

1 はじめに

自然言語の意味の理解において、テンス・アスペクト体系の分析は不可欠である。テンスは文中の出来事の時刻を発話時を基準として位置づける文法形式を指し、アスペクトは出来事の中の様々な局面を表す文法形式を指す。形式意味論による自然言語の研究はさかんに行われているが、テンス・アスペクトはほぼ全ての文に現れるにも関わらず、これらがもたらす言語現象を網羅するような理論は確立されていない。本研究では形式意味論の体系のひとつである依存型意味論 [1] に時間的な概念を導入し、日本語のテンス・アスペクト体系の分析を行った。

2 依存型意味論

依存型意味論 (DTS) は、依存型理論に基づいた自然言語の証明論的意味論である。従って、自然言語の意味について証明論的な推論が可能となる。文 A から文 B が推論可能であるという関係を、文 A の意味表示から文 B の意味表示への演繹可能性として捉える。

依存型理論は項に依存した型を記述する事が可能であり、重要な型の構成子として  $\Pi$  と  $\Sigma$  をもつ。  $\Pi$  は「すべての A であるような  $x$  について  $B(x)$  が成り立つ」ことを意味し、全称量化のように振る舞う。  $\Sigma$  は「 $B(x)$  であるような  $x$  が存在する」ことを意味し、存在量化のように振る舞う。それぞれ、型 B が項  $x$  に依存しない場合、一階述語論理における implication ( $\rightarrow$ ) と conjunction ( $\wedge$ ) に対応する。

$$(x:A) \rightarrow B(x) \quad \begin{matrix} x:A \\ B(x) \end{matrix}$$

$\Pi$  型  $\Sigma$  型

また、依存型理論では、投射  $\pi_1$ 、 $\pi_2$  が定義されており、以下のような規則をもつ。

$$y : \begin{matrix} x:A \\ B(x) \end{matrix} \quad (\Sigma E) \frac{y : \begin{matrix} x:A \\ B(x) \end{matrix}}{\pi_1(y) : A} \quad (\Sigma E) \frac{y : \begin{matrix} x:A \\ B(x) \end{matrix}}{\pi_2(y) : B[\pi_1(y)/x]}$$

自然言語の意味表示部分に依存型理論の型を対応づけることで、先行する文脈や句を考慮した意味表示を記述する事が可能となる。

3 日本語のテンス

日本語ではテンスに関わる言語表現として助動詞 [タ]、終止形活用語尾 [ル] が挙げられる。一般に、[タ] は過去、[ル] は非過去を表すとされる。Kamp ら [2] では eventuality (出来事) は event が state のどちらかであるとしている。中村 [3] によると、非過去の場合に、state の時刻が発話時と重なる事が出来るのに対し、event は発話時と重なることができないということから、event/state 間にはテンスにおける振る舞いにおいて違いがある。従って、文書中の出来事の時間的な振る舞いを把握するには、[タ][ル] などのテンスの情報をもたらす語句の分析だけではなく eventuality の判定が必要となる。以下 3.1 節では eventuality の判定方法を提案し、3.2 節ではテンスに係る語句のひとつである助動詞 [タ] について述べる。

表 1: eventuality の event/state の分類

品詞分類	eventuality <sup>1</sup>	特徴	例
動詞	event	動態動詞	動きを表す動詞
	state	状態動詞	状態を表す動詞
連用形複合動詞	event		
形容詞	state		
状詞	state		
動詞性接尾語	event	形式動詞	動きを表す形式動詞
	state		状態を表す形式動詞
	event	動詞の連用形に接続	アスペクトに関係
			完遂の意味を表す
			不遂行の意味を表す
			その他
	state	動詞のテ形に接続	アスペクトに関係
	*		授受に関係する
	*		その他
	event	形容詞/名詞の語幹に接続	
	event	普通名詞に接続	
	*	態	尊敬
	event		受身
event		使役	
state		可能	
状詞性接尾語	state		
形容詞性接尾語	state	否定の意味をもつ	
		動詞の連用形に接続	
		タロウ接続形に接続	
		イ形容詞/状詞語幹に接続	
		名詞に接続	
		動詞の終止形に接続	
		テ形に接続	
		その他	
助動詞	*	過去の意味を表す	
	state	丁寧/否定/推量の意味を表す	

<sup>1</sup> 表中の [event/state/\*] はそれぞれ、[event] は event を指定する語句、[state] は state を指定する語句、[\*] は直前の eventuality を保持するような語句に与えられている。

3.1 eventuality

一般に文書中の eventuality は述語によって表され、動態述語は event、状態述語は state を表すとしている。しかし、以下に示す 4 種類の要素が接続する度に eventuality は変化するため、実テキスト上には event/state の判定が自明でないケースが存在する。

