2B-2 多言語間機械翻訳システムにおけるテンスとアスペクト

Hartono 徳永 健伸 田中 穂積
東京工業大学

1 はじめに

多言語間の機械翻訳システムを考えると中間言語方式を採用することの利点は無視できない．我々は中間言語方式に基づく機械翻訳システムを開発するための基礎的研究として中間言語の設計をおこなっている．

中間言語には言語に依存しない中立的な表現が必要であり，そのためには各言語に対する洞察と言語間の対照研究が必要である．本稿では中国語，日本語，インドネシア語，タイ語，英語を対象言語として，中間言語を考える上で重要な要素のひとつであるテンスとアスペクトに関して，その言語間の相違，表現法に関して考察する．

2 テンス

テンスをどのように定義し，分類するかは，言語によって，また，言語学者によっても異なっている．たとえば，金田一は日本語のテンスを過去と非過去に分類している [2]．英語では，動詞の語形の変化により，過去，現在，未来を表現する[1]．たとえば，以下の例について考える．

1. He *went* to school yesterday.
2. He *goes* to school every day.
3. He will *go* to school tommorow.

この英文に対するインドネシア語と中国語の訳はそれぞれ以下のようになる．

1. Dia *pergi* ke sekolah kemarin.
2. Dia *pergi* ke sekolah setiap hari.
3. Dia *pergi* ke sekolah besok.
1. 他昨天去学校.
2. 他毎天去学校.
3. 他明天去学校.

インドネシア語と中国語ではテンスが変化しても動詞の語形が変化するという現象はない．テンスを表層レベルに陽に示さないのがこれらの言語の特徴の一つである．また，タイ語では時制副詞を用いてテンスを陽に表現することもあるが，省略される場合もあることから，英語ほど強くテンスが表層に現れない．

著者らは以上を考慮して，テンスを以下のように定義する．

> テンスとはある基準時点から眺めたイベント[1]が発生した時点の間の関係を示すものである．

[1]以下，動作，作用のことをイベントという．

A study on tense and aspect for a multilingual machine translation system
Hartono, TOKUNAGA Takenobu, TANAKA Hozumi
Tokyo Institute of Technology

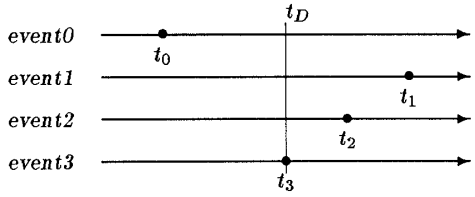

図 1: イベント間の発生時点の関係

時制を持たないインドネシア語，中国語などでは，イベントの発生時点は時を表す副詞または文脈によって示される．

著者らは中間言語におけるテンスを，単に過去，現在，未来というラベルで表現するのではなく，各イベントの発生時点の相対的な関係で表す(図 1)．発話時点を t_Dとし，イベントの発生時点を t_i $(i = 0, 1, 2, \ldots)$とする．抽出された各イベントの発生時点を図 1のような関係だとすれば，これらの関係は以下のように表現される．

$$t_0 < t_D = t_3 < t_2 < t_1$$

ただし，$t_x < t_y$は，t_xは t_yより以前であることを示し，また，$t_x = t_y$は，t_xと t_yが同時か，または部分的に重なっていることを示す．以下に例を挙げる．

1. He $\underline{went}_{t_0}$ to school. $\Rightarrow$ $t_0 < t_D$
2. He has $\underline{gone}_{t_0}$ to school. $\Rightarrow$ $t_0 < t_D$
3. 本を$\underline{読んで}_{t_0}$から$\underline{話した}_{t_1}$ $\Rightarrow$ $t_0 < t_1 < t_D$
4. 家に$\underline{到着した}_{t_0}$時，彼は本を$\underline{読んでいた}_{t_1}$
   $\Rightarrow$ $t_0 = t_1 < t_D$

上の例(1)と(2)は結果的にどちらも $t_0 < t_D$ で表されているが，(2)にはアスペクトとして後述の結果態が付加される．

3 アスペクト

前節で説明したテンスは，ある基準時点から眺めたイベントの発生時点間の関係を示している．これに対して，アスペクトは以下のように定義する．

> アスペクトとはイベントの発生時点を軸にしてそのイベントの進行様態を示すものである．

著者らはイベントを，瞬間的なものと非瞬間的なものに分け，これ基づいてアスペクトを図 2のように分類する．ここでいう，瞬間的なイベントとは，たとえば,「到着する」など，動作が持続しないものをいう．また，非瞬間的なイベントとは,「読む」のように動作がある期間持続するものをいう．

