

# デジタルアーカイブを用いた古典修辞構造検証手法の検討

村井源

東京工業大学 大学院社会理工学研究科

デジタルアーカイブと計量的手法を用いて、聖書の修辞構造の妥当性を数値的に検証するための手法を検討した。検証対象仮説としては聖書の多層集中構造を取り上げた。手法としてはまず、修辞構造での各テキストの箇所間の対応関係を、出現頻度の低い共通の単語・フレーズが含まれるかを基準として判定した。次にそれらのテキスト箇所間の対応関係の生起する確率がランダムに対応関係を想定した場合と比べて有意に多いかを正規分布を用いて判定した。結果として、対象とした聖書テキストの多くで、出現頻度 2 割以下の単語の対応関係において、検証対象仮説に有意差があることが明らかとなった。

## A study about validation methodology for classic rhetorical structure by utilizing digital archive

Hajime Murai

Graduate School of Decision Science and Technology  
Tokyo Institute of Technology

Utilizing digital archive and quantitative methods, a numerical validation method for rhetorical structure of the Bible was studied. "Parallel Concentric Structures within the Bible" was selected as a hypothesis for validation. As a methodology, at first, relationships between each part of text in rhetorical structure were validated based on occurrence of rare common word and phrases. Secondly, generation ratios of those relationships of each text were validated by compared with some random structures. As a result, in many of those texts of the Bible, it is found that hypothesis is validated in relationships of words which occurred less than 20% of part of text.

### 1. まえがき

古典文献の分野でも種々の電子化プロジェクトが進行している。電子化プロジェクトによって構築されつつある各種のデジタルアーカイブシステムはテキストの検索や比較（コンコルダンス等）などのさまざまな情報処理技術を取り込む形で進歩してきている。

このような中で筆者は、テキストのアーカイブ化のみでなく、情報の理解（テキストの場合解釈）のサポートまでを視野に入れた電子解釈システム（Digital Hermeneutic System: 以下 DHS）を提案し[1]、特にテキスト間の関係性（間テキスト性）からの意味処理をサポートするためのシステムの構築を行ってきた[2][3]。

間テキスト性をアーカイブに付与することによって、複雑に絡み合うテキスト間の関係性を数値的に扱うことが可能となり、歴史を通じて読み継がれてきた古典テキストとその周辺文書での解釈の分析に有効活用が可能と考えられる。

間テキスト性(inter-textuality)の中で、テキスト内の構造に特化したものを intra-textuality と呼ぶ[4]。Intra-textuality の具体例としては、テキストの内的なデザイン、構造、分割などが意図的に整然と配置されている物などが想定されている。これらは現代のテキストにも当然見出される物であるが、修辞学の発達したギリシャ・ローマ

時代から、様々な修辞的技法の実践としてテキスト内関係構造の精緻化が行われていた。聖書など古典テキストには、このような古代の修辞構造が多数用いられており、一つのテキストの中に膨大なテキスト箇所間の対応関係が含まれている。このため、人手による修辞構造の分析をサポートするため、情報処理システムを用いたアーカイブや意味処理が人文科学の発展に寄与する可能性は高いと考えられる。

本論文では特に、古典的な修辞構造の分析に常に付きまとう問題である構造特定の恣意性に対して、データベースと情報処理技術を用いた計量的アプローチを試み、古典テキストの意味解釈とその妥当性の検証という問題に対して、デジタルアーカイブが果たしうる役割をケーススタディ的に探っていくこととする。

### 2. 古典的修辞構造

#### 2.1. 古典的修辞構造とアーカイブ

前述のように、修辞学の発達したギリシャ・ローマ時代から、様々な修辞的技法の実践としてテキスト内関係構造の精緻化が行われていた。テキスト構造の修辞技法で広く知られているところでは、聖書学の分野で近年精力的に研究が進められている並行法・交差配列法・集中構造などを挙げられよう[5]。

交差配列法 (chiasmus) とは、ABB'A' などと記号で表現されることが多いが、テキストのある部分が異なりながらも対応している小部分の関係性による入れ子になっている構造である。もっとも単純な形の交差配列法の例を図 1 に示す。図中での対応する箇所を線でつなぐと X (ギリシア文字のカイ) の字になることから、カイアスムスやキアスムス (chiasmus) などと呼ばれる。

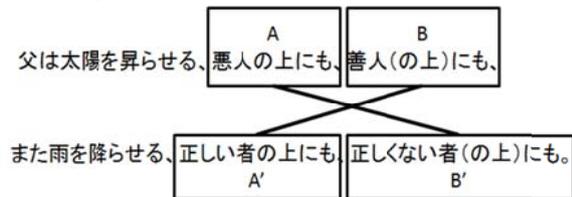


図 1 交差配列法の例 (Mt5:45)

