の主要 4 パタンについて、GIS デジタル世界言語地図上に現れた地理的分布を見ていく。まず、研究代表者のデータで約半数を占めるほど最も優勢であった場所的比較であるが、これは、地理的にも地球上で最も広範な地域に分布している。ただし、この地図上では、いわゆるアジア・ロシアの地域において小辞型比較が広く分布して見えるが、これは、本研究のデジタル地図が基にした Asher and Moseley(1994 & 2007)の地図において、この地域に関してはかなり広域にわたってロシア語を含めた分布図になっているためである。そこで、これらの地域からロシア語の小辞型比較による分断を捨象して考えれば、やはりアジア北部にも広大な場所的比較の連続体が形成されていると見ることが出来る。そのようにしてみると、場所的比較は、ヨーロッパおよび東アジアから東南アジアにかけての地域を除いたユーラシア大陸の全域に広く分布している様子が見えてとれる。そのほか、サハラ以北のアフリカ大陸、オーストロネシア諸語地域、また、前述したアメリカ大陸においても、おそらく最も広い分布を持つタイプと考えることができるだろう。次に、研究代表者のデータでは約 25%と、場所的比較に次いで優勢であった上回る型比較は、主に 2 つないし 3 つの地域に集中して分布している。特に明確にまとまった分布を示すのはアフリカ大陸と東アジアから東南アジアにかけての地域である。アフリカ大陸においては、北部のベルベル諸語を含んで西アフリカからサハラ以南の大半が、上回る型比較によって占められている。また、東アジアでは、日本列島と朝鮮半島を除く大半がこのパタンの比較によって占められ、その分布は一部ヒマラヤ地域にも及んでいる。そのほか、アメリカ大陸の分布は複雑であるが、北米中部や南米北部の一部において、このパターンがある程度のまとまりを示す可能性が考えられる。等位接続型比較は、特にオーストラリアとニューギニアで顕著である。また、データが十分でないため、地図にもあまり明確には現れていないが、ニューギニア島においてもこのパターンが顕著である可能性が考えられる。そのほか、先に述べたとおり、アメリカ大陸においても、等位接続型比較が世界のほかの地域に比べて広く分布する傾向が見られ、特に北米中部、北米の西海岸から中米にかけての地域、南米北部の大半の地域において、まとまった分布を示すように見える。小辞型比較については、Stassen(1985)が指摘したように、大部分がヨー

ロッパに集中している様子が見てとれるが、そのほかでは、マダガスカル島、フィリピン諸島、スマトラ島北部、ジャワ島、また、アメリカ大陸にも散在した分布が見られる。

(3) 得られた成果の国内外における位置づけとインパクト

国内外を問わず、従来、比較表現パターンに関しては、せいぜい百数十言語のデータを基にした研究しか存在していなかった。それに対して、本研究で収集した比較表現データは約 1,300 言語に上回る。これは、世界でも群を抜いて最大規模の比較表現データベースが構築されたということになり、今後の比較表現に関する研究にとって、きわめて貴重な資料となると考えられる。

また、比較表現の地理的分布についても、従来は *WALS* に収められているように、百数十言語について、小さな点として表現させた程度のもので存在しなかった。それに対して、本研究の GIS デジタル世界言語地図では、約 1,300 言語の規模で、比較表現パターンを地理的に連続した状態で表現するものである。それによって、今後は、世界全域にわたって、微視的にも巨視的にも、より正確かつ詳細に比較表現の地理的分布を分析することが可能となる。

(4) 今後の展望

本研究によって収集された比較表現のデータは、すべて、それまでに作成した、言語名（含別名）、系統、地域、位置情報、約 3,000 言語に関する様々な種類の語順特徴や約 1,800 言語の格標示に関する情報を含んだデータベースに統合されている。これらの情報は、データベースおよび GIS デジタル世界言語地図において、すべて連動している。本研究によって、GIS デジタル世界言語地図を充実させ、人類言語の基本的文法構造の両輪を成す語順と格標示に加えて比較表現についても、迅速かつ詳細に地理的な表現、検索、分析を行うことが可能になった。

今後は、さらに多くの重要な特徴に関する言語情報を収集し、データベースに組み込むことによって、GIS デジタル世界言語地図をいっそう充実させ、それらのデータベースと GIS デジタル世界言語地図を利用した研究をさらに進展させていく計画である。

< 引用文献 >

Asher, R. E. and Christopher Moseley, eds. 2007 *Atlas of the World's Languages*. 2nd ed. (1st ed., 1994) London: Routledge.

Haspelmath, Martin, Matthew S. Dryer, David Gil and Bernard Comrie, eds. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

5 . 主な発表論文等

〔図書〕(計 1 件)

山本 秀樹、(株)エイム、『地理情報システムによる世界諸言語の比較表現の言語類型地理論的研究』、2019 年、103 頁

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。