

# 語順地図作成に必要なデータ 及び語順地図に現れる語順分布

呉靱<sup>†</sup> 山本秀樹<sup>††</sup> 乾秀行<sup>†††</sup> 杉井学<sup>‡</sup> 松野浩嗣<sup>†</sup>

キーワード：GIS、語順類型論、語順地図、語順類型

## 1 はじめに

従来の類型論研究では、真の言語普遍性を見出す目的のために、地理的および系統的な要因は捨象され、主に言語内部の要因に着目して、言語普遍性や言語特徴の間に成り立つ含意ないし相関関係等を見出すことを目標として探求されてきた。しかし、近年、山本 (2003)の語順研究のように、地理的分布や歴史的ないし系統的分布を考慮した研究も次第に行われるようになり、それらの要因が言語現象に大きく関係していることが明らかになりつつある。

世界全域にわたり、種々の言語現象について言語地図を用いて地理的分布を考察する研究はまだ始まったばかりであるが、現在、山本等 (研究代表者：山本秀樹、研究分担者：乾秀行、研究協力者：松本克己)は、「地理情報システムによる世界諸言語の言語類型地理論的研究」(平成 18～21 年度科学研究費補助金)で、GIS (地理情報システム)によって利用可能なベクトル世界言語地図を作成し、種々の言語特徴データと連動させる計画を遂行中である。また、池田等 (研究代表者：池田潤、研究分担者：竹内茂夫、

---

<sup>†</sup> 山口大学大学院理工学研究科

<sup>††</sup> 弘前大学人文学部

<sup>†††</sup> 山口大学人文学部

<sup>‡</sup> 山口大学メディア基盤センター

乾秀行) は、「前 2-1 千年期における北西セム語の等語線の再確定：GIS による言語地理学的研究」(平成 18～21 年度科学研究費補助金)で、GIS により、前 2 千年期と前 1 千年期の北西セム語の GIS 等語線を重ね合わせて、その通時的変遷を研究している。

これらの研究との関連で、山本 (2006) および池田 (2006) は、WALS (The World Atlas of Language Structures) 等の言語特徴の地図化に利用できるツールに関する調査を行い、それらのツールに内在している問題点を指摘している。また、同時に GIS を利用するメリットを分析している。GIS の機能は多方面にわたるが、地図作成はその最も基本的な機能である。我々は GIS を用いた世界諸言語の言語特徴の地図化から出発し、言語類型論研究における GIS の応用の可能性を探ってみた。

言語類型論全体の中でも、語順類型論は、特に Greenberg (1963) の研究によって近年の類型論研究が活性化する大きな契機となった領域であり、すでに山本 (2003) によって多くの語順データが収集されている。そこで、我々は言語特徴として、まず語順に焦点をあてることにした。

本稿では、第 2 節において、まず語順地図の作成に必要なデータを述べた後、我々が行ったデータ調査の結果を報告する。第 3 節において、そのデータを改編・加工し、生成できた新たなデータによって作成した語順タイプの地図化の例を示す。

## 2 語順地図作成に必要なデータ

### 2.1 語順地図を作成ための必要なデータ

GIS を用いて語順地図を作成するためには、図 1 に示すように、世界諸言語の語順パラメータに関する属性データと言語の位置情報 (その言語が話されている地域の地理情報) を示す空間データが必要である。

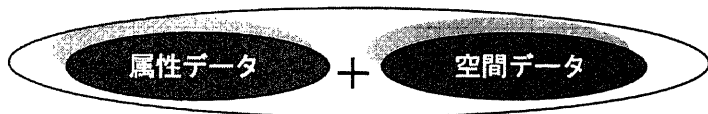


図 1：語順地図作成に必要なデータ

山本 (2003)は世界諸言語を対象として、約 3,000 の言語について、節語順、接置詞と名詞の順序、所有者と主要名詞の順序、形容詞と名詞の順序などの語順パラメータをデータ (以下「山本データ」と称する)として公表している。我々は山本データに基づいて世界諸言語の語順地図の作成を試みた。

空間データについては、池田 (2006)が、*Ethnologue* 第 15 版 (Gordon 2005) の GIS 向けデジタルデータパッケージ WLMS (The World Language Mapping System) <sup>2</sup> が一般発売されていることに言及している。そこで、我々は WLMS を入手し、その内容および語順地図作成に応用できる可能性について調査した。

WLMS には *Ethnologue* 第 15 版に掲載されている 7,299 言語の言語名や言語が話されている国、また言語の別名や方言など多数の情報が含まれている上、*Ethnologue* 第 15 版言語地図の作成に使ったと思われる空間データも含まれている。我々の管見の及ぶ限りにおいて、世界諸言語の GIS データが一般発売されているのはこのほかに例がなく、この既存のデータを利用できれば、メリットは大きい。

