

オセアニアにおける神話・民族誌テキスト分析の試み

後藤 明
(南山大学)

はじめに

最近、新型コロナウイルスに関する多数の論文テキストを A.I.を使って分析させ、キーワードの集積を見つけ出すような分析が行われたとの報道があった(NHK のニュース)。この原理は特定の言語コーパス(例 文、段落、論文、文学作品、等)の中に共存する語彙から、関連性の強い概念を引き出そうとするものである。このような分析はテキストの中に埋もれた語彙の関係を掘り出すという意味でテキストマイニング(mining=鉱山発掘)と呼ばれる(小林 2017; 牛澤 2018; 末吉 2019)。

文学研究では文の長さ、あるいは句読点の位置や頻度は書き手の癖を反映するので、作詞の個性を分析したり、作者不明の作品の作者を推測するために統計的な手法が使われてきた。英語の場合は使用されるアルファベットの頻度などの分析から作家の関心や価値観の変化などを考察する研究も行われている(小林 2017)。

人類学の文脈に置き換えると、テキストマイニング分析の可能性のひとつは、学史的な目的から民族誌や研究書を研究者が残した「作品」として分析することであろう。たとえばある研究者が初期と晩年ではどのように興味関心が変わったかを探る素材としてである。対象とした地域やテーマ(例 生業 vs. 宗教)が異なれば、使用される語彙、特に名詞は異なるであろうが、動詞や形容詞の使い方の違い方に関心の変化、あるいは日本語の場合は助詞や助動詞に特徴が現れる可能性がある(山田 2014; 小林 2017)。

近年、神話や民話テキストの計量分析にこのような手法が試みられ、『数学が神話と出会う』など論集が出されている(Kenna *et al* 017)。その中ではトンプソン・アールネ・ウター(Thompson -Arne-Uther)の神話・民話インデックス(ウター 2016)の語彙分析、古典的叙事詩(例 『イリアス』、『オディッセウス』、『ベーオウルフ』等)のテキストに読み取れる対人関係の分析などが試みられている。

さらに J.G.フレイザーのテキストを人類進化との関連で分析するような試みがなされている(D'Huy 2019; Thuillard *et al* 2018)。ここではフレイザー書の批判は繰り返さない。「神話としての神話の価値は、どんなに翻訳が悪かろうと、依然として存続する」(レヴィ=ストロース 1972: 233)と断言できるかはわからないが、一人の研究者が世界の事例をある視点から集成し、英語という一つの言語で記載したテキストの価値が今日再評価されているのである。

本稿ではオセアニア関係の人類学テキストのテキストマイニングを使った試験的分

析結果を示す。その前に背景を簡単に述べておきたい。

近年、現世人類の移動に関する学際的な議論が盛んに行われている。日本では理化学研究所（以下、理研）で「サピエンス学：文明的人間の起源」という研究グループが形成され、人類の脳構造や認知能力の進化と言語、神話あるいは神観念の発達などの議論が始まっている(後藤 2021a)。本稿は、この研究グループに参加する過程で筆者が行ってきた神話および民族誌テキスト分析の途中経過報告である。

無文字社会の神話自体は考古学的遺物としては残らないが、地球上の離れた地域に存在する神話の類縁性について、初期人類の移動の結果とする仮説が提唱され、遺伝学の分析結果との対応が問題にされている。たとえばアフリカのサハラ南部とオーストラリア・アボリジニ神話などにみられる共通性について Gondwana(Gondwana)型神話(後藤 2017)、あるいはメラネシアとアマゾンの神話の共通性にはメラゾニア(Malazonia)型神話などの名称も使われている(Krotayev and Khaltourina 2011)。このような仮説が成立するためには旧石器時代後期に遡る共通の神話の存在が必要であり、そしてその仮説の検証にはオセアニアが鍵となる地域のひとつなのである。

このような古層神話として天体に関する神話があげられる(大林 1999a; 後藤 2017)。たとえば位置関係を変えず天空を移動する星座は、永遠に獲物を追う狩人と獲物である、という宇宙狩り(cosmic hunt)神話である(e.g. D'Huy and Berezkin 2017)。このような問題意識から本稿ではまず、アボリジニ天体神話の英語テキストの試験的分析を報告する。

