

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 30 日現在

機関番号：11101

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25370456

研究課題名(和文) 地理情報システムによる世界諸言語の格標示体系の言語類型地理論的研究

研究課題名(英文) Areal-Typological Study of Case Marking System 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) 3,500,000円

研究成果の概要(和文)：研究代表者等は、2013年までに、GIS(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、研究代表者による約3,000言語のデータに基づく種々の語順特徴の地理的分布を表した。本研究では、格標示体系について、1,800を超える言語からデータを収集し、データ上の言語名をデジタル地図上の言語名に照合させつつ、それらのデータを同地図に組み入れていった。そして、世界諸言語の格標示体系の地理的分布を同地図上に表現、分析し、格標示体系についても、語順と同様、特に主要な格標示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。

研究成果の概要(英文)：By 2013, we made a digitalized version of maps on the world's languages by GIS (Geographical Information System), showing areal distribution of various features of word order based on my data from about 3,000 languages of the world. In the present study, I have collected data on case marking system from more than 1,800 languages. I have incorporated the data into the digitalized maps by matching the language names of my data against those on the digitalized maps. I have shown and analyzed the areal distribution of case marking system of the world's languages, demonstrating that major case marking patterns, in particular, occupy vast continuous areas on the globe as is the case with word order.

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 格標示

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

研究代表者等は、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した研究を開始した。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。

本研究は、上述した平成 24 年度までの科学研究費による研究をさらに推進すべく、研究代表者個人が申請した継続的な研究である。格標示は、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。そこで、本研究課題では、格標示に関して、世界の多くの言語から格標示データを収集し、それらに対してデジタル世界言語地図上で様々な地図表現を行い、世界全体にわたる格標示パタンの地理的分布を考察することにした。

(2) 先行研究

これまで格標示パターンについて、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、せいぜい数百言語のデータを基にしたものにすぎなかった。主なものとしては、Haspelmath et al. (2005) の *The World Atlas of Language Structures*. (通称 WALS) (Oxford UP) に収められているものがある。同書では、たとえば代名詞について 172 言語、完全名詞句について 190 言語、動詞について 380 言語が扱われている。

上述の WALS は、世界地図上に種々の言語特徴の分布を表現したものである。しかし、これは、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格標示データを抽出、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) しばしば語順と相補的な関係にあり、語順と並んで人類言語の根幹に関わる重要な文法特徴である格標示に関して、可能な限り世界の多くの言語から格標示データを抽出、収集し、世界最大規模の格標示データベースを構築する。

(2) 収集した格標示データを、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格標示パタンの地理的分布を、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) 格標示データの種類の種類と格標示パタンの認定

本研究では、1 項述語において自動詞主語に相当する S、2 項述語において、より動作主的で概ね他動詞主語に相当する A、より被動者の概ね他動詞目的語に相当する P をどのように表現するかに基づいて、格標示パターンを分類する。

S、A、P について論理的に可能な主な格標示パターンとしては、S と A が同一で P が異なる対格型、S と P が同一で A が異なる能格型、これら 3 項すべてを同一形にする中立型、3 項すべてを区別する 3 項型、A と P が同一で S が異なる二重斜格型がある。さらに、S の持つ意味的な特性に基づいて、より動作主的な S を A と同様に、より被動者の S を P と同様に表す動格型や、有生性や談話機能による階層にしたがって、A が P よりも階層上で上位にあれば順行形を、A が P よりも階層上で下位にあれば逆行形をとるような階層型もある。

そして、それらの格標示パターンが現れる項目については、基本的には WALS と同様、代名詞の格標示パターン、名詞の格標示パターン、動詞上に現れる格標示パターンに分けてデータを収集することにした。各言語に対し、これらの各項目について、対格型、能格型等の格標示パターンを認定していくことになる。

しかし、各言語に現れた格標示に対して、どのようなパタンのデータとして認定するかは、実際には、各研究者がとる基本的な認定方針によって異なってくる場合も少なくない。そこで、本研究では、WALS 等のデータをそのまま利用するのではなく、まず GIS デジタル世界言語地図上に簡潔、明確、効果的にデータを表現するための、各言語に対する格標示パタンの認定方針を自身で定めるこ

とにした。その上で、記述文法書等の情報を基に、各個別言語の格標示パターンに関する自身の認定方針にしたがって、格標示パタンのデータを収集していくことにした。

また、同一言語内の同一項目において複数の格標示パターンが分裂ないし共存する場合についても、どの格標示パタンのデータとして認定するか、あるいは、どのような分裂の場合に分裂パタンのデータとして認定するか等についても、自身の基本方針を定めた。さらに、分裂を引き起こす要因に関して、有生性の階層による分裂、定性等による分裂、時制や相による分裂に分けて、地図化するデータに含めることにした。

