概要

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．時間表現が表す概念には，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．これら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．

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．しかし，言語表現上の違いを克服するためには，表現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なかで捉え直す仕組みが必要である．

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．

目 次

1 はじめに **1**

2 時間表現について **2**

3 時間表現の翻訳手順 **3**

4 時間関係の表現形式 **5**

5 時間表現の解析 **7**

5.1 日本語動詞の解析 7

5.1.1 動詞分類 7

5.1.2 日本語動詞の時間的性質 7

5.2 時間副詞の解析 11

5.2.1 時間副詞の分類 11

5.2.2 時間副詞による拘束条件 12

5.2.3 時間副詞による時間関係の詳細化例 13

5.2.4 その他の時間副詞 14

6 時間表現の翻訳 **15**

7 時間表現の制約 **16**

7.1 結合価パターン 16

7.2 進行形に関する制約条件 18

7.3 完了形に関する制約条件 18

8 複文 (引用節) における時間関係 **19**

8.1 日本語複文における時間関係 19

8.2 英語への翻訳 20

9 評価実験 **22**

9.1 本手法の翻訳精度調査実験 22

9.2 本手法と市販ソフトとの精度比較実験 23

10 考察 **24**
10.1 各語尾形式ごとの精度調査実験について 24
10.1.1 タ形について 24
10.1.2 ル形について 25
10.1.3 テイル (テイタ) 形について 25
10.2 複文の精度調査実験について 27
10.3 本手法と市販ソフトとの精度比較実験について 28

11 おわりに **29**

図 目 次

1 時間表現の翻訳手順 4
2 発話時, 事象時, 参照時の時間関係 5
3 一般名詞意味属性体系 17
4 結合価パターンのマッチング結果 18
5 引用節を含む時間関係の例 (1) 20
6 引用節を含む時間関係の例 (2) 20
7 引用節を含む文の翻訳例 (1) 20
8 引用節を含む文の翻訳例 (2) 21
9 引用節を含む文の翻訳例 (3) 21

表 目 次

1 動詞分類表 7
2 発話時・参照時の時間関係決定規則 8
3 事象時・参照時の時間関係決定規則 8
4 時間副詞の分類表 11
5 時間副詞分類例 12
6 時間副詞による拘束条件 13
7 英語時間表現形式 15
8 結合価パターンの例 16
9 翻訳結果 (各語尾形式ごとの精度) 22
10 翻訳結果 (引用節を含む文) 23
11 翻訳結果 (本手法と市販ソフトとの精度比較) 23
12 本手法と市販ソフトとの正解・不正解の重なり 28

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．時間表現が表す概念には，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][2]，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 [3] 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言語に依存しない中間表現を介して英語への翻訳を行う方法を提案する．

本手法の精度評価のために，対訳データベースを用いて実験し，一意正解率 73.2%，複数選択正解を含めると 87.0%という結果を得た．また，本手法と市販の翻訳ソフトとの精度比較実験を行った結果，本手法の精度が 16%程度上回った．

以下，2 章では，時間表現の説明，3 章では時間表現の翻訳手順，4 章では，本稿で用いる時間関係の表現形式について述べる．また，5 章では日本語時間表現の解析，6 章では時間表現の翻訳、7 章で時間表現の制約について述べ、8 章では，複文における時間表現について検討する．9 章では評価実験を行い、10 章で考察、11 章で，今回の結論を述べる．

2　時間表現について

日本語と英語では，時間に関する認識の仕方が大きく異なることが，機械翻訳の品質低下の一因となっている．一般に、時間概念には，主にテンスとアスペクト 2 つのカテゴリーが存在する．テンスは，話者の発話時と対象とする事象の発生時との関係で、過去，現在，未来に分類される．アスペクトは，事象の内的時間に着目したカテゴリーで，未了，完了，継続，状態などに分類される．

これらの時間概念を言語表現から見た場合，英語では，テンスとアスペクトを明確に区別するが，日本語においては，テンスとアスペクトは深く結び付いている．

以下に例文を示す．

(例文 1) 列車が到着した．

The train have arrived.

