

アスペクト情報の処理について

中園 薫 古瀬 蔵 野村 浩郷
NTT 電気通信研究所

本稿では、アスペクト情報の処理をおこなう単純な手法とその機械翻訳への応用について報告する。ある事象の表現に関して、アスペクトとは、時間的経過の観点から見て、どの局面に着目点があたっているかを表している情報とする。各々の動詞は潜在的にアスペクトを持っており、補助動詞や助動詞などとの接続によりこの着目点が移動することによって、文自体のアスペクトが決まると考える。本稿では、LFGの枠組みをもちいて、特別なアルゴリズムを用意しなくともアスペクト情報の処理を実現できることを示す。さらに、F構造を中間表現として、原言語のF構造を目的言語のF構造へ変換する方法をもちいた機械翻訳への応用について述べる。

On Analysis of Aspect

Kaoru NAKAZONO, Osamu FURUSE, Hirosato NOMURA

NTT Electrical Communication Laboratories
3-9-11 Midori-Cho, Musashino-shi, Tokyo 180, Japan

This paper presents a simple method for processing aspect of sentence or verb and its application to machine translation. In the description of event, aspect can be considered as the information which shows the focus of attention in the passage of time. Each verb has its own aspect potentially, and the aspect of a sentence is determined by shifting the focus when it is combined with aspectualizers and auxiliary verbs. In the framework of LFG, the aspect can be easily analysed without using any special algorithms. Furthermore, the application to machine translation using the f-structure as the intermediate representations is also presented.

1 はじめに

アスペクトの情報は、表層表現に明示的な形で見られるとは限らず、また言語の相違による差異も大きい。このために、自然言語理解や機械翻訳をおこなう上で、アスペクト情報の処理は大きな問題となっている^{[1][2]}。

本稿では、アスペクトとは、ある事象(なんらかの動作や状態の変化、あるいは状態の叙述)を時間的経過の観点から見て、どの様な局面に着目しているかを表している情報と考える。そして動詞はそれぞれ単独で着目している局面を持ち、それが補助動詞や助動詞などとの接続により時間軸上をどの局面に移動するかによって文のアスペクトが決まると考える。このように、補助動詞や助動詞の働きを、もとのアスペクトに対して新しいアスペクトの値を与えるという関数的な働きと見なすことによって、アスペクト情報の処理を単純に形式化することができる。

次に、このアスペクト情報の解析を、LFGの枠組みをもちいて実現する。LFGでは、個々の語に対して、語義的な情報のみでなく、構文的な機能情報までも語彙項目に持たせている。本稿では、さらにアスペクトに関する関数的働きも同様な形で語彙項目に与える。これによって、特別なアルゴリズムを用意しなくとも、既にあるLFG処理系を使って、アスペクトの解析が実現できるようになる。

さらに、F構造を中間表現とした機械翻訳への応用について述べる。ここでは、原言語のF構造から目的言語のF構造への変換(LFTと呼ぶ)によって変換処理を実現する。この方法は、(1)語のレベルだけでなく、構造のレベルの変換までも、統一的な記述方法で、処理することができる、(2)一つの変換規則で両方向の変換が可能である(例えば日英、英日)、などの利点がある。特に、アスペクトの情報は語のレベルと構文のレベルの複合的な表現をとるため、LFTによる変換が有効である。

2 アスペクトの処理の問題点

自然言語を機械処理する場合に、特にアスペクトの処理に関して、以下のような問題点が考えられる。

(1)表層とアスペクト情報との対応

日本語では、アスペクトの情報は主に動詞に助動詞がついた形で表される。また、英語では、begin, finishなどの動詞や進行形、完了形などの形で表される。しかし、これらの表層と文のアス

ペクトは一对一には対応しない。この違いは日英間の翻訳を考えるとさらに顕著となる。

[1] 私は本を読んでいる。

I am reading a book.

[2] 私は彼を知っている。

I know him.

[3] 彼は死んでいる。

He has died. He is dead.

