

## テンス・アスペクトを考慮した現代日本語の複文の生成

西田元樹 松本裕治

奈良先端科学技術大学院大学情報科学研究科

意味表現によって記述された複数の事象群の中から時間的關係のあるふたつの有意な事象を現代日本語における複文の形式で生成する手法について考察する。時間に関する情報を言語化するにあたってふたつの事象の発生時区間関係、参照時区間と発話時点の関係、時間結合語、語彙アスペクトから得られる制約条件を元にして妥当なテンス・アスペクト組合せを計算することについて述べる。これによって意味表現の内容を明確に表した時間に関する複文を生成することが可能になる。またテンスや参照地点について従来時点として捉えられてきたが時区間に拡張したことによってより精確な表現が可能となった。

## Proper Treatment of Tense and Aspect in Japanese Complex Temporal Sentence Generation

NISHIDA Motoki and MATSUMOTO Yuji

Graduate School of Information Science, Nara Institute of Science and Technology

This paper presents a method to generate Japanese complex sentences that reflect the temporal relation calculated from duration of two events. Events are represented by an HPSG-style feature structure with time interval information. we propose a computational model of Japanese tense, aspect, and time-taxis which are necessary to generate appropriate complex sentences.

### 1 はじめに

法的推論システムなどは時間軸上に出現する複数事象から有意事象を抽出し因果関係を導出することによってある決断を下すことを目的としている。決断を支持する根拠として事象の時間的な発生順序関係が重要であり、その推論過程をユーザに提示する必要がある。我々は時間的順序関係で結合された複数の(本報告内では二つに限定)事象を複文の形式で記述(生成)する方法について考察している。勿論複数の事象を記述するに該り複文以外の形式もありえる(例えば二つ以上の単文で記述するなど)が時間的結束性を強調するには複文の方が適切であると考えたためである。したがって複文以外の形式での生成を排除するものではない。

これまで時区間における時間的關係については [Allen84] のイベント計算をはじめ [Kowalski86,

McDermott82, Kamp93] など様々な理論が提案されている。また現代日本語を対象としてテンスとアスペクト情報から時間關係を抽出する研究も [江良ほか 93, 東ほか 96] などによって行われ、「*る・た・ている・ていた*」の選択を時相状況を導入することによって決定する [池田ほか 91] らの研究がある。

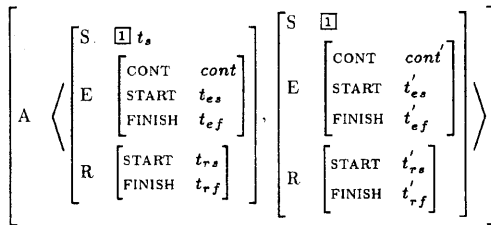
現代日本語において複文を生成するには、これらの時間關係に関する理論とテンス・アスペクトに関する研究を統合した上で日本語の複文構造に固有の制約条件を考える必要がある。

### 2 基本的な考え方

ここでは入力となる意味表現の記述、出力となる複文の形式及びテンス・アスペクト体系に対する基本的立場について述べる。

## 2.1 複数の事象の意味記述

複文を生成する為には何らかの意味表現によって記述された知識がシステムに与えられる必要がある。我々は[玉野ほか 96]で述べられている陳述の時間的構造をもとに図1のような構造から生成することを考察した。[玉野ほか 96]の時間的構造記述は[Reichenbach47]の体系を時区間に拡張した点に特徴があり事象の発生及び参照位置をより正確に把握することができるものである。各事象は発話時点(S)と事象発生時区間(E)及び照時区間(R)から構成される。EとRは時間軸上の始点と終点に関する情報が付加され、主たる事象・従たる事象の順にリストとして並べられるものとする。



(注)Sは発話時点,Rは参照時区間,Eは事象発生時区間。

図1: ふたつの事象の時間的構造を含む意味記述

## 2.2 時間従属複文

複文は従属文と主文により構成される。本研究では時間表現に限定した複文を対象とするので便宜上従属文は「いいおわり」の形の(狭義の)従属文と時間を表す結合語(時間結合語)から構成されるものと定義する。

