

学術利用・地域振興を目的とした民話データベース構築：
しまねアカデミアでの取り組み
Citizen-Oriented Database of Folktales: Attempt in Shimane

中分遥^{1,2†} 佐藤浩輔^{3‡} 宍戸俊悟⁴ 松本朝子⁵ 面坪紀久⁶ 安藤二香⁷ Christopher Kavanagh¹
浜崎浩⁸ 岩瀬峰代⁹ 小見波泰秀¹⁰ 田原敬一郎¹¹ 高見友理⁹ 吉澤剛¹² 佐藤鮎美⁹
Yo Nakawake^{1,2†}, Kosuke Sato^{3‡}, Shungo Shishido⁴, Tomoko Matsumoto⁵, Kiku Omotsubo⁶,
Nika Ando⁷, Christopher Kavanagh¹, Hiroshi Hamasaki⁸, Mineyo Iwase⁹, Yasuhide Kominami¹⁰,
Keiichiro Tahara¹¹, Yuri Takami⁹, Go Yoshizawa¹², Ayumi Sato⁹

¹オックスフォード大学 ²九州大学 ³明治大学 ⁴奥出雲町教育委員会 ⁵松江市地域おこし協力隊
⁶島根県古代文化センター ⁷政策研究大学院大学 ⁸一般社団法人スクナヒコナ ⁹島根大学
¹⁰Matsue-Works ¹¹科学コミュニケーション研究所 ¹²オスロ都市大学

¹Kyushu University, ²University of Oxford, Meiji University, ³Okuizumo Town Board of Education, ⁴Matsue
⁵Community Reactivating Cooperator Squad, ⁶Shimane Center for Studies of the Ancient Culture,
⁷National Graduate Institute for Policy Studies, ⁸Sukunahikona Inc., ⁹Shimane University, ¹⁰Matsue-Works,
¹¹Science Communication Research Institute, ¹²Oslo Metropolitan University

あらまし： 過去数十年間で、民話研究における計算機のおよびデジタル的なアプローチが急速に発展した。本発表では、近年の民話に対する学際的な研究手法や国内外の主要な民話データベースに関して概観した上で、我々がこれまで地域実践者と取り組んできた経過と本事例におけるアプローチ手法を紹介する。民話研究における研究者と地域実践者との協働によるデジタルアーカイブ化の方法および意義について議論する。

Abstract: In the last few decades, computational or digital approaches to the study of folktales have advanced rapidly. Here, we briefly review recent multi-disciplinary methods to study folktales and major digital databases of folktales and related fields. Then, we introduce our ongoing project of building a digital database, collaborating with citizens in Shimane. Based on the review, we discuss methods and benefits of collaboration between citizens and researchers.

キーワード： 民話, データベース, 持続可能性, デジタル人文学, 計算民話学

Keywords: folktale, database, sustainability, digital humanities, computational folkloristics

1. 緒言

情報処理技術の発達により、これまで人文学が取り扱ってきた質的なデータを計量的な観点から解析できる時代となった。本稿では、民話および物語研究の観点から昨今の質的なデータを計量的に扱う研究動向を紹介し、現在島根県において筆者らが取り組んでい

る、計量的な分析手法を用いて処理することを前提としたデータベースを官学民の連携によって作成する試みについて紹介する。

本稿では第一に、民話を計量的に取り扱うアプローチについて紹介する。まず、情報処理の観点から民話を取り扱う認知科学的なアプローチについて紹介し、

† School of Anthropology and Museum Ethnography, University of Oxford, 51/53 Banbury Road, Oxford. OX2 6PE. UK

‡ 明治大学 研究・知財戦略機構〒101-8301 東京都千代田区神田駿河台 1-1

Research Extension and Intellectual Property Headquarters, Meiji University, 1-1 Kanda-Surugadai, Chiyoda-ku, Tokyo, 101-8301

さらに生物学・生態学の方法論を導入する文化進化とよばれるアプローチ、テキストから様々な情報を抽出する自然言語処理的アプローチについて紹介する。

第二に民話をとりまくデジタルアーカイブの現況について概説する。近年、人類学や歴史学周辺において様々なデジタルアーカイブが誕生しており、これらを計量的に分析することで数多くの知見が生まれている。さらに国内外の民話を主題としたデータベースについて紹介する。

最後に、現在我々が取り組んでいる民話収集のプロジェクトについて紹介し、その意義について論じる。

2. 民話研究の認知科学的アプローチ

民話についての研究はこれまで主に人文系諸学間分野で行われてきた。樋口(2011)^[4]によれば20世紀の民話研究は主に①国際比較、②構造論、③精神分析、④文芸学、の4つの流れがあるという。①国際比較は、話型やモチーフといった民話を構成する要素に着目し、地域間の差異や特定の民話の起源を探る、といったアプローチであり、②構造論は、Propp(1928)^[2]に代表されるように、物語における構成要素が果たしている機能に着目し分析するアプローチ、③精神分析は精神分析学の立場から民話を解釈・分析するアプローチ、④文芸学は、口承文芸としての民話を解釈・分析するアプローチである。

