

国会議事録コーパスを用いたオノマトペの通時的分析

A diachronic analysis of onomatopoeia in the Diet record

中村聡史^{*1}
Satoshi Nakamura平田佐智子^{*2}
Sachiko Hirata秋田喜美^{*3}
Kimi Akita^{*1} 京都大学
Kyoto University^{*2} 日本学術振興会(東京大学)
Japanese Society for Promotion of Science^{*3} 大阪大学
Osaka University

There are few studies that examine the semantic change of onomatopoeic words in spoken discourse. This paper uses the Japanese Diet record as a massive longitudinal corpus of spoken Japanese, discussing the semantic change of some representative onomatopoeic words. We also comment on the importance of mutual contribution between informatics and other fields, such as psychology and linguistics.

1. はじめに

時間的推移によるオノマトペ(擬音語・擬態語)の意味変化の可能性に対して、オノマトペの通時的研究は少ない。本研究では大規模な話し言葉データである国会議事録データベースをもとに、発言中のオノマトペの頻度及び共起する語の分析により、オノマトペの意味がどのように推移したのかを検討する。また、これに基づいて、情報学と心理学・言語学という分野を大きく横断した研究において、各分野の研究者がどのように相互に貢献できるかといった点についての知見を述べる。

2. オノマトペの意味変化

2.1 オノマトペ

オノマトペは日本語語彙の重要な一部を占め、漫画や商品の宣伝などに幅広く使用される。また、様々な側面で一般語彙とはやや異なる性質を持つことも知られている。その性質の一つとして、「意味変化を引き起こしやすい」点が挙げられる。

2.2 意味変化に関する研究

オノマトペは、音や音以外を音声で表現する、すなわち意味と音声の間に類像的な関係が存在することが指摘されているが、

オノマトペの意味変化について扱った研究では、近代から現代の辞書の用例に含まれるオノマトペについて詳細な通時的分析を行っている[中里 2002, 2004a, 2004b]。その結果、「わくわく」や「まじまじ」、あるいは泣く行為／涙を意味するオノマトペには、時代によって意味の分化や他の語への意味移行による多義の縮小が見られることがわかっている。

また、通時的により短い範囲ではあるが、Twitter と呼ばれるマイクロブログサービスを用いた調査を行った研究も存在する[平田 2011]。これはユーザに直接問いかけることによりオノマトペの「現在」の意味を取り出し、既存の辞書と比較することでオノマトペの通時的意味変化を捉えることを試みた。調査の結果、「はんなり」などのオノマトペが既存辞書にはない新しい意味で主に用いられていることがわかった。

しかしながら、最もオノマトペが使用されやすい状況は日常会話である。よって、オノマトペの変化を捉えるためには、話し言葉に含まれるオノマトペの分析が必要である。Twitter におい

て使用される言葉は話し言葉に近い形ではあるが、基本的に文字によるやりとりであり、また、先行研究[中里 2002, 2004a, 2004b]が検討した例も元は文学作品であったため、話し言葉による分析は未だなされていないと言ってよい。そこで、本研究では、数十年単位で話し言葉のデータを蓄積している国会議事録データベースを元に、話し言葉中のオノマトペがどのような意味変化を起こすのか、検討することとした。

2.3 本研究の目的

本研究は、いままでオノマトペの通時的研究がほとんどなされていなかった話し言葉コーパスの中から、国会議事録データベースに注目する。そして、データベースに含まれるオノマトペと、オノマトペと共起する語を分析することによって、話し言葉の中でオノマトペの意味変化が起こっているかどうか、起こっているとすればどのような変化なのか、を検討する。

3. 方法

3.1 国会議事録データベース

本研究では、話し言葉のデータベースの中でも数十年単位でのデータ蓄積がなされている国会議事録データベースを分析対象として選択した。

国会議事録データベースは「国会会議録検索システム (<http://kokkai.ndl.go.jp/>)」(図 1 参照)として Web 上で一般公開されており、1947 年の第 1 回国会から最新の議事録まで 60 年にわたる議事録を検索することができる。また、選択閲覧機能を利用することで、すべての議事録について探索および閲覧を行うことが可能である。その範囲は衆議院・参議院本会議から両院協議会や合同審査会にまで及んでいる。

