

# 日本語連体修飾節の抽出による文分割の 可否判断に適した基準の提案

後藤 功雄

田中 英輝

NHK 放送技術研究所

## 1 はじめに

日本在住の外国人へニュースを分かりやすく伝えるために、ニュースをやさしい日本語で提供する研究を進めている。普通のニュースからやさしい日本語ニュースへの書き換えでは、主に、語彙の平易化、構造の平易化、文章の簡潔化、背景知識の補足という4種類の編集を行う。

これらのうち、本稿では、構造の平易化に着目する。この方法の1つとして、連体修飾節を抽出して被修飾名詞（主名詞と呼ぶ）と合わせて独立した文にする（主節化と呼ぶ）ことによる文の分割（主節化分割と呼ぶ）を人手により行っている [2, 1]。主節化分割による構造の平易化の例を以下に示す。

この遺伝子をマウスの脳の 記憶などをつかさどる海馬という部分に組み込みました。

この遺伝子をマウスの脳の海馬という部分に組み込みました。

海馬は 記憶などをつかさどります。

ただし、連体修飾節を含む文には次の例で示すように主節化分割しにくく不自然になる場合や主節化分割できない場合がある。

輸入する石油を減らす。

? 石油を減らす。その石油は輸入する。<sup>\*1</sup>

? 石油を輸入する。その石油を減らす。

このように書き換えてしまうと元の文より分かりにくくなってしまいます。このため、自動的に主節化分割してやさしい日本語ニュースに自動変換するためには、自然な文に分割できるかどうかを判断する必要があります。

連体修飾節はまず外の関係と内の関係に分類することができる [3]。これらのうち、本稿では、自然な文になるように主節化しやすいと考えられる内の関係を対象とする。なお、外の関係と内の関係との区別は、既存の自動判別手法 [8] などを用いて判別されている状況を想定する。

内の関係の連体修飾節を含む文を主節化分割できるかどうかの判断基準として、非限定的修飾は主節化分割が

可能で、限定的修飾は主節化分割しにくいというものがある [5] (p.35)。

内の関係の連体修飾節を含む文を主節化分割する研究として、あらかじめ非限定的修飾節のみを選択しておいて、それらを人手で主節化分割し、その書き換え結果を分析した研究がある [9]。この研究は、非限定的修飾が与えられた場合に、主節化分割がどのような書き換えで実現できるかを分析したものである。この研究で得られた書き換え方法を適用して主節化分割するためには、まず、非限定的修飾かどうかを判別する必要がある。

英語では制限用法（限定的修飾に相当する）と非制限用法（非限定的修飾に相当する）はカンマによって明示的に区別されているが、日本語ではこのような区別を示す表層的な標識は存在しない。そのため、日本語ではこれらの区別は自明ではない。また、どのような連体修飾節が非限定的修飾であるかという基準が複数存在する。さらに、非限定的修飾の定義は主節化分割の可否を考慮して決められたものではないので、非限定的修飾の基準を主節化分割の可否にそのまま適用できるかは検討が必要である。

我々は、内の関係の連体修飾節を含む文の主節化分割の可否判断に適した基準を提案する。まず、非限定的修飾の2つの定義を示し、本稿で対象とする課題を説明する（2節）。次に、非限定的修飾の各定義と主節化分割の可否との関係を調べる（3節）。そして、主節化分割の可否判断に適した基準を提案する（4節）。さらに、提案した基準で扱わない特徴で、主節化分割の可否に関係するいくつかの要因を示す（5節）。

## 2 対象とする課題

日本語の非限定的修飾（または非制限的修飾）と限定的修飾（または制限的修飾）の違いは表層的には自明ではなく、それらを分類する定義として2つの考え方があり。以下に2つの非限定的修飾の定義を示す。

- (i) 連体修飾節が主名詞全体を述べる場合 [7]

この考え方は次のように説明されている。「連体修飾節が被修飾名詞の表す物事を制限せず、その全体に関する叙述となっているものを非限定的修飾という。連体修飾節が主名詞の一部分のみに関する叙述となっている場合を限定的修飾と呼ぶ。」 [7] (pp.339-340)（これらの例を 3.1 節の (1) と (2) に

\*1 ? は不自然を示す。

示す。)

- (ii) 主名詞が特定されている場合 [5] (p.36), [6]

主名詞の指示対象が文中の述語によってあらかじめ具体的なある個体に決まっている場合に非限定的修飾と分類する。それ以外の場合が限定的修飾となる\*2。(これらの例を 3.2 節の (3) と (4) に示す。)

