

# 逆接の推論関係に着目した日本語談話関係アノテーション

窪田愛<sup>1</sup> 佐藤拓真<sup>1</sup> 天本貴之<sup>1</sup> 秋吉亮太<sup>1</sup> 峯島宏次<sup>1</sup>  
<sup>1</sup>慶應義塾大学

ai.kubota.m@gmail.com takuma.sato.keio@gmail.com  
 amamoto@keio.jp georg.logic@gmail.com  
 minesima@abelard.flet.keio.ac.jp

## 概要

本研究では、逆接の談話関係で前提される推論に着目し、同時進行と逆接の二つの用法を持つ接続詞「ながら」「つつ」に対して談話関係アノテーションを試みる。逆接を含む談話関係コーパス自体はすでに存在するが、前提される推論を実際に書き出し、逆接の推論関係を特定しようとする点に本研究の特色がある。本稿ではアノテーション手法とその結果、ならびに、アノテーションの際に判明した逆接の特徴について報告する。

## 1 はじめに

テキストの談話関係を適切に認識することは、人間の言語理解の重要な側面であり、自然言語処理でも重要な課題の一つである。しかし、談話関係を分類する基準、その関係を担う談話単位を設定する基準を明確にすることは必ずしも容易ではなく、言語学・言語処理で様々な試みが続いている。一つの有力なアプローチは、明示的な接続表現（談話標識）に着目して談話関係を判別するもの [1, 2] であるが、同一の接続表現が複数の談話関係を担う場合、その談話関係の判別は容易ではない。

このような背景から、本研究では逆接の意味をもつ接続表現に着目し、詳細な統語情報が付与されたツリーバンクを利用した談話関係アノテーションを試みる。具体的には、同時進行と逆接の二つの用法をもつ「ながら」「つつ」をとりあげ、「前提される推論」という観点から談話関係の分類を試みる。

## 2 関連研究

### 2.1 逆接・譲歩

ここでいう逆接（あるいは逆接・譲歩）とは、「何らかの含意関係の否定」[3]、あるいは、“denial of

expectation” [4, 5] を伴う接続関係のことであり、以下の二種類があると言われている。（例文は [4](37)より抜粋。）

- (1) 直接型 (Direct concessive)  
 Although John is poor, he is happy.
- (2) 間接型 (Indirect concessive)  
 The car is stylish and spacious, but it is expensive.

直接型と間接型の違いは、前提される推論の違いにある。間接型では「P (= John is poor) ならば、大抵 Q (= he is happy) ではない」という情報が前提されているのに対して、間接型では「P (= The car is stylish and spacious) ならば、大抵 Q (= it is expensive) ではない」という情報が前提されているとは考えにくい。間接型の場合、「P (= The car is stylish and spacious) ならば R (= we should buy it)、Q (= it is expensive) ならば R でない」といったように第三の項 (R) が隠れている点に特徴がある [6, 7, 8, 4, 5]<sup>1)</sup>。

### 2.2 「ながら」「つつ」

「ながら」は動詞の語基、非過去のイ形容詞、ナ形容詞の語幹に、また名詞及び副詞には直接後続し、「同時進行（付帯状況）」を表す用法と「逆接」を表す用法があるとされる [9, 10]<sup>2)</sup>。「つつ」は動詞の語基に後続し、やはり「同時／付帯状況」と「逆接」の二つの用法がある [9]。

典型的には、「食べる」「走る」など動的述語の場合は「同時進行」、「思う」「狭い」などの静的述語の場合は「逆接」の用法になることが多いものの、述語の種類だけで「同時進行」か「逆接」かが決まるわけではなく、以下の例 ([9] の 3.4 節より抜粋) が示すように主節との意味的な関係や文脈により解釈

1) 間接型は「迂回型」[6]、「疑似譲歩文」[7]とも呼ばれる。  
 2) 他に「様態」を表す用法もあるが、「涙ながら (に)」のような語彙化されたものであり [10]、本稿では対象外とする。

が影響される [9, 10]。

- (3) a. さびしいと思いながら毎日を過ごした。  
(同時進行)
- b. さびしいと思ながらそれを口にしな  
かった。  
(逆接)
- (4) a. 3時間、英語を勉強しながら、音楽を聞  
いた。  
(同時進行)
- b. 12年間も英語をながら、日常会話もろく  
にできない。  
(逆接)