本研究ではこのすべての議事録を対象として、オノマトペがどのような文脈で使用されるのか、またそれは時代によって変化するかを追った。



図 1 国会会議録検索システムの Web ページ

3.2 データベースの構築および分析

国会会議録のデータは、基本的に会議の開催された日(和暦)、会議の参加者、発話した人の名前、その発話自体からな

連絡先: 中村聡史

〒606-8501 京都市左京区吉田本町

京都大学大学院情報学研究科 社会情報学専攻

TEL:075-753-5940 Mail: nakamura@dl.kuis.kyoto-u.ac.jp

る。また、発話者や発話内容は上から下に時系列に沿って並んでいる。そこで、まず国会会議録の各会議録から日時情報を抽出し西暦に変換する。また、発話者および発話者の役職、発話内容を抽出し、発話内容を1文ずつに分解することで、発話ID(発話を一意に識別するもの)、発話日、発話者名、発話者の役職、発話文からなるデータベースを構築する。ここで、発言内容を1文ずつに分解するため、句点または国会会議録作成者により改行されている部分、話者が変わるまでの部分を1文として処理する。データベースに格納された発話文数は48,512,596文で、発話文字数は3,437,337,516文字となった。

構築したデータベースに対して、後述する分析のためのオノマトペリストから検索語を1つ取得し、検索語を含む発話文をすべて取得する。また、その発話文の発話年をそれぞれ集約することにより、年毎の検索語の発話量を計算するというものである。この処理を、すべての検索語について行った。

さらに、検索語が検出されたすべての文について、検索語がどのように使われているかを調べる。ここでは、その検索語を含む文をCaboCha[工藤 2002]を用いて係り受け解析を行い、検索語へと係っている語および語群、検索語から係っている語および語群を取得した。取得した語および語群を年ごとに集約し、そのランキングを作ることによって、年毎の変化を観察できるようにした。

3.3 対象としたオノマトペ

通時的分析をする上で、対象とするオノマトペは各年の発言にある程度出現している必要がある。そこで、先行研究[Kakehi 1996; 小野 2007]をもとに作成したオノマトペリストを複合し、作られたリストから各年20以上の出現数(先述の分析により取得)があるオノマトペを抜粋した。その中から検索時にオノマトペではないのにオノマトペであると判定されノイズとなってしまうやすいオノマトペ(例:「ばしばし」→後続の語によっては「しばしば」が誤って判定されやすい)をあらかじめ避けたものを検査対象とした。本研究ではさらに抜粋し、「どしどし」「すっかり」「さっぱり」の3つのオノマトペについての通時的分析の結果を報告する。なお、1947年から1954年までは旧仮名遣いによって議事録が作成されているため、「すっかり」「さっぱり」に関しては「すつかり」「さつぱり」を検索語として使用した。

4. 結果

4.1 オノマトペの出現傾向及び係る語の通時的变化

(1) 「どしどし」について

図2は1947年から2011年までの発言に含まれる「どしどし」の数を示している。

既存のオノマトペ辞書[小野 2007]によると、「どしどし」には二つの意味があり、

- 1) 何度も力強く踏みつけたり、たたく音。また、そのさま。
- 2) 遠慮無く物事を行うさま。同様の物事が次から次へと多く続くさま。

が挙げられる。各年代の発言において「どしどし」に係る動詞を検討した結果、1970年までの発言においては「(どしどし)やる」「(どしどし)する」など、代動詞が頻出する傾向が認められた。しかし、これ以降の年代では「(どしどし)推進する」「(どしどし)取り入れる」「(どしどし)公開する」など、具体性の高い動詞に係る傾向が見られた。具体性の高い動詞に対して「どしどし」が使用されるようになってきていることから、昔と比べて「どしどし」の意味合いが薄れていることがうかがえる。

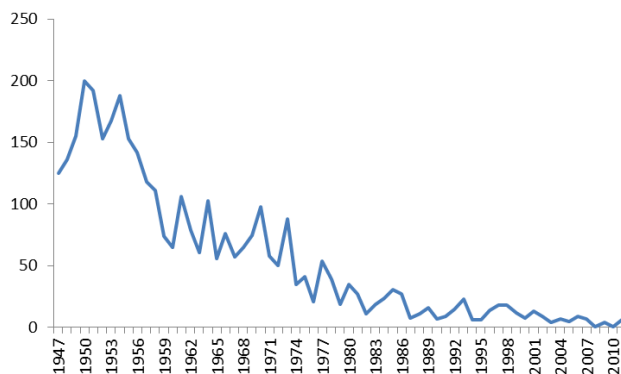


図2 国会議事録中の「どしどし」を含む発言数

(2) 「すっかり」について

図3は1947年から2011年までの発言に含まれる「すっかり」の数を示している。既存のオノマトペ辞書[小野 2007]によると、「すっかり」には4つの意味があり、