これらのアプローチは基本的には民話の質的な特性を明らかにしてきたが、近年、量的に民話を扱おうとするアプローチが登場している。たとえば認知科学の観点から民話や物語が伝達されるメカニズムを実証的に研究しようとする立場や、民俗学の中からも自然言語処理や統計的手法を用いたアプローチが登場し「計算民俗(民話)学」(Computational Folkloristics)という語も登場している^[3,4]。本節ではそのような民話に関する計量的・実証的な研究方法を紹介する。

2.1 物語変遷に関する認知科学的アプローチ

認知科学とは人々の学習や記憶といった知的活動を計算機による情報処理のアナロジーとして解釈する

アプローチである。民話に関しても人々に入力された情報が記憶として貯蔵され出力する過程として解釈でき、認知科学の対象として扱うことができる。

デジタル化の進んだ現在において、物語はデータとして完全に複製し伝達することが可能であるが、複写や口頭で伝承していた時代において、言葉が時代とともに変わっていくように物語も伝達の過程で変異してきた。Sperber & Hirschfeld (2007)^[5]は、物語の変異はランダムに引き起こされるのではなく、我々の記憶といった認知的制約を受けて引き起こされることを指摘した。この点は、民俗学者である Jack Zipes も支持しており、認知科学的観点から民話のサブカテゴリである童話に関する歴史的な変遷について考察している^[6]。これらの理論の基礎となるのが Bartlett (1932)^[7]による物語の伝達に関する一連の心理学実験である。

Bartlett は当時心理学において行われていた無意味語や数字を用いた記憶研究を批判し、物語を用いて社会的文脈や知識構造が記憶に与える影響を検討する研究を行った。彼は、物語を実験参加者に提示し、その後再生させるという手続きを行うことで実際の物語が継承される場面を模し、記憶による物語の変化を検討した。一連の実験によって明らかになったのは、物語が伝達されるにつれて、物語の順序に変化が生じたり単語や言い回しに変化することによって、つじつまの合わない内容が合理化される、あるいは実験参加者にとってより馴染みの深い内容へ改変されるということであった。

物語を伝達するというパラダイムは、その後認知心理学者や認知人類学者によって用いられ、実験によりゴシップのような社会的な内容や^[8]、天使や悪魔といった直観に反する内容が伝達されやすいことが示唆された^[9]。特に、後者に関しては実際に世界に存在する多くの民話にこうした反直観的な内容が含まれることが計量的に示されている^[10,11]。これは我々が反直観的な現象に対してより注意を向けることで学習が促進されるという乳児期から持つ認知的なバイアスに起因すると考えられる^[12]。

2.2 文化進化的アプローチ

文化は時とともに変化する。文化を「ある個体から別の個体へと社会的に伝達される情報」、その文化が変容する過程を生物進化のアナロジーとして捉える一連の研究アプローチは文化進化とよばれている^[13,14]。文化進化の研究はこれまで数理モデルを用いた理論研究が主流であったが^[15-17]、近年フィールドワークや実験^[18-20]、さらに矢尻・民話といった文化形質の計量的解析が盛んである^[21-26]。民話に関する研究は、生物学における系統学を文化研究へ応用した文化系統学とよばれるアプローチが用いられている。系統学とは存在する種に対して、種の形質(e.g.,くちばしの大きさ)および遺伝子情報に基づき、ある種が別の種に由来(継承)しているのか、種の分化がどの時点でおこったのかを統計的に推測し、系統樹という形で表現する方法である^[27]。この方法を民話研究に用いる利点は、現在の生物においてその過去を推測するように、現代の民話の由来を推測できるという点である^[23,24,28]。

文化系統樹を民話に導入した初期の研究である Tehrani (2013) ^[23]は、文化系統樹を用いて「赤ずきんちゃん」とその派生系であるとされている「虎老婆」といった、世界各地に存在する 58 の民話を「被害者が赤いフードを被っているか?」「悪役が食事後に睡眠をとったか?」といった複数の項目について民話の内容をコード化した。彼らは、その情報を遺伝子塩基配列かのように分析することで、系統樹を作成した。その結果、「西洋の『赤ずきんちゃん』の起源は東洋に見られる『虎老婆』である」という仮説に対して、むしろ「西洋のみられる『赤ずきんちゃん』の方が東洋に伝播した」という仮説の方が尤もらしいことが明らかとなった。

また、言語の系統樹を利用することによって da Silva & Tehrani (2016) は今日知られる民話のいくつかについて起源を推定した。^[24,28]その結果、いくつかの民話は、5000 年以上前から存在する可能性が高いことを見出している。このように文化進化アプローチは民話の過去を推測するうえで有用となるだろう。

2.2 自然言語処理アプローチ

これまで人文学のテキストは熟練した研究者の精読 (close reading) によって解釈されてきた。精読は、特定の時代や著者、地域といった限定された領域においてはひとりの研究者でも行うことができるが、世界中で収集され続ける民話や、日々増え続ける Web 上の情報といった膨大なテキストに対して個別の研究者がすべてを精読することには限界がある。

幸い、近年の計算機の能力の向上により、大量のデータを短時間に処理することが可能となった。特に、構造化されていないテキストから有意な情報を抜き出す自然言語処理(Natural Language Processing: NLP)手法の発展により、人文学で扱うような様々なテキスト群にネットワーク分析や統計解析といった計量的な手法が扱えるようになった。比較文学者の Moretti は人文学におけるこのようなアプローチを close reading に対して distant reading と呼び、提唱している^[29]。