日本語において「到着した」という動詞は，発話時より過去の事象を述べていると同時に，発話時現在までの結果というアスペクトの役割も担っている．それに対して，英語では，テンスを過去形である arrived が表し、アスペクトは，現在完了形が表している．

このような言語表現上の違いは、それぞれの言語の持つ認識の違いに起因していると考えられるため、機械翻訳において，言語表現上の違いを克服するためには，単に、両言語間の表現上の対応関係を規則化するのではなく、原言語の表現から対象間の客観的な時間の関係を抽出した後、得られた客観的な時間関係を目的言語の枠組みの中で表現し直すような仕組みを実現することが望まれる．

そこで，本稿では，動詞と時間副詞を時間的性質により分類し，日本語時間表現から言語に依存しない中間表現を介して英語への翻訳を行う方法を提案する．

3　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[4] によって導入された事象を捉える基準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、発話時と事象時の 2 時点では不可能だった動作継続と結果継続の区別が可能になる．時間関係抽出の際には，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

日本語から見て同じ時間表現であっても，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，本研究では、結合価パターン [5] 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が「 ~ テイル」は，どちらにも用いることができる .. 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．

以上の翻訳手順の流れ図を図 1 に示す．

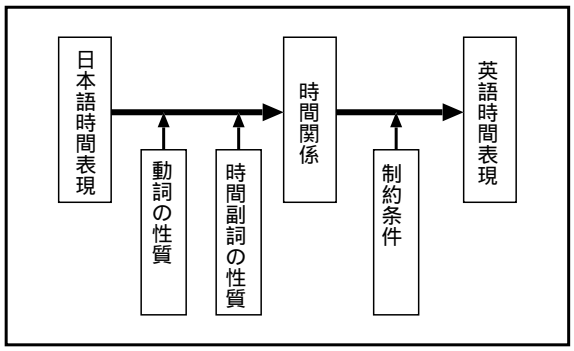

図 1: 時間表現の翻訳手順

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．

(例文 2) 雨が降る． < ル形 >

(例文 3) 雨が降った． < タ形 >

4　時間関係の表現形式

本稿では，文が表す時間関係を Reichenbach[4] によって導入された，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．S(発話時) とは，発話が与えられた時点であり，E(事象時) は，事象が発生した時点である．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 4) 私は家まで走った．

例文 4 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．

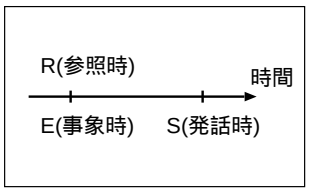

図 2: S(発話時),E(事象時),R(参照時) の時間関係

同時関係を「$=$」，順序関係を「$\rightarrow$」で表すと，図 2 は，以下のように表せる．

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

以降では上記の表現形式を用いることにする．

(例文 5) 私は彼と走っている．

例文 5 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 5 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．

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

(例文 6) 列車が着いている．

例文 6 では，発話時以前に完了した「着いた」という事象の結果を発話時から捉えている．よって時間関係は，以下のようになる．

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

以上の表現形式において，S(発話時) と E(事象時) は，直接関係しておらず，R(参照時) を介して関係を持っている．つまり，$\mathbf{S \rightarrow E \rightarrow R}$ と $\mathbf{S{=}E \rightarrow R}$ の間に違いはないものとする．

5　時間表現の解析

5.1　日本語動詞の解析

5.1.1　動詞分類

日本語において時間的性質は，動詞の語彙的内容とも深く関わっているため，時間関係を抽出するにあたって，動詞の分類が必要となる．本稿では，従来の分類法 [6] 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．形容詞も状態動詞に含める．動作動詞は，主体の動作を表す動詞である．本研究では，従来の分類に加えて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．

5.1.2　日本語動詞の時間的性質

日本語動詞における時間関係の解析においては，時間表現から発話時、参照時の関係，事象時，参照時の関係をそれぞれ抽出し，それを結合することにより行う．

以下の表 2 で語尾形式と動詞の分類から発話時と参照時の時間関係を決定する規則を示す．なお，表中の「＊」は，すべての動詞に適用可能である．発話時と，事象を捉える基準時である参照時との関係は、時間概念のテンスに相当する．

