

ネットワーク分析からみた共観福音書間の比較研究

——共観表のネットワーク描画——

三宅 真紀（大阪大学大学院言語文化研究科）

mamiyake@lang.osaka-u.ac.jp

要旨：本研究では、新約聖書における共観福音書を分析対象として、単語の共起情報からネットワークを構築し、ネットワーク分析を適用する。ネットワーク構造や視覚化されたグラフ図をもとに、文書間の類似性について考察する。

キーワード：新約聖書, 共観福音書問題, ネットワーク分析

1. はじめに

グラフ理論に基づいたネットワーク分析は、WWW やソーシャルネットワークのような実世界の複雑なネットワーク体系を直感的に把握しやすいデータ解析手法として有効である。近年、自然言語データにおいても、ネットワーク分析の観点から語彙や文書間の関係性を捉える応用研究が報告されている。Gfeller et al. (2005)や Dorow, B. et al.(2005)は対象とするコーパスは違うが、曖昧性をもつ単語に対してマルコフクラスタリングを適用し、クラスタリング精度を上げる手法を提案している。Steyvers & Tenenbaum (2005)は、英語コーパスによる意味ネットワークと複雑系ネットワークの共通性を示して、グラフ理論の有用性を提唱した。

新約聖書学においては、村井・徃住（2007）が、内容が重複するテキストに対して階層的クラスタリング手法を適用し、編集的中心メッセージを抽出することに成功した。三宅（2006）は、共起情報にもとづく福音書のネットワークを構築し、テキストデータの整形処理のためのネットワーク基本特性の活用に関して報告している。さらに、三宅（2008）では、登場人物や活動場所の固有名詞に限定したソーシャルネットワークを構築し、インタラクティブに登場人物の関係性が追跡可能な Web アプリケーションとして実装した。

本稿では、共通する部分が多いとされる共観福音書に対して、ネットワーク分析の適用を試みる。ネットワークグラフ図として視覚化された単語の関係性をもとに、文書間の類似性について考察する。

2. 共観福音書の類似性

2.1. 福音書

新約聖書には、福音書、書簡、歴史書（行伝）、黙示録といった4つの文学類型に分けられた27文書が、正典としておさめられている。福音書は、キリスト教会において新しく作り出された概念であり、宣教的意味を持つ。元来、福音書を表すギリシャ語 *εὐαγγέλιον* は、特有な文学類型を表すのではなく、救済のよい知らせそのものを意味するものであったが、キリスト教会により、一つの文学類型を意味するようになった (Conzelmann & Lindermann, 1998)。

福音書は、マルコ(Mk)、マタイ(Mt)、ルカ(Lk)、ヨハネ(Joh)の4つの文書から構成され、それぞれ別の著者によって書かれたものとされる。様々な口伝伝承、文献資料を用いて叙述されており、イエスの登場・活動を描き、受難と復活で終わっている。

原文である古典ギリシャ語『ネストレ=アールントの新約聖書 (Novum Testamentum Graece) 第27版 (Nestle-Aland, 1993)』を使用して、テキストの延べ語数や異なり語数などの基本統計量に関して、4福音書をそれぞれ個別に集計した情報を表1に示す。

各文書の長さを述べ語数から判断すると、マルコ福音書が一番短く、ルカ福音書が一番長いことがわかる。ここで、異なり語数は、形態素の処理を行わない生データのままで処理を行っており、同意語の名詞においても形態が異なれば別の語としてカウントしている。

語の延べ数や異なり数に注目すると、4福音書の中でもルカ福音書が一番長く、また多くの語を使用していることが分かる。一方で、文の区切りの単位である節の長さの平均をみると、4福音書とも大きな差はみられず、平均17語程度で節が作られていることがわかる。また、4福音書の異なり語数は、8361語であった。

表 1：基本統計量

	マタイ	マルコ	ルカ	ヨハネ
章数	28	16	24	21
節数	1068	673	1149	878
1節あたりの平均語数	17.4	17.0	17.1	17.8
述べ語数	18541	11427	19696	15635
異なり語数	3944	2859	4579	2572