WLMS の内容については以下で紹介するが、この WLMS のデータは、そのままでは語順地図の作成には利用できない。我々はその問題点を指摘し、語順地図作成のための GIS データの生成方法を提案した (呉等 2007)。本稿では、呉等 (2007)で得られた GIS データを用いて、山本データに基づく語順地図の作成例を示し、GIS の言語類型論研究への応用例を示す。

## 2.2 WLMS の入手方法

GIS 向けのデジタルデータパッケージ WLMS の発売元 GMI (Global Mapping International)は SIL インターナショナル (Summer Institute of Linguistic International)<sup>3</sup>の関連団体である。GMI と SIL インターナショナルが 17 年間に渡って協力した成果として、1995 年に GMMS (*the Global Ministry*

<sup>2</sup><http://www.gmi.org/wlms/>

<sup>3</sup>非営利のキリスト教信仰に基づくアカデミックな組織で、*Ethnologue* および Web サイト (<http://www.ethnologue.com/web.asp>)を通して言語研究資源を提供している。

Mapping System)の一部として WLMS が発売された。その後も *Ethnologue* の改訂に合わせてアップデートされてきた。

WLMS は、WLMS 単体での購入<sup>4</sup>、または Seamless DCW Base Map (Seamless Digital Chart of the World Base Map)<sup>5</sup> とのセットか Seamless DCW Base Map の中の一部のデータとのセットでの購入が可能である。WLMS には、上述したように言語に関する属性データおよび言語の空間データが含まれ、Seamless DCW Base Map には大陸や海洋、国境などの基本的な地理情報の空間データが含まれている。我々が 2007 年 3 月時点においてセットで購入した際の最新バージョンは、WLMS が 3.2.1、Seamless DCW Base Map が 3.2 となっていた。

## 2.3 WLMS のデータ構成

WLMS 及び Seamless DCW Base Map のパッケージ (以下まとめて WLMS という)は CD-ROM で提供され、ライセンスは 1 パッケージが 1 台の PC に付与される (CD-ROM ジャケットにライセンスが付与されている組織の名称、シリアル番号、ロック解除キーが印字されている)。

WLMS Version 3.2.1 は ESRI 社<sup>6</sup>の shapefile 形式<sup>7</sup>のファイルで提供され、ArcMap 9.1<sup>8</sup>に対応したものである。製品には Windows 対応のインストーラーがバンドルされていて、それを実行すればデータが PC (デフォルトではシステムドライブ) に格納される。インストール後のデータ構成を図 2 に示す (WLMS と Seamless DCW Base Map は別製品のため、図 2 ではそれぞれのデータが識別できるように示している)。

<sup>4</sup>発売元に直接コンタクトして購入することも可能のようだが、筆者らは ESRI ジャパン株式会社社の国内正規代理店を通して購入した。

<sup>5</sup>Seamless DCW Base Map は、GMI がアメリカ政府の提供している DCW (The Digital Chart of the World)を編集・加工し、ESRI 社 shapefile 形式に変換し、パッケージ化したものである。

<sup>6</sup>商用 GIS ソフトウェア製品 ArcGIS の発売元である (Web サイト <http://www.esri.com/>)。

<sup>7</sup>図形と属性の両要素をもった地図データファイル形式で、米国 ESRI 社が提唱するベクトルデータの業界標準フォーマットを指す。shapefile は仕様が公開されていて、多くの GIS ソフトウェアで利用可能である。また、shapefile 形式のファイルは複数のファイルがセットになって構成されており、その中の基本的なファイルの拡張子は次のようになる。 .shp : 図形情報ファイル、.dbf : 属性情報ファイル、.shx : インディックスファイル。