図 1 のように小さい構造の場合は単語やフレーズの単位での修辭的構造が作られるが、より大きな文単位で作られる交差配列法も存在する。また、対応関係も AB の二種類のみではなく、より多層の入れ子構造の形式が多数みられる。さらに文よりさらに大きな単位での構造として、聖書学のテキストの小区分である pericope 単位の交差配列法も多数存在する。pericope は長さ必ずしも一定していないが、数文から数段落程度の物語の切れ目で区切られた単位である。

次に集中構造(concentric structure)であるが、基本的な構造としては交差配列法と同じであるが、中心に核となる部分が存在する構造がこのように呼ばれる。中心に核が存在することで、テキスト中でどの部分が重要なかを明示的に示す効果を併せ持った技法であると考えられている。

マルコ福音書中にある pericope 単位での集中構造の例を図 2 に示す。図中の数字は pericope の範囲を示す聖書略記号である。この場合、中心に来るのは 9:30-32 にある二度目の受難の予告であり、当該箇所全体での中心的なテーマが受難の予告にあるということが解釈可能である[6]。

A	8:22-26	盲人癒される
B	8:27-30	ベト口の信仰告白
C	8:31-33	受難の予告一
D	8:34-38	イエズスに従うもの
E	9:1	死なないものの預言
F	9:2-8	イエズスの変容
G	9:9-13	エリヤが来る
H	9:14-29	悪霊に憑かれた子
I	9:30-32	受難の予告二
H'	9:33-41	一番偉いもの
G'	9:42-50	誘惑についての警告
F'	10:1-12	離婚の問題
E'	10:13-16	イエズスと幼子たち
D'	10:17-31	金持ちの青年
C'	10:32-34	受難の予告三
B'	10:35-45	ゼベダイの子らの願い
A'	10:46-52	エリコの盲人癒される

図 2 pericope 単位の集中構造の例 (Mk8:22-10:52)

最後に並行法(parallelismus)であるが、名称の通りテキストのある二つの小部分が並行的な関係にある修辭構造を指す言葉である。図 3 に文単位での並行法の例を示す。

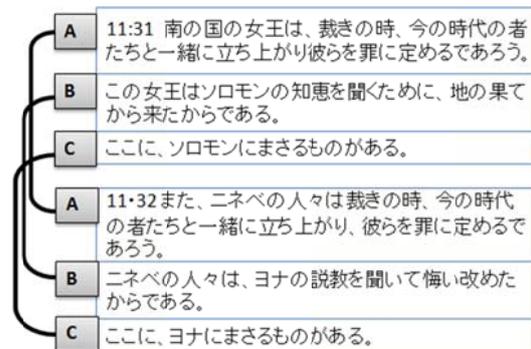


図 3 文単位の並行法の例 (Lk11:31-32)

これらの修辭構造は元来、伝承がテキスト化される以前の口頭伝承の際の記憶の補助的役割を担っていたと考えられるが、修辭学の発達や、テキスト自身が種々の編集段階を経ていく中で、テキストの意味解釈の上でも重要な役割を担う者になってきたと考えられている。上記のような修辭構造の分析・特定により、

- テキスト内の整合性の確認
  - テキストの核となるメッセージの発見
  - 対応によってメタファー等の解釈が明確化
  - 語義やテキスト範囲の確定
- などが可能になると考えられている[5]。

これらの修辭構造は、デジタルアーカイブにおいては対応する二つのテキスト箇所間の関係性を記述したデータとしてリレーショナルデータベースへの蓄積が可能である[3]。

## 2.2. 古典的修辭構造の検証の必要性

古典的な修辭構造にまつわる問題点として、あるテキスト箇所に対して複数の修辭構造の可能性が考えられ、かつそれらのうちのどの仮説の妥当性が高いかを客観的に検証することが困難であることが挙げられる。

