

令和 5 年 6 月 15 日現在

機関番号：10102

研究種目：基盤研究(C)（一般）

研究期間：2019～2022

課題番号：19K00591

研究課題名（和文）多義語に対するプロトタイプ義の量的分析 - クラウドソーシングによる大規模調査 -

研究課題名（英文）Quantitative analysis of prototypical meaning for polysemous words: through a large-scale crowdsourced survey

研究代表者

西内 沙恵 (Nishiuchi, Sae)

北海道教育大学・教育学部・講師

研究者番号：00804486

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）：本研究は心理実験的手法を通して、現代日本語形容詞の多義語を対象にプロトタイプ義を認定し、多義モデルを構築することを目的としている。本研究の新規性は、双方向評定によるプロトタイプ義認定の手法の確立にある。プロトタイプ義と拡張義の語義間における類似性が認められやすい方向に着目し、語義間の類似度を評定していただく心理実験の結果から、プロトタイプ義及び拡張関係の解明を試みた。調査手順として、先行研究の代表義リストと意味情報付与コーパスを統合的に活用しながら、クラウドソーシングで協力者を募集し双方向評定を調べた。調査結果は生データ・集計データ・デンドログラムの形式でウェブサイトに公開している。

研究成果の学術的意義や社会的意義

本研究を通して、クラウドソーシングを用いた大規模調査による均衡コーパスの用例間に対する類似度評定から多義語分析を行った。従来の研究では考案されていなかった、研究者の内省・直観によらない、複数の語に対する統一的な、再現性の高い手法の確立を試みた。クラウドソーシングを通して募集した協力者には、例文の類似度を双方向に評定していただき、評定された類似度の高低からプロトタイプ義が認定できることを実証した。また双方向の評定が、歴史的意味拡張の起点と現代の直観的プロトタイプの相違を調べる手法として活用できることを提案した。調査の結果はWebサイトで公開し、追試における比較検討に役立てられる。

研究成果の概要（英文）：The aim of this study is to establish a prototypical meaning for polysemy of contemporary Japanese adjectives through psycho-experimental methods, and to construct a polysemy model. The novelty of this study lies in the establishment of a prototypical meaning recognition method based on interactive grading. Focusing on the direction in which similarity between prototypical and extensional meanings is tend to be recognised, we attempted to clarify prototypical meanings and extensional relations based on the results of a psychological experiment in which the degree of similarity between the meanings was rated. As a research procedure, the bidirectional grading was investigated by crowdsourcing, integrating the representative sense lists and semantic information corpus of previous studies. The results of the survey are published on the website in raw data, aggregated data and dendrograms.

研究分野：認知意味論

キーワード：多義語 プロトタイプ 拡張 類似性 心理実験 言語学的テスト コーパス クラウドソーシング

1. 研究開始当初の背景

多義語とは、複数の関連する意味を持つ語であり、多義語の複数の語義はプロトタイプ義と拡張義に分類される。多義語について、プロトタイプ義と拡張義、またその拡張プロセスを明らかにしようとする研究は数多い。ただし、分析手法がさまざまに提案されてきたが、多数の語を統一的に調査・検証できる手法は編まれていなかった。具体的には、言語学的テストを用いた分析において研究者の内省・直観により再現性が低いという問題点があった。また心理実験的手法を用いた分析において、手続き上、多数の語を扱えないという問題もあった。本研究では、均衡コーパスの用例とクラウドソーシングを用いた大規模調査を行い、多義語分析を行う。これにより、多数の語を統一的な手法で分析する心理実験的手法を確立する。共時的な概念的な中心性は通時的な意味拡張の起点の一時点に過ぎず、両者は不可分な関係であるため、再現可能な調査手法としても確立する必要がある。

2. 研究の目的

本研究の目的は大きく次の2点である。

- (1) 双方向評定による例文間類似度がプロトタイプ義認定に役立てられることを検証する。プロトタイプ義と拡張義には方向性があることが論証されている (Lakoff 1987, Tuggy 1993)。プロトタイプ義から拡張義に類似度を認めるとき、反対方向よりも認知的負荷が軽いことが想定される。本研究では多義語間の類似度が認められやすい方向に着目し、双方向評定を活用し、派生関係を調べる (図1)。
- (2) 言語学的テストとの比較検討を行い、本研究が提案する手法の有効性を検証する。従来考案されてきた言語学的テストと本研究が提案する手法を比較し、それぞれが有効に働く範囲を調べる。

