

地方議会会議録におけるオノマトペの出現傾向に関する基礎的検討

—少数の自治体に高頻度で出現するオノマトペについて—

高丸圭一¹ 内田ゆず² 乙武北斗³ 木村泰知⁴^{*1}宇都宮共和大学 ^{*2}青山学院大学 ^{*3}福岡大学 ^{*4}小樽商科大学^{*1}takamaru@kyowa-u.ac.jp ^{*2}yuzu@ee.aoyama.ac.jp ^{*3}ototake@fukuoka-u.ac.jp ^{*4}kimura@res.otaru-uc.ac.jp

1. はじめに

筆者らは、地方自治体がウェブに公開している地方議会会議録を収集・整形し、関係データベースに登録することにより、コーパスとして学際的に利用することを目指した研究を進めている[1][2]。地方議会会議録コーパスは都道府県または市区町村議会における議員や首長、行政職員などの発言の記録である。属性(年齢・性別・肩書きなど)が明らかで、かつ、特定の自治体に居住する者の発言が、地域別・年度別に記録されている。地方議会会議録に何どのように記録されているかについて研究が進められている。政治学[3]、環境学[4]、図書館情報学[5]などの分野で、会議録を利用して地方議会の動向を探る研究が行われている。また、高丸ら[6]や井上[7]は会議録に含まれる文末表現の地域差や方言について指摘している。

本研究では、地方議会会議録におけるオノマトペの出現傾向について検討する。オノマトペ(擬音語および擬態語)は音、雰囲気、程度、様子を効果的に伝える手段であり、日本語の話しことばでは多用されることが知られている。国語学や音声学の分野では、語源や意味、構造などが古くから研究されている(例えば[8][9][10])。また、方言オノマトペの収集も行われており[11][12]、特に医療の現場において症状を表すオノマトペの地域差の問題が重視されている。近年、オノマトペの利活用に着目した研究が盛んになりつつある。平田ら[13]は国会会議録を対象として、議員の出身地別に使用するオノマトペの頻度を比較し、オノマトペ全体では出現頻度に地域差が見られないが、「CVQCVリ」型の4拍のオノマトペが関西で多く出現することを指摘した。

地方議会会議録は同じ地域に居住する者同士の議論を記録した資料であるため、オノマトペの地域差の分析にも有効であると考えられる。本稿ではまず、地方議会会議録というやや改まった話しことばの記録にどのようなオノマトペが出現するのかを確認する。その上で、出現頻度を調査し、特徴を分析する。どの自治体にも出現するオノマトペではなく、出現する自治体が限定的であり、かつ、一定の頻度で出現するオノマトペを分析対象とする。

2. 研究の対象と方法

2.1. 地方議会会議録

プログラムによる自動処理によって収集、データベース化した地方議会会議録コーパスを利用する。地域差のみを見るために、使用するデータは開催時期が2010年度のものに限定する。分析対象はすべての都道府県を網羅した402自治体の会議録(約3億語)である。

地方議会会議録は、議会での発言をすべて記録することを目的としている。しかし、議会を円滑に運営する目的で、議員の発言(質問)内容は事前に通告されており、読み上げ原稿が存在する発言が含まれる。また、正文の作業によって話しことばの特徴の一部が書きことば的に修正されている[14]。この2つの点において、会議録は厳密には自由会話の書き起こし資料であるとはいえない。議会会議録から話しことばの特徴を分析する際にはこの点に注意が必要となるが、正文の指針[15]によると、オノマトペは正文によって削除される対象ではないため、本研究における正文の影響は少ないと考える。

2.2. オノマトペ

『日本語オノマトペ辞典』[16]に掲載された4,500語のうち、意味分類索引に掲載された2,466語(異なり語数1,751語)を対象とする。これらのオノマトペは3種類の大分類、12種類の中分類、70種類の小分類に意味分類されている。意味分類の概要と該当するオノマトペの語数を表1に示す。

2.3. 形態素解析と出現確率の計算

2.1で述べた対象データに対して形態素解析処理を行った。形態素解析には日本語形態素解析システムJUMAN¹を用いた。また、オノマトペ用辞書はJUMANのユーザ形態素辞書に2.2の語を副詞として登録したものを利用した。その解析結果から、形態素として2.2のオノマトペが含まれている発言を抽出した。

都道府県ごとにコーパスに収録された発言数が異なるため(もっとも多い東京都は34,343,558語、最も少ない和歌山県は543,641語)、抽出したオノマトペの出現頻度そのものを都道府県間で直接比較する