<sup>8</sup>ArcGIS 製品群の中の 1 つで、地図の作成や編集を行う GIS ソフトウェア。

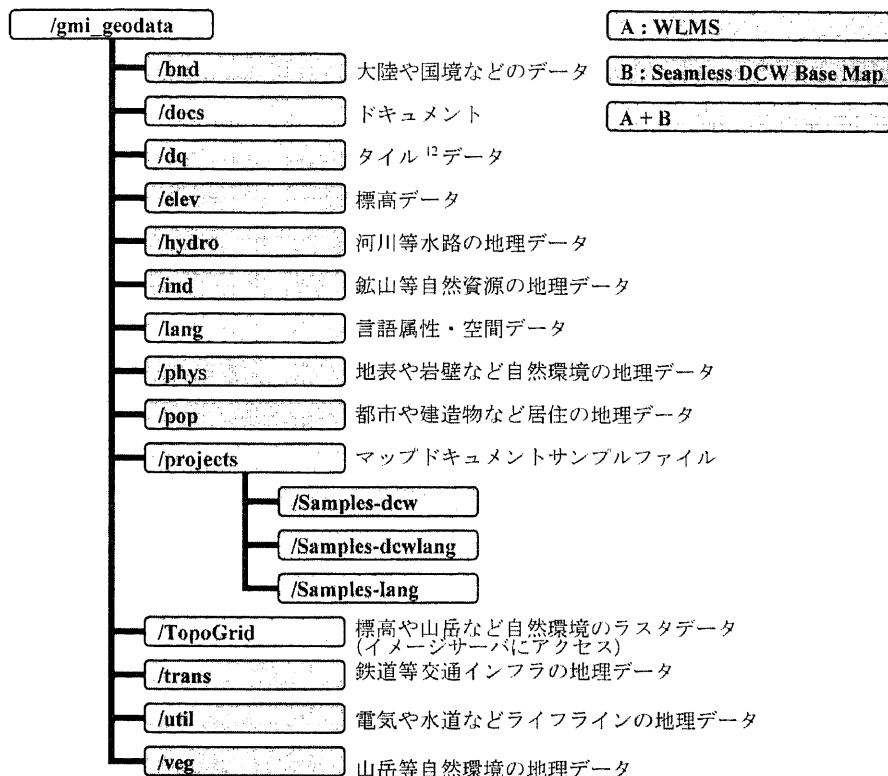


図 2:WLMS のデータ構成

## 2.4 WLMS 空間データの内容

フォルダ「c:/gmi\_geodata/lang/」に言語の地理情報に関する空間データが格納されている。空間データはポリゴンフィーチャー (面の図形) またはポイントフィーチャー (点の図形) の 2 種類があり、shapefile 形式のファイルで提供されている。表 1 に提供されている shapefile ファイルの一覧を示す。

<sup>12</sup>タイル (tile) とは、マップ・ライブラリに格納するための空間的な単位で、マップ・ライブラリがカバーする地域を分割し、位置によってライブラリ・データを組織化する (例として、県は、国全体のデータベースのタイルとして使用できる) (<http://www.esri.com/>により)。

表 1 : WLMS 空間データの内容 (shapefile ファイル一覧)

ファイル名 (shapefile 形式)	フィーチャー			
	形状	数	位置情報	クラス
Langa	ポリゴン	7,612	世界諸言語	L
mixlanga		869	多言語併存地域	L
Rother		12	<i>Ethnologue</i> に記述のない方言	L
lang_all_pt.pt	ポイント	9,239	すべてのポイントフィーチャー	すべて
lang_points_pt		7,612	世界諸言語	L
lang_extinct_pt		376	絶滅言語 (現在はもはや話されていない言語)	X
lang_widesprd_pt		313	広範囲言語 (例として標準中国語)	W
lang_widesprd_offset		313	上記広範囲言語につき表示上位置をずらしたものの	W
lang_other_pt		226	その他の言語	未定義
lang_only_pt		277	ポリゴンフィーチャーとして構成する位置情報が不足している言語	A
lang_unkwn_loc_pt		489	位置情報が未知な言語	U

表 2 : WLMS 空間データのクラス定義

クラス名表記	定 義
A	ポリゴンフィーチャーを持たないポイントフィーチャーを指す
L	ポリゴンフィーチャーを指す。クラス L のポイントフィーチャーはポリゴンフィーチャーの中心点に打点
U	位置情報が定まらないフィーチャーを指す。クラス U のポイントフィーチャーの位置は <i>Ethnologue</i> によって打点、または任意に打点
W	広範囲言語のフィーチャーを指す
X	絶滅言語のフィーチャーを指す
未定義 (Blank or Null)	未分類言語 (普通は <i>Ethnologue</i> に記述されていない移民言語) のフィーチャーを指す

地図作成・編集 GIS ソフトウェア (例として ArcMap) を使ってマップドキュメント (.mxd) に shapefile ファイルを追加すれば、1 つの shapefile ファイルは 1 つのレイヤを構成し<sup>10</sup>、多数のフィーチャーが表示される。

<sup>10</sup> 1 つの shapefile 形式のファイルは 1 ベクトルレイヤを構成する。shapefile 形式のファイルに格納可能な図形はポイント (点)、ライン (線)、ポリゴン (面) の 3 種類であるが、1 つの shapefile 形式のファイルには 1 種類の図形のみ格納可能である。

例えば、表 1 にある *langa.shp* には 7,612 のポリゴンフィーチャーが含まれている。また、すべてのフィーチャーはその空間データ生成に関しクラスが定義されている。表 2 に、そのクラスの一覧を示す。