複数の仮説が乱立しうる理由の一つとしては、修辭構造での各テキスト要素間の対応関係に様々な形式がありうるということが考えられる。

聖書での修辭構造の対応関係としても、

- 単語・フレーズが同じ
- 類義語が対応
- 反義語が対応
- 同じテーマを扱う
- 類似の物語構造を用いている

など様々な形での対応関係が考えられてきた。

このうちで単語やフレーズなど単語のレベルでの対応関係が明確なものに関しては計量的な分析への適性があると考えられる。

しかし類義語や反義語に関しては精度の良い辞書が必要であり、かつ辞書の妥当性も検証される必要が生じる。特に聖書などにおける神学的な意味での対応関係では一般的な辞書での類義語・反義語のリストは不十分な場合がある。例えば、出エジプト記においては、「神への祭儀」と「ファラオへの奴隷労働」が全体にわたって反義的な概念として機能するため（祭儀的な奉仕と奴隷労働はヘブライ語では同じく[obd]で表される）、「神」と「ファラオ」はしばしば物語全体の流れを決める重要な反義語として解釈する必要がある。このような、テキストに固有の反義語は一般的な語義の関係性の辞書には当然網羅されておらず、各テキストで個別の同義語・反義語の関係性を逐次抽出しなければ、精度の良い分析は困難である。

また、同じテーマを扱っているという関係や物語構造の類似は、テーマや物語を客観的数値に置き換えて扱う手法が未発達であるため、単語の対応関係以上に数値的な評価が困難である。

そこで本研究では修辞構造の仮説の妥当性の検証の手始めとして同語・同フレーズを対象とした数値的指標の検討を行う。

### 3. 修辞構造の妥当性検証手法の方針

ある修辞構造が著者の意図した構造である場合、それらの構造で対応するテキスト箇所間の共通性は、ランダムなテキスト箇所間のペアでの共通性よりも一般的に高いことが予測される。著者が意図せざる対応がランダムテキスト箇所間のペアで発見されることは当然考えられるが、ある程度以上の数のテキスト箇所のペアで平均をとった場合、著者によって意図的に組み合わせられたテキスト箇所間での共通性と、ランダムなテキスト箇所間の共通性には、統計的に有意な差が出るのが期待されるからである。

また、著者が意図的にテキスト箇所のペアを構成する場合に、共通の単語・フレーズは他のテキスト箇所の大部分にも出現する語ではないと考えられる。なぜなら、全体に頻出する語の場合、それがテキストの特定の二箇所にも共通に出現しても、読者側がそれらの対応関係を読み取るのが困難であることは自明だからである。

これより、出現頻度の低い単語・フレーズを対象に、検証対象の修辞構造でのテキスト箇所のペアとランダムなテキスト箇所間でのペアを比較して有意な差が出るかを、修辞構造の妥当性の数値的指標とする。

ランダムなテキスト箇所のペアと言っても様々な形態が考えられる。本論文での目的に沿ったペアをランダムに得るためには少なくとも下記を検討する必要がある。

1. 対象テキスト
2. テキスト箇所の分割
3. ランダムペアの抽出アルゴリズム

まず対象テキストは、語彙の分布が異なると結果が大きく異なると予測される。このため、検証対象仮説で用いるものと同じテキストを、ランダムなテキスト箇所のペアの作成においても用いる必要がある。

次にテキスト箇所の分割に関してだが、分割の位置を変えると各テキスト箇所の語の分布は変化する。結果として、出現頻度の低い単語の対応関係のできやすさや、対応関係になる語やフレーズ自体も変化する。このため本来は様々な考えうるテキストの分割に対してそれぞれ、語やフレーズのペアの構成の頻度を計量して、検証仮説の有意性を数値的に検証すべきである。しかしテキストの分割位置自体にも多数の仮説が乱立する状況であるため、それらの妥当性もまた別個に数値的に検証する必要がある。そこで本論文では検証対象仮説の分割にのっとった場合にどのような数値的指標が出るのみを計算することとし、テキスト分割の妥当性自体は問わない方針とする。つまり、手順としては、検証仮説の分割に基づいたランダムなペアとの比較で数値的に有意であるとの検証を本論文で行い、今後何らかの別の手法を用いてテキスト分割の有意性を検証するという方針をとる。

最後にテキスト箇所間のペアのランダム性であるが、一つの連続した物語である場合、一般的に近い位置にあるテキスト箇所間ほど類似性が高いことは自明である。このため、検証対象の修辞構造における箇所間のペアの平均距離と比較対象のランダムなテキスト箇所間のペアの平均距離は可能な範囲で近くする必要がある。この場合での距離は、テキストを分割した各箇所を1として数えた時に何箇所分離れているかを指標とする。分割されたテキストの小区分の数が9である場合の修辞構造とそこからの距離の算出の例を図4に示す。図中での各テキスト箇所の単位は聖書学での *pericope* としている。