さらにオーストロネシア世界は人類学、とくに文化進化論の議論において鍵を握る地域でもあった。かつて日本において、国立民族学博物館（以下、民博）では「東南アジア・オセアニアの文化クラスター」というデータベース構築が試みられたが(大林・杉田・秋道 1990)、近年はニュージーランドの研究者が中心となってオーストロネシア世界の宗教関係のデータベース Pulu が構築された(Watts *et al* 2015a)。また道徳的高神 MHG(moralizing high gods)や広義の超自然的罰 BSP(broad supernatural punishment)のような特徴と社会的側面との関連が分析されている(Watts *et al* 2015b)。筆者は同じような目的のために棚瀬襄爾氏による日本語テキスト(棚瀬 1966)の分析も行っている。この課題は理研プロジェクトと同時に、現在進行中の新学術領域研究科研や民博の共同研究(謝辞参照)とも親和性が高く、適宜報告を行ってきた。

オセアニア神話テキストの分析

テキストとして *A Tale-Type Index of Australian Aboriginal Oral Narrative* (Waterman 1987)を使った。この文献はアボリジニの神話のインデクスと同時に、各神話の数行程度の要約が掲載されている。

分析にあたりこの文献を OCR によってデジタルテキスト化をした。そして宇宙創

月には stone という語彙が伴う一方、太陽に earth および night という語彙の頻度が高い。太陽に night が伴うのは「(太陽は)夜には隠れる」という場面が強調されるからであろう。また頻度は多くないが太陽には bandicoot、月には hawk、kangaroo、opossum などの動物が伴い、さらにこれらの動物が排他的に出現するのに興味を湧く(図2)。

動詞においては come、become、go などが共通して頻度が高いが、これは一般的な動詞なので特徴を反映していないようだ。一方、月には kill や steal また throw や refuse あるいは drown(=洪水モチーフに伴う)という動詞の頻度が高い。これらの動詞は価値観を推定するネガポジ分析では「ネガティブ」と判断される傾向があり、月にはネガティブな語彙が太陽よりも多いといえる。

形容詞では太陽に hot が伴うは納得できるが、同時に cold や dark という語彙が伴うのはイメージと逆である。神話的要素の属性が現実と逆転ないし反転することはよくある現象であるが、ここでは太陽が夜の間どこを旅するかなどが語られるからであろう。

このようにアボリジニ神話における太陽と月に対する価値づけ、あるいはイメージの違いについての問題提起がなされ、あらためて原テキストへの回帰が促される。無論、これは著者が神話を英語に翻訳し、要約したテキストに関する限りの分析方法であり、人類学が得意とする個々の語りを現地語に即して詳細に分析する方法と対立するものではない。

今回、月は 38 話、太陽は 24 話のサンプルであるので、この量のデータなら従来のように表を作って粗筋や語彙の比較をすれば同じことがいえるであろうが、多数の神話テキスト比較をするためには、ある程度有効な方法ではないかと考える。

東南アジア・オセアニア民族誌の計量分析

次に、棚瀬襄爾氏の日本語テキスト『他界観念の原始形態』(1966)の分析を行う。棚瀬テキストの特長は：

- (1). 東南アジアからオセアニアまでの民族事例を一つの言語で要約している、希有なテキストである(ネグリティ、アボリジニとミクロネシア以外のオーストロネシアを含む)。
- (2). 葬法に関する以下の具体的データがある。
 - (a) 埋葬法：埋葬(土葬)、火葬、台上葬、洞窟葬、等)
 - (b) 遺体の格好(伸展葬、屈葬、蹲踞葬、等)や頭位方向
 - (c) 副葬品や持ち物の毀損、遷居などに関する記述。
- (3). (2)のような考古学的に観察可能な項目と「他界観」(例 死後の世界、等)のように直接観察できない観念との相関を見ることができる。

以上により、このテキストは現世人類の東南アジア島嶼部からオセアニアへの進出といった人類史の問題を論ずる上で基調な資料といえる。また(3)から、この資料は民族考古学の基礎資料としても再評価すべきであろう(後藤 2021b)。

従来、文字テキストの内容のような質的データは、何らかのコードを作って、その「ある・なし」を1/0のような形で入力し、統計処理するのが普通だった。この種の研究でオセアニア研究における先駆的なものは、大林太良が中心となって民博で行った「文化クラスター」研究である(大林・杉田・秋道 1990)。この研究ではマダガスカル、台湾、東南アジア、オセアニアにかけて 237 集団における、486 の文化要素の存否が入力された。文化要素の存否は過去と現在の両方が考慮された。