## 2.5 WLMS 属性データの内容

WLMS は空間データのほか、*Ethnologue* 第 15 版に掲載されている言語の情報も提供されている。フォルダ「*c:/gmi\_geodata/lang/*」の中の *eth\_wlms.dbf* というファイルに約 65 項目（項目名や内容などについては *Ethnologue* 第 15 版を参照されたい）の言語属性情報が包括的に格納されている。WLMS 属性データにおいても、言語の節語順（本稿第 3 章を参照）という語順パラメータが含まれている。本研究では山本（2003）の研究結果（山本データ）に準拠する<sup>11</sup>。

WLMS 空間データによる世界言語地図

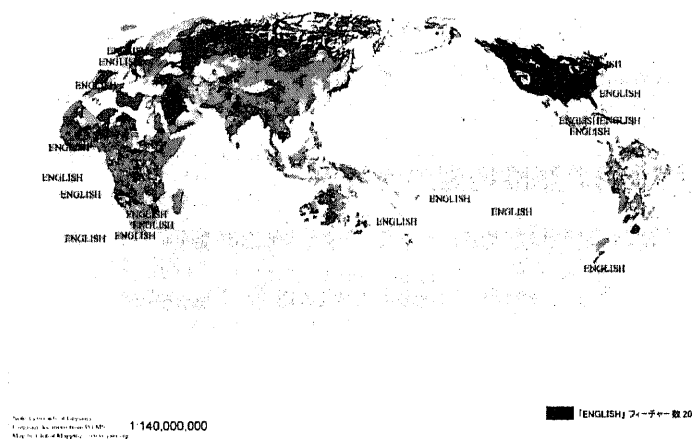


図 3： WLMS 空間データによる世界言語地図

<sup>11</sup>周知のように、語順類型論研究では、節語順だけではなく、接置詞と名詞の配列順序や所有者と名詞の配列順序など他の語順特徴も、言語普遍性を探求する重要な要素として取り上げられている。WLMS 属性データには、一部の言語について節語順に関するデータが含まれている場合があるが、一般にその他の語順特徴データは含まれていない。

また、WLMS Version3.2.1 の属性データは他の形式のデータから変換されたため、eth\_wlms.dbf という dBASE 形式<sup>12</sup>のファイルのフィールド長の制限により、項目によっては文字列が一部区切られたような不完全なものや、unicode<sup>13</sup>に対応していないため言語名の文字列に文字化けが発生している箇所が見られる。

## 2.6 WLMS データを用いた言語地図の作成例

WLMS データを基に、GIS ソフトウェア ArcMap9.1 を用いて作成した言語地図の例を図 3 に示す。

図 3 は前述の langa.shp のデータを基に作成した世界諸言語の地理的分布図となる。この図は、ここでは印刷の都合で白黒になっているが、実際には多色刷りで表示され、地球上で話される約 7,000 の言語の位置が、その系統グループごとに表されている。

なお、フォルダ「c:/gmi\_geodata/projects/」の中には、WLMS データ、Seamless DCW Base Map データ、WLMS データと Seamless DCW Base Map データを基にした地図サンプルファイル (拡張子が「.mxd」となるマップドキュメント) がそれぞれ「Samples-lang」、「Samples-dcw」、「Samples-dcwleng」のサブフォルダに格納されている。

## 3 世界諸言語の語順地図の作成例

### 3.1 言語類型論研究のための GIS データ生成の必要性

図 3 の世界諸言語の地理的分布図は WLMS 及び Seamless DCW Base Map のデータを用いて作成した言語地図の一例である。図 3 では、例えば北米等の地域において言語「ENGLISH」が話されている位置情報が示されているが、それは言語類型論的な観点からは問題があると思われる。特に大航

<sup>12</sup>1980 年代が全盛期だったデータベースソフト dBASE (dBASE II、dBASE III、dBASE IV) がサポートするファイル形式である。Microsoft Excel のワークブックを dBASE 形式で保存することが可能である (Excel2007 から dBASE II での保存はサポートされなくなった)。

<sup>13</sup>コンピュータ上で多言語の文字を単一の文字コードで取り扱うために提唱された文字コード体系である。



海時代以降、言語接触や通時的変化以外の原因によって急速に西洋語が拡張した結果、北米で話されている「ENGLISH」等の地理情報は、言語類型論的観点からみれば不要である。つまり、WLMS の空間データをそのまま語順類型研究のための言語地図の作成には利用できないということである。

しかしながら、我々は、たとえ WLMS 空間データを語順類型研究の言語地図作成に直接利用することはできないとしても、例えば北米において「ENGLISH」のデータが反映されないようにする等、これに必要な改編・加工を施せば必要な GIS データを作り出すことが可能であると考えた。

