

# 引用ト節(句)と共起する語彙と構文 -BCCWJコーパスに基づく語彙・構文彙の構築に向けて-

藤井 聖子

東京大学大学院 総合文化研究科 言語情報科学専攻

## 1. はじめに

本発表では、多機能性をもつ引用ト節(句)に関して、『現代日本語書き言葉均衡コーパス』(BCCWJ 2008 領域内公開版、2009 領域内公開版; 国立国語研究所)を用いて、引用ト節(句)と共起する語彙群と構文を明らかにし、さらに、その語彙または構文が喚起する各々の意味フレームにおいて引用ト節(句)がどのようなフレーム要素(5項説明参照)として機能するかを示す。本分析を通して、文法的機能語が標識する構文の、語彙・構文の統合分析の一事例とする。

## 2. 本研究で扱う問題、概要、目的

### 2.1 引用ト節(句)の用法の多様性

引用ト節(句)は、主節の文構成要素・主動詞の項構造の項として位置付けられる用法【内の関係】と、主節の項構造の外に位置付けられる副詞節的用法

【外付けの関係】とがある(また、両者の中間的用法も認められた)。さらに二つ目の類型指標として、引用助詞トは、発話思考の言語内容を引用する用法の他に、狭義の引用以外の用法もある。

### 2.2 引用ト節(句)の多様性の分析・記述への語彙的・構文的アプローチ

上記多様性を踏まえ、本研究は、多様な構文に参与する引用助詞トを一律・一様に捉えるのではなく、引用ト節(句)が参与する構文(タイプ)の多層性と広がり进行分析し明示することの重要性を尊重する立場をとる。【内の関係】と【外付けの関係】とを包括する構文類型を示すとともに、両者を峻別・記述する指標・方法をフレーム意味論に基づき提案する。

同時に、引用ト節(句)が参与する構文の多層性と広がり进行分析・記述するためには、(i)文法的機能語トが標識する構文(ここでは「引用ト標識構文」)进行分析・記述すること;(ii)この標的構文に参与する語彙群をコーパスに基づき構文ごとに示すとともに、その語彙相互の繋がり・体系を把握するために、フレーム意味論・フレームネットにおける概念・方法論が有益であること;(iii)文法的機能語と共起する語彙群をコーパスに基づき分析・記述する際、参与

する構文タイプごとに考察する必要があること、コーパスに基づく「構文彙」(8項説明参照)の分析と構築が有用であり必要であることを示す。

【内の関係】用法に関しては、引用ト節(句)が参与する構文を明示しつつ、引用ト節(句)を項・フレーム要素として喚起する述語語彙群を体系的に構築する。【内の関係】用法の場合、コーパスに基づく語彙分析と語彙記述が有益である。一方、後者の【外付けの関係】副詞節的用法に関しては、引用ト節(句)と共起する語彙自体を記述し構築することは、前者【内の関係】の場合ほど意味がなく、構文自体の構文的意味・統語的特徴・語用論的特徴を明らかにする必要はある。紙幅の関係で、後者副詞節的用法の別方法論での分析の詳細は他稿(藤田2000, 藤井2002, 2009等)に譲り、本発表では両者を横断する全体的分類や分析を示した上で、主に前者の分析を報告する。

## 3. 本分析に使用した BCCWJ サブコーパス

表 1. 分析対象とした BCCWJ のサブコーパス、総語数、引用ト構文の生起数、外付け用法用例数<sup>1</sup>

サブコーパス	コーパス語数	引用ト構文生起数	外付け用法生起数
白書	5,000,000	17,513	2
書籍: 文学 (文学全体の約72%)	6,324,175	84,149	272
書籍: 言語	398,497	6,244	27
書籍: 技術 工学	1,115,821	11,733	45
書籍: 自然 科学	1,074,332	13,684	58
書籍: 哲学	1,403,199	19,623	57
書籍: 歴史	2,141,841	27,846	118
書籍: 総記	521,436	6,515	25
合計	17,979,301	187,307	604

<sup>1</sup> 『BCCWJ領域内公開版』収録のテキストファイルを入力データとし、処理には、形態素解析ツール MeCab (<http://mecab.sourceforge.net/>、工藤拓、松本裕治)、及び、検索ツール ChaKi (<http://chasen.naist.jp/hiki/ChaKi/>、松本裕治等)を用いた。ChaKiによる出力データに対して、人手で形態素の誤解析削除と用例意味解釈とコーディングを加えたが、表1の引用ト生起数には、共格や接続助詞など形態素の誤解析検出トークンも含まれている。右欄外付け用法用例は、中欄データから手作業読解で一列ずつ抽出した。

#### 4. 所謂引用の助詞トが参与するト構文の構文類型（作業分類）と基軸指標

本稿で報告するコーパス分析では、国立国語研究所(1951)、藤田(1986)、特に山崎 (1993)の洞察・提案を鑑み、所謂引用の助詞「と」の用法の大別を以下の大分類とした。表2に示す作業分類に基づき手作業でコーディングをした。