表 2: 発話時・参照時の時間関係決定規則

語尾形式	動詞の種類	時間関係
ル形	状態動詞，内的動作動詞	**S=R**
	外的動作動詞，変化動詞	**S=R** or **S → R**
テイル形	*	**S=R**
タ (テイタ) 形	*	**R → S**

つまり「走った」では，過去の事象を捉えているため，発話時と参照時の関係は「R → S」となり「走っている」では，発話時において捉えているため「S=R」となる．

また，以下の表 3 で，語尾形式と動詞の時間的性質から事象時と参照時の関係を抽出する規則を示す．事象時と，参照時の関係は時間概念のアスペクトに相当する．

表 3: 事象時・参照時の時間関係決定規則

動詞の種類		完了相 (ル形，タ形)	継続相 (テイル，テイタ形)
状態動詞		**E=R**	
動作動詞	内的動作動詞	**E=R**	**E(P)=R**
	外的動作動詞	**E=R**	**E(P)=R**
変化動詞		**E=R**	**E → R**

状態動詞はテイル (テイタ) 形を取らないため，空欄となっている．

以下に例文を示す．

(例文 7) ここに本がたくさんあった．

例文 7 は，状態動詞のタ形であるから，発話時と参照時の関係は，表 2 から「R → S」．事象時と参照時の関係は，表 3 から「E=R」となる．よって，発話時、事象時，参照時の関係は「E=R → S」となる．

また，表には示していないが、ダロウ形は文法カテゴリーのモダリティーである推量を意味する．本来はモダリティーに関する規則も作成する必要があるが，モダリティーを担う助動詞は多数あるため，今回は「ル+ダロウ」のみ扱うことにする．[1]ダロウ形が，終止形と共に用いられると，未来の事象を推量をこめて述べる．つまり，発話時と，参照時の時間関係は次のようになる．

ル+ダロウ形　　**S → R**

[1]本稿で用いた発話時・事象時・参照時の時間関係では，ダロウ形が持つ推量の意味を表すことはできない．そのため，モダリティーを含めた時間表現の翻訳を考えるのであれば，新たな視点を導入する必要がある．

以下に、表 2，3 から決定される時間的性質を，例とともに説明する．

1. [ル形]
状態動詞，内的動作動詞：**S=R=E**
外的動作動詞，変化動詞：**S=R=E** or **S → R=E**

(例文 8) ここに本がたくさんある． ＜ 状態動詞 ＞

(例文 9) 私は君の成功を祈る． ＜ 内的動作動詞 ＞

(例文 10) 彼は $50m$ を泳ぐ． ＜ 外的動作動詞 ＞

(例文 11) 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 8 では「ある」という現在の状態，例文 9 は，現在の思考を発話時から捉えている．例文 10，例文 11 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]
＊：**E=R → S**

(例文 12) 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，動作動詞，変化動詞は，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 12 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：**S=R=E(P)**
変化動詞：**E → R=S**

(例文 13) 彼は本を読んでいる． ＜ 動作動詞 ＞

(例文 14) 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 13 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 14 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 15) 彼は本を読んでいた． ＜ 動作動詞 ＞

(例文 16) 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．つまり，テイル形において事象時と参照時を一つ過去にずらしたものとなる．例文 15 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 16 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 17) 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 17 は，発話時において未然の「来る」という事象を捉えている．

5.2 時間副詞の解析

5.2.1 時間副詞の分類

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，時間副詞をその属性により分類し，時間関係の詳細化に用いる．時間関係の詳細化についての説明を以下に示す．

(例文 18) 父は、明日仕事の関係で東京に行く．

例文 18 が与えられた場合、「行く」は，外的動作動詞のル形であるから 5.1.2 節の規則より，以下に示す通り現在の事実や習慣を表す場合と、未来の事象を表す場合の 2 通りの時間関係が得られる．

