

文脈に応じた Synset 選択に基づく単語の定量的意味抽出の検討 A Study on Context-Based Synset Selection for Quantitative Meaning Extraction of Words

佐々木 康佑[†], 西川 純平[†], 森田 純哉[†]
Kosuke Sasaki, Jumpei Nishikawa, Junya Morita

[†] 静岡大学

Shizuoka University

sasaki.kosuke.19@shizuoka.ac.jp

概要

本研究は、文脈に応じた Synset 選択に基づき単語の定量的意味を抽出する手法を検討した。入力単語のカテゴリを区別して定量的意味を抽出し、ロボットのジェスチャーに反映させることで、人間が持つ定量的な軸を抽出することを目指した。ジェスチャーと発話の自然さを評価する実験を行った結果、文脈を考慮したアプローチにおいて、カテゴリだけでなく各単語に着目するなど、さらなる文脈の考慮が必要であることが示された。

キーワード: 言語, 自然言語処理, ロボティクス

1. はじめに

人間の扱う言語は身体的な経験と密接に関連している。Lakoff and Johnson (2008) によると、人間の語彙はメタファーの利用、つまり過去の類似の経験に基づいて新しい状況にラベル（単語）を割り当てることによって構築されてきた。このような身体的な経験からの概念へのマッピングを“プライマリーメタファー”と呼ぶ。そして、人間の発話においてはプライマリーメタファーがアイコンニックジェスチャーとして現れる (Li et al., 2023; Saund et al., 2022)。

言語と身体の関係は他の認知科学の理論分野においても活発に議論されている。例えば、Hawkins (2021) は各概念の背後には概念に関する知識を内包する連続的な空間が存在し、言語などを介在することで身体動作に変換されるという座標系理論を提案した。このような言語と身体結びつきをより深く理解するには単語の記号の意味と数量的意味を変換できる計算モデルを確立する必要がある。

そのようなモデルに至るための具体的なアプローチとして、Grand et al. (2022) は、単語分散表現のベクトル空間上で「大きさ」や「危険性」などの定量的な文脈に基づく軸を定義する手法を提案し、単語間の関係を抽出した。しかし、このアプローチには、軸の定義における類義語の恣意的な選択、単語のカテゴリの

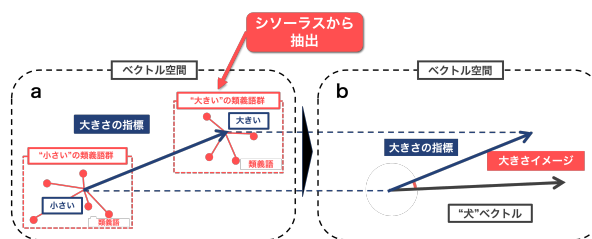


図1 提案手法の概要

考慮不足といった問題がある。この点から、著者らはシソーラスを利用して軸を定義し、カテゴリの違いを考慮した「大きさ」の数量的意味を算出する手法を検討してきた (Sasaki et al., 2023; 佐々木他, 2023)。ただし、著者らによる従来の手法においても、人間の有する定量的な軸を十分に抽出できたとは言い難い。そこで、本研究ではこれまでの著者らのアプローチについて再検討を進め、文脈を考慮することで単語の定量的な意味を抽出することを目的とする。

単語の意味は単体では曖昧であるが、文脈によって意味が明確になる。例えば、「犬」という単語に関する発話をする際、ペットという話題においては大きいイメージを持つが、生物全般に関する話題においては小さいイメージを持つ。このような事例を考慮し、本研究では単語の所属カテゴリが文脈として与えられた際の単語の定量的な意味の抽出を試みる。

2. 大きさイメージの抽出手法

筆者らの先行研究 (Sasaki et al., 2023) により提案された手法は、単語のカテゴリを考慮したシソーラスベースのアプローチである。この手法の概要を図1に示す。先行研究では、概念の定量的な軸として、「大きさ」に焦点を絞っている。そして、単語の“大きさイメージ”を検討する際に、「大きい」と「小さい」の両極からなる軸を“大きさの指標”にとらえ、“大きさの指標”と入力単語のベクトルのコサイン類似度を入力単語の“大きさイメージ”と定義する。

