

『分類語彙表』に対する反対語情報付与

加藤 祥[†]・浅原 正幸^{††}・森山奈々美^{†††}・荻原亜彩美^{††,†††}・山崎 誠^{††}

本論文ではシソーラス『分類語彙表』に対する反対語情報付与作業について示す。分類語彙表上では反対語対に対しても同じ分類番号が付与されている。この同じ分類番号が付与されている単語群から反対語対を抽出し、分類作業を行った。まず、人手により反対語候補となる単語対を網羅的に収集した。次に、一般的な日本語話者が反対語と感ずるかをクラウドソーシングにより収集し、50%以上の方が反対語とみなしたものを反対語と定義した。最後に、得られた反対語リストに対して、村木の「対義語」の分類を精緻化したものを付与した。分析にあたっては、反対語認識の非対称性・分類語彙表ラベル・分類・コーパス頻度・単語埋め込みなどの観点から検討を行った。言語学的な分析においては、閉じた反対語対に対してより反対語らしさを感じる傾向が確認された。言語処理的な分析においては、単語埋め込み上で反対語対の置き換え可能性の評定とコサイン類似度に中程度の相関があることが確認された。

キーワード：分類語彙表・反対語・対義語・クラウドソーシング・アノテーション

Opposite Information Annotation on ‘Word List by Semantic Principles’

SACHI KATO[†], MASAYUKI ASAHARA^{††}, NANAMI MORIYAMA^{†††}, ASAMI OGIWARA^{††,†††}
and MAKOTO YAMAZAKI^{††}

This paper presents research on opposite information annotation on the thesaurus ‘Word List by Semantic Principles (WLSP)’. The categorised words with the same label include antonym word pairs. We extract the opposite word pairs from the categorised group and classified the opposite word pairs. Firstly, the annotators manually extracted opposite word pair candidates. Secondly, we utilised Yahoo! crowdsourcing to evaluate how many people recognise the opposite word pair as the opposites. We defined ‘opposites’ as the words judged by greater than or equal to 50% people. Thirdly, we annotated the opposites types by Muraki for the word pairs. We analysed the opposite word lists by their asymmetry, the label of WLSP, the opposite types, their frequencies and word embeddings. In the linguistic point of view, the closed opposite word pairs tend to be regarded as ‘opposites’. In the natural language processing point of view, the distances between two words in the opposite pairs correlates their replaceability of the human judge.

Key Words: *Word List by Semantic Principles, Opposites, Crowdsourcing, Annotation*

[†] 目白大学 外国語学部, Mejiro University

^{††} 人間文化研究機構 国立国語研究所, NINJAL, Japan

^{†††} 津田塾大学, Tsuda University

1 はじめに

本稿では『分類語彙表』(国立国語研究所 2004)に対する反対語・対義語情報付与について論じる。分類語彙表は単語の意味をその語義に基づいて分類している。分類語彙表の初版の書籍は1964年に出版され、増補改訂版の書籍が2004年に出版された。この増補改訂版をもとにしたCSV形式のデータベース、分類語彙表増補改訂版データベース(以下『分類語彙表DB』)が公開されている。同データは区切り記号を含めて101,070 エントリからなる。

表1に分類語彙表の構造を示す。分類語彙表の単語・語義は、分類番号・段落番号・小段落番号・語番号が割り当てられている。これらの番号が、語の階層的なクラスを構成している。

分類番号は、語の統語分類を表す類と、意味分類を表す部門・中項目・分類項目により構成されている。分類語彙表のピリオドより前の数字が類を、ピリオドの後1ケタの数字が部門を、2ケタが中項目を、4ケタが分類項目を表す。表2に「最大」(分類番号1.1920)の例を示す。ここで最初の1は統語分類の類「体」を表す。ピリオドより後の.1920が意味分類を表し、最初の.1が部門「関係」、.19が中項目「量」、.1920が分類項目「程度・限度」を表す。

分類語彙表の意味の階層構造は反対語(opposites)・反義語(antonyms)も含めた類義語(synonyms)¹に対して同じラベルを割り当てる。例えば、表1の小段落番号1.1920-01-01には、「最大」⇔「最小」と「最多」⇔「最少」といった単語対が反対語だと考えられる。反対語は基本的には二項対立の単語対に対して認めるものであるが、三項対立のものに対して認めるものもある。

類	部門	中項目	分類項目	分類番号	段落番号	小段落番号	語番号	レコード番号	見出し
体	関係	量	程度・限度	1.1920	08	01	01	013631	最大
体	関係	量	程度・限度	1.1920	08	01	02	013632	最小
体	関係	量	程度・限度	1.1920	08	01	03	013633	最多
体	関係	量	程度・限度	1.1920	08	01	04	013634	最少

表1 分類語彙表の構造

「最大」: 分類番号 1.1920			
統語分類	意味分類		
類	部門	中項目	分類項目
体	関係	量	程度・限度
1.	.1	.19	.1920

表2 分類語彙表の項目「最大」

¹ 本稿では類義語(synonyms)と反対語・対義語(opposites)と反義語(antonyms)とする。反対語・対義語の定義については2.1節で論じる。以降、反対語・対義語を含めて「反対語」と呼ぶ。

本研究では、分類語彙表の類義語に対する反対語情報を大規模に付与する。まず、反対語と少しでも考えられる単語対を人手で分類語彙表から抽出した。次に、展開した反対語対候補とランダムに抽出した単語対を文字刺激として、反対語らしさの評定値をクラウドソーシングにより収集した。さらに、クラウドソーシングにより、正順呈示・逆順呈示のいずれかで50%以上の方が反対語であると認識した語対について、反対語の分類情報を付与した。これらのデータベースについて、反対語の分類ごとの評定値情報・コーパス頻度・単語ベクトルの傾向について調査した。

2節では、言語学における反対語の関連研究と、言語処理における極性分析などについて示す。3節に反対語情報の付与方法を示し、4節に分析結果を示す。5節にまとめと今後の課題について示す。

2 関連研究

2.1 関連研究：言語学の観点から

2.1.1 Cruse の反対語 (opposites) の分類

反対語 (opposites) は様々な定義があり、何を反対と考えるのかという基準や評価はいまだ明確ではない。Cruse (1994) の言及によると、反対語 (opposites) は相補語 (complementaries)・逆語 (converses)・反転語 (reversives)・反意語 (antonyms) を含む一般的な概念であり、

- (1) 対立的な反対性 (diametric opposition)
- (2) 二値性 (binarity)
- (3) 上位概念における包括性 (exhaustiveness of superordinate domain)
- (4) 対称性 (symmetry)

などが特徴としてあげられる。しかし、反対語と呼ばれるものが、これらのすべての特徴を備えているわけではないことに言及している。例えば、反対語対 (huge, tiny) は、(large, small) と比べて、上位概念における包括性は低い。

方向における反対語 (directional opposition) は、空間における動作も含めて定義しやすいとされる (Cruse 1986)。方向から発展して、連想により定義されるものも含まれる²。以下にさま

² 基準の示し方においては、内包的な定義もあれば、外延的な記述などさまざまな手法がある。例えば、Cruse (1994) においては、

- 相補語 (complementaries): 「X であること」が「Y でないこと」と同値である場合に、X と Y は相補と言う。
- 逆語 (converses): 変数 A, B をとる関数 $F^X(A, B)$ が成り立つ場合かつその場合にのみ $F^Y(B, A)$ が成り立つとき、X と Y は逆語と言う。