そのため、以下の例 ([10](31)) のように「同時進行」か「逆接」かの判断が難しいケースもある。

- (5) 加藤文太郎は秋の海を見ながら、思いは日本  
海に飛んでいた。

どのような要因により「同時進行」か「逆接」かの解釈が決まり、特にどのような場合に「逆接」の推論を読み取って解釈しているのかという問いが、本稿での焦点である。

## 2.3 談話関係コーパス

英語の談話関係コーパスの研究として代表的なものに、Penn Discourse Treebank (PDTB) [1, 11] がある。PDTB-3 [12] では、明示的な談話関係と非明示的な談話関係が区別され、明示的な談話関係については接続表現（談話標識）ごとに談話関係の単位 (Arg1, Arg2) と談話関係ラベルが付与されている。譲歩・逆接に対応する談話関係としては、CONCESSION と CONTRAST が区別され、COMPARISON 関係の下位クラスに分類されている。この CONCESSION と CONTRAST の判別は容易ではないことが報告されている [12]。

日本語の談話関係コーパスとしては、Web ページの冒頭3文を収集したコーパス（京都大学ウェブ文書リードコーパス）に対して、PDTB の方針を参考とし、談話関係タグを付与したものが [2] がある。談話関係タグとしては、「対比」と「逆接」が区別されているが、「が」のように同一の接続表現が対比と逆接の用法をもつ場合、その判別基準は必ずしも明確ではない。また、談話単位 (Arg1, Arg2) は自動認識によって設定されているため、言語学的な節単位とは必ずしも対応していない。

本研究は、詳細な統語情報が付与された既存のツリーバンクを利用することで、談話単位の認定という問題を回避し、より言語学的に妥当な談話関係ラ

ベルの設計に向けて、逆接の談話関係アノテーションを試みる。

## 3 アノテーションの方針

### 3.1 使用したデータ

今回使用した文は全て Kainoki Treebank [13] から抽出した。ツリーの例を付録の図 1 に示す。Kainoki Treebank は品詞タグ及び統語構造が付与されており、単に「ながら」「つつ」の文字列を含む文を抽出するだけでなく、接続詞「ながら」「つつ」が取ること二つの項（これを Arg1, Arg2 と呼ぶ）が、埋め込み節を含むような複雑な文でも正確に自動的に特定できることからこのコーパスを使用した。例えば、図 1 の例では Arg1 と Arg2 はそれぞれ「(女性が) 事務能力を事務能力を発揮し」「女性が一家の中心となって働く」であるが、Arg1 は「つつ」直上の IP ノード、Arg2 は「つつ」節を直接支配する IP ノードに支配される部分となっている。

### 3.2 談話関係ラベルと判断基準

接続詞「ながら」「つつ」を含む文（それぞれ 1,325 文、239 文）について筆者ら含め 11 名で下記の通りラベル付けを行った。各例に 2 名のアノテータがラベルを付与した。

- (6) a. **逆接・譲歩**（同時進行の意味を伴っていても逆接・譲歩の意味もあれば逆接・譲歩のラベル付けを優先する）例：(3a), (4a)
- b. **同時進行**（逆接・譲歩の意味を伴わないもの）例：(3b), (4b)
- c. **イディオム**（慣用句的表現）  
例：残念ながら／今更ながら
- d. **対象外**（後件が省略されているもの）  
例：窓の外を見ながら

「ながらも」「つつも」は全て逆接・譲歩とし、「ながら」「つつ」を「ながらも」「つつも」で置き換えて文意が変わらない場合も逆接・譲歩とした。更に、より正確に逆接・譲歩の意味を特定するために、**前提される推論** (2.1 節参照) を以下のように書き出す作業を行った。

- (7) [Arg1 たばこは体によくないと分かり]つつ  
[Arg2 なかなかやめられない。]  
(480\_textbook\_djg\_intermediate;page.543)

表現	分類ラベル	件数
ながら	A. 逆接・譲歩	146
	B. 同時進行	1,046
	C. イディオム	37
	A/B	68
	対象外	28
つつ	A. 逆接・譲歩	29
	B. 同時進行	209
	A/B	1

表1 アノテーション結果

**前提される推論**：たばこは体によくないと分かっているなら、なかなかやめられないことはない（やめられるはず）だろう。

## 4 アノテーション結果と考察

各分類ラベルの件数を表1に示す。2名のアノテータの判定ラベルが異なる場合、第3のアノテータがチェックし、ラベルを決定した。それでも決定できない曖昧なケースはA/Bと分類した。2名のアノテーション一致率はKappa値で、それぞれ、「ながら」は0.72、「つつ」は0.36であった。