まず“大きさの指標”の両極を定義するため、シソー

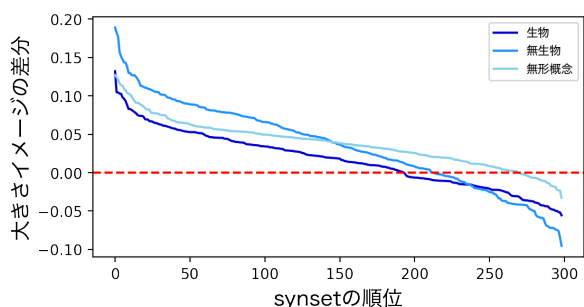


図2 “大きさの指標”の差分分布

ラスを参照して得られた類義語を利用する。シソーラスでは、類義語はそれらの意味ごとに集合 (synset) として定義されるため、集合の平均ベクトルを計算する。次に、これらの synset から最も人間の認識と整合する“大きさイメージ”を生成できる“大きさの指標”の極座標を定める。その際、人間に対する調査から得られる生物・無生物・無形概念における大きい・小さい単語の距離を最大化させる大きさイメージを調べる。これらの単語はクラウドソーシングによる調査により取得した。この大きさイメージが形成される synset の組み合わせを求めることで“大きさの指標”とする。大きい単語と小さい単語の大きさイメージの差分が大きいほど、人間が大きいと認識する単語を大きく、人間が小さいと認識する単語を小さく評価していることを示す。そのため、この差分が最大化された synset の組み合わせを利用することで概念の定量的な軸を抽出できると考える。

図2は調査により得られた大きい単語と小さい単語の距離を synset ごとに算出した結果を示す。本研究においてはカテゴリ間の違いを考慮するため、各カテゴリにおいて最も差分の大きい synset の組み合わせを用いることで“大きさの指標”を算出する。このような“大きさの指標”を用いて生成された“大きさイメージ”を身体動作の大きさにマッピングすることでジェスチャーが生成される。

先行研究 (Sasaki et al., 2023) に基づき生成された“大きさイメージ”を反映した、ロボットのジェスチャーの例を図3に示す。本研究においては単語分散表現から得られる“大きさイメージ”を利用した、より基本的な意味の軸に基づく大域的な動作に着目している。そのため、本研究で扱うロボットは腕と肩に関節を有しており基本的な動作を設計することが可能であれば良い。この条件を満たしたロボットとして、本研究では Vstone 社による小型コミュニケーションロボット Sota を用いた。

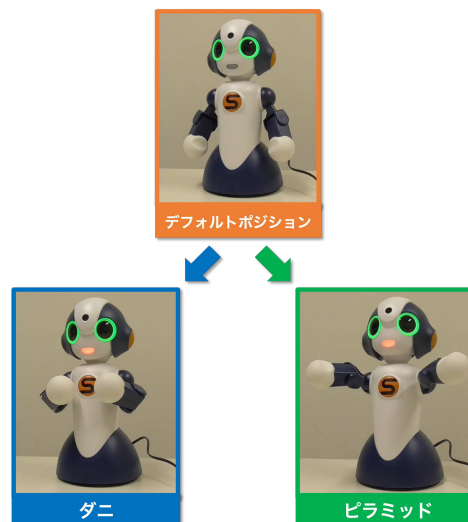


図3 ジェスチャーの生成例 (左下:ダニ, 右下:ピラミッド)

3. 方法

先行研究の実験をベースに、算出される“大きさの指標”を変更したジェスチャーを用いた実験を実施した。分析において、先行研究では生物・無生物・無形概念をまとめた汎用的な“大きさの指標”を定義していたのに対し、本研究では2章に示した手順によりカテゴリごとに独立の指標を定義した。そして、調査により得られた大きい単語と小さい単語の“大きさイメージ”の差分が最も大きい synset の組み合わせを実験群、最も差分の小さい組み合わせを対照群とした。

3.1 材料

本研究では、Sota の腕と肩と首の関節のパラメータを“大きさイメージ”によって制御することによりジェスチャーを生成する。本研究における Sota のデフォルトの姿勢は図3 (上) のように肩を落とし、腕を少し曲げた状態である。この状態から、腕と肩と首の関節のパラメータを変化させ、単語の大きさに対応したジェスチャーを生成する。“大きさイメージ”の値が最も大きい時には、図3右下のような姿勢を、値が最も小さい時には図3左下のような姿勢を取るよう設定した。最大値と最小値以外の値については、“大きさイメージ”の値に基づいてこれらの姿勢を最大最小とした範囲に紐づけられるよう設定した。