- 1) 残るところなくすべてにわたるさま。ことごとく。
- 2) (古語)思い切つてするさま。滞りのないさま。すがすがしいさま。
- 3) (古語)細身で格好のいいさま。すらし。
- 4) (古語)矢が狙った対象を射そこなうさま。

が挙げられる。各年代の発言において「すっかり」に係る動詞を分析した結果、1950年代では「(すっかり)調べる」「(すっかり)実行する」「(すっかり)増強する」など、中立あるいはポジティブな意味合いを持つ動詞に係っている。対して、それ以後は「(すっかり)変わってしまう」「(すっかり)忘れ去られる」「(すっかり)冷え切る」など、ネガティブな意味合いを持つ動詞に係る傾向が見受けられる。これは、元々ポジティブ・ネガティブどちらのタイプの動詞にも用いられていた「すっかり」の意味が縮小し、ネガティブな意味に限定されつつあることを示していると言える。

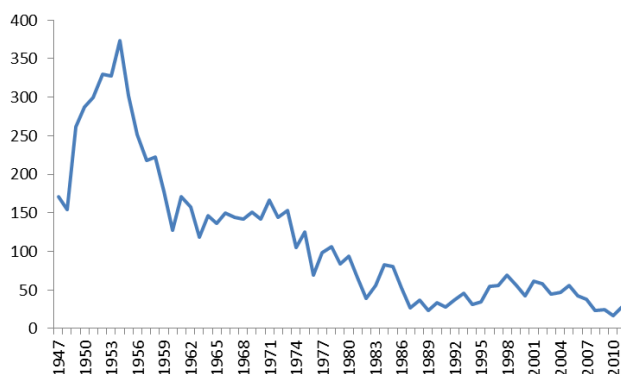


図3 国会議事録中の「すっかり/すつかり」を含む発言数

(3) 「さっぱり」について:意味変化が起こりにくい例として

図4は1947年から2011年までの発言に含まれる「さっぱり」の数を示している。

既存のオノマトペ辞書[小野 2007]によると、「さっぱり」には3つの意味があり、

- 1) 気持ちや気分が晴れて、さわやかになるさま。身なり・衣服、人や物の性質・態度、味などがくどくなく、清潔なさま。
- 2) 後に何も残らないさま。こだわりや後くされなどが無いさま。

3) まったく振るわないさま。すべてにあてはまるさま。まるで。少しも。

が挙げられる。各年代の発言において「さっぱり」に係る動詞を検討した結果、「(さっぱり)わからない」という特定の動詞に係る例が最も多く、またこの傾向は 1955 年から現在まで一貫していた。このことから、国会議事録内においては、「さっぱりわからない」という表現が定型的であり、他の動詞には係らないことが推測される。

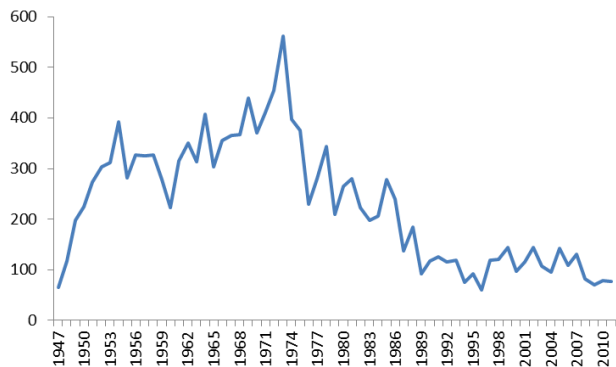


図 4 国会議事録中の「さっぱりさっぱり」を含む発言数

5. 考察

5.1 本研究のまとめ・問題点

本研究では、国会議事録データベースを話し言葉コーパスと見なすことで、1947 年から蓄積されている話し言葉データを通時的に分析し、オノマトペの意味変化を捉えることを目的とした。その結果、ごく一部ではあるがオノマトペの意味変化が起こっている事実を捉えることができたと言える。しかしながら、データベースの特性上いくつかの問題点も挙げられる。