民話や物語研究における活用例として、例えば Regan et al. (2016) は 1,327 件の小説のテキストを対象に時系列の感情スコアを算出し、感情の上下が描く曲線が6種類の型に分かれることを示した^[30]。また、筆者らも国際的な話型索引である Aarne-Thompson-Uther type index^[31]に収録されている動物の共起関係や動物物語カテゴリに属する物語群を対象に分析を行った^[32]。「ウサギとカメ」のように本来的には物語は自然の関係を反映している必要がないが、分析の結果、ブタとオオカミといった家畜動物とそれを襲う野生動物がペアとして登場する頻度が高いという結果が得られ、現実の自然環境における知識を民話が反映していることが明らかとなった。

自然言語処理手法を用いることの利点は、なぜそのような結果が得られたのかに関して第三者が客観的に評価可能である点である。ただし、これらの手法で得られた結果や解釈、特にテキストの中身に踏み込んだものについてどの程度妥当性・信頼性があるかについては議論の余地がある。例えば、感情分析を文章に対して行って得られた結果がどれほど妥当性のあるものなのかについて別途検証することが必要だろう。

3. 民話に関わるデータベース

民話の収集は主に民俗学者らや文化人類学者らによってなされてきた。こうして収集された膨大な資料を広く利用できるようにデータベースとして整備しようという取り組みがある。本節では国内外におけるデータベースの取り組みについて紹介する。

3.1 人類学・歴史学のデータベース

人類学における大規模なデジタルデータベースのひとつが eHRAF(Human Relations Area Files) World Culture^aである。このデータベースはエール大学により主導され、世界の多くの研究機関が参加している。このデータベースの特徴は地域や政治的単位に関するコーディングシステムである OWC (Outline of World Cultures) と 農耕・宗教・歴史といった主題ごとの文化に関するコーディングシステムである OCM (Outline of Cultural Materials)が存在することである。eHRAF は比較文化研究を行う際のデータセットとして用いられており、統計的な解析によって宗教的伝統や道徳性の文化の普遍性等が検討されている^[33,34]。eHRAF には民話(Folktale)という項目も存在しており 500 以上のデータがヒットする。もう一つ特徴的なデータベースは、マックス・プランク研究所の言語進化・文化進化に関する研究グループによる D-Place^bである^[35]。このデータベース上の変数はすべてコード化されており、すぐに統計的な分析を行うことができる。また、Web 上で言語の系統樹を元にした文化系統樹分析や、ある文化形質の有無を地図上にプロットするなどといった分析がインタフェース上で可能となっているのが大きな特徴である。

これらのデータベースは比較文化研究には有用であるが、時系列的なデータ分析を行うのに適していない。そこで人類学者・生態学者・歴史学者らが協力し構築された歴史学のデータベースが Seshat: Global History Databank^cである^[36]。このデータベースを利用した論文が 2018 年に初めて刊行された^[37]。論文は、

30 地域について 1 万年分のデータを用いて、これらの地域の社会的複雑性の進化に関連があるか計量的に検討したもので、分析の結果、社会の複雑性の指標である、人口サイズ・経済規模・インフラ設備・情報手段・文字システムといった複数の項目は互いに関連しており、これらが発展するパターンは地域を超えて共通であることを見出した。この研究はデジタルデータベースによって、これまで人文社会科学によって提唱されていた仮説や理論が体系的かつ計量的に検討可能となった一例である。こうしたデータベースは、例えば人類史における「宗教の社会的複雑性に与えた影響」といった複数の仮説が存在する分野において、その仮説の検証に有用であろう(宗教史を専門に扱ったデータベースとしては、The Database of Religious History が存在する^{d[38]})。

3.2 民話の専門のデータベース

3.2.1 国外の民話データベース

海外の民話データベースの著名なものとして、Meder らによるオランダ民話データベース^e(Dutch Folktale Database)がある^[39]。このデータベースは当初スタンドアロンの電子データベースとして作成され、2004 年にオンライン化されて現在に至る。2004 年の時点で 20,000 件、現在 46,000 件を超える民話が収録されており、各民話は語り部、採集者、採集場所、採集日時、サブジャンル、方言といったメタ情報のほか、モチーフや話型、キーワード、要約などが機械学習や自然言語処理技術を用いて自動的に付与されている。ビジュアルライゼーションのための機能も充実しており、類話や関係する民話の表示、キーワードに基づく時系列表示、地図上での表示などができる。

また、デンマークでは Evald Tang Kristensen Collection^fとして、19 世紀から 20 世紀初頭にかけて活動した民話収集家である Evald Tang Kristensen の 24,000 ページに渡る手稿を地理情報とともにデータベース化し、Web 上で閲覧できるようにしている。キーワ

^a <http://ehrafworldcultures.yale.edu/ehrafe/>

^b <https://d-place.org>

^c <http://seshatdatabank.info/>

^d <https://religiondatabase.org/landing/>

^e <http://www.verhalenbank.nl/>

^f <http://etkpace.scandinavian.ucla.edu/etkSpace/index.php>

ードや、Kristensen による独自のインデックスなどによる検索も可能である^[4,39]。

現在、上記のデータベースの作成者たちが主導して、欧州の複数地域の民話を横断的に利用できるような Intelligent Search Engine for Belief Legends (ISEBEL)プロジェクトが進められている。

