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 知能情報工学研究科
〒 680-0947 鳥取県鳥取市湖山町南 4-101
{matoba,ikehara,murakami}@ike.tottori-u.ac.jp

あらまし:
日英機械翻訳においては，時間表現の意味のずれが翻訳品質を低下させる原因の一つとなっている．従来の時間表現の翻訳では，表層上の対応関係から作成された翻訳規則が用いられていた．本稿では，動詞と時間副詞の時間的性質に着目し、日英翻訳を対象に，中間表現を介して翻訳を行う方法を提案する．本手法の評価実験を，和英辞典 600 文を対象に行なったところ、73.2%の一意正解率を得た．

キーワード：機械翻訳，時間関係，テンス・アスペクト

Semantic Analysis of Time Expressions for Japanese to English Machine Translation

Kazuyuki MATOBA　Satoru IKEHARA　Jin'ichi MURAKAMI

Graduate School of Information and Knowledge Engineering, Tottori University
4-101, Minami Koyamachou, Tottori city, 680-0947, Japan

Abstract:
In Japanese to English machine translation,the difference of meanings of time expressions is one of the difficult problems. Conventional method translates time expressions using the rules from surface relation. This paper proposes a method for translating Japanese time expressions into English through intermediate representation by time properties of verb and adverbs. The proposed method yields correct time expressions in 73.2% of the cases extracted from Japanese-English dictionary.

Keywords: machine translation,temporal relation,tense and aspect

1　はじめに

従来の日英機械翻訳においては，時間表現の意味のずれが翻訳品質を低下させる原因の一つとなっている．時間表現には，事象の時間的位置を表すテンス，動作や状態の時間的な局面を表すアスペクトなどのカテゴリーがある．これを言語表現から見た場合，例えば，日本語では，同じ表現で過去と完了を表すのに対して、英語では，過去と完了では，異なる表現を用いており、両言語間でずれが生じている．

時間表現の解析によって，客観的な意味を捉える研究には，日本語動詞の時間的性質を素性によって図式化しアスペクト解析を行ったもの [1]，また，数量表現を含む名詞句の時間表現を，意味によって分類し，形式表現に変換するアルゴリズムを提案したもの [2] などがある．しかし，これらの研究は日本語側から時間表現を解析したものであり，機械翻訳を対象とはしていない．

従来の機械翻訳における時間表現の翻訳では、表層上の対応関係から作成された翻訳規則が用いられていた．しかし，言語表現上の違いを克服するためには、表層上の関係を規則化するのではなく、原言語から客観的な時間関係を抽出し，それを目的言語の枠組みの中で捉え直す仕組みが必要である．

そこで，本稿では，動詞と時間副詞の時間的性質を手がかりに、日本語時間表現から中間表現を介して英語への翻訳を行う方法を提案する．

また，本手法の翻訳精度を対訳データベースを用いて確認する．

2　時間表現の翻訳手順

本稿では，日本語時間表現の英語への翻訳を，以下の手順に沿って行う．

(1)　時間関係の抽出

テンスの過去，非過去，現在，未来を表す一般的な語尾形式「タ」「ル」「テイル (テイタ)」「ル+ダロウ」の場合について、日本語時間表現から時間関係を抽出する．抽出された時間関係を，発話時，事象時，そして，Reichenbach[3] によって導入された事象を捉える時点である参照時の，3 つの時点の関係として表す．参照時を導入することで、動作継続と結果継続の区別が可能になる．その際，日本語動詞と時間副詞を時間的性質により分類し，時間関係を特定する手がかりにする．

(2)　中間表現を介した時間表現の翻訳

日本語側から抽出した時間関係を，英語の枠組みの中で捉え直すため，英語側からも現在形，過去形，未来形，進行形，完了形があらわす時間関係を発話時，事象時，参照時からなる時点の関係で表す．翻訳は，両言語から得られた時間関係の対応づけにより行う．

(3)　制約条件を用いた時間表現の修正

英語側の用法によって時間表現に違いが現れる場合があるため，結合価パターン [4] を用いて用法の推定を行い，文法的に正しい時間表現に修正する．例えば、進行形は，一時的な動作を表す場合に用いられ、長期的な動作や状態を表す場合には用いられない．そこで，結合価パターンを調べ，長期的な動作や状態をあらわす用法の場合は，進行形から単純形への修正を行う．