(2) 格標示パタンのデータ収集

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格標示パタン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格標示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、自身の立てた認定方針に基づいて、可能な限り多くの言語について格標示データを収集していった。また、データ収集の過程で、数度にわたって認定方針やデータの見直し、改良を行った。

(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格標示データを、24年度までに構築した、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

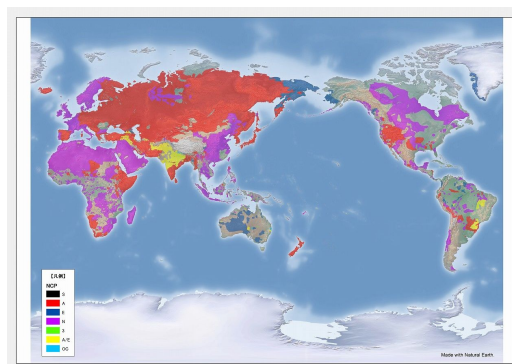
本研究によって作成された格標示のデータベースを基に、GIS デジタル世界言語地図上で分布を表現していった。

本研究では、代名詞の格標示パターン(PCP)、名詞の格標示パターン(NCP)、動詞上に現れる格標示パターン(VCP)、有生性の階層による分裂(AS)、定性等による分裂(DS)、時制や相による分裂(TAS)についてデータベースを作成している。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、PCP、NCP、VCP、AS、DS、TASのそれぞれについて、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、PCP、NCP、VCPについて、それぞれ 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、PCP、NCP、VCP、AS、DS、TASの分布を、それぞれ地勢と合わせて表現した。

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。本研究で表現した地図の一例をあげると、下は、名詞の格標示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、動格型、対格型、能格型、中立型、3項型、対格型と能格型との共存(分裂)型、その他の共存型を表している。



4. 研究成果

(1) 世界最大規模の格標示データベースの構築

本研究において結果的に格標示パタンのデータを抽出し、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1,822に上る。これは、世界でも最大規模の格標示データベースとなる。世界の言語全体から見れば、言語数としては3割程度と思われるが、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、収集した格標示データを 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格標示パタン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、格標示パタンの全体的な地理的分布の現れに関する簡潔な報告にとどめる。

格標示パタンの地理的分布は、少なくとも主要なパターンに関する限り、語順と同様、むしろそれ以上に、同一パターンが広大な範囲に連続して現れやすい。これは、一つには、基本的に形態法に属する格標示の方が、統語法に属する語順に比べて、より歴史的な変化を被りにくいことが地理的分布に反映されていると考えられるだろう。

例外的に格標示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、PCP や NCP についてオーストラリア、VCP について北米、PCP、NCP、VCP のすべてについて南米である。

オーストラリアは、名詞句における有生性による格標示パタンの分裂が顕著な地域であり、そのことが、特に代名詞や名詞の格標示パターンにおける多様な地理的分布に反映されているのであろう。北米には、高度な多総合性により動詞上に S、A、P の関係を表す言語が多く、それが、VCP における複雑な地理的分布を生み出しているものと思われる。また、南米も、北米ほど多総合性の度合いが高くないとしても、やはり動詞上で関係を表す言語が多い。さらに、南米は、節語順等についても多数の異なるタイプの言語が複雑に分布しており、そのことが、PCP、NCP、VCP すべての格標示パターンにおける複雑な地理的分布を生み出しているものと考えられる。

(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、格標示パターンに関しては、せいぜい数百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、本研究で収集した格標示データは 1,800 言語を上回る。これは、世界でも群を抜いて最大規模の格標示データベースが構築されたということになり、今後の格標示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、格標示パタン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、数百言語について、小さな点として表現させた程度のものしか存在しなかった。それに対して、本研究の GIS デジタル世界言語地図では、1,800 言語を上回る規模で、格標示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

かつ詳細に格標示パタン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 1,800 を上回る言語に関する格標示のデータは、すべて、24 年度までに作成した、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等の情報を含んだデータベースに統合されている。したがって、それらの情報は、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において 2 つの核を成す語順と格標示の両面にわたって、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。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

山本 秀樹、「地理及親縁関係視閥下的語序類型研究」、『日語学習與研究』、174 巻、7-12 頁、2014 年、査読有

[図書](計 1 件)

山本 秀樹、エイム、『地理情報システムによる世界諸言語の格標示体系の言語類型地理論的研究』、2017 年、104 頁

6. 研究組織

(1) 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