のように内包的に定義するものもあれば、

- 相補語の例 (dead, alive)
- 逆語の例 (is above, is below)

のように外延的な記述もある。但し、(Cruse 1994) は、すべての反対語の概念を内包的に定義もしくは外延的に記述することは難しいとも言及している。

ざまな反対語について示す。

- 対蹠語 (antipodal) : 何らかの軸の両端に位置するものから発展して, 両極として対になる表現 (例 : (top, bottom), (maximum, minimum))
- 対照語 (counterparts) : 位置的に対照的なものから発展して, 概念的・文化的に対になる表現 (例 : (ridge, groove), (hill, valley), (male, female), (heaven, hell), (yin, yang))
- 反転語 (reversives) : 反対方向への動作や変化を示す動詞など (例 : (rise, fall), (appear, disappear), (enter, leave))
- 逆語 (converses) : 2つの対象の相対的な関係を表す対 (例 : (above, below), (ancestor, descendant) (doctor, patient))

反対性に関わる極性 (polarity) には形態上あるいは論理上のほか, 何らかの特徴の有無, 評価的なポジティブ/ネガティブがあるとしている。Cruse (1994) によれば, 反対語は上記のような二項対立だけでなく, 三項以上の対立 (lack of congruence) も含まれる。

また, 反対語の中で, 反意語 (antonyms) は尺度の対立に関連する狭い範囲の語対を表す。Cruse (1986) の定義では,

- (1) 段階的なもの (gradable) で, ほとんどは形容詞であり一部動詞を含む (most are adjectives; a few are verbs)
 - (2) 何らかの変数の程度 (degrees of some variable property)
 - (3) 程度を表す尺度に沿って反対の方向性がある (opposite directions along the scale representing degrees)
 - (4) 語対により厳密に領域を2分しない (the terms of a pair do not strictly bisect a domain)
- と, 二項対立性や上位概念における包括性などを厳密に規定しない。つまり, 閉じた対ではなく, 開いた対であると言える。なお, (Murphy 2003) は, 反意語 (antonym) の下位分類に (1) contrary, (2) complementary (contradictory), (3) converse, (4) reversible としたうえで, (3), (4) に種々の雑多な例 (一般に方向関連) を含める。

2.1.2 村木の「対義語」の分類

日本語においては, 村木 (1987) が, 反意語 (antonyms) よりも広い意味で「対義語」を捉え

- (1) 相補関係に基づく反義対 : ある条件下で, 概念の領域を分割する対
- (2) 両極性に基づく反義対 : ものごとの対極を名付けた単語の対立
- (3) 程度性をもつ反義対 : 空間量, 重量, スピードといったものごとの属性や性質をあらわす単語の対立 (多くは形容詞)
- (4) 視点がらみの反義対 : (a) 同一の対象や過程が2つの異なる視点から名付けられた単語対, (b) 互いに他を前提としてなりたっている単語対

(5) 変化に関わる反義対：(a) ある空間を考慮して、互いに逆方向に移動することを含意する単語対、(b) 状態の変化に関与するもので、互いにもとの状態にもどる関係の単語対を挙げる。

Cruse の反意語 (antonyms) は概ね村木の「程度」にあたるといえる。三項以上の対立の単語対や相反するとはいえない側面の単語対も含める。このほか、対比が反対の意識になったと考えられる単語対を、開いた対義対と位置付け、二値性・全体-部分・二側面・文脈に依存する反義対・慣用句による反義対などを挙げる。また、村木が（閉じた対と区別して）開の対に分類する「都会：田舎」などが、Cruse では方向性で説明される（開閉の区別はない）ほか、「相補」関係と考えられる「表：裏」のような対や「変化」にあたる対は反転語 (reversives)、「視点がらみ」の対は逆語 (converses) の扱いとされ、反対を持つ対の分類は大きく異なる。詳細については、3.3 節に示す。

実際にどのような対が反対と認識されているのか、どのようなタイプの対が反対と認識されているのか、理論上の分類で反対の意識の説明が可能であるのか、実態からの調査や分析が待たれる。

2.1.3 反義性の調査

反対語において何が「反対」なのかを解明するため、辞書やコーパスから収集した反対語対の関係性や反義性の度合を調査する試みがなされている。

反対語対における反義性の強さについても、議論や調査がなされてきた。たとえば Herrmann et al. (1979) は、反意語 (antonyms) の程度の違いを実験的に調査し、関係性の類似が項目の類似よりも反対性に関わるとした。Herrmann et al. (1986) は、疑念構造を反意語 (antonym) 対の要因とし、速さ、明るさ、強さ、大きさ、幅、価値、厚さのような概念次元を基盤と考える。Cruse (1986) は、一次元で概念化できること、対立で意味が言いつくされていること、その他の意味に関して一致度が高いことなどが影響するという。村木 (1987) は、一次元の尺度で対立すること、つりあいが対立で言いつくされていること、文脈の不要性、同種類（品詞・語種・文体的特徴・共通語形部分）を挙げている。反義性の良し悪しを、コーパス調査、産出実験、心理実験などから検証 (Paradis et al. 2009) する試みも見られるが、いずれの要素も反義性の判断に影響している可能性が考えられるとはいえ、何がどれだけ影響しているのかは明確ではない。

日本語の反対語の実態的な調査の例には、荻野、野口 (1988, 1996) による「通常日本語話者が考えている」反対語の報告がある。165 反対語対の反義性判断実験を行ったもので、結果から、一般的な「対比」の意識は雑多で個人差があるものの、意味関係よりも現実世界のモノの関係でとらえられ、二項対立の単語対に反義性の意識が強い傾向が分析されている。

また、日本語の反対語の程度性については、松本 (2007) が、138 反対語対の反義性判断実験の結果に基づき、反義性の程度を方向性と極性の概念的な対立の観点から分析している。対比

的な比較において方向性が顕著な単語対で反義性が高く判断される傾向と、肯定性・否定性の程度が関わる傾向が示された。但し、Cruse (2004) は、(doctor, patient) のような単語対について内在する二項対立を考えるが、松本 (2007) は、極性の明確さが要因と分析する。また、方向性に関する単語対では「両端開放型スケール」³で反義性判断程度が高いという結果も得られている。これらの日本語を用いた先行研究はいずれも百数十単語対規模の実験であるが、現実世界のモノの関係・比較において方向性が顕著という点において一致しているといえる。村木 (1987) が「程度」に分類する形容詞対 (いわゆる Cruse の「反意語」(antonyms)) で反対性の強い可能性も考えられる。反対語と認識されている対を網羅的に収集した検証が求められよう。

なお、荻野, 野口 (1996) では、提示順を逆転させた調査において、165 単語対中 15 単語対において反対語らしさの大小が異なっていた。この現象の理由として、先に読んだ語に複数の反対語が意識される場合、単語対の反対語意識が低下する傾向を分析している。Paradis et al. (2009) の調査でも、反対語対間の結びつきの強さが連想の起点とする語によって変わる例が示され、想起される使用文脈との関係によると指摘されている。しかし、提示順の違いによる反対語らしさの違いが、先に示された語の多義性のみで全て説明されるのではない。荻野, 野口 (1996) は、提示順の逆転で反対語らしさが異なっていた (以前, 以後), (上, 下) などの単語対はあてはまらないとし、「前後」「上下」のような熟語の影響を否定している。反対性の意識が語の提示順で異なる例について、どの程度どのような種類として存在するか調査し、反対性の意識に関わる要因を確かめる必要がある。

