

国立国語研究所学術情報リポジトリ

<全文>Evidence-based Linguistics Workshop 2024 発表論文集

メタデータ	言語: Japanese 出版者: 国立国語研究所 公開日: 2024-10-11 キーワード (Ja): Evidence-based Linguistics キーワード (En): Evidence-based Linguistics 作成者: 国立国語研究所 メールアドレス: 所属:
URL	https://doi.org/10.15084/0002000329

This work is licensed under a Creative Commons Attribution 4.0 International License.



Evidence-based Linguistics Workshop 2024

発表論文集

2024年9月21日(土)

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 編

目次

- 「構成の反復」における反復単位の認定基準に関する考察
伊藤薫（九州大学） 1
- BE 動詞と HAVE 動詞の統語的特性: 自己ペース読み課題と容認性判断実験による検証
木村一馬（筑波大学：学生） 13
- 多義語の語義間類似度の地域差に見る多義ネットワークの構成：現代日本語形容詞「あまい」
を題材に
西内沙恵（北海道教育大学） 27
- 編集後記
浅原正幸（国立国語研究所） 36

「構成の反復」における 反復単位の認定基準に関する考察

伊藤 薫 (九州大学)[†]

Discussion of an Annotation Guideline for the Repetition Unit in a *Structural Repetition*

Kaoru Ito (Kyushu University)

要旨・既発表の有無

本研究では、「構成の反復」に関する言語資源を構築するために必要な反復単位の認定基準について考察する。構成の反復とは、同じ言語表現のパターンを繰り返す修辞表現である。構成の反復を収集した言語資源を構築するには、同一構成の反復とみなされる表現の範囲をテキスト上にアノテーションする必要がある、そのための基準を定めなければならない。比較的単純な例では範囲を定めることに大きな困難はないが、複雑な例では認定方法によって得られる結果やそこから得られる示唆が異なってくる。そのため、本研究では具体例に対し複数の基準でアノテーションした結果を比較しつつ、各々の利点と欠点について考察する。また、範囲の認定基準によっては、全ての要素に共通する特徴は見られないが、部分的に特徴を共有することで形成されるまとまり、つまり家族的類似性が形成されうることを示す。既発表無。

1. はじめに

「構成の反復」は同じ構成が繰り返されることを特徴とする修辞表現の1つである。本研究では伊藤に引き続き、構成の反復に関するデータセットの構築を目指し、アノテーション方法について議論する。本研究で議論の対象となるのは、構成の反復の範囲を定める方法である。簡単な構成の反復の例としては、イギリスの元首相チャーチルの発言 “I am fond of pigs. Dogs look up to us. Cats look down on us. Pigs treat us as equals.” の一部である、“Dogs look up to us. Cats look down on us.” が挙げられる。ここでは反復する2つの文が同じ語数、同じ品詞の並びで構成されており、共通性が明確である。このような簡単な例を見る限り範囲アノテーションガイドラインの作成は簡単に見えるが、実際には複雑な例もある。

本研究で念頭においている構成の反復とは、佐藤ほか (2006:55) が下位分類なども含め包括的に定義しているもので、「言語表現における同じかたち、もしくはパターンの繰り返し。パターンの基本は構文、すなわち文、節、句におけるかたちだが、これに音声や意味の繰り返しを加味されると、かたちは鮮明なものとなる。」とされている。Farnsworth (2011:74) も isocolon (同形節反復) について “the use of successive sentences, clauses, or phrases similar in length and parallel in structure.” と定義しており、佐藤ほかの定義と類似している。Farnsworth の

[†] ito@flc.kyushu-u.ac.jp

分類では他にも佐藤ほかの言う構成の反復に該当する修辞表現があり、上記研究以外にも関連する修辞表現や分類の方法は多数ある。本研究では前掲のチャーチルの発言などを典型例とし、実際の言語使用では様々な周辺の例が広がっていることを想定する。

構成の反復の範囲を定めることの難しさは、反復の際に生じる変異をどの程度まで容認するかという問題と関係している。つまり、何らかの構成が反復していると認めるには、ある表現と別の表現の構成が同一であるということが前提になるが、何を以て同一の構成とするかという基準は一意に決められるものではない。上述の「周辺の例」は、その同一性の判断の難しさによって生まれる。そこで本研究では、まず複数の範囲認定方針を決め、それらの方針に従って少数の複雑な例をアノテーションし、方針ごとの結果を比較する。その結果、方針によっては全ての反復間に共通する特徴が見られるわけではなく、一部の反復間に特徴が共有され、それらがゆるやかに繋がる家族的類似性 (family resemblance, Wittgenstein 1953) が見られることを示す。

2. 反復単位の認定方法

本節では、反復単位の認定方法を示す。まず、基本的な方針として下記の2つを定める。

(i) 反復単位は2語以上とする

(ii) 語や品詞の線形順序の共通性を認定の基礎とする

(i) は「構成の反復」を別の修辞表現である「語の反復」と区別するために設けた方針である。(ii) は構成の同一性を認定する際に用いる単位についての基準である。何かが反復すると認定するためには、それらの同一性を認定する必要がある。「構成の反復」は修辞学の概念であるため、何を以て「構成」とするかは言語学的な概念によって明示されていない。アノテーションのためには何らかの基準を設けなければいけないが、過度な抽象化は判断を難しくし、かえってアノテーションの労力を上昇させてしまう可能性がある。そのため、本研究では語とその品詞を認定の基礎とする。

語や品詞の並びに一定の共通性を認定した後は、それに付随する要素の扱いが問題となる。各例の分析で細かな差異はあるが、反復単位の認定についての基本的な方法を提案し、それらと比較する。比較対象となる具体的な方針は下記の3種である。

(a) 反復単位はできる限り細分化し、一連の反復を比較したときに反復単位内に部分構造が認められないよう分割する。1語に対し複数の反復単位の割当を許す。

(b) 一連の反復を比較したときに、反復単位内の部分構造が認められる場合でも1つの単位として認定する。1語に対し唯一の反復単位の割当のみを許す。

(c) 一連の反復を比較したときに、反復単位内の部分構造が認められる場合でも1つの単位として認定する。1語に対し複数の反復単位の割当を許し、部分構造の反復もアノテーションする。

(b), (c) の「一連の反復を比較したときに、反復単位内の部分構造が認められる場合でも1つの単位として認定する」とは、たとえば、反復とみなしたい一連の表現が $A_1B_1, A_2B_2, A_3B_3C_1, B_4C_2$ の4つであった場合、 A, B, C それぞれに分割せず、 AB や ABC を1つの反復と認定するということである。

なお、本研究の目的はアノテーションの大方針として、構成の反復を分割する際、上記の方針それぞれの性質を議論することであるため、細かい分割の基準については恣意的な例も含まれうるがご容赦いただきたい。

本研究の中心的な目的はあくまでも構成の反復単位を認定することであり、実際の作業で抽象的な構成をアノテーションすることは想定していない。むしろ、反復単位内の構成は様々なツールを用いた分析が可能だと思われるので、反復の範囲を主なアノテーション対象とするデータセットでは先入観を避けるために付与しない方が望ましいと考えている。しかし、反復単位を認定するためには構成を意識することが不可欠であるため、本研究では便宜上構成の内容も記述する。

基本的な記述方針は以下の通りである。まず、ここでの最小単位は語とし、語および、それらを抽象化することで得られるレンマや品詞を中心として記述する。その際、語の表層形とレンマは小文字、それよりも抽象化された単位は大文字で記述する。略号は付録に示す。語の連なりとして形成される構成は角括弧「[]」で囲むことで表示し、その中に中核となる要素を記述する。記述内容は全ての反復に共通することを条件として、できる限り具体的なレベルで記述する。つまり、表層形が共通していれば表層形で記述し、より抽象的なレベルでしか共通性が見いだせない場合は、レンマ、品詞、句など抽象化したレベルで記述する。複数の反復を比較した際、ある品詞が構成の内部に含まれる場合と含まれない場合があるときは、その品詞に括弧「()」を付けて示す。

3. 実例の分析

本節では、構成の反復の具体例を分析し、反復単位を定める方法を検討する。用例はいずれも、Farnsworth (2011) で isocolon (同型節反復) として引用されているものを使用した。

3.1 具体例 1: Grattan's Speech

1 つ目の分析対象は下記に示す例 (1) である。特に分析の対象としたい文は第 2 文であるが、代名詞 *them* の照応先が含まれるため前方の文脈も含め引用した。第 1 文でも “the foundation of trade,” “the charters of the land,” ... と構成の反復が含まれるが、第 2 文では “the work is imperfect,” “the foundation is wanting,” ... と構成の反復が続く。

- (1) But liberty, the foundation of trade, the charters of the land, the independency of Parliament, the securing, crowning, and the consummation of everything are yet to come. Without them the work is imperfect, the foundation is wanting, the capital is wanting, trade is not free, Ireland is a colony without the benefit of a charter, and you are a provincial synod without the privileges of a Parliament.

(Grattan, 1796, speech in the Irish Parliament)

(1) の第 2 文を本説冒頭で挙げた 3 つの方針で分割し、結果について比較する。各表では、列ごとに構成をまとめ、列の最下部には抽象的な構成を示した。抽象的な構成を示した際の括弧は、反復間に品詞や語数、依存する語について変異があることを示す。まず、表 1 は方針 (a) によりできる限り反復の単位を細分化して認定した場合を示している。反復している構成は 3

種類となり、構成 1 と構成 2 では be 動詞が重複している。これは反復単位が 2 語以上と定められており、この重複を認めなければ一部の例 (構成 1 の Ireland, you や構成 2 の imperfect, wanting) を反復単位として認められなくなり、特に構成 2 が反復として成り立たなくなってしまうためである。

表 2、表 3 は方針 (b) による分割であるが、表 2 では表 3 よりも構成の同一性を認める上での変異の許容度を低くしている。つまり、表 2 では構成 1[(N) is (ADJ)] と構成 2[N be a (N) without the N of a N] という 2 つの構成に分かれるが、表 3 では後半にみられる変異の許容度を高く取り、掲載されている全ての反復が [(N) be COMP] という 1 つの構成にまとめられる。

表 4 は方針 (c) に則り、構成とその部分の分割を示している。[(N) be COMP] という構成は表 3 と共通だが、前半の 4 つは部分構成 1[is (ADJ)] を持ち、後半の 2 つは部分構成 2[a (N) without the N of a N] を持つ。なお、部分構成 1 では be 動詞を含め、部分構成 2 では含めないという差異があるが、これは部分構成 1 では be 動詞も含めなければ 1 語のみの反復となり方針 (i) に反するためである。