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

しかし，例文 18 の時間副詞「明日」から時間関係をさらに詳細化できる．この時「明日」は，事象「行く」の時間が発話時以後に起ったことを表している．そのため時間関係が「$\mathbf{S \rightarrow R{=}E}$」の一つに絞ることができる．本稿で分類した属性を表 4 に示し、各属性の説明を以下に述べる．

表 4: 時間副詞の分類表

属性	属性値
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

a. 基準時

時間副詞の指す時を特定する際に、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．

b. 基準時とのテンス

基準時が相対表現の場合には，基準時から見た設定時とのテンスを以前、以後，同時関係に分類する．同時とは、発話時を含んでおり以前か以後か文脈によって決定される表現である「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．

c. 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，間隔を表す表現に分類する「翌日に」「翌日から」「翌日まで」「3 日毎」は，それぞれ時点，開始時，終了時，間隔を指す．

その他に、「よく」や「いつも」「毎日」のような習慣，反復を表す副詞も対象とする．以下の表 5 に分類例を示す．

表 5: 時間副詞分類例

	明日	翌日から	昨日まで
a. 基準時	相対表現 (発話時)	相対表現 (内容時)	相対表現 (発話時)
b. 基準時とのテンス	以後	以後	以前
c. 時間区間	時点	開始時	終了時

表 5 では「明日」は，発話時の一日後を指しており、時間区間は時点である「翌日から」では，ある内容時の翌日以降を表すために，基準時は基準時は内容時基準であり，時間区間は開始時である「昨日まで」では，発話時の一日前までを表している．

また表 4 で分類できない，習慣や反復を表す時間副詞にも，拘束条件を適用する．

5.2.2 時間副詞による拘束条件

表 4 で分類した時間副詞に拘束条件を指定することにより時間関係を詳細化する．拘束条件には，S(発話時)，R(参照時)，E(事象時) の間の関係を指定する．以下の表 6 に拘束条件を示す．基準時が絶対表現・内容時基準・不明，また，基準時とのテンスが同時に分類される時間副詞は，時間関係を指定できないため拘束条件からは除外する．

表 6: 時間副詞による拘束条件

		基準時とのテンス	
		以前	以後
時間区間	時点	$\mathbf{E \rightarrow S}$	$\mathbf{S \rightarrow R}$
	開始時	$\mathbf{E \rightarrow R}$	$\mathbf{S \rightarrow R}$
	終了時	$\mathbf{E \rightarrow S}$	?
	間隔	$\mathbf{E{=}R}$ ($\mathbf{E(P)}$ のときは $\mathbf{E}$ に)	

基準時とのテンスが「以後」の場合、時間区間が「時点」・「開始時」とも，発話時以後を基準にして未来の事象を捉えている．そのため，拘束条件を「$\mathbf{S \rightarrow R}$」とする．時間区間が「終了時」の時は，「明日まで勉強をしている．」のように現在も継続中の可能性があるため、時間関係の判別が不可能である．

基準時とのテンスが「以前」であり，時間区間が「時点」・「終了時」の時は、発話時以前の事象を述べているため拘束条件を「$\mathbf{E \rightarrow S}$」とする．ここで，拘束条件を「$\mathbf{R \rightarrow S}$」でなく「$\mathbf{E \rightarrow S}$」としたのは、「彼は 2 年前に死んでいる．」のように，現在を基準に過去の事象を捉える場合があるためである．時間区間が「開始時」の時は、過去に開始した事象の基準時までの継続を表しているため，「$\mathbf{E \rightarrow R}$」とする．

時間区間が「間隔」の場合は，基準時とのテンスが「以前」・「以後」に関わりなく、習慣的な事象をひとまとまりとして捉えているため拘束条件を「$\mathbf{E{=}R}$」とする．テイル形であっても，事象の継続性が失われるため，事象時が $\mathbf{E(P)}$ をとることはないとする．

5.2.3 時間副詞による時間関係の詳細化例

すでに述べたように，時間副詞には時間関係を詳細化する役割がある．表 2，3 で求めた時間関係を，表 6 で指定した拘束条件を用いて詳細化する方法を説明する．以下に例を示す．