また、山本データに基づいて WLMS 空間データを語順地図の作成に利用するには、山本データと WLMS の属性データとの間で言語の同定が必要になる。言語名の照合によって言語を同定することになるが、言語名は、言語学者によって同じ言語が異なる言語名で呼ばれていたり、あるいは異なる言語が同じ言語名を持つこともしばしばある。つまり、同じ言語でも山本データの言語名と WLMS 属性データのそれとは必ずしも一致しない。

我々は、上記の問題点を踏まえた上で、山本データに基づく語順地図の作成に着目して、WLMS 空間データを語順類型論研究に利用する上での間

表 3: 語順パラメータ

語順 パラメータ	タイプ	意 味
節語順	SOV SVO VSO VOS OVS OSV	他動詞構文の 3 つの構成素である主語(S)、目的語(O)、動詞(V)の配列順序
接置詞と名詞 の順序	PR	前置詞(接置詞が名詞の前に置かれる)
	PO	後置詞(接置詞が名詞の後に置かれる)
所有者と主要 名詞の順序	GN	所有者(G)―主要名詞(N)の順序
	NG	主要名詞(N)―所有者(G)の順序
形容詞と名詞 の順序	AN	修飾形容詞(A)―被修飾名詞(N)の順序
	NA	被修飾名詞(N)―修飾形容詞(A)の順序

題解決の手法を提案した (呉等 2007)。その手法は語順に関する言語特徴だけでなく、他の言語特徴に対しても応用できる。次の 3.2 において、呉等 (2007) で生成したデータを基にした世界諸言語の語順地図の作成例を示す。

## 3.2 語順地図の作成例および語順類型

### 3.2.1 節語順タイプの地理的分布と SO 語順の普遍性

語順類型論研究において、従来から表 3 に示す 4 つの語順パラメータの相対順序に関心が向けられてきた。その中でも特に第 1 のパラメータの「節語順」という節の主要構成素 (主語 S、目的語 O、動詞 V) の配列順序は語順類型論の最も重要なパラメータの 1 つとして取り上げられてきた。

統計的手法は語順類型論研究において従来から用いられてきた有用な分析方法の 1 つであると考えられる。そこで、まず、本稿において言語地図の作成に用いられた呉等のデータ (呉等 2007) を山本データ (山本 2003) および松本データ (松本 2006) と対照させた形で、6 つの節語順タイプにつき、それぞれのデータにおける言語数および出現率を表 4 に示す。

なお、どんな言語も語順が厳格に決まっているわけではなく、山本データでは大文字同士の共存 (例えば「SOV/SVO」は SOV と SVO の優位差が

表 4: 節語順タイプの統計値の比較

型	節語順 タイプ	山本データ		松本データ		呉等のデータ	
		言語数	出現率 (%)	言語数	出現率 (%)	言語数	出現率 (%)
1	SOV	1,231	48.5	771	49.3	831	48.3
2	SVO	981	38.7	547	35.0	674	39.2
3	VSO	234	9.2	175	11.2	145	8.4
4	VOS	61	2.4	44	2.8	46	2.7
5	OVS	18	0.7	16	1.0	14	0.8
6	OSV	12	0.5	10	0.6	9	0.5
合計		2,537	100.0	1,563	100.0	1,719	100.0

必ずしも明確ではない場合)、または小文字の共存 (例えば「SOV/svo」は、2 つの語順が共存するが、明らかに SOV の方が優勢と認められる場合)の形でデータがまとめられている。表 4 の山本データおよび呉等のデータの言語数は大文字同士の共存を持つ言語を除き、小文字の共存を捨象して集計した数である (後に表 6 であげるその他の語順パラメータの数値についても同様である)。また、山本データでは節語順が空白や「VO」または「OV」となっている言語もあるが、それらは集計に入れていない。

図 4 は節語順タイプの地理的分布を示した言語地図である。この図では、カラーの原図をやはり白黒に印刷しているために識別しにくくなっているが、山本 (2003) が指摘したように、SVO 語順はヨーロッパ、サハラ以南のアフリカ大陸、東南アジアの 3 つの地域に特にかたまって存在し、それ以外の地域では、アジア北部、インド亜大陸等、広い範囲にわたって SOV 語順の地域が連続するという分布が顕著に現れている。