¹ <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

表1 意味分類の概要とオノマトペの語数

	大分類			総数
	「自然」	「人間」	「事物」	
中分類	「天気」 「温度」 「水・液体」 「火・土」	「動作・状態」 「感情・感覚」 「性格・性質」 「体格・姿」	「音・道具・金銭」 「形・状態」 「程度」 「動き・変化」	3
小分類の数	12	27	31	70
延べ語数	261	939	1,266	2,466
異なり語数	228	801	1,050	1,751

ことは適当ではない。オノマトペの用例を比較するため、出現確率を計算した。オノマトペ_wの都道府県gにおける出現確率 $P_{g,w}$ は、発言の総単語数 N_g 、オノマトペの出現頻度 $C_{g,w}$ を用いて以下の式によって求めた。

$$P_{g,w} = \frac{C_{g,w}}{N_g}$$

3. 全体的な傾向

3.1. 形態素解析結果の概要

形態素解析の結果、約3億語の地方議会会議録コーパスから186,416語（異なり語数982語）のオノマトペが抽出された。出現確率の全国計上位10語を表2に示す。「しっかり」「どンドン」「はっきり」「だんだん」などが多く出現した。特に「しっかり」の出現確率は非常に高く、抽出したオノマトペ全体の約41%を占める。このように会議録には、施策の推進（「しっかり」「どンドン」等）や適切な判断（「はっきり」等）などを表すとみられるオノマトペが高頻度で出現する。これらの語を手がかりに文の目的格を抽出することで、議論の焦点などを分析することができる可能性がある。これは今後の検討課題である。

内田ら[17]によると、Yahoo!の商品レビューに出現する大分類「事物」のオノマトペ上位10語は「ぴったり」「ふわふわ」「さらさら」「きらきら」「さっぱり」「さらっ」「ぎりぎり」「ふんわり」「つい」「ふかふか」である。文書の特性によって、出現するオノマトペの種類が大きく異なることがわかる。

3.2. 意味分類別の出現確率

抽出されたオノマトペの出現確率を中分類ごとに合計した。これを図1に示す（ $\times 10^7$ で表示）。大分類「事物」のオノマトペがもっとも多く抽出された。表2においても、上位10語中8語が「事物」のオノマトペであった。ついで「人間」のオノマトペが多かった。「自然」のオノマトペはもっとも少なく、全体の2.5%であった。中分類を見ると、「事物：程度」に該当するオノマトペ（「しっかり」「もろもろ」「そろそろ」「ゆっくり」「じっくり」等）が最も多く、抽出されたオノマトペの48.7%を占めた。ついで、「事

表2 全会議録における出現確率上位10語

オノマトペ	大分類	出現頻度	出現確率 全国計($\times 10^7$)
しっかり	事物	77,464	119,550
どンドン	事物	20,680	29,763
はっきり	人間	19,382	28,122
だんだん	事物	5,679	8,509
かん	事物	5,171	7,647
びっくり	人間	2,910	4,721
つい	事物	2,728	4,544
もろもろ	事物	2,372	3,688
そろそろ	事物	2,021	2,906
だっ	事物	1,984	2,809

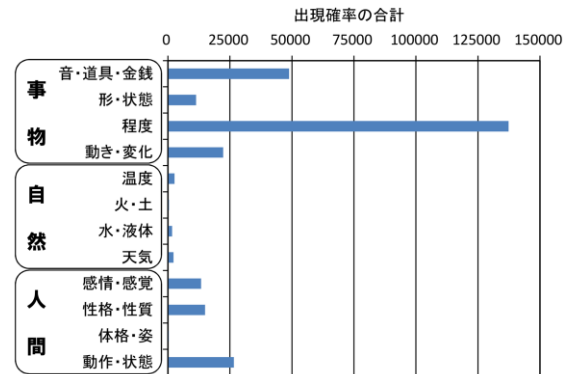


図1 意味分類別の出現確率の合計

物：音・道具・金銭」（17.3%）、「人間：動作・状態」（9.4%）の順であった。

4. 少数の都道府県に出現するオノマトペ

一定の頻度で出現するが、出現地域に偏りのあるオノマトペに着目し、傾向を分析する。本章では、出現確率の合計が 50×10^7 以上かつ、出現都道府県数が20未満の61語²を分析の対象とする。この61語（延べ語数4,164語）がどのように出現しているのかを具体的に確認する。このため、オノマトペを含む文を人手ですべて読み、共著者相互の合意により表3に示す8種類の特徴に分類した。