(例文 19) 明日運動会がある．

例文 19 の時間関係は，状態動詞のル形より表 2，3 から

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

となる．ここで時間副詞「明日」を用いると，表 6 より，拘束条件は「$\mathbf{S \rightarrow R}$」であるから，詳細化された時間関係は次の通りになる．

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

(例文 20) 昨日から雨が降っている．

例文 20 の時間関係は，外的動作動詞のテイル形から

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

となる．時間副詞「昨日から」は，基準時とのテンスが「以前」，時間区間は「開始時」となり，拘束条件は表 6 から「$\mathbf{E \rightarrow R}$」となる．よって，時間関係は，次のように詳細化される．

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

5.2.4　その他の時間副詞

表 4 で分類できない時間副詞についても拘束条件を指定する．本稿では，習慣と完了を表す副詞を対象とする．

・「習慣」を表す時間副詞
(例) いつも，よく，つねに，しょっちゅう，このごろ
(例文) 彼は，いつも 廊下を走っている．
(拘束条件) $\mathbf{E{=}R}$，$\mathbf{E(P)}$ のときは $\mathbf{E}$ に

表 6 で，時間区間が「間隔」を表す時間副詞と似た性質を持つ副詞である．つまり，例で示した副詞がテイル形と共に用いられても，進行中の動作ではなく，習慣や反復性を持つ事象をひとまとまりとして捉えている．拘束条件には表 6 における時間区間が「間隔」を表す時間副詞と同じ条件を指定する．

・「完了」を表す時間副詞
(例) もう，かって，すでに，ちょうど，とっくに
(例文) 雨は とっくに やんでいる．
(拘束条件) タ形・テイル形 ： $\mathbf{E \rightarrow R{=}S}$
テイタ形 ： $\mathbf{E \rightarrow R \rightarrow S}$

完了を表す時間副詞がタ形・テイル形の文に用いられたときは，過去の事象の完了を発話時から捉え，テイタ形の文と用いられたときは，過去の事象の完了をある過去時から捉えている．上の条件では，発話時，事象時，参照時を 3 つとも条件として指定しているため，どのような動詞であっても時間関係は一つに絞られる．

6　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．単文における英語時間表現形式を以下の表 7 に示す [4]．

表 7: 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．空欄は，対応する時間表現がないことを表す．

以下に例を示す．

(例文 21) 彼らは通りを歩いていた．

例文 21 では，動作動詞がテイタ形で使われているため、5.1.2 節の規則から時間関係は次の通りである．

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

表 7 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．

7　時間表現の制約

時間表現の翻訳にあたっては，原言語である日本語の動詞，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，目的言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 22) 娘は高等学校にかよっている．
　　My daughters goes to the high school.

例文 22 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，5.1.2 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．

目的言語側の用法に依存した誤りを避けるために，本稿では，目的言語の文法に沿った時間表現に修正する制約条件を導入する．

7.1　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 [5] 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．以下に進行形に関する制約条件の説明をする．

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 [5] 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 8 に結合価パターンの例を示す．

表 8: 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．一般名詞意味属性体系の概略図を図 3 に示す．

