スタイロメトリによるコプト語文献の著者帰属の再検討

宮川創 [1]　Eliese-Sophia Lincke [2]　Heike Behlmer [3]
[1] 筑波大学　[2] ベルリン自由大学　[3] ゲッティンゲン大学
miyagawa.so.kb@u.tsukuba.ac.jp

概要

本稿では，コプト語説教文書「On Christian Behaviour」(OCB) の著者帰属を，スタイロメトリ（文体統計学）を用いて再検討する．OCB は写本題辞で修道院長シェヌーテ（Shenoute, 紀元後 4〜5 世紀）の名を冠するが，実際にはシェヌーテの特徴である修道士向けの規律に関する厳しい忠告や厳しい断罪が乏しく，早くから「偽シェヌーテ文書」と疑われてきた．本研究では，OCB の K. H. Kuhn の校訂版テキストと，白修道院文学（シェヌーテ, ベーサ, ヨハンネス）および新約聖書，コプト語聖人伝，新学書，説教などを対象に，R 言語の stylo パッケージを用いて文体比較を行った．結果として，OCB がシェヌーテの著作群よりも新約聖書の書簡のクラスタに近いという従来の示唆を定量的に裏付け，OCB を「偽シェヌーテ文書」の範疇に位置づける客観的根拠を提供した．

1　はじめに

「On Christian Behaviour」（以下，OCB）は，エジプトの白修道院で作られたとみられる長大な説教文書であり，主に 5〜6 世紀頃の成立が推定される．ニューヨーク Pierpont Morgan Library 所蔵の Codex M 604 が代表的証本であり，そこには冒頭から「アパ・シェヌーテの聖なる説教」と明記されている．一方で，Walter E. Crum や Paul E. Kahle，Karl Heinz Kuhn らが指摘してきたように ([1] の pp. v–vi 参照), OCB は修道士向けの具体的な懲戒規定や厳格な修道生活の描写に乏しく，むしろ一般信徒や聖職者の振るまいを説く平明な倫理説教に近い．こうした特徴は既知のシェヌーテの著作群（いわゆる *Canons* や *Discourses* 所収の書簡や説教文など [2]）とはかなり趣を異にするため，早くから「偽シェヌーテ文書」なのではないかとの議論が交わされてきた．

シェヌーテは 4〜5 世紀に白修道院を指揮し，多数の説教文書・書簡を残したとされる．その文体は，激しい罪の告発や繰り返しを多用した説得レトリック，独特の修道院規律への言及などで彩られ，後世のコプト教会にも大きな影響を及ぼした．しかし OCB は，こうした厳しい修道院的モチーフをあまり含まず，むしろ新約書簡に類似する愛や謙虚さ，自制を説く内容が目立つ．このため，Shisha-Halevy [3] は OCB 内にシェヌーテ起源の断片的表現を認めつつも，テキスト全体としては別の作者が編纂した複合物ではないかと論じた．一方で，Kosack [4] は OCB をシェヌーテの真作の延長にあると主張するなど，結論は必ずしも定まっていない．

本研究では，白修道院文学に分類されるシェヌーテ (Shenoute)，ベーサ (Besa)，(アパ・) ヨハンネス (Apa John) の諸テキストと，福音書やパウロ書簡・公同書簡を含むコプト語訳新約聖書，そのほかに，コプト語最大のタグ付き多層コーパス [5] である Coptic SCRIPTORIUM [6][7] コーパスがオープンソースで提供している聖人伝や神学書などのコプト語サイード方言テキスト対象として，OCB の文体がいずれのグループに近いかをスタイロメトリで分析する．具体的には，形態素単位での語彙頻度を R 言語の stylo [8] パッケージで比較し，階層的クラスタリングによってテキスト間の文体距離を可視化した．OCB が新約書簡クラスタ側に大きく寄ることが繰り返し示されれば，「偽シェヌーテ文書」である可能性を改めて強化できる．

2　研究背景

白修道院（White Monastery / Deir el-Abyad）は，上エジプト（エジプト南部）にあり，4 世紀末から 5 世紀にかけて修道院制度の発展とともに修道院長シェヌーテの下で大きく繁栄した．シェヌーテと彼の後継者にあたるベーサやヨハンネスらが記した説教や書簡は，しばしば「罪の糾弾」「修道士の行動規範」「不従順な修道士や異端への厳罰」などを強調する独自の文体をもつ．実際，シェヌーテ著作群を読むと，断罪的な語彙や律法的な表現の多用が際立つ．

