

日英時間表現の意味と対応関係の解析

的場 和幸 池原 悟 村上 仁一

鳥取大学大学院工学研究科

{matoba,ikehara,murakami}@ike.tottori-u.ac.jp

1 はじめに

従来の日英機械翻訳において、日英両言語間の時間に関する概念把握の仕方の違いが、翻訳品質を低下させる原因の1つとなっている。一般に時間概念にはテンス、アスペクトの二つのカテゴリーがあり、これを言語表現から見た場合、日本語のタ形はテンスでは過去を表しアスペクトでは完了を表すのに対し、英語の過去形はテンスの過去を表し、完了形はアスペクトの完了を表すというように異なる表現を用いており、日本語と英語ではずれが生じている。

そこで本研究では両言語の時間表現を解析し、共通の意味表現に置き換えることにより、単文、引用節・関係節を含む文について、日本語から英語への適切な時間表現の生成を試みる。

2 テンス・アスペクト体系

2.1 テンス

テンスとは、話者がある事象を時間の流れの中における一つの点として捉え、発話時からみて以前か以後かを問題にした概念である。

日本語においてテンスを担う文法形式は、ル形とタ形である。ル形、タ形とは次の例文のように、動詞の語尾がル、タで終わるものを指す。

(例文1) 私は次の月にまた 来る。 ル形

(例文2) 彼は大声を 上げた。 タ形

英語の現在形と日本語のル形はテンスの非過去、英語の過去形と日本語のタ形は過去を表す。非過去とは現在、未来を含んでいる。また英語では助動詞の「will」、日本語ではダロウ形を用いて、未来の事象を推量の意味を込めて表す。

(例文3) 彼はあとから 来るだろう。 ダロウ形

2.2 アスペクト

アスペクトとはある事象を時間的長さを持ったものとして捉え、その内的時間を問題にする概念である。日本語のタ形と英語の完了形はアスペクトの完了を表し、日本語のル形と英語の現在形は未了を表す。また、日本語のテイル形と英語の進行形は継続を表す。

(例文4) 私は何冊かの本を 失った。(完了)

I have lost a number of books.

2.3 動詞の分類

テンス・アスペクトは、以上のような文法形式だけで決まるのではなく、動詞の語彙的内容とも深く関わっている。本研究では従来の分類法 [1][2] を参考に、動詞の中にある時間的性質によって、表1に示す通り動詞を3つに分類した。

表1: 動詞分類表

動詞の種類		例
状態動詞		ある, いる
思考・知覚動詞		思う, 考える
動態動詞	動作動詞	走る, 書く
	結果動詞	着く, 終わる

状態動詞は存在や特性など、時間的な位置づけができない動詞である。思考・知覚動詞は思考や知覚・感覚といった人の内的事象を捉える動詞である。動態動詞は時間の中の運動を捉える動詞である。動態動詞はテイル形のアスペクトの違いによって、動作継続を表す動作動詞と、結果継続を表す結果動詞に分類する。

3 時間表現の翻訳

3.1 意味表現への置き換え

本研究では時間表現の翻訳のために、文法形式と動詞の時間的性質から、テンスとアスペクトを意味表現

に置き換える。

まず日本語文のアスペクトを以下の表 2 から決定する。

表 2: 日本語動詞のアスペクト

ル形 (タ形)		テイル形 (テイタ形)	
— (状態)	状態動詞	〜 (動作継続)	思考・知覚動詞、 動作動詞
● (完了、未了)	思考・知覚動詞、 動態動詞	●— (結果継続)	結果動詞

表 2 では動詞の語尾形式と時間的性質からアスペクトを 4 つに分類する。動作継続とは、ある動作がまだ進行中であることを表す。結果継続は、ある動作が基準時以前に終り、その結果生じた状態が基準時まで続いていることを表す。

次に図 1 から動詞の示す時間と発話時との時間関係を決定する。

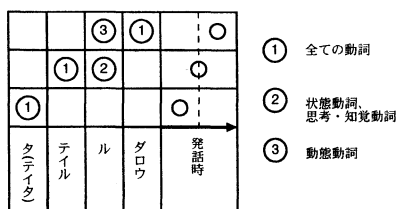


図 1: 日本語動詞のテンス

図 1 では、動詞の語尾形式から、発話時との時間関係を決定する。①は全ての動詞 (状態動詞, 思考・知覚動詞, 動態動詞) を指し、②は状態動詞, 思考・知覚動詞, ③は動態動詞をそれぞれ指している。図 1 から①の動詞がタ形をとると、発話時より過去、②の動詞がル形をとると発話時現在ということを表す。従属節を含む文については、図 1 を拡張して、発話時、主節、従属節の表す時間関係を決定する。従属節を含む文の例として、引用節を含む文のテンスを図 2 に示す。