(2)アスペクトの多義

表層上は全く同じ文でも、アスペクトの情報が多義となる場合がある。たとえば、

[4] 私は彼女に手紙を書いている。

という文は、「書いてしまっている」という意味にも、「書いてるところだ」という意味にも取り得る。しかし、副詞と共に起して、

[4'] 私はいま彼女に手紙を書いている。

私はすでに彼女に手紙を書いている。

となると、そのような多義は解消する。これらの現象をうまく説明できるような枠組みを作る必要がある。

(3)動詞、助動詞の接続によるアスペクトの変化

日本語では、動詞が接続したり、さらに助動詞がつながったりすることにより、アスペクトが次々と変化していく。

[5] 私は走る。

[6] 私は走り終わる。

[7] 私は走り終わっている。

(4)アスペクトに関する非文

次のような非文は、単純な句構造規則による解析だけでは排除できない。アスペクトも含めた解析が必要となる。

[8]* 彼は黙り終わる。

[9]* 私は本を読み終わり続ける。

3 動詞の分類とアスペクト

なんらかの事象について述べている文は、その事象の(時間軸上の)どこかの局面に焦点があたっていると考えることができる。そこで、アスペクトとは、その文が、事象のどの局面に焦点をあてているかという情報であると考えられる。

動詞は、その動詞単独でどの局面に焦点があたっているかの情報、すなわち動詞単体でのアスペクト情報をプロトタイプとして持っていると考えられる。このプロトタイプが持っているアスペクト値により、動詞の分類ができる。そして、動詞が実際の文の中で使われたとき、すなわちインスタンティエイトされたときに「～てい

る」などの助動詞や、「始める」などの動詞(補助動詞と呼ぶ)と接続することによって、その焦点が時間軸上を移動し、全体で新たなアスペクトを与えることになる。このような働きをする助動詞や補助動詞をアスペクト形式素と呼ぶ。

3.1 アスペクトカテゴリー

アスペクトのカテゴリーを開始前;開始;継続;終了;結果の5つに分類する。たとえば、ある文がアスペクト値として「終了」を持つとは、ある事象が終了している局面に焦点を当てて表現している、という意味である。これらのカテゴリーの概念を図3-1に示す。開始と終了に対してはそれぞれの直前、直後というカテゴリーも考えられるがここでは省略する。

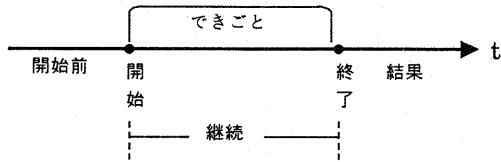


図3-1. アスペクトカテゴリーの概念

3.2 動詞の分類

動詞を表3-1に示すように5つに分類する。この分類の基準は、その動詞が(アスペクト形式素が付かない)それ自体で、事象のどの局面について焦点をあててのべているかという、アスペクトの意味を中心としたものである。したがって、各分類に属する動詞はそれぞれ同一の潜在的アスペクト値を持っている。

表3-1. 動詞の分類

分類	典型例	アスペクトの特徴	潜在的アスペクト値
状態動詞	ある 美しい	状態を静的に述べる	
動作動詞	走る 読む	動作終了後の状態を陽に意識しない	継続
変化動詞	過渡動詞 開く 沈む	変化中の局面に焦点を当てている	継続
	終了動詞 知る 死ぬ	変化の達成点に焦点を当てている	終了
結果動詞	愛する 黙る	変化後の状況に焦点を当てている	結果

変化動詞は、その動作によって引き起こされた状態の変化を明らかにしているものである。たとえば、動作動詞と過渡動詞は、ともに動作性という性格をもつが、動作動詞が、動作終了後の状態を陽に意識しないのに対し、「開

く」などの過渡動詞は変化動詞の一種で、動作終了後の状態を+openのように明確に把握できる。

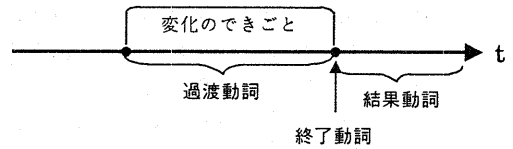


図3-2. 変化動詞の焦点のちが

3.3 アスペクト形式素によるアスペクトカテゴリーの変化

動詞のプロトタイプが潜在的に持っていたアスペクトは、アスペクト形式素と接続して文中に現れることにより、新たなアスペクト値となって顕在化する。また、複数のアスペクト形式素が接続した場合も同様にもとのアスペクト値から新たなアスペクト値に置き変わる。このアスペクトカテゴリーの対応を表3-2に示す。動詞が文の中に単独で現れたときは、潜在的アスペクトがそのまま顕在化する場合や「人は死ぬ」のように抽象的表現でアスペクトを持たない場合などがある。