このように複数の定義がある状況では、非限定的修飾なら主節化分割が可能でそうでないなら不可能であるという判断基準は曖昧である。さらに、非限定的修飾の定義は主節化分割の可否を考慮して決められたものではないので、その可否判断に適しているかは検討が必要である。そのため、主節化分割の可否を判断するのに適した基準を明らかにすることが課題である。

### 3 非限定的修飾の各定義と主節化分割との関係

#### 3.1 (i) の定義と主節化分割との関係

まず、2 節の (i) の定義と主節化分割との関係について調べる。この定義では、次の例は文の後の括弧内に示すように分類される。

- (1) (学部生と違って、) 高度な知識を身につけた大学院生は専門性を必要としない職場では敬遠されることがある。(非限定的修飾)

大学院生は高度な知識を身につけている。大学院生は専門性を必要としない職場では敬遠されることがある。

- (2) (大学院生には高度な知識を身につけていないものもいるが、) 今日は高度な知識を身につけた大学院生が参加している。(限定的修飾)

今日は大学院生が参加している。その大学院生は高度な知識を身につけている。

(1) の例は、一般に大学院生とは高度な知識を身につけているものであると解釈しているため、連体修飾節は大学院生全体に関して述べている。そのため、非限定的修飾に分類される。そして、この例は主節化分割が可能である。

(2) の例では、連体修飾節は、大学院生のうち一部の大学院生に関して述べている。そのため、限定的修飾に分類される。しかし、この例も主節化分割が可能である。

これらのことから、(i) の定義による非限定的修飾の範囲と主節化分割が可能範囲は一致していないことが分かる。

#### 3.2 (ii) の定義と主節化分割との関係

次に 2 節の (ii) の定義と主節化分割との関係について調べる。主名詞があらかじめ具体的なある個体に定まっ

ていることを、[6] で用いられている表現を用いて「特定指示されている」と表現する。ある個体とは、例えば「本」が指しているものは本以外の机などと区別される型あるいは種類であるのに対し、「この本」が指しているものは具体的なある個体である。主名詞が特定指示されているとは、主名詞が具体的なある個体を指しているということである。

- (3) このソフトウェアを使える学生が助けてくれた。

学生が助けてくれた。その学生はこのソフトウェアを使える。

この例の分割前の文中の「学生」が指示するものは具体的なある個体である。なぜなら主節の述語で学生を特定しているからである\*3。そのため、(ii) の定義では非限定的修飾に分類される。そして、この例の場合は、主節化分割が可能である。

一方、次の例は主節化分割しにくい。

- (4) このソフトウェアを使える学生に助けて欲しい。

? 学生に助けて欲しい。その学生はこのソフトウェアを使える。

? 学生がこのソフトウェアを使える。その学生に助けて欲しい。

この例の分割前の文中の「学生」は特定指示されていない。そのため、(ii) の定義では非限定的修飾ではない。そして、この例の場合は、主節化分割しにくい\*4。

(3) と (4) の違いとして、主名詞が特定指示されているかないかという違いがある。このことから主名詞が特定指示されていれば主節化分割できる [5] と考えられる。

ここで、次の例について考える。

- (5) 植物プランクトンが持っている遺伝子に着目する。

植物プランクトンが遺伝子を持っている。その遺伝子に着目する。

分割前の文の「遺伝子」が指示しているものは型や種類であって、「この遺伝子」のようなある個体を指してはいない。そのため、(ii) の定義では、この例は非限定的修飾には該当しない。しかし、主節化分割が可能である。これにより、特定指示されていない場合でも分割可能な場合があるということが分かる。すなわち、(ii) の定義による非限定的修飾の範囲と主節化分割が可能範囲は一致していない。

\*2 [5] (p.39) では「主名詞の指示対象があらかじめ定まっているかどうかによって集合限定の機能が発動されるか否かが決まる」と述べている。この集合限定とは限定的修飾を意味する。

\*3 特定していることを示す表層的な特徴には、主節で主名詞を項に持つ述語がタ形、またはそのアスペクトがテイル形、またはその述語が存在を表す動詞(例: いる)などがある。

\*4 この場合は、「学生がこのソフトウェアを使える、そういう条件に合う学生に助けて欲しい。」という意味である [4] (p.91)。

表 1 非限定的修飾の各定義と提案する定義による判別結果と主節化分割の可否との関係

番号	連体修飾節を含む文	分割可 (判別結果) <sup>a</sup>			主節化分割 <sup>b</sup>
		(i)	(ii)	提案	
(1)	(学部生と違って、) 高度な知識を身につけた大学院生は専門性を必要としない職場では敬遠されることがある。	✓		✓	
(2)	(大学院生には高度な知識を身につけていないものもいるが、) 今日高度な知識を身につけた大学院生が参加している。		✓	✓	
(3)	このソフトウェアを使える学生が助けてくれた。		✓	✓	
(4)	このソフトウェアを使える学生に助けて欲しい。				?
(5)	植物プランクトンが持っている遺伝子に着目する。			✓	