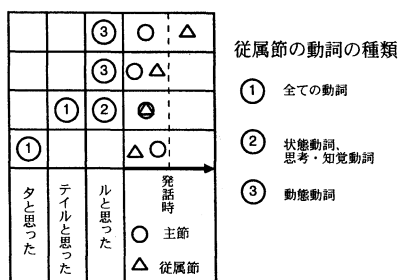


図 2: 引用節を含む日本語文のテンス

以下に日本語文から意味表現へ置き換える例を示す。

(例文 5) 大きな家がある。

(例文 6) 雲が動いていた。

例文 5 は状態動詞「ある」がル形で使われているため、アスペクトは状態、テンスは発話時現在となる。例文 6 は動作動詞「動く」がテイタ形で使われているため、アスペクトは動作継続、テンスは過去となる。これらの例文のテンスとアスペクトを意味表現に置き換えると次の図 3 のようになる。

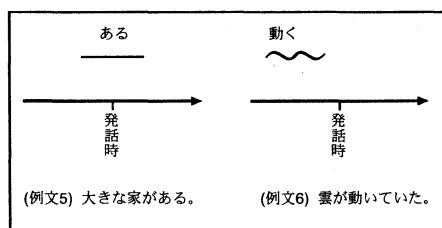


図 3: 意味表現の例

3.2 英語への変換

日本語時間表現を翻訳するため、英語についてもアスペクトとテンスを以下の表 3、図 4 から決定する。

表 3: 英語動詞のアスペクト

〜s	have 〜ed	be 〜ing
— (状態)	●— (結果継続)	〜 (動作継続)
● (完了、未了)	●— (結果継続)	〜 (動作継続)

英語の完了形には基準時以前に事態が実現し、基準時までその事態の結果が残っていることを示す。

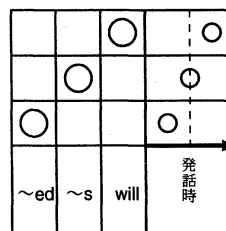


図 4: 英語動詞のテンス

英語側のテンスとアスペクトを表す図表を用いて、日本語意味表現を英語時間表現に変換する。例えば、図3の例文6、「雲が動いていた。」という文では、アスペクトが動作継続であるから、進行形となり、テンスが発話時以前であるから、過去形となる。よって英語時間表現は、過去進行形となる。

また、時間表現を日本語から英語に翻訳するにあたり、日本語動詞と英語動詞の時間表現が1対1に対応していない場合や、異なった概念で用いられる場合があるため、変換が必要になる。図5に単文における変換ルールを示す。

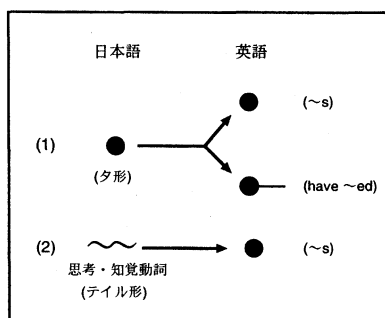


図5: 変換

タ形はアスペクトの完了とテンスの過去を表すが、英語の過去形はテンスの過去、完了形はアスペクトの完了と結果継続を表す。そのため(1)の変換規則で両者に対応させる。

(2)では思考・知覚動詞のテイル形の変換を行う。日本語においては、思考・知覚動詞がテイル形で用いられることにより動作継続を表すが、英語では動詞自体が継続状態を表す場合が多いためである。

3.3 翻訳規則表

以上のような時間表現の翻訳をもとに翻訳規則を作成する。規則の一部を表4に示す。文が表す時間的な性質においては、時間副詞も重要な役割を果たしている。そのため本研究では、時間副詞を英語への翻訳のための判断材料の一つとする。

4 実験結果

本研究で提案した、翻訳方法の精度を以下の2種類の対訳コーパスを用いて評価した。

- (1) 機能試験文集¹ (クローズドテスト)
- (2) 和英辞典 (オープンテスト)

1つの日本語表現から複数の候補が得られる場合があるため、評価を次の3つに分類する。

- 1, 候補が1つで、それが正解の場合
- 2, 複数の候補の中に正解がある場合
- 3, 候補の中に正解が含まれない場合

機能試験文集から単文700文を取り出して規則に適用した結果を表5に示す。括弧内はそれぞれの度数である。なお、機能試験文集のデータは、翻訳規則の作成に利用したため、クローズドテストである。

表5: 翻訳結果 (機能試験文集)

		1	2	3
単文	タ形 (312)	45.2%(141)	52.2%(163)	2.6%(8)
	ル形 (149)	49.0%(73)	47.6%(71)	3.4%(5)
	テイル形 (200)	70.5%(142)	0 (0)	29.5%(58)
	ダロウ形 (39)	69.2%(27)	0 (0)	30.8%(12)
平均 (700)		54.7%(383)	33.4%(234)	11.9%(83)