- a. アスペクトに関する表現
- b. 否定表現
- c. モダリティに関する表現
- d. テンスに関する表現

本研究ではこれらの各要素が述語と接続した際に event/state のどちらを指定するかを判断するために、戸次 [4] の品詞分類に基づき eventuality の event/state 分類表を作成した。表 1 はその抜粋である。分類表を作成したことにより、動詞、形容詞、状詞<sup>1</sup> などの単一の要素からなるフレーズだけでなく、述語語幹に様々な要素が接続するような複雑なフレーズに対しても event/state を特定することが可能となった。

3.2 [タ]

一般に助動詞 [タ] は eventuality (出来事) の時刻を過去に位置づけるはたらきをもつ。以下の例文 (1)(2) では、動態動詞 [食べる] により指定された event と、状詞 [サンマだ] により指定された state を過去に位置づけている。

- (1) 今日は 7 時に朝ご飯を食べた。
- (2) 昨日の晩ご飯はサンマだった。

しかし、助動詞 [タ] は eventuality を過去に位置づけるだけに留まらず、様々な意味を持つことが日本語学において知られている。

<sup>1</sup> 戸次 [4] では、状詞の概念は、形容動詞と、名詞 + 判定詞、副詞、連体詞にまたがる分類であるとしている。

表 2:[タ] ガイドラインのカテゴリと例文

直前の eventuality	カテゴリ	例文
event	過去	「8時に朝ご飯を食べた」「明日の決勝で勝ったチームがオリンピックへの出場権を得る」
	完了	「私の父は、ガンでもう死んだ」
	動作開始時	「野球の実況中継」イチロー、走った!」「(赤ちゃんを眺めていたら、たまたま笑い出した)あ、笑った!」
	ムード	「飯に走ったとします」
state	仮定	「さあ、行った行った!」「買った、買った!」
	命令	「太郎は自宅にいた」「昔は英語が話せた」
	過去	「父は既に3時間前に死んでいた」
	完了	「あ、ここにあった」「あや、こんなところに銭があった」
	発見	「(今日、CDを買った)ベートベンの「第九」だった」
	判明	「(誰だろう、と振り向いて)何だ、井之上さんでしたか?」
	思い出し	「(誰だろう、と振り向いて)何だ、井之上さんでしたか?」
	修正補強	「(誰だろう、と振り向いて)何だ、井之上さんでしたか?」
	仮定文	「ここに100万円あったとします。あなたなら何に使いますか?」
	正解	「正解は3番でした」
他	「ありがとうございました」「ご苦労様でした」	

- (3) さあ行った 行った!
- (4) あ、そういえば午後からミーティングだった。

これらは eventuality を過去に位置づける意味を持たない [タ] の例であり、(3) は命令のニュアンスを、(4) は思い出しのニュアンスをもたらす。つまり、統語上では同じ振る舞いをする助動詞 [タ] であるが、その用法によってそれぞれ異なる意味表示を与える必要がある。

本研究では、[タ] のもつ様々な用法を区別するために、MCN コーパスのガイドライン設計 [5] と井上による「タ」の分析 [6] に基づき、[タ] のガイドラインを作成した。全カテゴリ数は 14 種類であり、event に後接する [タ] には 5 種類、state に後接する [タ] には 9 種類のカテゴリを設けた。このうち eventuality を過去に位置づけるはたらきを持つカテゴリは、event に後接する [event-過去][event-完了]、state に後接する [state-過去][state-完了] の 4 種類であり、過去の意味を持たない [タ] については、[event-動作開始時] やムード (話し手の心的態度) を表すようなカテゴリ 10 種類を設けている。(3)(4) はどちらもムードを表す例であり、(3) は [event-命令]、(4) は [state-思い出し] のカテゴリに属する。表 2 は、各カテゴリに属する例文の一覧である。[タ] ガイドラインの詳細については宇津木ら [7] を参照されたい。

作成した [タ] ガイドラインについて、現代日本語書き言葉均衡コーパス<sup>2</sup> の書籍ドメインのテキストに含まれる [タ] に対しアノテーションを実施した。表 1 に従い、あらかじめ [タ] の直前の eventuality の属性情報 (event/state) を付与したテキストを使用した。アノテーション件数は 533 件、アノテーターは言語学専攻でない大学院生 2 名である。本作業における単純一致率は 0.80、Kappa 値は 0.65 であり、これらの値から、今回のアノテーションは比較的高い一致率で作業できたと考えられる。