61語のうち、抽出されたオノマトペがすべて実際にオノマトペであったものは28語、すべてが形態素解析誤りなどでオノマトペがひとつも含まれていなかったものが15語、混在していたものが18語あった。各4.1節以降に具体例を述べる。

4.1. 抽出されたオノマトペ

分析対象のうち、1,321例（異なり数は46語）がオ

² 「えへん」「じゃっ」「しとり」「びっ」「さんさん」「はった」「かったん」「びしゃっ」「きらら」「じゃん」「かんから」「とことこ」「ぼたん」「しとっ」「かちっ」「さっ」「びー」「くすり」「にっ」「ふっ」「ぼくっ」「ごん」「びちっ」「ぼっぼっ」「すいすい」「ぶらぶら」「ぐんぐん」「ちん」「ぞろぞろ」「どかん」「じゃんじゃん」「どろどろ」「とくとく」「ぼこぼこ」「ずきん」「ぶっぶっ」「じりじり」「ころり」「ちよん」「からっ」「きちっ」「つつる」「ふわふわ」「ふわっ」「きっ」「どさくさ」「はっし」「ばりっ」「ぼっきり」「ぶんぶん」「ざわざわ」「ちゅっ」「すい」「ぼーん」「ぼっかり」「ぶくぶく」「どっさり」「ぐいぐい」「もくもく」「しゃんしゃん」「じっ」の61語。

表3 分類の基準と該当するオノマトペの頻度

オノマトペであるか	分類	頻度
オノマトペである	① オノマトペとして発言されている	1,321
	② 名称・呼称に利用されている	334
オノマトペでない	① 同音異義語	28
	② 方言に起因する解析誤り	1,527
	③ 他のオノマトペの一部	58
	④ 名称・固有名詞の一部	720
	⑤ 言い間違い・入力ミス等	27
	⑥ その他の形態素解析誤り	149

ノマトペとして会議録中に出現した。このうち、「びしゃっ」の出現頻度が最も高く、12県に112例出現した。このうち90例が九州地方で使用されている。「びしゃっ」のうち、戸口などを閉めるさまから派生して、「ある基準で完全にやめる、閉める、止めるさま」を表す用例は以下のように全国に見られる。

- 「市長はお金がないというたった一言で**びしゃっ**と切っちゃったんです。」(和歌山県)
- 「それで**びしゃっ**とやめるとなれば」(長野県)
- 「水が入るのを**びしゃっ**と入り口でとめていただいた」(山口県)

一方、九州地方には、「きちんと」「しっかりと」「はっきりと」などの意味合いをもつ例が多く見られる。

- 「そういうものについて**びしゃっ**と明らかにして示す必要がある」(福岡県)
- 「赤字が出ようが**びしゃっ**とボーナスも出ておる。」(長崎県)
- 「資料の提出方法を、もうちょっと**びしゃっ**と定めた方がいい」(熊本県)
- 「こんな問題は、**びしゃっ**とした方がおれはいいんじゃないかねえかと思うんだけど」(天分県)
- 「こちら辺を**びしゃっ**と整備しなければいけないというふうに思っております」(宮崎県)

このほか、69例出現する「ばくっ」³は会議録では、「(棒状のものが) **ばくっ**と折れる」(岩手県) 1例のほかは、すべて以下のような「漠然とした、大ざっぱな」の意味で用いられた。

- 「議会としての**ばくっ**とした考え方なりを聞きたい」(愛知県)
- 「この請願の項目については、割に**ばくっ**と書いてあって、」(兵庫県)
- 「ちょっと**ばくっ**とした聞き方になりますが」(熊本県)

また、「ふわっ」は柔らかいさま(3例)や浮かんでいるさま(5例)よりも「何か**ふわっ**とした決着」(長崎県)「何となく**ふわっ**とした組織」(滋賀県)のように曖昧なさまを表す例が多く(26例)見られる。これも会議録の特徴の一つであろう⁴。

「つるつる」は「表面が滑らかなさま」を表すオノマトペであり、北海道から沖縄県までの14都道府県に37例出現する。このうち、北海道に出現する「つるつる」(11例)はすべて道路の凍結を表すものであった。このように一つのオノマトペがほぼ同一の語