2.2 関連研究：言語処理の観点から

以下では、主に単語埋め込み (Mikolov et al. 2013a, 2013b) を含めた分布意味論における反対語の扱いについて紹介する。Mohammad et al. (2013) は、分布意味論で構成される単語の類義性評価において反対語同士の分布意味論的な距離が近いところに出現するとしており、Ono et al. (2015), Rothe and Schütze (2015) は単語の対義関係を単語埋め込みから判断するのは難しいとしている。一方、Mikolov et al. (2013c) が word2vec 上の linguistic regularity の例に出した

$$\vec{k\acute{i}ng} - \vec{m\grave{a}n} + \vec{w\acute{o}m\grave{a}n} = \vec{q\acute{u}\acute{e}e\acute{n}}$$

における male/female 関係は閉じた対義関係にあるとも言えるだろう。日本語においては、単語埋め込み中にも反対語間で文脈的な差異があるとして、単語埋め込み上での識別を試みた研究もある (佐藤 他 2016; 中村 他 2018, 2019; 川島 他 2019) が、主として評価的なポジティブ／ネガティブを対象としており、言語学的に精緻な反対語の調査は行われていない。

³ 両側開放型の例として (大きい, 小さい), (長い, 短い), 片側開放型の例として (無害の, 有害の), (同じ, 違う), 両側閉鎖型の例として (満杯の, 空の), (真っ白な, 真っ黒な) があげられる。

村木 (1987) においては, (man, woman) や (king, queen) ⁴は相補 (complementation) に相当する閉じた反対語対である. Mikolov et al. (2013c) の例では male/female の例として, (uncle, aunt) を示しており, これも相補に相当する反対語対⁵である. 一方, 視点 (viewpoints) に相当する uncle の反対語として, nephew や niece が考えられる. 松本 (2007) においては (man, woman) は非関係的であるが, (king, queen) · (uncle, aunt) · (uncle, nephew) · (uncle, niece) は関係的である. これらの先行研究から, はたして単語埋め込み上で反対語対がどのように表現されるのか, 適切に収集した反対語データベースに基づく検証が求められよう. さらに村木 (1987) の「対義語」の分類ごとに異なる差ベクトルを持つのかを検証したい.

3 データ構築方法

日本語の反対語対を収集し, 一般的な日本語話者による反対語かどうかの判断を取得する. 反対語の定義として, 一般的な日本語話者に対して質問紙調査を行った際の「反対語である」の判定が 50% 以上のものを「反対語」とする. また, 反対語として得られた対の分析を行い, 先行研究で示された傾向を大規模なデータで検証するとともに, 反対語とされる単語対のタイプ分類と, 反対語と認識されやすいタイプについての調査も試みる.

具体的なデータ構築方法は次のとおりである. 最初に分類語彙表から反対語に相当する候補を手により収集した (3.1 節). 次にクラウドソーシングにより反対語であるか否かの評定値情報を収集した (3.2 節). 最後に手により村木 (1987) による反対語のタイプ (村木においては「対義語」) を付与した (3.3 節).

3.1 反対語対候補の抽出

4 人の作業者により, 分類語彙表の階層構造において同じラベルを持つものを対象に, 反対語対候補の抽出を行った. まず, 4 人の作業者により, 小段落番号が同じ単語群の組み合わせ 162,990 対から反対語候補対を抽出した. 次に, 2 人の作業者が段落番号が同じ単語群の組み合わせ 842,459 対から反対語候補対を抽出した. この際, 小段落番号が同じ単語群において漏れがあった場合には追加した. 抽出にあたっては, 「三省堂反対語対立語辞典」に採録されている単語対についても参照した. 抽出作業は 2017 年 6 月に開始し, 2018 年 11 月に終了した.

本作業で 7,658 の反対語対を抽出した. 3,405 語対は小段落番号が同じ単語群から, 4,253 語対は段落番号が同じ単語群から抽出したものであった.

⁴ (king, queen) は, 「最高支配者」という領域において相補.

⁵ (uncle, aunt) は「父母のきょうだいもしくは父母のきょうだいの配偶者」という領域において相補.

設定した設問ID: inv1

以下の2つの語句が対義語・反対語か否かを判定してください。
また、テキストに出現した場合に、2つの語句を置き換えても文法的に正しい（格助詞が変わらないなど）か否かの判定もお願いします
(文法的に置き換え不可な例: 「AにBを」 「加算する」 ⇔ 「AからBを」 「減算する」)
(文法的に置き換え可能な例: 「北」 ⇔ 「南」)

【弱い】 と 【強い】

対義語・反対語でない

対義語・反対語であるが、置き換え不可

対義語・反対語であり、置き換え可

100102-100109

図 1 クラウドソーシングによる評定値情報収集画面

3.2 クラウドソーシングによる評定値情報

3.1 節で抽出した 7,658 反対語対候補に、ランダムで追加した 4,342 の単語対⁶をあわせた 12,000 単語対を対象として、一般の方が反対語として認識するかどうかの評定値を Yahoo! クラウドソーシングにより収集した。実験対象者は、20 歳以上の Yahoo! Japan ID を所持する者とした。実験に際し、不適切な回答を行う実験不適切者の排除を行ったが、性別・居住地などによる統制は行っていない。図 1 に評定値情報収集画面の例を示す。反対語の程度を「対義語・反対語でない」・「対義語・反対語であるが置き換え不可」・「対義語・反対語であり、置き換え可」の 3 択から選択させた。「対義語・反対語でない」は、単語対に対義語・反対語性を見いだせないときに選択する。「対義語・反対語であるが置き換え不可」は、単語対に対義語・反対語性があるが、格交代などにより文脈によっては置き換えができないときに選択する。置き換え不可の例として、「(AにBを) 加算する」 ⇔ 「(AからBを) 減算する」を示した。「対義語・反対語であり、置き換え可」は、単語対に対義語・反対語性があり、置き換えができるときに選択する。置き換え可の例として、「北」 ⇔ 「南」と「暑い」 ⇔ 「寒い」を示した。

単語対の呈示順として、分類語彙表上の ID の昇順に示したもの（正順呈示）と降順に示したもの（逆順呈示）の 2 種類を設定した。クラウドソーシングは 2 回に分けて実施した、1 回目は 2018 年 12 月 17 日 8:03～21:25 に行い、1,597 人が参加し、18 時間 21 分を要した。2 回目は 2019 年 11 月 22 日 8:03～24 日 7:40 に行い、1,753 人が参加し、47 時間 43 分を要した。費用は税別で 40 万円を要した。各単語対は順方向 20 人・逆方向 20 人評価を行った。

3.3 反対語タイプアノテーション

クラウドソーシングにより、実験協力者 50% 以上の方が正順呈示・逆順呈示のいずれかで

⁶ 小段落番号が同じ単語群からランダムに 2 語をフィラー候補対として抽出した。

「対義語・反対語である」(置き換えの可不可問わず)の5,538単語対を評価対象とした。ラベル分類は閉じた対・開いた対を確認しながら村木(1987)の分類にならった。但し、反対語と評価された対には、村木(1987)では示されていない関係性の対が存在することや、反対性が見いだせない(分類不能)ことなどが予想されたため、閉開ともに「その他」のラベルを設定した。閉じた対のタイプは以下の通り。