以下、方針 (a)~(c) の比較について述べる。まず、方針 (a) による認定では、他の方針に比べ反復間に共通の構成が明確になる一方、認定範囲が狭くなり重複も発生するため、全体像が見えづらくなっている。方針 (b) のうち変異の許容度を低くした表 2 では、方針 (a) より反復の範囲を広く認定するため、全体像は見えやすくなっている。一方、他の方法では捉えることができている 6 つの反復の共通性を捉えることができず、構成 1 を 4 つの反復、構成 2 を 2 つの反復としてそれぞれ別個に記述することになる。方針 (b) のうち、変異の許容度を高くした表 3 では、範囲を広く認定することで構成の反復の全体像や流れが見えやすくなる一方、品詞の範疇に収まらない補部というレベルで抽象的な構成を記述する必要に迫られる。方針 (c) では認定される上位の構成が表 3 と共通しているが、部分構成の情報が追加されることにより、6 つの反復の共通性と、それらが更にグループ化できることを表現できている。ただし、範囲の重複を許すことでデータ構造は複雑になり、実際、補部の内部構造も (b) の場合よりも複雑になっている。なお、(1) を対象とした分析では、方針 (b) のうち変異の許容度を高く取った場合と方針 (c) の違いは上位の構成の認定に関してさしたる違いをもたらさないが、次節で述べる例では影響が大きくなる。このように、範囲の広さと変異の許容度はトレードオフの関係にある。どの方針が望ましいかはデータ構築の目的によって異なってくる。

構成 1	構成 2	構成 3
the work is	is imperfect,	a colony without the benefit of a charter,
the foundation is	is wanting,	a provincial synod without the privileges of a Parliament.
the capital is	is wanting,	
trade is	is not free,	
Ireland is		
you are		
[(N) be]	[is (ADJ)]	[a (N) without the N of a N]

表 1 方針 (a) による (1) の反復単位

構成 1	構成 2
the work is imperfect,	Ireland is a colony without the benefit of a charter,
the foundation is wanting,	you are a provincial synod without the privileges of a Parliament.
the capital is wanting,	
trade is not free,	
[(N) is (ADJ)]	[N be a (N) without the N of a N]

表 2 方針 (b) による (1) の反復単位 (変異の許容度低)

構成 1
the work is imperfect, the foundation is wanting, the capital is wanting, trade is not free, Ireland is a colony without the benefit of a charter, you are a provincial synod without the privileges of a Parliament.
[(N) be COMP]

表 3 方針 (b) による (1) の反復単位 (変異の許容度高)

構成 1	部分構成 1	部分構成 2
the work is imperfect, the foundation is wanting, the capital is wanting, trade is not free, Ireland is a colony without the benefit of a charter, you are a provincial synod without the privileges of a Parliament.	is imperfect is wanting is wanting is not free	a colony without the benefit of a charter, a provincial synod without the privileges of a Parliament.
[(N) be COMP]	[is (ADJ)]	[a (N) without the N of a N]

表 4 方針 (c) による (1) の反復単位

3.2 具体例 2: Sheil's speech

2つ目の分析対象は下記に示す例(2)である。対象としたいのは“— the church of minority...”以降だが、(1)と同様に理解しやすいよう前方の文脈も含めている。また、対象とする一連の反復の中には、反復ではない部分(“but evil ... poured out”)による断絶が含まれる。

- (2) For what into all these affrighting perils are we to rush? For what into those terrific possibilities are we madly, desperately, impiously to plunge? For the Irish church! — the church of the minority, long the church of the state, never the church of the people — the church on which a faction fattens, by which a nation starves — the church from which no imaginable good can flow, but evil after evil in such black and continuous abundance has been for centuries, and is to this day, poured out — the church by which religion has been retarded, morality has been vitiated, atrocity has been engendered; which standing armies are requisite to sustain, which has lost England millions of her treasure, and Ireland torrents of her blood.

(Sheil, 1835, speech in the House of Commons)

表5は方針(a)によって(2)を分割したものである。この場合は構成1[the church PP NOMINAL]、構成2[(PP) which NP VP]、構成3[N has been V-pp]、構成4[N N of her N]という4つの構成に分かれる。範囲の重複を許しているため、構成1と構成2ではPPとその直後のwhichが重複することが多く、構成2と構成4でも重複が見られる。

表6は方針(b)によって(2)を分割したもののだが、全ての反復に共通する構成を抽出するのは難しい。そのため、表では反復ごとに抽象的な構成を示し、行ごとにテキストと抽象的な構成を示した。抽象的な構成を抽出する際には、複数の反復間で共通の構造ができるだけ多くなるように構成を分割しており、その結果一部の反復間には[(ADJ) the church][PP DET N][PP REL][N N of her N]など、部分的な構成の共有が見られる。この点については4節で詳しく考察する。

表7は方針(c)による分割で、構成1は表6と同様に共通する構成を抽出するのが難しいため、最下段に構成を示していない。表7における構成1の分割は表6よりも1つ1つの反復を長く認定しており、“the church”とそれに続く“(PP) which”を核としたゆるやかな共通性を持っている。一部の反復はその内部に構成の反復を持ち、それらを内部反復として第2~4列に記載している。内部反復1は4つ目、内部反復2は6つ目、内部反復3は最後の反復の一部となっている。なお、表4と同様に、複数の構成間に共通する部分反復についてアノテーションすることも可能だが、表が冗長になってしまうため割愛している。

方針(a)~(c)による差異は基本的に1の分析と共通しているが、2では方針(b)と方針(c)による分割で差が生じている。2の分析では、方針(b)によれば反復数は12であったが、方針(c)では内部反復1~3に挙げた部分が併合されており、上位構成の反復数は9に減少している。これにより、反復間で共有されている部分の一致度は、“the church”とそれに続く“(PP) which”が明確である分方針(c)による分割の方が高いが、内部にも構成の反復を含むため変異はより多くなっている。

構成 1	構成 2	構成 3	構成 4
the church of the minority, the church of the state, the church of the people the church on which —the church from which the church by which [the church PP NOMINAL]	by which a nation starves on which a faction fattens, from which no imaginable good can flow, by which religion has been retarded, which standing armies are requisite to sustain, which has lost England millions of her treasure, [(PP) which NP VP]	religion has been retarded, morality has been vitiated, atrocities have been engendered; [N has been V-pp]	England millions of her treasure, Ireland torrents of her blood. [N N of her N]

表 5 方針 (a) による (2) の反復単位

構成 1	抽象的な構成
—the church of the minority, long the church of the state, never the church of the people— the church on which a faction fattens, by which a nation starves — the church from which no imaginable good can flow, the church by which religion has been retarded, morality has been vitiated, atrocities have been engendered; which standing armies are requisite to sustain, which has lost England millions of her treasure, and Ireland torrents of her blood.	[(ADJ) the church][PP DET N] [(ADJ) the church][PP DET N] [(ADJ) the church][PP DET N] [(ADJ) the church][PP REL][DET N V] [PP REL][DET N V] [(ADJ) the church][PP REL][DET N V] [(ADJ) the church][PP REL][N has (V-pp)] [N has (V-pp)] [N has (V-pp)] [REL N be ADJ to V] [REL has (V-pp)][N N of her N] [N N of her N]

表 6 方針 (b) による (2) の反復単位

構成 1	内部反復 1	内部反復 2	内部反復 3
<p>— the church of the minority, long the church of the state, never the church of the people — the church on which a faction fattens, by which a nation starves</p> <p>— the church from which no imaginable good can flow, the church by which religion has been retarded, morality has been vitiated, atrocity has been engendered; which standing armies are requisite to sustain, which has lost England millions of her treasure, and Ire- land torrents of her blood.</p>	<p>on which a faction fattens, by which a nation starves</p>	<p>religion has been retarded, morality has been vitiated, atrocity has been engendered;</p>	<p>England millions of her treasure, Ireland torrents of her blood.</p>
	[PP which a N V]	[N has been V-pp]	[N N of her N]

表 7 方針 (c) による (2) の反復単位

4. 反復間の家族的類似性

3節で観察した事例では、反復の単位認定の際に見られる変異を広く許容しつつアノテーションすると、表6のように全ての反復に共通する構造が見られない場合があった。本節では、表6の分析結果を参照しつつ、これらの反復が緩やかに特徴を共有しつつ形成する緩やかな「家族的類似性」の 카테고리であることや、その形成過程を示す。

まず、それぞれの反復および、その反復が持つ抽象的な構成のを表8に示す。なお、最左列の数字は表8の行数に対応し、その行の「構成1」列の反復について言及していることを示す。各セルの”1”は反復がその構成を持つことを示し、空所持たないことを示す。なお、構成 [N has (V-pp)] および [REL has (V-pp)] についてはスラッシュを用いてその位置に名詞または関係代名詞が入るということを示しつつ、[N/REL has (V-pp)] に統合した。

	[(ADJ) the church]	[PP DET N]	[PP REL]	[DET N V]	[N/REL has (V-pp)]	[REL N be ADJ to V]	[N N of her N]
1	1	1					
2	1	1					
3	1	1					
4		1	1	1			
5			1	1			
6	1		1	1			
7	1		1		1		
8					1		
9					1		
10						1	
11					1		1
12							1

表8 各反復が持つ構成

この表からは、全ての反復に共通する構成は見られないが、いくつかの反復が部分的に構成を共有しつつ、全体として緩やかなまとまりを形成していることが読み取れる。[(ADJ) the church] が反復1~3, 6, 7, [PP DET N] が反復1~4, [PP REL] が反復4~7, [DET N V] が反復4~6といったように、前半の7つはそれらの間全てに共有される構成はないものの、複数の構成を共有している。これらと後半のまとまりをつなぐのが反復7であり、反復7~9, 11に共有されている構成 [N/REL has (V-pp)] を介して前半との関係を形成している。最後の反

復 12 は、[N N of her N] によって構成 11 との共通点を持ち、全体とのつながりを持つ。

なお、反復 10 は表中に示した構成を基準にすると他のどの反復とも構成を共有していないが、直前 3 つの反復と be 動詞の使用が共通していたり、主格関係代名詞 which の使用が反復 7 と共通しているなど、他の反復との間に要素間の共通性が見られる。2 節でも触れた通り、構成の記述は便宜的なものであり、構成の共通性に関する基準を厳しく設定すれば反復 10 のような事例は一連の反復に含めないという立場も可能である。しかし、メタファーコーパスのアノテーションガイドライン MIPVU (Steen et al. 2010) でも採用されている WILDII (When In Doubt, Leave It In) の方略に従って、判断に迷う境界例にもできる限りアノテーションを施し、必要があればデータセット使用者がフィルターにより排除することが望ましいと考える。

方針 (b), (c) を採用した場合、ある構成を基準として、そこに新たな構成の付加やある構成の削除を許すことになる。表 8 の一連の事例は、反復の回数が多くなり付加や削除が繰り返されると、ある段階で全ての反復に共有される要素がなくなり、個々の反復間で共有される特徴のネットワークが形成されることを示している。そして、このネットワーク全体をカテゴリーとして見た場合、家族的類似性にもとづく緩やかなカテゴリーであるとみなすことができる。

5. おわりに

本研究では、構成の反復の範囲認定について事例をもとに検討し、複雑な例では様々な認定方法を取りうることを示した。また、認定方法によっては、付加や削除を繰り返すことで一連の構成の反復が共通の構成を持たず、家族的類似性にもとづく緩やかなカテゴリーを形成することを示した。最適な認定方法はデータセットの目的によって異なるため、今回の分析結果を吟味しつつ選択したいと考えている。