表3-2. アスペクト形式素の接続によるアスペクト値の変化

もとのアスペクト値	「ている」		始める	終わる	続ける
	継続化	結果化			
継続	継続	結果	開始	終了	継続
終了	×	結果	×	×	×
結果	×	結果	終了	×	結果
開始	×	継続	×	×	×

×:接続不可

この表について、いくつか補足しておく。まず、「結果動詞+始める」は結果の状態の開始に相当する。これは、その状態をもたらす変化(というできごと)が終了することを意味する。例えば「住み始める」は、「住む(=住んでいる)」という状態の開始、すなわち「そこに住むようになる」という事象がちょうど終了した時点を示している。このようすを、図3-3に示す。次に「開始動詞+ている」であるが、この場合は開始している時点で焦点が当たっている動詞が結果化することによって、その事象が継続していることを表す。

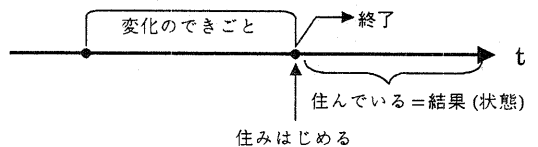


図3-3「住みはじめる」(「結果動詞+始める」)のアスペクト

この表から、2章で述べたアスペクト処理に関するいくつかの問題が説明できる。[4]の「書いている」は、「動作動詞+ている」であるから、継続または結果の両方のアスペクトを持つ。[6]の「走り終わる」は、継続アスペクトを持つ動作動詞に「終わる」がついているので、全体として終了アスペクトを持つ。さらに[7]のように「ている」がついて「走り終わっている」となると、表から結果アスペクトを持つことがわかる。[8]は「黙る」という結果アスペクトを潜在的に持つ動詞には「終わる」がつかないことから非文であることがわかる。同様に[9]も終了アスペクトを持つ「読み終わる」に「続ける」は接続しないので非文である。

4 LFGによるアスペクト解析

LFG^{[3][4]}はBresnanとKaplanによって考案された文法理論である。LFGでは、文を表現する構造として2つのレベルをもちいている。ひとつは、句構造文法理論の句構造木に相当するもので、C構造(Constituent Structure)と呼ばれている。これに対して文要素の機能的構造をより深いレベルで表現したものをF構造(Functional Structure)という。これは、文を文法機能に関する“属性名-属性値”の対の入れ子構造で表現している。

LFGは、種々の文法現象を説明する手段として、それまでの句構造文法理論で扱われてきた変形操作のかわりに、語自身がさまざまな機能を持つとして、これを辞書の中の語彙項目に与えている。文はまず、句構造規則に基づいて解析されC構造で表現される。このC構造から、各語(すなわち木構造の葉)に与えられた機能を構文木の上方にむかってまとめあげていく(ユニフィケーションと呼ばれる)操作によって、文全体のF構造を得る。このように、LFGは、木構造の中の個々の情報をまとめあげていく際の手続き(プログラム)に相当する部分はユニフィケーションという操作にだけ制限し、情報の大部分を語彙項目(データ)の中に押し込めている。この方法によって、さまざまな文

S	→	NP *	(SCOMP)	VP
		(↑(↓ CASE-NAME))=↓	(↑ XCOMP)=↓	↑=↓
NP	→	N	PART	
VP	→	ADV	VP	
VP	→	VP	AUX	
VP	→	V		
SCOMP	→	S		

図4-1. C構造規則(日本語)の例

法現象が静的な記述で説明され、非常に見透しがよくなっている。

この章では、このLFGの枠組をもちいて、特別な解析アルゴリズムをもちいることなしに入力文からF構造を作る手続きの中で、3章に述べた日本語文のアスペクト解析が実現でき、F構造に埋め込まれた形でアスペクトの情報が得られることを示す。