また、Greenberg (1963) の研究以来、広く知られているように、節語順の 6 つのタイプの出現率には著しい偏りがあり、上位 2 つのタイプが他を圧

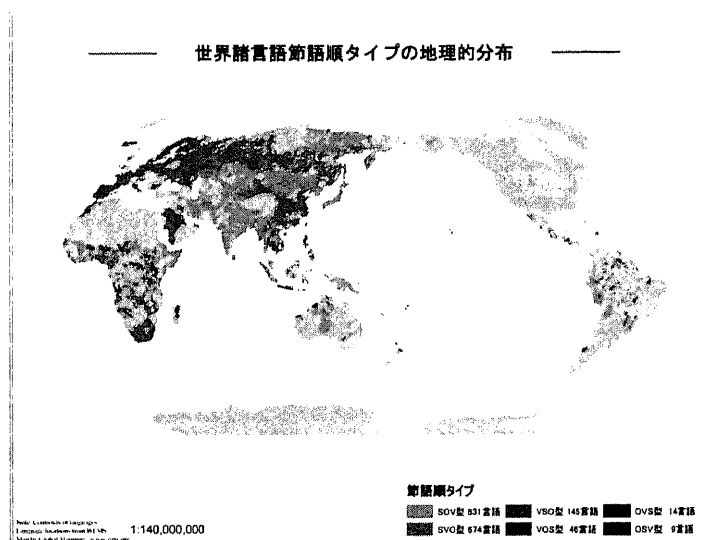


図 4: 世界諸言語節語順タイプの地理的分布

表 5: SO 順序の相対頻度

型	節語順 タイプ	山本データ		松本データ		呉等のデータ	
		言語数	出現率 (%)	言語数	出現率 (%)	言語数	出現率 (%)
1	SO	2,446	96.4	1,493	95.5	1,650	96.0
2	OS	91	3.6	70	4.5	69	4.0
	合計	2,537	100.0	1,563	100.0	1,719	100.0

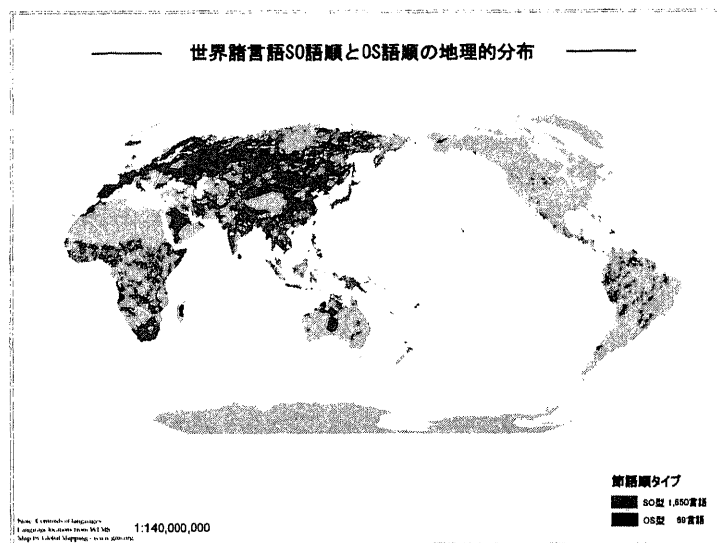


図 5 : SO 語順の普遍性

しているが、特に主語(S)が目的語(O)に先行する上位3つのタイプが世界諸言語の大半を占めているという様子が、表4と図4にも現れている。

さらに、表5のように、表4のデータからSとOの相対順序だけの統計を示してみると、この普遍的傾向が、いずれのデータにおいても明瞭にみとれる。また、実際にこのSO語順とOS語順の分布を、図5のように我々の地図上に表してみると、そもそもOS語順の言語が少数の弱小言語のためにOS言語自体の分布が見えにくくなってしまっているものの、図4の大

半の言語が SO 言語として図 5 の中に現れていることで、この普遍的傾向を容易に確認することができる。

### 3.2.2 4 大特徴による基本語順タイプの地理的分布と語順パラメータ間の相関関係

語順類型論では、しばしば表 3 にあげた 4 つの語順パラメータを語順の主要な 4 大特徴として扱ってきた。これらの 4 大特徴の組み合わせからは、理論的に 48 通りのタイプがあり得る。節語順の場合と同様に、山本データ (山本 2003) および松本データ (松本 2006) と対照させる形式で、これらの基本語順タイプについて、言語数および出現率の統計値を表 6 に示す。

表 6: 基本語順タイプの統計値の比較

型	基本語順 タイプ				山本 データ		松本 データ		呉等の データ	
					言語 数	出現 率 (%)	言語 数	出現 率 (%)	言語 数	出現 率 (%)
1	SVO	PR	NG	NA	359	27.0	323	26.6	258	28.1
2	SOV	PO	GN	NA	316	23.8	231	19.0	212	23.1
3	SOV	PO	GN	AN	268	20.2	262	21.6	191	20.8
4	VSO	PR	NG	NA	93	7.0	95	7.8	61	6.7
5	SVO	PO	GN	NA	66	5.0	50	4.1	38	4.1
6	SVO	PR	NG	AN	50	3.8	42	3.5	29	3.2
7	SVO	PR	GN	NA	21	1.6	8	0.7	18	2.0
8	VOS	PR	NG	NA	17	1.3	18	1.5	16	1.7
9	SVO	PO	GN	AN	22	1.7	30	2.5	16	1.7
10	VSO	PR	NG	AN	18	1.4	20	1.6	15	1.6
11	SVO	PR	GN	AN	21	1.6	26	2.1	13	1.4
12	SOV	PO	NG	NA	18	1.4	12	1.0	12	1.3
13	VOS	PR	NG	AN	15	1.1	16	1.3	9	1.0
14	SOV	PR	NG	NA	9	0.7	30	2.5	7	0.8
15	SOV	PO	NG	AN	4	0.3	3	0.2	3	0.3