## 2.3 時間結合語

時間結合語とは「まえに」「のちに」「ときに」「瞬間に」など従属文の文末を示すマーカーのうち時間に関する語を指す。時間幅を有するふたつの事象は[Allen84]に従えば13に分類にできる。この13の分類を[工藤 95]の時間順序関係(=タクシス)に準じて共起的な時間関係と継起的な時間関係に大別した(表1)。

## 2.4 現代日本語のテンスとアスペクト

従来現代日本語のテンス・アスペクトとして考察されてきたのは「する・した・している・していた」のパラダイグマティックな語彙対立であった(表2)。

	非過去	過去
完成相	する	した
継続相	している	していた

表2: 従来のテンス・アスペクト体系

ところが参照時区間と事象発生時区間との関係を考えると(図2)アスペクトを「完成相・開始相・継続相・完了相」の対立で捉えなければ事象の時間記述に対して不十分である。したがって我々は表3に示すテンス・アスペクト体系に従い事象の時間表現をめざした。

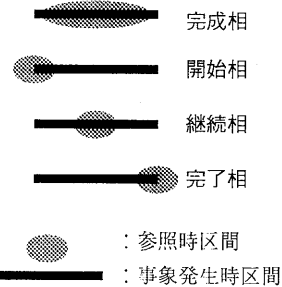


図2: 参照時区間と事象発生時区間の関係

	非過去	過去
完成相	する	した
開始相	しはじめる	しはじめた
継続相	している	していた
完了相	してしまう	してしまった

表3: 対象とするテンス・アスペクト体系

## 3 現代日本語固有の制約条件

ここでは2で述べたことを前提として現代日本語に固有な制約条件について考察する。現代日本語

タクシス	時間関係	時間結合語
継起	before	前(に)、前から
	after meets	後(に)、から、以来 直後(に)、瞬間(に)、なり やいなや、とたん
共起	met by	直前(に)、まぎわ(に)
	overlaps	間(に)、うち(に)、最中(に)
	overlaped by	◇
	during	時(に)、間(に)、うち(に) ころ(に)、瞬間(に)、最中(に)
	contains	◇
	starts	同時(に)
	started by	◇
	finishes finished by	同時(に)、まで(に)、時まで(に)
equal	◇ 時(に)、同時(に)	

表 1: 時間結合語とその分類

において時間従属複文を生成しようとした場合以下のような現代日本語固有の制約条件が挙挙できる。  
**継起タクシス制約** 継起タクシスは従属事象のテンスを一意に決定しアスペクトは継続相以外を要求する。

**共起タクシス制約** 共起タクシスは現代日本語の語彙アスペクトによって等価な別のテンス・アスペクトに置換することができる。

**時間結合語制約** 個々の時間結合語は主文に対して一定のテンス・アスペクトを要求する。

**参照時区間・発話時点の配置制約** ふたつのR時区間の間にSを配置することはできない。

以下ではこれらの制約条件について順に考察する。

### 3.1 継起タクシス制約

継起タクシスは従属事象のテンスを一意に決定してしまう。その変更は不可能である。例えば  $E_1$  「被申込者が承諾の通知を發した」という事象が  $E_2$  「申込者による申込の取消の通知が被申込者に到達した」という事象より後に発生している場合に

\* 「被申込者が承諾の通知を發した前に申込者による申込の取消の通知が被申込者に到達した。」

と記述することは不可能であり非過去の「發する」を使用しなければならない。

継起タクシスにおいては従属する事象のテンスは発話時点を基準として決定することはできず、ふたつの事象の時間関係によって決定される。before, met by の関係の場合には非過去のテンスが after, meets の関係の場合には過去のテンスが使用さ

れるのである(表 4)。現代日本語にはこのような継起タクシスにおけるテンスの制約が認められるので従属文として表現されるべき事象は正確に記述することが不可能となっている。