- **相補 (complementation)**: ふたつの単語 X と Y がある条件下で概念の領域を分割する対。一方が肯定されれば他方が否定される。中間段階を認めない。
例) (男, 女), (父, 母), (上, 下), (裏, 表), (真, 偽), (等しい, 異なる), (同じ, 違う), (出勤, 欠勤), (合格する, 落第する), (引き受ける, 断る)
但し、人間の判断にかかわる関係については、どちらも否定可能で厳密ではない場合がある⁷。
- **両極 (bipolar)**: 連続的であれ不連続的であれ、ものごとの両極に位置する部分の対。中間段階が存在する。
例) (満点, 零点), (最高, 最低), (最大値, 最小値), (頂上, ふもと), (天, 地), (始発駅, 終着駅), (開会, 閉会), (入学, 卒業), (始まる, 終わる)
- **程度 (degree)**: ものごとの属性や性質をあらわし(多くは形容詞)、程度を特徴づける副詞類(「とても」「少し」など)と共起する対。対立するふたつの語の意味領域は厳密には別れない(差は相対的)。中間点は多くの場合存在しない(≠両極)が、一方の否定は他方の肯定を意味しない(≠相補)。例) (大きい, 小さい), (高い, 低い), (重い, 軽い), (早い, 遅い), (暑い, 寒い), (きれいな, 汚い), (安全, 危険), (積極的, 消極的), (うれしい, 悲しい), (太った, やせた)
- **視点 (viewpoints)**: 現実の対象世界に属さず、人間の対象のとらえかたに関わる対。
 - 同一の対象や過程がふたつの異なる視点から名づけられた単語のあいだに成立。
例) (上り坂, 下り坂), (入口, 出口), (スタート地点, ゴール地点), (いく, くる), (売る, 買う), (貸す, 借りる), (教える, 教わる)
 - 互いに他を前提として成立。
例) (親, 子), (夫, 妻), (先生, 生徒), (医者, 患者), (師匠, 弟子)
- **変化 (change)**:
 - ある空間を考慮して、互いに逆方向に移動する(位置の変化が逆方向)関係。
例) (あがる, さがる), (つく, 離れる), (到着, 出発)
 - 状態の変化に関与し、互いにもとの状態に戻る(可逆的)関係。
例) (現れる, 消える), (結ぶ, ほどく), (覚える, 忘れる), (積む, 下ろす), (生産,

⁷ 例: 賛成-反対の中間に「保留」、信じる-疑うは「信じもしないし疑いもしない」があり得る。

破壊), (着衣, 脱衣), (暖める, 冷やす), (のばす, ちぢめる)

開いた対は, ものごとのふたつの側面をとりだして対比し, 反義性を感じる対である. 文脈に依存するが, 二値性や両極性が認容される.

- **典型 (2 値)** (representative): ある対象にふたつしかないなど, ふたつがものごとの典型である対.
例) (和室, 洋室), (都会, 田舎)
- **部分全体** (whole-part): さししめしている対象の関係が全体と部分である対.
例) (往復, 片道), (両手, 片手)
- **2 側面** (two sided): ものごとのふたつの側面に注目した対.
例) (たて, よこ), (水平, 垂直), (一般, 特殊)
- **慣用** (idiom): 慣用句として反義性をもつと感じられる対.
例) (気を失う, 我にかえる), ((気が) はれる, (気が) ふさぐ), (骨をおる, 骨をおしむ)
- **その他開** (others): 上記以外のもの.

最初に 3 人の作業者が独立に作業を行った. 関係性の判定においても, 反対性の考え方には個人差が生じることになる. 3 人の判定が完全に一致したのは 1,539 対 (27.8%), 2 人の判定が一致したのは 2,625 対 (47.4%), であり, 3 人とも一致しなかったのは 1,374 対 (24.8%) であった. 2 人以上一致した 4,164 対 (75.2%) については, 一致したものを最終結果として採用した. なお, 特に判定が一致しなかったものは閉の「相補」「両極」「程度」(597 対) と, 開 (425 対) の判定である. 判定が一致しなかったものについては, 4 人目の判定を追加するとともに, 「その他開」は傾向により次の 3 つのタイプを導入し, 最終結果とした.

- **終了** (finished): 継続していた状態とその終了を示す語の対. 終わった状態であれば「相補」関係が成立するため問題はないが, (授乳, 離乳) のような反対語対においては判定が一致していなかった. ほかに (進捗, 停滞) などがある.
- **主副 (別)** (main-sub): 一方が主であるものの「相補」や「両極」, 「部分全体」とも判定し難い (本流, 支流), (直行する, 迂回する) などで判定揺れが生じていた.
- **因果** (cause-effect): 一方が原因となる単語対に付与した. (におう, 脱臭する) や, 因果関係を示す (起因する, 結果する) のような単語対で判定が一致していなかった.

最終的に「その他開」が付与されている対は, その他の反義でない (が反対語と半数以上に判定された) 類義語対, 反義ではあるが態が異なり反対と認められなかった単語対 (例: (首がつながる, 首にする)) である.

4 分析

4.1 対応する反対語数の分析

まず、反対語対は必ずしも 1 対 1 対応しているわけではなく、1 対多対応している場合がある。表 3 に 1 つの反対語に対応する反対語数の分布（反対語であるとの認定が正順呈示・逆順呈示のいずれかで 50% 以上のもののみ）を示す。9,286 語⁸中、88% の 8,184 語が 1 対 1 対応しているが、残りの 12% は複数の反対語を持つ。

Cruse (1986) は、反対語を共有するという合同関係に基づく合同類 (congruence variants) についての検討を行っている。Cruse は反対語の合同類を 上位下位型 ('hypo-/super' type) と逆元なし型 ('semi-' type) の 2 種類に分けたうえで、前者のみが頻出するとしている。日本語では上位反対語として「脱ぐ」の例（「着る」「かぶる」「履く」「はく」「穿く」）を示しており、本データベース上でも確認できた（表 3）。本データベース上で見られたそのほかの例として、性差・視点のものがあつた。例えば、表 4 に示すとおり、「ホステス」「ホスト」「ゲスト」は、それぞれ互いに反対語の関係にあるが、性差の関係としての相補と、主客の関係としての視点の関係がある。動詞の例としては「婿取りする」が、視点として「婿入りする」や、典型として「嫁とりする」などの反対語を持つ。

表 5 に対応する反対語数が多い用例を示す。分類語彙表に含まれる類義の慣用表現が多い場合に、1 つの反対語に対応する語数が多い傾向が確認できた。また、動詞において、自他交代した事例が出現する傾向も見られた。

対応する反対語数	左の反対語数を持つ単語数
1	8,184
2	761
3	191
4	86
5	33
6	16
7	2
8	7
9	4
10	1
12	1

表 3 1 つの反対語に対応する反対語数の分布

対象語	タイプ	反対語
脱ぐ	(変化)	着る
		かぶる
		はく, 穿く
		はく, 履く
ホスト	(相補)	ホステス
		ゲスト
ホステス	(相補)	ホスト
		ゲスト
ゲスト	(視点)	ホステス, ホスト
婿取りする	(視点)	婿入りする, 婿に入る
	(典型)	嫁取りする, 嫁をもらう, めとる