##### 4.1 指標1：【内の関係】vs.【外付けの関係】

概要で述べたとおり、引用ト節(句)は、主節の文構成要素・主動詞の項構造の項として位置付けられる用法【内の関係】(例：1)と、主節の項構造の外に位置付けられる副詞節的用法【外付けの関係】(例：3)とがある。また、本研究のコーパスデータの分析では、これら両者の中間に位置づけられるべき用法も認められた(例：2)。

- (1) Q1. 内の関係：「言語表現補語引用型」  
 分かったと言った。分かったと書いた。  
 分かったと思った。分かったと喜んだ。
- (2) Q1-2. 中間的用法：「言語表現引用同一事態型」  
 分かったと頭を下げた。分かったと涙を流した。

- (3) Q2. 外付けの関係：「言語表現外付け引用事態型」  
 分かったと飛び出て行った。  
 分かったとバナナを手渡した。

★名詞が主なフレーム喚起語となる場合

- (4) 分かったと期待をかけた。ーと歓声をあげた。  
 分かったと手紙を送った。

動詞の項構造の項かどうかという問題に留まらない。

「期待」等の事態性名詞は項構造をもつ名詞として、当然ト標識引用部をフレーム要素として含むフレームの喚起語となる。さらに、「分かったと手紙を送った」における「手紙」のように、通常名詞がフレーム喚起語として、ト標識引用部をフレーム要素として喚起することもある。

##### 4.2 指標2：引用用法と狭義の引用ではない用法

山崎(1993)が指摘するとおり、所謂引用助詞トには、狭義の引用でない重要な用法がある。表2では、この区別をQ(上)とP(下)とで示す。

- (5) P1. 酒席は作戦会議の場と化した。  
 (6) P2. 財産被害が、19.7%と多くなっている。  
 (7) P3. 朝倉、島田、木下、山田… 渡辺と続いて…

表2. 所謂引用の助詞「と」が標識する構文(引用ト構文)の用法の本分析における作業分類

引用「ト」構文の用法大別	構文の類型	ト標識引用部	述部	第一段階コーディング	述部の例	
Q. 引用の用法	Q1.	言語表現補語引用型	言語表現の内容	言語表現を内容とする事態	発話思考 発話思考:発話 発話思考:思考 感情	いう、話す 主張、回答、説明 思う、考える 判断、予想 驚く、喜ぶ 心配、後悔
	Q1とQ2との中間	Q1-2. 言語表現引用同一事態型	言語表現の内容 言語表現の内容	言語表現を内容とする事態	行為:感情 行為:言動	笑う、泣く 微笑み、号泣 うなずく (電話、メールはQ1とした)
	Q2.	言語表現外付け引用事態型	言語表現の内容	言語表現を内容としない事態	事態:外付け	様々 (様々な述語(なつ、手渡す、等))
	Q3.	非言語引用事態型	非言語的表現の内容	非言語的表現を内容とする事態	事態:他	様々な述語 様々
	P. 狭義の引用ではない用法	P1.	変化結果提示型	非言語表現	変化結果 移動結果	結果:変化結果 移動:移動結果
P2.		尺度属性型	非言語表現	尺度・測定・比較	尺度関連	高まる、増える 増加、減少、低迷
P3.		列挙型	非言語表現		列挙	続く、並ぶ
参考:形態素解析(MeCab)において		一貫して「引用助詞」と誤解析された形態素 頻繁に「引用助詞」と誤解析された形態素			誤:共格 誤:接続助詞	

#### 5. フレーム意味論・フレームネットにおける【内の関係】【外付けの関係】の捉え方

##### 5.1 フレームネット

フレームネット(FrameNet: FN; Ruppenhofer他 2006; <http://www.icsi.berkeley.edu/~framenet/>)は、フレーム意味論(Fillmore 1982, Petruck 1996)に基づき、コーパスデータを参照しつつ、語彙の意味および参与する構文の

分析を行い、その意味・形式の記述・情報を電子語彙体系として構築・資源化していく。<sup>2</sup> FrameNetでは、語彙項目(Lexical Unit: LU)の意味的・構文的特質や語彙項目間の関係を、フレームやフレーム要素(Frame Element: FE)という概念を用いて記述する。

<sup>2</sup> 英語のFrameNetプロジェクトとの連携のもとに進行中の「日本語フレームネット」プロジェクトに関しては、小原他(2005a)、小原他(2005b)、etc.), <http://jfn.st.hc.keio.ac.jp/ja/index.html>等を参照。

本研究では、語彙または構文が喚起する各々の意味フレームにおいて引用ト節(句)がどのようなフレーム要素として機能するかを分析した。「喚起される意味フレームにおいて引用ト節(句)がどのようなフレーム要素として機能するか」という観点での分析は、フレーム要素の二つの分類軸で行った。