ふたつの事象の時間関係		従属文のテンス
before	met by	非過去
after	meets	過去

表 4: 継起タクシスにおける従属文のテンス

また継起タクシスにおいて従属文のアスペクトは継続相を採ることができない。この制約については現代日本語に固有の制約条件であるかどうか現在調査中である(少なくとも現代日本語と英語に対しては適用できそうである)。

### 3.2 共起タクシス制約

共起タクシスにおいてある種のテンス・アスペクト表現が他のテンス・アスペクト表現と置換可能である場合がある。たとえば

「家に帰った時、申込書が到達した」  
「家に帰っている時、申込書が到達した」  
「物品を引き渡す時に、書類を交付した」  
「物品を引き渡している時に、書類を交付した」

という文を考えると「帰った」と「帰っている」が「引き渡す」と「引き渡している」が等価であることが分かる。「帰った」は結果段階を捉えているが

ゆえに「帰っている」に置換することができ「引き渡す」は動作過程段階を捉えているがゆえに「引き渡している」に置換することができる。これらの語が置換可能であるのは現代日本語の語彙アスペクトに依拠している。

### 語彙アスペクト

表層に出現する用言自体は静的な状態 (states), 動的な活動 (activities), 一定幅の動作を得て達成されるもの (accomplishments), 瞬間に発生しその状態が保持されるもの (achievements) など様々な状況に関する情報を有している [Vendler67]. 語彙アスペクト (動作相, Aktionsart) はそのような用言の意味に出現する状況を規定したものである。これは用言が表示する状況 (表示状況), その前状況及び後状況からなりたっている。用言の表示状況は持続的なものと瞬間的なものに二分できそれらとその境界から語彙アスペクトを分類すると表 5 にしめす 10 種類に分類できる [金子 95, 東ほか 96].

前状況 before	開始点 start	表示状況 indicate	終了点 end	後状況 after
V	nil	V	nil	V
V	nil	V	end	not V
not V	start	V	end	not V
not V	start	V	end	not V(eff)
not V	start	V	nil	V(res)
not V	start	V	nil	V(eff)

前状況 before	表示状況 indicate	後状況 after
not V	V	not V
not V	V	not V(eff)
not V	V	V(res)
not V	V	V(eff)

表 5: 語彙アスペクト

表 5 に掲げられたもののうち後状況において効果が持続するものや結果状態にあるものについてはテンス・アスペクトを置換することが可能となる。ただし厳密に言えばテンス・アスペクトを置換した後ではそれ以前のものと同価ではない。厳密性が要求される文生成においては置換すべきではないが文生成においてはある程度の自由度が許されても良い。表現形態のひとつとして許容できる範囲内であろう。

### 3.3 時間結合語制約

タクシスは主に従属文におけるテンス・アスペクトを制約するものであった。これに対して時間結合語制約とは主文におけるテンス・アスペクトを制約するものである。またタクシスは時間結合語群全体に渡る制約条件であるのに対し時間結合語制約は個々の語が有している制約条件である。たとえば *during* の時間結合語中の非常に似通った「あいだ」と「あいだに」において「あいだに」は主文が完成相であることを要求するのに対し「あいだ」は継続相や存在動詞または完成相のうちの非内的限界性 (activities と等価) の動詞を要求する。個々の語に依存した制約であるのですべてを列挙することはここではしないが現代日本語の語彙アスペクトとの関連性を考察する余地は残っている。

### 3.4 参照時区間・発話時点の配置制約

$R_1$  時区間と  $R_2$  時区間の間に  $S$  時点が存在してはならない。西洋語では基本テンス構造 (BTS) における  $S \cdot R \cdot E$  の線型順序は時間結合法則 (RTC) によって複合テンス構造 (DTS) になった場合においても保持されていなければならない (複合テンス構造制約, CDTS) という制約がある [Honstein90] がこれと同様の制約である。ただし現代日本語においては時間結合法則による参照時点の関連付けは行わず単に  $R_1$  と  $R_2$  の間には  $S$  が配置されないだけである。これは西洋語に観察される「時制の一致」の現象を現代日本語にあてはめることができないからである。論理的にはふたつの  $R$  時区間の間に  $S$  を配置することは何ら問題はない。しかし現代日本語の枠組においてそのような配置に基づく複文を生成することは不可能なのである。