3.2.2 国内の民話データベース

本邦における民話データベース化の取り組みについて、Web 上で閲覧可能なものをいくつか紹介する。

東アジア民話データベース：「東アジア民話データベース[§]」は東アジア民話データベース作成委員会によるもので、日本民話・韓国民話・中国民話・沖縄伝承話の各データベースを統合したものである^[41-43]。音声データを中心に 63,500 件の民話情報が収録され、音声データ以外のメタ情報のデータベースが専用のファイルでダウンロードできるほか、音声データの一部がサンプルとして Web 上で公開されている。

怪異・妖怪伝承データベース：国際日本文化研究センターが公開している「怪異・妖怪伝承データベース^{||}」は、怪異や妖怪の登場する民話に関して記述された国内の日本民俗学の文献をデータベース化したものである。1997 年に作成が開始され、2002 年 6 月から Web 上で利用できるデータベースとして公開されており^[44]、2019 年 1 月現在約 35,800 件の書誌情報が採録されている。採集された各民話について、怪異(妖怪)の呼称・出典・地域といった書誌情報に加え、100 字程度の要約がついている。

国立民族学博物館データベース：国内外の民族資料を収集している国立民族学博物館には、研究者が世界各地で収録したテープ・CD などの音源が音響資料として収蔵されている。2018 年の時点では 62,651 点の音響資料が存在し^[45]、この音響資料の中に世界各地の民話の語りの音声も含まれ、目録が Web 上で検索できる。また、国内の民話に関しては、「日本昔話資料：稲田コレクション」がある。これは民俗学者の稲

田浩二らのグループが日本の 29 道府県で行った調査のテープの内容を、一話を一項目として整理したものであり、3,696 件分の項目が存在する。Web 上には目録が公開されており、題名・時間・収録年月日・収録地・収録者・話型分類・概要といった情報が利用できる。音声データについては民族学博物館の内部からのみアクセスできる。

3.3 民話データベースの課題

このように国内外で様々な民話データベースが構築されているが、課題も明らかになっている。

一番大きな課題はコストで、国内の民話データベースに関してはいずれもデータ整理やタグ付けといった作業が人手で行われており、聞き取った音声データの民話の分類や整理をする研究者・専門家の不足が指摘されている^[41]。

一方で、デンマークの Tang Kristensen Collection の作成に携わった Tangherlini は、人手による分類はコストが大きいだけでなく、コレクションの増加につれて維持するのが困難であると指摘している^[41]。Meder らや Tangherlini らのグループは民話のモチーフや話型の判定に機械学習や自然言語処理を用いた自動化処理を採用している^[3,4,39,40,46,47]。今後は、これら情報技術をどう採用していくかが鍵となるだろう。

「怪異・妖怪伝承データベース」の利用状況について報告した松村(2014)^[48]は、活用事例として①「下調べとしての活用」、②「計量分析を行う際の活用」、③「妖怪事典を作成するための活用」、④「展示構成を考えるための活用」を挙げ、人文学へ与えた影響として①「事例収集の時間の縮小」、②「計量分析の可能性」、③「新たな視点の発見」、④「教育の場における資料収集ツールの提供」、⑤「事典の作成への寄与」、⑥「当該分野の研究者以外による活用」があると指摘している。特筆すべきは、④「教育の場における活用」や、⑤「事典の作成への寄与」、⑥「当該分野の研究者以外による活用」であろう。これらは、当初

[§] <http://minwadata.fm.senshu-u.ac.jp/EastAsiaMinwaDB/about/index.html>

^{||} <http://www.nichibun.ac.jp/YoukaiDB2/index.html>

ⁱ <http://htq.minpaku.ac.jp/databases/av/audcat.html>

^j <http://htq.minpaku.ac.jp/databases/inada/>

データベースを作成したときに想定していたような利用法ではなく、利便性の高い Web 上のデータベースとして公開したことにより、潜在的な利用者によるアクセスが可能になり、新たな利用法が見出されたといえる。

またオランダやデンマークの民話データベースでは、それ自体をコーパス(言語資源)として使った研究も行われており^{39,46,47,50}、狭義の民話研究にとどまらない幅広い可能性が民話データベースにはある。

このような様々な可能性を活かすために、潜在的な利用者にリーチしやすい、利便性の高いデータベースをいかに構築するかが重要な課題である。

4. CODAMA プロジェクト：地域と連携した民話データベースの試み

「しまねアカデミア」は、既存の大学や学協会の補完的な主体として、学際的な学術研究と地域振興とを一体化させた開かれた場を作ろうとする運動である。

その取り組みの一環として 2018 年に CODAMA: *Citizen Oriented Digital Archive of Mythology and Anthropology* (「市民志向の神話学および人類学のデジタルアーカイブ」)プロジェクトを開始した。

