

令和元年6月5日現在

機関番号：32404
 研究種目：基盤研究(C) (一般)
 研究期間：2015～2018
 課題番号：15K02185
 研究課題名(和文)京阪座敷音曲の中の舞踏的身体 テキストマイニングによる花柳界の地唄と上方唄の分析

研究課題名(英文)Diuta and Kamigata-mai - Identification and classification of lyrics using text-mining technique -

研究代表者
 矢島 ますみ (YAJIMA, Masumi)

明海大学・経済学部・教授

研究者番号：80220135

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究は、座敷舞(上方舞)の特徴を、舞踏で使用する地唄・上方唄の歌詞の視点からテキストマイニングによる手法を使って明らかにすることを目的としている。収集した地唄舞の歌詞のテキストデータ約600曲に対して、舞踏付きと無しの差異、流派ごとの曲目の異についてその特徴分析を行った。一部混乱しているところもあるが、おおまかに舞踏付きと舞踏無しのグループにクラスタリングされること、および流派ごとの曲風の差異を確認した。

研究成果の学術的意義や社会的意義

本研究では、舞踏や伝統芸能などの無形文化財を、テキストマイニング等の数理的な手法を応用することで、言語文化と密接にリンクする日本の豊かな芸能諸側面の特徴を俯瞰的に浮かびあがらせようとする、および、数理的な分析技術、音楽学、舞踏学の3者が交わる学際的なアプローチであることが学術的に意義深い。特に、歴史が深く、芸妓文化として特徴的な京阪座敷の舞踏文化の数理的な言語分析は、「おもてなし日本」の“お座敷文化に潜む舞踏的身体”を明らかにする新たな読み解きである。

研究成果の概要(英文)：The main object of this research is to identify and classify lyrics from the point of view of whether it is used for dance in DIUTA(classical Shamisen song) of 600 songs, by using text-mining techniques. Even some of analysis results were confusing, they were generally clustered into dance music and non-dance music groups. Furthermore, the lyrics of the "Kamigata-mai" schools was clearly identified to each other.

研究分野：教育学、社会学をベースとした舞踏表現論

キーワード：地唄 上方舞 テキストマイニング 頻度情報 クラスタリング

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

京阪の花柳界は、「地唄」や「上方唄」の他、端唄・小唄等の演奏とともに、<舞>の世界を守り発展させてきた。「地唄」は伝統的な声楽としての側面をもち、「上方唄」は、軽快で明るい曲調の京阪都市部の三味線音楽である。「寛延以降地歌唄本収録曲作曲者索引(1973)」で文献上確認できる「地唄」は1,000以上あり、その中で現存する曲は、約500曲(「現存する地歌・箏曲の曲目総覧(1991)」)と言われる。舞踊伴奏曲の数は、『日本舞踊曲集成 2 京舞・上方舞編(演劇出版社 2005)』によると、「地唄」は127曲、「上方唄」は88曲 合計215曲が舞踊曲となっている。上方舞(地唄舞)の流派である井上流・吉村流・山村流・煤茂都流・神崎流・篠塚流等によって、演じられている。

本研究では、京阪独自の座敷を支える音曲や舞踊の歴史的発展の概略を明らかにし、そして、<舞踊唄>および<演奏唄>の約600-700曲のデータの整理、歌形式・唄内容の分類を行った後、テキストマイニング手法を用い“唄ことば”の視点から、京阪の座敷舞踊の世界観の特徴を発見していく。

2. 研究の目的

2012-2014の科学研究では、江戸(東京)に発展した「江戸小唄」を対象に“定型化されていない文章の集まりを自然言語解析の手法を使って分析して有用な情報を抽出”するテキストマイニング手法(text mining)を使い、“舞踊小唄”と“小唄”の唄ことばの比較によって、意気(いき)でいなせな江戸情緒性があらわれる舞踊的身体世界の特徴分析を試みた。

今回の研究でも、同様な視点から、京都・大阪の座敷音曲である「地唄と上方唄」を対象に、“舞踊伴奏用に演奏される音曲(以後<舞踊唄>)”と、“純粋に演奏される音曲(以後<演奏唄>)”の特徴をテキストマイニング分析によって抽出し、その“ことば世界”を比較することによって、京阪の花柳文化の中核をなす座敷舞踊文化に表れる上方の粋(すい)な世界観の特徴を客観的な指標によって明らかにしようとするものである。

