様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

# 科学研究費助成事業　　研究成果報告書

令和　２ 年　 ６ 月 １７ 日現在

機関番号：３４３１６
研究種目：挑戦的萌芽研究
研究期間：2016～2019
課題番号：１６Ｋ１２５４４

研究課題名（和文）インド古典のフレーズインデックス付き統合アーカイブ構築とフレーズ分析

研究課題名（英文）Construction of archives of Indian classics with phrase index by means of corpus based extraction of formulaic sequences

研究代表者
中谷　英明（Nakatani, Hideaki）

龍谷大学・公私立大学の部局等・研究員

研究者番号：２０１４０３９５

交付決定額（研究期間全体）：（直接経費）　　2,300,000円

研究成果の概要（和文）：グーグルのMapReduceを用いて本研究分担者が開発した新しいフレーズ自動抽出法は、従来扱えなかった大規模データベースからのすべての「フレーズ」（連続する単語群）の自動抽出を可能にした。制作年代不詳の文献が多く、また一本の中に年代の異なる複数の隠れた層を含むインド古典からのフレーズ自動抽出は、固有の「言い回し」の所在を明確にすることによって、諸文献の関係性及び一文献内の隠れた諸層の存在を明らかにし、インド古典理解の画期的精密化を約する。インド古典のフレーズ分析研究は、インド古来の口頭伝承の特異な正確さと、諸古典の相互関係及び内部構造を明らかにし、古典理解を飛躍的に向上させるであろう。

研究成果の学術的意義や社会的意義
インド古典の二特徴は、数千年に渡る正確無比の口頭伝承と、歴史意識の欠如から来る制作年代の不詳である。また一書が制作年代の異なる数層から成ることも稀ではない。従って、インド古典学にとって、諸文献の相対年代及び一文献内の諸層の確定は必須の前提条件であるが、残念ながら現在そうはなっていない。
芝野が考案したフレーズ自動抽出法は、諸文献の関係及び文献内諸層の確定に力を発揮し、従来研究者が長年かけて得た理解を数カ月で獲得することを可能にした。
21世紀には世界の最重要国の一つとなるインドの伝統的心性の理解には、インド古典の理解が必須であり、フレーズ分析による古典理解の画期的進展が期待されるゆえんである。

研究成果の概要（英文）：The all new automatic phrase extraction algorithm developed by Kohji Shibano using Google's MapReduce has enabled the automatic extraction of all "phrases" (groups of consecutive words) from a large-scale database that could not be processed before.
Not only do Indian classics contain a large number of works of unknown author and date, but a single work often contains several hidden layers of different dates. Automatic phrase extraction, however, indicating all locations of unique "phrases", will clarify the relationship of multiple works, as well as the existence of layers of different date within the same work.
Thus, phrase analytic research of the Indian classics will prove the accuracy of oral textual transmission from ancient India and will allow a far precise understanding of Indian classics.

研究分野：インド古典学・宗教学・中期インド語学

キーワード：インド古典　フレーズ抽出　テキスト内階層　リグ・ヴェーダ　Ngram　MapReduce　Formulaic Sequence　テキストデータベース

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

## １．研究開始当初の背景

(1) インド古典の二大特徴は、数千年に渡る正確無比の口頭伝承と、歴史意識の欠如による制作者・制作年代の不詳である。これに加えて、一書が時代の異なる数層から成り、その事実が判明していない場合も多い。従って、インド古典学にとって、諸文献の相対関係・相対年代及び一文献内の諸層の特定は、必須の前提のはずであるが、現在そうなっていないことも少なくない。
(2) 研究分担者芝野耕司がグーグルの MapReduce を用いて開発した大規模データベースからのフレーズ自動抽出法は、従来扱えなかった巨大データベースからのすべての「フレーズ」（連続する単語群）の自動抽出に成功した。この方法は、諸文献の関係性及び文献内諸層の確定に大きな力を発揮し、研究者が多年かかった分析を数カ月で終了することを可能にすると期待された。