この CODAMA プロジェクトは、地域の伝承や民話の聞き取りを行いデジタルアーカイブを整備するとともに、地域の文化資源として活用するもので、出雲神話に代表される島根県と、日向神話に代表される宮崎県を主なフィールドワークの地とし、住民参加の市民科学として展開することを目指している。2018 年 3 月に奥出雲町にてプレ・ワークショップを行い、同年 8 月にはプロジェクトのキックオフイベントにあわせ神話の伝承地を巡るエクスカージョンを行った。

本プロジェクトは研究者と住民とで組織されており、単なる研究者主導のプロジェクトではなく、両者の参与・協働により、お互いにとって価値のある結果を生み出すことを目的としている。研究者が専門知を、住民が現場知を提供することで、研究者にとってはより学術的に価値の高い研究成果を得られることを、住民にとってはより地域に還元可能な形で成果を活用できる

ことを目的とする。本節では、CODAMA プロジェクトの沿革と、取り組みの意義について議論する。

4.1 市民と研究者の協働によるアーカイブ化

CODAMA プロジェクトはもともと、しまねアカデミア内で「島根県の地域実践者から民話や神話を活用するような研究ができないか」という問いかけがあり、その具体的な方法として民話のデジタルアーカイブ化が提案されたことが発端となり始まった。

民話・伝承の収集に関しては、これまで研究者が主体となって行われてきた。それは主に、喪われゆく文化の収集・保存といった学術的な価値に重きが置かれており、実際にそれら口承の文化の担い手である地元の人々にとってどのような価値を持ちうるかについては十分に顧みられてこなかった。

地域の人々が情報提供者として受動的な存在としてのみ関わるのではなく、収集・保存作業に主体的に参与し、自らが持つ文化遺産の活用法を模索していくという点で本プロジェクトはユニークであるといえる。

協働の利点 研究者と地域実践者との協働による利点は、第一に地域の豊富な現場知を活用することができることである。通常、研究者は外部者として地域に関わる。その際、地名や特有の言い回し、地域の誰が何を知っているか、などといった地域固有の知識にアクセスすることは容易ではない。地域実践者と協働が可能な場合、そのような知識のギャップをある程度補完した状態で作業を開始することができる。聞き取りの際の基盤となる信頼関係の醸成も、研究者が単独で乗り込む場合よりも促進される可能性が高い。

また後述するように、地域にとってのメリットがあるならば地域実践者にとって活動を推進するモチベーションとなる。研究者によるトップダウンの活動では、ともすれば、失われ滅びゆく伝統の一瞬を切り取るだけに過ぎないかもしれない。一方で、地域の草の根のボトムアップの保存活動となれば、継続的に記録していくことが可能になり、自らの取り組みによって地域の文化資源を保全・活用していくことができるだろう。

協働が研究者にとってメリットがある一方で、では逆に、地域実践者にとって研究者と協働するメリットは何であろうか。もし、地域実践者のみでこれらの活動を完結できるのであれば地域の外部者である研究者が携わる必要性はないはずである。

地域実践者にとってのメリットは、ひとつには、研究者の持つ方法論的枠組みを活用できるという点が挙げられる。聞き取りの手順や集めたデータの管理、データベースの設計といった形式に関わる様々な作業は、ゼロから策定すると非常に煩雑なことになる。また仮に独自の方式を策定したとしても、その独自性ゆえに外部の者にとって使いようがないものになりかねない。そこにノウハウを持つ研究者が参与することで、より標準的な形でデータベースを作成することができる。

それ以外にも、他の地域とをつなぐ役割を研究者が果たせる可能性がある。一般に、地域を拠点として生活する地域実践者の活動は少なからず地理的に制約され、離れた地域と接点を持つことは容易ではない。アーカイブの作業を通じて、研究者がハブとして介入することによって、これまで交流の少なかった地域との接点を作ることができる。他地域の文化資源を知ることによって、自地域の文化資源の独自性だけでなく、地域を超えた普遍性を確認することができる。このことは自分たちの住む地域の文化資源の価値を、学術的な成果だけによってだけではなく、自分たち自身で発見、再評価、発信することにつながるだろう。

このように地域実践者と研究者が互いの長所を活かしあうことによって、片方だけでは生み出せない、そして両者にとって価値のある成果を生み出せるという点で、協働の意義がある。

文化資源の活用 民話や伝承といった口頭伝承は、語り手と語りの場、聞き手の喪失という危機に直面している^[41-43]。収集・記録していくことはこうした無形の文化資源を保全する第一歩として重要である。しかしながら、一度記録されればよいというものではない。保全というからには、常にアクセス可能な形で維持がなされなければならない。さもなければ、せっかく集めた貴

重な資料も埃をかぶって埋もれてしまうことになる。こうした民話や伝承それ自体が当該の地域の人々にとって継承し維持・発展させていく価値のあるものであると認識される必要がある。そのためには、収集した民話を文化資源として活用することを考えねばならない。