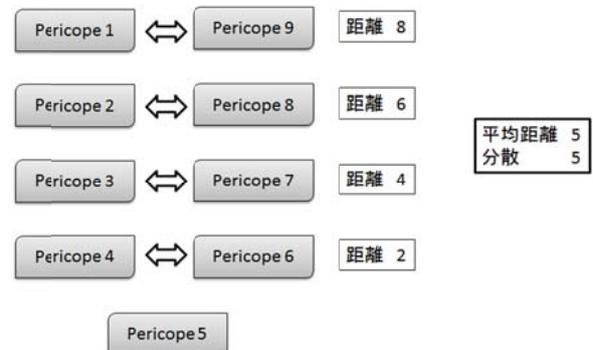


図4 検証仮説からの距離の算出

図4の場合に、検証のために比較対象とするランダムな箇所間のペアは例えば図5のような形となる。ランダムペアは平均距離が検証仮説と近くなるように、テキストの小区分の位置をランダムに入れ替えて作成する。

実際に検証に用いるランダムデータは、検証対象仮説に含まれる対応間の各距離に対して、正規分布の乱数を加えた距離を作成し、その距離を満たす二か所の対応を、重複がないようにランダムに選択して作成していく。結果として得られたランダムな対応全体の平均距離と分散が検証対象仮説に近くなるようにパラメータを調整し補正を加えたものを用いた。

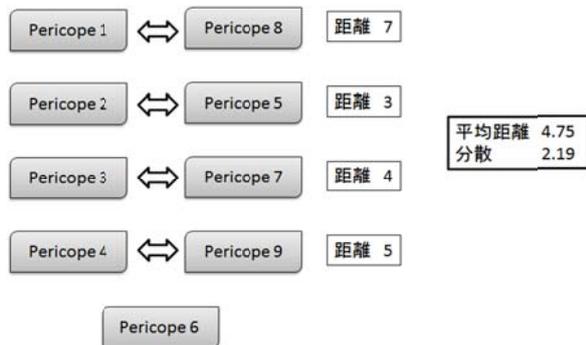


図 5 ランダムな構造からの距離の算出

修辞構造妥当性検証のための基本的数値指標としては出現頻度の低い単語・フレーズが共通に含まれるテキスト箇所間のペアの数を用いる。出現頻度が低いにも関わらず想定する対応関係のペア上で共通な単語・フレーズ数が多いほど、意図的な語彙配置が行われている可能性が高いと考えられるからである。出現頻度が高い語彙の場合は、前述のように、たとえ筆者が意図的に対応関係のペアを構成するように配置したとしても、ランダムな箇所間のペアでの共通性と数値的に大きな差が出るとは期待できないため分析対象とはしない。単語の出現頻度が異なる場合に、どのように共通語彙を持つペアの成立確率が変化するかを示すために、先ほどと同じ9分割での場合の図を図6と図7に示す。

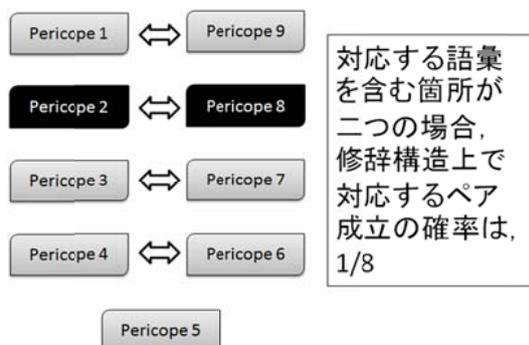


図 6 出現頻度 2 の場合の ペア成立確率

統計的な有意性を確認するため、ランダムな試行 100 回を行った場合の平均と分散を計算し、検証対象仮説での単語・フレーズを共通に持つペアの生起頻度との比較を行う。

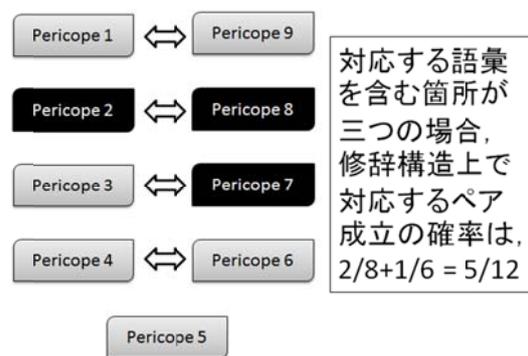


図 7 出現頻度 3 の場合の ペア成立確率

対象となる単語・フレーズの出現頻度であるが、分析対象テキストの筆者がどのような基準で共通語彙を選択したか不明であるため分析者が一意に閾値を決めることができない。そこで、全テキスト箇所中の 10%未満、20%未満、30%未満の三種類の閾値を用い、閾値以下の箇所数でのみ出現する単語とフレーズを対象に検証を行うこととした。フレーズとしては前置詞・冠詞・接続詞を除いて 2 語の連続フレーズと 3 語の連続フレーズを対象とした。また、これらに合わせて、少数の語を間に挟みつつも関連して用いられることの多い単語のペアもフレーズと同様に分析に含めるため、5 単語以内の距離にある単語の組（サイズ 5 のウィンドウ法により得られる単語の組）も分析対象とした。