民博の「文化クラスター」研究では文化要素の分布をモニター上に表示するというデータベース化の他に、文化要素の存否から民族集団間の距離あるいは類似度を算出し、その値から多変量解析を行って、民族集団のグルーピングが行われた。そして東南アジア島嶼部とオセアニアにおけるオーストロネシア系と非オーストロネシア系という言語集団の対比が文化要素から出てくるグルーピングと対応するのか、「メラネシア」という文化複合が規定できるのか、などが議論された(大林 1999b)。

分析の目的と手順

テキストマイニングが有効に使われる事例として、自由記述アンケートの分析がある。たとえばスマートフォンを選ぶ基準をアンケートして、その結果を性別や年齢を基準にどのような語彙が集積するかを見るのである。年齢によって「機能」「デザイン」「最新」などと「値段が安い」「使い方が簡単」など、異なった語彙表現が集積する傾向が見つかれば、年齢に合わせた宣伝を行うための戦略が導かれる(牛澤 2018; 末吉 2019)。性別や年齢(=外部変数)を地域や民族集団に置き換え、自由記述アンケートを民族誌の記述あるいはフィールドノートの記載に置きえれば、民族誌分析に応用できるのではないかと筆者は考えた。

本稿で分析の対象にしたのは 17 枚の付表のテキストである(図 3)。表では民族集団ごとに、哀哭、死体、葬法、体位・向位、死者の財物、副葬品、死穢・払穢、服喪、定時儀礼、死因・病因、生霊、死霊、他界への出発・旅、帰来、他界、他界の性質、再生・転生、その他、資料(引用)の記載がある。この表を OCR によってデジタル化を行った。

分析に入る前にいくつか検討事項があった。まず表の作り方が独特である。民族集団名を単位にして、A はオーストラリア、M はメラネシア、N はニューギニア、P はポリネシア、E は東インド諸島として、各集団に数字をつけてある。たとえばオーストラリアであれば、民族ごとに A1 とか A2 など順序に他界観念の特徴が記載されている。

このような立場から、民族集団を基本的な単位として、哀哭から死因・病因までを「埋葬法」、生霊から再生・転生までを「他界観」としてセルを結合した。すなわち、各民族集団には「埋葬法」と「他界観」の二つのテキストがともなうように処理した。

このあと前処理として形態素分析を行い¹、その後分析をしたが、いくつか予想外の変な結果が出てきた。たとえば意味のありそうな語彙に紛れて「む」という語彙が抽出された。原表に戻って検討すると、それは「遺体を穴に埋む」とか「副葬品を墓に収む」という表現の「む」が、現代日本語を分析するために開発された形態素分析ソフトでは名詞として抽出されてしまった。あるいは「死者の骨を母親がしばらく持ち歩く」のような場合の「母」は意味ある単位であるが、「母指縛って埋葬」とか「真珠母貝を副葬する」の「母」が同じ意味の形態素として抽出されてしまった。

試験的分析を繰り返し、前処理としてこれらを適当な表現になるように「置換」するなどのデータクレンジングを行った²。

¹ テキストマイニングを施すための前処理として形態素分析が必要となる。日本語の場合はフリーソフトのメカブ(MeCab)や茶筌(ChaSen)を使ってそれを行うことができる。たとえば、「私は仙台市で生まれた人類学者です。」を MeCab で分析すると：

| | |
|-----|--------------------------------|
| 私 | 名詞,代名詞,一般,*,*,*,私,ワタシ,ワタシ |
| は | 助詞,係助詞,*,*,*,は,ハ,ワ |
| 仙台 | 名詞,固有名詞,地域,一般,*,*,仙台,センダイ,センダイ |
| 市 | 名詞,接尾,地域,*,*,*,市,シ,シ |
| で | 助詞,格助詞,一般,*,*,*,で,デ,デ |
| 生まれ | 動詞,自立,*,*,一段,連用形,生まれる,ウマレ,ウマレ |
| た | 助動詞,*,*,*,特殊・タ,基本形,た,タ,タ |
| 人類 | 名詞,一般,*,*,*,人類,ジンルイ,ジンルイ |
| 学者 | 名詞,一般,*,*,*,学者,ガクシャ,ガクシャ |
| です | 助動詞,*,*,*,特殊・デス,基本形,です,デス,デス |
| 。 | 記号,句点,*,*,*,。,,。 |