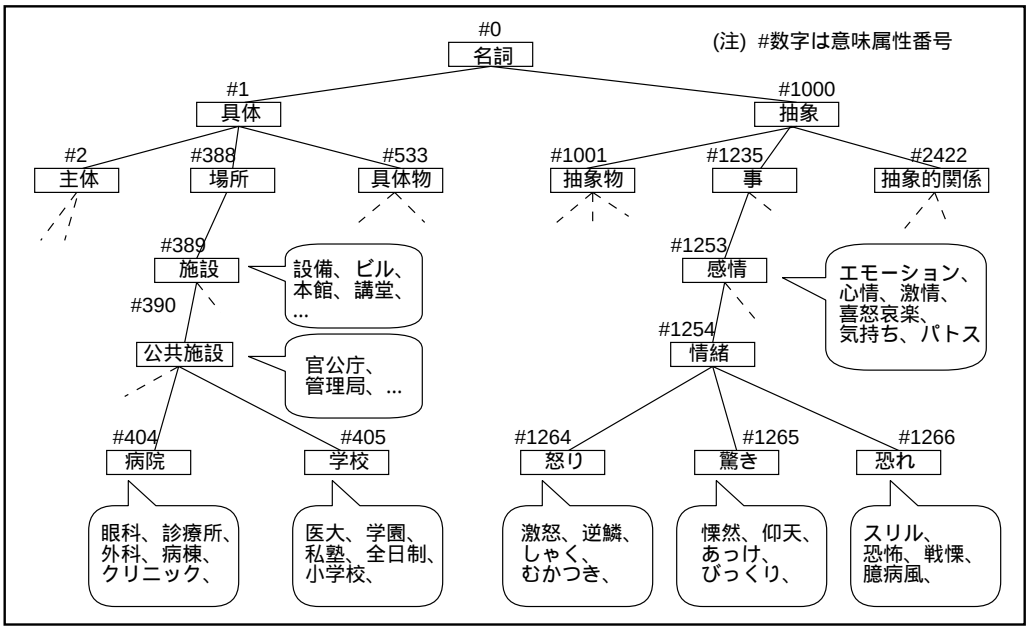

図 3: 一般名詞意味属性体系

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．本論文で使用する一般名詞意味属性体系は、約 40 万語の名詞を、最大 12 段の木構造を構成する 2,710 の意味属性に分類されている。

また、一般名詞意味属性体系は、木構造を基本構成としているため、上位の意味属性の性質を、下位の意味属性の性質に伝搬・継承できるという性質がある。この性質により、下位の意味属性を定義するための記述量を減らしている。結合価パターンは、この上位下位の関係を利用してパターンの照合を行う。

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．

7.2　進行形に関する制約条件

例文 22 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 22 に適用すると「かよう」は，以下の図 4 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 4: 例文 22 のマッチング結果

7.3　完了形に関する制約条件

日本語では「変化動詞+テイル (テイタ) 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，完了形ではなく「be 動詞 + adj」型や状態動詞で表すことが多い．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

「濡れている」の場合日本語では「濡れた」という結果と「濡れている」という状態の両方を捉えるのに対して，英語では「be wet」で，もともと濡れた状態を表すことが多い．従って，事象時と参照時の時間関係が完了形を表す「$\mathbf{E \rightarrow R}$」の場合，英語文型パターンが「be 動詞 + adj」型，または状態動詞の時は，制約規則により「$\mathbf{E{=}R}$」とする．この制約規則は，変化の結果を捉えるのではなく、状態そのものを捉えることを意味する．

8　複文 (引用節) における時間関係

8.1　日本語複文における時間関係

引用節とは「彼が来ると思う」などの「と」に先行する節のことである．引用節を取る動詞は「思う」「考える」などの思考動詞、「言う」「聞く」などの伝聞を表す動詞がある．

日本語における引用節の時間関係を決定するには，主節の時間を基準に決定すればよい．

(例文 23) 私は，彼がごはんを食べに来ると思った．

(例文 24) 私は，彼がごはんを食べに来たと思った．

(例文 25) 私は，彼がごはんを食べに来ると思う．

例文 23 では，主節の「私は思った」時点で引用節の「彼がごはんを食べに来る」という事象が以後に行われることを表している．例文 24 では，同じ時点で「食べに来る」という事象が過去に行われたことを表している．例文 25 では，主節の「私は思う」という発話時現在の時点で「食べに来る」という以後の事象を表している．つまり，引用節が表す時間は，主節の時間を基準に考えれば良い．

引用節を含む日本文から時間関係を決定する手順を以下に示す．

- 主節，引用節の時間関係を個別に決定する
- 引用節の S(発話時) を主節の E(事象時) に一致させる
- 主節の S(発話時) を複文の S(発話時) とする

なお，主節、引用節の時間関係を決定する際には，単文と同じ 5.1.2 節の表 2，3 に示した規則を用いる．ただし、表 2 において外的動作動詞・変化動詞の場合，発話時と参照時の関係が $\mathbf{S{=}R{=}E}$ または $\mathbf{S \rightarrow R{=}E}$ となっているが，習慣をあらわす副詞と用いられている場合を除いて，引用節については，$\mathbf{S \rightarrow R{=}E}$ とする．なぜなら，引用節の動詞のル形は，習慣的な事象というよりも，主節を基準にした未来の事象について想像したり，伝聞しているからである．