また、本研究は構文の拡張とも関連する。構文文法では、話者の言語知識として長期的に定着した構文のネットワークについて研究されているが、本研究は具体的な構成体 (construct) に対し、それらの共通性を記述する試みであるともいえる。構成の反復のような短期的に形成される局所的なネットワークと、構文ネットワークのような長期的に形成される大局的ネットワークの研究が好循環をもたらすことを期待している。

アノテーションガイドラインの策定では、作例ベースの言語学の分析と異なり、データに出現する全ての事例を網羅的に判断しなければならない。アノテーション作業を進める中でその時点のガイドラインでは対応できない事例に遭遇する可能性が非常に高い。そのため、MAMA (Model-Annotate-Model-Annotate) cycle のようなサイクルを回して改良を重ねていくのが一般的である (Pustejovsky et al. 2017:24)。構成の反復に関するデータ構築はまだガイドラインの作成が始まったばかりであるが、今後も多くの事例を観察しつつ改良を重ねていきたい。

付録 A. 略号一覧

ADJ	形容詞
COMP	補部
DET	決定詞

N	名詞
NP	名詞句
NOMINAL	名詞類
PP	前置詞
REL	関係詞
VP	動詞句
V-pp	過去分詞

謝 辞

本研究は国立国語研究所基幹型プロジェクト「実証的な理論・対照言語学の推進」
・サブプロジェクト「アノテーションデータを用いた実証的計算心理言語学」
によるものです。

また、本研究は JSPS 科研費 23K12164 の助成を受けたものです。

文 献

- 伊藤薫 「「構成の反復」の並行性についての構文文法的記述の試み」 Evidence-based Linguistics Workshop 2023 発表論文集, pp. 39–46.
- 佐藤信夫・松尾大・佐々木健一 (2006). 『レトリック事典』 大修館書店, 東京.
- Ward Farnsworth (2011). *Farnsworth's classical English rhetoric.*: Godine Publisher.
- Ludwig Wittgenstein (1953). *Philosophical investigations.* New York: Macmillan.
- Gerard Steen, Lettie Dorst, J. Herrmann, Anna Kaal, Tina Krennmayr, and Trijntje Pasma (2010). *A method for linguistic metaphor identification: From MIP to MIPVU.* Amsterdam: John Benjamins.
- James Pustejovsky, Harry Bunt, and Annie Zaenen (2017). “Designing Annotation Schemes: From Theory to Model.” Nancy Ide, Christian Chiarcos, Manfred Stede, and Steve Cassidy (Eds.), *Handbook of Linguistic Annotation.* Dordrecht: Springer. pp. 21–72.

BE 動詞・HAVE 動詞の統語的特性 自己ペース読み課題・容認性判断課題による検証

木村 一馬 (筑波大学大学院) †

The Syntactic Properties of BE and HAVE: Insights from Self-Paced Reading and Acceptability Judgement Studies

Kazuma KIMURA (University of Tsukuba)

要旨・既発表の有無

本研究は、BE・HAVE 動詞の統語的特性とその心的実在に関して、文処理実験の観点から検証を行うことを目的としている。理論言語学では、BE・HAVE 動詞は、2つの名詞句を取る際（コピュラ文、所有文）、それらが動詞補部位置に埋め込まれており、片方の名詞句が主節へと移動することで派生されるいわゆる繰り上げ動詞 (raising verb) の一種として分析されてきた。本研究では、この理論的仮定の妥当性を、自己ペース読み課題 (Self-paced Reading, SPR) を通して検証する。自己ペース読み課題では、埋め込み構造を持たない自動詞や他動詞に比べ、BE・HAVE 動詞文がそれぞれ痕跡を持つと考えられる領域での読み時間が優位に伸びるという結果が得られた。また、BE 動詞の統語構造に関して、構造的複雑性 (痕跡の数) が wh 句の抜き出しの容認度に影響することを容認性判断実験 (Acceptability Judgement Task, AJT) を用いて検証した。いずれの実験結果も、BE・HAVE は語彙的な自動詞や他動詞と比べて統語構造が特殊であり、その統語的特性が文理解時に処理負荷をかけていることを示唆している。これは、理論言語学における仮説を支持するものとなっており、心理言語学的アプローチが理論言語学の仮説を検証する妥当な手法であることも示唆する。

(発表無し)

1. はじめに

理論言語学では、BE・HAVE は統語構造上 2つの名詞句 (以下, DP) を埋め込む (1) のような派生をすると分析されてきた。(1a) の BE 文では、どちらか一方の DP が、(2a) では高い位置に生起する所有者 DP が主節へと移動し、(1b)、(2b) のようにそれぞれ BE・HAVE 文を派生する (Myler (2016))。

- (1) a. A picture of the wall was the cause of the riot.
b. [BE [[DP a picture of the wall] [DP the cause of the riot]] (Moro (1997))

† s2320038@u.tsukuba.ac.jp

(2) a. John has a brother.

b. [*HAVE* [[**DP** John]][**DP** a brother]]

しかし、埋め込み構造の内部に着目すると、同じ繰り上げ動詞でも BE・HAVE 文は異なる派生をすると考えられてきた (Belvin and Den Dikken (1997), Freeze (1992), Kayne (1993))。

(3) a. [*BE* [[**DP** a picture of the wall]][**DP** the cause of the riot]]

b. [[**DP** a picture of the wall]_i [*BE* [_{ti} [**DP** the cause of the riot]]]]

(Moro (1997), 一部修正)

(4) a. [*HAVE* [**PP** [**DP** a brother]][**TO** [**DP** John]]]]

b. [[**DP** John]_i [*HAVE* [**PP** [**DP** a brother]][**TO** _{ti}]]]]

(Kayne (1993), 一部修正)

(3) の BE 動詞の派生は、埋め込み節内部で名詞句が並列する構造を持つ一方、(4) の HAVE 動詞は、2 つの名詞句の間に抽象的な前置詞句 PP が介在する構造をとっており、一般的には前置詞の補部位置にある名詞句が主節へと移動することで所有文が派生されると考える⁽¹⁾。埋め込み節からの移動に関して、生成文法で仮定されている移動した要素の元位置を記す痕跡の位置が BE・HAVE 文では異なる⁽²⁾。BE 動詞文は、動詞の補部位置に痕跡を含むのに対し、HAVE は動詞が補部にとる (抽象的な) 前置詞句の補部位置にあることになる。

このような BE・HAVE 文が持つ統語構造や派生の違いが心的実在であれば、それら諸特性が文処理時にも影響し、文の読み時間や容認性判断といった行動反応に反映されると考えられる。本研究では、この問題に対して、(a) 自己ペース読み課題を用いて、BE・HAVE 文が痕跡が含まれると想定されている領域において読み時間が伸びること、(b) 容認性判断実験を用いて、BE 文が特定の統語位置からの wh 句の抜き出しに対して、統語構造の複雑さに沿った容認度の低下を見せること⁽³⁾、この 2 点を検証する。

本論文の構成は、以下の通りである。2 節で BE・HAVE 動詞の埋め込み構造に関する理論的分析と関連する文処理実験を概観し、本研究が扱うリサーチクエスチョンを明らかにする。3、4 節では、BE・HAVE 動詞の痕跡位置と読み時間の関係を検証する自己ペース読み課題の結果を報告し、5 節では、wh 句移動を伴う BE 動詞文の容認度を調査した結果を報告する。6 節では、これらの結果を踏まえ、理論的仮説との整合性や今後の課題について議論する。7 節では本論文のまとめを述べる。

(1) この HAVE 動詞の埋め込みを認可する前置詞句がどのような範疇であるかについて一致した見解はなく、先行研究で様々な議論、提案がなされている。本研究では、この前置詞句がどのような範疇であるかについては、立ち入らないこととする。また、Belvin and Den Dikken (1997) は、have の経験者用法においても同様の派生を仮定しているが、本研究では所有文のみを分析または実験の対象とする。

(2) 本論文では、GB (Government and Binding) 理論の慣習に倣い、移動要素の痕跡を t として表す (Chomsky (1973), Fiengo (1977), Chomsky (1993))。また、痕跡に下付きで記されているインデックスは移動した後の要素と移動元の要素が統語計算において同一指標であることを示す。

(3) 上で仮定している HAVE の統語構造と wh 句の抜き出しに関しては、本研究では検証できなかったため、今後の課題としたい (詳細は、4 節を参照)。

2. 文処理実験を用いた検証

2.1 埋め込み構造と読み時間

前節で概観した BE 文の埋め込み構造に関連し、文処理実験の観点からは Samar and Berent (1991) の行った実験がある。Samar and Berent (1991) は、刺激文に含まれる主語名詞句の有無を問う内容理解課題において、BE 動詞文はその埋め込み構造のために自動詞や受動文よりも主語名詞句の情報に速くアクセスするという表 1 のような実験結果を提示している。表 1 の Intransitive・Transitive 文とコピュラ文のターゲット（刺激文の主語名詞句）へアクセスするまでの反応時間 (RT) を比較すると、痕跡を埋め込みに含む BE 動詞文の方が短いことが分かった (Intransitive・Transitive vs copular; 被験者; Scheffe's $F(1, 22) = 4.64, p < 0.001$, 項目; Scheffe's $F(1, 40) = 29.23, p < 0.0001$)。

表 1 Mean Target Response Times (RTs) for the Four Different Sentence Types

Reference	Mean RT	Copular analog	Mean RT	Difference RT
Uaccusative	706.5	Unaccusative analog	708.9	-2.4
Passive	689.9	Passive analog	743.3	-53.4
Intransitive	725.7	Intransitive analog	656.5	69.2
Transitive	707.1	Transitive analog	627.4	79.7

しかし、Samar and Berent (1991) の行った実験は、HAVE 動詞を射程に入れていないことも相まって、あくまで BE 文が補部に埋め込み構造を持つ間接的な証拠を提示したに過ぎず、本研究の問題設定を踏まえば不十分である。本研究が問題にする、BE 文のみならず HAVE 文も埋め込み補部を持つこと、BE・HAVE 文において、オンライン処理時に文中のどの領域に処理負荷がかかり、反応時間が伸びるのかについては依然として経験的証拠がない状態にある。

2.2 BE 動詞の埋め込み補部からの wh 句移動

ここでは、BE 動詞の埋め込み構造に関連して、その埋め込みに含まれる痕跡が、wh 句の移動の可否に関連することを確認する。以下に示すように、英語において、名詞句を 2 つ認可するコピュラ文は、論理的意味をほとんど変えることなく、主述関係を入れ替えることが可能である。

- (5) a. A picture of the wall was the casue of the riot.
b. The cause of the riot was a picture of the wall.