という結果が瞬時に出力される。一方、英語のように単語が分離して記載される言語では、形態素の分離は比較的容易である。

² 原データを変更しないためには「My 辞書」にそのような語彙を登録して適切に処理することができる。たとえばスマートフォンに関する自由記述アンケートで出てくる「スマートフォン」と「スマホ」を同じ語彙だと登録することができる。

これらの問題はわかりやすいが、やっかいなのはテキストに潜む半角文字がプログラミング上の「メタ文字」だと解釈されるときである。このようなデータの「ゴミ」は、正規表現に対応するエディターなどで削除や置換する必要がある。このような作

- (2). マレー、ニューギニア、メラネシアは地域的に近接しているためか、南島語族か否かに関わらずかなり密な集積が見られる。ネグリトはマレー系集団と長い間の接触のため、このグループの影響を強く受けている結果となったようだ。そしてこのグループはいわゆる「首狩り文化複合」(山田 2015)に重複し、遺体展示(あるいは顕示)風習とも相関する傾向が強い。
- (3). 首長制が発達するポリネシアは、埋葬法は比較的地味で、骨を秘匿したり、モニュメント的な施設を作らない傾向があるが、特に他界観は他の集団から突出する傾向がある。アボリジニとポリネシアが他の集団とやや離れるのは世界神話学仮説と関連する可能性がある(後藤 2017)。

この分析結果はあくまで棚瀬氏が多数の異なった言語の民族誌を集成して日本語で要約した限りにおいてである。このような結果は結論ではなく、遡及すべき原著の再検討や新たな視点の開拓につなげるべきだろう。

今回の地域分けも問題である。たとえばメラネシアと一括した地域を近オセアニア(Near Oceania)と遠オセアニア(Remote Oceania)などに区分して解析する、また東南アジアの集団の区分も別途行って分析を試みるべきだろう。

おわりに

筆者がハワイ大学大学院で言語学を学んでいた1980年代、まず学ばされたのはチョムスキーの生成言語学(変形文法)であった。それは少数の法則の変形で多様な言語の文法や意味を説明する演繹的モデルであった。一方、近年の自然言語処理ではむしろ多数の事例から帰納的に進化する「深層学習」の考え方が主流となり、それに基づく機械翻訳技術が格段に進歩した(黒橋 2019)。このような技術を背景にしてテキストを分析し特徴が把握できるのであれば、神話・民話テキストだけではなく、民族誌や人類学的文献の分析にも応用できるのではないかと筆者は考えた。

本稿で示したような、棚瀬テキストに依拠した分析結果は、他の文献でも同じような結果が出るかを検証する必要もある。棚瀬テキストにはミクロネシアが含まれていなかった。一方、東南アジアは入っていないが類似のテーマを扱った文献にフレイザーの『不死の信仰と死者の崇拜』の3巻本がある(Frazer 1913-1924)。その第1巻はオーストラリア・アボリジニとメラネシア、第2巻がポリネシア、第3巻がミクロネシアである。1、2巻はeBookテキストがフリーで入手でき、第3巻もOCRによるデジタル化を完了している。また同様に大林太良著『銀河の道・虹の架け橋』(1999a)のデジタル化が終わっているので、今後これらのテキストの分析も進めていきたい。

<謝辞>

本稿で分析したテキストは、理化学研究所の理事長裁量経費に依ってデジタル化した。特にテキストの処理については中分遥氏(高知工科大学)と佐藤浩輔氏(明治大学)のご協力を得た。そして本稿の内容は理研のプロジェクト「サピエンス学・文明的人間の起源」(代表・入来篤史氏)、新学術領域科研「出ユーラシアの統合的人類史—文明創出メカニズムの解明」(代表・松本直子氏)、および民博の共同研究会「島世界における葬送の人類学—東南アジア・東アジア・オセアニアの時空間比較」(代表・小野林太郎氏)で漸次、途中経過を報告してきた。これらの研究会に関係するすべての研究者に御礼申し上げたい。

<参考文献>

D'Huy, Julian

2019 Folk-tale networks: a statistical approach to combinations of tale types.
Journal of Ethnology and Folkloristics 13(1): 29-49.

