

時間関係に基づくテ形節の用法分類

野口咲帆 戸次大介

お茶の水女子大学

{noguchi.sakiho, bekki}@is.ocha.ac.jp

概要

テ形節は日本語において頻出する形式である。テ形節の用法を判定できれば、テ形節と後続する節の意味関係を考慮した意味解析が可能となるが、その基礎となるのはテ形節の用法分類と、それに基づくテキストアノテーションである。日本語学においてはこれまで多様なテ形節の用法分類がなされてきた。ところが先行研究の用法分類をもとにアノテーションを試みると、一つの表現に複数の用法が当てはまってしまう、といった問題がある。そこで本研究ではアノテーションを念頭に置いたテ形節の用法分類を提案する。本研究の用法分類はテ形節と主節の時間関係を軸に行い、その中で継続動詞／瞬間動詞のテ形で形成されたテ形節が「動作の継続」／「結果残存」を表す場合に着目した分類を設けた。

1 はじめに

テ形節には複数の用法が存在することが知られている。

- (1) コップが机から落ちて割れた
- (2) このイチゴは赤くて大きい

文(1)では「机から落ちて」、文(2)では「このイチゴは赤くて」の部分でテ形節である。テ形節と主節の意味関係を比較すると、文(1)はコップが落ちたことと割れたことの間で順序関係が成り立ち、対して文(2)はイチゴが赤いことと大きいことに並列関係が成り立っている。このようなテ形節と主節の意味関係を考慮すれば、より適切な意味解析を行うことができる。その意味において、テ形節の用法分類は自然言語処理分野の重要なタスクの一つと言える。

本研究の背景にある目標は、このような意味関係を判別した上で意味合成を行い、テ形節を含む文のより適切な意味表示を得ることである。その部分タスクとしてテ形節の自動用法判定器を作成するために、本研究では用法分類およびアノテーションガイドライン作成に取り組んでいる。

2 先行研究

2.1 森田 (1975) [1]

日本語学におけるテ形節の用法分類の例として、森田 (1975) [1] を挙げる。森田 (1975) [1] では、テ形節の用法は前件と後件の間の意味関係によって定めるとしている。各用法の説明として、その特徴と例文、加えてその用法のテ形節と同じ意味になる表現が述べられている。森田 (1975) [1] で提案された 8 分類を表 1 に示す。

表 1 森田 (1975) [1] の用法分類と同じ意味の表現

用法	同じ意味の表現
並列	ソシテ
対比	ソシテ一方、～ガ
同時進行	～ナガラ
順序	ソレカラ、～テカラ
原因・理由	～カラ、～ノデ
手段・方法	～ニヨッテ、～デ
逆接	～ノニ、～ガ、～ニモカカワラズ
仮定の結果	～ト、～バ

用法アノテーションにおける課題点 テ形節についての先行研究は必ずしも実テキストのアノテーションを目的としたものではないため、用法分類の判断基準が明確に示されていないことがある。先行研究をもとに用法を判断する際の課題点を 2 点述べる。

1 点目は、非専門家にとって用法を判断することが難しい場合があることである。森田 (1975) [1] では、〈同時進行〉(以降、用法名を〈〉内に記す)のテ形節を含む文の前件と後件の間には「～ナガラ」の関係が成り立つ(文(3a))としている。

- (3) a. 飛行機が煙を吐いて墜落した〈同時進行〉
b. 飛行機が煙を吐きながら墜落した

その上で、「～ナガラ」は状態性を表す前件に接続しない点でテ形節と異なり、言い換えができない場合があると述べている。例として、文(4a)を文

(4b) に言い換えると不自然になる。

- (4) a. 立って話しましょう〈同時進行〉
b. 立ちながら話しましょう

したがって〈同時進行〉のテ形節であっても、「～ナガラ」へ置き換えられるかという基準で判断すると正しく〈同時進行〉に分類されない場合がある。

2つ目は、1つのテ形節に複数の用法の特徴が当てはまる場合があることである。森田(1975)[1]ではそれぞれ「～テカラ」が〈順序〉、「～ノデ」が〈原因・理由〉のテ形節と同じ意味の表現として挙げられている。「～テカラ」「～ノデ」への言い換えができるかどうかでテ形従属節を異なる用法へ分類しようと試みるとする。その場合、「雨が降って、水たまりができた」という文のテ形従属節はどちらの言い換えも可能であるため、用法は1つに特定できない。

用法分類とアノテーションガイドラインの作成にあたって、以上の問題点を解決する必要があると考えた。

2.2 野口ら(2023)[2]