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

一方，OCB は「On Christian Behaviour」という表題が示すとおり，修道士個人の生活全般よりも広く世俗聖職者や一般信徒のモラル・礼拝態度などを説き起こす内容が中心である．Kuhn [1] はこの一般教会的性格を，シェヌーテ伝統からかけ離れたものと捉え，修道院で読まれたからこそシェヌーテ名義がつけられた「偽作」と結論づけた．もっとも，Shisha-Halevy [3] は文中にシェヌーテを想起させる表現が存在するため，全編が架空ではなく，後代の編者がシェヌーテ断片を組み合わせて増補したという観点を示している．

3 方法

OCB のテキストは Kuhn [1] の校訂版を OCR にかけて抽出した，コプト語 Unicode のデジタル・テキストに基づく．この OCR は，Eliese-Sophia Lincke により RNN を用いた OCR エンジン calamari をコプト語文献のために訓練したものである [9]．この過程で出力されたデジタル・テキストのエラーを修正したのち，テキストをさらに Coptic NLP Service にかけ，語分割およびトークナイズを行った [10]．その結果を用いて，語境界に半角スペースを挿入した．同様の措置を，他の比較テキストに行った．他の比較テキストは Coptic SCRIPTORIUM プロジェクトが GitHub 上で提供しているオープンソースの TEI XML ファイル [11] から，語間半角スペース付きデジタル・テキストを抽出したものであり，次のものがある．白修道院文学としては，シェヌーテの確実な著作（例: “Shen_Thundered”, “Shen_GF”, “Shen_Abraham” など）を参照し，同修道院に関わるベーサの文書（“Besa_Nuns”, “Besa_Aphthonia” 等）やヨハンネスの文書（“Johannes_Canons”）を用いた．新約聖書文書としては福音書（“Gos_Matthew”, “Gos_Mark”, “Gos_Luke”, “Gos_John”），使徒行伝（“NT_Acts”），パウロ書簡（“P.Letter_Romans”, “P.Letter_1Cor” 等），公同書簡（“C.Letter_James”, “C.Letter_1Pet”, “C.Letter_Jude”, “C.Letter_1John” 等），黙示録（“Rev”）を全て含めた．そのほか, 修道制関連文献（“Inst_Pachomius”, “Apophthegmata_Patrum” 等），聖人伝文学（“Life_Aphou”, “Life_Eustathius” 等）およびグノーシス主義の神学書 “Pistis_Sophia” などの諸文献も加えた．

スタイロメトリは，テキストの言語的特徴（単語や形態素の頻度分布など）を統計的に比較し，類似度を可視化する手法である．ここでは R 言語の stylo パッケージ [8] を用い，Eder’s Delta による階層的クラスタリングを実施した．補助的に主成分分析（PCA）を行い，二次元上で各テキストの相対位置をプロットし，OCB がシェヌーテ著作群か新約書簡群かのどちらに近いかを確認した．

4 結果と議論

4.1 OCB の明確な新約書簡寄りクラスタリング

図 1 に示したクラスタリングの結果，大まかに「白修道院文学（シェヌーテ, ベーサ, ヨハンネス）+ 新約聖書書簡」と「新約聖書福音書・使徒行伝＋聖人伝＋その他の説教・神学書など」が二大クラスタを形成し，OCB は前者グループに属することが繰り返し確認された．特に 前者グループはさらに，「白修道院文学＋新約聖書書簡の一部」と「新約聖書書簡の大部分＋黙示録＋ OCB」のグループに分かれた．そして，「ペテロの手紙 1」や 「ヤコブの手紙」 が位置するブランチに OCB が接近する傾向が強く，シェヌーテの著作からは明らかに隔たった距離をもつ（デンドログラム上の枝長が長い）状態が観察された．これは「OCB は修道院的文体に収まらず，むしろ新約書簡風の語彙やテーマが中心である」という Kuhn [1] の指摘を定量的に補強する結果となる．Shisha-Halevy [3] が主張するように，一部にシェヌーテ的フレーズが残る可能性は否定されないが，全体的作風は明らかに異なり，後代編者による多層的編集を示唆する．

