

多次元尺度法を用いた語順パラメータの間の関係付け

Relations among Word Order Parameters Analyzed by Multi-Dimensional Scaling

江原 輝将

Terumasa EHARA

NHK放送技術研究所

NHK Science and Technical Research Laboratories

eharate@strl.nhk.or.jp

1. はじめに

世界には多くの言語があり、それらの間には類似点と相違点がある。言語類型論は、この点を明らかにする研究分野である。言語類型論の分野の中でも、語順に関する類型についての研究は初期のころから行われており、様々な事実が分かっている。しかし、従来の研究結果は、定性的なものであり、また、語順に関する様々な性質を、主として、個別に検討するものであった。本研究は、従来から調べられている語順に関する諸性質を、定量的かつ総合的に捉え、諸性質の間の関係を明らかにするとともに、語順類型から見て、世界の言語を特徴付けることを目的としている。そのために、多次元尺度構成法を用いる。

2. 語順に関する諸性質とその数量化

本文で用いる語順に関する諸性質としては、[角田]で用いられている20の性質を用いた。表1にこれらの性質を示す。[角田]は、130の言語に対して、これらの諸性質がどのように実現されているかを定性的に示した。ここでは、定量的な解析を行うために、これらの諸性質を語順パラメータとして、数量化する。パラメータの値は-10ないし+10の範囲であり、各性質に対して、+の方向は、その性質が日本語で、どのように実現しているかに応じて定めた。例えば、第1番目の語順パラメータP1である、平叙文における主語(S)と動詞(V)の相対順序については、日本語と同一のSVの語順が+であり、VSの語順が-である。次に、注目している言語Lに対して、パラメータの値は、以下のように定める。

(1) Lの典型的な表現に対して、当

該の性質が片方のみに実現しているときには、+10または-10の値を与える。例えば、日本語では、典型的にSVであるから、P1の値が+10である。また、ウェールズ語では、典型的にVSであるから、P1の値が-10である。

(2) Lに対して、当該の性質が片方のみに実現していない場合は、その程度に応じて、+10と-10の間の数値を与える。例えば、英語では、固有名詞と普通名詞の順序に対して、N-PrとPr-Nの両方の場合があるが、N-Prの方が多いため、英語に対するP9の値は-5とした。

(3) Lに対して、当該の性質がほぼ同等に生ずる場合には、値0を与えた。また、Lに対して、当該の性質が不明の場合や、その性質を考えることができない場合も0を与えた。例えば、中国語では、平叙文における目的語(O)と動詞(V)の語順として、OV

表1 語順に関する性質とそのパラメータ化

No	語順に関する性質	+	その他	-
1	平叙文での主語(S)と動詞(V)	SV		VS
2	平叙文での目的語(O)と動詞	OV		VO
3	名詞(N)と側置詞(Ap)	N-Ap		Ap-N
4	所有格(G)と名詞	GN		NG
5	指示詞(Dm)と名詞	Dm-N		N-Dm
6	数詞(Nu)と名詞	Nu-N		N-Nu
7	形容詞(A)と名詞	AN		NA
8	関係節(Re)と主名詞	Rel-N		N-Rel
9	固有名詞(Pr)と普通名詞	Pr-N		N-Pr
10	比較の基準(C)と形容詞	CA		AC
11	本動詞と助動詞(X)	VX		XV
12	副詞(D)と動詞	DV		VD
13	副詞と形容詞	DA		AD
14	疑問の印	SF		SI
15	一般疑問文でのS V倒置	no		yes
16	疑問詞の位置	SF	PiD	SI
17	特種疑問文でのS V倒置	no		yes
18	否定の印(Ng)と否定の対象(On)	On-Ng		Ng-On
19	条件節(Cc)と主節(Cl)	Cc-Cl		Cl-Cc
20	目的節(Cp)と主節	Cp-Cl		Cl-Cp

SI: 文頭 SF: 文末 PiD: 平叙文の位置

と VO が、ほぼ同等に生ずるので、P2 の値が 0 となる。

(4) P16 については、疑問代名詞が平叙文と同一の位置に現われる場合、+1 を与える。

以上的方法で 126 の言語に対して 20 の語順パラメータの値を設定した。この設定に当たっては、多く [角田] によったが、一部 [亀井] も参考にした。また、筆者独自の調査による部分も一部ある。付録 1 にその結果を示す。付録 1 では、値 10 を a で示してある。