### 4.1 逆接・譲歩の特徴

「ながら」「つつ」が逆接を表すのは、典型的には述語が動静的か静的かが関与するものの、例外もある(2.2節)。ここでは今回のアノテーションで逆接・譲歩とラベル付されたものに共通してみられる傾向があるかどうかについて考察する。

「ながら」「つつ」の用法に影響を及ぼしうると考えられる以下の特徴について、ラベルごとにそれぞれ当てはまるものの件数を数え、表にまとめた(付録の表2)。なお、それぞれの特徴はKainoki Treebankから自動的に抽出してカウントした<sup>3)</sup>。

- (8)
- 状態性アスペクト (ASP) : Arg1 の述部が「(～て) いる」「(～て) ある」「～ておく」
  - 否定 (NEG) : Arg2 が否定を含む
  - モーダル (MOD) : Arg2 が「はず」「べき」等モーダル要素を含む
  - 接続詞 (CNJ) : Arg2 が「しかし」「一方」などの接続詞を含む

3) 例えば、ASPの特徴をもつ「ながら」文の場合、Kainoki Treebankのウェブサイトにある検索インターフェイスでP-CONN <(ながら, おき|い|あり)>という検索式を用いて抽出できる。

- 取り立て詞 (TOR) : Arg2 が「も」「は」「さえ」などの取り立て詞を含む
- 疑問文 (QUE) : Arg2 が疑問文
- 命令文 (IMP) : Arg2 が命令文

まず、Arg1の述部が「(～て) いる」「(～て) ある」「～ておく」など静的状態を表すものである場合、逆接の用法である傾向が強いことが分かる(31/146対3/1046)<sup>4)</sup>。また、Arg2に否定が現れる場合も、比較的逆接の用法になりやすい傾向が見られる(21/146対13/1046)。その他には特に目立った特徴は見られないが、疑問(QUE)に関しては、逆接・譲歩の場合ほとんど全てが以下のような修辞(反語的)疑問文であった点が特徴的である。

- (9) [Arg1 敵軍撃滅という偽情報を国民に流し続けて煽動し]ながら [Arg2 何をいうか]、と怒りをおさえがたいが、国民(女も)が戦意に燃えていたのは事実であろう。(1037\_fiction.onna.to.senso)

「つつ」に関しては全体件数も少なく、目立った特徴は見られなかった。

### 4.2 前提される推論が書き出しにくい例

逆接・譲歩とは前提される推論(直接型の場合「Pならば-Q」、間接型の場合「PならばR、かつQならば-R」)が書き出せるものとした(2.1節)。しかし、実際に実例を見ていくと、判定が難しいものや、一見すると逆接と思われるもののいざ前提される推論を書き出そうとすると難しいものなど、興味深い例が見られた。

**A. 前提のモーダルの種類** 前提される推論を考える際、明示的には表れていない隠れたモーダルを補いながら考えなくてはならない場合がある。例えば、以下の例で、機械的に「Pならば-Q」というテンプレートに当てはめて前提される推論を書き出そうとすると(10a)となるが、「大抵～」というような物事の典型的な起こり方を示す(stereotypical)モダリティというよりは、(10b)のような義務的(deontic)なモダリティの意味合いが強い。

4) 今回、典型的な動作動詞ではない「思う」「知る」など思考や認識を表す動詞についてはカウントしなかったが、これらの動詞は「(～て) いる」の形をとらなくても静的状態を表しているとも考えられるため、それらの動詞についても「逆接」「同時進行」のいずれかになりやすいなどの傾向を調べる必要がある。動詞の種類の詳細に関しては今後の課題とする。

(10) [Arg1 遺家族とあれば一層保護を加うべき任にあり] ながら、[Arg2 色と慾の二筋道をかける] など実に言語道断です。(454\_aozora\_Oda-1976-2)

- a. 一層保護を加うべき任にあるならば、**大抵**色と慾の二筋道をかけない。
- b. 一層保護を加うべき任にあるならば、色と慾の二筋道をかけるべきではない。

**B. 同時進行と逆接の二分類の妥当性** 「同時進行」と「逆接」は相互排他的なものではなく、「同時進行」でもあり「逆接」でもあるケースもある。例えば、以下の例では「ながら」に「も」が後続する逆接用法であるが、叫んだのは逃げ終わったあ後ではなく逃げている最中とも解釈可能である。