表 4 複数の反対語に対応する例

⁸ 分類語彙表番号が異なる語を別語とした場合の、反対語対のいずれかの要素になる異なり語数。

対象語	反対語数	対象語の反対語のリスト
気を楽しむ	12	緊張する 堅くなる 気が引き締まる 気を引き締める 意気込む 手綱(たづな)を締める 気が張る 気張る 気を張る 氣勢を張る 精神を集中する 気迫を込める
リラックスする	10	緊張する 上がる 堅くなる 気が張る 気張る 気を張る 氣勢を張る しゃんとする 精神を集中する 気迫を込める
一定する	9	移り変わる 変移する うつろう 変遷する 有為転変(ういてんぺん)する 変動する 流動化する 変転する 転変する
衰える	9	盛(さか)る 盛んになる 興る 興す 興隆する 勃興する 繁栄する 繁盛する 栄(さか)える
口を滑らす	9	箝口(かんこう)する 緘黙(かんもく)する 口を閉ざす 口をつぐむ /を結ぶ/を緘する 口をつぐむ 口を結ぶ 口を緘する 黙止する 黙秘する
口が滑る/を滑らす	9	箝口(かんこう)する 緘黙(かんもく)する 口を閉ざす 口をつぐむ /を結ぶ/を緘する 口をつぐむ 口を結ぶ 口を緘する 黙止する 黙秘する

表 5 対応する反対語数が多い用例

4.2 反対語データベースを構成する語の多義性

次に反対語の多義性について検討する。多義性を検討するにあたり、WLS2UniDic (分類語彙表番号-UniDic 語彙素番号対応表)⁹(近藤, 田中 2020)を用いる。反対語対に出現する分類語彙表の語に対して、UniDic の語彙素番号と対応をとったうえで、UniDic 側からみて分類語彙表の語義が何種類割り当てられているかを集計したものとその例を表 6 に示す。語義数の欄が「-」とあるのは、分類語彙表と UniDic の単位が異なるために対応する要素がないもの(3,050 例)である。残りの 6,232 例中、59% の 3,651 例が UniDic の語彙素単位に単義の語であった。残りの 41% は多義性を持ち、多いものでは「掛ける」が 12 の語義を持つ。ここでは、複数の反対語に対応する多義語の例として「甘い」を表 7 を示す。「難易」「待遇」「味」の 3 つの語義に対して、「厳しい」と「辛い」が反対語として対応していることがわかる。

4.3 クラウドソーシングによる評定値の分析

まず、表 8 にクラウドソーシングによる評定値の集計結果について示す。候補外の単語対に対して 83.8% の回答が「反対語でない」である一方、反対語候補の単語対に対しては 34.3% (小段落一致候補)・35.9% (段落一致候補)であった。「反対語でない」の判断においては、小段落一致と段落一致の差は 1.6% (35.9% - 34.3%)であった。しかしながら、「置き換え可」か否かについては、小段落番号一致のほうが段落番号一致よりも 5.6% (42.4% - 36.8%) 大きな値で

⁹ <https://github.com/masayu-a/WLS2UniDic/>

語義数	頻度	例
—	3,050	上の上（じょう） 悪条件 オンライン 新体制
1	3,651	前項 真性 類似 メリット 帰り道
2	1,743	実（じつ） 偶然 上がり目 良性 アップ
3	477	真実 本流 プラス 全部 得手
4	207	こちら イレギュラー 不純 上がり
5	70	あちら ダウン ショート 下（くだ） る
6	41	純粹 清潔 休め 落ちる 薄い
7	22	先（さき） 上（のぼ） る やる
8	14	整う 乗ずる 明ける 難しい
9	4	変わる 上がる 悼む
10	4	内（うち） 上げる
11	2	下ろす 降ろす
12	1	掛ける

表 6 反対語データベースを構成する語の多義性

対象語	タイプ	反対語	分類語彙表番号	分類項目
甘い	(程度)	厳しい	3.1346	相-関係-様相-難易・安危
甘い	(程度)	厳しい	3.3680	相-活動-待遇-待遇・礼など
甘い	(程度)	辛（から）い [点が一]	3.3680	相-活動-待遇-待遇・礼など
甘い	(程度)	辛（から）い	3.5050	相-自然-自然-味

表 7 多義語「甘い」の反対語

	反対語でない		置き換え不可		置き換え可		総計
小段落一致候補	46,746	34.3%	31,716	23.3%	57,738	42.4%	136,200
段落一致候補	61,109	35.9%	46,421	27.3%	62,590	36.8%	170,120
候補外	145,467	83.8%	14,758	8.5%	13,455	7.7%	173,680
総計	253,322	52.8%	92,895	19.4%	133,783	27.9%	480,000

表 8 クラウドソーシングによる評定値の集計結果

あった。小段落番号一致のほうが、置き換え可能性が高い傾向がみられる。100% 反対語であると判断された例は、可決 ⇔ 否決（置き換え可 87.5%）、寒流 ⇔ 暖流（置き換え不可 85%）であった。

表 9 に反対語対候補毎の評定値の分布を示す。割合は 10% 単位に四捨五入した。なお、アノテーション対象である正順呈示・逆順呈示のいずれかが 50% 以上反対語である要素数は、小段落一致候補 2,465 対、段落一致候補 3,008 対、候補外 65 対であった。

クラウドソーシングの評価は呈示順序を正順と逆順の 2 パターン（それぞれ 20 人）評価した。表 10 に、正順と逆順で「反対語でない」の判断が分かれる事例を示す。たとえば「ぷりぷ

小段落一致候補	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
反対語でない	203	785	522	419	315	264	268	282	233	110	4
置き換え可	32	565	1,201	1,138	391	68	10	0	0	0	0
置き換え不可	44	366	414	376	428	569	638	427	130	13	0
段落一致候補	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
反対語でない	225	823	671	554	438	382	354	366	306	127	7
置き換え可	23	553	1,058	1,342	964	274	37	2	0	0	0
置き換え不可	85	572	621	657	662	683	595	309	67	2	0
候補外	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
反対語でない	1	2	3	7	20	26	70	292	1,420	2,195	306
置き換え可	823	2,888	568	57	6	0	0	0	0	0	0
置き換え不可	1,047	2,820	396	48	19	7	3	2	0	0	0

表 9 反対語対候補毎の評定値の分布

比較対象 A	比較対象 B	正順：反対語でない	逆順：反対語でない	差分
ぶりぶりする	にこにこする	70%	10%	60%
官立	私立 (しりつ・わたくしりつ)	20%	75%	55%
水勢	火勢	25%	80%	55%
粗放農業	集約農業	25%	80%	55%
治癒	回復	35%	90%	55%
手を施す	手をつかねる	100%	45%	55%
にぎわしい	寂しい	0%	50%	50%
従兄	従弟	85%	35%	50%
稚拙	老巧	25%	75%	50%
晴れ上がる	ぐずつく	30%	80%	50%

