

## 話し言葉翻訳のための日本語アスペクト処理

1F-7

久米雅子 豊島孝之 永田昌明

ATR自動翻訳電話研究所

### 1.はじめに

自然言語、特に話し言葉を解析・翻訳する場合、対象言語のテンス(事態の時を発話時を基準に位置づける文法形式)やアスペクト(動きの展開の様ざまな局面を表す文法形式)の処理が問題となる<sup>[1][2]</sup>。本稿では、SL-TRANS<sup>[3]</sup>の日本語解析部で行っているアスペクト処理について報告する。日本語解析は、HPSG-JPSCに基づいている<sup>[4]</sup>。

### 2.日本語アスペクト処理の概要

日本語と英語では、テンス・アスペクトの表現体系が異なるので<sup>[5]</sup>、翻訳を行う場合は、両言語間の相違に対処する必要がある。例えば、日本語の動詞の「た」形は、英語では過去(ex1)と過去完了(ex2)に訳し分ける必要があり、「る」形は、未来の事態として表現すべき場合(ex3)もある。

(日本語) (英語)

ex1. もう登録しました。 I've already registered.

ex2. 昨日登録しました。 I registered yesterday.

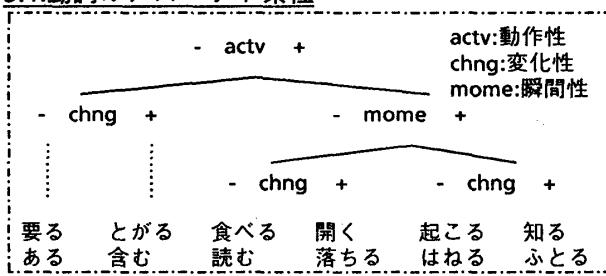
ex3. 来週登録します。 I will register next week.

また、複文の場合、英語には、主節の動詞と従属節の動詞との間に時制の関係を保つ、「時制の一致」があるのに対し、日本語には、時制の一致の制約はない。従って、日本語では、日本語の主文のアスペクトは相対的な意味しか持たないので、英語の「時制の一致」を達成するのに必要な、事態間の時間的順序を解析する必要がある。

SL-TRANSでは、以上の問題を解決するため、單文内の制約から述部ごとに解釈する 1) イベント単位のアスペクト(以下、イベントアスペクト)と、1)のイベントアスペクトを基に、発話時を基準に時系列に関する制約 2) 事態の時間的順序関係(以下、時間的順序)、を解釈している。

### 3.イベントアスペクトの処理

#### 3.1.動詞のアスペクト素性



動詞が潜在的にもつアスペクトに関する素性を図1のように分類し語彙記述として付与する。

#### 3.2.イベントアスペクトの制約

##### 3.2.1.述部承接の制約

一般に、述部のアスペクトは、動詞に、アスペクトを表す助動詞が接続して表現される。従って、イベントアスペクトは、動詞が個々に持つアスペクトに関する情報(図1参照)と、助動詞の組合せにより決定できる。動詞が、「る」形、「た」形、「ている」形、「ていた」形の場合のアスペクト解釈を、表1に示す。

表1 イベントアスペクトに関する制約

	actv-		actv +			
	chng -	chng +	mome -		mome +	
			chng -	chng +	chng -	chng +
teiru - ta -	stat	X	unrl	unrl	unrl	unrl
teiru - ta +	past	stat	past	past	past	past
teiru + ta -	X	stat	prog	rslt prog	iter	rslt
teiru + ta +	X	pstat	pprog	prslt pprog	expr pprog	prslt

(Xは、起こりえない組合せを示す。)

例えば、[[actv +][mome +][chng -]]のアスペクト素性をもつ「振り込む」は、「る」形で、unrl(未完了)、「た」形で、past(完了または過去)、「ている」形でrslt(結果の状態または継続)、「ていた」形でprslt(過去における結果の状態または継続)を表す。

特に、話し言葉における述部の構造は複雑であり、待遇関係(致す、する、なさる)、丁寧度(ます、です)、共感度関係(やる、もらう)など、語用論的な情報を表す助動詞が種々接続する。

ex4. 登録料が振り込まれていませんでした。

ex5. 案内書をお送りさせてくださいています。

ex4、ex5に見られるように、動詞とアスペクトを表す助動詞は必ずしも連接しない。このような長距離依存関係を扱うために、語彙規則中のアスペクトに関する情報を、一般的なCFG規則により单一化操作で親ノードに継承し、一つのイベントに関する記述のまとめを表す終止形などの統語情報を参照し解釈を行っている。

##### 3.2.2.副詞による制約

アスペクト副詞(もう、すでに)やテンス副詞(明日、昨日)との共起制約は、イベントアスペクトの多義の解消に役立つ。

ex6. もう会議は開催されていますでしょうか。  
 ex7. 今日会議は開催されていますでしょうか。  
 ex6とex7において、「開催する」は、[[actv +][mome -][chng +]]のアスペクト素性をもち、「ている」との接続ではrs1tとprogの解釈が可能であるが、「もう」と「今日」の副詞からの制約によりex6はrs1t、ex7はprogに解釈される。