## ２．研究の目的

(1) 本研究の目的は、インド古典学研究者と計算機科学研究者が共同して、新情報処理技法「巨大データベースからのフレーズ自動抽出法」を、巧緻な口頭伝承法（クラマ・パータ）によって伝承されてきたインド古典に適用して､インド諸古典の関係性と､一古典内の多層構造を解明し、インド古典理解の飛躍的向上を図ることである。
(2) また、千年～三千年以上正確に口頭伝承されてきた、世界的に比類のないテキストに、このフレーズ自動抽出という新技法を適用し、新技法の有効性や特性を確認し、その技法をさらに改善する余地を探ることである。

## ３．研究の方法

(1)　従来の言語データ処理は、中間処理データが膨大となることによる計算上限があった。芝野が開発した新技法はグーグルの MapReduce を用いてこれを回避し、すべての「フレーズ」の抽出を可能にした。すなわちその「統合 Ngram 分析」は、１）一文中の全 Ngram を生成し、２）その中から重複を排除し、３）同一の出現文リストの場合には最長の Ngram 以外を削除する、という手順によって、フレーズを自動的に抽出し、フレーズインデックスを作る。
(2)　インド古典は前 1200 年頃インドで初めて編纂された聖典、『リグ・ヴェーダ』(約 1 万詩)を起点とし、そこから他の 3 ヴェーダ、ブラーフマナ、ウパニシャッド等の祭官制作のバラモン聖典が次々に派生し、さらに仏教、ジャイナ教等の聖典や、『マハーバーラタ』（7 万 5 千詩）等の武士階級が制作の中心的役割を担った古典が制作される。中世には、これらの古典を基にプラーナ（主要作品のみで 42 万詩）等の神話・歴史物語集や哲学書が現れる。従って、すべての古典の淵源となった『リグ・ヴェーダ』から始めて、上記の順にフレーズ分析を行うこととする。

## ４．研究成果

(1)　統合文脈 N-gram 処理による定形表現抽出

一般の N-gram 統計では、各 N 単位で連続する単語の頻度統計を用いるのが基本である。新考案の統合文脈 N-gram 処理では、頻度の単位を文と統一し、一つの文に含まれる異なる N を含む全ての N-gram を集計対象とする“統合”に加えて、それぞれの N-gram が出現する文の ID リストを記録する“文脈”処理を加え、同一文脈に出現し、より長い N-gram に含まれる短い N-gram を取り除くこと、すなわち統合文脈 N-gram 処理によって定形表現の自動抽出を行うこととした。しかし、この処理ではすべての N-gram を扱う“統合”処理によって大幅にデータ量が増えるだけでなく、頻度数並びに対象となる集合の文 ID の記録によってもデータ処理量が爆発的に増え、メインメモリ処理が不可能となる。また OS ファイルシステムの上限を超えるデータ処理が必要となる。このため 2012 年に Google が発表した、ビッグデータを分割処理し、その後に集計を行う MapReduce アルゴリズムを用いることによって、この処理を実現した。

しかし、本来の意味でのビッグデータ処理が必要となるため、先ずは N-gram を Google のコーパスに準じて 7-gram に制限し、処理を行った。

その後、文脈 ID リストを一定単位ごとにリスト ID を付してリスト ID アブストラクションを行うことによって、処理量を削減する方法を考案し、7-gram の制限を外した。

また、これまでの N-gram 処理を形態素単位とする処理は形態素解析を前提としたことから、新規語彙や方言などの特殊語彙処理に対応することができなかったが、上記のリスト ID アブストラクションによる処理の効率化によりこの制限を取り除くことができた。この結果、これまでの統合文脈単語 N-gram 処理を拡張し、形態素解析によらない文字単位での処理を行う統合文脈文字 N-gram 処理を行えるようになった。この処理によって、既存の形態素解析が単語辞書に依存している依存性を除去できただけではなく、形態素解析が既存文法規則に依存している部分をも脱却することができた。