4.1 構文規則

LFGでは、まず、C構造を得るために、C構造規則と呼ばれる句構造規則をもちいる。これは、普通の文脈自由型の規則と相違ないが、各規則の範疇記号にはあとでF構造を作っていく際に参照する等式が付記されている。この等式はその節点自身の文法範疇に相当するF構造(下向き矢印↓で表される)と、これを支配する文法範疇に相当するF構造(上向き矢印↑で表される)との関係を表している。

ここでは、単純な例として図4-1に示す日本語のC構造規則をもちいる。この規則は複文を扱うことはできないが、単文に関するアスペクトを解析することができる。

4.2 語彙項目

語彙機能に関する情報は辞書の中に等式の形で記述されている。

文のアスペクトを決定する要素は、文法機能の観点から以下のように分類できる。

(1) 他の文法機能を下位範疇化して、新たなアスペクトを与えるもの。

- a. それ自体で文のF構造を作れるもの。一般の動詞、形容詞。
- b. 他の動詞の後に接続して、アスペクトを与える動詞。このような動詞を補助動詞と呼ぶ。ここでは、補助動詞の前に接続している動詞を主語を持たない補文(XCOMP)の動詞と見なして、補助動詞がこのXCOMPを下位範疇化していると考えられる。

【例】始める、終わる、続ける

(「彼は本を読み始める」ならば「本を読む」がXCOMP)

(2) 他の文法機能と結合してアスペクトを与える助動詞。

【例】た、ている

(3) それ自体で新たなアスペクトを与えることはせず、共起できるアスペクトを制約すること

少年	N	(↑ PRED) = '少年'
は	PART	(↑ CASE-MARKER) = は (↑ CASE-NAME) = SUBJ
を	PART	(↑ CASE-MARKER) = を (↑ CASE-NAME) = OBJ
に	PART	(↑ CASE-MARKER) = に (↑ CASE-NAME) = OBJ2
走る	V	(↑ PRED) = '走る <(↑ SUBJ)>' (↑ NATIVE-ASPECT) = 継続
始める	V	(↑ PRED) = '始める <(↑ SUBJ)(↑ XCOMP)>' (↑ XCOMP SUBJ) = (↑ SUBJ) (↑ XCOMP NATIVE-ASPECT) = c 継続 (↑ ASPECT) = 開始
始める	V	(↑ PRED) = '始める <(↑ SUBJ)(↑ XCOMP)>' (↑ XCOMP SUBJ) = (↑ SUBJ) (↑ XCOMP NATIVE-ASPECT) = c 結果 (↑ ASPECT) = 終了
終わる	V	(↑ PRED) = '終る <(↑ SUBJ)(↑ XCOMP)>' (↑ XCOMP SUBJ) = (↑ SUBJ) (↑ XCOMP NATIVE-ASPECT) = c 継続 (↑ ASPECT) = 終了
ている	AUX	(↑ AUX-FORM) = 'ている'
すでに	ADV	(↑ ADV-FORM) = 'すでに' (↑ ASPECT) = c 結果
今	ADV	(↑ ADV-FORM) = '今' (↑ ASPECT) = c 継続

図 4-2. 語彙項目の例

によって、フィルターのな役目をするもの。
副詞、副詞句など。

【例】明日、今、....まで

これらの語彙の機能をLFGの語彙項目としてどのように表現するかを以下に述べる。

(1-a) 一般の動詞

一般の動詞の語彙項目には、その語の意味述語を表すPRED属性のほかに、その動詞が本来持っているアスペクト素性を与えるNATIVE-ASPECT属性の値が記述されている。

(PRED属性に対応している値をsemantic formという。<>の中の並びはその述語が下位範疇化する文法機能を表している。)

(1-b) 補助動詞

アスペクトを与える補助動詞は、制約等式によって、その補助動詞が下位範疇化するXCOMPの動詞のNATIVE-ASPECTの種別を規定し、ASPECTの値を与える。

(制約等式は '=' の代わりに '='c' で表される。通常の等式が、単にF構造の中にその「属性名-属性値」の対が存在することを定義しているのに対して、制約等式は、その式で表された「属性名-属性値」の対の存在が他の等式によって定義されていることを要求する。)