#### 4. 時間的順序の処理

##### 4.1. 形式副詞の制約

日本語の場合、副詞節内のイベントアスペクトは、主文に対して相対的な意味しか表さない<sup>[6]</sup>。例えば、未完了、すなわち、未来の動作を述べているex.8-10の副詞節のアスペクトは、「電話する」動作と「振り込む」動作の生起順序を表す。

- ex8 振り込んだあと、電話して下さい。  
 (振込み=>電話)  
 ex9 振り込むのまえに、電話して下さい。  
 (電話=>振込み)  
 ex10 振り込むのとき、電話して下さい。  
 (振込み=>電話)または(電話=>振込み)

動詞句を補語にして副詞節をつくる接続助詞または形式副詞(以下、形式副詞)のSUBCAT素性は表2のように分類する。

表2 形式副詞のSUBCAT素性

形式副詞	補語
あと、結果、	た形 (IOPPT-)
まえ、まで、うち、あいだ、と	る形 (IPAST-)
とき、場合、のに、ので、が、	る形、た形 (IEVID-)

注 助動詞の承接順序は、MODL素性のハイアラキーをテンプレートとして記述した。!OPTT-は、テンプレートの値で、「た」が承接可(文献[3]参照)。

また、これらの形式副詞は、主文と従属文との時間関係を制約する。例えば、「あと」は、主文の動作・作用より以前に完了している動作・作用を導き、「まえ」は、まだ完了していない動作・作用を導く。一方、「とき」は、同時または総称的な時間関係の動作・作用を導く。このような、事態の時間関係の制約は、語用論的素性としてPRAG素性に記述する。「あと」の語彙記述を以下に示す。

```
[[syn [[head [[pos ADV]
  [form ato]
  [coh [[syn [[head [[pos V]
    [modl !OPTT-]]]]]
  [sem ?Z]]]]]
  [subcat [[[syn [[subcat {}]
    [head [[pos V]
      [modl [[past +]]]]
    [morph [[cform SENF]]]]]
  [sem ?X]]]]]
  [sem [[parm ?Y]
    [restr [[relnあと]
      [iden ?X]
      [entity ?Y]]]]]
  [prag [[aspe [[[reln AFTER]      ;;;事態の
    [refpoint ?Z]        ;;;順序関係
    [event ?X]]]]]]    ;;;に関する制約
```

#### 4.2. アスペクト解析

ex8の発話について、単文内のイベントアスペクトと時間的順序を解析した例を示す。イベントアスペクトをSEM素性のaspt値として、順序関係をPRAG素性のaspe部に表している。

ex8 振り込んだあと、電話して下さい。  
 [[sem ?X[[reln ください] ;;;?Xは、タグを示す
 [aspt UNRL] ;;;イベントアスペクト
 [obje [[reln 電話する]
 [agen ?Z]
 [obje []]
 [recp ?Q]
 [cond [[reln あと]
 [iden ?Y[[reln 振り込む]
 [aspt PAST];;;イベント
 [agen ?Z]
 [obje []]
 [sloc []]]]]]]]
 [prag [[speaker ?Q[[label \*SP\*]]]
 [hearer ?Z[[label \*HR\*]]]
 [aspe [[[reln AFTER] ;;;順序アスペクト
 [refpoint ?X] ;;;
 [event ?Y]]]]]]]];;;

#### 5. アスペクト情報の利用

日本語のアスペクトを上述のように解析することによって、英語への適切な翻訳が可能となる。例えば、「振り込むのときに電話しました。」では、ある過去の時点において、「振込み」と「電話」という動作が同時平行的に行われたことが解析されるので、“I called you when I made bank-transfer.”という時制の一致が可能となる。また、複文のイベント間の時間関係を明らかにすることにより、英文生成の際、談話の流れに応じて、主文、従属文の順番を入れかえたり、単文構造に変換したりすることができる。

また、話し言葉においては、話者の主張の強さや丁寧度がアスペクト表現の微妙な使い分けに反映される。例えば、「電話したと思ったんですが」では、「電話したと思うんですが」と比較して、「た」を用いることによって、主張を弱め、丁寧度を上げている。発話意図は、文化的背景に依存して異なる言語形式で表現されるので、話し言葉の翻訳では、発話意図を抽出することが重要な課題である<sup>[7]</sup>。発話意図とアスペクトの関係には研究すべき部分がまだ多く残されている。

#### 6. おわりに

今後は、本稿で述べた枠組みを精密化し、発話意図の解析に利用する予定である。

#### 参考文献

- [1] Tsujii: "The Transfer Phase in an English-Japanese Translation System", Proc. of Coling 82, 1982.
- [2] 中園他:「アスペクト情報の処理について」, NL研究会資料61-2, 1987.
- [3] 井上編:「日本文法小事典」、大修館書店、1989
- [4] 森元他:「音声言語日英翻訳実験システム(SL-TRANS)の概要」, 第39回情処全国大会, 1989
- [5] 吉本、小暮:「日本語端末間会話のための句構造文法」, 第37回情処全国大会, 1988
- [6] Nakau: "Tense, Aspect, and Modality", Syntax and Semantics 5, 1976
- [7] M. Kume et al.: "A Descriptive Framework for Translating Speaker's Meaning", in 4th EACL, 1989.