4.2 「偽シェヌーテ文書」説の評価

OCB の書名や写本題辞でシェヌーテ名義が付されている以上，かつてはこれをシェヌーテの真正作とする見解もあったが，本研究が示すような大きな文体差を考慮すると，この文書をシェヌーテの著作に組み入れることは難しいと言わざるを得ない．むしろ「修道院外での一般説教」と「新約書簡を重視する公教会的文書」をシェヌーテ名義で転用・拡張した可能性が高まる．アラビア語版 OCB や複数写本の異同にもとづく研究でも，同様の結論が支持されるか今後注目される．

加えて，シェヌーテの後継者であるベーサやヨハンネスとも大きく離れる点は，単に「シェヌーテではなく後継者が書いた」というシナリオでは説明しきれないことを示唆する．白修道院内部だけでなく，外部の修道院や教区司祭，あるいは信徒が OCB

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

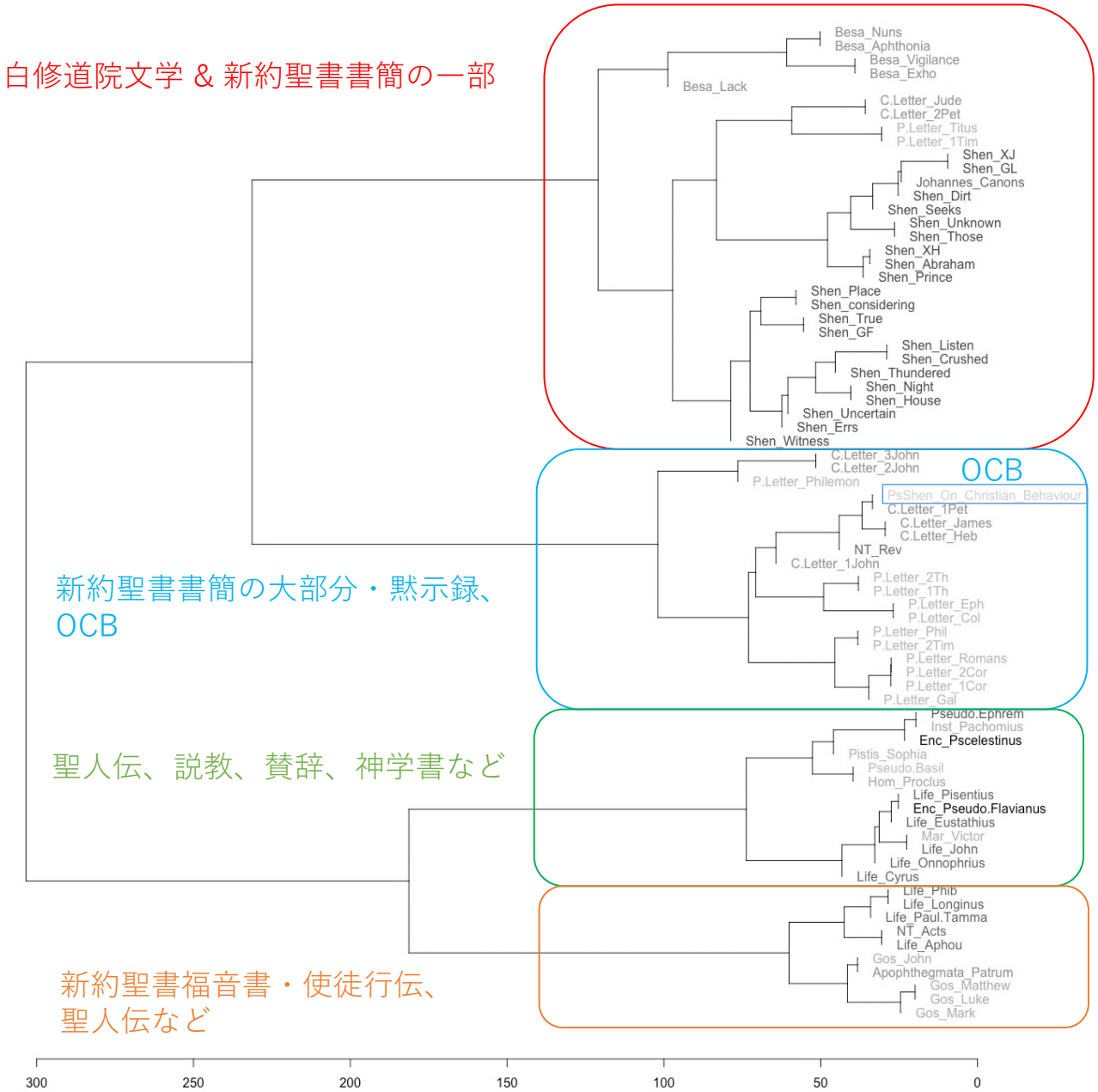

図 1　本研究で用いたコプト語諸テキストの stylo によるクラスタリング

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

を編纂・加筆した可能性もある．こうした多元的編纂プロセスを綿密に分析するためには，スタイロメトリと従来の文献学的比較（写本校訂・異読対照など）を併用することが有益である．

4.3 OCB と新約書簡の関係

OCB は，キリスト教倫理を平易に説き，教会共同体での振る舞いや悔い改めの重要性を強調する．こうした特徴は新約聖書書簡，特に「ペテロの手紙 1」(“C.Letter_1Pet”) や「ヤコブの手紙」(“C.Letter_James”) といった公同書簡の実践的・勧告的スタイルと非常に近い．実際，図 1 のデンドログラムでも，赤と青で囲った白修道院文学と新約聖書書簡のグループと緑とオレンジで囲った新約聖書福音書・使徒行伝・聖人伝・説教・新学書のグループが大きく分かれ，OCB が前者の中で青で示されている公同書簡の大部分と同じクラスタにあることが示された．