2.2. 共観福音書問題

福音書のうち最初の3福音書 (Mt, Mk, Lk) は、ヨハネ福音書と比較して、全体構成や内容が似通っているほか、言い回しが一致している箇所も多い。共観福音書の共通部分の比率をみると、マルコ福音書全体の約 95 パーセントが、マタイ・ルカ福音書のいずれかと共通している。その共通部分は、マタイの約 58 パーセント、ルカの約 41 パーセントに相当している。また、マタイ、ルカ福音書において、マタイ・ルカにのみ共通している部分については、それぞれ約 20 パーセントの割合である (小林, 1996)。

このように共通性が著しいことから、18 世紀以降「共観福音書 (Synoptic Gospels)」と一般に呼ばれている。18 世紀の終わりには、文書間の文学上の関係性、共通部分の要因を探り、文書間の成立上の相互関係を整合的に説明しようとする、「共観福音書問題」の提起にいたった。これまで、原福音書説、断片説、伝承説など様々な仮説が立てられ、長い間議論されている。その中で、「二資料説」は、最も説得力のある説とされている。

「二資料説」は、マルコ福音書が最も古く、マタイとルカが個別にマルコを資料として用いたと考える「マルコ優先説」を前提とし、マタイ・ルカ福音書のみ共通して現れる箇所が頻出することから、マタイとルカは、マルコ福音書とは別の資料 (Q 資料) を用いていたと想定した。つまり、マタイ・ルカ福音書は、共通の資料としてマルコ福音書と「Q 資料」の二資料をそれぞれ用いたと考える説である。

2.3. 共観表

共観表(Synopsis)は、福音書の共通箇所を並べて表にしたものであり、「共観福音書問題」を扱うにあたって欠かせないツールである。古いものでは、Griesbach (1976) の”Synopsis Evangeliorum Matthaei, Marci et Lucae (共観福音書対観表)”がある。近代聖書学研究では、Aland (1983) の Synopsis がよく参照されているが、最近では、『岩波版新約聖書』を使用した、佐藤研 (2005) の原文の基づき共通語を彩色した、『福音書共観表』が秀でていいる。

共観表の一例として、「2 資料説」で説明されうる典型的な 2 つの共通部分 (3 福音書、マタイ=ルカ共通) を含んだ並行箇所を、表 2 と表 3 に示す。いずれも、佐藤 (2005, 2006) の福音書共観表から転載した。平行箇所の小見出しは「来るべき者の告知」(マタイ 3, 11-12、マルコ 1, 7-8、ルカ 3, 15-18) の部分である。表 2 は岩波版の日本語訳共観表であり、それに対応するギリシャ語原文が表 3 にあたる。紙面の都合上、共通部分が表れない節は省略した。共通部分の色に関しては、佐藤氏の彩色原則に従い、原文の単語が一致している場合は「網掛け」にし、形態が異なる語は「2 重下線」で色をつけ

た。3福音書共通部分(Mt=Mk=Lk)は水色、マタイ・マルコ共通(Mt=Mk)は明るい緑色、マタイ・ルカ共通(Mt=Lk)は黄色、マルコ・ルカ共通(Mk=Lk)はピンク色で表している。

表 2：平行箇所「来るべき者の告知」（佐藤研訳）

マタイ 3: 11-12	マルコ 1: 7-8	ルカ 3:16-18
<p>この私はお前たちに、 回心に向け、 水によって浸礼を施している。 しかし私の後から来たるべき者は 私よりも強い。 〔私は〕皮ぞうりを 脱がす値打ちもない。 彼こそは、 お前たちに聖霊と火とによって 浸礼を施すだろう。</p> <p>彼はその箕を 手に持ち、 その脱穀場を隅から隅まで掃き清め、 その麦を倉に集めるだろう。 しかしみ殻は、消えない火で 焼き尽くすだろう。</p>	<p>そして彼は宣教して 言った、 私よりも強い者が私の後から来る。 〔私は〕その者の皮ぞうりの紐を かがんで解く値打ちもない。 この私はお前たちに、</p> <p>水で浸礼を施した。</p> <p>しかし彼こそは、 お前たちに聖霊 によって 浸礼を施すだろう。</p>	<p>〔そこで〕ヨハネは皆に答えて 言うのであった、</p> <p>この私はお前たちに</p> <p>水で浸礼を施している。 しかし 私よりも強い者が来る。 〔私は〕その者の皮ぞうりの紐を 解く値打ちもない。 彼こそは、 お前たちに聖霊と火とによって 浸礼を施すだろう。</p> <p>彼はその箕を 手に持ち、 その脱穀場を隅から隅まで掃き清め、 その麦を倉に集めるだろう。 しかしみ殻は、消えない火で 焼き尽くすだろう。</p>