3. 研究の方法

本研究の目的は、京阪の「地唄・上方唄」の<舞踊唄>と<演奏唄>の“唄ことば”の特徴をテキストマイニングによって分析することである。まず、京阪の座敷に特有な音曲の内容を整理するために関係者へのヒアリングと資料収集、そして、収集した資料から600-700曲の歌を整理、分類し、データベースを作成した。全体約600曲で、その内舞踊で利用される曲約150曲であった。

分析の視点は2つ

- (1) 舞踊唄と演奏唄の歌詞の違い(特徴)
- (2) 流派(家元)の違いによる歌詞の違い(特徴)

まず、利用している語彙について、名詞と形容詞を対象を絞った上で頻度情報をRMeCabを用いて分析した。RMeCabとは統計解析環境であるRから、日本語形態素解析器MeCabを利用するためのインターフェイスであり、形態素解析と統計解析処理を統合的に行えるテキストマイニングツールである。

4. 研究成果

(1) 舞踊唄と演奏唄の歌詞の違い

まず、舞踊唄と演奏唄の歌詞の違いについてのクラスター分析について説明する。クラスター分析とは、データの各個体(本研究においては歌詞)間の内容の類似度を距離とみなし

て、その距離を基準にグループ分けする手法である。

本分析においては、まず舞踊唄と演奏唄それぞれについて14個の文書ファイルを作成し、RMeCabにより名詞、形容詞を抽出した。また、データ間距離についてはキャンベラ距離を導入した。キャンベラ距離は、次式で定義されるものである。

$$\sum_{i=1}^n \frac{|x_i - y_i|}{|x_i + y_i|}$$

ここで、 i は抽出されたターム(形態素)の番号、 x_i, y_i はそのタームのそれぞれの文書での出現頻度である。また、クラスター作成方法については分散を最小にするよう併合するウォード法を用いた。ウォード法は、クラスター内での分散が最も小さくなるように、結合を順次決定する手法である。本研究では抽出したデータに対して階層的クラスター分析を適用した。

RMeCabを用いて分析を行った結果は図1の通りである。buyouで始まるものが舞踊唄のファイルであり、diutaで始まるものが演奏唄のファイルである。舞踊唄がおおよそ左に、演奏唄がおおよそ右側に集まっており、それぞれのファイルがおおまかではあるがクラスターを構成している。

