

国立国語研究所学術情報リポジトリ

多義語のプロトタイプの意味認定における脱文脈化の活用：クラウドソーシングを用いた量的調査から

メタデータ	言語: Japanese 出版者: 公開日: 2020-02-06 キーワード (Ja): キーワード (En): Word List by Semantic Principles, NINJAL Web Japanese Corpus (NWJC), Balanced Corpus of Contemporary Written Japanese (BCCWJ), BCCWJ-WLSP 作成者: 西内, 沙恵 メールアドレス: 所属:
URL	https://doi.org/10.15084/00002550

多義語のプロトタイプの意味認定における脱文脈化の活用 ークラウドソーシングを用いた量的調査からー

西内 沙恵 (筑波大学 [院])[†]

Utilization of the De-context in the Identification of the Prototypical Meanings of Polysemic Words : A Quantitative Investigation Using the Crowdsourcing

Sae Nishiuchi (University of Tsukuba)

要旨

多義語のプロトタイプの意味の認定には、意味的出現の高頻度・想起の容易さ・用法上の制約の少なさ・歴史的出現の順序・習得段階など、様々な手法が提起されている。本研究では、意味の移り変わりを前提とした、再調査可能なデザインの量的調査による認定手法を提案する。調査では、多義的形容詞の実例と脱文脈化した語の類似度を調べ、その結果に基づいてプロトタイプの意味の認定を行う。発表では、この手法の妥当性を多角的に検討する。

1. この研究の背景と目的

プロトタイプの意味とは、意味的に関連する2つ以上の語義が同一の音形に結びつく多義語(国広 1982: 97)の多義のうち、より基本的な語義である。田中(1990: 90)は、プロトタイプの意味を軸にそうでない語義が広がりを見せると述べている。多義語のプロトタイプの意味の認定は、語義の派生関係や通時の変化の解明といった研究の根幹をなす基礎的課題である。また、共時的な概念中心性は、通時的な意味拡張の一時点である(木下 2019)ことから、変遷の可能性を考慮することも重要である。本研究では、形容詞を題材に、意味の移り変わりを前提とした再調査可能なデザインの量的調査による認定手法を提案する。

山崎(2011)は、語形と意味との対応について『現代日本語書き言葉均衡コーパス』(以下、BCCWJ)で頻度を調査し、形容詞「甘い」の事例も報告している。「甘い」に①<糖度が高い>、②<愛情のある>、③<厳しくない>という3つの語義を認めたとき、「甘かった」という語形では111例中89例と、8割もが③の意味使用に偏っていた。また、「甘すぎる」という語形でも③の意味使用に偏ることが報告されている¹。『日本国語大辞典』に見る歴史的出現の順序、及び山崎(2011)に報告される意味的な出現頻度に基づき①<糖度が高い>をプロトタイプの意味に認めるとき、上の計量的調査結果から、非過去形の語形にプロトタイプの意味が見て取れやすい可能性がある。また、この可能性は語と意味の認識における認知的負荷という観点からも導き出されることから、(1)のように仮説設定する。

(1) 非過去形の語形である終止形・い落ち形に語のプロトタイプの意味が認められやすい。

2. 調査対象の選定と多義認定

2.1 調査対象の選定

理論的に基礎的な意味タイプに属する語を題材にすることで、より妥当な一般化が期待される。Dixon(1982)の基礎的な意味要素の分類に基づき、次元特性を有する「長い」、「短い」を扱う。

[†] snishiuchi[at]ninjal.ac.jp

¹ 名詞に転成した「甘め」に①の意味使用が多くなることも報告されている。本研究では、異なる品詞に転成したとき特定の意味に特化される可能性を考慮し、扱わないこととする。たとえば、「うまい<上手だ / 美味>」から転成した「うまみ<美味>」は、<美味>に特化している。

2.2 調査対象の多義認定

まず、「長い」が多義であることは、類義語の違い(靱山 1993)によって(2)のように認められる。＜空間的隔たりが大きい＞という意味と＜時間的な継続が長い＞という意味である。次に、「短い」の多義をくびき語法などの修辞技法による逸脱の知覚(小松原 2016)によって認めるのが、(3)である。これらの意味のほか、「気が長い/短い」のように決まった名詞との述定用法で＜のんびりした/せっかちな＞という性格・態度を表すことがある。

(2)a. 長い髪 (類義語：伸びた) b. 長い歴史 (類義語：長期の)