(11) [Arg1 逃げていき] ながらも、[Arg2 「ひゃあ！」と叫んだ。] (421\_aozora\_Harada-1960)

典型的には「ながら」は同時進行と逆接のいずれかであるものの、「ながら」の時間性（同時か継起か）と表現主体の判断（順接か逆接か）はそれぞれ独立した軸であるため、それらの組み合わせの結果、(ア)「同時・順接」(イ)「同時・逆接」(ウ)「継起・順接」(エ)「継起・逆接」の4パターンあるとの見方 [10] もある。今回のアノテーションにおいても、上記のような例が見つかることから、同時進行と逆接という二分類が単純な相互排他的な二つのカテゴリーではないということが窺える。

**C. 対比** 逆接は対比の意味も伴っていることが多く、その二つを切り分けるのは容易ではない (2.3 節)。PDTB-3 においては CONCESSION と CONTRAST のラベル付けに関する判断基準が決められており、CONCESSION でもあり CONTRAST でもある場合は、前者を優先してラベル付けすることになっている [14]。

今回のアノテーションにおいても、そのような逆接と対比の切り分けの難しさが窺えた。例えば、以下の例では「*P* ならば  $\neg Q$ 」というテンプレートに当てはめて前提を書き出すと (12a) のようになるが、そもそも「おだてる」が本心から相手を尊敬して褒め称えるというよりは、何か下心があってことさらに褒めて相手をいい気分させるというニュアンスがあることから、前提は不自然に感じられる。

(12) 「本土の人って、[Arg1 沖縄をおだて] ながら、[Arg2 ほんとうはお人好しのバカだと思ってるでしょ]」 (494\_fiction\_onna\_to\_senso)

- a. 沖縄をおだてるなら、**大抵**はほんとうはお人好しのバカだと思わない。
- b. (表面的には) 沖縄をおだて、**ほんとう**はお人好しのバカだと思っている。

これを純粋な対比であるとするれば、(12b) のように「表面的には」vs. 「ほんとうは」、「おだてる」vs. 「お人好しのバカだと思っている」という二点において対立する対比の関係と見ることができる。

**D. 累加** 最後に累加の例を挙げる。

(13) 東北大学院工学研究科の風間基樹教授(地盤工学・地震工学)は「[Arg1 行政が作る防災マップを参考にし] つつも、[Arg2 普段から地域の地形的特性を把握して、非常時の対策を考える必要がある]」と指摘する。(27\_news\_KAHOKU\_184;K201406140A0T10XX00001)

- a. 行政が作る防災マップを参考にするなら、**大抵**は普段から地域の地形的特性を把握して、非常時の対策を考える必要がない。
- b. 行政が作る防災マップを参考にし、**そしてそれだけではなく**普段から地域の地形的特性を把握して、非常時の対策を考える必要がある。

上記の例では、テンプレートに当てはめただけの前提 (13a) は不自然であり、(13b) のように累加の意味があると考えたほうが自然に解釈できる。

逆接と累加の関係については、例えば、英語の接続詞 *but* で接続された節の後件に累加的表現の *too* が現れうるということが知られており、*too* の持つ前提がどのように *but* の表す前提や前件の意味と相互的に作用するのかについての研究がある [5]。今回のアノテーションでも、逆接と累加とが言語学的に興味深い関係を示す例が見られた。

## 5 おわりに

本研究では、前提される推論という観点から、談話関係が曖昧である「ながら」「つつ」という接続表現に着目し、逆接の談話関係アノテーションを行った。前提される推論を明示的に書き出すことで、逆接の談話関係とその用法の広がりが見えやすくなった。今後の課題としては、「が」「のに」など様々な逆接・譲歩の意味を伴う接続表現を対象としたアノテーションを行うことなどが考えられる。