直前態は瞬間的なイベントが発生する直前を，直後態は発生した直後を表わす．結果態は瞬間的なイベントが発生した結果

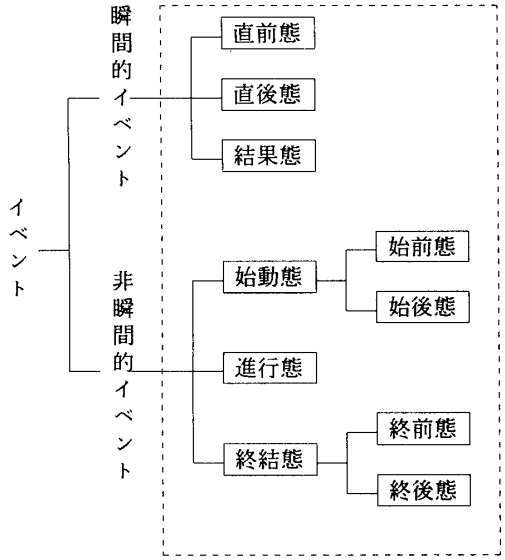

図 2: アスペクトの分類

表 1: アスペクトの例

アスペクト	金田一の分類	例
直前態	将然態, 将現態	着くところ, 読むところ
直後態	既然態, 既現態	着いたところ, 読んだところ
結果態	既然態	着いている
始動態	始動態	読みはじめる
始前態	将然態	読みはじめるところ
始後態	既然態	読みはじめたところ
進行態	進行態	読んでいる
終結態	終結態	読みおわる
終前態	将然態	読みおわるところ
終後態	既然態	読みおわったところ

が残っている状態を表わす．始動態以降は非瞬間的なイベントに関するアスペクトを表わす．始動態はイベントが始まることを，進行態は進行中であることを，終結態は終わることを表わす．この内，始動態と終結態については，イベントが始まる，終わる瞬間を瞬間的なイベントと考えることによって，直前態と直後態に相当するようなアスペクトをそれぞれ考えることができる．これが，始前態，始後態，終前態，終後態である．金田一は，動詞の種類を，状態，瞬間，継続，第 4 種の動詞の 4 種類に分類した上で，それぞれにつくアスペクトを分類しているが，ほとんどの動詞が 2 種類以上の動詞，特に瞬間動詞と継続動詞を兼ねているものが多いと述べている [2]．金田一の瞬間動詞，継続動詞は，ここでいう瞬間的なイベント，非瞬間的なイベントに対応している．

日本語のアスペクトの分類も,言語学者によってさまざまではあるが，ここでは，金田一の分類を取り上げて，著者らの分類との対応関係を表 1に示す．金田一の分類の中で，単純状態態，単純動作態，継続態が表のからもれているが，著者らはこれらをこれらをアスペクトとは考えず，単純状態態，単純動作態については概念のレベルで考え，継続態についてはイベントの属性としてアスペクトとは別扱いにする．

英語，インドネシア語，中国語についても各アスペクトに対応する表現が存在する．

タイ語場合は，結果態のアスペクトが存在しない．例えば，He has read the book. と He read the book.の区別はしない．つまり，どちらでも発話時点より前に発生したイベントであるということしか考えないということになる．テンスの表現で表すと，どちらも $t_0 < t_D$となる．

インドネシア語や中国語のアスペクトは，アスペクトマーカーによって表現される．一例を挙げると以下のようなものがある．インドネシア語には，mulai (begin), selesai (finish), sudah (already), baru (just), sedang (in-progress), hampir (almost,nearly), akan (will)などがある．中国語には，開始(begin),完了(finish),已経(already),剛(just), 正在(in-progress), 要(will)などがある．

COMRIE は英語において完了形として表現される経験もアスペクトに含めているが，著者らはこれを金田一の継続態と同様にイベントの属性として考える [6]．また，英語の完了形もテンスの一形式とする考え方もあるが，著者らはこれをアスペクトとして扱う．

4 おわりに

本稿では，中国語，日本語，インドネシア語，タイ語，英語を対象言語とし，テンスとアスペクトに関して比較考察した．また，それをもとに機械翻訳のための中間言語における表現方法について述べた．これらの表現を計算機の上でインプリメントする予定である．

謝辞

本稿をまとめるにあたり，活発な議論をしてくださった田中研中間言語ワーキンググループの Vises Vorascha，Quek Chee Huei，崔進，田村直良，知野哲郎，根本治朗，各氏に感謝します．

参 考 文 献

[1] 安井稔．英文法総覧．開拓社，1982．

[2] 金田一春彦編．日本語動詞のアスペクト．むぎ書房刊, 1976．

[3] 松村文芳訳　湯　延池著．中国語変形文法研究．白帝社，1982．

[4] 草薙裕．テンス・アスペクトの文法と意味．，文法と意味 I，朝倉書店，1983．

[5] 北京語学院編．中国語教科書．光生館，1958．

[6] Bernard Comrie. *Aspect.* Cambridge, 1976.

[7] Soenjono Dardjowidjojo. *Indonesian Syntax.* PhD thesis, Georgetown University, 1967.

[8] Philip N.Johnson-Laird George A. Miller. *Language and Perception.* Harvard University Press, 1976.

[9] Samsuri. *Tata Kalimat BAHASA INDONESIA.* Sastra Hudaya, 1982.