<sup>a</sup> (i) と (ii) のチェックは非限定的修飾を示す。

<sup>b</sup> は主節化分割が可能を示す。

## 4 提案する主節化分割の可否判断に適した基準

我々は主節化分割できる (分割可と呼ぶ) か主節化分割しにくい (分割不可と呼ぶ) かの判断に適した基準として、次の定義を提案する。

**分割可の場合** 文の書き手が文中において連体修飾節と主名詞が表す内容を事実として述べている場合。

**分割不可の場合** 文の書き手が文中において連体修飾節と主名詞が表す内容を事実としてではなく条件を述べている場合。

ここで、事実として述べられているかどうかは文の意味で決まる。特に文中の述語<sup>\*5</sup>によって決まる。ここで、事実とは、特定された個別の事態のみでなく、書き手が事実として扱っている事態も含む<sup>\*6</sup>。

この定義は (ii) の定義による非限定的修飾の対象範囲を含み、さらにその対象範囲を拡張するものである。

ここで「事実」がどのようなものであるかを例を用いて説明する。

(6) ピアノを弾ける学生が話しかけてきた。

(7) ピアノを弾ける学生を探す。

(6) の連体修飾節は、ある学生がピアノを弾けるという「事実」を説明している。一方、(7) の連体修飾節では、ある学生がピアノを弾けるという事実を述べているわけではなく、その連体修飾節はピアノが弾けるという「条件」を示している。

ここで、提案した定義および非限定的修飾の定義によって、(1)~(5) の例が分割可か分割不可かのどちらに判別されるかを調べる。なお、(i) と (ii) では非限定的修飾に該当する場合に分割可、そうでない場合に分割不可と判別する。結果を表 1 に示す。

可と判別する。結果を表 1 に示す。

まず (3) と (4) について説明する。これらの違いとして、(3) の連体修飾節は、ある学生がこのソフトウェアを使えるという事実を述べていて、(4) では、学生がこのソフトウェアを使えるという事実は示されておらず、このソフトウェアを使えるという条件を述べている。提案する定義と (ii) の定義は (3) を分割可、(4) を分割不可と判別する。これらの例のうち主節化分割が可能であるのは (3) であるため、これらの 2 つの定義による判別結果と主節化分割の可否は一致している。

次に、(5) について説明する。この例では「植物プランクトンが遺伝子を持っている」ことを文中で事実として述べている。そのため、提案する定義はこの例を分割可に判別する。一方で、3 節で説明したように、(ii) の定義ではこの (5) は非限定的修飾には該当しないので分割不可に判別される。そして、この例は主節化分割が可能である。この例を分割可に判別する定義は、比較した 3 つのうち提案する定義のみである。

最後に、(1) と (2) について説明する。(1) では、書き手は大学院生が高度な知識を身につけていることを一般的な事実として説明している。(2) では、具体的なある大学院生が高度な知識を身につけているという事実を述べている。そのため、提案する定義は (1) と (2) の両方を分割可に判別する。一方で、3 節で説明したように (i) の定義は (1) のみを分割可に判別する。そして (1) の主名詞が指示しているものは型や種類であり、主名詞が特定指示されているのは (2) のみであるので、(ii) の定義では (2) のみが分割可に判別される。これらの例はどちらも主節化分割が可能であり、比較した 3 つの定義のうち提案する定義のみがどちらの例も分割可に判別する。

以上から、提案する定義が示す範囲は、非限定的修飾の (i) と (ii) の定義に比べて、主節化分割が可能なもの範囲に近いと考えられる。

## 5 提案した基準で扱わない特徴と主節化分割との関係

主節化分割した文がニュースとして自然な文になるかどうかは、4 節で提案した基準以外の要因にも影響され

\*5 主に主名詞を項とする述語や連体修飾節の述語である。

\*6 例えば (5) の場合に実際に植物プランクトンが遺伝子を持っているかどうかを科学的に検証する必要はなく、書き手が事実として扱っていればよい。

る。そこで本節では、このような要因となるいくつかの特徴を示す\*7。ここで着目する特徴は排他的なものではない。

#### ● 大きさなどを表す主名詞

大きさや量など数値で示される概念を表現する名詞が主名詞の場合。

この場合、主節化するとニュースとしては自然にならないと考えられる。

- (8) 5キログラムある重さのスイカを買った。  
? 重さが5キログラムある。その重さのスイカを買った。
- (9) 渋滞で車がゆっくりとした速度で進んだ。  
? 速度がゆっくりとしていた。渋滞で車がその速度で進んだ。

