

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 10 日現在

機関番号：14401

研究種目：若手研究（B）

研究期間：2010～2012

課題番号：22700245

研究課題名（和文）

ネットワーク分析に基づく新約聖書校訂本の比較研究

研究課題名（英文）

A Comparative Study of the Editions of the New Testament Based on Network Analysis

研究代表者

三宅 真紀（MIYAKE MAKI）

大阪大学・大学院言語文化研究科・准教授

研究者番号：80448018

研究成果の概要（和文）：

本研究は、聖書学におけるギリシャ語新約聖書の主要な校訂本を対象にして、写本の異読情報を付帯した意味ネットワークを構築し、グラフ理論に基づいたネットワーク分析を適用した。また、共通箇所が多い校訂本に埋もれている微小な違いに焦点をあてながら、類似テキストの差異の抽出を試みた。さらに、類似性を表す統計値と TEI (Text Encoding Initiative) エンコードガイドラインに準拠して作成したデジタル批判本を組み合わせ、異本比較研究ツールを開発した。

研究成果の概要（英文）：

The study applied some network analyses based on graph theory to the influential modern editions of the Greek New Testament and constructed semantic networks with the data of the variant readings among the editions. We also developed a comparable tool combining the statistical results and the digital critical editions based on the Text Encoding Initiative Guidelines.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010 年度	1,200,000	360,000	1,560,000
2011 年度	1,100,000	330,000	1,430,000
2012 年度	700,000	210,000	910,000
年度			
年度			
総計	3,000,000	900,000	3,900,000

研究分野：総合領域

科研費の分科・細目：情報学・図書館情報学・人文社会情報学

キーワード：計量情報学、科学計量学、新約聖書学、グラフ理論

1. 研究開始当初の背景

本研究代表者は、文理融合研究の一環として、新約聖書解釈学の問題を対象とし、従来の文献学とは異なる新規な解釈方法論を、計量的な視点から提示してきた。特に、新約聖書学の様式・伝承史的批判で議論されてきた、福音書の成立の相互関係を扱う「共観福音書問題」の計量分析の結果から、テキストの類

似概念の抽出や潜在的な意味構造をネットワーク図で表すことに成功した。

さらに計量分析を深めて、計算工学からの「新仮説」を打ち出すためには、校訂本の写本情報を考慮した、文献学の本文批評に匹敵する精密な分析が必要である。

しかし、現在聖書学で参照されている校訂本は、脚注に莫大な量の写本情報を羅列して

記載されており、扱いにかなりの熟練度が要求されるばかりか、テキストから写本の体系や傾向を掴むことは難しいといった問題があげられる。

これまでに、近代に印刷された校訂本の影響関係を表す文献は多数存在しているが、参照写本の傾向や、校訂本間の類似・相違性を統計的な立場から議論した研究は少ない。

2. 研究の目的

本研究は、19世紀以降に出版された主要なギリシャ語聖書校訂本を対象として、写本間の異読情報を付帯したネットワークを構築し、校訂テキストを意味ネットワーク図で表現する。そして、グラフ理論に基づいたネットワーク分析を適用して、膨大で複雑な校訂テキストの体系をネットワーク構造として捉えて、諸校訂本の比較研究を行う。

校訂テキストと脚注の写本情報を使用して、時系列に沿った校訂本の影響関係を、Van Dongen (2000)が提唱する Markov Clusteing などのグラフクラスタリング手法を用いて、校訂本の類似性・相違性について解析する。

校訂本ネットワークの構築は、市販されているギリシャ語聖書校訂本の電子版をもとにして、TEI(Text Encoding Initiative)エンコード形式で作成する。

複雑な校訂本の体系をネットワーク構造として捉えることで、従来の文献学の方法論では困難であった校訂本の影響関係や写本の選択傾向を計量的に分析し、諸校訂本の比較を行うことを目的としている。

3. 研究の方法

諸校訂本の比較分析に適用する方法論は、情報検索におけるベクトル空間モデルに基づく類似度計算法や大規模実データ解析で用いられるネットワーク分析を適用しながら、テキスト間の構造や明らかにする。分析結果は、テキスト体系を直感的に把握しやすいようにインタラクティブなネットワーク図として視覚的に表し、出現単語の検索やテキスト間の比較が可能とする Web アプリケーションを開発した。