表 10 正順と逆順とで「反対語でない」の判断が分かれる事例

りする」「にこにこする」の正順で呈示した場合には 70% の回答が「反対語でない」であったが、逆順で呈示した場合には 10% にまで小さくなった。ほかの事例として、性差（(淑女, 紳士) のほうが (紳士, 淑女) よりも反対語性がある）、大小（(小鼓, 大鼓) のほうが (大鼓, 小鼓) よりも反対語性がある）などの傾向がみられた。

4.4 分類語彙表番号による分析

本データベースは分類語彙表をもとにして構成されており、分類語彙表の意味情報に基づく検討が可能である。表 11 に分類語彙表の分類番号ごとの集計結果（頻度上位 10 位まで）を示す。頻度最上位の 2.3100 言語活動は「口を開く」などの慣用表現が多く含まれているために頻度が高くなっている。それ以外については、2.1580 増減（変化（閉））・1.1730 方向（視点（閉）、両極（閉）、相補（閉）、典型（開）など）・1.1770 内外（相補（閉））の表現が多く反対語として見られた。

分類番号	分類項目	例	頻度	反対語でない	置き換え不可	置き換え可
2.3100	用-活動-言語-言語活動	ささやく-叫ぶ, 言い出す-言い終わる, 口を開く-黙る	60	30.0%	36.5%	33.5%
2.1580	用-関係-作用-増減・補充	増える-減る, 増加する-減少する, 空ける-埋める	57	14.4%	37.7%	47.9%
1.1730	体-関係-空間-方向・方角	こちら側-あちら側, 南向き-北向き	46	17.4%	21.4%	61.3%
1.1770	体-関係-空間-内外	内-外, 中側-外側, 館内-館外	45	11.7%	25.2%	63.1%
2.1500	用-関係-作用-作用・変化	変わる-定まる, 強まる-弱まる	43	24.5%	37.8%	37.7%
1.2340	体-主体-人物-人物	偉人-凡人, 善人-悪人, 暑がり-寒がり	42	20.7%	25.2%	54.1%
1.2120	体-主体-家族-親・先祖	父-母, パパ-ママ, しゅうと-しゅうとめ	41	33.7%	20.7%	45.7%
2.3790	用-活動-経済-貧富	盛る-衰える, 華やぐ-寂れる	39	29.6%	33.1%	37.2%
2.1527	用-関係-作用-往復	出勤する-退勤する, 上る-下る	38	25.6%	36.1%	38.3%
2.1520	用-関係-作用-進行・過程・経由	来航する-帰航する, 大回りする-近道する	38	28.5%	35.5%	36.1%

表 11 分類語彙表ごとの集計 (頻度上位 10 位まで)

4.5 反対語のタイプによる分析

アノテーションした反対語のタイプによる集計結果を表 12 に示す。頻度で最も多いものは相補であり、次いで変化・程度・両極と閉じた反対語対が多くみられた。表は「反対語でない」の回答が少ない順にならべたが、閉じた対が開いた対よりも「反対語である」と判断されやすい傾向が確認できた。開いた対の中では 2 側面が「反対語である」と判断されやすい傾向がみられた。置き換え可能性は程度 (閉)・相補 (閉)・両極 (閉)・2 側面 (開) などで高い傾向が確認された。一般に、閉じた反対語対のほうが置き換え可能性が高い傾向があるが、変化 (閉) や視点 (閉) などは、格交替を行う可能性が高く置き換え可能性が低くなったと考えられる。また、正順と逆順とでの「反対語でない」の判断の差分は主副 (別) (開) で若干高くなる傾向がみられた。表 10 の結果とあわせて考えると、副から主への方向のほうが、反対語として認定されやすい傾向がうかがえる。

4.6 コーパスの Uni-gram 頻度による分析

次に『現代日本語書き言葉均衡コーパス』(BCCWJ) 上の Uni-gram 頻度との対照比較を行う。国語研短単位 (SUW) に基づく頻度比の対数「SUW 対数頻度比」¹⁰と国語研長単位に基づ

¹⁰ SUW 対数頻度比: 単語対中の短単位に基づく単語 A の頻度 (1 以上) を短単位に基づく単語 B の頻度 (1 以上) で割ったものの自然対数をとる。絶対値をとったもの。LUW 対数頻度比は長単位に基づいて同様の計算を行ったもの。母集団が同じため対数尤度比に相当する。2 単語の頻度が同じ場合は 0、頻度の差があるほど大きな値になる。

タイプ	開閉	頻度	反対語でない	置き換え不可	置き換え可	正順と逆順の差分
程度	閉	803	18.4%	29.0%	52.6%	8.6%
相補	閉	1,686	19.9%	28.0%	52.0%	9.0%
変化	閉	991	20.0%	35.2%	44.8%	9.2%
両極	閉	756	20.1%	28.8%	51.1%	8.6%
慣用	開	44	20.4%	36.8%	42.8%	10.3%
視点	閉	434	22.1%	30.6%	47.2%	9.3%
2 側面	開	72	22.9%	25.5%	51.7%	9.8%
典型 (2 値)	開	452	32.5%	23.2%	44.3%	11.6%
部分全体	開	75	35.8%	26.5%	37.7%	12.0%
終了	開	44	37.1%	33.6%	29.3%	12.4%
主副 (別)	開	29	41.0%	25.3%	33.6%	16.6%
その他開	開	124	44.6%	25.2%	30.2%	14.8%
因果	開	28	47.6%	26.9%	25.5%	14.1%

表 12 反対語のタイプに基づく集計 (「反対語でない」昇順)

く頻度の対数比「LUW 対数頻度比」により, 単語対の 2 単語の頻度に差があるか否かを算出する. 算出された頻度の差が, クラウドソーシングによる「正順と逆順の差分」と相関があるのかを検証する. 対象は「反対語である」が 50%以上であり, 単語対の双方が BCCWJ の短単位・長単位に基づく頻度情報を持つ (頻度 1 以上の) 2,747 対とする. まず「SUW 対数頻度比」と「正順と逆順の差分」の Spearman 相関係数を単語対ごとに評価したところ, 相関係数が 0.051 ($p < 0.01$) と相関がみられなかった. 「LUW 対数頻度比」と「正順と逆順の差分」についても, Spearman の相関係数 0.054 ($p < 0.01$) と同様の結果であった. このことから, 反対語性判断における呈示順序の選好性を表す「正順と逆順の差分」は頻度に基づく差異と関連がないということがわかる. なお, 対数頻度比に方向性を与えた設定でも検討したが, 相関係数 SUW 0.004, LUW 0.003 (いずれも $p < 0.01$) と相関は見られなかった. 表 13 にタイプごとの対照結果を示す. 視点・典型・2 側面などの反対語対は頻度差がない傾向がみられるが, 変化・部分全体・主副 (別)・因果については頻度差がある傾向がみられた.

4.7 コーパス中の生起順序による検討

次に『国語研日本語ウェブコーパス』(NWJC) 上の生起順序との対照比較を行う. 反対語対の 2 要素が 1 文中に共起する用例中において, その生起順序の頻度比の対数¹¹により, 生起順序に差があるか否かを算出する. 算出された頻度の差が, クラウドソーシングによる「正順と逆順の差分」と相関があるのかを検証する. 対象は「反対語である」が 50%以上であり, 単語

¹¹ 1 文中に反対語対の要素 A と B がこの順に出現した頻度を, 要素 B と A の順に出現した頻度で割ったものの自然対数をと, 絶対値をとったもの.