(Moro (1997))

しかし、それぞれの統語構造に着目すると、Den Dikken (2006) が仮定するように、(5b) の方が複雑な派生をしていると考えられる (埋め込み節は、Small Clause を省略し、便宜上 SC と表記する)。(6a), (6b) は、それぞれ (5a), (5b) に対応する統語構造である。

- (6) a. [TP [DP a picture of the wall]_i [SC t_i [DP the cause of the riot]]]] (canonical)
 b. [TP [DP the cause of the riot]][FP t_i [SC [DP a picture of the wall]t_i]]] (inverse)

(Den Dikken (2006))

(6a) の主述関係が入れ替わる inverse 構造における統語派生には、FP と呼ばれる機能範疇が投射しており、埋め込み節では述部であった名詞句 (the cause of the riot) は、この位置で主語としてのステータスを得ると仮定されている⁽⁴⁾。(6) の構造で重要なのは、BE 動詞は埋め込み構造を持ち、痕跡を必ず含むが、主述関係の認可によって、痕跡の数が異なる、という仮定である。

この構造を踏まえ、(7) のような wh 句の一部抜き出しに関して、ある予測をすることができる。以下に、(6a-b) のそれぞれの構造から、低い位置にある名詞句の中から要素を抜き出した後の構造を示す。なお、wh 句は最終的に CP 領域に移動することを前提としている。

- (7) a. *Which riot do you think that the cause of was a picture of the wall?
 [CP Which riot_i[TP ... [FP [DP the cause of t_i]][SC t_j [DP a picture of the wall]]]]
 b. *Which wall do you think that the cause of the riot was a picture of?
 [CP Which riot_i[TP ... [SC [DP the cause of t_i]][DP a picture of the wall]]]

(6b) の構造で、それぞれ名詞句の中から wh 句 (which wall, which riot) を主節へ移動させると、一般的に両方とも容認されないことが分かっている。しかし、ここまでの構造の違いを仮定すると、この 2 つの移動の容認度の低下度合いは異なると予測する。この違いの所在は、統語部門での移動後、意味部門で再構築が起こることを前提にすると、明確になる。移動後の wh 句が構造上最も近い元位置で解釈を受けるとした場合、(7a) の wh 句は問題なく再構築できるが、(7b) の場合は、主語名詞句が FP と TP に残した痕跡が介在し、埋め込み内部にある痕跡を見つけることができず再構築が不可能となる。一方、canonical 構造は、FP のような投射がないため、痕跡の数が inverse に比べて少ないため、統語計算及び文処理にとって負荷をかけないことが予測される。したがって、wh 句移動を伴う BE 動詞文の容認度に関する予測は、「inverse + 主節主語名詞句からの一部抜き出し (= (7a)) < canonical + 主語名詞句からの一部抜き出し (= (7b))」のようになる。

2.3 リサーチクエスチョン

以上を踏まえ、本研究で掲げるリサーチクエスチョンを 2 点以下に示す。

- (8) a. BE・HAVE 文は埋め込み構造のうち、痕跡を含む領域周辺での読み時間が伸びるのか (BE: 後動詞領域, HAVE: 後目的語領域)。
 b. BE 文のうち、canonical 構造と比較して、inverse 構造からの wh 句の抜き出しは、容認度を低下させるのか。

⁽⁴⁾ 先述の HAVE の埋め込みを認可する前置詞句と同様、ここでは、FP が具体的にどのような機能範疇であるかについての議論は行わない。重要なのは、inverse 構造には、通常の構造に存在しない機能範疇が投射していることであり、構造的に複雑である、という想定である。

まず, (8a) に関して, Moro (1997) や Myler (2016) の分析を踏まえれば, 文処理時に, 埋め込み構造を認可するタイミングで処理時間が伸びると考えられる。文処理時に統語構造を構築することを前提した研究は, 埋め込み (主語名詞句のコピーまたは PRO 要素) を認可する要素がある領域で読み時間が伸びるという結果を報告している (Bever and McElree (1988), McElree and Bever (1989), Yamaguchi and Ohta (2023))。上で議論した BE・HAVE 動詞の埋め込み構造における痕跡位置の違いを踏まえ, BE 動詞と HAVE 動詞では, 埋め込み構造を認可する領域が異なり, それらが文中の各領域における読み時間に反映されるのかを自己ペース読み課題を用いて, 実験 1, 実験 2 で段階的に検証する。

次に, (8b) に関して, BE 動詞文のうち, inverse 構造において痕跡の数が構造的複雑性を生み, wh 句の移動をブロックすると考えるのであれば, inverse 構造自体が wh 句の抜き出しにとって問題であり, 容認性判断に反映されると考えられる。Greco et al. (2020) の考察では, BE 動詞文における wh 句移動の可否は抜き出す要素の元位置がどのような文法的ステータスを持つかに左右されることになるが, 本研究は, wh 句が主節に到達するまでに残す痕跡の数が容認度に影響する仮説を採用する。

3. 実験 1: 自己ペース読み課題

3.1 刺激文・要因設計

刺激文は, (9) に示す通りである。Samar and Berent (1991) に従い, 1 要因 (動詞タイプ) の 4 水準 (BE, HAVE, 受動文, 自動詞) で設定した。動詞領域は各タイプの動詞, 後動詞領域は BE・HAVE 文には名詞句, 受動文には過去分詞, 自動詞には副詞を用いた。

- (9) a. Airline pilots with great experience are rolemodels for aspiring pilots.
 b. Airline pilots with great experience have opportunities for career growth.
 c. Airline pilots with great experience are involved in this company.
 d. Airline pilots with great experience talk friendly in this room.

表 2 に示すように, ターゲット文において, 動詞領域は領域 6, 後動詞領域は領域 7 となるよう主語の単語数はどの水準においても均一になるよう設定した⁽⁵⁾。ターゲット文は, 主語・後動詞領域の語彙項目を変更した 6 項目用意し, 合計で 24 文作成した。

表 2 ターゲット文の領域

Region→	1	2	3	4	5	6	7	8	9	10
BE sentence (=8a))	Airline	pilot	with	great	experience	are	role-models	for	aspiring	pilots.
HAVE sentence (=8b))	Airline	pilot	with	great	experience	have	opportunities	for	career	growth.

⁽⁵⁾ 各単語の音節数などはなるべく統一するよう配慮したが, 多少のばらつきは見られる。しかし, 読み時間に影響する範囲ではないと判断した。

3.2 方法

この実験の参加者は、Amazon Mechanical Turk で募集したアメリカ合衆国に在住の英語母語話者 24 人である。オンライン実験用プラットフォーム PCIbex Farm (Zehr and Schwarz (2018)) 上で、上で挙げた 24 の刺激文に加え、36 のフィラー文を含めた合計 60 文をランダム化した順番で 1 文ずつ単語区切りで読ませ (moving window 法)、各領域の読み時間 (RT) をミリ秒 (ms) 単位で計測した。また、実験参加者には、各刺激文を読んだ後に刺激文の真偽を問う comprehension question を解いてもらった。学習効果を防ぐため、各質問文とその答えは正解のものとは不正解のものとの比率が同じになるよう設定した⁽⁶⁾。

3.3 結果

4 つの条件文の各領域の RT は、図 1 に示す通りである。各動詞タイプの領域ごとの RT の影響を調べるために、線形混合モデル (Linear Mixed Model, LMM) を用いて解析を行なった。解析に際し、受動文をベースラインに設定し、従属変数を着目する領域 (領域 6, 7, 9) と文全体の対数変換した読み時間とし、説明変数を動詞タイプ要因の各水準とした。解析には R (R Core Team (2018)) の lmerTest パッケージを使用した。なお、2000ms 以上及び 100ms 以下の RT は「間違い値」として分析から除外した⁽⁷⁾。

動詞周辺領域である領域 6, 7 の動詞タイプごとの RT の分析結果は表 3 に示す通りである。この領域では、BE 動詞のみが優位に RT を伸ばすことが分かった。また、表 4 に示す通り、刺激文全体の RT は、BE が最も長くなることが分かった ($t = 1.862$, $p < 0.0774$)。さらに、表 5 に示すように、動詞後ろの付加要素が現れる領域 9 では、HAVE 動詞の RT が最も長くなることが分かった ($t = 1.673$, $p > 0.095$)。

表 3 領域 6, 7 における RT の線形混合効果モデルによる分析

	Estimate	Std. Error	t value	P Value
(Intercept)	-0.02018	0.14424	-0.140	0.8897
BE	0.10859	0.04534	2.395	0.0264 *
HAVE	0.02465	0.04520	0.545	0.5916
Intran	0.04370	0.04528	0.965	0.3460

モデル式: $\text{lmer}(\text{RT}_{67} \sim \text{BE} + \text{HAVE} + \text{Intran} + (1 | \text{subj}) + (1 | \text{item}))$

3.4 考察

この予備の実験で明らかになったのは、表 3, 4 のように BE 動詞は特に動詞周辺領域で読み時間が伸びることであり、文全体での RT ($t = 1.826$, $p < 0.0774$) よりも動詞周辺領域での

⁽⁶⁾ 実験 1 で用いた質問文は、理解度を促す目的で設定したため、以下に示す文の真偽を問う単純な Yes/No 疑問文を用いた。

(i) Airline pilots with great experience are role models for aspiring pilots.
→ Do airline pilots have great experience? (Yes)

⁽⁷⁾ これらの値は、認知処理を適切に反映しているが「外れ値」として捉えずに、実験を行う際のエラー (e.g., 集中力低下や誤って連続でスペースキーを押してしまった、等) によるものだと考える。また、データポイント除外の基準は、先行研究が採用するものに従った (新井・Roland (2016))。

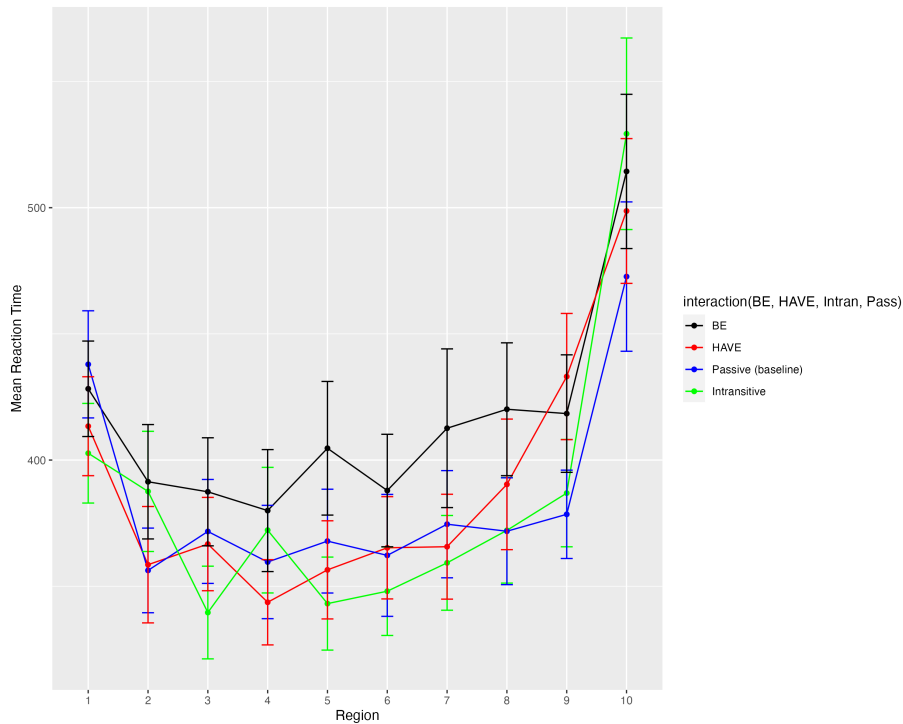


図 1 動詞タイプ別の各領域の読み時間（エラーバーは標準誤差）

表 4 刺激文全体の RT の線形混合効果モデルによる分析

	Estimate	Std. Error	t value	P Value
(Intercept)	0.084999	0.125717	0.676	0.5043
BE	0.065051	0.034943	1.862	0.0774
HAVE	0.009972	0.034905	0.286	0.7781
Intran	0.008141	0.034936	0.233	0.8181