表 3：平行箇所「来るべき者の告知」（原文）

Matthew 3, 11-12	Mark 1,7-8	Luke 3,16-18
<p>ἐγὼ μὲν ὑμᾶς βαπτίζω ἐν ὕδατι εἰς μετάνοιαν ὁ δὲ ὀπίσω μου ἐρχόμενος ἰσχυρότερός μου ἐστίν οὐ οὐκ εἰμὶ ἰκανὸς τὰ ὑποδήματα βαστάσαι αὐτὸς ὑμᾶς βαπτίσει ἐν πνεύματι ἁγίῳ καὶ πυρί</p> <p>οὐ τὸ πύον ἐν τῇ χειρὶ αὐτοῦ, καὶ διακαθαριεῖ τὴν ἄλωνα αὐτοῦ, καὶ συνάξει τὸν σίτον εἰς τὴν ἀποθήκην τὸ δὲ ἄχυρον κατακαύσει πυρὶ ἀσβέσῳ</p>	<p>καὶ ἐκήρυσσεν λέγων Ἔρχεται ὁ ἰσχυρότερός μου ὀπίσω μου, οὐ οὐκ εἰμὶ ἰκανὸς κύψας λῦσαι τὸν ἱμάντα τῶν ἐγὼ ἐβάπτισα ὑμᾶς ὕδατι</p> <p>αὐτὸς δὲ βαπτίσει ὑμᾶς ἐν πνεύματι ἁγίῳ</p>	<p>ἀπεκρίνατο λέγων πᾶσιν ὁ Ἰωάννης</p> <p>Ἐγὼ μὲν ὕδατι βαπτίζω ὑμᾶς: ἔρχεται δὲ ὁ ἰσχυρότερός μου οὐ οὐκ εἰμὶ ἰκανὸς λῦσαι τὸν ἱμάντα τῶν ὑποδημάτων αὐτὸς ὑμᾶς βαπτίσει ἐν πνεύματι ἁγίῳ καὶ πυρί:</p> <p>οὐ τὸ πύον ἐν τῇ χειρὶ αὐτοῦ διακαθαίρει τὴν ἄλωνα αὐτοῦ καὶ συναγαγεῖν τὸν σίτον εἰς τὴν ἀποθήκην αὐτοῦ τὸ δὲ ἄχυρον κατακαύσει πυρὶ ἀσβέσῳ</p>

表 4：平行箇所「来るべき者の告知」のテキスト情報

	マタイ 3:11-12	マルコ 1:7-8	ルカ 3:1-5
述べ語数	58	30	58
異なり語数	43	28	49

表 2,3 から、平行箇所「来るべき者の告知」は、3 文書共通部分が内容の中心部分であり、マタイ・ルカ共通箇所が続いており、マルコ文書には表れない” $\pi\upsilon\rho\iota$ (火)”に関する内容となっている。前半部分には、“ $\delta\pi\acute{\omicron}\sigma\ \mu\omicron\upsilon$ (私の後から)”といったマタイ・マルコだけに出現する単語もみられ、マルコ・ルカに共通する文章が3 文書共通文章に入り込む形をしているところもある。表 4 に、ギリシャ語の並行テキストに対して、各文書の出現単語の情報をそれぞれ示す。

3. ネットワーク分析

3.1. テキストのネットワーク表示

ネットワークは、ノードとエッジによって表すことができる。言語データをネットワーク分析に応用する場合、ノードの単位は、出現単語や章などの区分がとして考えられる。さらに文書間を比較するときは、文書全体を1つのノードをして表すことができる。エッジは、2つのノード間になんらかの関係があるときに、それらのノードを線で結んで表す。