対の双方が正順と逆順ともに NWJC 中 1 文以上共起する 2,725 対とする。対数頻度比と「正順と逆順の差分」の Spearman 相関係数を評価したところ、相関係数が 0.011 ($p < 0.01$) と相関がみられなかった。このことから、反対語性判断における呈示順序の選好性を表す「正順と逆順の差分」はコーパス中の生起順序と関連がないということがわかる。なお、対数頻度比に方向性を与えた設定でも検討したが、相関係数 -0.008 ($p < 0.01$) と相関は見られなかった。表 14 にタイプごとの対象結果を示す。程度・2 側面・終了などの反対語対は頻度差がない傾向がみられるが、両極・視点・典型 (2 値)・慣用・その他開については頻度差のある傾向がみられた。

タイプ	開閉	度数	SUW 対数頻度比	LUW 対数頻度比	正順と逆順の差分
相補	閉	940	0.641	0.643	0.090
両極	閉	431	0.625	0.647	0.086
程度	閉	432	0.615	0.632	0.086
視点	閉	223	0.516	0.500	0.093
変化	閉	338	0.708	0.676	0.092
典型 (2 値)	開	227	0.500	0.496	0.116
部分全体	開	46	0.778	0.862	0.120
2 側面	開	39	0.353	0.359	0.098
慣用	開	5	0.398	0.640	0.103
終了	開	4	0.376	0.527	0.124
主副 (別)	開	14	0.741	0.864	0.166
因果	開	4	0.980	0.867	0.141
その他開	開	44	0.806	0.808	0.148

表 13 反対語タイプとコーパス上の Uni-gram 頻度差

タイプ	開閉	度数	NWJC 対数頻度比	正順と逆順の差分
相補	閉	954	0.308	0.084
両極	閉	427	0.392	0.082
程度	閉	442	0.252	0.078
視点	閉	229	0.377	0.079
変化	閉	340	0.317	0.082
典型 (2 値)	開	220	0.366	0.104
部分全体	開	30	0.303	0.115
2 側面	開	41	0.236	0.100
慣用	開	4	0.698	0.038
終了	開	3	0.237	0.050
主副 (別)	開	8	0.334	0.206
因果	開	2	0.240	0.125
その他開	開	25	0.378	0.110

表 14 反対語のタイプとコーパス上の反対語対の生起順序の頻度差

4.8 単語埋め込みによる分析

本小節では、『国語研日本語ウェブコーパス』(Asahara et al. 2014) から訓練した単語埋め込みモデル NWJC2vec (Asahara 2018) において, 反対語対がどのように表現されるかを検証する。「反対語である」が50%以上であり, 単語対の双方が NWJC2vec に基づく単語埋め込みを有する 2,809 単語対を分析対象とした。NWJC2vec のモデルとして, fastText (Bojanowski et al. 2017) で訓練した 300 次元の skip-gram のものを用いた。

まず, 単語対間の距離を評価した。置き換え可の回答と相関分析(反対語対ごと)を行ったところ, Spearman 相関係数は 0.286 ($p < 0.05$) であった。このことから, ヒトが感じる置き換え可能性と, 単語埋め込み上の類似度とに中程度の相関があることがわかる。表 15 にタイプごとのコサイン類似度の平均とクラウドソーシングによる置き換え可能性の評定値の平均の比較を示す。反対語対の2単語の類似度が高いのは2側面・慣用・典型(2値)など開いた反対語対であった。このことから, 狭義においては反義でなくても, 単語ベクトルで確認できる前後文脈の類似性から, ヒトは反義性を感じるのではないかと考えられる。閉じた対の中では, 視点のコサイン類似度が高かった。表現として格交替が考えられるものの, 単語埋め込みが捉えられる範囲においては前後文脈の類似性が高かった可能性がある。

また, 反対語対をなす各語のベクトルや, 反対語対の差ベクトルについても, 主成分分析・クラスタ分析・tSNE 法による可視化などにより分析を行ったが, 反対語タイプごとの違いについて特段の傾向が見られなかった。今後, 他の単語埋め込み手法などを用いて, さらなる調査を進めたい。

タイプ	開閉	度数	コサイン類似度 (平均)	置き換え可 (平均)
相補	閉	910	0.599	54.6%
両極	閉	425	0.563	55.4%
程度	閉	480	0.544	55.3%
視点	閉	218	0.645	53.3%
変化	閉	398	0.544	49.3%
典型 (2 値)	開	243	0.633	47.2%
部分全体	開	35	0.495	43.4%
2 側面	開	40	0.720	55.1%
慣用	開	6	0.686	44.2%
終了	開	1	0.554	32.5%
主副 (別)	開	13	0.565	35.4%
因果	開	5	0.443	33.5%
その他開	開	35	0.527	33.6%
総計		2809	0.582	52.8%

表 15 反対語対のコサイン類似度とクラウドソーシングによる置き換え可能性

5 おわりに

本稿では『分類語彙表』に対する反対語情報付与について示した。人手により『分類語彙表』より反対語候補を収集したうえで、クラウドソーシングによりヒトが反対語と判定するか否かを収集し、正順呈示・逆順呈示のいずれかで50%以上の方が反対語と判定したものを反対語と定義した。反対語として得られた単語対に対して、村木(1987)の反対語タイプを付与した。開いた反対語対については、新たに「終了」「主副(別)」「因果」を導入し、精緻化を行った。

分析においては、認知言語・言語処理の両面の分析を進めた。まず、クラウドソーシングにおいて順序を変えて呈示した場合に反対語対と認識する割合が異なるものを中心に分析し、主副・性差・大小において非対称性がみられることを確認した。分類語彙表の分類項目による分析においては、増減・方向・内外・変化・往復・進行などの反対語表現が多いほか、言語活動・人物・家族などにも反対語表現が多いことを確認した。反対語のタイプごとの分析においては、閉じた反対語対のほうが、開いた反対語対よりも反対語として認識されやすい傾向を確認した。コーパス類度の分析においては、反対語対としての認識に非対称性がみられるものが類度によるものではないということを確認した。最後に単語埋め込みによる分析においては、置き換え可能性がベクトル空間上のコサイン類似度と中程度の相関があることを確認した。

本データは <https://github.com/masayu-a/WLSP-antonym> より、クリエイティブコモンズライセンス BY-NC-SA 3.0 で公開する。

謝 辞

本研究の一部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「コーパスアノテーションの拡張・統合・自動化に関する基礎研究」、JSPS 科研費 17H00917, 18H05521, 18K18519, 19K00591, 19K00655 の助成を受けています。

参考文献

- Asahara, M. (2018). “NWJC2Vec: Word Embedding Dataset from ‘NINJAL Web Japanese Corpus.’” *Terminology: International Journal of Theoretical and Applied Issues in Specialized Communication*, **24** (2), pp. 7–25.
- Asahara, M., Maekawa, K., Imada, M., Kato, S., and Konishi, H. (2014). “Archiving and Analysing Techniques of the Ultra-large-scale Web-based Corpus Project of NINJAL, Japan.” *Alexandria: The Journal of National and International Library and Information Issues*, **25** (1–2), pp. 129–148.

- Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). “Enriching Word Vectors with Subword Information.” *Transactions of the Association for Computational Linguistics*, **5**, pp. 135–146.
- Cruse, A. (1986). *Lexical Semantics*. Cambridge University Press.
- Cruse, A. (1994). “Prototype Theory and Lexical Relations.” *Rivista di linguistica*, **6**, pp. 167–188.
- Cruse, A. (2004). *Meaning in Language: An Introduction to Semantics and Pragmatics* (2nd edition). Oxford University Press.
- Herrmann, D. J., Chaffin, R., Daniel, M. P., and Wool, R. S. (1986). “The Role of Elements of Relation Definition in Antonym and Synonym Comprehension.” *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie*, **194** (2), pp. 133–153.
- Herrmann, D. J., Chaffin, R. J. S., Conti, G., Peters, D., and Robbins, P. H. (1979). “Comprehension of Antonymy and the Generality of Categorization Models.” *Journal of Experimental Psychology: Human Language and Memory*, **5** (6), pp. 585–597.
- 川島寛乃, 松田寛, 毛利研 (2019). 対義語対の差分ベクトルを使用した評価極性辞書の拡張. 言語処理学会第 25 回年次大会発表論文集, pp. 1335–1338. [H. Kawashima et al. (2019). Taigotui-no Sabunbekutoruwo Shiyoushita Hyoukakyokuseijisho-no Kakuchou].
- 国立国語研究所 (編) (2004). 分類語彙表一増補改訂版一. 大日本図書. [National Institute for Japanese Language and Linguistics (ed.) (2004). Word List by Semantic Principles – Revised and Enlarged Edition.].
- 近藤明日子, 田中牧郎 (2020). 「分類語彙表番号 – UniDic 語彙素番号対応表」の構築. 国立国語研究所論集, **18**, pp. 77–91. [A. Kondo and M. Tanaka (2020). Construction of an Alignment Table between ‘Word List by Semantic Principles’ and UniDic. NINJAL Research Papers, **18**, pp. 77–91.].
- 松本曜 (2007). 反義性に関する認知意味論的考察. 日本言語学会第 135 回大会. [Y. Matsumoto (2007). Hangisei-ni Kansuru Ninchiimirontekikousatsu, Nihongengogakkai Dai 160 Kai Taikai].
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). “Efficient Estimation of Word Representations in Vector Space.” In *International Conference on Learning Representations*.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013b). “Distributed Representations of Words and Phrases and their Compositionality.” In *Advances in Neural Information Processing Systems*, pp. 3111–3119.
- Mikolov, T., Yih, W., and Zweig, G. (2013c). “Linguistic Regularities in Continuous Space Word Representations.” In *Proceedings of the 2013 Conference of the North American Chapter of*

- the Association for Computational Linguistics: Human Language Technologies*, pp. 746–751.
- Mohammad, S. M., Dorr, B. J., Hirst, G., and Turney, P. D. (2013). “Computing Lexical Contrast.” *Computational Linguistics*, **39** (3), pp. 555–590.
- 村木新次郎 (1987). 対義語の輪郭と条件. 日本語学, **6** (6), pp. 29–41. [S. Muraki (1987). Taigigo-no Rinkaku-to Jouken. *Nihongogaku*, **6**(6), pp. 29–41.].
- Murphy, L. (2003). “Antonymy and contrast.” In *Semantic Relations and the Lexicon*, Chap. 5, pp. 169–215. Cambridge University Press.
- 中村拓, 田然, 乾健太郎 (2018). 単語の極性を埋め込んだ分散表現. 言語処理学会第 24 回年次大会発表論文集, pp. 348–351. [T. Nakamura et al. (2018). Tango-no Kyokusei-wo Umekonda Bunsanhyougen, Gengoshorigakkai Dai 24 Kai Happyouronbunshuu, pp. 348–351.].
- 中村拓, 田然, 佐々木翔大, 乾健太郎 (2019). 単語埋め込みにおける複数視点の対義語判定. 人工知能学会第 33 回全国大会論文集, 1N3-J-9-04 号, pp. 1–4. [T. Nakamura et al. (2019). Distinguish Antonyms in Multiple Viewpoints on Distributed Word Representations. The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019, 1N3-J-9-04. pp. 1–4.].
- 荻野綱男, 野口美和子 (1988). 辞書における反対語記述の問題点. ソフトウェア文書のための日本語処理の研究 – 9 – IPAL (Basic Verbs) をめぐって, pp. 119–171. 情報処理振興事業協会. [T. Ogino and M. Noguchi (1988). Jishoniokeru Hantaigokijutsu-no Mondaiten. Sofutoueabunsho-no Tame-no Nihongoshori-no Kenkyu – 9 – IPAL (Basic Verbs) wo-megutte. pp. 119–171.].
- 荻野綱男, 野口美和子 (1996). 反対語意識の構造. 日本語研究: 東京都立大学国語学研究室, **16**, pp. 78–111. [T. Ogino and M. Noguchi (1996). Hantaigoishiki-no Kouzou. *Nihongokenkyuu: Toukyoutoritsudaigaku Kokugogakukenyuushitsu*. **16**, pp. 78–111.].
- Ono, M., Miwa, M., and Sasaki, Y. (2015). “Word Embedding-based Antonym Detection using Thesauri and Distributional Information.” In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 984–989.
- Paradis, C., Willners, C., and Jones, S. (2009). “Good and Bad Opposites: Using Textual and Experimental Techniques to Measure Antonym Canonicity.” *The Mental Lexicon*, **4** (3), pp. 380–429.
- Rothe, S. and Schütze, H. (2015). “AutoExtend: Extending Word Embeddings to Embeddings for Synsets and Lexemes.” In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 1793–1803.

佐藤貴俊, 高村大也, 奥村学 (2016). 分散表現を用いた単語の感情極性抽出. 情報処理学会研究報告 2016-NL-228, 12号, pp. 1-6. [T. Sato et al. (2016). Bunsanhyogen-wo Mochiita Tango-no Kanjokkyokuseichuushutsu. IPSJ SIG Technical Report, 2016-NL-228(12), pp. 1-6].

略歴

加藤 祥：2011年神戸大学人文学研究科博士後期課程修了。2012年より国立国語研究所コーパス開発センタープロジェクトPDフェロー。同プロジェクト非常勤研究員。2020年より目白大学外国語学部専任講師。博士（文学）。日本語学会，日本認知言語学会，日本認知科学会各会員。

浅原 正幸：2003年奈良先端科学技術大学院大学情報科学研究博士後期課程修了。2004年より同大学助教。2012年より国立国語研究所コーパス開発センター特任准教授。2019年より同教授。博士（工学）。言語処理学会，日本語学会，日本語学会各会員。

森山奈々美：2018年津田塾大学学芸学部英文学科卒業。2018～2020年国立国語研究所コーパス開発センター技術補佐員。

荻原亜彩美：2019年津田塾大学学芸学部英文学科卒業。現在，同大学院文学研究科修士課程に在学中。2019年より国立国語研究所コーパス開発センター技術補佐員。

山崎 誠：1984年筑波大学大学院文芸・言語研究科言語学専攻（日本語学）単位取得退学。同年国立国語研究所研究員。現在，同言語変化研究領域教授。博士（学術）。言語処理学会，日本語学会，計量国語学会各会員。

(2020年7月1日 受付)

(2020年9月11日 再受付)

(2020年10月14日 採録)