今後の見通し： 従来の言語研究が欧米での文法規則の演繹的適用を主としてきたのに対し、統合文脈文字 N-gram 処理による大量の言語データ処理は、機能的に規則を再構築する事によって、現行言語学の限界を超克することができる。例えば言語教育ではビジネス英語や法律英語、医学英語など文脈ごとの表現の重要性は従来から広く理解され､ESP（English for Specific Purpose）と呼ばれる分野の研究が行われてきたが、この研究でも従来の定形表現研究と同様の限界が残っている。統合文脈文字 N-gram 処理を用いれば、対象とする言語集合ごとの特徴的表現の直接抽出が可能となり、言語テキスト研究、並びに言語教育学の新地平を開くことが可能となるであろう。

(2)　『リグ・ヴェーダ』のフレーズ分析

インド古典すべての源泉となり、現代に至るまで常に参照されてきた『リグ・ヴェーダ』のフレーズ分析を行った。底本は GRETIL 公開版（Input by Barend A. Van Nooten and Gary B. Holland, converted and revised by Detlef Eichler）を用い、下記のような 2 種の表（インデックス）を作った。

表 1 は 1 行に、A 列 : 抽出フレーズ、B 列 : 抽出フレーズの単語数、C 列 : 所在詩節数、D 列 : 所在詩節番号、を示す。約 1 万詩の『リグ・ヴェーダ』からこのような「フレーズ」が 23,705 行抽出された。インデックスにおいて語数の降順に並べた時、先頭に来るのが下記の 1 行である。

表 1

| A | B | C | D |
|---|---|---|---|
| yathā ha tyad vasavo gauryaṃ cit padi ṣitām amuñcatā yajatrāḥ evo ṣv asman muñcatā vy aṃhaḥ pra tāry agne prataraṃ na āyuḥ | 22 | 2 | RV_10,126.08;RV_4,012.06 |

『リグ・ヴェーダ』10 巻の成立は、Michael Witzel (‘The Development of the Vedic Canons and its Schools’, 1997)によれば、次の 4 層に分けられる。

I. B.C.1700-1450：2〜6 巻の一部
II. B.C.1450-1300: 2, 4, 5, 6 巻及び 3 巻、7 巻、8 巻 1〜66、1 巻 51〜91
III. B.C.1300-1200：8 巻 8.67〜103、1 巻 1〜50、10 巻 1〜84（Atharvaveda like）、10 巻 85〜191
IV. 9 巻：2〜7 巻から神酒ソーマ讃歌を抜粋

表 1 の詩節は編纂の最終段階で付加された 10 巻が 4 巻から借用したもので、10 巻にはこのように古い詩篇から 1 詩全体あるいは常套句を借用したものが多い。また 9 巻も古い詩篇からソーマ讃歌のみを抜粋して 1 巻としたため、元の詩篇との共通フレーズが少なからず見つかる。

表 2 は一フレーズの出現個数を各巻ごとに示したインデックスの一行。単語数 5、出現個所総数 82 である。7 巻の 76 カ所、9 巻と 10 巻のそれぞれ 3 カ所にこのフレーズが在ることを示す。

表 2

| Phrases | Ngr | Frq | I | II | III | IV | V | VI | VII | VIII | IX | X |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| yūyam pāta svastibhiḥ sadā naḥ | 5 | 82 | 0 | 0 | 0 | 0 | 0 | 0 | 76 | 0 | 3 | 3 |

表 2 の 1 行は、7 巻の主要部を占めるヴァシシュタ家の讃歌のこの反復句、「あなたさま方（神々）は常に我らを平安のうちにお護り下さいますよう」が、9 巻に集められたソーマ讃歌 3 詩にも見られ、また後代には 10 巻に 3 度借用されていることを示す。自動抽出された全フレーズを納めるこのインデックスは、このように諸巻の関係を鮮明に示すことができる。