モデル式: $\text{lmer}(\text{RT} \sim \text{BE} + \text{HAVE} + \text{Intran} + (1 | \text{subj}) + (1 | \text{item}))$

RT ($t = 2.395, p < 0.0264$) の方が優位であったことを踏まえると、やはり動詞領域における埋め込み構造を認可する特性がこの領域の RT に強く影響していると考えられる。これは先行研究の理論的仮定や Samar and Berent (1991) の実験結果と概ね合致すると考えられる。BE 動詞は繰り上げ動詞であり、動詞の補部位置に名詞句を 2 つ埋め込む構造をしており、BE 動詞はいわゆる虚辞要素として現れると考えることができる。

一方、HAVE 動詞に関しては、表 5 に示したように、10 パーセント水準で、目的語よりも後ろの領域 9 で RT が優位に伸びている ($t = 1.673, p < 0.095$)。これは、Freeze (1992) や Kayne (1993) の理論的な予測と概ね合致しており、HAVE は同じ繰り上げ動詞でも痕跡の位置が BE 動詞とは異なっており、目的語よりも低い位置にある主語名詞句の痕跡がこの領域の RT の伸長に影響している可能性が示唆される。

この実験では、痕跡の位置に対応する領域の RT が動詞ごとに異なるという結果が得られたが、

表 5 領域 9 における RT の線形混合効果モデルによる分析

	Estimate	Std. Error	t value	P value
(Intercept)	0.12855	0.14998	0.857	0.399
BE	0.06765	0.04839	1.398	0.163
HAVE	0.08020	0.04793	1.673	0.095
Intran	-0.01680	0.04837	-0.347	0.729

モデル式: $\text{lmer}(\text{RT9} \sim \text{BE} + \text{HAVE} + \text{Intran} + (1 | \text{subj}) + (1 | \text{item}))$

不足している点として、動詞の意味役割や項構造構築の影響が考慮されていないことが挙げられる。また、痕跡の位置が異なることによる派生的な効果を検証する必要がある。

4. 実験 2: 自己ペース読み課題

そこで実験 2 においては、実験 1 で検証しなかった動詞タイプである他動詞を要因に組み入れ、意味役割付与や項構造の構築が RT に関与するのかを検証するとともに、Samar and Berent (1991) に従い、BE・HAVE 動詞の痕跡位置をより緻密に特定するために、comprehension question における主語名詞句へのアクセス時間を動詞タイプごとに計測する。

4.1 刺激文・要因設計

実験文は、前節の実験 1 と同様、動詞タイプを要因としたが、本実験では、他動詞を追加し、合計で 5 タイプの動詞を用いた。

- (10) a. College students in small dorms are participants in study groups. (BE)
- b. College students in small dorms have access to study resources. (HAVE)
- c. College students in small dorms are provided with study resources. (Passive)
- d. College students in small dorms may gather for study sessions. (Intransitive)
- e. College students in small dorms obtain access to study resource. (Transitive)

4.2 方法

実験参加者は、Amazon Mechanical Turk で募集した英語母語話者 26 人であり、実験は先ほどと同様 PClbex Farm 上で実施した。本実験では、自己ペース読み課題に加え、comprehension question による主語名詞句へのアクセスにかかる反応時間を計測した。被験者は、自己ペース読み課題を行なった後、文の真偽を問う質問に回答する。正解と判断した場合は F キーを、不正解と判断した場合は J キーを押し、質問が提示されてからキーを押すまでの時間をミリ秒単位で計測した。今回は、主語名詞句へのアクセス時間を計測することから、質問文は、(11) に示すように、主語名詞句を焦点とする分裂文を用いた。質問に対する回答が不正解のものは、主語名詞句の修飾要素及び主要部が刺激文と異なるものに置き換わっている。実験 1 と同様、学習効果を防ぐため、回答は正解と不正解の比率が等しくなるように設定されている。

- (11) a. Airline pilots with great experience are rolemodels for aspiring pilots.
 → It was **airline pilots with great experience** that are role models for aspiring pilots. (Yes)
- b. Active adults in good health have memberships in fitness clubs.
 → It was **inacitive adults in good health** that have membership in fitness clubs. (No)

正解・不正解いずれの場合も、分裂文中の主語名詞句が直前に読んだ刺激文に含まれていたものと同一であるかどうかを判断する際に、主語名詞句にアクセスすると仮定し、埋め込みを動詞補部位置に持つ動詞はその主語名詞句の痕跡をヒントに、埋め込みを持たない動詞よりも回答時間が短くなると予測する。

4.3 結果

計測したターゲット（主語名詞句）への動詞タイプ別のアクセス時間を図2に示す。また、一般化線形混合モデル (Generalized Linear Mixed Model, GLMM) を用いて分析した。従属変数に生の反応時間 (ミリ秒), 説明変数に各動詞タイプを用い, リンク関数に恒等関数を, 確率分布に逆ガウス分布を用いた⁽⁸⁾。解析に対し, 150000 ミリ秒以上のデータポイントは, 実験1と同様「間違い値」として除外した。解析は, 実験1, 2と同様 R Studio (R Core Team (2018)) を用い, lme4 パッケージの glmer 関数を用いた (モデル式は, 表6下部に記載)。その結果を表6に示す。刺激文が BE 動詞文だった場合の反応時間は, 優位に他の動詞タイプよりも短いという Samar and Berent (1991) の実験結果と概ね同様の結果が得られた ($t = -2.70, p < 0.0069$)。一方, 有意差こそなかったが, 図2から明らかであるように, 刺激文が HAVE 動詞だった場合の反応時間は他の動詞タイプと比べても長くなることが分かった ($t = 1.37, p < 0.1717$)。これは, 本研究の予測とは合致せず, 動詞の痕跡を目的語よりも後ろの位置に含むと考えられていた BE 動詞がむしろ痕跡を含まない自動詞や他動詞といった動詞タイプと比べて短くなっているという結果が得られた。また, 痕跡を BE 動詞よりも低い位置に含むとされていた HAVE 動詞のアクセス時間が最も長いという結果になった。

表6 一般化線形混合モデルによるターゲット（主語名詞句）へアクセスするまでの反応時間の分析

	Estimate	Std. Error	t value	Pr(> z)
(Intercept)	38.29	12.19	3.14	0.00168 **
BE	-6.99	2.59	-2.70	0.00697 **
HAVE	14.99	10.97	1.37	0.17174
Tran	4.48	5.72	0.78	0.43357

⁽⁸⁾ ここで, 一般化線形混合モデルを用いているのは, ターゲットへのアクセス時間を対数変換するのを避ける目的のためである。読み時間や反応時間を対数変換することは, 各要因の影響を調べる際に適切ではないという見方がある中で, 今回の実験に関しては, 各動詞タイプのターゲットへのアクセスという認知処理への影響を考慮したため, 反応時間を対数変換することは望ましくないと判断し, 線形混合モデルは用いなかった。

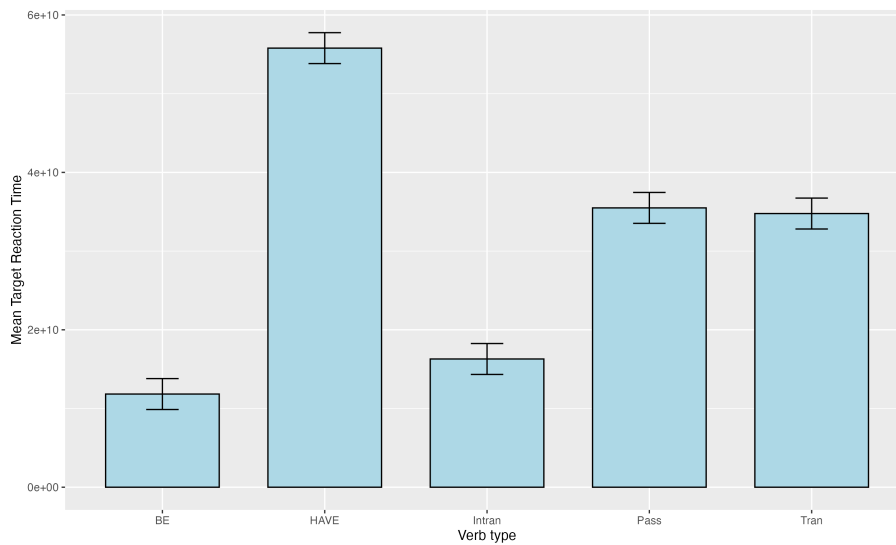


図2 動詞別のターゲットへのアクセス時間 (エラーバーは標準誤差)

4.4 考察

この結果を、理論的に解釈するために、以下に BE 動詞と HAVE 動詞の想定される統語構造を再掲する。

- (12) a. [[DP a picture of the wall]_i [BE [ti [DP the cause of the riot]]]]
 b. [[DP John]_i [HAVE [PP [DP a brother]][TO ti]]]

実験結果を踏まえると、HAVE 動詞の痕跡は、確かに動詞よりも低い前置詞句に含まれている。しかし、主語名詞句アクセスする際に、その痕跡位置から見て、再構築可能な名詞句は 2 つ ([DP John], [DP a brother]) 存在することになる。さらに、その痕跡に取って、処理負荷がかからないのは、一番近い、いわゆる局所的な統語位置に生起する名詞句であるが、その名詞句を痕跡位置に再構築してしまうと、[DP a brother] が痕跡位置に存在するという誤った意味構造を作り出してしまう。そのため、処理負荷のかかる非局所的な主語名詞句である [DP John] を再構築する必要がある。本実験で観測された反応時間の伸びは、この非局所的な名詞句の再構築にかかる処理負荷を反映していたのではないかと考えられる。この反応時間の伸びは、痕跡が前置詞句の補部位置にあることで生じると考えられるため、(12b) の構造を支持していると結論づける。

5. 実験 3

この実験では、2.3 節で議論した BE 動詞文統語構造の複雑性と埋め込み構造からの wh 句の抜き出しに関して、inverse 構造が canonical 構造よりも wh 句の抜き出しに対して重い制約の違反をすることを、容認度判断実験を通してを検証する。この実験によって、BE 動詞の埋め込み構造は、inverse 構造の時にのみ wh 句の移動が動詞補部に痕跡が存在することによって

ロックされるという理論的仮定に対し経験的な証拠を与えることが期待される。

5.1 刺激文・要因設計

刺激文は, Greco et al. (2020) に倣い, 2×2 要因の wh 句の移動を伴う BE 動詞文を用いた。2.2 節での議論を踏まえ, BE 動詞の構造 (inverse vs. canonical) と wh 句の抜き出し位置 (主語 vs 述部) を要因に設定し, これらが文の容認度にどの程度影響するかを検証する。

(13) a. 主語抜き出し + canonical

Which wall do you think that a picture of was the cause of the riot?

b. 述部抜き出し + canonical

Which riot do you think that a picture of the wall was the cause of?

c. 主語抜き出し + inverse

Which riot do you think that the cause of was the picture of the wall?

d. 述部抜き出し + inverse

Which wall do you think that the cause of the riot was a picture of?

