

神話のトピック構造分析による比較研究

東北大学工学部 情報知能システム総合学科 太田 明

東北大学 文学研究科 山田 仁史

東北大学 情報科学研究科 橋本 和夫

1 はじめに

神話が古代の未開社会において占める位置の重要性は、現代からは考えられないほど大きいものであった。故に神話は古代の思想や文化を知る鍵となり得るものであり、そのような神話を言語学、文献学、民族学などの方法を用いて研究分析するのが神話学という学問である。特に神話学において比較は極めて重要な分析手法である。しかし比較という作業は極めて知的かつ意味的であり、資料そのものと同時に研究者個人の能力に依存する部分が大きく、誰にでも共有できる視座を提供することは難しい。本研究はそのような状況において、客観的かつ定量的に神話を比較することを試みる。

クロード・レヴィ＝ストロース (1908-2009) の構造主義的神話研究は、神話を構成要素に分解し、それらの関係が作り出す構造に重点を置いた。我々は、この神話の構造という概念と相性がいいと考え、トピックモデルの概念に基づいた分析手法を用いることにした。今回の実験ではトピックモデルのうち、Multiscale Topic Tomography (MSTT) というモデルを用いた。

2 トピックモデル

2.1 トピックモデルの原理

トピックモデルは統計言語学的なテキスト解釈モデルである。これはまずトピックという抽象概念を設定する。テキストに複数のトピックが確率的に分布し、さらにトピックから単語が確率的に生成されると仮定し、ベイズの法則に基づいて、テキストに分布する単語から逆にどのような確率分布のトピックが存在しているのかを推測する。

2.2 Multiscale Topic Tomography

MSTT は Latent Dirichlet Allocation (LDA) という、トピックを語彙に対する多項分布と定義するモデルに、巨視的な時間の概念を持ち込んだものである。しかし今回の実験で扱う文書データは単一のドキュメントであり、コーパスを成していないので、この巨視的な時間要素は無視し、基本的には LDA として使用している。

3 比較実験

3.1 比較対象となる神話とその説明

本研究では、『日本書紀』に記録された保食神 (ウケモチノカミ) の神話、『古事記』に記録された大気都比売神 (オホゲツヒメノカミ) の神話を比較する。これらの神話は、女神が口から食べ物を出しある神に供するが、その神が女神を殺し、その死体の各部から、当時の生活に不可欠なものが初めて得られたと語る、極めて似通った構造を有している。大気都比売神の神話は上述の骨組み以外の要素は持っていないが、保食神の神話には太陽と月が別々に現れることの起源も説かれている。今回の実

験は、両神話が「女神が体から有価なものを出すこと」を E_1 、「女神が殺害されること」を E_2 、「女神の死体の各部位から五穀などが得られること」を E_3 と呼び、この3つの出来事に着目し、これら3つの出来事を象徴するトピックがこの順番の通りに出現するトピック構造を持つという仮説のもとに行う。

3.2 実験内容

3.2.1 使用するソフトウェアと動作環境

今回使用するのは LDA と MeCab というソフトウェアである。これはテキストデータを単語ごとに分割して、単語の品詞や活用形などを解析する、形態素解析と呼ばれる処理をするものである。システムの動作時には主に4つのステータスを設定する。単語出現回数の閾値は、語彙を構成する単語を出現回数により足切りするための閾値である。トピック数は、コーパス (ここではテキストデータである神話) に含まれるトピックの全数である。チャンク数とチャンクサイズはコーパスの時間的処理にかかわるものであるため、ここでは割愛する。

まずテキストデータは MeCab により処理され、含まれる単語を語彙とする。なおこのとき名詞・動詞・形容詞・形容動詞以外の単語は除外される。それ以外の助詞や助動詞などはトピックに影響しないと考えられるためである。出現回数が上述の閾値より少ない単語も除外される。また Bag of Words を作る。これは単語表にある単語のうち、それぞれの行にどの単語がどの回数現れるかを示すものである。語彙と Bag of Words が LDA で処理され、結果を出力する。出力ファイルのうち特に重要なのは行ごとのトピック分布を示す γ ファイルとトピックごとの単語分布を示す μ ファイルである。