#### ● 具体性が低い主名詞

具体性の度合いが低い名詞が主名詞の場合。

「もの」や「こと」は最も具体性が低い名詞であるが、これらに次いで具体性が低いと考えられる名詞が主名詞の場合、これらが修飾節や修飾語なしに単独で用いられると不自然に感じられる。

- (10) お正月をふるさどで過ごした人が戻ってくる。  
? 人がお正月をふるさどで過ごした。その人が戻ってくる。

#### ● 修飾語がある主名詞

主名詞に修飾語がある場合。

主名詞の具体性が低い場合でも、修飾語によって具体性が高くなれば分割後の文が自然になる。

- (11) お正月をふるさどで過ごした都会の人が戻ってくる。  
都会の人がお正月をふるさどで過ごした。その人が戻ってくる。

#### ● 原因や対比などを表す語を含む主節

「だから」など原因・理由・条件を示す語、「のみ」など限定を示す語、「より」など対比を表す語が主節にあって、これらの語が連体修飾節の内容を受けられる場合。

分割後の文の順番に制限が生じると考えられる。抽出した連体修飾節を独立させた文の順番が先となる必要があると考えられる。

- (12) 輸入した大豆だから安かった。  
? 大豆だから安かった。その大豆は輸入した。  
大豆を輸入した。その大豆だから安かった。

例の他に、「ので」「ために」「ならば」「だが」「のみ」「だけ」「しか」「と比べて」「より」「もっと」「さらに」「ま

だ」などがある。

#### ● 「主名詞+だ」の構文

主節が「主名詞+だ」の構文の場合。

主名詞が主題の上位語だと、意味のない文になって不自然になる(例14)ので、主題と主名詞の関係が抽出可・不可に影響する。また、主題がないなどで主文が短いと意味のない文になる場合がある(例15)。

- (13) 彼は宇宙を研究している研究者だ。  
彼は研究者だ。その研究者は宇宙を研究している。
- (14) 彼は宇宙を研究している人だ。  
? 彼は人だ。その人は宇宙を研究している。
- (15) 宇宙を研究している研究者だ。  
? 研究者だ。その研究者は宇宙を研究している。

## 6 おわりに

2つの非限定的修飾の定義と内の関係の連体修飾節を含む文の主節化分割の可否との関係を調べ、どのような条件において主節化分割できるかを検討した。そして、主節化分割の可否判断に適した基準を提案した。さらに、提案した基準で扱わなかった特徴で主節化分割の可否に影響するいくつかの要因を示した。今後は提案した基準に関係する特徴量を用いて主節化分割の可否を機械学習で自動判別する予定である。

## 謝辞

文献を紹介いただいた IR-ALT の佐藤氏に感謝する。

## 参考文献

- [1] Isao Goto, Hideki Tanaka, and Tadashi Kumano. Japanese News Simplification: Task Design, Data Set Construction, and Analysis of Simplified Text. In *Proceedings of MT Summit XV*, pp. 17–31, 2015.
- [2] 後藤功雄, 熊野正, 田中英輝. 一般のニュースからやさしい日本語ニュースへの書き換えの分析. 言語処理学会第20回年次大会, pp. 15–18, 2014.
- [3] 寺村秀夫. 連体修飾のシンタクスと意味(1)~(4). 「寺村秀夫論文集I」くろしお出版1992所収, 1975–1978.
- [4] 大島資生. 朝倉日本語講座5第5章 連体修飾の構造. 朝倉書店, 2003.
- [5] 大島資生. 日本語連体修飾構造の研究. ひつじ書房, 2010.
- [6] 南美英. 連体修飾節の「制限・非制限」の規定について. *Journal of Japanese Language and Culture*, No. 15, pp. 21–37, 2009. DOI: <http://dx.doi.org/10.17314/jjlc.2009.15.002>.
- [7] 日本語文法学会(編). 日本語文法事典. 大修館書店, 2014.
- [8] 阿辺川武, 奥村学. 日本語連体修飾節と被修飾名詞間関係の解析. 自然言語処理, Vol. 12, No. 1, pp. 107–123, 2005.
- [9] 野上優, 藤田篤, 乾健太郎. 文分割による連体修飾節の言い換え. 言語処理学会第6回年次大会, pp. 215–218, 2000.

\*7 ここでは「ある」という修飾語を必要とする場合はニュースとして自然ではないとして扱う。