まず考えられるのは観光資源としての活用である。本プロジェクトのキックオフイベントを行った際、民話・神話にまつわる箇所を巡るスタディーツアーを試験的に行い、ヤマタノオロチ伝説にまつわる地域などを巡った。神話や民間伝承にふれることで、文化的要素を求める旅行者の満足度を上昇させることができる。既に実際に観光業界と協力することで、ツアーガイド等に民話や神話に関する知識の伝達を行っている。今回の研究プロジェクトにおける中心地である奥出雲および宮崎県高原町においては、数多くの神話が存在しているが、知識が分散しており、観光ガイドでも知らない民話が多いため、これらの資源を活用する余地がまだ十分にあると考えられる。神話に関わる名所・名跡は、例えば寺社仏閣といったその建築物自体が分かりやすい観光的価値を持つものだけではなく、背景となる物語を知ることによって初めて観光資源としての価値が生まれるものもある。今回の民話のアーカイブ化によって、地域における民話を整理・分類し、地域の文化資源をより利用しやすい形にすることで、有効に活用することができるだろう。

また、文化資源を整備することは、地域の人々が自らの持つ資産を対象化し、認識する機会となる。言い換えれば、地域の価値を地域の人々が再発見するプロセスであるといえる。民話や神話、および関連する諸事項についての知識を集める中で地域のルーツや歴史についてより深く知ることができ、自分たちの住む地域へのアイデンティティを高めることが可能となる。このことは郷土教育などへの直接の活用だけではなく、間接的に、地域のつながりといった社会関係資本の強化にもつながるだろう。

このような文化資源の活用は、地域実践者の主体的な取り組みにつながり、持続的な民話の収集・保存

活動につながる。住民自身が主体になって民話を収集・保存し文化資源として整備・活用できるようになることが、目指すべきひとつの目標であると考えられる。

4.2 特定地域における民話収集のメリット

今回 CODAMA プロジェクトでは、民話を島根(と宮崎)という比較的小規模の地域で活動を行う。第3節で述べたように、国際的な民話データベースや、東日本民話データベースのような全国的なデータベースがある中で、あえて地域固有のデータベースを用意することの意義は何だろうか。

ひとつは、近年の民話研究の関心に応じたデータセットを用意するということである。このデータベースが目指すのは「深く狭く」ではなく「広く浅く」を目的としたデータセットで、狭義の「語り手」のみではなく、広く聞き取りを採集する。これにより、ひとつひとつのデータは断片的であっても、たとえば文化進化アプローチのような系統学的な分析が可能になる。

この点については、昆虫の標本採集をアナロジーとして考えたい。例えばチョウの標本を作る際の一つの目的は、世界に存在するあらゆる蝶の種を集め並べることである。こうした目標を設定した場合、一地域において得られるバリエーションには限りがある。一方、別の方法として、特定の種のチョウを比較的小規模の地域から複数サンプリングしてくる方法がある。こうした例えば山陰地方や島根県といった比較的小さい地域において、例えばチョウの羽の長さにどのような違いがあるのか、そのチョウが存在しない地域があるかどうかといったことが明らかとなる。この違いが地形や捕食者、あるいは餌となるような種の有無によって説明が可能となる場合がある。民話や伝承についても同様に考えることができる。比較的小規模の地域において同じような民話を収集することで、たとえば同じ話であってもどのような要素は比較的忠実に模倣され、あるいはどのような要素が変容するかといったことが明らかになる。系統樹分析を行うことでその民話の起源がどこに存在するのかを推定可能になるほか、様々な計量的な分析を行うことが可能である。

もうひとつは、前項で述べたように、地域住民にとっても価値のあるデータベースを作成することにより、地域住民自身が民話の収集・保存活動に取り組むことができることである。対象地域を比較的小規模なものに設定することで、大規模な場合に比べアーカイブ作業の意義が明確になり、チームワークも取りやすくなる。

既存の(少なくとも国内の)民話データベースにおいて、増え続ける膨大な民話データに対して有効な処理の方法を提示できていない以上、現在の研究者主導のトップダウンな取り組みだけでは処理能力に限界があることは明らかである。もしこのような民話収集・データベース化の取り組みが地域住民との協働によって可能なのであれば、有効な打開策となりうる。もしこの取り組みが広がり、各地でボトムアップ的にデータベース化を進めることができれば、民話研究の新地平を拓くことができるであろう。

4.3 民話データベースに向けて

最後に、我々の民話データベースが満たすべき要件について指摘して本節を終わりにする。

ひとつは共通のフォーマットを確立することである。樋口^[41-43]も述べているように、複数の地域のデータベースを統合するためには、コアの部分で共通のフォーマットを採用していることが不可欠であり、それぞれが独自のフォーマットを作成した場合、利便性が失われる。地域の関心に応じて拡張可能にするとしても、コアの部分は統一されているべきであろう。

また、各データベースはマシンリーダブル、すなわち計算機で直接読み取りが可能な形式で Web 上に公開されるべきである。ここでいうマシンリーダブルとは、単にテキスト形式で機械が読めるというだけでなく、タグなどで構造化され情報を系統的に取り出すことができるよう整備されていることを指す。ただし、標準的な Web ページで用いられる HTML は構造化されてはいるものの、利用者が情報を活用するためにはひと手間かかる。より活用を図るならば、専用の Web API (アプリケーションプログラミングインターフェース)を備えているべきであろう。利用者からのクエリ(問い合わせ

せ)に応じて必要な情報を JSON などプログラムで扱いやすい形式で返すことで、利便性は飛躍的に向上する。各データベースが小規模でも、それぞれが検索 API を用意すれば、横断して検索することのできるメタ検索エンジンなどを作成することも可能である。

