

国立国語研究所学術情報リポジトリ

「構成の反復」における反復単位の認定基準に関する考察

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): 修辞表現, アノテーション, 言語資源, 反復 キーワード (En): figurative expression, annotation, language resources, repetition 作成者: 伊藤, 薫 メールアドレス: 所属:
URL	https://doi.org/10.15084/0002000331

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



「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、上記研究の他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、それらを比較する。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	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,	religion has been retarded, morality has been vitiated, atrocities have been engendered;	England millions of her treasure, Ireland torrents of her blood.
[the church PP NOMINAL]	[(PP) which NP VP]	[N has been V-pp]	[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」のカテゴリーであることや、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.