野口ら(2023)[2]では、テ形節の用法分類と言語学的テストを用いたアノテーションガイドライン[3][4]を設計した。用法は〈並列〉〈逆接〉〈仮定〉〈原因〉〈手段〉〈目的〉〈状態〉〈継起〉の8分類である。言語学的テストとは、文に対して対象表現の置き換えや語順の入れ替えといった操作を行い文の容認性や適切性を判定するもので、アノテーション時の判断の揺れが軽減されることが期待される。野口ら(2023)[2]のアノテーションは、用法ごとに設定された言語学的テストを指定の順番で試し、言語学的テストを通った時点でその用法をアノテーションするという手順で行われる。そのため必ず1つの用法がアノテーションされる。

言語学的テストの導入により、非専門家にとって用法の判断が難しいという課題点は克服できたと考えられる。しかし1つのテ形節に複数の用法の特徴が当てはまるという課題点は十分に解消されたとは言えない。野口ら(2023)[2]では1つの用法だけがアノテーションされるものの、実際には複数の用法の特徴が当てはまることに変わりないからである。その原因として、用法の判断軸が時間関係や因果関係などにまたがっており1つに定まっていないことが考えられる。この課題点の解決には野口ら(2023)[2]のようなアノテーション方法の工夫だけ

ではなく、用法の設計を見直すことが必要だと思われる。

2.3 金子(2020)[5]

金子(2020)[5]は日本語談話関係の分類体系と客観的な決定手順を定義している。談話関係には「命題間の時間的な関係」「明示的に因果を主張する関係」「論理的な関係」というカテゴリーが設けられており、カテゴリーごとに用意されたラベルが述部対に付与される。

金子(2020)[5]の特徴の1つは、解釈が揺れる「意図に関する関係」は排除し、「明示的な情報により解釈が一意に定まる関係」のみに絞って談話関係を定義していることである。「意図に関する関係」の例として「毎日ブドウにありがとう、と言って、立派なブドウが育った」という文の節間に因果関係があるかの判断は人それぞれであることを述べている。さらにこのような判断の揺れは談話関係の分析を難しくすると指摘している。この指摘と野口ら(2023)[2]の課題点を受け、本研究ではテ形節の時間関係を軸とした用法分類を行うこととした。

また、金子(2020)[5]では談話関係アノテーションの前段階として述部に event, state, modality のいずれかを付与する方法が定義されている。動詞には event または state を付与するとし、その判定には金田一(1950)[6]による動詞の分類に基づいた動詞の eventuality 判定テストを作成している。

2.4 金田一(1950)[6]

金田一(1950)[6]は日本語動詞をアスペクトに関する特徴から4分類している。主文における接尾語「ている」との接続可能性、または「ている」形式をとる際の意味について以下の特徴が述べられた。

- 状態動詞。「ている」を接続できない「ある」「いる」
- 継続動詞。動作の継続を表す「読む」「歩く」
- 瞬間動詞(変化動詞)。結果の残存を表す「死ぬ」「消える」
- 第四種動詞。常に「ている」を接続して用いられる「そびえる」「ありふれる」

金子(2020)[5]の動詞の eventuality 判定テストはこの特徴を用いて作成されており、継続動詞と瞬間動詞は event だと判定される。

ここで、テ形節の述部が継続動詞である文(5)と瞬間動詞である文(6)の意味を考える。

表2 述部の分類とテ形節の用法分類の対応

主節\テ形節	event	state	継続動詞+teiru	瞬間動詞+teiru
event	動作の並列/前後	前後/背景	同時	背景
state	前後	状態の並列	状態の並列	状態の並列

- (5) 歌を歌って帰る [1]
 (6) 腕を組んで話す [7]

文(5)は帰る間に歌を歌うという動作が継続しているという意味で、文(6)は腕を組んだ状態で話すという意味である。さらにそれぞれのテ形節の述部に「ている」を接続した意味を考える。すると「歌を歌う」は「歌を歌っている」となり、歌を歌うという動作の継続を表す。また、「腕を組んでいる」は腕を組む動作の結果として腕を組んだ状態であることを表す。このことから、文(5)と文(6)は前件がテ形節の述部に「ている」を接続した意味で捉えられていると考えられる。本研究ではこの考察を取り入れた用法分類を行う。