³ オノマトペ辞典によると「勢よく食いつく」「大きく開く」の意味。

⁴ 商品レビュー文などでは、柔らかいことを示す用法が多い。

義で使用されていても、係り先に地域差が見られる場合がある。

オノマトペとして使用されている1,321例の出現確率(都道府県計)は宮崎県、高知県、岡山県、和歌山県、兵庫県の順に高く、出現確率上位20都道府県のうち17府県が近畿地方以西であった。この61語の範囲では、西日本においてオノマトペの出現が多い傾向が確認された。

4.2. 名称・呼称に使用されているオノマトペ

分析対象のうち、334例(異なり語数11語)は名称・呼称に使用されたオノマトペとして会議録に出現した。このうち「きらら」の使用例が最も多く26種類の名称で出現した。例えば「海きらら」(長崎県・17例)、「きらら交流館」(山口県・16例)、「きらら化粧室」(岩手県・7例)、「きらら397」(北海道・5例)などである。次いで、「さんさん」が15種類の名称に使用されていた。例えば、「塚口さんさんタウン」(兵庫県・43例)、「地域活動支援センターさんさん」(沖縄県・18例)、「複合施設さんさんセンター紀の川」(和歌山県・11例)などである。このほか、3種類以上の名称として使用されていたオノマトペは「ぐんぐん」8種類、「ぶんぶん」4種類、「ぶらぶら」、「ふわふわ」3種類であった。

これらは、それぞれの地域で独自に命名したイベント、施設、取り組み等の名称であるため、出現地域は限定される。「明るくまぶしく輝き続けているさま。」を表す「きらら」や、「あざやかで美しいさま。きらきらと輝いて美しいさま。」を表す「さんさん」は、明るい印象を与えることを意図して名称に多く使用されていると考えられる。このほか、「子育ての駅ぐんぐん」(施設名称/新潟県)、「ぶんぶんテレビ」(第三セクターのケーブルテレビ局名/佐賀県)、「ぶらぶら歩きがこちよいまち」(まちづくりの理念/東京都)、「ふわふわドーム」(公園遊具の名称/埼玉県)というようにオノマトペの持つ意味を利用して、直感的に理解しやすい名称を付与している例が多く観察される。名称におけるオノマトペの利用実態については、社会言語学的な観点からさらに調査を進めたい。

4.3. オノマトペではないもの

分析対象のうち、2,509例(60.3%)は、形態素解析においてオノマトペとして抽出したものの、オノマトペではない文脈で出現していた。出現する場面は主に表3に示す6種類に分類される。

このうち「②方言に起因する解析誤り」に該当する例が最も多い。「えへん」は京都府、大阪府、滋賀県、三重県等に出現しており、会議録中の「えへん」(512例)すべてが、咳を表すオノマトペではなく動詞の否定形(「せえへん」「出えへん」等)であった。

このほか、以下の5語が100例以上出現した。

「じゃっ」(例「〜じゃった。」/山口県, 岡山県等)
「しとり」(例「〜をしとります。」/岡山県, 島根県等)
「はった」(例「〜くれはった。」/大阪府, 京都府等)
「かったん」(例「〜できへんかったんかな」/大阪府, 兵庫県等)
「かんから」(例「〜せなあかんから」/兵庫県, 大阪府等)

方言に起因する解析誤りは上述のような文法に関わるものが大半であるが、「きっぽし」(干し芋)「ごじゃっぺ」(出鱈目)なども見られた。

このほか、「③他のオノマトペの一部」では、「くちゅっ」「ちんたら」「びちん」などの例が見られた。「④名称・固有名詞の一部」では、「あびっ子クラブ」「ふっつ」「あさびー」「あそびっ子」「ごんごバス」「ばりっ子」などの例が見られた。「①同音異義語」には、「くすり」, 「ずきん」, 「どかん」, 「ぼたん」などの例が見られた。

このように方言文法, 名称等におけるひらがなの文字列が表記の上でオノマトペと一致するため, ユーザ辞書を用いた形態素解析だけでは, オノマトペ以外も広く抽出してしまうことが分かった。会議録におけるオノマトペの出現傾向についてより詳細に検討を進めるためには, 抽出手法の検討も必要である。今後, 内田ら[18]のオノマトペ抽出手法などを参考に検討を進めていきたい。

5. まとめ

本研究では, 地方議会会議録コーパスに出現するオノマトペについて基礎的な検討を行った。2,466語のオノマトペを辞書登録した上で形態素解析を実行した結果, 約3億語のコーパスから186,416語(異なり語数982語)のオノマトペが抽出された。全体的な傾向としては, 地方議会会議録には事物の程度を表すオノマトペ, 特に「しっかり」「どんどん」「はっきり」など政策の推進や適切な判断などに関わるとみられるオノマトペが高頻度で出現することが分かった。