表 6: 基本語順タイプの統計値の比較 (続き)

型	基本語順 タイプ				山本 データ		松本 データ		呉等の データ	
					言語 数	出現 率 (%)	言語 数	出現 率 (%)	言語 数	出現 率 (%)
16	OVS	PO	GN	NA	6	0.5	6	0.5	3	0.3
17	VSO	PO	GN	NA	4	0.3	3	0.2	2	0.2
18	SVO	PO	NG	NA	2	0.2	2	0.2	2	0.2
19	SOV	PR	GN	AN	5	0.4	13	1.1	2	0.2
20	OVS	PR	NG	NA	2	0.2	1	0.1	2	0.2
21	OSV	PO	GN	NA	2	0.2	4	0.3	2	0.2
22	VSO	PO	NG	NA	1	0.1	0	0.0	1	0.1
23	VSO	PO	NG	AN	1	0.1	0	0.0	1	0.1
24	VSO	PO	GN	AN	1	0.1	4	0.3	1	0.1
25	VOS	PR	GN	NA	1	0.1	0	0.0	1	0.1
26	VOS	PO	GN	AN	1	0.1	0	0.0	1	0.1
27	SOV	PR	NG	AN	2	0.2	2	0.2	1	0.1
28	VSO	PR	GN	NA	0	0.0	2	0.2	0	0.0
29	VSO	PR	GN	AN	0	0.0	3	0.2	0	0.0
30	VOS	PR	GN	AN	0	0.0	2	0.2	0	0.0
31	VOS	PO	NG	NA	0	0.0	0	0.0	0	0.0
32	VOS	PO	NG	AN	0	0.0	1	0.1	0	0.0
33	VOS	PO	GN	NA	0	0.0	1	0.1	0	0.0
34	SVO	PO	NG	AN	0	0.0	3	0.2	0	0.0
35	SOV	PR	GN	NA	3	0.2	2	0.2	0	0.0
合計					1,328	100.0	1,215	100.0	917	100.0

表 6 のいずれのデータにおいても、例えば従来の類型論で言われてきた整合的タイプの優位性等がみてとれるが、我々の地図システムでは、これら種々のタイプの地理的な分布を、自由な組み合わせで地図上に表すことが可能である。ここでは、やはりカラーの原図を白黒で印刷したものには

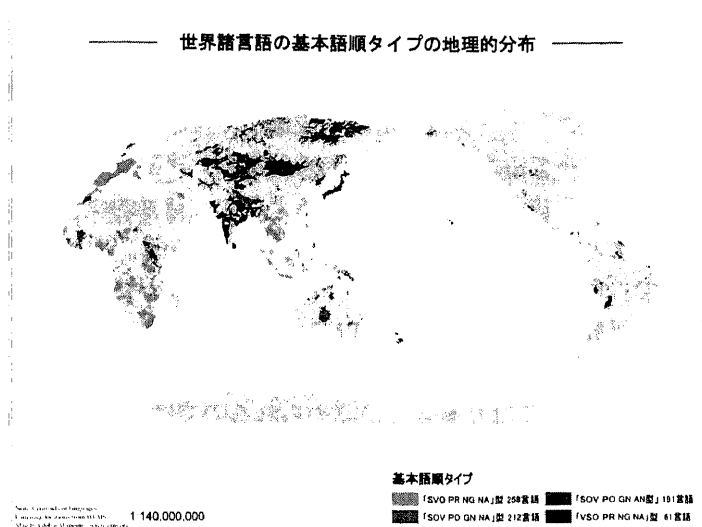


図 6: 世界諸言語の基本語順タイプの地理的分布

なるが、一例として、表 6 における出現率の上位 4 つのタイプについて、それらの地理的分布を我々の地図上に表した図 6 をあげておく。

整合的 VO 型に属する「VSO PR NG NA」タイプは、これら 4 つのタイプの中では最も少数であり、またオセアニア地域の島上に分布する言語も少なくないことなどから、この地図上に明瞭な分布として現れているとはいえない。しかし、整合的 VO 型の中で最も多数を占める「SVO PR NG NA」タイプは、特にサハラ以南のアフリカ大陸や東南アジアにおいてまとまった分布を持つ様子が、この地図からも十分に観察できる。また、従来、山本 (2003) や松本 (2006) 等によって指摘されたように、整合的 OV 型の代表である「SOV PO GN AN」タイプの言語は、ユーラシアの大部分の言語を含む広大な言語圏を領有し、形容詞のみがずれる「SOV PO GN NA」タイプの言語は、一部のチベット・ビルマ諸語地域を除けば主としてユーラシア以外の地域に分布する様子も、この地図上に反映されている。

