

第125回 人文科学とコンピュータ研究会発表会

2021年2月13日（土）於：Zoom

中世日本密教資料「聖教」のNgram解析——作者の特定と思潮の流れの把握手段として

京都大学人文科学研究所
白眉特定准教授 ラポーガエタン

【1】資料紹介：「聖教」とはどのような資料か？

【2】方法論：N-gram 分析について

NGRAM=テキストにおける言葉や文字の羅列

N=言葉や文字の数（例えば、2 gram = 2 単語、2 文字）

聖教の場合、先行研究と仏教学におけるNGRAMの手法から、ワード単位ではなく、文字単位のNGRAMを使用

→Corpusや複数のテキストにおいて、ngramの頻出度を比較（単独、共通など）

【3】分析の対象：真福寺本『大師御入定記』（aka. 太神宮本地）「入定全身像」

14世紀成立

奇妙な弘法大師空海像の存在

真言密教と中世神道を融合した内容

文観とは：

西大寺出身僧、1316年頃醍醐寺に入る

後に、真言僧として、後醍醐天皇の側近となる

後醍醐天皇と南朝のため、様々な儀礼を作り、積極的に宗教政策に関与

文観=『太神宮本地』の作者？

現存する写本では、書写者（石山寺の光賢）が分かるが、作者不詳

歴史状況の考察によって、文観（1278～1357）という真言僧が作者である可能性

従来の当該分野での作者不詳の場合の推定方法：多数の関連テキストを読了した専門家の経験と、関連人物の史的状況を考慮にいたした上での状況的な個別的判断に基づく

- ・歴史的な背景：文観の弟子（宝蓮）と『太神宮本地』の書写者（光賢）の関係
- ・文観は、既に『御遺告大事』（1327年）において、奇妙な大師像を創造するなどの実績有
- ・『太神宮本地』の内容も、文観の著作と共通する



テキスト解析による経験的類推の検証

【4】データの準備：Character ngramを行う上での資料準備

- ① Corpusの断定：空海の著作/平安時代の真言密教テキスト/文観の著作と中世神道のテキスト/真言密教の基本文献/平安時代のテキストと同時代の僧侶によるテキスト

(歴史的検証によって、文観と『太神宮本地』の関係が既に推察)

- ②翻刻テキストのデータ化：異体字を正字として打ち込む。
- ③元の写本において、日本語の略字が用いられている場合には、旧漢字に打ち込み直す。訓読点等を削除
- ④文字エンコーディングを Unicode (UTF-8) に変換
- ⑤梵字を解読して、相当する漢字に変換／或いは、Unicodeにある程度の総合性のあるフォントを使う

【5】N—GRAMの分析方法

Marcus BingenheimerとJen Jou Hengによって開発された方法：もともと、中国仏教における翻訳者の判定に使う

Variable length ngram “変動式 ngram”

長いngramに既に含まれている短いngramの累計における重複を避けるための方法

Gram	Text1	Text2	Text1	Text2
三尊合行法	2	0	2	0
三尊合行	2	0	0	0
三尊合	2	0	0	0
尊行法	2	0	0	0
三尊	7	3	5	3
尊合	2	0	0	0
合行	3	1	1	1

頻出 (variable length ngram)

Ngram	御遺告秘決 簡略本	御遺告大事	御遺告秘決	秘密源底口決
浄菩提心	2	5	1	1
浄菩提心寶珠	0	6	1	0
菩提心論	1	3	0	0
菩提心論云	0	0	0	1
菩提	0	2	1	2

頻出度

頻出をテキストの文字数で割ったもの（利便性のため、x100にした）

Ngram	御遺告秘決 簡略本	御遺告大事	御遺告秘決	秘密源底口決
浄菩提心	0.09881423	0.05609783	0.01385617	0.025445293
浄菩提心寶珠	0	0.0673174	0.01385617	0
菩提心論	0.04940711	0.0336587	0	0
菩提心論云	0	0	0	0.025445293
菩提	0	0.02243913	0.01385617	0.050890585

【6】データの可視化

IBMの「SPSSStatistics」によるComponent式グラフ

(テキストの間のそれぞれのngramの頻出度の比較によって計られたテキストの距離)

Factor dimension reduction with a varimax rotationを使用

[Hung, Bingenheimer, and Wiles 2010を参照]