(2) 助動詞

アスペクトに関する助動詞として、「ている」と「た」などがある。これらの助動詞は一般の動

<継続の動詞に継続化のテイルがついた場合>

Apply if: (↑ NATIVE-ASPECT) = 継続
Add: (↑ AUX-FORM) = c 'ている'
(↑ ASPECT) = 継続

<継続, 終了, 結果の動詞に結果化のテイルがついた場合>

Apply if: (↑ NATIVE-ASPECT) = 継続 or 終了 or 結果
Add: (↑ AUX-FORM) = c 'ている'
(↑ ASPECT) = 結果

<開始の補助動詞に結果化のテイルがついた場合>

Apply if: (↑ ASPECT) = 開始
Rewrite: (↑ ASPECT) = 開始 ==> (↑ ASPECT) = 継続
Add: (↑ AUX-FORM) = c 'ている'

<終了の補助動詞に結果化のテイルがついた場合>

Apply if: (↑ ASPECT) = 終了
Rewrite: (↑ ASPECT) = 終了 ==> (↑ ASPECT) = 結果
Add: (↑ AUX-FORM) = c 'ている'

図 4-3. 語彙項目書換え規則の例

走る V (↑ PRED) = '走る <(↑ SUBJ)>'
(↑ NATIVE-ASPECT) = 継続
(↑ ASPECT) = 継続
(↑ AUX-FORM) = 'ている'

走る V (↑ PRED) = '走る <(↑ SUBJ)>'
(↑ NATIVE-ASPECT) = 継続
(↑ ASPECT) = 結果
(↑ AUX-FORM) = 'ている'

始める V (↑ PRED) = '始める <(↑ SUBJ)(↑ XCOMP)>'
(↑ XCOMP SUBJ) = (↑ SUBJ)
(↑ XCOMP NATIVE-ASPECT) = c 継続
(↑ ASPECT) = 継続
(↑ AUX-FORM) = 'ている'

終る V (↑ PRED) = '終る <(↑ SUBJ)(↑ XCOMP)>'
(↑ XCOMP SUBJ) = (↑ SUBJ)
(↑ XCOMP NATIVE-ASPECT) = c 継続
(↑ ASPECT) = 結果
(↑ AUX-FORM) = 'ている'

図 4-4. 語彙規則によって書き換えられた語彙項目

詞または補助動詞のいずれかに直接接続する。ここでは、助動詞の語彙項目としては、単にAUX-FORMという属性名で、その助動詞の形を与えるだけとする。そして、動詞、補助動詞の側の語彙項目として、助動詞がついた場合に相当するASPECTの値を持つようなものを設けておく。このような助動詞の接続を仮定した語彙項目はそれぞれの動詞に対して最初から用意するのではなく、「語彙規則」という規則を用意しておき、もとの動詞の語彙項目から書き換えることによって得られる。

(3) 副詞

制約等式を利用して、アスペクトを制約する。

語彙項目の例を図4-2に示す。また、「ている」などの助動詞に関する、動詞と補助動詞の語彙規則を図4-3に、さらにそれによって書き換えられた新たな語彙項目の例を図4-4に示す。

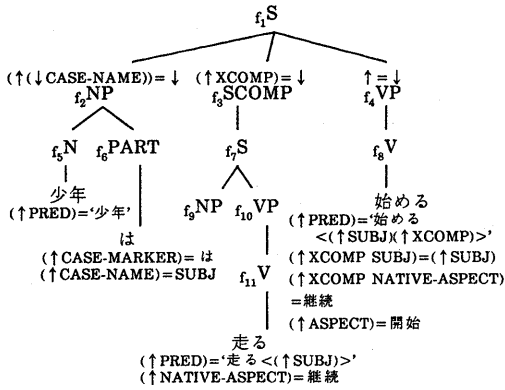


図4-5. 「少年は走り始める」のC構造

4.3 C構造とF構造

これまでの節で示した構文規則と語彙項目をもちいて構文解析をおこない、C構造を得る。この段階では、構文規則や語彙項目は非決定的に選ばれるだけである。したがって、一般に一つの文に対して複数個のC構造が得られ、それらの多くは誤った解析結果を与えている。また、非文に対してもC構造を与えてしまう。