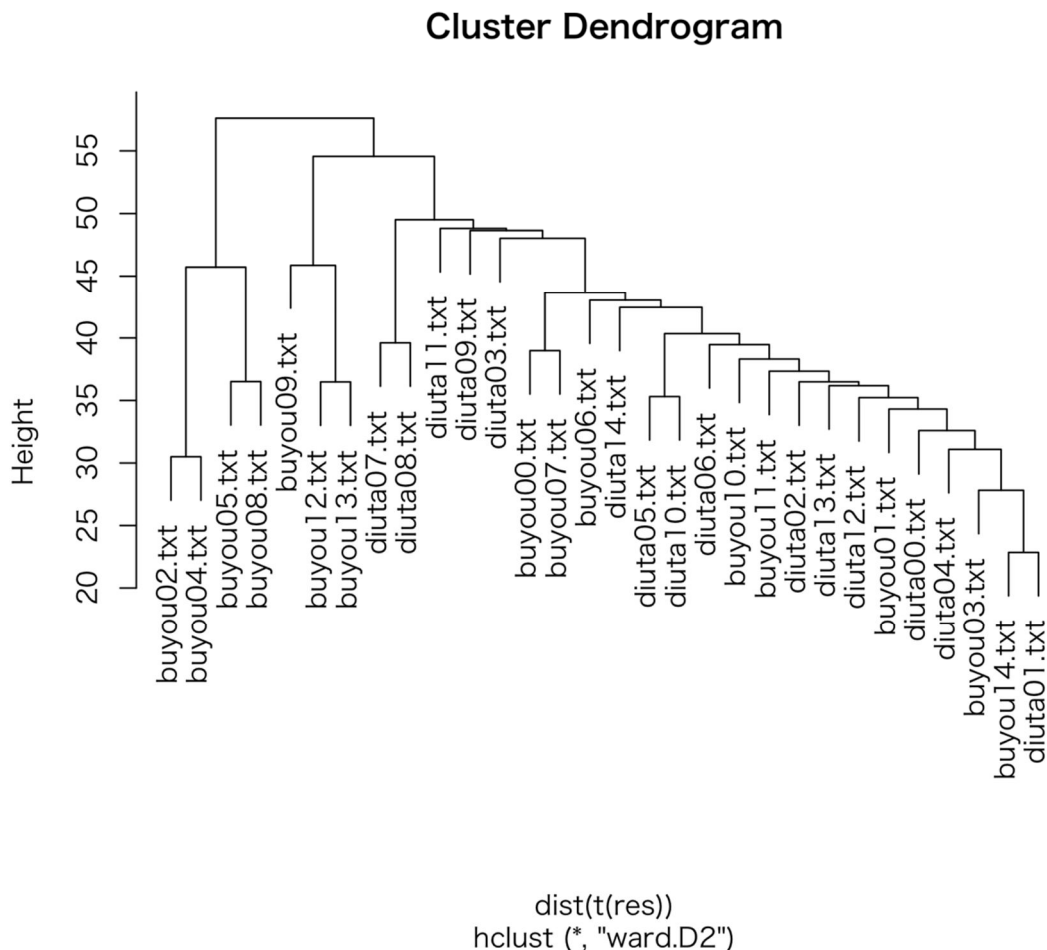


図1：クラスター分析出力(舞踊唄と演奏唄)

(2) 流派(家元)の違いによる歌詞の違い

次に、流派(家元)の違いによる上演曲の歌詞の特徴(違い)について、クラスター分析を試みた。上演曲データは、独立行政法人日本芸術文化振興会「文化デジタルライブラリー」に掲載されている演目である(国立劇場、国立文楽劇場において、これまでに上演された舞踊・邦楽公演の自主公演の演目データ)。

本分析においては、吉村流四世家元 吉村雄輝、吉村流六世家元 吉村輝章、吉村流五世家元 吉村雄輝夫、井上流四代目井上八千代、井上流五代目井上八千代、山村流三代目山村若のそれぞれの家元の文書ファイル(歌詞データファイル)を作成し、RMeCabにより名詞、形容詞を抽出した。また、4.1 同様データ間距離についてはキャンベラ距離を導入し階層的クラスター分析を行った。その結果を、図2に示す。