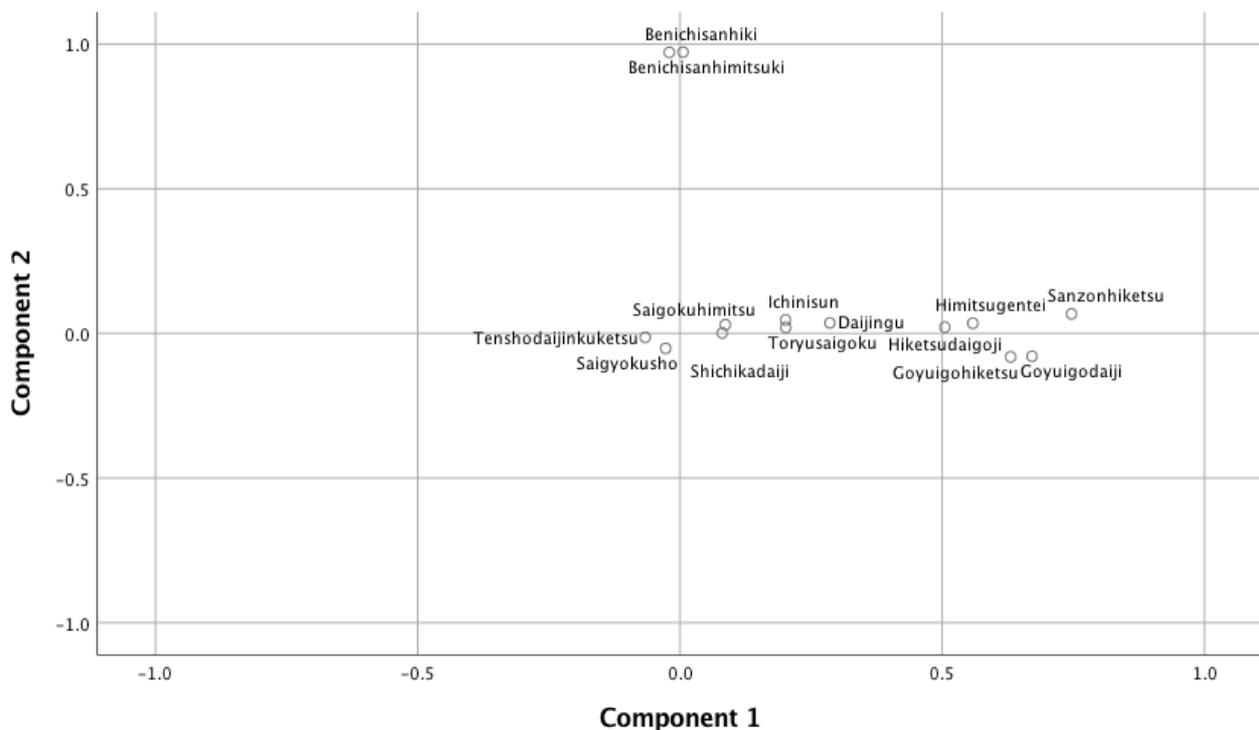
Bingenheimer・Hungの方法論を使い、二つのグループを比較

→結果『大神宮本地』：他の真言のテキストや中世神道よりも、文観の著作に近い

最後の段階の図：

グループ①『一山秘記』・『一山秘密記』・『天照太神口決』という三つの中世神道書

グループ②：文観の著作



→『大神宮本地』：文観の著作に近い

(『天照太神口決』は、西大寺の僧侶に作成され、文観と同時代)

結果として、解析によって、歴史的考察を立証する材料を提供できると分かる

【7】他のNGRAM分析方法論とツールとTEIとの併用の可能性

Michael Radichによる中国仏教における翻訳者

従来蓋然視されてきた宗派感の見直しを客観的に提唱できる

TEIでmarkupを使用することによって、テキストの各部分の性質(例: speaking voice)を考

慮にいたれた制度の高い分析

日本の宗教学において：神道と密教の関連性の証明（従来は、垂迹などの曖昧な流れで捉えられてきたもの）

仏教学における思想史の流れを解明

【8】主な参考文献

ラポー・ガエタン (2020) 「『太神宮本地』 解題・翻刻」、阿部泰郎・大東敬明編『真福寺善本叢刊〈第三期〉神道篇 第4巻 中世神道資料集』（臨川書店）。

Rappo, Gaétan. (2018) “Data Mining in the Works of the Shingon Monk Monkan (1278–1357): Using Digital Methods to Assess the Contested Authorship of Three Religious Texts From Medieval Japan.” IN: *Journal of the Japanese Association for Digital Humanities* 3, no. 1 : 114–149.

Radich Michael and Anālayo Bhikkhu. (2017) “Were the Ekottarika-āgama and the Madhyama-āgama Translated by the Same Person? An Assessment on the Basis of Translation Style.” In Dhammadinnā (ed.), *Research on the Madhyama-āgama* (Taipei: Dharma Drum Publishing Co.): 210-237.

Bingenheimer, Marcus, Jen-Jou Hung, and Cheng-en Hsieh. (2017) “Stylometric Analysis of Chinese Buddhist Texts - Do Different Chinese Translations of the Gaṇḍavyūha Reflect Stylistic Features That Are Typical for Their Age?” IN: *Journal of the Japanese Association for Digital Humanities* 2 : 1–30.

Hung, Jen-Jou, Marcus Bingenheimer, and Simon Wiles. (2010) “Quantitative Evidence for a Hypothesis Regarding the Attribution of Early Buddhist Translations.” IN: *Literary and Linguistic Computing* 25, no. 1: 119–34.

Stamatatos, Efstathios. (2009) “Intrinsic Plagiarism Detection Using Character n-Gram Profiles.” IN: *PAN-09: 3rd Workshop on Uncovering Plagiarism, Authorship and Social Software Misuse and 1st International Competition on Plagiarism Detection. Proceedings of the SEPLN'09 Workshop on Uncovering Plagiarism, Authorship and Social Software Misuse*, edited by Benno Stein, Paolo Rosso, Efstathios Stamatatos, Moshe Koppel, and Eneko Agirre, 38–46. <http://ceur-ws.org/Vol-502/>.

師茂樹 (2001) 「XML と NGSM によるテキスト内部の比較分析実験—『守護国界章』研究の一環として」 IN: 『漢字文献情報処理研究』 2: 62–67.