D'Huy, Julian and Yuri Berezkin

2017 How did the first humans perceive the starry night?: on the Pleiades.
The Retrospective Methods Network, Newsletter 12/13: 100-122.
(researchgate.net/publication/321979171_2017_How_Did_the_First_Humans_Perceive_the_Starry_Night_On_the_Pleiades_-_The_Retrospective_Methods_Network_Newsletter_12-13_100-122)

エリアーデ、ミルチア

2013 『アルカイック宗教論集：ルーマニア・オーストラリア・南アメリカ』、
国書刊行会。

Frazer, James G.

1913-1924 *The Belief in Immortality and the Worship of the Dead*. 3 Volumes.
MacMillan.

後藤 明

2017 『世界神話学入門』、講談社現代新書。

2021a 「オセアニアへの人類進出と認知論的構造」『科学』91(2): 171-172

2021b 「島嶼世界の埋葬法と他界観：民族誌情報のテキスト・マイニング的分析の
試み」新学術領域科研「出ユーラシア文明の人類史」全体集会発表ポスター。

樋口耕一

2020 『KH Coder 3 リファレンス・マニュアル』(ソフト付属のPDF マニュアル)

Kenna, Ralph, M. MacCarron and P. MacCarron (eds.)

2017 *Maths Meets Myths: Quantitative Approaches to Ancient Narratives*.
Springer.

小林雄一郎

2017 『Rによるやさしいテキストマイニング』、オーム社。

Krotayev, Andrey, and D. Khaltourina

2011 Genes and myths; which genes and myths did the different waves of the
peopling of Americas bring to the New World? Translation of Chapter 3 of
Myths and Genes: A Deep Historical Reconstruction, pp. 1-63. Moscow.
(https://www.sociostudies.org/almanac/articles/genes_and_myths-which_/)

黒橋禎夫

2019 『自然言語処理』、NHK 出版。

- レヴィーストロー、クロード
 1972 『構造人類学』、みすず書房。
- 大林太良
 1999a 『銀河の道・虹の架け橋』、小学館。
 1999b 「オーストロネシアの文化複合」中尾佐助・秋道智彌(編)『オーストロネシアの民族生物学：東南アジアから海の世界へ』、pp.37-54.平凡社
- 大林太良・杉田繁治・秋道智彌(編)
 1990 『東南アジア・オセアニアにおける諸民族文化のデータベース作成と分析』、『国立民族学博物館研究報告別冊』11。
- 末吉美喜
 2019 『テキストマイニング入門：Excel と KH Coder でわかるデータ分析』、オーム社。
- 田村光平
 2020 『文化進化の数理』、森北出版。
- 棚瀬襄爾
 1966 『他界観念の原始形態：オセアニアを中心として』、東南アジア研究双書1。
- Thuillard, Marc, J-L. Le Quellec, J. D'Huy, Y. Berezkin
 2018 A large-scale study of world myths. *Trames* 22(72/67): 407-424.
- ウター、ハンス=イェルク
 2016 『国会昔話話型カタログ：分類と文献目録』、小澤昔ばなし研究所。
- 牛澤賢二
 2018 『やってみようテキストマイニング：自由回答アンケートの分析に挑戦!』、朝倉書店。
- Waterman, Patricia P.
 1987 *A Tale-Type Index of Australian Aboriginal Oral Narrative*. Academia Scientiarum Fennica.
- Watts, J., O.Sheehan, S.J.Greenhill, S.Gomes-Ng,Q.D.Atkinson,J.Budulia,and R.D.Gray
 2015a Pulu: database of Austronesian supernatural beliefs and practices. *PlosOne*: 1-17.
 (<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0136783>).
- Watts, J., S.J. Greenhill, Q.D. Atkinson, T.E. Currie, J. Bulbulia and R.D. Gray
 2015b Broad supernatural punishment but not moralizing high gods precede the evolution of political complexity in Austronesia. *Royal Society Proceedings B* 282: 1-7. (https://www.researchgate.net/publication/273122350_Broad_supernatural_punishment_but_not_moralizing_high_gods_precede_the_evolution_of_political_complexity_in_Austronesia)
- 山田仁史
 2015 『首狩りの宗教民族学』、筑摩書房。
- 山田敏弘
 2014 『あの歌詞は、なぜ心に残るのか：J ポップの日本語力』、祥伝社。