(3) スカートとスピーチは短い＜空間的隔たりが小さい / 時間がかからない＞ほうがいい。

靱山(1995)は、「短い」について、(4)に見る用法上の制約から＜空間＞をプロトタイプの意味に認定している。しかし、「長い」については、同様の副詞化で＜空間＞も＜時間＞も表せるため、共時的な言語事実からプロトタイプの意味が認定できないと結んでいる。

(4)a. 花子は髪をミジカク切った。＜空間＞

b. ×私はアメリカにミジカク滞在した。＜時間＞(靱山 1995: 634. <>内は筆者による。)

3. 実験調査のデザイン

実験調査では、分類語彙表番号付与済み BCCWJ (加藤ほか 2019) から取得した例文、及び BCCWJ から抽出した例文を用いた。実例を用いることで、ある程度文脈が読み取りやすく、任意の一意が限定・想起されやすくなることが期待される。これらの実例を述定・装定の用法と多義で組み合わせ、データセットを作成した。このセットを用いて、ある1文を指標文にして、ほかの文(判定文)が類似しているかどうかを6段階でチェックしてもらう、クラウドソーシングを通じた大規模被験者実験(浅原 2019)を「長い」、「短い」各36例の全順列1260対、対毎に50人(延べ2100人：異なり「長い」723人、「短い」733人)実施した(図1)。実験協力者は、Yahoo!日本語IDを持つ20歳以上の男女である。意味は、『日本国語大辞典』及び『分類語彙表』から語源と複数の語義を付与した先行研究(山崎・柏野 2017, 加藤ほか 2019)を参照した。なお、タスクの性質上、指示文や例文を読まずに回答することが可能である。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設問を設けている。「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²。

図1 実験調査画面

4. 実験調査の結果と分析

調査結果に基づき、まず多義性を(5)のように確認する。次に、プロトタイプの意味を(6)と(7)のように認定し、(8(1再掲))の手法を検討する。

(5) 2節で認めた別義を表す用例間で類似が低く評定される。

(6) 認知的負荷によりプロトタイプの意味(仮)→派生義(仮)の類似度が高く評定される。

(7) プロトタイプ性により、いずれの語義間にも類似度が高く評定される(図2)。

(8) 非過去形の終止形・い落ち形に語のプロトタイプの意味が認められやすい。(1再掲)

² 「落選」した作業者は、「長い」に5人、「短い」に3人いた。

(5)を「長い」の結果から確認するのが表1である。かかる名詞が細長い対象で、＜空間的な隔たりが大きい＞ことを意味しているとき、同様の判定文に高類似を評定しやすい。また、(6)を「長い」の結果から確認するのが表2である。各語義の平均から、＜時間＞が指標のとき＜空間＞が指標のときより評定が低く見積もられやすい。最後に(7)を「短い」の結果から表3のように確認できる。

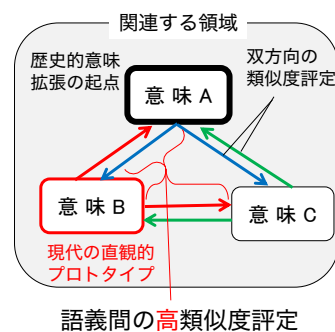


図2 語義間類似度高評定による検証

表1 「長い」の語義間高類似度

[illegible]

表2 「長い」の双方向評定の平均

	空間	空間/時間	時間	態度・性格
空間	3.17	1.87	1.40	1.69
空間/時間	1.88	2.41	2.29	1.60
時間	1.35	2.32	2.91	1.62
態度・性格	1.70	1.50	1.63	2.56

表3 「短い」の双方向評定の平均

	空間	空間/時間	時間	態度・性格
空間	3.05	2.48	1.65	1.51
空間/時間	2.59	2.45	1.77	1.44
時間	1.55	1.75	2.67	1.90
態度・性格	1.40	1.45	1.90	

以上の結果分析から、「長い」と「短い」に＜空間的な隔たり＞がよりプロトタイプの的であることが確認された。なお、「長い」では、「短い」ほど明快に語義間での派生義による認知的負荷が確認されなかった。しかしながら、靱山 (1995) では認められなかったプロトタイプの意味が量的調査から検討できる点で有用であると考えられる。この結果を踏まえ、終止形とい落ち形との類似度評定を「長い」を例に表4・5に見る。

表 4 終止形の類似度評定の平均

	終止形	空間	空間/時間	時間	態度・性格
終止形		3.05	2.94	2.84	2.67
空間	3.26				
空間/時間	2.98				
時間	2.83				
態度・性格	2.72				