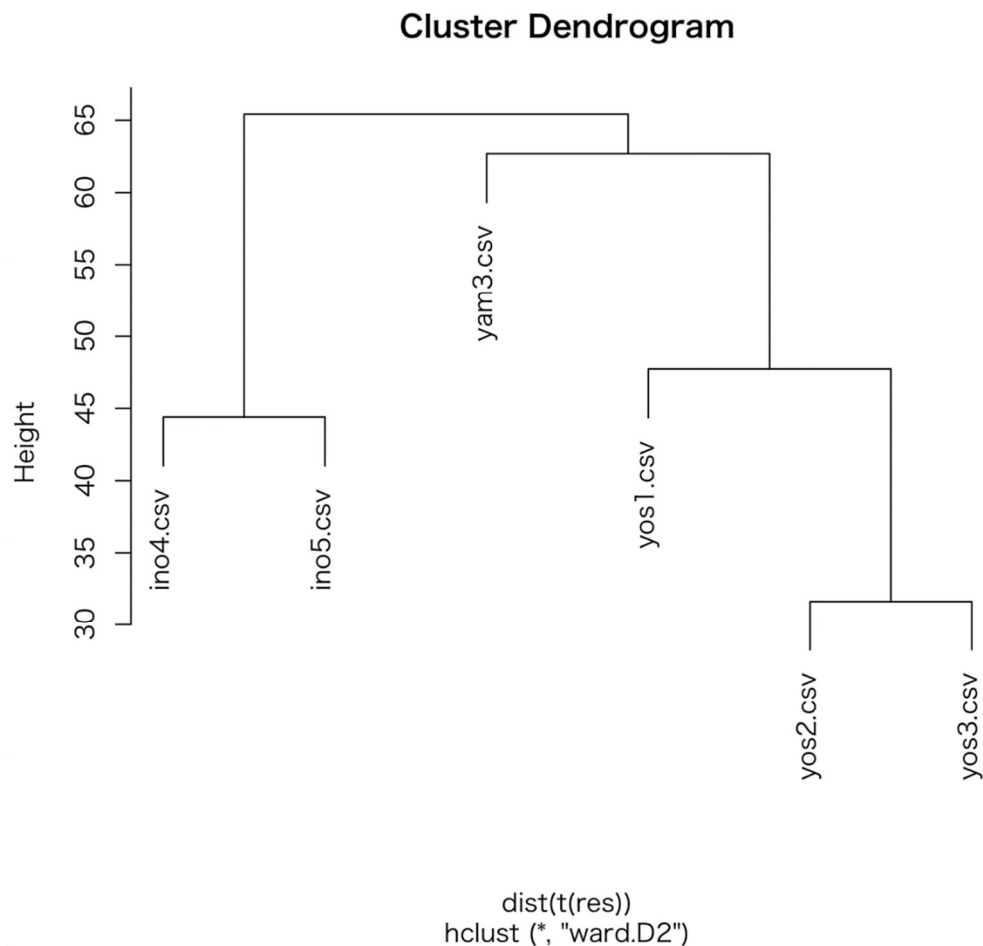


図2: クラスター分析出力(流派)

図の中のファイル名の対応関係は以下の通りである。

yos1: 吉村流四世家元吉村雄輝 / yos2: 吉村流六世家元吉村輝章 /
yos3: 吉村流五世家元吉村雄輝夫 / ino4: 井上流四代目井上八千代 /
ino5: 井上流五代目井上八千代 / yam3: 山村流三代目山村若

吉村流3名、井上流2名、山村流の3つのクラスターにきれいに分類されており、流派内の類似度は明らかである。これにより、テキストマイニング技術による自動分類の可能性が認められた。

(3) 結果・考察・結論

結果の抜粋(単語 品詞 属性 頻度)を以下に示す。

舞踊唄と演奏唄の歌詞の違い(特徴)

・舞踊付き:花 名詞 一般 114 / 心 名詞 一般 83 / 身 名詞 一般 76 / 夜 名詞 副詞可能 69 / さ 名詞 接尾 66 / ない 形容詞 自立 71 / よい 形容詞 自立 41 / 憂い 形容詞 自立 23 / 辛い 形容詞 自立 18 / 深い 形容詞 自立 18

・舞踊無し:花 名詞 一般 218 / 心 名詞 一般 173 / 身 名詞 一般 163 / 月 名詞 一般 121 / 春 名詞 一般 121 / ない 形容詞 自立 113 / 無い 形容詞 自立 75 / 憂い 形容詞 自立 53 / 深い 形容詞 自立 36 / 嬉しい 形容詞 自立 35

全体的に大きな差異は確認できなかったが、例えば舞踊唄に関しては形容詞として「辛い」、演奏唄に関しては「嬉しい」といった単語がもう一方に比べて特徴的に使われていることがわかった。さらに、収集したテキストデータを舞踊唄と演奏唄についてワード法による階層的クラスタリングを行ったところ、一部混乱しているところもあるが、おおまかに舞踊唄と演奏唄のグループにクラスタリングされることを確認した。

流派(家元)の違いによる歌詞の違い(特徴)

家元別の上演作品の歌詞の特徴を表にすると、以下のようである。

| | 井上流四代目 井上八千代 (11曲) | 井上流五代目 井上八千代 (22曲) | 山村流三代目 山村若 (30曲) | 吉村流四世家元 吉村雄輝 (27曲) | 吉村流六世家元 吉村輝章 (20曲) | 吉村流五世家元 吉村雄輝夫 (17曲) |
|-----|--------------------|--------------------|------------------|--------------------|--------------------|---------------------|
| 名詞 | 夜 9 | 花 23 | 花 34 | 夜 20 | 夜 17 | 夜 14 |
| | 声 7 | やま 19 | 夜 24 | 花 17 | 音 15 | それ 12 |
| | 誰 6 | 心 18 | やま 21 | 心 17 | 身 14 | 身 12 |
| | 音 6 | 今 15 | 月 19 | 水 16 | 子 14 | 水 11 |
| | 花 6 | 夜 14 | 心 19 | やま 15 | 花 13 | 音 10 |
| 形容詞 | ない 4 | ない 16 | ない 20 | よい 14 | ない 8 | ない 8 |
| | 高い 2 | よい 9 | よい 17 | ない 14 | 恐ろしい 5 | めでたい 5 |
| | 長い 2 | 深い 6 | よい 17 | 良い 4 | めでたい 5 | 恐ろしい 4 |
| | 薄い 2 | 目出度い 5 | 憂い 6 | 憂い 4 | 憂い 4 | 安い 4 |
| | 清い 2 | 憂い 4 | 恐ろしい 5 | やさしい 4 | 安い 4 | よい 3 |
| 動詞 | する 15 | する 36 | する 34 | する 42 | する 33 | する 32 |
| | なる 9 | 見る 22 | なる 27 | 見る 22 | なる 21 | なる 14 |
| | 見る 7 | なる 20 | 見る 21 | なる 16 | れる 16 | れる 14 |
| | 思う 5 | れる 20 | 散る 16 | れる 15 | 見る 13 | 追う 10 |
| | れる 5 | 知る 11 | れる 15 | 思う 14 | ける 12 | ける 9 |