『リグ・ヴェーダ』の讃歌は祭官の各家系がそれぞれに伝承してきたため、10 巻や 9 巻のような特殊な場合を除き、独自に工夫、創案した表現を原則として遵守する。それが他家系の巻と共通する場合には、両家系の特別な関係か、あるいは後代における一方の借用が示唆される。10 語以上の抽出フレーズの中で、10 巻、9 巻を除く重複は、次の両巻に起こっている。1＝3、1＝4（2 回）、1＝5、1＝6、1＝9、2＝6、3＝4、3＝6（２回）、3＝7（３回）、4＝8、6＝7、7＝8。これは 1 万という総詩節数からすると極めて少なく、祭官家が独自性を保つことに腐心したことがこれによっても判明する。しかしまさにその故に、抽出された共通フレーズは各巻、各詩篇、各家系の関係を示唆する貴重な情報であり、この表を基に今後の『リグ・ヴェーダ』研究は一層効率的に、また理解精度を高めて遂行することが可能となった。

本研究は、新開発の統合文脈文字 N-gram 処理が、正確な口頭伝承、作者・年代の不詳、一文献の多層構造等の特徴を有するインド古典の分析に極めて有効であり、この技法をインド古典学に取り入れた時の極めて高い効果を示し得たと考える。

インド古典は『リグ・ヴェーダ』（前 18〜前 13 世紀）、『パーリ仏典』（前 5〜後 6 世紀）、『マハーバーラタ』（前 3〜後 4 世紀）等々、長期に亘って作り続けられたものが多い。従来の研究は各文献中の諸篇が長期に亘って順次制作されたことを十分考慮せず、一括して扱うことが多かった。しかし、フレーズ分析が、韻律分析と共に、諸古典の相対年代や相互関係、一文献中の諸篇の関係等の解明に資することが判明した以上、今後のインド古典研究は、それらの相対年代や関係性を特定した上で遂行する必要があろう。すなわちフレーズ分析と韻律分析は、今後のインド古典学に必須の手法となるであろう。

ただし、付言するならば、インド古典データベースが内包する問題も今回初めて明らかになった。現在オンライン公開されているテキストデータベースは、同一テキストの中で一つの表記、例えば á を、しばしば異なる方式で入力している。これは多数の研究者が順次部分的修正を加えた結果であるが、画面上は同一に見えるので、その標準化作業に時間を要した。この予測しなかった事態のため、研究に若干の遅れが生じ、現在は『リグ・ヴェーダ』の処理のみが終了している。この研究成果とインデックスは Harvard 大学の電子雑誌に発表する予定である。

５．主な発表論文等

〔雑誌論文〕　計12件（うち査読付論文　10件／うち国際共著　0件／うちオープンアクセス　2件）

| | |
|---|---|
| １．著者名<br>Minako Nakamura, Kohji Shibano | ４．巻<br>1 |
| ２．論文標題<br>Mining formulaic sequences from a huge corpus of Japanese TV closed caption | ５．発行年<br>2020年 |
| ３．雑誌名<br>DH(Digital Humanities), Budapest 2019 | ６．最初と最後の頁<br>未定 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| | |
|---|---|
| １．著者名<br>中谷英明 | ４．巻<br>43 |
| ２．論文標題<br>インド・チベット古典学と日仏東洋学会 | ５．発行年<br>2020年 |
| ３．雑誌名<br>日仏東洋学会通信 | ６．最初と最後の頁<br>29-46 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>無 |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である） | 国際共著<br>－ |

| | |
|---|---|
| １．著者名<br>Kohji Shibano | ４．巻<br>1 |
| ２．論文標題<br>Analyzing formulaic sequences in spoken Japanese from a large Japanese TV closed caption corpus | ５．発行年<br>2017年 |
| ３．雑誌名<br>The 18th World Congress of Applied Linguistics (AILA 2017), 23-28 July 2017 | ６．最初と最後の頁<br>- |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| | |
|---|---|
| １．著者名<br>Kohji Shibano | ４．巻<br>1 |
| ２．論文標題<br>Developing Intimacy by Style-shifting in Japanese: A TV Subtitle Corpus-based Study, XIAO Tingting, | ５．発行年<br>2017年 |
| ３．雑誌名<br>The 2017 conference of the American Association for Applied Linguistics (AAAL 2017), 18-21 March, 2017 | ６．最初と最後の頁<br>- |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| | |
|---|---|
| 1．著者名<br>Hajime Mochizuki, Kohji Shibano | 4．巻<br>1 |
| 2．論文標題<br>Analyzing Usefulness of Dialogues from Closed Caption TV Corpus as an Example of Can-do Statements for Language Learning | 5．発行年<br>2018年 |
| 3．雑誌名<br>2018 Hawaii University Conference, Arts, Humanities, Social Sciences & Education (AHSE), Hawaii, USA | 6．最初と最後の頁<br>- |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスとしている（また、その予定である） | 国際共著<br>－ |