ターゲット文は, 合計で 16 文で, 48 のフィラー文を含む, 合計で 64 文を実験文とした。

5.2 方法

実験参加者は, Amazon Mechanical Turk で募集した英語母語話者 20 人で, 実験は実験 1, 2 と同様 PCIbex Farm 上で行なった。実験参加者は, 刺激文を 1 文ずつ提示され, 各刺激文の容認度を 7 段階のリッカート尺度で評価する⁽⁹⁾。

5.3 結果

各要因ごとの対数化した容認度は, 図 2 に示す通りである。また, 線形混合効果モデルを用いて, 各要因の容認度に与える影響を分析した。対数変換した容認度を従属変数, 各要因 (埋め込み構造, 主語抜き出し) を説明変数とした。解析の結果, inverse 構造が優位に容認度を低下させることが分かった ($t = -1.747, p < 0.0817$)。また, 比較対象として, wh 句移動を伴う HAVE 動詞を BE 動詞と並行的にその構造特性に応じた要因で容認性判断を行ったが, いずれの要因でも有意差はなかった (SubjExtract: $t = 0.029, p > 0.977$, Inverse; $t = -0.015, p > 0.988$)⁽¹⁰⁾。

5.4 考察

上で示した実験結果から読み取れることとして, 表 7 において, 主語抜き出しは優位な影響を与えるわけではない, ということに注目されたい。あくまで inverse 構造になることそれ自体が wh 句の (一部) 抜き出しにとって問題であり, これは, 抜き出す要素の文法的なステータ

⁽⁹⁾ なお, 実験の段階で刺激文の「文法性」を問う指示はしなかった。実験の指示の際には, 「容認性」という語を用いた。これは, 「文法性」という概念が一般に定着していないことで, 実験参加者間の判断基準の揺れが起こることを防ぐためである。

⁽¹⁰⁾ Belvin and Den Dikken (1997) を参考に, HAVE 動詞にも inverse 構造を許すと仮定されている経験者用法, canonical 構造であると仮定される使役用法を実験に用いた。

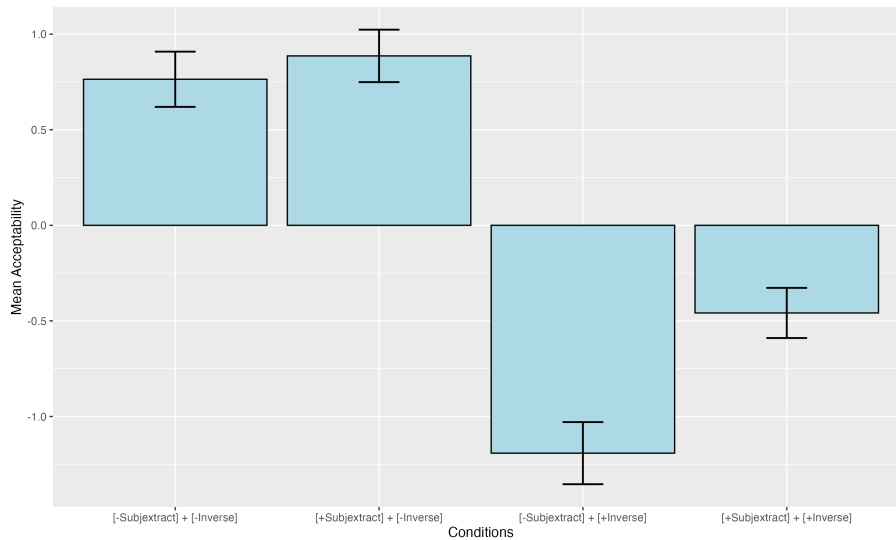


図3 wh 句の移動を伴う BE 動詞文における各条件ごとの容認度 (エラーバーは標準誤差)

表7 wh 句の移動を伴う BE 動詞文の容認度の線形混合効果モデルによる分析

	Estimate	Std. Error	t value	P Value
(Intercept)	-0.05995	0.17900	-0.335	0.7413
SubjExtract	0.01751	0.03867	0.453	0.6509
Inverse	-0.06755	0.03867	-1.747	0.0817.
SubjExtract:Inverse	0.01253	0.03873	0.324	0.7465

モデル式: $\text{lmer}(\text{Value} \sim \text{SubjExtract} + \text{Inverse} + \text{SubjExtract} * \text{Inverse} + (1 | \text{subject}) + (1 | \text{item}))$

スによって違反する統語論上の制約が異なるために容認度が変化するという想定からは予測できない (Greco et al. (2020)) ため、少なくとも抜き出す位置に関する統語制約が容認度に大きく影響するわけではないと暫定的に結論づけられる。

ただ、この実験結果から分かったのは、あくまで抜き出す位置が容認度には影響しないということであって、inverse 構造のどのような特性が容認度に影響するのかわからない。そのため、ここでは、本研究の理論的仮定に従い、inverse 構造が canonical 構造よりも多く痕跡を含むために、2.3 節で議論した再構築効果の阻害が起こり、容認されなくなると考える。この仮定をより直接的に検証するには、移動先と移動元の構造関係が読み時間に影響するのかわかなどを自己ペース読み課題を用いる必要があるだろう。しかし、このような再構築効果は、意味部門で起こると考えられるため、自己ペース読み課題などのオンライン処理時に処理負荷として反映されるかどうかについては慎重に検討しなければならない。

6. 結語

本研究では、BE 動詞または HAVE 動詞の統語的特性が文処理時に反映されるのかについて、自己ペース読み課題と容認性判断実験を用いて検証を行った。以下は、実験から得られた結果

とそれに対応する理論的な考察である。

(14) 実験結果から示唆されること

- a. BE 動詞・HAVE 動詞は、それぞれ動詞後領域 (領域 6, 7), 目的語後ろの領域 (領域 9) での読み時間が優位に伸びた。
- b. HAVE 動詞は、主語名詞句へのアクセスするための反応時間が他の動詞と比べて、著しく長い。
- c. BE 動詞は、その埋め込み構造が複雑である文でのみ、wh 句の一部抜き出しを許さない。

(15) 理論的考察

- a. BE 動詞は、動詞後ろの領域に、HAVE 動詞は、目的語よりも後ろの領域に、それぞれ主語名詞句の痕跡を持つ。
- b. HAVE 動詞の主語名詞句の痕跡は、動作先の要素が非局所的な位置にあり、別の名詞句が介在するため処理負荷をかける。
- c. BE 動詞の inverse 構造は、canonical 構造よりも痕跡を 1 つ多く残すことで、wh 句の再構築を阻害する可能性がある。

本研究の実施した実験により、BE 動詞・HAVE 動詞の構造的特性が処理負荷として文処理時にも反映されることから、理論的に仮定されてきた埋め込み構造またはその内部構造は一定の妥当性を持つものであるということが明らかになった。しかし、上記の実験はいずれも、依然として予備実験の域を出ないため、要因設計やデータの扱い等に関して、先行研究の手法や理論研究の問題意識をもとに、より精査していく必要がある。

謝 辞

この研究を進めるにあたり、実験の企画から実施、分析に至るまで非常に有益な助言をくださった小野雄一氏には感謝を申し上げる。また、本研究の理論的分析に関して数々の有益なコメントをくださった山村崇斗氏・橋本龍弥氏に感謝申し上げます。

文 献

- Neil Myler (2016). *Building and Interpreting Possession Sentences.*: MIT Press.
- Andrea Moro (1997). *The raising of predicates: Predicative noun phrases and the theory of clause structure.* Vol. 80.: Cambridge University Press.
- R Belvin, and M Den Dikken (1997). “There, happens, to, be, have.” *Lingua*, 101:3-4, pp. 151–183.
- Ray Freeze (1992). “Existentials and other locatives.” *Language*, pp. 553–595.
- Richard S Kayne (1993). “Toward a modular theory of auxiliary selection.” *Studia linguistica*, 47:1, pp. 3–31.

- Noam Chomsky (1973). “Conditions on transformations.” *A Festschrift for Morris Halle/Holt, Reinhart and Winston*.
- Robert Fiengo (1977). “On trace theory.” *Linguistic Inquiry*, 8:1, pp. 35–61.
- Noam Chomsky (1993). *Lectures on government and binding: The Pisa lectures*:9: Walter de Gruyter.
- Vincent J. Samar, and Gerald P. Berent (1991). “Be is a raising verb: Psycholinguistic evidence.” *Journal of Psycholinguistic Research*, 20:5, pp. 419–443.
- Marcel Den Dikken (2006). *Relators and linkers*. Vol. 10.: MIT Press.
- Thomas G Bever, and Brian McElree (1988). “Empty categories access their antecedents during comprehension.” *Linguistic Inquiry*, 19:1, pp. 35–43.
- Brian McElree, and Thomas G Bever (1989). “The psychological reality of linguistically defined gaps.” *Journal of Psycholinguistic Research*, 18, pp. 21–35.
- Koki Yamaguchi, and Shinri Ohta (2023). “Dissociating the processing of empty categories in raising and control sentences: A self-paced reading study in Japanese.” *Frontiers in Language Sciences*, 2, p. 1138749.
- Matteo Greco, Paolo Lorusso, Cristiano Chesi, and Andrea Moro (2020). “Asymmetries in nominal copular sentences: Psycholinguistic evidence in favor of the raising analysis.”
- Jeremy Zehr, and Florian Schwarz (2018). *PennController for internet based experiments (IBEX)*.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing Vienna, Austria.
- 新井学・Douglas Roland (2016). 「言語理解研究における眼球運動データ及び読み時間データの統計分析」 *統計数理* = *Proceedings of the Institute of Statistical Mathematics*, 64:2, pp. 201–231.