以降では，語尾形式が「ル」をとるものをル形，「タ」をとるものをタ形，同様にテイル形，ダロウ形と表記する．

3　時間関係の表現形式

本稿では，文が表す時間関係を S(発話時)，E(事象時)，R(参照時) の 3 つの時点の関係として表す．R(参照時) とは事象を捉えている基準時点である．以下に例文を挙げる．

(例文 1)　私は家まで走った．

例文 1 では，「走った」という過去の事象をその時点までさかのぼって捉えている．そのため，E(事象時) と R(参照時) は，一致しており、S(発話時) 以前に位置する．つまり，時間関係は，次の図のようになる．

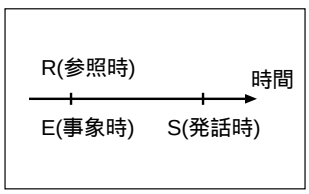

図 1: S(発話時),E(事象時),R(参照時) の時間関係

同時関係を「＝」，順序関係を「→」で表すと，図 1 は，以下のように表せる．

$$\mathbf{E{=}R \rightarrow S}$$

(例文 2)　私は彼と走っている．

例文 2 の「走っている」のように事象が動作の継続を表す場合には、事象時 E を E(P) とする．つまり，例文 2 では，現在進行中の動作を発話時から捉えているため，E(P) と R,S は，一致し次のようになる．

$$\mathbf{E(P){=}R{=}S}$$

4　時間表現の解析

4.1　動詞分類

テンス・アスペクトは，動詞の語彙的内容とも深く関わっているため，動詞の分類が必要となる．本稿では，従来の分類法 [5] を参考に，動詞を時間的性質から状態動詞，動作動詞，変化動詞に分類する．

表 1: 動詞分類表

動詞の種類		例
状態動詞		ある，いる
動作動詞	内的動作動詞	思う，考える
	外的動作動詞	走る，書く
変化動詞		着く，終わる

状態動詞は，存在や属性など，主体の状態を表す動詞である．動作動詞は，主体の動作を表す動詞である．動作動詞をさらに思考や感覚など人の内的な事象を捉える内的動作動詞と「走る」や「飛ぶ」のように人や物の動態的な運動を捉える外的運動動詞に分類する．変化動詞は，主体の変化を表し，運動終了後，主体の状態が変わる動詞である．

3 節で述べた S，E，R の関係は，動詞の時間的性質に依存している．さきほど分類した動詞それぞれに対する時間関係規則を以下に示す．なお「＊」は，すべての動詞にあてはまることを示している．

1. [ル形]
状態動詞，内的動作動詞：$\mathbf{S{=}R{=}E}$
外的動作動詞，変化動詞：$\mathbf{S{=}R{=}E}$ or $\mathbf{S \rightarrow R{=}E}$

(例文 3)　ここに本がたくさんある．
＜状態動詞＞

(例文 4)　私は君の成功を祈る．
＜内的動作動詞＞

(例文 5)　彼は 50m を泳ぐ．
＜外的動作動詞＞

(例文 6)　彼は、八時すぎに家を出る．
＜変化動詞＞

状態動詞のル形は，現在の状態，内的動作動詞のル形は，現在の思考・感情を表している．例文 3 では「ある」という現在の状態，例文 4 は，現在の思考を発話時から捉えている．例文 5，6 のように外的動作動詞・変化動詞がル形をとると，現在，または未来の事象を表す．

2. [タ形]
＊：$\mathbf{E{=}R \rightarrow S}$

(例文 7)　彼らは彼のことを笑った．
＜動作動詞＞

状態動詞のタ形は，過去の状態，動態動詞は，過去の動作などの過去の出来事を，事象時から捉えている．例文 7 では「笑った」という過去の事象を，その時点にさかのぼって捉えている．