ランダムな組み合わせの分析結果の平均値と標準偏差より、検証仮説での共通語彙を含んだペア数が偶然に生起する確率を、ペア数の分布を正規分布と仮定して算出する。その結果より統計学的な有意水準を超えているかを判定する。

## 4. 修辞構造の検証

### 4.1. 検証対象修辞構造

本論文での修辞構造の数値的な妥当性検証の対象仮説として、聖書の多層集中構造を用いる [6][7][8][9]。これは例えば出エジプト記の場合、全テキストを従来の聖書学的なテキスト分割の位置にほぼ沿って 49 のテキスト箇所 (pericope) に分割した場合、25 番目を中心とする集中構造がテキスト全体に現れると同時に、全テキストを 3 等分 (1~17 番, 17~33 番, 33~49 番), 4 等分 (1~13 番, 13~25 番, 25~37 番, 37~49 番) など整数で分割した場合にも同様に各領域で集中構造が表れるという仮説である。

検証仮説におけるテキスト分割と対応関係の概要は引用文献 [9] に掲載されている。

検証対象仮説では 49 に分割した出エジプト記テキストから 23 の修辞構造 (一部重複を含む)、122 組のテキスト箇所のペアが得られる。そこで比較対象となるランダムなテキスト箇所間の

ペアでも、テキスト箇所間の距離の平均と分散が近くなるように調整した 122 組のランダムなペアを作成し、出現頻度の低い単語とフレーズが共通に出現する組をカウントする。これを 100 回繰り返し、ランダムな 122 組中での共通性のあるペア出現の回数の平均と分散を算出する。

検証対象仮説はテキストの分割位置および各テキスト箇所間の対応関係が、すでに開発したデジタルアーカイブシステムに入力されており、分析はこれらのデータに基づいて新たに分析用のプログラムを付加して行った。

本論文においては、性・数・格などの様々な要因で変化する単語の変化形で数えるのではなく、原形が同じ単語を一つの単語として計量を行うために、各テキストの単語を形態素解析した結果のデータを分析に用いた。分析対象として用いたのは Bibleworks 8 [10]所収の、単語の形態素解析データである。旧約聖書に関しては Groves-Wheeler Westminster Morphology and Lemma Database[11]を、新約聖書に関しては BibleWorks Greek New Testament Morphology[12]を用いている。

表 1 に検証対象として用いた聖書のテキストとそれらの多層集中構造仮説による修辞構造を記す。統計的な検証を行うため、テキスト量が多く、かつ対応するテキスト箇所のペアも多いテキストを聖書中から選択している。また、表 2 から表 5 にテキストの分割数が 49, 81, 105, 145 のそれぞれの場合に、どの長さの修辞構造がどれだけ含まれるかを示す。なお表中の数字は重複する構造を含んだものとなっている。

表 1 検証に用いたテキストと構造

テキスト分割	修辞構造	対応箇所数	テキスト名
49	23	122	出エジプト記, 民数記, 申命記, エゼキエル書
81	32	242	創世記, サムエル記, 列王記, エレミヤ書, マルコ, ヨハネ, 使徒行録, 黙示録
105	30	328	イザヤ書
145	41	536	マタイ, ルカ

表 2 49 分割の場合の各構造の長さとお数

長さ	49	25	17	6	8	7
個数	1	2	3	4	6	7

表 3 81 分割の場合の各構造の長さとお数

長さ	81	41	27	21	17	11	9
個数	1	2	3	4	5	8	9

表 4 105 分割の場合の各構造の長さとお数

長さ	105	53	35	27	21	15	14
個数	1	2	3	4	5	7	8

表 5 145 分割の場合の各構造の長さとお数

長さ	145	73	49	37	29	25	19	13
個数	1	2	3	4	5	6	8	12

## 4.2. 旧約聖書での検証

以下表 6 から表 14 に旧約聖書のテキストでの検証結果を順に示す。検証の結果 1%有意水準を\*\*, 5%有意水準を\*, 10%有意水準を+で右欄に示した。