このC構造から、付記された等式と語彙項目の情報をもとにF構造を作っていく。誤った解析結果に基づいたC構造はこの過程で失敗して、廃棄される。この手順を以下に示す。

「少年は走り始める。」

この文を解析した結果得られるC構造を図4-5に示す。このC構造の各節点の文法範疇に対するF構造を f_1, f_2, \dots として、それぞれの節点に付記されている等式のメタ変数(\uparrow, \downarrow のこと)を、その矢印がさす節点の具体的なF構造の名前に置き換える。この操作をInstantiationと呼ぶ。Instantiateして得られた等式の集合をF表現(F description)という。図4-5に示したC構造をもとに得られたF表現を図4-6に示す。

このF表現の式を連立方程式と見なして解いてゆく。この途中で矛盾が生じた場合は、失敗としてそのC構造を廃棄する。こうして得られた f_1, f_2, \dots の解がそれぞれ、対応するF構造となる。したがって、文SのF構造は f_1 の解に相当する。例文のF構造を図4-7に示す。

実際には、C構造として、補助動詞「始める」の部分に、図4-2で示した語彙項目および、それらを図4-3の語彙規則で書き換えた語彙項目のすべてが入りうるが、(XCOMP NATIVE-ASPECT)に関する制約等式によって誤ったアスペクトを与えるようなものは廃棄される。

- $(f_1 (f_2 \text{ CASE-NAME})) = f_2$
- $(f_1 \text{ XCOMP}) = f_3$
- $f_1 = f_4$
- $f_2 = f_5$
- $f_2 = f_6$
- $f_3 = f_7$
- $f_4 = f_8$
- $(f_8 \text{ PRED}) = \text{'少年'}$
- $(f_6 \text{ CASE-MARKER}) = \text{'は'}$
- $(f_6 \text{ CASE-NAME}) = \text{SUBJ}$
- $f_7 = f_9$
- $f_7 = f_{10}$
- $(f_9 \text{ PRED}) = \text{'始める <(\uparrow \text{SUBJ})(\uparrow \text{XCOMP})>'}$
- $(f_9 \text{ XCOMP SUBJ}) = (\uparrow \text{SUBJ})$
- $(f_9 \text{ XCOMP NATIVE-ASPECT}) = c \text{ 継続}$
- $(f_9 \text{ ASPECT}) = \text{開始}$
- $f_{10} = f_{11}$
- $(f_{11} \text{ PRED}) = \text{'走る <(\uparrow \text{SUBJ})>'}$
- $(f_{11} \text{ NATIVE-ASPECT}) = \text{継続}$

図4-6. 「少年は走り始める」のF表現

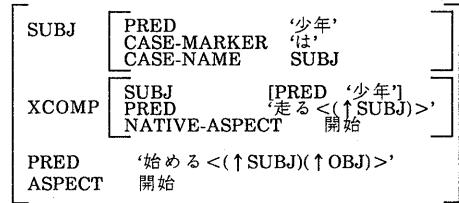


図4-7. 「少年は走り始める」のF構造

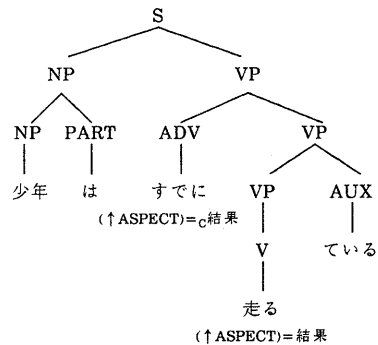


図4-8. 「少年はすでに走っている」のC構造

次に副詞によるアスペクトのフィルターの効果を見る。「少年はすでに走っている」を例にとると、副詞「すでに」の語彙項目の制約等式により、図4-8に示したC構造だけが正しいF構造を生じ、結果相を取ることがわかる。