近年、Web API を活用したオープンイノベーションが加速している。これは企業や自治体といった API を公開し、開発者や利用者が自由に利用できるようにすることで、新たな活用の可能性を開くものである。データベースは作成者の想定にとどまらない価値と可能性を持っている。使い方を狭いものに規定するのではなく、利用者が価値を発見しやすい、より利用しやすい形でデータを公開することが必要であろう。

また、マシンリーダブルであることは、テキストを自然言語処理や機械学習のデータとして扱えることを意味する。国内の民話データベースに関する文献では、いずれもアーカイブ作業には専門的な知識と判断能力が必要、という記述がある^[41,48]が、その専門性の要請がデータベースの公開を妨げている部分もある。国立民族学博物館に収蔵されている資料について述べた守屋(1998)^[49]では水島・山口文庫なる昔話の録音テープ群の存在について言及されているが、未整理のため検索の対象になっていないという。

先行するオランダやデンマークのデータベースにおいては、話型やモチーフの判定を、統計的な分類問題へと帰着することで、タグ付けそれ自体がひとつの研究対象となっている(e.g. [39,46,47,50])。

公開されない「完全な」データベースよりも、部分的にでも利用可能な「不完全な」データベースの方が価値が高いと考えられる。最初から完璧な状態での公開を期するよりも、事実に関する最低限必要なメタ情報が実装された段階で公開し、アップデートしていく方がより活用可能なデータベースとなるだろう。

5. 結語

本稿では、近年隆興しつつある民話に計量的な視点を導入した研究を紹介した。デジタル化が進むことによって、より多くの対象を分析で扱えるようになり、民話

の起源の推定等これまでの民話研究において困難であった多くの知見が明らかになっている。我々の取り組みは国内外の既存のデジタルアーカイブと比較して小規模なものである。しかし、地域の市民と一体となってデータベースを作成し、観光や教育に反映していくことによって、研究の成果が目に見える形で社会へと還元されるプロジェクトとなることを目指している。また、小さい地域において密度の高い民話を収集することは計量的な観点からは様々な仮説を検証するデータセットとして学術的な利用価値も十分に高いと考えられる。今後、こうした地域にフォーカスしたデータベース構築が複数地域で行われることで、地域を超えた市民交流や研究が行われることを期待する。

参考文献

- [1] 樋口淳. (2011). 民話の森の歩き方. 春風社.
- [2] Propp V.Y. (1968) *Morphology of the Folktale: Second Edition, Revised and Edited with Preface by Louis A. Wagner, Introduction by Alan Dundes*: University
- [3] Abello, J., Broadwell, P., & Tangherlini, T. R. (2012). Computational folkloristics. *Communications of the ACM*, 55(7), 60
- [4] Tangherlini, T. R. (2013). The Folklore Macroscope: Challenges for a Computational Folkloristics. *Western Folklore*, 72(1), 7–27.
- [5] Sperber, D., & Hirschfeld, L. (2007). Culture and modularity. *The innate mind: Culture and cognition*, 2, 149–164.
- [6] Zipes, J. (2012). *The irresistible fairy tale: The cultural and social history of a genre*. Princeton University Press.
- [7] Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University
- [8] Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, 97(3), 405–431.
- [9] Barrett, J. L., & Nyhof, M. A. (2001). Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. *Journal of Cognition and Culture*, 1(1), 69–100.
- [10] Barrett, J., Burdett, E. R., & Porter, T. (2009). Counterintuitiveness in Folktales: Finding the Cognitive Optimum. *Journal of Cognition and Culture*, 9(3), 271–287.
- [11] 中分遥, Burdett, E., Jong, J., & Whitehouse, H. (2012) 民話における超自然的存在と道徳的示唆 日本人間行動進化学会第 10 回大会 pp. 28-29
- [12] Stahl, A. E., & Feigenson, L. (2018). Violations of Core Knowledge Shape Early Learning. *Topics in Cognitive Science*, 1–18.
- [13] Mesoudi, A. (2011). *Cultural evolution: How Darwinian theory can explain human culture and synthesize the social sciences*. University of Chicago Press.