| | |
|---|---|
| 1．著者名<br>Hajime Mochizuki, Kohji Shibano | 4．巻<br>1 |
| 2．論文標題<br>Searching Discourse Segments for Formulaic Sequences in a Closed Caption TV Corpus for Language Learning, | 5．発行年<br>2017年 |
| 3．雑誌名<br>World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education | 6．最初と最後の頁<br>19-27 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| | |
|---|---|
| 1．著者名<br>Hajime Mochizuki, Kohji Shibano | 4．巻<br>1 |
| 2．論文標題<br>Discourse Segment Clustering with Word Embedding based on Formulaic Sequences for Language Education | 5．発行年<br>2017年 |
| 3．雑誌名<br>International Conference on Education and Multimedia Technology (ICEMT 2017) | 6．最初と最後の頁<br>- |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| | |
|---|---|
| 1．著者名<br>Hajime Mochizuki, Kohji Shibano, | 4．巻<br>1 |
| 2．論文標題<br>The Acquisition of a Japanese Practical Formulaic Sequences List from a Closed Caption TV Corpus | 5．発行年<br>2017年 |
| 3．雑誌名<br>Hawaii University Conferences, STAM/STEAM Education Conference | 6．最初と最後の頁<br>- |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| 1．著者名<br>Hajime Mochizuki, Kohji Shibano | 4．巻<br>1 |
|---|---|
| 2．論文標題<br>Extracting Formulaic Sequences Containing Useful Expressions for Language Learning from Closed Caption TV Corpus | 5．発行年<br>2016年 |
| 3．雑誌名<br>World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, E-Learn 2016, Alexandria, USA | 6．最初と最後の頁<br>29-37 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| 1．著者名<br>Hajime Mochizuki, Kohji Shibano | 4．巻<br>1 |
|---|---|
| 2．論文標題<br>Modification of word2vec by Formulaic Sequences and Extraction of Useful Expressions for Language Learning from Closed Caption TV Corpus | 5．発行年<br>2017年 |
| 3．雑誌名<br>The IAFOR International Conference on Language Learning, Hawaii 2017, IICLLHawaii 2017, Honolulu, USA | 6．最初と最後の頁<br>ポスター発表 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| 1．著者名<br>Kohji Shibano | 4．巻<br>1 |
|---|---|
| 2．論文標題<br>Developing Intimacy by Style-shifting in Japanese: A TV Subtitle Corpus-based Study, XIAO Tingting | 5．発行年<br>2017年 |
| 3．雑誌名<br>The 2017 conference of the American Association for Applied Linguistics (AAAL 2017) | 6．最初と最後の頁<br>印刷中 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>有 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

| 1．著者名<br>Kohji Shibano | 4．巻<br>1 |
|---|---|
| 2．論文標題<br>Analyzing formulaic sequences in spoken Japanese from a large Japanese TV closed caption corpus | 5．発行年<br>2017年 |
| 3．雑誌名<br>The 18th World Congress of Applied Linguistics (AILA 2017), 23-28 July 2017, Rio de Janeiro, Brazil | 6．最初と最後の頁<br>印刷中 |
| 掲載論文のDOI（デジタルオブジェクト識別子）<br>なし | 査読の有無<br>無 |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難 | 国際共著<br>－ |