以下に例を示す．

(例文 26) 私は、彼が来ると思った．

主節の時間関係　$\mathbf{R{=}E \rightarrow S}$　　引用節の時間関係　$\mathbf{S \rightarrow R{=}E}$

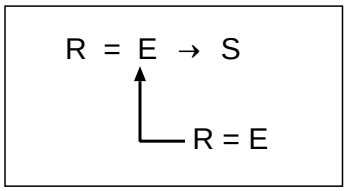

図 5: 引用節を含む時間関係の例 (1)

(例文 27) 私は、彼が来たと思った．

主節の時間関係　$\mathbf{R{=}E \rightarrow S}$　　引用節の時間関係　$\mathbf{R{=}E \rightarrow S}$

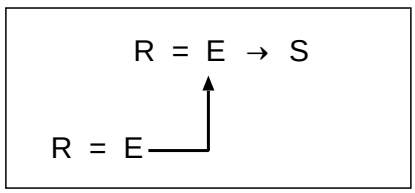

図 6: 引用節を含む時間関係の例 (2)

8.2　英語への翻訳

前節で決定した時間関係を英語へ翻訳する場合，15 ページの表 7 の規則を用いて，主節，引用節の個別に翻訳を行う．しかし英語には，主節の動詞が過去形の場合に、従属節の動詞もそれに応じて形が変化する．これを時制の一致という．

時制の一致とは、従属節の事象を，主節の事象と同じ時間基準で捉えることである．従って，本稿では，時制の一致が要求される場合，引用節の R(参照時) を主節の R に一致させるという制約条件を導入する，これにより時制の一致を満たすことができる．

例として例文 26 から抽出した時間関係を翻訳する．主節が過去形であるため，時制の一致が生じる．引用節の R(参照時) を主節の R に一致させると，図 5 の時間関係は次のようになる．

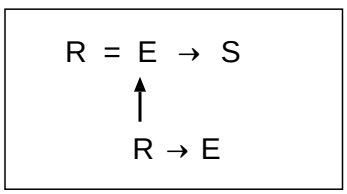

図 7: 引用節を含む文の翻訳例 (1)

よって，15 ページの表 7 の規則を用いることで，主節は単純過去形，引用節は過去以後 (would) となる．図 7 からは，引用節の「来る」という事象は、発話時以後か以前か判断できないが、時制の一致が生じ，時間表現の翻訳には，問題が生じない．

また，例文 27 から抽出した時間関係を翻訳する際も，時制の一致が生じる．つまり，図 6 は次のように変化する．

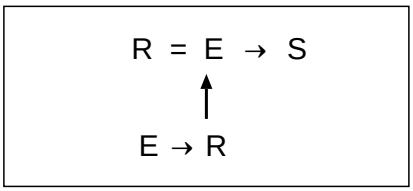

図 8: 引用節を含む文の翻訳例 (2)

図 8 から，主節は過去形，引用節は過去完了形に翻訳できる．

次に，例文 28 のような主節が未来の事象を表し，引用節の動詞が夕形の例を示す．

(例文 28) 私は、彼が来たと思うだろう．

主節の時間関係　$\mathbf{S{=}R \rightarrow E}$　　引用節の時間関係　$\mathbf{E{=}R \rightarrow S}$

主節の動詞が過去形ではないため時制の一致は生じない．よって複文全体の時間関係は次のようになる．

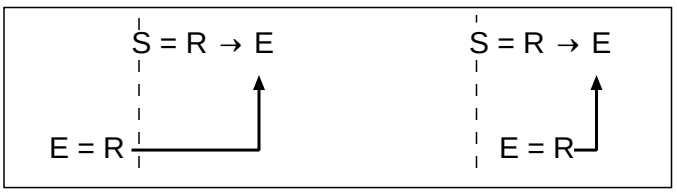

図 9: 引用節を含む文の翻訳例 (3)