3. [テイル形]
動作動詞：$\mathbf{E(P){=}R{=}S}$
変化動詞：$\mathbf{E \rightarrow R{=}S}$

(例文 8)　彼は本を読んでいる．
＜動作動詞＞

(例文 9)　窓が開いている．
＜変化動詞＞

動作動詞のテイル形は，現在進行中の動作，変化動詞では，過去に発生した動作を発話時から捉えている．例文 8 では，発話時において進行中の「読む」という動作を捉えている．例文 9 では，過去に発生した「開く」という事象の結果を発話時から捉えている．

4. [テイタ] 形
動作動詞：$\mathbf{E(P)}{=}\mathbf{R} \rightarrow \mathbf{S}$
変化動詞：$\mathbf{E} \rightarrow \mathbf{R} \rightarrow \mathbf{S}$

(例文 10)　彼は本を読んでいた．
＜動作動詞＞

(例文 11)　窓が開いていた．
＜変化動詞＞

動作動詞のテイタ形は，進行中の動作，変化動詞は，完了した動作の結果を，過去のある時点から捉えている．例文 10 では「読む」という進行中の動作を過去のある時点から捉えている．例文 11 は，過去に発生した「開く」という事象の結果を過去のある時点から捉えている．

5. [ル+ダロウ形]
＊：$\mathbf{S} \rightarrow \mathbf{R}{=}\mathbf{E}$

(例文 12)　彼女は必ず来るだろう．
＜変化動詞＞

動詞のル形にダロウがつくと，発話時点で事象が未然であることを推量の意味をこめて表す．例文 12 は，発話時において未然の「来る」という事象を捉えている．

4.2　時間副詞

文が表す時間関係には，時間副詞も重要な役割を果たしている．本稿では，時間副詞を 4 つの属性により分類し，時間関係の詳細化に用いる．分類表を表 2 に示し、各属性の説明を以下に述べる．

a. 設定時
時間副詞が設定する時点である．E (事象時) を指す表現と，R (参照時) を指す表現に分類する．例えば「翌日に」では，翌日のある事象を指しているが、「翌日までには」では，ある事象が翌日までに完了していることを表すため、参照時を指す．

b. 基準時
設定時を特定する際に、時間副詞のみで位置を決定できる絶対表現と，基準時を必要とする相対表現に分類する．例えば「2001 年に」は，絶対表現である．相対表現は，さらに発話時を基準時とする表現、文脈によって決定される内容時を基準とする表現に分類する「昨日」は，発話時の 1 日前を指しているため、発話時基準「翌日」は，文脈によって示される内容時の 1 日後を指しているため，内容時基準である．また基準時が判別不可能な表現は，不明に分類する「3 時から」などがあてはまる．

c. 基準時とのテンス
基準時が相対表現の場合には，基準時から見た指示時とのテンスを以前、以後，同時関係に分類する「昨日」「明日」「今日」は，発話時基準でそれぞれ以前，以後、同時を指す．

d. 時間区間
時間副詞が表す区間を，時点を明示する表現，開始時・終了時のみを明示する表現、間隔を明示する表現に分類する「翌日に」「翌日から」「翌日まで」は，それぞれ時点，開始時，終了時を指す．間隔を明示する表現には「3 日間」などがある．

その他に、「よく」や「いつも」「3 日おき」や「毎日」のような習慣，反復を表す副詞も対象とする．

4.1 節で得られた時間関係は，特定の時間副詞を伴った場合に詳細化できる「動作動詞 + テイル」において，時間副詞から時間関係を詳細化する規則を表 3 に示す．表 3 は該当する時間副詞の属性と，文全体が表すアスペクト、そして特定された時間関係からなる．なお表中の空欄は，特に指定する必要がない場合である．3 つの規則それぞれに例文を示す．

表 2: 時間副詞の分類表

属性	属性値
設定時	事象時，　参照時
基準時	絶対表現，　相対表現 (発話時，内容時)，　不明
基準時とのテンス	以前，　以後，　同時
時間区間	時点，　開始時，　終了時，　間隔

表 3: 時間関係決定規則 (動作動詞+テイル形)