3.2.2 実験手順

まず解析するテキストデータを実験システムで処理し、 γ ファイルに示されたトピック分布をグラフで視覚化する。上述の注目する出来事に該当する行において大きな値を持つトピックを特徴付ける単語を μ ファイルから確認し、その出来事とトピックを特徴付ける単語との間に関連があるかどうかを見る。

3.3 実験結果

3.3.1 保食神の神話

保食神の神話は、単語の出現回数の閾値は0、すなわち全ての単語が含まれるようにして実験した。これはこの神話がとても短いからである。大気都比売神の神話も同様である。トピック数は25とした。

比較する他の神話との関連をわかりやすくし、また求めるトピックが出やすくなるように、表1のように単語を変換して解析を行った。

テキストにおけるトピックの分布を図1にグラフで示す。横軸は行、縦軸はトピックの値である。なお簡略化のため、 E_1, E_2, E_3 において特筆すべき値を持たないトピックは表示させていない。

表 1: 保食神の神話 単語の変換

| | | | |
|-----|-----|------------|------|
| 変換前 | 保食神 | (口から出す食べ物) | (五穀) |
| 変換後 | 女神 | 食べ物 | 穀物 |

| | | |
|-----|--------|--------|
| 変換前 | 大気都比売神 | 大宣津比売神 |
| 変換後 | 女神 | |

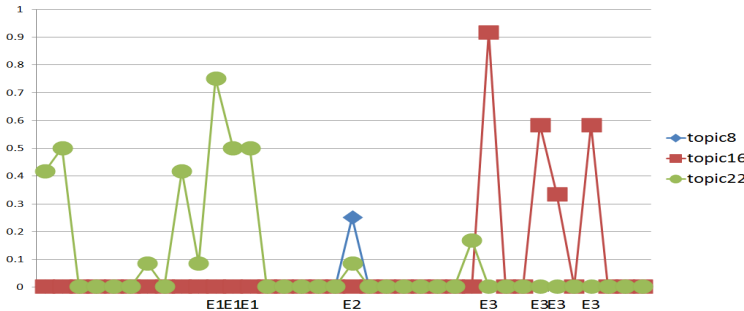


図 1: 保食神の神話

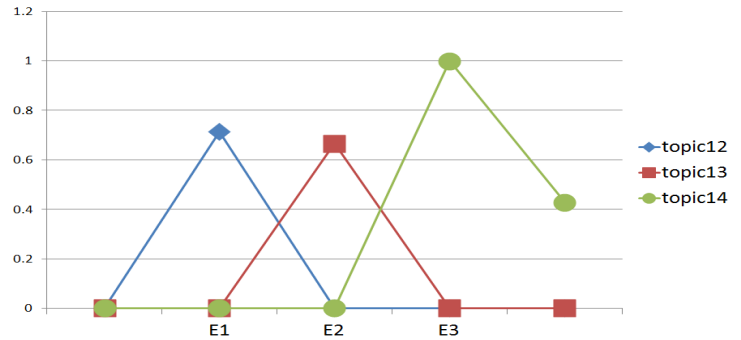


図 2: 大気都比売神の神話

E_1, E_2, E_3 で大きな値を持つトピック 22, トピック 8, トピック 16 を特徴付ける単語をそれぞれ表 2 に表示する。「単語」の欄はそのトピックを特徴付ける単語である。