表5 い落ち形の類似度評定の平均

	い落ち形	空間	空間/時間	時間	態度・性格
い落ち形		2.89	2.58	2.39	2.02
空間	2.99				
空間/時間	2.65				
時間	2.46				
態度・性格	2.20				

終止形では、いずれの語義とも高類似が評定されたが、中でも、プロトタイプの意味の＜空間＞に顕著である。い落ち形では、いずれの語義とも満遍なく高類似が認められたが、低く評定されるものもあった。このことは、(9)のような意外性のある実例がい落ち形「長っ！」と相性がよく、文脈の解釈が類似度評定に反映された結果だと考えられる。

(9) 仕事中は、動きやすくて脚が【長く】見えるパンツが大活躍！（類似度評定 3.5）

終止形とく空間的隔たりが大きい>ことを意味する用法のつながりの強さが高類似度(3.2以上)に確認されることを可視化したものが図3である。終止形が語のプロトタイプの意味に認められやすいことが検証された。

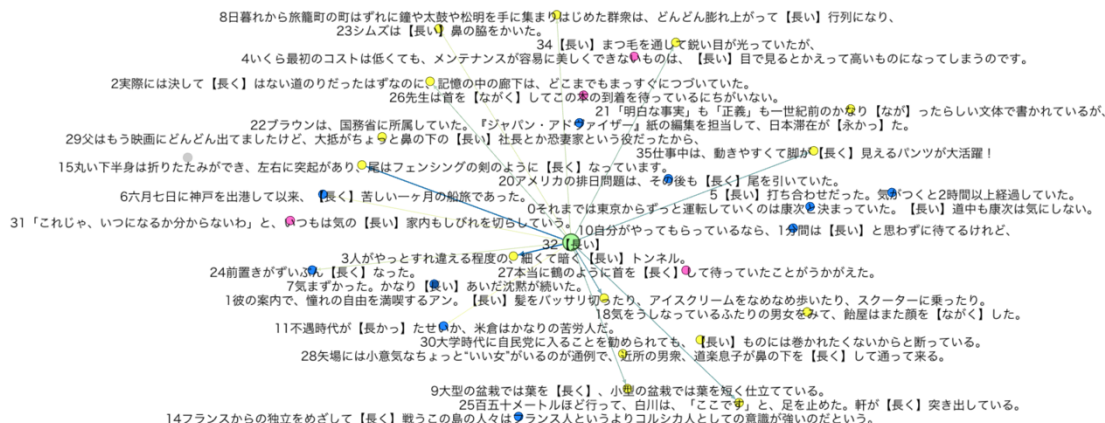


図3 高類似度による終止形と実例とのつながりの可視化

5. おわりに

多義語のプロトタイプの意味の認定において、本研究で提案する量的調査の分析手法が従来提案されてきた多様な手法を補うものになることを論じた。分析の観点として、脱文脈化した非過去形の語形が活用できることを仮定し、い落ち形が実例の文脈に依存し、終止形が適切に機能することを検証した。以上を(10)にまとめる。

(10) 脱文脈化した非過去形の語形、終止形が量的調査を通してプロトタイプの意味の認定に活用できる。

謝 辞

本研究は、国立国語研究所コーパス開発センター共同研究プロジェクト及び国立国語研究所所長裁量経費 2018, JSPS 科研費 19K00591 によるものです。

文 献

- 浅原正幸 (2019) 「クラウドソーシングを用いた言語分析」『日本言語学会第 158 回大会予稿集』 359-384.
- 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): To Appear.
- 木下りか (2019) 「多義動詞の意味拡張の起点と直観的プロトタイプ」『日本認知言語学会論文集』 19: 519-524.
- 国広哲弥 (1982) 『意味論の方法』大修館書店.
- 小松原哲太 (2016) 『レトリックと意味の創造性—言葉の逸脱と認知言語学』京都大学学術出版会.
- 田中茂範 (1990) 『認知意味論—英語動詞の多義の構造—』三友社出版.
- 榎山洋介 (1993) 「多義語分析の方法—多義的別義の認定をめぐる—」『名古屋大学日本語・日本文化論集』 1: 35-57.
- 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—意味転用の一方向性：空間から時間へ—」『東京大学言語学論集』 14: 621-639.
- 山崎誠 (2011) 「多義語を構成する意味の使用傾向—品詞と活用形による違い—」『言語処理学会第 17 回年次大会発表論文集』 659-662.
- 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.

Dixon, R. M. W. (1982) *Where have all the adjective gone? and other essays in semantics and syntax*. Berlin: Mouton.

関連 URL

コーパス検索アプリケーション『中納言』 <https://chunagon.ninjal.ac.jp/>