## 謝辞

本研究は、JST CREST、JP-MJCR2114 の支援を受けたものである。

## 参考文献

- [1] Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Milt-sakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The Penn Discourse TreeBank 2.0. In **Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)**, Marrakech, Morocco, May 2008. European Language Resources Association (ELRA).
- [2] 岸本裕大, 村脇有吾, 河原大輔, 黒橋禎夫. 日本語談話関係解析: タスク設計・談話標識の自動認識・コーパスアノテーション. 自然言語処理, Vol. 27, No. 4, pp. 889–931, 2020.
- [3] 衣畑智秀. 日本語の「逆接」の接続助詞について: 情報の質と処理単位を軸に. 日本語科学, Vol. 17, pp. 47–64, 2005.
- [4] Mitsuko Narita Izutsu. Contrast, concessive, and corrective: Toward a comprehensive study of opposition relations. **Journal of Pragmatics**, Vol. 40, pp. 646–675, 2008.
- [5] Grégoire Winterstein. What but-sentences argue for: An argumentative analysis of but. **Lingua**, Vol. 122, No. 15, pp. 1864–1885, 2012.
- [6] 野矢茂樹. 「しかし」の論理. 言語, Vol. 32, No. 3, 2003.
- [7] 坂原茂. 日常言語の推論. 東京大学出版会, 1985.
- [8] 小泉保. 譲歩文について. 言語研究, Vol. 91, pp. 1–14, 1987.
- [9] 日本語記述文法研究会. 現代日本語文法 6. くろしお出版, 2008.
- [10] 村木新次郎. 語彙論と文法論と. ひつじ書房, 2019.
- [11] Rashmi Prasad, Bonnie Webber, and Aravind Joshi. Reflections on the Penn Discourse Treebank, comparable corpora, and complementary annotation. **Computational Linguistics**, Vol. 40, No. 4, pp. 921–950, 2014.
- [12] Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. The Penn Discourse Treebank 3.0 annotation manual, 2019. University of Pennsylvania.
- [13] Ed Kainoki. The Kainoki Treebank – a parsed corpus of contemporary Japanese, 2022. <https://kainoki.github.io> (accessed September 2022).
- [14] Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Milt-sakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The Penn Discourse TreeBank 2.0. In **Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)**, Marrakech, Morocco, May 2008. European Language Resources Association (ELRA).

# 付録

Kainoki Treebank のツリーの例を以下に示す。

ID@Kainoki\_185\_essay\_Hepburn\_no\_shoogai\_to\_nihongo2

女性が事務能力を発揮しつつ、一家の中心となって働くさまは、日本の少女少女たちを驚かせた。

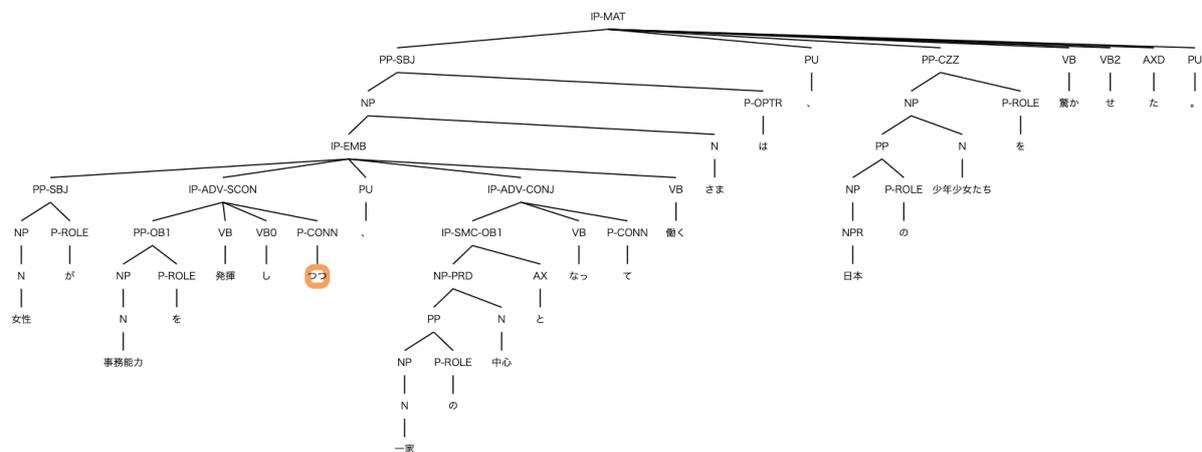


図1 Kainoki Treebank のツリーの例

各分類ラベルが付与された言語的特徴の統計情報を以下に示す。

表2 アノテーション結果の言語的特徴

表現	分類ラベル	総数	ASP	NEG	MOD	CNJ	TOR	QUE	IMP
ながら	A. 逆接・譲歩	146	31	21	3	8	3	6	0
	B. 同時進行	1046	3	13	7	4	6	2	1
	C. イディオム	37	0	0	1	1	0	0	0
	A/B	68	1	2	2	2	0	0	0
つつ	A. 逆接・譲歩	29	0	1	1	1	0	0	0
	B. 同時進行	209	0	1	0	0	0	0	0
	A/B	1	0	0	0	0	0	0	0