〔学会発表〕　計14件（うち招待講演　5件／うち国際学会　10件）

１．発表者名
中谷英明

２．発表標題
ブッダの「自省利他」

３．学会等名
龍谷大学創立380周年記念講演会（招待講演）

４．発表年
2019年

１．発表者名
中谷英明

２．発表標題
世界の人々のアイデンティティとしての「自省利他」の研究－社会実装を視野に入れ て

３．学会等名
科研費挑戦的研究（開拓）「世界の人々のアイデンティティとしての「自省利他」の研究－社会実装を視野に入れ て」第1回研究会（招待講演）

４．発表年
2019年

１．発表者名
Minako Nakamura, Kohji Shibano

２．発表標題
Mining formulaic sequences from a huge corpus of Japanese TV closed caption

３．学会等名
DH(Digital Humanities), Budapest 2019, Hungary（国際学会）

４．発表年
2019年

１．発表者名
中谷英明

２．発表標題
自省利他ー仏陀の哲学 Introspective Altruism－ Philosophy of the Buddha －

３．学会等名
仏光山大学校長論壇、台湾（招待講演）（国際学会）

４．発表年
2019年

１．発表者名
中谷英明

２．発表標題
自省利他の思想 ―『スッタニパータ』八頌品における釈尊の教え

３．学会等名
駒澤大学成道会法要記念講演（招待講演）

４．発表年
2019年

１．発表者名
中谷英明

２．発表標題
インド古典学・チベット古典学と日仏東洋学会

３．学会等名
日仏東洋学会（招待講演）

４．発表年
2019年

１．発表者名
中谷英明

２．発表標題
八頌品（アッタカ・ヴァッガ）の韻律と思想

３．学会等名
日本印度学仏教学会第68回学術大会（国際学会）

４．発表年
2017年

１．発表者名
Kohji Shibano

２．発表標題
nalyzing formulaic sequences in spoken Japanese from a large Japanese TV closed caption corpus

３．学会等名
The 18th World Congress of Applied Linguistics（国際学会）

４．発表年
2018年

１．発表者名
Kohji Shibano

２．発表標題
eveloping Intimacy by Style-shifting in Japanese: A TV Subtitle Corpus-based Study, XIAO Tingting,

３．学会等名
The 2017 conference of the American Association for Applied Linguistics（国際学会）

４．発表年
2017年

１．発表者名
Hajime Mochizuki, Kohji Shibano

２．発表標題
Analyzing Usefulness of Dialogues from Closed Caption TV Corpus as an Example of Can-do Statements for Language Learning,

３．学会等名
2018 Hawaii University Conference, Arts, Humanities, Social Sciences & Education（国際学会）

４．発表年
2017年

１．発表者名
Hajime Mochizuki, Kohji Shibano

２．発表標題
Discourse Segment Clustering with Word Embedding based on Formulaic Sequences for Language Education

３．学会等名
International Conference on Education and Multimedia Technology（国際学会）

４．発表年
2017年

１．発表者名
Hajime Mochizuki, Kohji Shibano

２．発表標題
The Acquisition of a Japanese Practical Formulaic Sequences List from a Closed Caption TV Corpus,

３．学会等名
2017 Hawaii University Conferences, STAM/STEAM Education Conference（国際学会）

４．発表年
2017年

| |
|---|
| １．発表者名<br>Kohji Shibano |
| ２．発表標題<br>Developing Intimacy by Style-shifting in Japanese: A TV Subtitle Corpus-based Study, XIAO Tingting |
| ３．学会等名<br>The 2017 conference of the American Association for Applied Linguistics (AAAL 2017)（国際学会） |
| ４．発表年<br>2017年 |

| |
|---|
| １．発表者名<br>Kohji Shibano |
| ２．発表標題<br>Analyzing formulaic sequences in spoken Japanese from a large Japanese TV closed caption corpus |
| ３．学会等名<br>The 18th World Congress of Applied Linguistics (AILA 2017), 23-28 July 2017, Rio de Janeiro, Brazil（国際学会） |
| ４．発表年<br>2017年 |

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

| | 氏名<br>（ローマ字氏名）<br>（研究者番号） | 所属研究機関・部局・職<br>（機関番号） | 備考 |
|---|---|---|---|
| 研究分担者 | 芝野　耕司<br>(Shibano Kohji)<br>(50216024) | 東京外国語大学・その他部局等・名誉教授<br>(12603) | |