少数の都道府県に高頻度で出現するオノマトペの分析から, 多義的に用いられ, かつ, 地域によって語義の異なるオノマトペが確認された。「ぴしゃっ」や「ばくっ」は議会の文脈で使用されやすい方言的な語義を持つ特徴的な例である。また, 「つるつる」のようにほぼ同一の語義でも, 係り先の語に地域差のあるオノマトペが観察された。方言語彙には共通語のオノマトペ辞典には掲載されていない方言オノマトペが多数存在する。これらの取り扱いについても今後の検討が必要である。

オノマトペを手がかりとして地方議会会議録の分析を進めることで, 表現技法やその地域差という側面だけでなく, 議題や主題の理解にも役立つことが

示唆された。ただし, これには係り先や文の目的格など構文の情報も必要となるだろう。また, 本稿では限られた範囲での分析に留まっているため, 会議録におけるオノマトペの地域差, 用法, 意味などについてさらに検討を進める必要がある。

謝辞 本研究の一部は科研費 No.22300086 および, No.23700256, No.25370524による。福岡女学院大学二階堂整教授には九州方言について有益な助言をいただいた。

参考文献

- [1] 木村泰知他(2012)「地方議会会議録コーパスの構築とその利用」第26回人工知能学会全国大会, 3B3-NFC-4-3
- [2] 齋藤誠他(2011)「地方議会会議録の収集とコーパスの構築」言語処理学会第17回年次大会論文集, P2-21
- [3] 増田正(2012)「地方議会の会議録に関するテキストマイニング分析:高崎市議会を事例として」地域政策研究 15(1), pp.17-31
- [4] 上田翔, 八木田浩史(2012)「地方議会議事録における環境用語の出現頻度に基づく自治体の環境問題対応の解析」環境情報科学術研究論文集(26), pp.283-288
- [5] 松本直樹(2008)「地方議員の図書館への関心に関する予備的考察:埼玉県市議会の議会会議録分析をもとに」日本図書館情報学会誌 54(1), pp.39-56
- [6] 高丸圭一, 乙武北斗, 渋木英潔, 木村泰知, 森辰則(2013)「形態素N-gramを用いた地方議会会議録コーパスの地域変異検出の試み-文末表現を例に-」言語処理学会第19回年次大会発表論文集, pp.737-740
- [7] 井上史雄(2013)「去った〇日」『ことばの散歩道』明治書院, pp.154-155.
- [8] 齋藤義七郎(1952)「擬声語擬態語の音韻論的考察」第79号, 音声学会会報, pp.11-17
- [9] 宮地裕(1978)「擬音語・擬態語の形態論小考」国語学, 115号, pp.33-39
- [10] 那須昭夫(2007)「オノマトペの語末促音」音声研究, 第11巻, 第1号, pp.47-57
- [11] 方言研究ゼミナール(1992)「身体感覚を表すオノマトペ」, 方言資料叢刊, 第2巻
- [12] 竹田晃子(2012)「円滑な医療コミュニケーションのための方言集-『東北方言オノマトペ用例集』の取り組み-」日本方言研究会研究発表会発表原稿集, 94, pp.55-58
- [13] 平田佐智子, 中村聡史, 小松孝徳, 秋田喜美(2013)「国会会議録コーパスを用いたオノマトペ使用の地域比較」第27回人工知能学会全国大会, 3N4-OS-01c-2
- [14] 高丸圭一(2011)「規模の異なる自治体における地方議会会議録の整文の比較」社会言語科学会第27回研究大会発表論文集, pp.256-259
- [15] 野村稔・鶴沼信二(1996)『地方議会実務講座 第3巻』ぎょうせい
- [16] 小野正弘(編)(2007)『日本語オノマトペ辞典』小学館
- [17] 内田ゆず, 荒木健治, 米山淳(2013)「商品レビュー文に含まれるオノマトペに基づく商品カテゴリの分類」第29回ファジィシステムシンポジウム(FSS2013)講演論文集, pp.762-777, 2013.
- [18] 内田ゆず, 荒木健治, 米山淳(2012)「ブログ記事からのオノマトペ用例文の自動抽出手法」Journal of Japan Society for Fuzzy Theory and Intelligent Informatics 24(3), pp.811-820