	設定時	基準時	テンス	時間区間	文のアスペクト	時間関係
規則 1	事象時	相対表現 (発話時)	以前	開始時	発話時までの動作継続	E(P) → R=S
規則 2	事象時			間隔	発話時までの動作継続	E(P) → R=S
規則 3	事象時	相対表現 (発話時)	以前	時点	過去の動作の結果	E → R=S

(例文 13)　彼は 昨日から 走っている．< 規則 1>

(例文 14)　彼は 3 時間 走っている．　< 規則 2>

(例文 15)　彼は 昨日 走っている．　< 規則 3>

一般に「動作動詞+テイル」は，発話時点における動作の継続を表す．しかし，例文 13，14 のように，開始時や間隔を表す時間副詞を伴った場合は、ある過去時から発話時までの「走る」という動作の継続を表す．また，例文 15 のように，過去のある時点を指す副詞を伴った場合は，その時点で完了した動作の結果が発話時まで継続していることを表す．

5　時間表現の翻訳

本稿では，日本語側から抽出した時間関係を，英語に変換することで時間表現の翻訳を行う．英語時間表現形式を以下の表 4 に示す．

表 4: 英語時間表現形式

	R=S	R → S	S → R
E=R	単純現在	単純過去	単純未来
E → R	現在完了	過去完了	未来完了
R → E	単純未来	過去以後	

なお，事象時が E(P) となっているものに関しては，進行形を取るものとする．

以下に例を示す．

(例文 16)　彼らは通りを歩いていた．

例文 16 では，動作動詞がテイタ形で使われているため、4.1 節の規則から時間関係は，E(P)=R → S である．表 3 から対応する英語時間表現は，単純過去形となり、事象時が E(P) であるから，過去進行形と翻訳される．

6　時間表現の制約

時間表現の翻訳にあたっては，日本語の動詞，副詞から時間関係の抽出を行い，英語側の時間表現を決定してきた．しかし，英語側の用法によって時間表現に違いが出る場合がある．

(例文 17)　娘は高等学校にかよっている．
My daughters goes to the high school.

例文 17 の動詞「かよう」は，本稿の動詞分類では，動作動詞に属するため，4.1 節から英語時間表現は進行形となる．しかし，実際の文では，単純現在形となっている．これは，一時的な動作も長期的な動作も表す日本語のテイル形と，一時的な動作にしか用いられない英語の進行形の違いにより起こる．

6.1　結合価パターン

用法に依存した英語時間表現を正しく翻訳するためには，前もって用法を知っておく必要がある．そこで，本稿では，結合価パターンを用いて用法の推定を行い，得られた情報をもとに英語時間表現を修正する．

結合価パターンは，名詞と用言の関係を記述したものであり，文法的，意味的な情報を含んでいる．本稿では，日本語語彙大系 [4] に掲載されている「構文意味辞書」の結合価パターンを使用する．一般の文型と慣用的表現の文型をあわせて約 14,800 件の日本語文型パターンにまとめられている．以下の表 5 に結合価パターンの例を示す．

表 5:結合価パターンの例

属性変化 (動作受け身不可)	N1(962 機械) が N2(*) で/から疲れる N1 be fatigued with N2
身体変化 (状態受け身不可)	N1(4 人 535 動物) が N2(*) で疲れる N1 get tired of N2
属性 (状態受け身不可)	N1(4 人 535 動物) が N2(*) で/に 疲れるている N1 be tired from N2

結合価パターンには，意味属性と，その用言が持つ日本語の文型パターンを示し，対応する英語の文型パターンが記載されている．意味属性の右側には，変形情報が “()” で囲って示してある．変形情報とは，英語の文型パターンが進行形，受け身形に変形しうるのかを示している．(動作) が進行形可，(状態) が進行形不可を表す．また，特定の助動詞をともなって英語の文型パターンが変化する場合には，助動詞を小書きにして後ろに付記してある．

この結合価パターンでは，格要素の名詞の代わりに一般名詞意味属性で記述されている．一般名詞意味属性とは，名詞の意味的用法を整理，体系化したシソーラスである．

なお，結合価パターンの選択には，格要素が最も多く一致したパターンを選択する．

6.2 動作動詞+テイル形に関する制約条件

例文 17 で示した誤りを避けるには，英語側の動詞が進行形にならない動詞か知る必要がある．そこで，日本語側から得られた時間関係の事象時が E(P) であり，かつ結合価パターンの変形情報が (状態) の場合は，E に修正する．