主要近代校訂本の比較研究には、類似するテキストに対してテキストマイニングの手法を適用した。共通する文章に埋もれている微小な違いに焦点をあてながら、類似テキストの差異の抽出を試みた。

分析テキストとして、ギリシャ語新約聖書の代表的な近代校訂本を選定し分析の対象とした。そして、新約聖書学で定本とされている Nestle-Aland のギリシャ語新約聖書を基準テキストとして使用しながら、諸本間の相違点について観察した。

ミクロな視点からのアプローチとして、各文書の特徴語彙を手がかりとして文章を比

較参照し、使用語彙の違いを調べた。

一方、マクロ的な観点からのアプローチとして、テキスト全体に対して小単位の単語類似度計算を行い、その計算結果である俯瞰図からテキストの異同部分を観察した。

4. 研究成果

意味ネットワークは、様々な設定条件のもとで複数の意味ネットワークを構築し、ネットワーク基本特性量からデータの構造を考察した。意味ネットワークの構築は、節を単位とした単語の共起情報、主要キーワード周辺の単語情報などの設定条件のもとで、一つの校訂本に対して複数の意味ネットワークを構築した。

構築した諸本ネットワークの体系・構造の特性を調査する方法論は、Steyvers & Tenenbaum(2005)が、大規模自然言語データとして WordNet(Fellbaum, 1998)から構築した意味ネットワークを用いて、実世界の複雑系ネットワークにみられる特性を確認した手法を導入した。基本統計量として、スケールフリー・スモールワールド性を明らかにするために欠かせない指標として、次数分布とクラスタリング係数を算出し、スケールフリー・スモールワールド等のネットワークの体系を示す特性を調べ、各諸本の構造の類似・差異を確認した。諸本比較研究のプロトタイプとして、小規模なネットワークデータでの比較研究を行った。

また、本研究の手法の汎用性を示すために、研究対象の新約聖書とは異なるテキストを使用して、語彙の特徴度指数による文書の特徴分析を行った。対照テキストとして、ヴィクトリア朝の女性向け定期行誌である English Woman's Journal を用いた。そして、同じテーマを扱っている項目に対して、多変量解析から分類される文書・単語間の類似性を調べた。その結果、English Woman's Journal の語彙特徴を、多変量解析やネットワーク分析を組み合わせながら抽出する手法の有効性を示した。(図1参照)

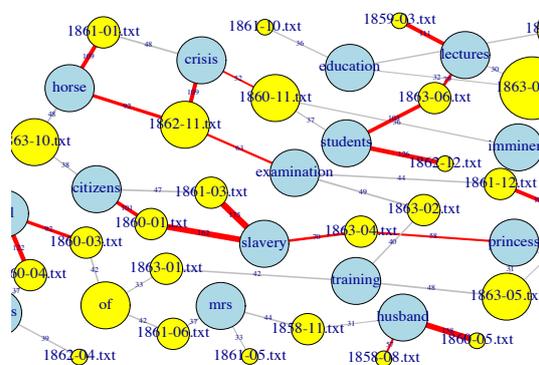


図1：特徴語彙のネットワーク図

ギリシャ語新約聖書の校訂本研究に関しては、異本比較研究のための使用可能なツールの開発を行った。

電子テキストの作成には、TEI(Text Encoding Initiative)エンコード方式に準拠した形式を採用し、テキストの計量分析から得られた諸本間の類似度結果のグラフ図とデジタル批判本を組み合わせた異本データを作成した。この構造化電子テキストは、共有資源としての意義もあり、実際に、本研究で構築したテキストの異本箇所を視覚的に表現するアプリケーション(図2参照)において実装した。

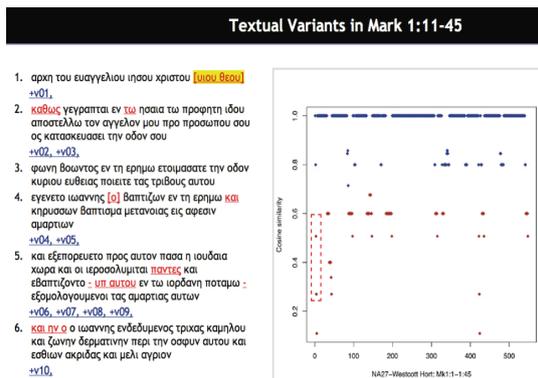


図2: 異本間比較ツール

本研究成果は、従来の文献学的方法論では考察が困難であった校訂本の複雑な異本箇所の傾向を統計指標と結びつけることで、校訂本間の類似性を視覚的に観察することを可能にした。

【参考文献】

Van Dongen (2000), Graph Clustering by Flow Simulation. PhD thesis, University of Utrecht.