3 本研究の用法分類の提案

本研究ではテ形節と主節の時間関係を軸として、テ形節の用法を〈動作の並列〉〈状態の並列〉〈前後〉〈背景〉〈同時〉の5つに分類する。用法の詳細は後述する。

テ形節の用法分類のために、テ形節と主節の述部を分類し、節間の時間関係を考察した。述部の分類とテ形節の用法の対応を表2に表す。

1行目の event, state, 継続動詞+teiru, 瞬間動詞+teiru はテ形節の述部の分類であり、1列目の event, state は主節の述部の分類である。本章では述部の分類について説明した後、テ形節の用法の特徴を説明する。

3.1 述部の分類

event/state は述部の eventuality が event/state であることを表す。eventuality の判定方法は金子(2020) [5]で定義された方法を用いる。ただし、金子(2020) [5]においてはモダリティ表現に対して独立して modality というタグが付与されるが、現時点では、構造が [[テ形節 主節] モダリティ] か [[テ形節 [主節 モダリティ]]] かを判断する方法を提示できていないという問題が残っており、ここではモダリティの分析は今後の課題とする。

継続動詞+teiru, 瞬間動詞+teiru は本研究で新しく設けた分類である。継続動詞+teiru と瞬間動詞+teiru

をまとめて「動詞+teiru」と表記する。動詞+teiru はテ形節の述部が継続動詞や瞬間動詞である場合、かつ、「動作の継続」もしくは「結果残存」の意味で捉えられている場合に付与する。テ形節の述部が継続動詞や瞬間動詞であっても、「ている」形式の意味で捉えられていなければ event に分類される。

テ形節の述部が継続動詞または瞬間動詞である場合の述部の分類の判断方法は4章で議論する。

3.2 テ形節の用法

テ形節の用法の特徴を説明する。以降で、テ形節と主節が表す内容を単に前件・後件とする。前件/後件の出来事時とは、述部の分類が event であるときはその event が生じてから終わるまでの時区間、state であるときはその state が成立する時区間、継続動詞+teiru であるときは継続動詞が表す動作が継続している時区間、瞬間動詞+teiru であるときは瞬間動詞が表す動作が終わりその結果が残存している時区間を指すとする。

動作の並列 与えられた時間幅において前件と後件が成り立つ。いわゆる Parallel な関係である。前件と後件がどちらも成立することを表し、2つの事象の時間的配置については言及しない。テ形節の述部と主節の述部はどちらも event である。

- (7) おじいさんは山へ芝刈りに行って、おばあさんは川へ洗濯に行った

状態の並列 与えられた時点において前件と後件が成り立つ。テ形節の述部と主節の述部はどちらも state であるか、テ形節の述部が動詞+teiru で主節の述部が state である。

- (8) リンゴがあって、ミカンがある
 (9) 交通は混乱して、人心は陰悪である [8]
 (テ形節の述部が継続動詞+teiru)

前後 前件の出来事時の開始時点が後件の出来事時の開始時点に先行する。テ形節の述部と主節の述部が event または state である。ただし前件と後件がどちらも state である場合を除く。

- (10) 彼は立ち上がって外出のしたくを始めた [1]

背景 前件の出来事時が後件の出来事時を包含する。テ形節の述部が state または瞬間動詞+teiru で、主節の述部が event である。テ形節の述部が瞬間動詞+teiru であるとき、瞬間動詞で表される動作が後件の動作より時間的に先に位置する。

- (11) お金がなくて困った [9]
- (12) 眼鏡をかけて本を読む
(テ形節の述部が瞬間動詞+teiru)

同時 後件の出来事時が前件の出来事時を包含する。テ形節の述部が継続動詞+teiru で、主節の述部が event である。

- (13) テープを聞いて練習します [1]

4 用法の判断

テ形節を3.2章で提案した用法に分類する方法について議論する。テ形節の述部が継続動詞または瞬間動詞でない場合は、金子 (2020) [5] で定義された eventuality 判定方法に従ってテ形節の述部と主節の述部に対して event/state の判定を行い、前件と後件の時間関係を判断すれば用法が一意に定まる見込みである。しかしテ形節の述部と主節の述部の組が以下の場合には、テ形節の述部が動詞+teiru, event のどちらであるか判断が必要になる。

- 1. (継続動詞, 分類が event)
- 2. (瞬間動詞, 分類が event)
- 3. (継続動詞, 分類が state)
- 4. (瞬間動詞, 分類が state)

現時点でケース 1~4 のテ形節の述部の分類は、動詞+teiru であるという仮定の元で当てはまる用法が、文の意味と照らし合わせて適切であれば動詞+teiru であり、そうでなければ event であるとしている。例としてケース 1 の場合、つまりテ形節の述部が継続動詞かつ主節の述部が event である以下の 2 文のテ形節の述部を分類する。