例文 17 に適用すると「かよう」は，以下の図 2 の結合価パターンにマッチし，変形情報は，(状態) であるから，事象時を E(P) から E に修正する．

通う (かよう)
物理的移動 (状態 受け身不可)
N1(4 人 535 動物) が N2(388 場所 2610 場) から/より
N3(388 場所 2610 場) に/へ/まで 通う
N1 go from N2 to N3

図 2:例文 17 のマッチング結果

6.3 変化動詞+テイル形に関する制約条件

日本語では，「変化動詞+テイル形」で変化の結果状態を表すのに対して，英語では，最初から状態として捉え，「be 動詞 + adj」型で表すことが多い．以下に例を示す．

【日本語】	【英語】
濡れている	be wet
慣れている	be used to

従って「変化動詞+テイル形」の英語文型パターンが「be 動詞 + adj」型の時は，制約規則により英語時間表現を現在形に修正する．

7 評価実験

本稿で提案した翻訳方法の精度を対訳コーパスを用いて評価する．コーパスには，アンカー和英辞典を用い，ル形・タ形それぞれ 200 文，ル+ダロウ形・テイル (テイタ) 形それぞれ 100 文をランダムに取り出し人手で実験を行った．取り出した文はすべて単文である．

ル形については，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．

- ○．候補が一つで，それが正解の場合
- △．複数の候補の中に正解が含まれる場合
- ×．候補の中に正解が含まれない場合

以下の表 6 に結果を示す．括弧内は該当する度数である．なお実験で用いたデータは，規則の作成には無関係なためテストはオープンテストである．

表 6: 翻訳結果 (和英辞典 600 文)

	○	△	×
タ形 (200)	92.5%(185)		7.5%(15)
ル形 (200)	51.5%(103)	41.5%(83)	7.0%(14)
テイル (タ) 形 (100)	65.0%(65)		35.0%(35)
ル+ダロウ形 (100)	86.0%(86)		14.0%(14)
平均 (600)	73.2%(439)	13.8%(83)	13.0%(78)

表 6 の実験結果から正しい英語時間表現が，タ形については，92.5%，ル+ダロウ形では，86.0%求

まるという結果を得た．また，候補が一つに決定できない場合があるル形については，一意正解率が51.5%であったが，候補の中に正解が含まれる割合は，93.0%となった．テイル (テイタ) 形については，一意正解率が 65.0%にとどまった．

8 考察

8.1 タ形について

(1) 正解率について

タ形において，本手法の正解率は 92.5%であったが，タ形 200 文をすべて過去形と翻訳した場合、184 文が正解となり，正解率は 92.0% となった．本手法の正解率とあまり差がなかった原因は，時間副詞から完了形と判断した文が，実際には，単純過去形となっていた文が 3 文存在したためである．以下に例を示す．

(例文 18) かつて国鉄には多くの赤字路線があった．
JNR had many deficit-ridden lines.
正解 : 単純過去形
選択した訳 : 現在完了形

本手法では，副詞「かつて」から，完了形と判断したが，実際の訳では，過去形である．訳では，副詞「かつて」が訳出されておらず，現在とは切り放した過去の事象を述べる過去形が用いられている．

(2) 不正解文について

タ形において，時間表現を正しく翻訳できなかった文は，15 文存在するが，そのうち 7 文は，現在形とすべきところを過去形と判断した．以下に例文を示す．

(例文 19) 朝晩めっきり涼しくなった．
It is quite cool these days mornings and evenings.
正解 : 単純現在形
選択した訳 : 単純過去形

「涼しくなった」という過去の事象を直接的に捉えているのではなく、過去の事象の現在の状態を表しているために，現在形が使われていると思われる．

8.2 ル形について

(1) 正解率について

ル形については，現在の事象か未来の事象か判断できない場合があるため，評価を一意選択正解，複数選択正解，不正解の 3 つに分類した．一意に判断できない文は，86 文存在したが、すべて単純現在形と翻訳すると 76 文が正解となり，それを含めたル形 200 文の正解率は，89.5%であった．