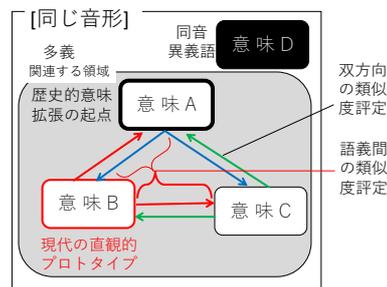


図1 語義の双方向評定

3. 研究の方法

目的(1)(2)に対応する調査として(1)(2)の方法を取った。

- (1) 例文間に対する双方向評定類似度によるプロトタイプ義認定
 - ①の手順でデータセットを作成し、②の方法で実施した。調査協力者は、20歳以上のYahoo! Japan ID 所有者である。
 - a. 『現代日本語書き言葉均衡コーパス』および『国語研日本語ウェブコーパス』から用例を抽出する。
 - b. 代表義情報付与済み多義語リスト (山崎・柏野 2017) と意味情報付与済み均衡コーパスのデータ (加藤ほか 2019) を援用しながら、日本国語大辞典・分類語彙表から語源と語義を付与する。
 - c. 述定用法・装定用法・副詞的用法の用法別に分類する。
 - d. 語義と用法別の用例分布を整理し、データセットを作成する。
 - ② a. クラウドソーシングで数千名規模の調査協力者を募集し、調査を実施する。
 - b. データセットの用例について、ある1文 (指標文) から見てほかの文 (判定文) が似ているかどうかを6段階でチェックしていただく (図2)。
 - c. 回答を語義・用法別に集計する。

クラウドソーシングとは、不特定多数の作業者への簡単な仕事の外注である。Yahoo!クラウドソーシングやランサーズなどのクラウドソーシングサイトで、非常に小さい単位のタスクを数千人規模で依頼できるサービスが提供されている。大量の語彙・用例に対する大人数のアンケート調査を安価に実施できる (浅原 2019)。なお、タスクの性質上、指示文や用例を読まずに回答できる。このような不適切な回答を排除するために、調査協力の同意確認を兼ね、「同意する」・「同意しない」をランダムに配置したチェック設問を設けた。「同意しない」を選択した回答者は「落選」となり、回答が回収されない。

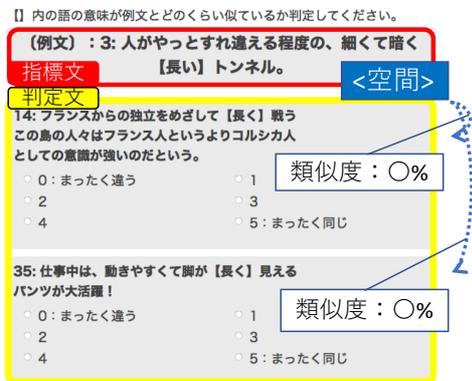


図2 調査画面の一部

- (2) 言語学的テストとの比較検証

先行研究で考案されてきた、多義性を認める言語学的方法を、語彙テスト・文法テスト・論理テストに区分して一覧し、有効範囲を検討する。それぞれのテストの仕組みと特徴を分析し、本研究が提案する手法と比較する。

(3) 多義語と文法の関係

本研究を通して多義語の実態を明らかにするために、多義語と統語的特徴の関係を分析する必要性が生じた。従来、多義語が語義ごとに異なる統語的特徴を持つのは辞書意味論の立場から語義が意味的に充足するために要求される項構造が要因だと考えられていた（仁田 1977, Pustejovsky 1995）。本研究の量的調査を通して、統語的特徴が語義の拡張プロセスに起因している可能性が認められた。多義語はプロトタイプ義からの拡張によって発生し、その拡張プロセスにおいて拡張元との識別のために統語的特徴を要請している場合があることを定性的な分析から検討した。多義語の語義と統語的特徴の関係は語義という要素のみに還元できない、多義語が持つ拡張プロセスによる全体論的なあり方によって説明可能になる。本研究では形容詞が取る二格名詞句、二重主語構文、語幹構文という部分的な現象ではあるが、多義語の語義と統語的特徴の関係に対して拡張プロセスを動機付けとした説明を試みた。多義語の語義と統語的特徴は拡張元、また拡張先の意味に応じて文法的な制約が課される点で文法的な知識に支えられている。一方、統語的特徴もまた多義語の拡張プロセスによって説明可能な、意味の実現に資する制約が現象している。今後様々な統語的特徴を射程に事例を幅広く扱うことで、多義語の実態が明確にできると考えられる。