5 機械翻訳への応用

これまで、LFGの枠組みをもちいて、F構造の形でアスペクト情報を取り出す方法について述べた。この章では、その応用のひとつとして機械翻

訳を取り上げる。LFGでは、言語表現の持つ文法的な関係と語彙的な関係を、F構造という統一的な枠組みのもとで表現している。そこで、このF構造を中間表現と見なし、原言語のF構造から目的言語のF構造への変換(LFT)をおこなうことによって、構造変換と語彙変換とを統一的に扱った機械翻訳をおこなうことができる(図5-1)。特にアスペクト情報は、構文的な構造と語彙との複合的な表層形式で表現され、言語間での相違も大きいいため、LFTによる変換は有効である。

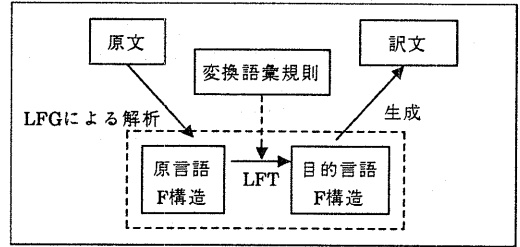


図5-1. LFTに基づく翻訳モデル

表5-1.各動詞のアスペクト値に対する表層形式(英語)

対応する英語 の動詞の種別	アスペクトの値	
	継続	結果
状態動詞	×	時制のみ
動作動詞	進行形	完了形
過渡動詞	進行形	完了形
終了動詞	×	完了形
結果動詞	×	時制のみ

5.1 アスペクトカテゴリーと英語の語形

異なった言語間で相当する文を比べてみると、言語によってさまざまな表現形式でアスペクト情報が表現されていることがわかる。日英翻訳を例とすると、同じ「動詞+ている」の文でも、英語側では進行形、完了形などいろいろな形を取ることは2章ですでに述べた。

特に、2章の[2]と[3]の英語側の構造を比べてみると、日本語側のアスペクト解析の結果はいずれも「結果」を与えるのに、[2]では動詞の現在形だけで表現され、[3]では現在完了形あるいは形容詞文の形で表現されている。これは、英語側の動詞の潜在的アスペクト値が異なるため、構文形態の調整によって文のアスペクトを合わせているのである。

このように、原言語のアスペクト情報を目的言語側で表現するためには、動詞の潜在アスペクト値を考慮して構文構造などの表現形態を選択する必要がある。そのために、英語の動詞についても日本語の場合と同様に表3-1に基づいた動詞分類をおこない、それぞれの動詞についてアスペクトカテゴリーと表層形式の対応を規定する。「継続」と「結果」アスペクトに対する英語の表層形式を表5-1に示す。

この表により例の[1]から[3]が説明される。[1]の日本語は「継続」アスペクトを持ち、「read」は動作動詞に属する。したがって、英語では進行形を取る。[2]の文は「結果」アスペクトを持ち、「know」は結果動詞である。したがって進行、完了などのアスペクト形式は取らない。[3]の文も同様に「結果」アスペクトを持つが、これに対応する英語の「die」、「be dead」はそれぞれ終了動詞、状態動詞なので「die」ならば完了形になり、「be dead」ならば特別なアスペクト形式は取らないわけである。

5.2 F構造の変換

LFGのF構造を中間表現と見なし、原言語のF構造から目的言語のF構造に変換する方法(LFT, Lexical-functional Transfer)が提案されている^{[5][6]}。この方法を簡単に説明する。一般のトランスファー方式における変換規則に相当するものとして、原F構造の中の式が目的F構造の中のどの式に対応するかを表した対照表を用意する。変換過程では、もとのF構造をInstantiateし、この対照表を参照しながら、目標とするF構造の式をInstantiateされた形;すなわちF表現の形に変換する。これから先は、通常のLFGにおけるF表現からF構造を作るときと全く同じ手順で目的言語のF構造が得られる。

まず、日本語文「私は彼を知っている」を英語に翻訳する場合を考えよう。解析した結果得られるF構造の概略を図5-2に示す。これに対して、図5-3に示す変換規則を適用すると(ここでは、「知っている」に相当する規則のみ示す)、図5-4に示すような、「I know him」に相当するF構造が得られる。

次に、日本語文「少年は走り始める」を英語に翻訳する場合を考える。この場合、「The boy begins to run」のようになる。このように、「～始める」から「begin to」のように、両方の言語で開始相の情報がXCOMPを取る補助動詞の構文の形で表れている場合は、ASPECT属性の値を参照する必要はなく、構文対構文の変換ですむ。