また，アラビア語版 OCB（Paris BN MS 144 など）の存在は，この文書が後世においても広く流布し，一般教会説教として機能していた可能性を示唆している．対してシェヌーテの真作は，明確に修道士集団や修道院外部の敵対者に向けた内容が多く，いわゆる「修道院的世界観」が先行する．

5 結論と拡張

本研究では，OCB と，シェヌーテ著作群を含む白修道院文学，さらに新約聖書や聖人伝，説教などの様々なコプト語テキストを文体統計学的に比較し，「OCB がシェヌーテ著作群ではなく新約聖書書簡クラスタに近い」という結果を確かめた．これは従来の仮説を一段と定量的に支持し，OCB の著者性がシェヌーテから大きく外れる可能性を示す大きな根拠となる．一方，OCB にはシェヌーテ由来と思われる単語表現や言い回しが部分的に存在するため，完全に「別作者が創作した偽作」と断定するのではなく，後世の編者がシェヌーテのフレーズを引用・再編集し，新約書簡風の文書に仕立てた複合テキストという見方も残る．アラビア語版 OCB や複数のコプト語写本の編集史をさらに突き詰めれば，これらの断片をどの段階で組み込んだのか，OCB がどう伝承されてきたかが明らかになるだろう．

また，シェヌーテの全著作や白修道院文学全体のデジタル化が進めば，テキストのほぼ全域を対象とするスタイロメトリがより精緻化できるはずである．さまざまなその他の聖人伝や外典系テキストとの比較や，白修道院文学の再評価も視野に入れれば，コプト語文学全体の文体地図を描き直すことが可能となる．いずれにせよ，OCB が現在の形でシェヌーテ伝統から大きく外れ，新約書簡的性格を持つ説教として記憶されている事実は，本研究のスタイロメトリ的分析を通じて再確認された．

6 課題と今後の展望

OCB のように複数写本や言語間翻訳（アラビア語版など）が存在し，かつ著者名義が後世の編者によって付された可能性が高い文書の場合，定量分析だけで著者を断定するのは困難が伴う．今後の展望としては，以下のような点が挙げられる．