多義語の語義間類似度の地域差に見る多義ネットワークの構成：現代日本語形容詞「あまい」を題材に

西内 沙恵（北海道教育大学）[†]

The construction of polysemous networks in regional variations in the degree of similarity of polysemous words: the case of the contemporaneous Japanese adjective *amai*

Sae NISHIUCHI (Hokkaido University of Education)

要旨・既発表の有無

本発表では現代日本語の多義語の語義間類似度を地域別に調べた結果を報告し、多義ネットワーク構成の地域的な差異について検討する。従来、個人間の差異の抽象(国広 1982)に基づき、大多数の人々の共有部分に対する分析が作業原則とされ、共通語を対象に調査・分析がなされることが多かった。しかし、多義語の運用には地域的な差異があることが方言辞典などで報告されている。本発表で報告する調査では現代日本語形容詞「あまい」が「関西のみそしるはあまくてどうも口に合わない。」(飛田・浅田 1991: 31-32)といった例文において〈塩分がたりない様子〉を表す際、「うすい」、「みずくさい」などの異なる形式で代替する地域があること(石垣 1983)に着目し、題材とした。Yahoo! クラウドソーシングで日本の 11 地域ごとに「あまい」の例文に対する語義間類似度を調査した結果に基づき、各地域で構成される多義ネットワークの差異を検討する。

本発表は既発表無しの内容である。

1. 背景と目的

多義語は単一の形式に関連する複数の意味が結びついた語と定義され(国広 1982), (1) を主要な課題に研究の蓄積がある。

- (1) a. 何らかの程度の自立性を有する複数の意味の認定
- b. プロトタイプの意味の認定
- c. 複数の意味の相互関係の明示
- d. 複数の意味すべてを統括するモデル・枠組みの解明

(靱山 2019: 34)

従来、大多数の人々が共有する部分を分析する個人間の差異の抽象(国広 1982)を作業原則に、異なる言語の話者同士が意思疎通のために用いる共通語を対象として、調査・分

[†] nishiuchi.sae@a.hokkyodai.ac.jp

析が行われることが多かった。本発表では多義語に対する言語知識の地域的な差異を調べることを目的に、地域別の語義間類似度の調査結果を報告する。対象とする「あまい」は(2)のように複数の語義が記述される(飛田・浅田 1991: 31-32)が、石垣(1983)によれば(2b)の語義〈塩分がたりない様子〉は関東・中部・近畿・四国では「うすい」、
「みずくさい」など異なる形式で使用されるという。

- (2)
- a. 夜あまいものを食べると虫歯になりますよ。(砂糖のような味がする様子)
 - b. 関西のみそしるはあまくてどうも口にあわない。(塩分がたりない様子)
 - c. あまい雰囲気酔う。(芳醇で快い様子)
 - d. あの先生は採点があまい。(厳格さがたりない様子)
 - e. 三番目のねじがあまくなっていた。(程度が低く、不十分である様子)
- (飛田・浅田 1991: 31-32 下線は発表者による。)

〈塩分がたりない様子〉を「あまい」という形式で表さない地域において、話者が持つ「あまい」の言語知識は共通語と異なる構造を持っている可能性がある。本発表では「あまい」の5区分の語義ごとに3種の例文を用意し、2文ずつ組み合わせ、類似度を評定していただくアンケートを11地域分類で実施した調査に関して報告を行う。

2. 先行研究

個人間の差異の抽象(国広 1982)は意味の抽象化に関して三種類あげられるうちのひとつである¹。図1はAからHが個人を表し、L、M及びWからZまでが個人が同一の単語について持つ意味の範囲を表している。全員の共通部分X-Yを抽出すると、最も言語経験・学習活動の貧困な個人が代表されることになる一方、言語経験の他、専門知識や特殊グループでの用法を含むL-Mを意味範囲とすると、全員の言語知識に欠陥があることになるとしている。いずれも実情にそぐわないため、仮想的な共通部分W-Zを大多数の人々が共有する部分と仮定して、分析対象とする方針を示している。

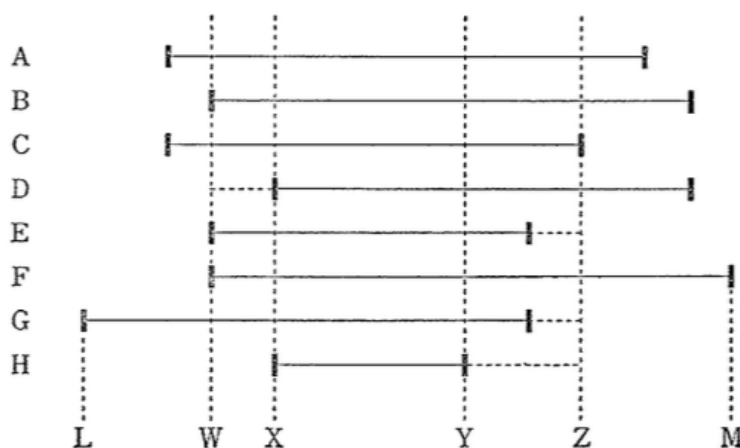


図1 個人が同一の単語について持つ意味の範囲(国広 1982: 40)

¹ 個人間の差異の他、文脈の影響及び指示物との相対的關係が含まれている。

共通語を対象に意味分析が推進されてきた背景には、個人間の差異の抽象の作業原則があったためと考えられる。個人間の差異の抽象を分析の方針とすることで、言語話者が共通基盤として有する言語知識の解明につながる。本発表で取り組む、地域的な差異を明らかにする試みもまた、地域差の範囲が特定される点で、言語知識の共通基盤の解明に寄与できると考える。地域差は表面的な違いと考える立場もあるが、言語知識が経験と相互作用しながら動的に形成される（Johnson 2005）と考える立場から、共通語と異なる多義語の知識の構成を探究する試みは、言語知識の動的なプロセスのあり方に関する（1d）複数の意味すべてを統括するモデル・枠組みの課題として位置付けられる。

3. 調査

多義語の語義間の関係を検討するために、例文に対する類似度評定を調べた。調査の概要は次の通りである。「あまい」が有する〈塩分がたりない様子〉の用法に地域的な差異が認められるかを調べるために、飛田・浅田（1991）による5区分の語義ごとに3種の異なる主語を取る例文のデータセットを作成し、例文間の類似度を調べるアンケートを実施した。調査にご協力くださる参加者をYahoo! クラウドソーシングを通して11地域ごとに300名ずつ募集し、調査協力者に外部サイトで類似度を評定していただいた。

3.1. 調査の材料

調査の材料は（2）の語義5区分×人物・場所・時間の第一主語3種及び異なる第二主語＋〈塩分がたりない様子〉のみ第二主語に「の塩加減」を付加した3種による18刺激文（表1）を異なる語義別に組み合わせた270対である。

表1 刺激文の組み合わせ

意味	第一主語	第二主語	例文
塩分がたりない様子	この店主は	おつゆ	この店主はおつゆ（の塩加減）があまい
	関西は	煮物	関西は煮物（の塩加減）があまい
	今日は	みそ汁	今日はみそ汁（の塩加減）があまい
砂糖のような味がする様子	この人は	卵焼き	この人は卵焼きがあまい
	この店は	ケーキ	この店はケーキがあまい
	今年は	みかん	今年はみかんがあまい
芳醇で快い様子	この人は	声	この人は声があまい
	この店は	香り	この店は香りがあまい
	付き合いだしてから	雰囲気	付き合いだしてから雰囲気があまい
厳格さがたりない様子	この先生は	採点	この先生は採点があまい
	この会社は	判断	この会社は判断があまい
	最近では	親	最近では親があまい
程度が低く、不十分である様子	この新人は	ねじ	この新人はねじがあまい
	この店は	焼き	この店は焼きがあまい
	今は	ピント	今はピントがあまい

刺激文を二重主語文にしたのは、例文に対する解釈がゆれないようにするためである。〈塩分がたりない様子〉を「あまい」という形式で表さない話者は、第二主語が飲食物のとき〈塩分がたりない様子〉ではなく〈砂糖のような味がする様子〉と解釈し、〈塩分がたりない様子〉の例文と〈砂糖のような味がする様子〉の例文の類似度を高く評定することが予想される。ただし、類似度が高く評定されるとき、〈塩分がたりない様子〉と解釈しているが、〈塩分がたりない様子〉と〈砂糖のような味がする様子〉が近い派生関係にあるために高く評定されたとも考えられる。そこで、〈塩分がたりない様子〉と解釈しているかを検証するために、第二主語に「の塩加減が」を付加し、〈砂糖のような味がする様子〉と解釈しにくい例文も刺激文に含めることとした。調査結果から〈塩分がたりない様子〉を表す例文と〈砂糖のような味がする様子〉を表す例文に対する類似度に地域差が認められたとき、地域ごとに異なる多義構造が形成されている可能性が示唆される。

3.2. 調査の設計

調査の材料に対する類似度評定を調べた手続きは次の通りである。2024年7月にYahoo!クラウドソーシングで年齢を20歳以上に、また地域を11地域ごとに指定し、各地域300名ずつ参加者を募集した。参加者は個人のパソコンでYahoo!クラウドソーシング上の図2のような募集画面から参加し、図3の調査協力の同意確認を兼ねた、「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に回答する。「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。

【20歳以上関東エリア限定】2文が似ているか判断してください。〈あまい〉

終了

報酬獲得の条件：あり [レポート](#)

応募条件

**20歳以上限定
関東エリア限定**

【東京都|神奈川県|埼玉県|千葉県|茨城県|栃木県|群馬県|山梨県】

タスクの種類

[アンケート](#)

報酬

5

タスクにかかるめやす時間

5分00秒

残り件数

0件
(全300件)
タスク実施中のユーザー：0人

100%完了

実施できる件数の上限 **1件**

制限時間 [レポート](#) **30分00秒**

残り日数 **終了**
(2024/07/03 10:25)

オーナー **非公開**

このタスクって、どんな内容？

日本語に関する簡単なアンケートです。
20歳以上の方にのみ回答をお願いいたします。

この調査では、日本語の文を二つ読み、どのくらい似ていると感じるか答えていただきます。
リンク先で提示された二つの文が全く似ていないと感じる場合は「0」を、非常に似ていると感じる場合は「6」を、中間の場合は引っかかりの程度に応じて「1」～「5」を選択してください。

回答に要する時間は5分程度です。

注意事項

※パソコンでご参加くださいようお願いいたします。
スマートフォンでは操作できないことがあります。

※文がうまく表示されない、動作が遅いなど、問題がありましたら、一度ブラウザを閉じて、アンケート実施者（西内 nishiuchi-s[a]kanda.kuis.ac.jp）まで[a]を@に変えてご連絡ください。

※このタスクは日本語母語話者の直感を調査することを目的としています。
得られたデータは研究目的のみに利用します。
集計データを専門の学会や資料などで公表する可能性があります。個人情報を含むことはありません。

※設問を読んでいると見られる回答は作業の承認を拒否させていただく場合があります。
参加を自由に取りやめることができますが、その場合ポイントは進呈されません。

[一覧に戻る](#)

図2 参加募集画面

下記の注意事項を読み、実験参加に同意していただける場合、「同意する」を選択してください。「同意しない」を選択された場合、ご参加いただけません。

注意事項

※パソコンでご参加くださいますようお願いいたします。

スマートフォンでは操作できないことがあります。

※このタスクは日本語母語話者の直感を調査することを目的としています。

得られたデータは研究目的のみに利用します。

集計データを専門の学会や資料などで公表する可能性があります。個人情報を含むことはありません。

※参加を自由に取りやめることができます。

※設問を読んでいないと見られる回答は作業の承認を拒否させていただく場合があります。

※文がうまく表示されない、動作が遅いなど、問題がありましたら、一度ブラウザを閉じて、アンケート実施者（西内 nishiuchi-s[a]kanda.kuis.ac.jp）まで[a]を@に変えてご連絡ください。