これが「少年は走り始めている」になると、ASPECT属性の値は継続であるが、変換処理では

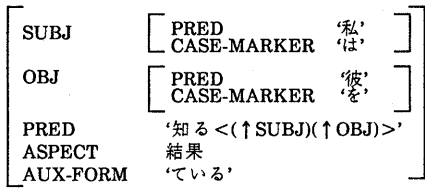


図5-2. 「私は彼を知っている」のF構造

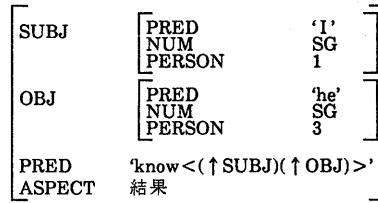


図5-4. 'I know him' のF構造

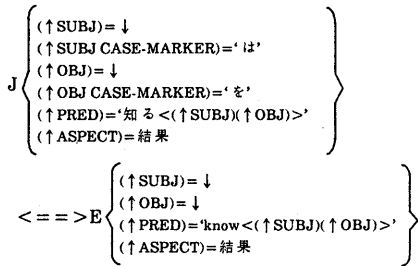


図5-3. 「知っている」の変換規則

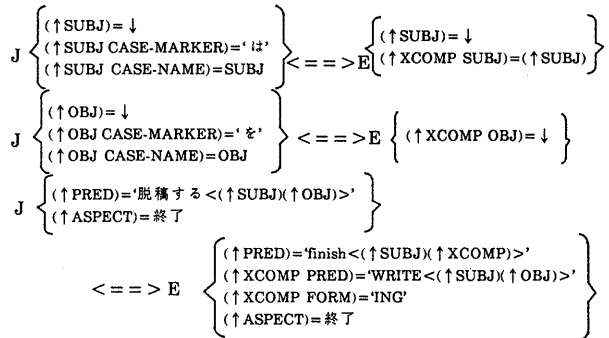


図5-5. 「脱稿する」 ↔ 'finish writing' の変換規則

「始める」に対して「ている」がつく形を優先させて、"have begun to ~" という構文を取る。

しかし、「彼は小説を脱稿する」という文を考えると、動詞の「脱稿する」は終了動詞で、単独で終了アスペクトを持つ。しかし、英語では、これに対して一語で相当するような表現はないため、動作動詞の "write" と補助動詞の "finish" を組み合わせて、"finish writing" に変換する必要がある。この変換規則を図5-5に示す。

このように、LFTでは、語のレベルと構文構造のレベルの情報に加えて、ASPECT属性値の情報も適宜組み合わせる適切な翻訳ができる。

5.3 生成過程

生成過程はF構造からまずC構造を作り、その終端ノードの式をもとに形態素合成をおこない、目的文を生成する。ここでは、詳細は省略する。

6 あとがき

文のアスペクト情報を、動詞本来が持つアスペクト的性質と、アスペクト形式素が持つ関数的性質によって解析する方法について述べた。また、このアスペクト情報は、LFGの枠組みをもちいることによって特別な解析アルゴリズムを使わずに解析でき、F構造に埋め込まれた形で得られることを示した。さらに、F構造を中間表現とした機械

翻訳に応用することによって、アスペクト情報を含んだ文に対しても適切な翻訳ができることを示した。

【参考文献】

- [1] 井上、変形文法と日本語(下)、大修館(1976)
- [2] 草薙、「日本語文解析におけるテンス・アスペクトの問題」情報処理学会自然言語研究会資料、34-11(1982)
- [3] Bresnan, J. (Ed.), *The Mental Representation of Grammatical Relations*, MIT Press (1982)
- [4] Ishikawa, A., "Complex Predicates and Lexical Operations in Japanese", Doctorial Dissertation, Stanford Univ. (1985)
- [5] 工藤、野村、成田「LFGのスキーマを用いた機械翻訳」情報処理学会自然言語研究会資料 48-8 (1985).
- [6] Kudo, I., and Nomura, H., "Lexical-Functional Transfer: A Transfer Framework in a Machine Translation System Based on LFG", *COLING* 86 (1986)