3. 多次元尺度法による解析

語順パラメータを用いて、言語 L_i と L_j ($i, j = 1, \dots, N$; $N = 126$) の間の距離が、R 次元 ($R = 20$) ユークリッド空間での距離として計算できる。このようにして得られた $N \times N$ の距離行列 D に基づいて、Torgerson の多次元尺度構成法によって、内積行列 B が求められる。このとき、Torgerson の単純法で加算定数を計算し、用いた。この値は -0.89 であった。B の固有値を図 1 に示す。第 2 主成分まで累積寄与率 53% を達成している。図 2 は N 個の言語を第 2 主成分までの固有空間に配置した結果である。この空間を「語順空間」と名付ける。各言語に対する、第 1 および第 2 主成分の値を付録 1 の列 A と B に示す。

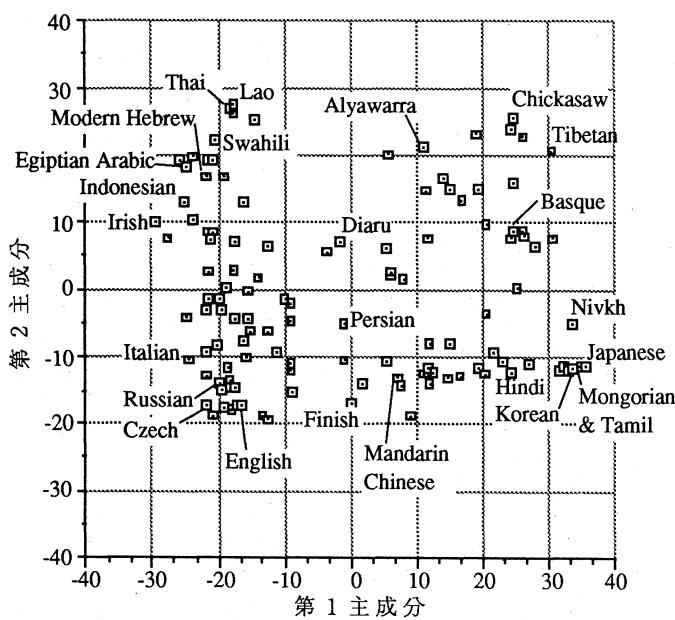


図 2 語順空間内での126言語の分布

4. 語順パラメータの間の関係

語順空間での各言語の配置から、以下の様にして語順パラメータの間の関係を求めることができる。パラメータ P_k に対して、 $L = \{L_i : i = 1, \dots, N\}$ を次の 3 個の部分に直和分割する。 $L_{k+} = \{L_i : P_k(L_i) > 0\}$ 、 $L_{k-} = \{L_i : P_k(L_i) = 0\}$ 、 $L_{k-} = \{L_i : P_k(L_i) < 0\}$ 。語順空間での L_{k+} と L_{k-} の平均位置を q_{k+} 、 q_{k-} とするとき、パラメータベクトルを $q_k = q_{k+} - q_{k-}$ と定義する。ベクトル q_k と横軸との成す角 (α_k) を求めて、 L_{k+} と L_{k-} の要素数と共に表 2 に示す。表 2 から、語順パラメータの集合 $P = \{P_k : k = 1, \dots, R\}$ を以下の 3 個の部分に分割する。 $P_1 = \{P_k : 0 \leq \alpha_k \leq 30\}$ 、 $P_2 = \{P_k : 60 \leq \alpha_k \leq 90\}$ 、 $P_3 = \{P_k : 30 < \alpha_k < 60\}$ 。 P_1 の要素は第 1 主成分と相関が強く、13 個の語順パラメータを含む。また、 P_2 の要素は第 2 主成分と相関が強く、4 個の語順パラメータを含む。また、各部分に属するパラメータ同士はお互いに相関が強い。 P_1 に属する代表的なパラメータとして、P2 (目的語と動詞の順序) を選ぶことができ、 P_2 に属する代表的なパラメータとしては、P7 (形容詞と名詞の順序) を選ぶことができる。

これによって、言語は、目的語と動詞の順序によって第 1 に特徴付けられ、形容詞と名詞の順序によって第 2 に特徴付けられることが分かる。前者は、従来から言われていたことであるが、後者は、必ずしも明確ではなかった。[\[Comrie\]](#) によれば、目的語 (O) と動詞 (V) の順序と形容詞 (A) と名詞 (N) の順序には強い相関があるとされ、