- **引用句と編集層の分析**：OCB 中の聖書引用句や，「修道院的フレーズ」「新約書簡的フレーズ」を個別に抽出し，どの部分が後世に挿入されたかをテキスト・リユース分析で追跡する．Shisha-Halevy [3] が提唱した「断片編集」説を細部で検証する材料となる．
- **アラビア語版との比較**：Paris BN MS 144 などで断片的に伝わるアラビア語版 OCB との文脈比較を行い，どの箇所が翻訳され，どの段階で書き換えられたかを体系的に調べると，編纂史に大きな手がかりが得られる可能性がある．
- **写本間の差異と年代推定**：Codex M 604 以外の写本（Bodleian B39, Berlin P.22115, BL Or.12689 など）を含め，同じ文書でも差分があるなら，スタイロメトリに適用し，「どの写本がよりシェヌーテ的・あるいは書簡的要素を強く示すか」を段階的に評価できる．そうした差異からテキストの伝承年代をある程度推測できる．
- **文体変化と時期区分**：シェヌーテ自身が長い生涯を送り，文体や執筆対象が変化した可能性を考慮し，シェヌーテ著作群を執筆推定時期ごとに分割して OCB との距離を改めて計測する．もし OCB がシェヌーテ後期文体と近いなどの結果が得られれば，解釈の幅が広がる．

これらのアプローチを組み合わせれば，OCB がいかなる経緯でシェヌーテ名義を得て，新約書簡的要素を取り込みながら広く流通していったのか，より具体的なテクスト史を描き出せるはずである．いずれにせよ，修道院文学と教会文書が交錯する 5〜6 世紀エジプトの文化的土壌を考察するうえで OCB の著者帰属問題はきわめて重要な事例である．

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

謝辞

本研究は JSPS 科研費 JP21K00537, JP23K21993 の助成を受けて行われた．Coptic SCRIPTORIUM プロジェクトをはじめ，白修道院文学をはじめとするコプト語文学のデジタル・コーパス化に尽力された多くの方々に深く感謝する．また，Pierpont Morgan Library & Museum など写本所蔵機関がデジタル画像や書誌情報を公開していることが，このような大規模比較を実現した大きな要因である．さらに，Kuhn, Shisha-Halevy, Kosack ら先行研究の知見に学んだ点は計り知れない．

参考文献

[1] Karl Heinz Kuhn. **Pseudo-Shenoute on Christian Behavior**, Vol. 206 of **Corpus Scriptorum Christianorum Orientalium**. Secréteriat du CorpusSCO, Louvain, 1960.

[2] So Miyagawa. **Shenoute, Besa and the Bible Digital Text Reuse Analysis of Selected Monastic Writings from Egypt**. SUB Göttingen, Göttingen, 2022.

[3] Ariel Shisha-Halevy. Two New Shenoute Texts from the British Library. **Orientalia**, Vol. 44, pp. 149–185, 1975.

[4] Wolfgang Kosack. **Shenoute of Atripe "De vita christiana": M 604 Pierpont-Morgan-Library New York/Ms. OR 12689 British-Library/London and Ms. Clarendon Press b. 4, Frg. Bodleian-Library/Oxford. Introduction, edition of the text and translation into German**. Christoph Brunner, Basel, 2013.

[5] Amir Zeldes. **Multilayer Corpus Studies**. Routledge, London, 2018.

[6] Amir Zeldes and Caroline T. Schroeder. Computational Methods for Coptic: Developing and Using Part-of-Speech Tagging for Digital Scholarship in the Humanities. **Digital Scholarship in the Humanities**, Vol. 30, No. suppl_1, pp. i164–176, 2015.

[7] Caroline T. Schroeder and Amir Zeldes. Raiders of the Lost Corpus. **Digital Humanities Quarterly**, Vol. 10, No. 2, pp. 1–13, 2016.

[8] Maciej Eder, Mike Kestemont, and Jan Rybicki. Stylometry with R: A package for computational text analysis. **The R Journal**, Vol. 8, No. 1, pp. 107–121, 2016.

[9] Eliese-Sophia Lincke, Kirill Bulert, and Marco Büchler. Optical Character Recognition for Coptic Fonts: A Multi-Source Approach for Scholarly Editions. In **Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage**, pp. 87–91, 2019.

[10] Amir Zeldes and Caroline T. Schroeder. An NLP Pipeline for Coptic. In **Proceedings of LaTeCH - The 10th SIGHUM Workshop at the Annual Meeting of the ACL**, pp. 146–155, Berlin, 2016.

[11] So Miyagawa, Amir Zeldes, Marco Büchler, Heike Behlmer, and Troy Griffitts. Building Linguistically and Intertextually Tagged Coptic Corpora with Open Source Tools. In Chikahiko Suzuki, editor, **Proceedings of the 8th Conference of Japanese Association for Digital Humanities**, pp. 139–141, Tokyo, 2018. Center for Open Data in the Humanities.

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).