(4) 多義語の習得プロセス

母語話者の言語直感を調査する心理実験を通して多義語の心的実在性を実証する必要性が生じた。人間の認知能力に基づいて、多義語のメカニズムを解明しようとする研究には蓄積があり、本研究もこれに位置付けられる。一方で多義語の分析はアーティファクトであるという批判が根強い。語に多義性を認めない立場に、大別して、語に意味を認めず人が具体的な用法をその都度解釈していると考える用法説と、日常的な言語運用で話者が多義性を意識することがないため、多義性が研究者による創作に過ぎないと考えるアーティファクト説がある（図 3）。多義語の心的実在性に関する議論が未だ続いている中で、本研究の心理実験的手法が役立てられると考えた。心理実験では私たち一人一人が習得した言語の経験をもとにカテゴリー化のプロセスが追究できるためである。本研究の成果から、母語話者の多義語に対するカテゴリー化のプロセスを示し、多義性の心的実在性の実証を試みた。今後は大人の言語直観だけでなく、子どもが現代日本語の多義語を習得するプロセスを調査することで、多義語の言語知識を形成する基盤となる、比較する認知能力によるカテゴリー化が検証され、多義語の心的実在性が検討できる。

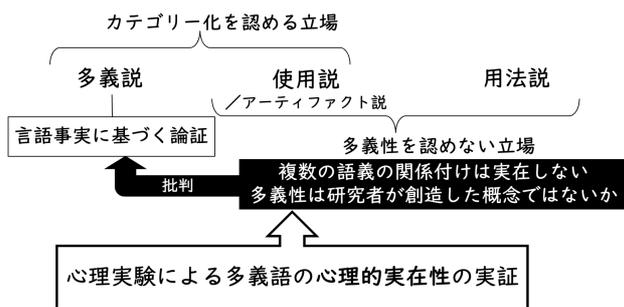


図 3 多義語の心的実在実証の必要性

参考文献

- 浅原正幸 (2019) 「クラウドソーシング結果の可視化手法と統計処理」『日本言語学会第 158 回大会予稿集』 379-384.
- 加藤祥・浅原正幸・山崎誠 (2019) 「分類語彙表番号を付与した『現代日本語書き言葉均衡コーパス』の書籍・新聞・雑誌データ」『日本語の研究』 15(2): 134-141.
- 国広哲弥 (1982) 『意味論の方法』大修館書店.
- 仁田 義雄 (1977) 「多義性を有する用言についての二三の考察 : Lexico-Syntax の姿勢において」『日本語・日本文化』 5: 79-94.
- 早瀬尚子 (2018) 「言語表現の意味とその指示対象」早瀬尚子 (編) 『言語の認知とコミュニケーションー意味論・語用論, 認知言語学, 社会言語学ー』 29-44. 開拓社.
- 榎山洋介 (1995) 「多義語のプロトタイプの意味の認定の方法と実際ー意味転用の一方向性: 空間から時間へー」『東京大学言語学論集』 14: 621-639.
- 山崎誠・柏野和佳子 (2017) 「『分類語彙表』の多義語に対する代表義情報のアノテーション」『言語処理学会第 23 回年次大会発表論文集』 302-305.
- 李在鎬 (2018) 「動詞「流れる」の語形と意味の問題をめぐって」『計量国語学』 26(2), 64-74.
- Lakoff, G. (1987) *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press.
- Pustejovsky, J. (1995). *The generative lexicon*. MIT Press.
- Tuggy, D. (1993) Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4(3): 273-290.