3.2. 平行箇所の共起情報

小規模なネットワークの作成例として、2 節で取りあげた並行箇所「来るべき者の告知」をもとにして、共観表ネットワークを構築する。ここで、ノードは文書の出現単語とし、出現単語に共起する単語をエッジで結ぶ。表 2, 3 からわかるように、共観表は、文書間で一致する記事の部分なるべく並べてみられるように単語の配置が巧妙に工夫されている。この配置に従い、共観表の行を単位として共起情報の範囲を設定し、各行に同時に出現する単語を共起単語としてみなした。出現単語には、共通部分のカテゴリ情報を付随情報として加え（便宜上、Mt=Mk=Lk, Mt=Mk, Mt=Lk, Mk=Lk, Mt, Mk, Lk の順に、1-7 までの番号を単語の後に付加）、同じ単語でも出現する共通カテゴリが異なる場合は、別の単語として認識している。このようにして、共起単語ペアの頻度リストを作成し、その一部を表 5 に示す。それぞれの文書におけるペア延べ数は、マタイが 103、マルコが 50、ルカが 107 ペアであり、共観表としての全体の延べ数は、107

ペアであった。

表 5 : 共起単語ペアリスト (抜粋)

単語1	単語2	ペア頻度数
ἀγίω_1	πυρί_3	2
ἀπεκρίνατο_7	Ἰωάννης_7	1
ἀπεκρίνατο_7	λέγων_4	1
ἀπεκρίνατο_7	ὁ_7	1
ἀπεκρίνατο_7	πάσιν_7	1
ἀποθήκην_3	αὐτοῦ_3	1
αὐτός_1	βαπτίσει_1	2
αὐτοῦ_3	ἄλωνα_3	2

各文書のネットワークに関して、ネットワーク基本特徴量のノード数、平均次数、クラスタリング係数の平均値を順に表 6 に示す。ここで、次数は単語の繋がり具合を表す指標であり、Watts & Strogatz (1998)のクラスタリング係数は、ノード間の繋がり具合を表す重要な指標である。テキストのネットワーク分析では、単語の曖昧性を抽出するための指標として、クラスタリング係数を活用することもできる。

表 6 から、マタイとルカの 2 つのネットワークは、ほぼ同じ大きさであり、平均次数やクラスタリング係数からみても類似した構造をしていることがわかる。マルコは、最も小さなネットワークを形成し、そのクラスター性は極めて高いことが、平均クラスタリング係数からわかる。

表 6 : 基本ネットワーク特徴量

	マタイ 3:11-12	マルコ 1:7-8	ルカ 3:1-5
ノード数 (カテゴリー別異なり数)	50	30	53
平均次数	4.0	3.3	3.9
平均クラスタリング係数	0.74	0.93	0.77