3.2 参加者

本実験にはクラウドソーシングサイト Lancers で募集した300名が参加した。それぞれの参加者は、依頼画面に表示された説明をに同意した上で実験に参加した。報酬は100円であった。

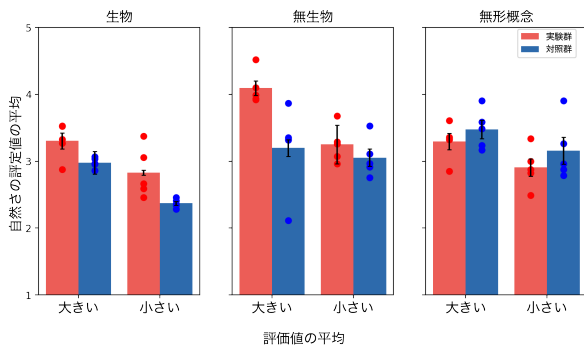


図4 自然さの評定値（ドットは各単語の評定値）

3.3 手続き

先行研究における人間に対する調査において得られた29の単語について生成されたジェスチャーをキャプチャした動画（約4秒）を実験群・対照群双方で作成した。なお、29の単語としたのは大きい無形概念と小さい無形概念の双方で「心」が出現したためである。

実験は、参加者が専用のサイトを訪問することで行った。サイトには、実験群と対照群の58本の動画から、参加者ごとにランダムに選択された14本の動画が配置された。これらの動画の間に、指定された数字を回答するように指示するダミーの動画が1本挿入された。各動画の下部には、「自然さ」を問う質問項目を設け、ジェスチャーを行うロボットの自然さを単語ごとに5段階で評価してもらった。なお、自然さとは言葉と動作の意味が対応しているかということと参加者に教示した。これらの手続きにより、提案手法が言語と身体の結びつきを表現できるのかを検証した。

4. 結果

各カテゴリにおける自然さの評定値の平均値を図4に示す。図に示されるように、実験群と対照群の差は大きい無生物というカテゴリのほかは顕著ではない。この結果は、先行研究と類似するものである。これらの結果の原因を調査するため、各単語の評定値に着目すると、「アリ」「ノミ」（小さい生物）「宇宙」（大きい無形概念）「嫉妬」「妬み」（小さい無形概念）において対照群が実験群を上回っていた。詳細な分析のため、図2で示した各synsetの組み合わせにおける各単語の“大きさの指標”を比較した。結果は図5に示されており、各図におけるプロットが各synsetにおける“大きさの指標”，青色の線は図2に示した差分を表す。

「アリ」や「ノミ」は調査において「小さい」単語として回答された。そのため、大きい単語と小さい単語の差分を最大化した上位のSynsetにおいて、“大きさの指標”は小さい値となるべきである。だが、プロットからは、これらの単語の上位のsynsetの方が下位

のsynsetよりも“大きさの指標”の値が大きい傾向があることがわかる。また、「宇宙」においてはsynsetの順位に関わらず全体的に“大きさの指標”が大きく、最上位のsynsetと最下位のsynsetの差分が顕著ではない。「嫉妬」や「妬み」においては他の単語と比較して非常にばらつきが大きい。

5. まとめ

本研究の結果は、文脈に応じたsynsetの選択においては単語間の違いを考慮する必要があることを示している。「アリ」、「ノミ」、「嫉妬」、「妬み」においては上位のsynsetで想定通りの結果が得られておらず、カテゴリ間の違いだけでなく単語間の違いを考慮することが有効である可能性を示唆する。「宇宙」においては他の単語と比較しても規模の大きい単語であるにも関わらず大きな差分が表れていないことから、本研究のシソーラスベースのアプローチでは大きさを判別することが困難であることを示唆する。

今後はLLMを活用した大きさイメージの生成方法を検討することで文脈を考慮することを目指す。現在Meta社の提供するLLMであるllama3を用いることで人間の認識と同様に大きさを判別できることを確認している。今後はllama3を提案手法に組み込み、よりインタラクションに踏み込んだ検討を進める。