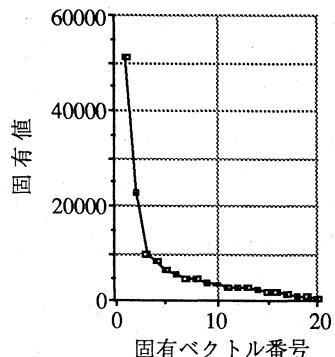


図 1 内積行列の固有値

OV N-Ap GN ANの組とVO Ap-N NG NAの組の2つが語順に関する主要な型であるとされているが、ここでの結果は、必ずしもそうとは言えないことを示している。

5. 語順と思考パターンの関係

本節では、語順と思考パターンの関係についての仮説を述べる。前節で述べたように言語はOV型であるか、VO型であるかで、第1に特徴付けられる。そして、OV型である言語は、N-Apであり、GNである傾向が強い。逆に、VO型言語はAp-N、NGである傾向が強い。日本語は典型的なOV型であり、欧州の言語の多くはVO型である。

一方、文を主要要素と従要素との関係を中心に解析する手法に依存解析（依存文法）がある。VO型言語の文を依存解析すると、主要要素が従要素の前にくることが多く、OV型言語の文を依存解析すると、主要要素が従要素の後方にくることが多い。依存解析と同様な手法である係受け解析が日本語に対して存在する。依存解析も係受け解析も、解析の結果、従要素と主要要素を有向弧で結んだ解析木を得るが、図3に示すように、依存解析では、弧の方向が主要要素から従要素の方向を向き、係受け解析では従要素から主要要素の方向を向く。この違いは単なる記法の差で

あろうか。筆者は、これを語順類型の影響ではないかと考える。つまり、弧の方向を示す矢印は本物の矢と同じに、前方から後方に向かって付けられるのが自然であるので、VO型言語の多い欧州で生まれた依存解析では、弧の方向が、主要要素から従要素に向かい、OV型言語である日本語に対する係受け解析では従要素から主要要素に向かったのではないかと考える。そして、この差は、実は思考パターンにも影響しているのではないか。つまり、VO型言語の話者は、トップダウン（演繹的）に思考する傾向が強く、OV型言語の話者はボトムアップ（帰納的）に思考する傾向が強いのではないかと考える。あるいは、既に、このような関係のないことが証明されているのかもしれないが、もし、そうでないとすれば、こうした関係があるか否かを実証することは興味あることと思われる。

6. おわりに

これまでに得られている語順類型に関する研究成果を、多次元尺度構成法を用いて定量的に調べた。その結果、第2主成分までで、累積寄与率53%が得られた。第1主成分は「目的語と動詞の順序」、第2主成分は「形容詞と名詞の順序」として意味付けられることが分かった。最後に、語順類型と思考パターンの間の関係に関する仮説を述べた。

参考文献

- [Comrie] Comrie, B. : Language Universals and Linguistic Typology, University of Chicago Press, 1981.
- [亀井] 亀井孝ほか（編著）：言語学大辞典1～4、三省堂、1988～92年。
- [角田] 角田太作：世界の言語と日本語、くろしお出版、1991年。

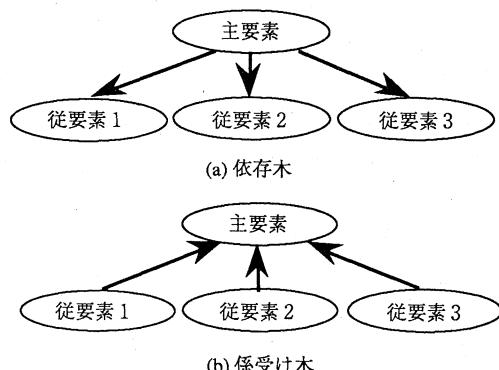


図3 依存木と係受け木

Group No.	Param. No	ak (deg.)	L k+	L k-
1	9	0.0	29	75
	3	0.9	51	60
	10	1.7	33	53
	2	3.6	55	68
	12	4.6	38	18
	11	4.7	39	50
	18	5.4	30	75
	8	6.3	21	83
	4	8.1	73	45
	20	10.2	35	15
	16	11.6	30	95
	1	17.9	99	26
	14	19.7	50	47
3	15	37.6	112	14
	17	40.4	109	17
	19	54.4	121	3
2	6	61.4	96	29
	5	64.0	83	38
	13	72.6	85	28
	7	83.8	65	60