「同意する」のみを選択してください。

同意しない

同意する

同意する

同意しない

同意する

同意しない

同意しない

同意する

同意する

同意しない

確定して次へ

図3 チェック設問画面

続いて、チェック設問の後に表示される図4のタスク画面の「アンケートのリンク」から外部サイト（PCibex Farm）にアクセスし、図5のような調査画面で例文に対する類似度をリッカート尺度により最大値6から最小値0で評定する。参加者は18グループに分けられ、刺激文がラテン方格法で分配され、1回の回答に刺激文15文が割り当てられ、フィラー30文とともに擬似ランダム順で計45文が表示される。全ての文を評定した後に表示されるランダムコードを、Yahoo! クラウドソーシングのタスク画面に戻り「確認コード入力欄」にコピーアンドペーストし、「エントリーする」をクリックすると作業完了となる。

この調査では、日本語の文を二つ読み、どのくらい似ていると感じるかを答えていただきます。提示された二つの文が非常に似ていると感じる場合は「6」を、全く似ていないと感じる場合は「0」を、中間の場合は引っかかりの程度に応じて「5」～「1」を選択してください。パソコンでご参加くださいますようお願いいたします。スマートフォンでは操作できないことがあります。

アンケートのリンク

✓ チェックするページを見る

手順

- ① 上記リンク先よりアンケートに回答してください。
- ② アンケートの設問がすべて終了すると最後に確認コードが表示されます。
- ③ アンケートページで確認コードをコピーし、下の入力欄にペーストしてください。

文がうまく表示されない、動作が遅いなど、問題がありましたら、一度ブラウザを閉じて、アンケート実施者（西内 nishiuchi-s[a]kanda.kuis.ac.jp）まで[a]を@に変えてご連絡ください。

確認コード入力欄

エントリーする

図4 タスク画面



A このみかんは【あまい】ね。 B あの先生は採点が【あまい】。

【】内の語の意味が、BがAに非常に似ていると感じる場合は「6」を、全く似ていないと感じる場合は「0」を、中間の場合は引っかかりの程度に応じて「5」～「1」を選択してください。

全く似ていない 0 1 2 3 4 5 6 非常に似ている

数字をクリックしてください。

図5 調査画面の一部

リンクからアクセスした外部サイトの調査画面では、まず(3)のような3文の練習を行う。その後、刺激文とフィラーに対する類似度評定に進む。練習では多義語の近い語義の例文、反義語の例文、多義語の遠い語義の例文を含めている。(3a)では類似度を高く、(3b)では低く、(3c)では中程度に評定することを想定した練習である。刺激文が連続しないよう提示するフィラーも、練習と同様に(4)のような「全く似ていない」から「非常に似ている」まで評定が分かれることが期待される例文を提示している。また評定の押し間違いをなくすために、評定ごとに「スペースキーを押すと次の文に進みます。疲れを感じたら休憩してください。」という画面を挟み、調査画面が連続しない形で文を提示している。

- (3) 【】内の語の意味について、BがAに非常に似ていると感じる場合は「6」を、全く似ていないと感じる場合は「0」を、中間の場合は引っかかりの程

度に応じて「5」～「1」を選択してください。

- a. A このテーブルは【まるい】。 B 地球は【まるい】。
 - b. A このテーブルは【まるい】。 B このテーブルは【しかくい】。
 - c. A このテーブルは【まるい】。 B 父は年をとって【まるく】なった。
- (4)
- a. A 今年は雪が【すくない】。 B この辺りは雨が【すくない】。
 - b. A この人は心が【あたたかい】。 B この人は心が【つめたい】。
 - c. A この国は料理が【からい】。 B あの先生は採点が【からい】。

3.3. 調査の結果

回答の収集が完了した地域の順に、概要を表2に示す。表2の終了日時から、回答の収集に要する時間が地域ごとに異なることがわかる。Yahoo! クラウドソーシング作業員数より PCIbex Farm 収集データが少ないのは、PCIbex Farm で類似度評定を回答せず、Yahoo! クラウドソーシングで「エントリーする」をクリックした作業員が存在するためである。PCIbex Farm 上で生成されるランダムコードと Yahoo! クラウドソーシングの「確認コード入力欄」のコードが一致しなかった参加者は今後タスクに参加できないよう処置している。

表2 収集データの概要

地域	開始日時	終了日時	クラウドソーシング作業員数	落選者数	PCIbex Farm 収集データ	平均回答時間 (ms)
関東	2024/7/3 8:13	2024/7/3 10:25	300	2	293	5716
近畿	2024/7/3 8:14	2024/7/3 11:10	300	1	295	6048
東海	2024/7/3 8:08	2024/7/3 15:55	300	1	290	6114
九州	2024/7/3 8:09	2024/7/3 21:40	300	1	288	5806
東北	2024/7/3 8:10	2024/7/4 16:40	300	3	279	6310
中国	2024/7/3 8:02	2024/7/6 15:55	300	0	287	6293
北海道	2024/7/3 8:13	2024/7/8 9:10	300	3	285	6093
信越	2024/7/3 8:07	2024/7/13 20:25	300	1	275	6439
四国	2024/7/3 8:08	2024/7/17 8:10	289	0	255	5926
北陸	2024/7/3 8:00	2024/7/17 8:10	235	3	216	5741
沖縄	2024/7/3 8:03	2024/7/17 8:10	78	0	72	6736

評定の平均値を語義別に全地域で集約すると表3のようにまとめられる。第二主語に「の塩加減が」を付加せず〈塩分がたりない様子〉を表す例文は〈砂糖のような味がする様子〉を表す例文と類似度が高く評定された。第二主語に「の塩加減が」を付加した例文では〈砂糖のような味がする様子〉を表す例文との類似度が、「の塩加減が」を付加しない場合より低く評定されている。このことから、「の塩加減が」を付加せず〈塩分がたりない様子〉を表す例文が〈砂糖のような味がする様子〉を表す例文と類似度が高く評定された要因は、〈塩分がたりない様子〉を表す例文が〈砂糖のような味がする様子〉と解釈された場合があったためと考えられる。

表3 語義別類似度平均

例文 A \ 例文 B	塩分がたりない様子	「の塩加減」付加	砂糖のような味がする様子	芳醇で快い様子	厳格さがたりない様子	程度が低く、不十分である様子
塩分がたりない様子		2.74	4.21	1.62	1.08	1.26
「の塩加減」付加	2.98		2.35	1.34	1.83	1.98
砂糖のような味がする様子	4.54	2.16		1.92	1.19	1.02
芳醇で快い様子	1.64	1.26	1.81		1.65	1.11
厳格さがたりない様子	0.97	1.66	1.08	1.56		2.50
程度が低く、不十分である様子	1.02	1.85	0.97	1.10	2.70	

類似度が高く評定された、第二主語に「の塩加減が」を付加せず〈塩分がたりない様子〉を表す例文と〈砂糖のような味がする様子〉を表す例文の地域ごとの評定を表4に示す。表4では表3で示した平均値より低く評定された地域の値を太枠で示している。

表4 〈塩分がたりない様子〉と〈砂糖のような味がする様子〉の11地域別類似度評定

例文 A \ 例文 B	砂糖のような味がする様子										
	北海道	東北	北陸	信越	関東	東海	近畿	中国	四国	九州	沖縄
塩分がたりない様子	4.26	4.19	4.32	4.02	4.25	4.22	4.39	3.93	3.72	4.48	4.56
例文 A \ 例文 B	塩分がたりない様子										
	砂糖のような味がする様子	4.23	4.10	4.47	4.15	4.33	4.55	4.77	4.66	4.65	4.85

〈塩分がたりない様子〉から〈砂糖のような味がする様子〉への評定では、東北・信越・中国・四国において例文間の評定が比較的低かった。関東・中部・近畿・四国では「うすい」、「みずくさい」など異なる形式で使用されるという石垣（1983）の記述から、関東・中部・近畿・四国では「あまい」に〈塩分がたりない様子〉の直感がないと予測したが、調査結果からは四国でも〈塩分がたりない様子〉と〈砂糖のような味がする様子〉の語義の区分がなされることが見て取れ、石垣（1983）の記述に沿わない結果となった。一方、九州・沖縄で高い評定が得られたことから、当該の地域で〈塩分がたりない様子〉の用法が〈砂糖のような味がする様子〉と解釈されやすいことがわかった。

〈砂糖のような味がする様子〉から〈塩分がたりない様子〉への評定では、北海道・東北・北陸・信越・関東で収集された回答で例文間の評定が比較的低かったが、いずれも4以上と、表3で示した他の語義を表す例文間での評定より高い。〈砂糖のような味がする様子〉という他の語義の拡張元である「あまい」のプロトタイプの意味を表す例文が指標文であったことが、類似度の評定に影響したと考えられる。

4. おわりに

本発表では現代日本語形容詞「あまい」の〈塩分がたりない様子〉の地域的な差異を11地域で調べた結果を報告した。発表では地域ごとに形成される多義ネットワークについて検討する。

今後の課題として、参加者の居住地と言語形成期の異なりの解消と、参加者の少なかった地域における回答の収集があげられる。

謝 辞

本研究は国立国語研究所基幹型プロジェクト「実証的な理論・対照言語学の推進」サブプロジェクト「アノテーションデータを用いた実証的計算心理言語学」によるものです。

文 献

- 石垣福雄 (1983) 『北海道方言辞典』北海道新聞社.
国広哲弥 (1982) 『意味論の方法』東京：大修館書店.
飛田良文・浅田秀子 (1991). 『現代形容詞用法辞典』東京堂出版.
粂山洋介 (2019) 「多義語分析の課題と方法」プラシヤント＝パルデシ・粂山洋介・砂川有里子・今井新悟・今村泰也(編) 『多義動詞分析の新展開と日本語教育への応用』32-50. 開拓社.

Johnson, M. (2005) The philosophical significance of image schemas. In BeateHampe (Ed.), *From perception to meaning: Image schema in cognitive linguistics*: 15-3.

関連 URL

- | | |
|------------------|---|
| PCibex Farm | https://farm.pcibex.net |
| Yahoo! クラウドソーシング | https://crowdsourcing.yahoo.co.jp |

編集後記

本イベント『Evidence-based Linguistics Workshop』も3回目になりました。理論言語学・対照言語学においても言語学のオープンサイエンス化を目的として、査読のない自由な発表の場として、本ワークショップを運営しております。既存の学会の発表審査で通らなかった内容や、アイデア段階の内容でも発表申込をしていただければと考えております。地方の学生には旅費支援なども行っておりますので、発表を検討していただければと思います。

今回は招待講演1件・口頭発表2件・ポスター発表4件を9月21日に国立国語研究所2F講堂にて行います。国立国語研究所からは、2件のプロジェクト紹介を行います。また、2024年8月に新しく設置された「次代言語科学研究センター」についても紹介いたします。

浅原正幸（国立国語研究所）