文献

- Grand, G., Blank, I. A., Pereira, F., and Fedorenko, E. (2022). “Semantic projection recovers rich human knowledge of multiple object features from word embeddings,” *Nature human behaviour*, 6(7), 975–987.
- Hawkins, J. (2021). *A thousand brains: A new theory of intelligence*, New York: Basic Books.
- Lakoff, G. and Johnson, M. (2008). *Metaphors we live by*: University of Chicago press.
- Li, M., Aussems, S., and Kita, S. (2023). Is adults’ ability to interpret iconicity shared between the spoken and gestural modalities?, *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46.
- Sasaki, K., Nishikawa, J., and Morita, J. (2023). “Body movement construction based on distributionally represented size image,” in *Virtual MathPsych/ICCM 2023*.
- Saund, C., Matuszak, H., Weinstein, A., and Marsella, S. (2022). “Motion and Meaning: Data-Driven Analyses of The Relationship Between Gesture and Communicative Semantics,” in *Proceedings of the 10th International Conference on Human-Agent Interaction*, 227–235.
- 佐々木康佑・純平西川・森田純哉 (2023). 単語分散表現を用いた概念の数量的意味獲得”, 2023年度第40回日本認知科学会大会論文集, 511–514, 日本認知科学会.

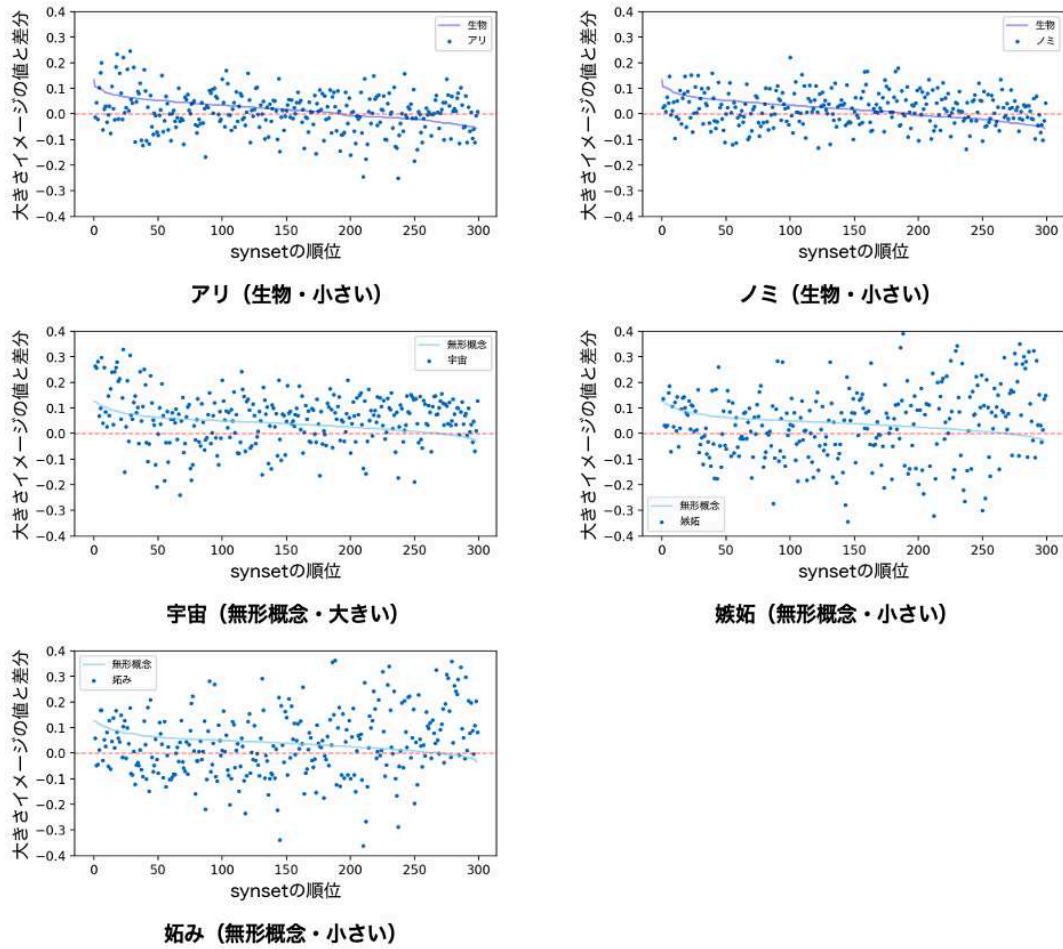


図5 各単語の大きさイメージ分布