結果より、創世記、イザヤ書とエゼキエル書を除いて、単語の出現閾値を 20%にした場合に検証仮説は有意にペア数が多いことが分かった。

表 6 創世記での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	222	211.2	4.53	**
	20%	221	210.5	5.01	*
	10%	205	197.7	4.42	*
単語	30%	219	213.0	3.90	+
	20%	195	187.8	5.96	
	10%	140	129.9	6.73	+
二語連続フレーズ	30%	53	69.1	6.37	
	20%	53	69.0	7.03	
	10%	44	54.3	5.84	
三語連続フレーズ	30%	6	6.1	2.27	
	20%	6	5.6	2.02	
	10%	6	5.8	2.34	

表 7 出エジプト記での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	108	110.5	2.92	
	20%	107	109.2	2.68	
	10%	89	92.5	3.92	
単語	30%	110	103.3	3.21	*
	20%	98	91.2	4.16	+
	10%	38	39.8	5.07	
二語連続フレーズ	30%	44	40.0	4.29	
	20%	40	34.5	4.55	
	10%	29	27.8	4.08	
三語連続フレーズ	30%	8	10.2	2.60	
	20%	8	9.9	2.41	
	10%	6	6.7	2.30	

表 8 民数記での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	105	109.2	2.73	
	20%	103	107.5	3.36	
	10%	96	94.8	3.95	
単語	30%	112	106.6	3.07	*
	20%	96	85.6	4.34	**
	10%	39	44.6	4.45	
二語連続フレーズ	30%	48	48.3	5.26	
	20%	45	44.8	4.49	
	10%	28	29.8	4.15	
三語連続フレーズ	30%	18	15.5	3.35	
	20%	18	16.1	3.75	
	10%	12	11.0	2.50	

表 11 列王記での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	231	230.7	3.22	
	20%	224	224.1	3.87	
	10%	206	209.0	4.21	
単語	30%	220	213.5	3.97	*
	20%	204	188.1	5.66	**
	10%	126	127.8	6.71	
二語連続フレーズ	30%	123	116.4	6.17	
	20%	108	103.6	6.26	
	10%	92	85.6	6.69	
三語連続フレーズ	30%	49	51.9	4.64	
	20%	38	37.4	4.37	
	10%	27	22.1	3.37	+

表 9 申命記での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	117	115.1	1.88	
	20%	116	113.1	2.70	
	10%	97	96.4	3.77	
単語	30%	104	102.2	3.45	
	20%	90	83.6	4.77	+
	10%	47	36.3	3.97	**
二語連続フレーズ	30%	67	63.4	4.63	
	20%	63	60.5	4.98	
	10%	45	37.6	5.06	+
三語連続フレーズ	30%	19	16.0	3.68	
	20%	19	16.4	3.77	
	10%	15	10.9	2.68	+

表 12 イザヤ書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	264	257.8	7.60	
	20%	251	243.8	7.60	
	10%	246	233.9	7.39	+
単語	30%	312	312.6	3.13	
	20%	291	289.6	5.19	
	10%	229	222.0	6.78	
二語連続フレーズ	30%	92	88.1	8.20	
	20%	61	56.1	5.02	
	10%	61	55.5	6.50	
三語連続フレーズ	30%	7	5.8	2.08	
	20%	7	6.4	2.22	
	10%	7	5.6	2.22	

表 10 サムエル記での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	224	225.6	3.57	
	20%	219	222.0	3.53	
	10%	212	214.6	3.95	
単語	30%	222	221.6	3.54	
	20%	207	195.7	4.71	**
	10%	152	141.0	6.78	+
二語連続フレーズ	30%	75	70.7	6.32	
	20%	75	69.4	6.50	
	10%	75	70.3	6.63	
三語連続フレーズ	30%	7	4.6	1.83	+
	20%	7	4.5	1.81	+
	10%	7	4.8	2.09	

表 13 エレミヤ書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	229	230.8	2.66	
	20%	227	228.4	3.21	
	10%	221	220.5	3.47	
単語	30%	228	219.4	3.91	*
	20%	214	199.3	5.02	**
	10%	138	132.8	6.14	
二語連続フレーズ	30%	122	125.4	5.95	
	20%	115	117.3	6.96	
	10%	86	80.7	6.63	
三語連続フレーズ	30%	28	31.5	4.76	
	20%	28	32.0	4.21	
	10%	23	23.0	3.61	

上記の三テキストを除き、単語での出現率 10%以下と 30%以下も合わせると、少なくとも一つは 5%有意水準を満たしている。なお創世記、イザヤ書とエゼキエル書に関してはウィンドウ