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.

Fellbaum, C., WordNet: An electronic lexical database, Cambridge, MA: MIT Press, 1998.

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8件)

- [1] 三宅真紀、トークン・文字単位距離に基づく文書間の類似度計算法の考察、統計数理研究所共同研究レポート298:統計学的マイニング技術を応用したテキスト研究、査読無、(2013)、29-36

- [2] 三宅真紀、TEIに準拠した異本間比較ツールの開発、言語文化研究プロジェクト:電子化言語資料分析研究2011-2012、査読無、(2012)、19-28
- [3] Maki Miyake, Classifying Documents Using Keyness Values of Words, In Proceedings of the 11th International Conference on Textual Data Statistical Analysis (JADT 2012), 査読有、(2012)、729-736
- [4] 三宅真紀、語彙の特徴度指数による文書の特徴分析—English Woman's Journalの計量分析の多変量解析—、統計数理研究所共同研究レポート278:マイニング技術を応用したテキスト分析研究、査読無、(2012)、1-16
- [5] Hiroyuki Akama, Maki Miyake, and Jaeyoung Jung, Automatic Extraction of Hidden Keywords by Producing 'Homophily' within Semantic Networks, Digital Humanities 2011, 査読有、(2011)、71-74
- [6] 三宅真紀、クラスター分析による文書分類の類似性を探る—English Woman's Journalの計量分析に向けて—、言語文化研究プロジェクト:電子化言語資料分析研究2010-2011、査読無、(2011)、31-48
- [7] 三宅真紀、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、文字と非文字のアーカイブズ/モデルを使った文献研究、査読無、(2011)、23-30
- [8] 三宅真紀、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—、言語文化研究プロジェクト、電子化言語資料分析研究2009-2010、査読無、(2010) 41-63

〔学会発表〕(計8件)

- [1] 三宅真紀、トークン・文字単位距離に基づく文書間の類似度計算法の考察、統計数理研究所言語系共同研究グループ合同発表会:言語研究と統計2013、2013年03月27日、統計数理研究所
- [2] Maki Miyake, Developing a TEI Based Analytical Tool for Textual Variants, The 2nd Conference on JADH (The Japanese Association for Digital Humanities) 2012, 2012年09月16日、東京大学
- [3] Maki Miyake, Classifying Documents Using Keyness Values of Words, The 11th International Conference on Textual Data Statistical Analysis (JADT 2012), 2012年06月13日、リエージュ大学

- [4] 三宅真紀、語彙の特徴度指数による文書の特徴分析、統計数理研究所言語系共同研究グループ合同発表会：言語研究と統計 2012、2012年03月07日、統計数理研究所
- [5] Maki Miyake, A Pilot Analysis of Textual Variants Based on TEI - Encoding, Digital Humanities Australasia 2012, 2012年03月29日、オーストラリア国立大学
- [6] Maki Miyake, Statistical text analysis on English Woman's Journal: Investigating Keynes, Osaka Symposium on Digital Humanities 2011, 2011年09月13日、大阪大学
- [7] 三宅真紀、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—第一回シンポジウム「文字と非文字のアーカイブズ／モデルを使った文献研究」、2011年2月18日、京都大学
- [8] Maki Miyake, Mining the landscape of stylistics: A multivariate analysis of PALA papers, The 30th International Conference of the Poetics and Linguistics Association, 2010年7月22日、University of Genoa
〔図書〕(計1件)
- [1] 三宅真紀、共立出版、「13章 類似テキストの異同を計る：新約聖書校訂本の比較研究」, 『コーパスとテキストマイニング』石田基広, 金明哲(編)、(2012) 155-165

6. 研究組織

(1) 研究代表者

三宅 真紀 (MIYAKE MAKI)

大阪大学・大学院言語文化研究科・准教授

研究者番号：80448018