(2) 不正解文について

ル形において不正解となった文は，14 文存在し．一意不正解文が 11 文であった．以下に誤った例を示す．

(例文 20) あすから定期試験が始まる．
Our term exams begin tomorrow.
正解 : 単純現在形
選択した訳 : 単純未来形

例文 20 では，発話時以後の事象を示す時間副詞「あすから」から，英語時間表現は，未来形となったが，実際の訳では，現在形となっている．例文 20 のように現在形と訳すべき文を，未来の事象を表す時間副詞から未来形と訳した文が 5 文存在した．未来の確定した予定を表すときに英語では，現在形が使われるためである．

8.3 テイル (テイタ) 形について

(1) 正解率について

テイル (テイタ) 形については，一意正解率が65.0%と，最も低い結果となった．日本語のテイル形は，英語では，現在進行形，現在完了形，単純現在形など様々に訳されるためである．なお，テイル (テイタ) 形の実験データ 100 文で，最も頻度の高かった英語時間表現は，単純現在形 (テイタ形では単純過去形) であり，100 文すべてを単純現在形 (単純過去形) と翻訳した場合、正解率は 56.0%になる．

(2) 結合価パターンの効果について

英語側の動詞の推定に結合価パターンを用いるこ

とで，正しい英語時間表現に翻訳できた文が，100 文中 23 文存在した．以下に例を示す．

(例文 21)　あいにく店は閉まっていた．

To my disappointment,the shop was closed.

正解　　：　単純過去形
選択した訳：　単純過去形

例文 21 がマッチした英語文型パターンは「N1 be closed」となるため，6.3 節の制約条件から単純過去形と翻訳した．

しかし，結合価パターンを用いたために，誤った英語時間表現に翻訳した文も 3 文存在した．

(例文 22)　連日の雨で大会関係者は頭を痛めている．

This spell of rainy days is giving the organizers of the tournament a big headache.

正解　　：　単純現在形
選択した訳：　現在進行形

例文 22 では，実際の英語時間表現は，現在進行形となっているが，結合価パターンから得た英語動詞は「N1 be worried about N2」(状態) となるため，単純現在形と翻訳した．結合価パターンでは「大会関係者」が頭を痛めているという状態を捉えているのに対して、実際の訳では「連日の雨」が「大会関係者」に頭痛を与えているという動作を捉えているため，誤った時間表現を選択した．しかし，この 2 文の違いは，主格が「連日の雨」か「大会関係者」かの違いであって，現在形で「N1 is worried about N2」と訳しても間違いではないと思われる．

9　おわりに

本稿では，語尾形式が「タ」「ル」「テイル (テイタ)」「ル+ダロウ」の場合における日本語時間表現の英語への翻訳を，発話時，事象時，参照時の 3 つの時点の時間関係からなる中間表現を介して行う方法を提案した．その際，日本語側の動詞，時間副詞を時間的性質により分類し、時間関係特定の手がかりとした．また，英語側の用法によって時間表現に違いが現れることがあるため，結合価パターンを用いて用法の推定を行い、正しい時間表現への修正を行った．

提案した手法を和英辞典 600 文に適用した結果，一意正解率 73.2%，複数選択正解を含む正解率 87.0%を得た．

今後は対象とする語尾形式を増やし，時間関係を詳細化することで正解率を向上させる必要がある．また，計算機に実装することで，客観的な評価を行う予定である．

参考文献

[1] 古瀬，中園，野村 (1987)：アスペクト情報の素性図式化と日本語解析への応用，情報処理学会 87-NL-63-2.

[2] 溝渕，住友，泓田，青江 (1999)：日本語時間表現の一解釈法，情報処理学会論文誌 vol.40，No.9，pp.3408-3419.

[3] Reichenbach,H (1947)：*Elements of Symbolic Logic*，Collier-Macmillan，London，England.

[4] 池原，宮崎，白井，横尾，中岩，小倉，大山，林 (1997)：日本語語彙大系，岩波書店

[5] 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