法で取得した単語の対に対して有意水準を満たしている。これらより、検証仮説でのテキストの修辞構造によるテキスト箇所間の対応関係は、一つの単語あるいは少し離れて出現する二つの

関連単語の組によって構成されていると考えられよう。また単語の場合には各テキスト箇所中での出現比率 20%以下程度の単語が、テキスト箇所間の関係性を構成する単語として用いられやすい傾向にあると言えよう。ただし 5%有意水準が一つも出ていないイザヤ書に関しては、この結果のみから仮説の有意性を主張することは困難であると考えられる。

イザヤ書での単語やフレーズの共通なペアが少ない理由としては、聖書学での有力な学説によればイザヤ書は第一イザヤ（1章から39章）、第二イザヤ（40章から55章）、第三イザヤ（56章から66章）の三グループの筆者達によって個々に記されたテキストが編集されたと言われており、筆者グループの相違が語彙の共通性の低さに影響したことも考えられる。

表 14 エゼキエル書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	120	116.0	1.99	*
	20%	119	115.5	2.61	+
	10%	110	106.5	3.16	
単語	30%	116	115.8	2.11	
	20%	100	100.5	3.18	
	10%	59	58.3	4.96	
二語連続フレーズ	30%	49	48.1	4.90	
	20%	49	48.3	4.92	
	10%	36	36.9	4.63	
三語連続フレーズ	30%	17	12.1	2.93	*
	20%	17	11.8	2.74	*
	10%	11	6.4	2.32	*

#### 4.3. 他の同構造テキストでの検証

次に、新約聖書中のテキストでの同様の分析結果を表 15 から表 20 に示す。有意差の記号も先と同様に記す。

表 15 マタイ福音書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	453	437.9	7.90	*
	20%	436	424.0	9.46	
	10%	388	370.3	9.69	*
単語	30%	460	453.6	7.70	
	20%	397	395.0	9.20	
	10%	240	244.6	10.76	
二語連続フレーズ	30%	96	103.7	10.17	
	20%	96	103.4	7.84	
	10%	88	95.3	8.23	
三語連続フレーズ	30%	12	12.1	2.97	
	20%	12	12.7	3.71	
	10%	12	12.7	3.47	

表 16 マルコ福音書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	208	207.2	4.95	
	20%	200	202.3	4.35	
	10%	177	175.5	5.27	
単語	30%	216	210.4	4.61	
	20%	195	184.5	5.47	*
	10%	118	111.8	5.91	
二語連続フレーズ	30%	54	56.2	6.03	
	20%	54	57.3	5.81	
	10%	45	47.3	6.00	
三語連続フレーズ	30%	2	4.0	1.93	
	20%	2	4.4	1.82	
	10%	2	4.2	2.15	

表 17 ルカ福音書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	436	437.6	7.37	
	20%	418	419.7	8.39	
	10%	368	374.6	9.41	
単語	30%	453	445.4	7.27	
	20%	407	390.8	9.74	*
	10%	286	265.4	9.38	*
二語連続フレーズ	30%	97	108.0	8.48	
	20%	97	105.8	7.79	
	10%	67	79.5	7.77	
三語連続フレーズ	30%	4	5.4	2.04	
	20%	4	5.9	2.25	
	10%	4	5.9	2.44	

表 18 ヨハネ福音書での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	233	232.5	2.97	
	20%	226	228.0	3.36	
	10%	209	209.7	4.18	
単語	30%	211	214.8	4.13	
	20%	182	187.0	6.00	
	10%	107	97.9	6.78	+
二語連続フレーズ	30%	79	72.0	6.33	
	20%	79	71.1	6.11	+
	10%	60	53.5	5.43	
三語連続フレーズ	30%	4	5.2	2.00	
	20%	4	5.6	2.05	
	10%	4	5.3	2.00	

表より、マルコ、ルカと黙示録においては旧約聖書の場合と同様に単語の出現頻度 20%以下の閾値で 5%有意水準を達成している。マタイではウィンドウ法で 5%の水準で有意差がある。

表 19 使徒言行録での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	220	220.1	4.22	
	20%	218	218.8	4.14	
	10%	206	204.8	5.00	
単語	30%	219	220.8	3.63	
	20%	211	212.8	4.44	
	10%	143	139.3	6.60	
二語連続フレーズ	30%	67	64.3	6.23	
	20%	58	54.8	6.83	
	10%	54	54.6	5.53	
三語連続フレーズ	30%	3	2.7	1.48	
	20%	3	2.2	1.32	
	10%	3	2.4	1.36	