## 5.2 フレーム要素のアノテーション (その1)

一つは、参与項目のフレームにおける意味役割(一般的な「意味役割」に相当するものであるが、フレーム意味論/FN においては、各々の意味フレームにおいて想定され定義されるきめ細かな事態構成要素の意味)の記述である。この意味役割的アノテーションでは、【内の関係】では、引用ト節(句)が MESSAGE, CONTENT, DECISION, LABEL, REASON 等のフレーム要素として機能することが認められた。

## 5.3 フレーム要素のアノテーション (その2)

もう一方は、該当フレーム要素がフレーム喚起語や喚起されるフレームとどのような関係にあるかに関する分析である。フレームネットでは、この観点において、三種類のフレーム要素の関与レベルを設定している: Core frame elements (FE), peripheral FE, extrathematic FE である。フレーム喚起語に直接喚起され、喚起されるフレーム事態に内在的に参与するフレーム要素が Core FE または peripheral FE である。中でも Core FE は、喚起されるフレーム事態において概念的に必須の参与要素であり(常にではないが)多くの場合統語的にも選定される要素である(項構造の必須項)。一方、extrathematic FE は、フレーム喚起語に直接喚起される要素ではなく、様々なフレームと共起し、該当フレームとは別の事態を導入しつつ、該当フレームの参与者や状況を補足的に描写するフレーム要素である(Ruppenhofer et. al. 2006: 135-136)。

## 5.4 【内の関係】【外付けの関係】の捉え方: フレーム要素としての引用ト節/句のバリエーション

本研究では、このフレーム要素の喚起フレームとの関わり方(Core FE, peripheral FE, extrathematic FE)という観点で、引用ト節(句)の【内の関係】と【外付けの関係】との峻別を捉え直し、引用ト節(句)が文全体・構文全体の中で、(主節の)フレーム喚起語や喚起されるフレームとどのような関係にあるかを、コーパスデータから抽出した用例で分析した。

### 5.4.1 【内の関係】

【内の関係】の引用ト節(句)は、フレーム喚起語が喚起するフレームの主要フレーム要素 Core FE として位置づけられる。その Core FE の代表例が、「話す」「述べる」「唱える」等が喚起する Statement フレームにおける MESSAGE フレーム要素; 同様に Communication フレームの下位フレームである Communication\_manner フレームや Request フレームにおける MESSAGE; 「喜ぶ」「驚く」「困る」等が喚

起する Experiencer\_subj フレームにおける CONTENT フレーム要素; 「決める」等が喚起する Deciding フレームにおける DECISION 等である。

### 5.4.2 【外付けの関係】

一方、副詞節的用法では、ト標識引用節を extrathematic FE として位置づけることができる。副詞節的引用節は、意味的に主節に関与しているが、主節の内在的構成要素ではなく、主節述語によって要請されたフレーム構成要素ではない。

(8) 「やれるだけやってみます」ともうすぐ早朝点呼の始まる寮室へ戻っていった。

### 5.4.3 中間的用法

しかし、抽出したコーパスデータを吟味すると、extrathematic FE とみなしてよい用例の他に、中間的なもの(Q1-2)「言語表現引用同一事態型」に多々遭遇した。(9)の例は、主節の述部が喚起するフレーム事態とは別の事態をト引用部が導入しているわけではなく、言語表現を補足的に添えることにより、主節の提示する行為の様態描写を精巧化している。従って、このような中間的な Q1-2 でのト標識節は、extrathematic FE ではなく、peripheral FE (ここでは MANNER や CIRCUMSTANCE フレーム要素)と位置付けられる。

(9) 美女が「ニイハオ」と出迎えてくれる  
「どうぞ」と席を譲った。  
「お先に失礼します」と頭をさげた。  
「じゃあな」と片手をあげて走りさっていった。  
「よし、それでよい」と合格点をつけたはずだ。

### 5.4.4 引用ト節(句)のバリエーション

以上、引用ト節/句が、参与する構文によって、Core FE, peripheral FE, extrathematic FE の3種類のいずれにもなりうることを明らかにした。このことから、文法的機能語の記述に構文的アプローチが不可欠である。

## 6. 共起語彙とそれらが喚起するフレームの分析: Q1 および Q1-2 の場合

### 6.1 コーパスにおける共起語彙の分析

引用助詞トの後続5語以内に共起する動詞および事態性名詞の語彙リストを(誤解析削除等)手作業で修正し、動詞・事態性名詞それぞれ頻度順に共起語彙表を作成した(表3・表4は紙幅制限のため本稿では割愛)。用例を語彙ごとに分析したところ、共起語彙リストにおいて上位にあがる語の殆どが【内の関係】の言語表現引用型 Q1(または Q1-2)であった

### 6.2 高頻度語彙が喚起する意味フレームの分析

述部のフレーム喚起語が喚起するフレームに関して、白書、書籍文学、書籍言語、書籍技術工学4つのコーパスでの、動詞・事態性名詞それぞれ頻度上位100語(延べ800語)が喚起する意味フレームを分析した(図1の下位フレーム参照)。さらに喚起するフレームごとに分類語彙表を作成した(紙幅のため本稿では割愛)。