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 5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 西内 沙恵	4. 巻 22
2. 論文標題 多義語のカテゴリー拡張における項構造との関係 - 第一種二重主語文を取りうる形容詞を事例に -	5. 発行年 2022年
3. 雑誌名 日本認知言語学会論文集	6. 最初と最後の頁 125-137
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス 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 -

1. 著者名 西内 沙恵	4. 巻 21
2. 論文標題 多義的形容詞に対する類似度評定を用いた多義認定の有効性 - 現代日本語次元形容詞の結果から -	5. 発行年 2021年
3. 雑誌名 日本認知言語学会論文集	6. 最初と最後の頁 358-364
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス 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 -

1. 著者名 西内 沙恵	4. 巻 5
2. 論文標題 類似度評定を用いた多義間の相互関係の分析-「鋭い」を事例に-	5. 発行年 2020年
3. 雑誌名 言語資源活用ワークショップ発表論文集	6. 最初と最後の頁 33-42
掲載論文のDOI（デジタルオブジェクト識別子） 10.15084/00003144	査読の有無 無
オープンアクセス 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 -

1. 著者名 西内 沙恵	4. 巻 26
2. 論文標題 談話における形容詞の対人関係調整機能：「いい」の振る舞いから	5. 発行年 2020年
3. 雑誌名 筑波応用言語学研究	6. 最初と最後の頁 30-43
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス 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 -

1. 著者名 西内 沙恵, 加藤 祥, 浅原 正幸	4. 巻 20
2. 論文標題 語義間類似度の双方向評定に基づくプロトタイプの意味の解明 クラウドソーシング を用いた量的調査による多義的形容詞分析	5. 発行年 2020年
3. 雑誌名 日本認知言語学会論文集	6. 最初と最後の頁 256-268
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス 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 -

1. 著者名 西内 沙恵	4. 巻 24
2. 論文標題 多義性を認める言語学的テストの有効性 現代日本語の名詞・形容詞・動詞を対象に	5. 発行年 2023年
3. 雑誌名 国立国語研究所論集 = NINJAL research papers	6. 最初と最後の頁 133-152
掲載論文のDOI (デジタルオブジェクト識別子) 10.15084/00003691	査読の有無 有
オープンアクセス 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 -

[学会発表] 計9件 (うち招待講演 0件 / うち国際学会 1件)

1. 発表者名 Sae Nishiuchi, Sachi Kato, Masayuki Asahara
2. 発表標題 Polysemy and De-contextualisation: A Quantitative Investigation Based on Similarity for the Identification of Prototypical Meaning
3. 学会等名 UK COGNITIVE LINGUISTICS CONFERENCE 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 加藤 祥, 西内 沙恵, 浅原 正幸
2. 発表標題 多義語用例の類似度による語義の分類: 「遠い」と「近い」を例に
3. 学会等名 日本認知言語学会第20回全国大会
4. 発表年 2019年

1. 発表者名 浅原 正幸
2. 発表標題 クラウドソーシング結果の可視化手法と統計処理
3. 学会等名 日本言語学会第158回大会
4. 発表年 2019年

1. 発表者名 加藤 祥
2. 発表標題 クラウドソーシングによる語義調査
3. 学会等名 日本言語学会第158回大会
4. 発表年 2019年

1. 発表者名 西内 沙恵
2. 発表標題 クラウドソーシングによる述定・装定の用法分析
3. 学会等名 日本言語学会第158回大会
4. 発表年 2019年

1. 発表者名 西内 沙恵, 加藤 祥, 浅原 正幸
2. 発表標題 ヒトによる多義的形容詞に対する類似性の評価データベース構築-「長い」と「短い」の事例から-
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 浅原 正幸, 西内 沙恵, 加藤 祥
2. 発表標題 NWJC-BERT: 多義語に対するヒトと文脈化単語埋め込みの類似性判断の対照分析
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 西内 沙恵
2. 発表標題 現代日本語の多義語に対する心理実験によるカテゴリー化の解明 - 習得モデルの構築に向けて -
3. 学会等名 令和4年度語学文学会学術研究会
4. 発表年 2022年

1. 発表者名 西内 沙恵
2. 発表標題 現代日本語形容詞の多義性と構文の関係 - 感動文を手がかりに -
3. 学会等名 関西言語学会第47回大会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>多義結び https://sites.google.com/view/tagimusubi/%E5%A4%9A%E7%BE%A9%E7%B5%90%E3%81%B3</p>
--

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	加藤 祥 (Kato Sachi) (40623004)	目白大学・外国語学部・専任講師 (32414)	
研究分担者	浅原 正幸 (Asahara Masayuki) (80379528)	大学共同利用機関法人人間文化研究機構国立国語研究所・言語資源開発センター・教授 (62618)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関