## 4 むすび

本研究では、山本データから抽出された呉等のデータを用いて言語地図の作成を行った。このデータは、その生成過程で言語の同定ができなかったという原因によって山本データの一部の言語が欠落している。しかしながら、本稿で示した統計的数値に見られるように、結果的に、山本データ及び松本データとほぼ同様な傾向を示していることから、言語地図作成に利用する上で、呉等のデータは、かなりの程度有効性が認められるであろう。また、本研究において用いた WLMS 空間データについては、今後さらなる検証が必要かもしれないが、本稿で作成を試みた語順地図に見られるように、巨視的にはやはり有効性が認められるのではないかとと思われる。

言語地図によって言語特徴の地理的分布はとらえやすくなる。GIS は地図の表現力や、データの共有および再利用の面から言語地図作成の強力なツールとなるであろう。言語類型論において GIS を応用するメリットは言語地図の作成には限らない。今後は、GIS のもつ属性検索および空間検索の機能を世界諸言語の通言語的な特徴の考察に活かしていきたい。さらには、言語の多様性を生み出していると考えられる言語の内的および外的な種々の要因に関する定量的な分析手法を見出し、GIS を用いた多次元空間における言語普遍性のモデルを樹立する可能性を探ってみたい。

### 【参考文献】

- 呉靱・乾秀行・杉井学・松野浩嗣 (2007) 「言語研究のための GIS データの生成について : *Ethnologue* GIS データを言語特徴の地図化に用いる一手法」『人文科学とコンピュータシンポジウム論文集 (2007)』253-258.
- Gordon, R.G. (ed.) (2005) *Ethnologue: Languages of the World*, 15th ed. Dallas: SIL International.
- Greenberg, Joseph H. (1963) 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements.' *Universals of Language*. Ed. Joseph H. Greenberg. Cambridge, MA: MIT Press, 58-90.



池田潤 (2006)「GIS と言語研究」『一般言語学論叢』9:1-10.

松本克己 (2006)『世界言語への視座-歴史言語学と言語類型論』三省堂.

山本秀樹 (2003)『世界諸言語の地理的・系統的語順分布とその変遷』溪水社.

山本秀樹 (2006)「GIS と言語類型論:世界言語地図に基づく言語研究」『一般言語学論叢』9:31-40.

# Essential Data for Word Order Mapping and Word Order Distribution Revealed on the Map

Ren WU, Hideki YAMAMOTO, Hideyuki INUI,

Manabu SUGII & Hiroshi MATSUNO

For the purpose of mapping linguistic features of the world's languages, we have proposed a method of generating GIS data by combining Yamamoto's data with the data of WLMS (The World Language Mapping System).

In this paper, we verify the usefulness of our generated GIS data in studying word order. First, we introduce WLMS to show its validity for mapping word order of the world's languages. Then, we discuss the geographic distribution of word order types based on our GIS data. Finally, we verify the usefulness of our GIS data by mapping the word order distribution and comparing our data with Yamamoto's data and Matsumoto's data.

*Graduate School of Science and Engineering  
Yamaguchi University*

*1677-1 Yoshida, Yamaguchi, Yamaguchi 753-8540, Japan*

*E-mail: wu@ib.sci.yamaguchi-u.ac.jp (Wu)*

*Faculty of Humanities*

*Hirosaki University*

*1 Bunkyo-cho, Hirosaki, Aomori 036-8560, Japan*

*E-mail: hideyama@cc.hirosaki-u.ac.jp (Yamamoto)*

*Faculty of Humanities*

*Yamaguchi University*

*1677-1 Yoshida, Yamaguchi, Yamaguchi 753-8540, Japan*

*E-mail: inui@yamaguchi-u.ac.jp (Inui)*

*Media and Information Technology Center*

*Yamaguchi University*

*1677-1 Yoshida, Yamaguchi, Yamaguchi 753-8540, Japan*

*E-mail: manabu@yamaguchi-u.ac.jp (Sugii)*

*Graduate School of Science and Engineering*

*Yamaguchi University*

*1677-1 Yoshida, Yamaguchi, Yamaguchi 753-8540, Japan*

*E-mail: matsuno@sci.yamaguchi-u.ac.jp (Matsuno)*