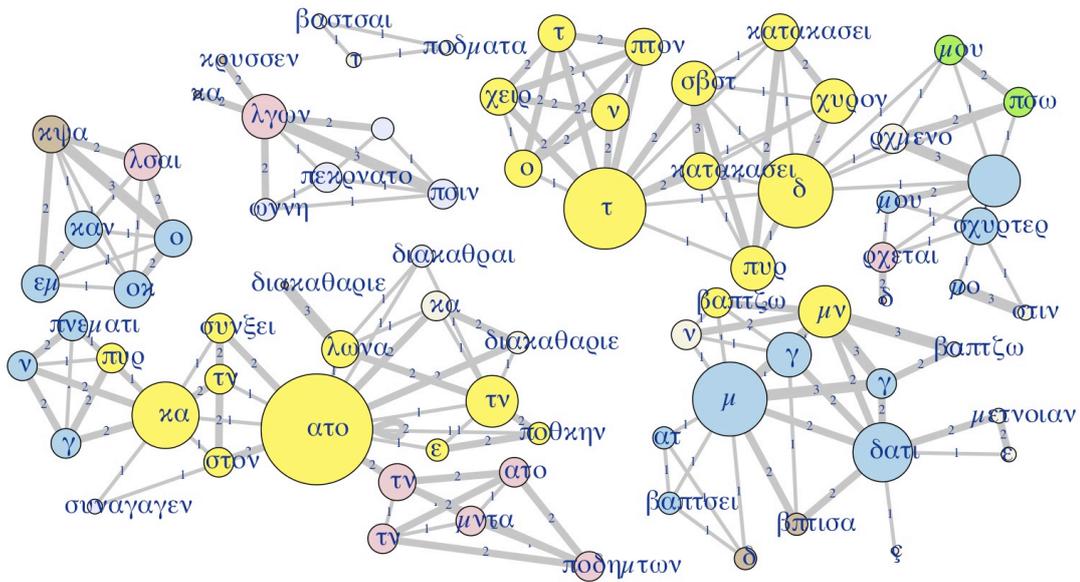


図 1：共観表ネットワーク図

3.3. ネットワーク描画

文書構造の体系をグラフ図として視覚化することにより、単語や文書の関係性を直感的に捉えることができる。ネットワーク視覚化には、統計解析ソフトウェア R の `collonet` パッケージを使用した（三宅, 2010）。このパッケージは、`igraph` パッケージを言語データ用にカスタマイズすることを目的として作成したものである。図 1 に共観表ネットワーク図を示す。共観表のカテゴリー分けに従って、ノード色を区別した。4つの共通カテゴリーに加えて、各福音書に独立して表れる単語についても文書別に識別し、文書特有の使用単語を確認できるようになっている。次数の大きさに比例して、ノードの大きさを変化している。単語と単語の結びつけているエッジの太さは、共起単語ペアの頻度数に比例している。図 1 から、3福音書共通部分とマタイ・ルカ共通部分のクラスターが分かれていることが読み取れる。そのクラスターに派生する形で、マルコ・ルカ共通部分のノードの小クラスターが繋がっている部分もあり、ネットワーク構造をもとにして、共観表の並行箇所をさらに細分化することも可能である。

4. おわりに

本稿では、共観福音書の 3 福音書に対して、共観表の並行箇所から単語の共起データを抽出しネットワークを構築する手法を提案した。共通カテゴリーの情報が付加された単語を使用したグラフはクラスター性を高め、文書の共通部分が把握しやすいネットワーク図として描画された。内容一致を重視したクラスター抽出のための単語の形態素

処理や、クラスタリング係数を閾値としたクラスター形成については、今後の課題としたい。

参考文献

- Aland, K. (1985), *Synopsis Quattuor Evangeliorum 15th edition*, Deutsche Bibelgesellschaft.
- Conzelmann, H. & Lindermann, A. (1998), *Interpreting The New Testament*, Hendrickson Publishers.
- Dorow, B. et al.(2005), Using Curvature and Markov Clustering in Graphs for Lexical Acquisition and Word Sense Discrimination, *MEANING*.
- Gfeller, D., et al. (2005), Synonym Dictionary Improvement through Markov Clustering and Clustering Stability, *ASMDA*, 106-113.
- Greisbach (1776), *Synopsis Evangeliorum Matthaei, Marci et Lucae*, Helle.
- Nestle-Aland (1993), *Novum Testamentum Graece 27th edition*, German Bible Society Stuttgart.
- Steyvers, M., Tenenbaum, J. (2005), The Large Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth, *Cognitive Science*, 29 (1) pp.41-78.
- Watts, D. and Strogatz, S. (1998), Collective dynamics of ‘small-world’ networks, *Nature*, 393:440-442.
- 小林稔 (1996), 「福音書問題」, 『現代聖書講座』, 日本基督教団出版社, 1996.
- 佐藤研 (2005), 『福音書共観表』, 岩波書店.
- 佐藤研 (2006), 『ギリシャ語彩色共観表』, <http://www.rikkyo.ne.jp/web/msato/GrSynAll.pdf>.
- 村井 源, 往住 彰文 (2007), 正典テキスト群から編集的中心メッセージを抽出するネットワーク解析法, 『情報知識学会誌』 17(3), pp.149-163.
- 三宅真紀 (2006), グラフクラスタリングに基づく共観福音書意味ネットワークの実装, 『人文科学とコンピュータシンポジウム論文集 じんもんこん-2006』, pp.161-165.
- 三宅真紀 (2008), 福音書ソーシャルネットワークにおけるコミュニティ構造の考察, 『人文科学とコンピュータシンポジウム論文集 じんもんこん-2008』, pp.319-324.
- 三宅真紀 (2010), 「ネットワーク分析のための R パッケージの開発—テキストマイニングへの応用に向けて—」, 『電子化言語資料分析研究 2009-2010』, 大阪大学大学院言語文化研究科, pp.41-63.