## 4 システムの概要

試作したシステムの概要図を図 3 に示す。システムは三つのモジュールとひとつのチェックルーチンより構成される。モジュールはタクシス計算・時間結合語を選択するモジュール, テンス・アスペクト計算を行い妥当なテンス・アスペクト組合せ群を出力するモジュール, BCG [Haruno et al.93] に基づいて単文の生成を行うモジュールから構成される。その他入力群が RTC 違反を犯していないかどうかを



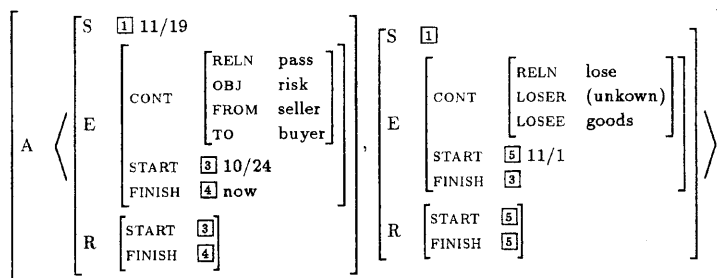


図 5: システムの入力例

## 5 まとめ

事象発生の時間的表現および参照を時点から時間幅に広げたことによって、時間に関する情報をより精確に入力することが可能となった。事象発生時区間関係(タクシス関係)、参照時区間関係、参照時区間・発話時点関係、時間結合情報、語彙アスペクト情報からさまざまな制約条件を導き出すことが可能となり複文における妥当なテンス・アスペクト組合せ、語彙選択が可能となった。また HPSG に基づく素性構造表現を採用したことによって解析と生成両方に共有できる文法体系が使用できるという利点が得られた。

時間関係は継起タクシスなどにおいては因果関係を暗に示す場合があり、その関連性については今後の課題とする。また時間副詞を含んだ文においては直接的なテンス指定が行われるが時間副詞と複文との関係についても今後の課題とする。

## 参考文献

- [Reichenbach47] Hans Reichenbach. *Elements of symbolic logic*. Collier-Macmillan, 1947.
- [Vendler67] Zeno Vendler. *Varbs and Times. Linguistics in Philosophy*. Cornell University Press, 1967.
- [Allen84] James F. Allen. Towards a General Theory of Action and Time. *Artificial Intelligence*. pp.123—154, 1984.
- [Honstein90] Norbert Hornstein. *As Time Goes By*. The MIT Press, 1990.
- [Haruno et al.93] Masahiko Haruno, Yasuharu

Den, Yuji Matsumoto, and Makoto Nagao. Bidirectional Chart Generation of Natural Language Texts. *AAAI-93*. pp.350—356, 1993.

- [池田ほか 91] 池田光生, 東条敏, 安川秀樹. 心的時間の分割とテンス・アスペクト表現. 情報処理学会研究報告 NL83-9. pp.63—70, 1991.
- [江良ほか 93] 江良浩一, 甲斐郷子, 中村順一, 吉田将. テンス・アスペクトを考慮した物語文章の時間関係の抽出. 情報処理学会研究報告 NL95-6. pp.41—48, 1993.
- [Dorr et al.95] Bonnie J. Dorr, and Terry Gaasterland. Selecting Tense, Aspect, and Connecting Words In Language Generation. *IJCAI-95*. pp.1299—1305, 1995
- [工藤 95] 工藤真由美. アスペクト・テンス体系とテキスト. ひつじ書房, 1995.
- [金子 95] 金子亨. 言語の時間表現. ひつじ研究叢書(言語編) 第七巻. ひつじ書房, 1995.
- [東ほか 96] 東正造, 山田篤, 松本裕治. テキストにおけるイベント間の時間関係の抽出. 情報処理学会研究報告 NL112-19. pp.133—138, 1996.
- [玉野ほか 96] 玉野健一, 松本裕治. 制約条件を用いた事象の時間構造の記述. 情報処理学会研究報告 NL115-2. pp.9—14, 1996.