一つは、国会議事録データベースは「国会における発言内容」というある種特殊な状況下の発言であることを考慮しなければならない点である。国会で発言される内容は、すなわちその時代における日本の状況をそのまま反映している。すなわち、国が問題を抱えていたり、情勢があまり良くない時代にはネガティブな発言内容が増え、その影響がデータベースの内容にもおよぶのである。よって、今回見られた意味変化が、オノマトペ自体の純粋な意味変化ではなく、情勢の変化によるものである可能性が考えられるのである。この性質は、データ採取期間に幅のあるデータベースにはつきものではあるが、他のコーパスと比較することによってこれらの影響を軽減する必要があると考える。

もう一つは、公の発言であるために生じる性質である。国会での発言は公のものであるため、くだけた表現が得られにくい。オノマトペは公的文書や新聞などにはあまり現れないため、国会での発言にも含まれにくく、含まれるとしても感性的表現(例えば、「さっぱり」における 1)の意味など)は使用されにくいことが推測される。この点に関しても、上に挙げた問題と同様、他のフィールドで採取されたデータとあわせて変化を見ていく必要があると考える。

5.2 国会議事録データベースの応用可能性

前節では国会議事録の性質による問題を挙げたが、このコーパスの持つ性質を別の角度から見ることで、興味深い可能性も見いだすことができる。例えば、その時代の情勢をそのまま反映

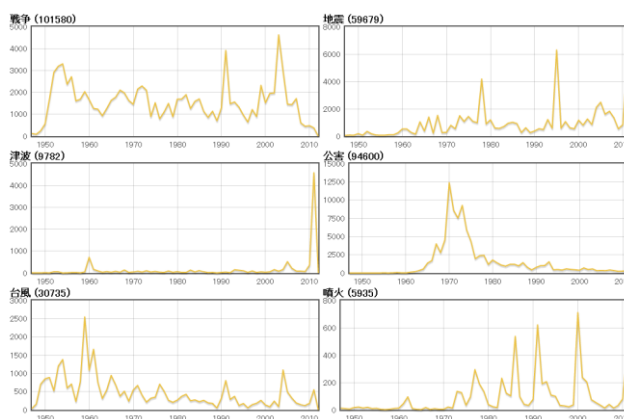


図 5 通時変化観察のための比較インタフェース

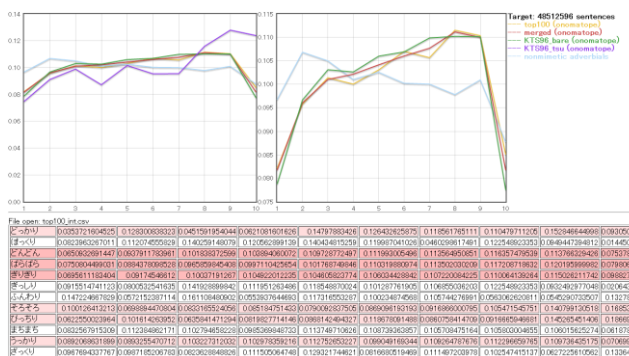


図 6 リアルタイム視覚化システム

しているという特性から、特定の言葉から各時代の傾向を捉えることが可能となる。予備的な分析ではあるが、他のオノマトペは現在に近くなるにつれて出現数が減っていく傾向があるのに対し、「しっかりと」というオノマトペは現在に近づいても出現数が一時的に多くなる現象が観察された。これは、「しっかりと」を含む発言が多く発言された、すなわち日本が当該時代に「しっかりと」言わなければならない状況下にあったのではないかと推察される。

また、オノマトペは公的発言には含まれにくいという点と、初期の発言中にオノマトペが多く含まれており、だんだん減っていくという点をあわせると、昔の国会はくだけた表現が多い場だったのではないかと考えられる。このように、言葉の性質と、国会議事録の性質を合わせることで言語研究に対しても、歴史的・社会的研究に対しても新しい切り口を提供できるのではないかと考える。

5.3 分野横断的研究のすすめ

本研究及びオノマトペに対する意識の地域差比較[平田 2012]、オノマトペのインタラクション性に関する量的考察[秋田 2012]という 3 つのオノマトペを中心とした研究は、2011 年の人工知能学会全国大会のオーガナイズドセッション「オノマトペの利活用: ユーザの曖昧な意図をどう扱うか」を切っ掛けとして生まれたものである。今回は、主に情報学、心理学、言語学といった異なる分野で連携し、オノマトペを中心としてテーマを選定するとともに、相互協力することで研究を推進してきた。