- [14] Morin, O. (2016). *How traditions live and die*. Oxford University Press.
- [15] Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago press.
- [16] Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach* (No. 16). Princeton University Press.
- [17] Henrich, J., & Mcelreath, R. (2003). The Evolution of Cultural Evolution. *Evolutionary Anthropology*, 123–135.
- [18] Derex, M., Beugin, M., Godelle, B., & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature*, 503(7476), 389–391.
- [19] Mace, R., Thomas, M. G., Wu, J., He, Q., Ji, T., & Tao, Y. (2018). Population structured by witchcraft beliefs. *Nature Human Behaviour*, 2(January).
- [20] Henrich, J. (2015). *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
- [21] O'Brien, M. J., Lyman, R. L., Mesoudi, A., & VanPool, T. L. (2010). Cultural traits as units of analysis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1559), 3797–3806.
- [22] Kamilar, J. M., & Atkinson, Q. D. (2014). Cultural assemblages show nested structure in humans and chimpanzees but not orangutans. *Proceedings of the National Academy of Sciences of the United States of America*, 111(1), 111–5.
- [23] Tehrani, J. J. (2013). The phylogeny of little red riding hood. *PLoS ONE*, 8(11).
- [24] da Silva, S. G., & Tehrani, J. J. (2016). Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales. *Royal Society Open Science*, 3(1), 150645.
- [25] Ross, R. M., & Atkinson, Q. D. (2016). Folktale transmission in the Arctic provides evidence for high bandwidth social learning among hunter-gatherer groups. *Evolution and Human Behavior*, 37(1), 47–53.
- [26] Acerbi, A., Kendal, J., & Tehrani, J. J. (2017). Cultural complexity and demography: the case of folktales. *Evolution and Human Behavior*, 38(4), 474–480.
- [27] Mace, R., & Holden, C. J. (2005). A phylogenetic approach to cultural evolution. *Trends in ecology & evolution*, 20(3), 116–121.
- [28] Pagel, M. (2016). Anthropology: The Long Lives of Fairy Tales. *Current Biology*, 26(7), R279–R281.
- [29] Moretti, F. (2013). *Distant reading*. Verso Books.
- [30] Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1), 31.
- [31] Uther, H. J. (2004). *The Types of International Folktales: A Classification and Bibliography. Animal tales, tales of magic, religious tales, and realistic tales, with an introduction* (Vol. 1). Academia scientiarum Fennii
- [32] 中分遥, 佐藤浩輔 (2018). 国際民話抄録の計量的分析 ~ 文化知識の伝達の観点から ~ 信学技報, vol. 118, no. 122, NLC2018-1, pp. 1-6.
- [33] Curry, O. S., Mullins D, A, & Whitehouse, H. (in press). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*.
- [34] Winkelman, M. J. (1990). Shamans and Other “Magico-Religious” Healers: A Cross-Cultural Study of Their Origins, Nature, and Social Transformations. *Ethos*, 18(3), 308–352.
- [35] Kiry, K. R., Gray, R. D., Greenhill, S. J., Jordan, F. M., Gomes-Ng, S., Bibiko HJ, Blasi DE, Botero CA, Bowers C, Ember CR, Leehr D, Low BS, McCarter J, Divale W, Gavin MC. (2016). D-PLACE: A global database of cultural, linguistic and environmental diversity. *PLoS ONE*, 11(7), 1–14.
- [36] Turchin, P., Brennan, R., Currie, T.E., Feeney, K.C., François, P., Hoyer, D. Manning, J. G., Marciniak, A., Mullins, D., Palmisano, A., Peregrine, P., Turner, E.A.L., & Whitehouse, H. (2015). Seshat: The Global History Databank. *Cliodynamics: The Journal of Quantitative History and Cultural Evolution*, 6(1).
- [37] Turchin, P., Currie, T. E., Whitehouse, H., François, P., Feeney, K., Mullins, D et al. (2018). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences*, 115(2), E144–E151.
- [38] Slingerland, E., & Sullivan, B. (2017). Durkheim with Data: The Database of Religious History. *Journal of the American Academy of Religion*, 85(2), 312–347.
- [39] Meder, T. (2017). The Technological Developments of the Dutch Folktale Database (1994–2016). *Studies in Oral Folk Literature*, (5), 45.
- [40] Tangherlini, T. R. (2016). Big Folklore: A Special Issue on Computational Folkloristics. *Journal of American Folklore*, 129(511), 5–14.
- [41] 樋口 淳. (2012). 民話データベースの課題と可能性. *口承文藝研究*, (35), 16–33.
- [42] 樋口 淳. (2015). 講演 東アジア民話データベースと新しい語り. *山陰民俗研究*, 20, 1–13.
- [43] 樋口 淳. (2016). 口頭伝承のデータベース化: その戦略・実践・国際協力の可能性について. *専修人文論集*, (98), 163–192.
- [44] 小松和彦, 常光徹, 山田奨治, 中山和久. (2003). 異界へのいざない: 怪異・妖怪伝承データベースの試み. *総研大ジャーナル*, 3, 42–43.
- [45] 国立民族学博物館 (2018). 国立民族学博物館要覧 2018
- [46] Meder, T., Karsdorp, F., Nguyen, D., Theune, M., Trieschnigg, D., Everhardus, I., & Muiser, C. (2016). Automatic Enrichment and Classification of Folktales in the Dutch Folktale Database. *Source: The Journal of American Folklore*, 129(511), 78–96.
- [47] Karsdorp, F., Van Der Meulen, M., Meder, T., & Van Den Bosch, A. (2015). MOMFER: A search engine of thompson’s motif-index of folk literature. *Folklore (United Kingdom)*, 126(1), 37–52.
- [48] 松村薫子. (2014). 「怪異・妖怪伝承データベース」による妖怪研究の展開. *研究報告人文科学とコンピュータ*, 2014(10), 1-5
- [49] 守屋祐子. (1998). 民博のデータベースたち: その作成と利用. *情報の科学と技術*. 48(2), 87-94.
- [50] Tangherlini, T. R., & Leonard, P. (2013). Trawling in the Sea of the Great Unread: Sub-corpus topic modeling and Humanities research. *Poetics*, 41(6), 725–749.