#### 4 DTS による意味表示

$$\text{time} \stackrel{\text{def}}{=} \mathbb{Q} \times \mathbb{Q}$$

本研究では依存型意味論に時間的な概念を導入するため、新たに time 型を定義した。time 型は有理数の直積であるとし、time 型の項は、開始時刻 pivot と区間長 length からなる 2 つ組であるとする。従って、投射  $\pi_1$  と  $\pi_2$  を用いる事で、 $t: \text{time}$  について  $t$  の開始時刻と区間長を取り出す事ができる。つまり、 $\pi_1(t) = \text{pivot}(t)$ 、 $\pi_2(t) = \text{length}(t)$  となる。また、time 型に関する関数も定義した。表 3 にその一部を示す。

<sup>2</sup>[http://www.ninjal.ac.jp/corpus\\_center/bccw/](http://www.ninjal.ac.jp/corpus_center/bccw/)

表 3:時間に関する関数

$\text{dur} : \text{eventuality} \rightarrow \text{time}$	eventuality 型をもつ項を受け取ったらその項の時間を返す
$\text{end}(t) \equiv \text{pivot}(t) + \text{length}(t)$	$t$ の終了時刻
$t < t' \equiv \text{end}(t) < \text{pivot}(t')$	時間 $t$ が時間 $t'$ よりも過去に位置する
$t \sqsubseteq t' \equiv \begin{cases} \text{pivot}(t') < \text{pivot}(t) \\ \text{end}(t) < \text{end}(t') \end{cases}$	時間 $t$ が時間 $t'$ に含まれる

本研究では、時間的な概念を導入した依存型意味論を用いて、時間副詞、過去・非過去の意味をもたらす [タ][ル]、接尾語 [テイル][ハジメル] などテンス・アスペクトに関わる重要語句の語彙項目に意味表示を与えた。ここで、各語彙項目は戸次 [4] に従い、[音韻表示・統語範疇・意味表示] の 3 つ組で表現するとし、統語範疇には組み合わせ範疇文法を使用する。

#### 4.1 [タ]

ここでは [タ (過去)] に着目し解説する。以下に示す [タ (過去)] は 3.2 節での [タ] ガイドラインで「eventuality を過去に位置づける意味をもつ」と判断された [タ] である。ムードの意味をもたらす [タ] などの「eventuality を過去に位置づける」という意味を持たないカテゴリの [タ] については今後の課題とする。

$$\text{タ (過去)} \vdash S \setminus S : \lambda p. \lambda t. \lambda c. \left[ \begin{array}{c} ptc \\ t < st \end{array} \right]$$

st は発話時刻を指す。述語  $p$ 、時間  $t$ 、文脈  $c$  を受け取り、時間  $t$  が発話時よりも過去に位置することを意味している。たとえば、「太郎が走った」の意味表示は以下のように導出される。最終的な意味表示では、[太郎が走る] というイベントの時間  $t$  が発話時よりも過去に位置づけられていることを表している。

$$\frac{\frac{\frac{\text{太郎が}}{T/(T \setminus NP)} : \lambda p. \lambda t. \lambda c. \left[ \begin{array}{c} \text{走っ} \\ t = \text{dur}^*(e) \end{array} \right]}{<B} \quad \frac{\frac{\text{た}}{S \setminus S : \lambda p. \lambda t. \lambda c. \left[ \begin{array}{c} ptc \\ t < st \end{array} \right]}}{>}}{S : \lambda t. \lambda c. \left[ \begin{array}{c} e: \text{run}(taro) \\ t = \text{dur}^*(e) \\ t < st \end{array} \right]}$$

#### 5 おわりに

日本語のテンス・アスペクト体系を依存型意味論を用いて分析することが本研究の目的である。本稿では、助動詞 [タ] に注目し、時間に関わる語句の分析を行い、依存型意味論による意味表示を与えた。本研究では [タ] や [ル][テイル] の代表的な読みの意味表示を与えたが、その他の読みを含めて今後も引き続き様々な言語現象の分析と意味表示の構築を行っていく予定である。

#### 参考文献

- [1] Daisuke Bekki. Representing anaphora with dependent types. In *Logical Aspects of Computational Linguistics*, pp. 14–29. Springer, 2014.
- [2] Hans Kamp and Uwe Reyle. *From discourse to logic*. Dordrecht, 1993.
- [3] 中村ちどり. 日本語の時間表現, 第 14 巻. くろしお出版, 2001.
- [4] 戸次大介. 日本語文法の形式理論 - 活用体系・統語構造・意味合成 -, 第 24 巻. くろしお出版, 2010.
- [5] 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. *NLP2012*, pp. 401–404, 2012.
- [6] 井上優. 現代日本語の「タ」—主文末の「...タ」の意味について—. 「た」の言語学. ひつじ書房, 2001.
- [7] 宇津木舞香, 稲田和明, 金子貴美, 戸次大介, 乾健太郎. 形式意味論に基づく出来事関係認識に向けてリソース構築の展望とテンス「タ」のアノテーション. *NLP2015* (発表予定).