今回、情報学側から連携において工夫した点は、完了までかなりの時間を要する分析を、多少不正確でもよいので目に見える形でリアルタイムに視覚化するという点である(図 5, 図 6 はその一例)。データの偏りなどが考えられるため、すべてを処

理するまでは正確な値であるとは言えないが、ある程度目に見える形で提示することによって心理学、言語学の研究者に対して気づきを与え、そこから新たなアイデア、分析案へと繋げることができた。また、簡易的な Web インタフェースを用意し、そのインタフェースを利用して少し深い分析を行えるようにすること、さらに、詳細な結果は CSV などで出力し、視覚化データなどを利用して分析のためのとっかかりを見つけてもらい、詳細な分析をしてもらったということも特筆に値する。空間、時間を共有することが難しかったため、ChatWork などのチャットベースのコミュニケーションシステムを利用することによって、随時コミュニケーションをとったということもプロジェクトの円滑な進行にとってプラスに働いた。それでもなお、大きな進展を迎えるタイミングは、節目ごとの直接ミーティングの直後であり、未来の共同プロジェクトにおいても伝統的な方法は軽視できないことを再確認した。

情報学から見て今回のような分野横断的研究の面白い点は、情報学が役に立つ領域というのはまだまだあり、そこに貢献することができるということである。言語学や心理学などの領域における未解決の問題や証明されていない仮説を、情報系の技術によって解決または明らかにできる可能性があり、そこから色々な方向への研究の発展性を見出すことができた。言語学や心理学において、問題は明確化されつつも、大規模処理などを専門としていないために解決できずにいた根本的問題が、その分析の難しさを考えず情報学の研究者に対してリクエストを出すことで解決できることが判明したのは、双方にとって嬉しい驚きであった。以上のように、3つの分野を連携した横断型研究というのはとても刺激的かつ発展的であり、また何より面白いものである。今後も規模を拡張しつつ、様々なテーマに関してこうした共同プロジェクトを推進していきたいと考えている。

謝辞

本研究は、科学研究費補助金：学術研究助成基金助成金（若手研究(A)）（課題番号：23680006）（中村）、科学研究費補助金：特別研究員奨励費（課題番号：23・4301）（平田）、科学研究費補助金：学術研究助成基金助成金（若手研究(B)）（課題番号：24720179）（秋田）の補助を受けて行われた。

参考文献

- [中里 2002] 中里 理子: オノマトペの多義性と意味変化—近世・近代の「まじまじ」を例に—。上越教育大学研究紀要, Vol.22, pp.11-25 (2002)
- [中里 2004a] 中里 理子: オノマトペの意味縮小—「わくわく」を例に—。上越教育大学研究紀要, Vol.23, pp.15-27 (2004a)
- [中里 2004b] 中里 理子: 「泣く」「涙」を描写するオノマトペの変遷—中古から近代にかけて—。上越教育大学研究紀要, Vol.24, pp.314-303 (2004b)
- [平田 2011] 平田 佐智子、澤井 大樹、藤井 弘樹、喜多 伸一: Twitter を用いたオノマトペ記述データの収集システム。第 25 回日本人工知能学会大会論文集 (2011)
- [平田 2012] 平田 佐智子、秋田 喜美、小松 孝徳、中村 聡史、藤井 弘樹、澤井 大樹: オノマトペに対する意識の地域比較。第 26 回日本人工知能学会大会論文集 (2012)
- [Kakehi 1996] Hisao Kakehi, Ikuhiro Tamori, and Lawrence Schourup: Dictionary of Iconic Expressions in Japanese. Berlin/New York: Mouton de Gruyter. (1996)
- [小野 2007] 小野正弘(編): 擬音語・擬態語 4500 日本語オノマトペ辞典。小学館:東京。(2007)

[工藤 2002] 工藤 拓、松本 裕治: チャンキングの段階適用による日本語係り受け解析。情報処理学会論文誌, Vol. 43, No.6, pp. 1834-1842. (2002)

[秋田 2012] 秋田 喜美、中村 聡史、小松 孝徳、平田 佐智子: オノマトペのインタラクション性に関する量的考察。第 26 回日本人工知能学会大会論文集(2012)