ケース 1

- (5) 歌を歌って帰る
- (14) インスタント食品にお湯を注いで、2, 3分待ちます [10]¹⁾

文 (5) のテ形節の述部を継続動詞+teiru だと仮定する。すると文 (5) のテ形節が同時に分類されるはずである。文 (5) は帰る間に歌っているという同時進

1) BCCWJ サンプル ID:PB26.00004 山田 みどり 『「この人から買いたい」と思わせるプロの接客!』

行を表し、テ形節の用法が同時で適切であるため、テ形節の述部は継続動詞+teiru に分類される。一方で文 (14) のテ形節の述部を継続動詞+teiru だとすると、仮定の上で文 (14) のテ形節が同時に分類され、2, 3分待つことの開始時点がお湯を注ぐ動作より先ということになる。文 (14) はお湯を注いでから 2, 3分待つという意味であり矛盾が生じるため、テ形節の述部は event に分類される。

ケース 2

ケース 1 の場合には、以上のように前件と後件の開始時点の前後を考慮することで比較的容易にテ形節の述部を判定できると見られる。しかしその他のケース 2~4 については現段階で判定方法を検討中である。ケース 2 のときにテ形節の述部を判定する際の難しさを述べる。

難しさの原因は、テ形節の用法が〈背景〉と〈前後〉であるときに前件と後件の時間的配置は同じになることだと考えられる。なぜならテ形節の述部を event に分類することが適切である場合でも、テ形節の述部が継続動詞+teiru でテ形節の用法が〈背景〉だという仮定を拒否しづらいからである。以下の文を例にする。

- (15) a. A さんは今日図書館へ行って
b. 帰り道で友達に会った

節 (15a) の動作の結果維持は、図書館に A さんが到着した状態の維持、つまり図書館にいたことだと考えられる。節 (15b) によると A さんが友達に会ったのは図書館の外であるため、この文のテ形節の用法は〈背景〉ではないと判断でき、テ形節の述部は event に分類される。しかしながら、もし節 (15b) に「帰り道で」という情報がなかったら以上のように述部の分類を判断できるだろうか。その場合はテ形節の動作自体と動作の結果維持のどちらに焦点が置かれているかで判断をする他なく、テ形節の用法や述部の分類を判断することは難しい。

5 まとめ

本研究では「述部の分類」と「テ形節と主節の時間関係の判断」を利用したテ形節の用法分類を提案した。テ形節の述部の分類として、非明示的に「ている」形式で意味を解釈することを表す「動詞+teiru」を新たに設けた。述部の分類の判定方法には定まっていない箇所があるため、今後は判定方法の定義を行いアノテーションを実施する予定である。

謝辞

本研究は、JST CREST JPMJCR20D2, および JSPS 科研費 JP23H03452 の助成を受けたものである。

参考文献

- [1] 森田良行. 複文の文型練習—「たら」「て」を含む文型を中心に—. 講座日本語教育, No. 11, pp. 1–15, 1975.
- [2] 野口咲帆, 田中リベカ, 戸次大介. テ形従属節の用法判定のための用法分類アノテーション. 人工知能学会全国大会論文集 第 37 回 (2023), pp. 2E6GS601–2E6GS601. 一般社団法人人工知能学会, 2023.
- [3] 田中リベカ, 小池恵里子, 戸次大介, 川添愛. 言語学的テストに基づく意味アノテーションのガイドライン設計—確実性判断に関わる表現を中心に—. 言語処理学会第 18 回年次大会発表論文集, pp. 401–404, 2012.
- [4] 川添愛, 田中リベカ, 戸次大介. Mcn コーパス：モダリティ関連表現の曖昧性解消のためのアノテーションと言語学的テストの利用. テキストアノテーションワークショップ・コンテスト, 国立情報学研究所, 2012.
- [5] 金子貴美. 日本語談話関係認識のための理論とコーパス構築. PhD thesis, お茶の水女子大学, 2020.
- [6] 金田一春彦. 國語動詞の一分類. 言語研究, Vol. 1950, No. 15, pp. 48–63, 1950.
- [7] 森田良行. 基礎日本語 2. 角川小辞典, 1980.
- [8] 仁田義雄. シテ形接続をめぐって. 複文の研究, pp. 87–126, 1995.
- [9] 益岡隆志, 田窪行則. 基礎日本語文法-改訂版, 1992.
- [10] Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. Balanced corpus of contemporary written Japanese. **Lang Resources & Evaluation**, Vol. 48, pp. 345–371, 2014.