表 20 黙示録での仮説の検証結果

	閾値	共通ペア数		標準偏差	有意差
		検証仮説	ランダム平均		
ウィンドウ	30%	214	201.8	5.58	*
	20%	214	198.4	5.57	**
	10%	199	173.8	6.26	**
単語	30%	223	214.7	4.06	*
	20%	197	176.8	5.37	**
	10%	128	110.0	8.38	*
二語連続フレーズ	30%	90	75.5	7.23	*
	20%	90	73.7	5.65	**
	10%	80	59.2	6.34	**
三語連続フレーズ	30%	16	13.4	3.28	
	20%	16	13.3	3.27	
	10%	16	13.1	3.45	

ただし、ヨハネは 10%の有意水準のみであり、使徒言行録は有意水準を達成している項目がない。

これらより、使徒言行録を除き、検証対象の仮説は少なくとも、単語・フレーズレベルでの数値的妥当性が想定できると言えよう。ただしイザヤとヨハネに関しては 10%の有意水準しか達成しておらず、さらなる精査が必要である。

聖書の筆者は各テキストで異なるが、基本的に同じ伝統に所属しており、修辭的技法も民族の中で代々受け継がれてきたものと考えられる。そのため、大多数の聖書テキストで類似の特徴を持った有意差が検出された多層集中構造は、聖書に共通する基本的な修辭構造であった可能性があると見られる。つまり、聖書各テキストの筆者が単語やフレーズを多層集中構造に対応する形で配置した可能性がありうると言えよう。

## 5. 結論と今後の課題

本論文においては従来多数の仮説が乱立して客観的な検証の困難であった修辭構造の妥当性

に対して、数値的な検証手法を提案し、聖書における多層集中構造を検証仮説として数値的な検証を行った。結果として、検証仮説では出現頻度の低い単語とウィンドウ法での語彙の組で、共通性を持つテキスト箇所間のペアがランダムな場合より有意に多いことが分かった。よって、検証対象仮説は妥当である可能性がある。

ただし本論文で得られた結果はあくまで共通の単語・フレーズに基づく対応関係のみの分析であり、同義語や反義語やその他の種々の対応関係まで含めた妥当性の検証には至っていない。また、テキスト箇所の分割の妥当性に関しても現段階では未着手である。数値的な修辭構造の妥当性検証手法の確立のためには今後これらの課題を順次克服することが必要である。

## 謝辞

本研究は科研費「レトリカルデータベースシステムの構築による計量的修辭分析手法の確立」(22700256)の助成を受けた。

## 参考文献

- [1] 村井源, 往住彰文, “古典文献用 Digital Hermeneutics System の提案と実装”, IPSJ SIG Computers and the Humanities Symposium 2007, Vol. 2007, No. 15, pp. 101-106, 2007.
- [2] 村井源, 三宅真紀, 赤間啓之, 中川正宣, 往住彰文, “テキスト間参照情報を考慮した古典文献デジタルアーカイブ”, IPSJ SIG Computers and the Humanities Symposium 2005, Vol. 2005, No. 21, pp. 21-25, 2005.
- [3] 村井源, “デジタルアーカイブへの intra-textuality の導入に向けて”, IPSJ SIG Computers and the Humanities Symposium 2010, Vol. 2010, No. 15, pp. 131-136, 2010.
- [4] Alison Sharrock, Helen Morales, “Intratextuality: Greek and Roman Textual Relations”, Oxford Univ Pr on Demand, 2001.
- [5] 森彬, “新・聖書の集中構造”, ヨルダン社, 2001.
- [6] 村井源, “マルコ福音書の多層集中構造”, 日本カトリック神学会誌, Vol. 20, pp.65-95, 2009.
- [7] 村井源, “聖書における多層集中構造の妥当性の検討”, 日本カトリック神学会 第21回学術大会, 2009.
- [8] Hajime Murai, “The Parallel Concentric Structures within Exodus”, Society of Biblical Literature Annual Meeting, 2010.
- [9] 村井源, “聖書の修辭構造”, [http://www.valdes.titech.ac.jp/~h\\_murai/bible/bible.html](http://www.valdes.titech.ac.jp/~h_murai/bible/bible.html).
- [10] Michael S. Bushell, Michael D. Tan, and Glenn L. Weaver, “Bibleworks 8”, 2008.
- [11] The Westminster Theological Seminary, “JDP - Groves-Wheeler Westminster Morphology and Lemma Database (WTM) release 4.10”, 2008.
- [12] Michael Bushell, Jean-Noel Aletti, SJ and Andrzej Gieniusz, “BibleWorks Greek New Testament Morphology (BNM)”, 2001.