表4: 翻訳規則表 (単文)

動詞の語尾の形	機能的意味	時間副詞	動詞の種類	英語での動詞の時制
タ形	過去完了・結果	モウ、チョウド、スデニ、など	外的動詞 (結果動詞)	過去完了形
	現在完了・結果	モウ、チョウド、スデニ、タッタ今、など	外的動詞 (結果動詞)	現在完了形
	過去の事象	過去のある時点 (昨日、去年、など)		過去形
テイル形	動作継続		動作動詞	進行形
	結果継続		思考・知覚動詞	現在形
	現在の状態		結果動詞	現在完了形
ル形	現在の感情、感覚		状態動詞	現在形
	習慣、真理	頻度を表す副詞	内的動詞	
	未来の事象、状態	未来を指す副詞 (明日、来年など)		
ダロウ形	推量未来			未来形

¹日英機械翻訳機能を試験するための対訳文約6200文が収録されているデータベース

次に和英辞典から、単文 400 文、引用節、関係節を含む文それぞれ 200 文を取り出して実験した結果を表 6 に示す。なお、テストはオープンテストである。

表 6: 翻訳結果 (和英辞典)

			1	2	3
単文	タ形	(200)	47.0%(94)	51.5%(103)	1.5%(3)
	ル形	(100)	54.0%(54)	42.0%(42)	4.0%(4)
	テイル形	(100)	43.0%(43)	0 (0)	57.0%(57)
引用節		(200)	68.0%(136)	14.0%(28)	18.0%(43)
関係節		(200)	16.0%(32)	68.5%(137)	15.5%(31)

上記の 2 つの実験結果から、単文において動詞の時間的性質・時間副詞句より、英語の時間表現が一意に 50% 程度決まることがわかった。しかし関係節については効果が見られなかった。

5 考察

評価実験において、英語時間表現が一意に決定できなかった例と、誤った時間表現を選択した例を示す。

1. 英語時間表現が一意に決定できなかった例

(例文 7) 私は、何冊かの本を失った。

例文 7 は、まだ本を失った状態が続いている場合は完了形となり、見つかっている場合は過去形となる。このように 1 文だけでは判断ができない文が機能試験文集の単文では、約 47% 存在した。

また、特に一意正解率が低かった関係節を含む文の例を以下に示す。

(例文 8) 犬を探してくれた方には礼をします。

一般に従属節を含む日本語における従属節のル形とタ形は、発話時において事態が未了か完了かといったアスペクトを表す。そのため例文 8 では、主節の動詞が示す時点で、その行為が完了したことを示しており、発話時との時間関係を表していない (図 6)。そのため従属節でも発話時との時間的関係を示す英語での候補が 1 つに絞れない。

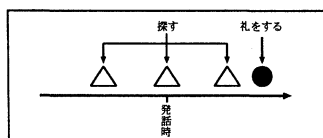


図 6: 関係節を含む例

例文 7, 8 において、正しい英語時間表現を選択するためには、前後の文脈や分野依存性などから正しいテンス・アスペクトを決定する必要がある。

2. 誤った英語時間表現を選択した例

(例文 9) 線路が私の家のそばを通っている。

The tracks passes by my house.

本研究の規則では、動作動詞「通る」のテイル形は現在進行形と訳すが、正解の訳は現在形となっている。これは、テイル形が一時的な動作も永続的な状態もどちらも表せるのに対して、英語の進行形が一時的な動作しか表さないことから来ると思われる。例文 9 のような、テイル形と進行形のアスペクトのずれによって誤った候補を選択した文が、機能試験文集 200 文の中に 9 文存在した。

この例文を正確に訳すためには、動詞の時間的性質だけでなく、名詞と動詞の意味的關係などの、他の判断材料の導入が必要である。

6 おわりに

本研究では、日本語時間表現の英語への変換を、意味的な表現を介して行う方法を提案した。この方法に対訳コーパスを用いて精度評価実験を行い、単文で一意の英語時間表現が 50% 程度求まることがわかった。しかし、半数程度の文において候補が一つに絞れないという結果が出た。今後は複数の候補が出る場合の翻訳文決定のため、前後の文の時間関係から対象文の時間表現を決定したり、時間関係の分野依存性、名詞との意味的關係などの新しい判断基準を検討していく必要がある。

参考文献

- [1] 工藤真由美 (1995) : アスペクト・テンス体系とテキスト, ひつじ書房
- [2] 国立国語研究所 (1985): 現代日本語動詞のアスペクトとテンス, 秀英出版
- [3] 吉川千鶴子 (1995) : 日英比較-動詞の文法, くろしお出版