考察・結論

上方舞は心の内面を動きで表現する芸能であり、西形節子氏の言う「関東の踊には拍子にのるリズム的な踊と、物真似の振の要素が強く、関西の舞は表現を内にこめた舞の伝統が土台となって展開してきたものである。¹⁾」という点で、関西の三味線歌である地唄のなかでも地唄舞に使用される曲の歌詞に「辛い」という単語の特徴があることは、秘めた心情の表現に適している曲が使用されていることが考察できる。また、流派(家元ごと)の比較では、上記表の歌詞の一覧からは明確に特徴を読み取ることは難しいが、吉村流3名、井上流2名、山村流の3つのクラスターにきれいに分類されたことは、各流派が確固たる芸風(演目特徴)

¹⁾ <引用文献>

西形節子、「昭和62年度春季第23回 舞踊学会講演記録 上方舞の系譜」、1987年度春季第23回 舞踊学会 『舞踊學』,10号,1987, p55-61.

を持っていることを示すものであり、女性のみにて伝承してきた井上流、近年男性を家元としながらも女性舞を伝承してきた吉村流、そして歌舞伎の流れがある山村流、それぞれの独自性をもった芸風があることが確認された。

今後の課題としては、吉村流・井上流・山村流以外の流派のデータ分析も手掛け、京阪の舞の世界観を浮き彫りにしていきたい。

5 . 主な発表論文等

〔雑誌論文〕(計 3 件)

T.Maeda, Y.Fukushige, M.Yajima and A.Wakatani, Multi-Lingual Pseudo Natural Language Programming Environment for Beginners, Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2018), 査読有, 1,2018, in USB stick.

T.Maeda and M.Yajima and A.Wakatani, Personalized Mobile Rapid Sentence Reading with Multi-lingual Extensio, The 28th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE PIMRC 2017), IEEE PIMRC 2017, 査読有, 2017, in USB memory.

T.Maeda and M.Yajima, Motion Frequency Data Analysis for Sports Skill, Proceedings of 2017 Computing Conference (CC 2016), 査読有, CC 2016, 2017, 113-116.

〔学会発表〕(計 3 件)

Toshiyuki MAEDA and Masumi YAJIMA, Sports Skill Discrimination with Motion Frequency Analysis, The 2016 Future Technologies Conference (国際学会), 2016年12月07日, San Francisco, CA (USA).

Toshiyuki MAEDA and Masumi YAJIMA, Skill Identification Using Time Series Data Mining, The 2016 International Conference on Data Mining (国際学会), 2016年07月27日, Las Vegas.

T.Maeda and M.Yajima, Personal Skill Identification Using Time Series Motion Picture Data, The 41th Annual Conference of the IEEE Industrial Electronics Society (IECON 2015)} (国際学会), 2015年11月09日~2015年11月12日, Yokohama (Japan).

〔図書〕(計 0 件)

6 . 研究組織

(1)研究分担者

研究分担者氏名：前田 利之

ローマ字氏名：(MAEDA, toshiyuki)

所属研究機関名：阪南大学

部局名：経営情報学部

職名：教授

研究者番号(8桁)：70320041

研究分担者氏名：長谷川 慎

ローマ字氏名：(HASEGAWA, makoto)

所属研究機関名：静岡大学

部局名：教育学部

職名：准教授

研究者番号(8桁)：00466971

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。