表 2: 保食神の神話 トピック

| | | | | |
|---------------------------------|---------|-------------|--------|---------|
| 保食神の神話 トピック 22 | | | | |
| 出来事: E_1 女神が体から有価なものを出す | | | | |
| 単語 | 女神 0.13 | 食べ物 0.08 | 出 0.08 | 口 0.08 |
| 保食神の神話 トピック 8 | | | | |
| 出来事: E_2 女神の殺害 | | | | |
| 単語 | 抜い 0.02 | 撃ち殺さ 0.02 | 剣 0.02 | 女神 0.00 |
| 保食神の神話 トピック 16 | | | | |
| 出来事: E_3 女神の死体の各部位から五穀などが得られる | | | | |
| 単語 | 穀物 0.32 | (体の部位) 0.02 | 蚕 0.02 | 牛馬 0.02 |

3.3.2 大気都比売神の神話

同様に大気都比売神の神話について示す。単語の閾値を 0, トピック数を 14 とした。また特に長い 1 文があり, E_1 と E_2 をともに含んでいたため, その切れ目で 2 行に分けた。

単語の変換は表 3, トピックの分布は図 2, トピックと単語はそれぞれ表 4 である。

3.3.3 トピック構造の比較

両神話について, E_1, E_2, E_3 の出来事に対し対応するトピックがこの通りの順番で得られたため, 類似したトピック構造を抽出できたとと言える。

4 結論と考察

実験結果より, 2 つの神話に共通の構造があることを, トピックモデルの手法を用いて示すことができた。

しかし本研究は, 確かに神話研究に対してある程度寄与するものではあるが, それは神話の比較に対する量的基準の適用に留まっており, 神話研究に対する新たなパラダイムの提唱とはなっていない。なぜなら, 仮説の段階で既に従来の比較方法を使用しており, トピックモデルはその証明にのみ使用されているからである。この手法を用いて神話研究にさらなる発展をもたらすためには, いくつかの神話を並行して分析し, そ

これらのトピック構造を一望俯瞰的に比較し, 思いもかけない新たな神話間のつながりなどが発見できるようにならねばならない。そのためにはまずトピック分析の信頼性の向上が必要である。またトピックモデルは言語処理技術であるため, 本研究のような人文科学領域の研究にも大きく貢献し得ると考えられるが, 人文科学研究者からのユーザビリティを考えるならば, テキストの電子化, 各種設定, トピック分布の可視化操作, トピックを特徴付ける単語やそれが特徴的に現れる箇所の閲覧, といった各種インターフェースの改善も重要である。

参考文献

- [1] Nallapati,R. Cohen,W. Dittmore,S. Lafferty,J. Ung,K. "Multi-scale Topic Tomography"
- [2] 宇治谷孟 訳『日本書紀(上) 全現代語訳』(1988年) 講談社(講談社学術文庫)
- [3] 次田真幸 訳注『古事記(上) 全訳注』(1977年) 講談社(講談社学術文庫)
- [4] 大林太良 著『神話学入門』(1966年) 中央公論社(中公新書)
- [5] 松村一男 著『神話学講義』(1999年) 角川書店
- [6] 渡辺公三・木村秀雄 編『レヴィ=ストロース『神話論理の森へ』』(2006年) みすず書房

表 4: 大気都比売神の神話 トピック

| | | | | | |
|---------------------------------|-------------|-----------|-------------|---------|-----------|
| 大気都比売神の神話 トピック 12 | | | | | |
| 出来事: E_1 女神が体から有価なものを出す | | | | | |
| 単語 | 女神 0.16 | 鼻 0.16 | 口 0.16 | 尻 0.16 | 取り出し 0.16 |
| 大気都比売神の神話 トピック 13 | | | | | |
| 出来事: E_2 女神の殺害 | | | | | |
| 単語 | 速須佐之男命 0.16 | | | 女神 0.16 | 殺し 0.16 |
| 大気都比売神の神話 トピック 14 | | | | | |
| 出来事: E_3 女神の死体の各部位から五穀などが得られる | | | | | |
| 単語 | 生り 1.16 | (五穀) 0.16 | (体の部位) 0.16 | | |