例文 28 において引用節の示す時間は，主節の時間以前となるが，主節の時間は発話時以後の事象を表しているため，引用節の事象と発話時の関係が図 9 に示すように過去か未来か判断できない．実際の機械翻訳システムにおいては、時間表現がただ一つ求まることが必要とされるが，本稿では時間表現が一意に決定できない例として扱うことにする．

9　評価実験

本稿で提案した翻訳方法の精度評価のために，以下の 2 通りの実験を行った．

1. 本手法の翻訳精度調査
2. 本手法と市販ソフトとの精度比較調査

1 の実験では，対象とした語尾形式ごとの評価と，複文 (引用節を含む文) についての評価を行った．また，2 の実験では，精度比較対象としてオムロンソフトウェア株式会社の「翻訳魂」を使用した．なお，どちらの実験も本手法の翻訳は人手で行い，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．

9.1　本手法の翻訳精度調査実験

(1) 対訳コーパス (アンカー和英辞典) から，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，英訳の時間表現と比較し正解率を求めた．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 9 に結果を示す．括弧内は該当する度数を意味する．

表 9: 翻訳結果 (各語尾形式ごとの精度)

	○	△	×
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 9 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．

(2) 対訳コーパス (外国人のための日本語例文・問題シリーズ 15 テンス・アスペクト [8]，アンカー和英辞典) から，引用節を含む日本語文 200 文を取り出し英訳の時間表現と比較し正解率を求めた．結果を以下の表 10 に示す．

表 10: 翻訳結果 (引用節を含む文)

正解	不正解	一意決定不可能
79.5%(159)	20.0%(40)	0.5%(1)

なお，一意決定不可能の一文は，8.2 節の例文 28 で述べた引用節の示す時間が，発話時以後か以前か判断できない文である．

9.2 本手法と市販ソフトとの精度比較実験

対訳コーパス (外国人のための日本語例文・問題シリーズ 15 テンス・アスペクト [8]) から，本稿で対象とした語尾形式を持つ単文 151 文を取り出し，本手法と市販ソフトとの精度比較実験を行った．本手法では，ル形において現在形と，未来形の 2 つの候補が得られる場合があるが，市販ソフトはただ一つの時間表現を出力する．よって，実験結果比較のために 9.1 節の実験結果から[2]，ル形について複数の候補が出た場合については、デフォルトとして単純現在形を選択することにする．以下の表 11 に結果を示す．括弧内は該当する度数を意味する．

表 11: 翻訳結果 (本手法と市販ソフトとの精度比較)

	○	×
本手法	68.2%(103)	31.8%(48)
市販ソフト	51.7%(78)	48.3%(73)

本手法のル形において，現在形と未来形の 2 つの候補が求まった文は 151 文中 6 文存在した．デフォルトとして単純現在形を選択した場合，3 文が正解となり，残る 3 文は，単純未来形であった．

[2]ル形について，時間表現を一意に選択できない文は 86 文存在したが，すべて単純現在形と翻訳すると 76 文が正解となった.

10　考察

10.1　各語尾形式ごとの精度調査実験について

10.1.1　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 29) 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 30) 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．

10.1.2　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 31) 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 31 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 31 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．

10.1.3　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 32) 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 32 がマッチした英語文型パターンは「N1 be closed」となるため，7.3 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 33) 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 33 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，例文 33 と結合価パターンから得た文の違いは，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．

10.2　複文の精度調査実験について

複文 (引用節を含む文) の精度評価実験の結果，正解率は 79.5%と単文の平均一意正解率 (73.2%) より高い値となった．原因として単文において正解率を下げる要因となった語尾形式テイル (タ) を含む文が少なかったことが考えられる．

不正解となった文は 40 文存在したが，9 文が主節，引用節の動詞ともにタ形であるにも関わらず，英語の引用節の時間表現が単純過去形となっている．以下に例を示す．

(例文 34) 医者は娘が危ない状態を脱したと言った．

The doctor told us that our daughter was out of danger.

正解　　：　主節 単純過去形　引用節 単純過去形
選択した訳：　主節 単純過去形　引用節 過去完了形

タ形における評価実験の考察でも述べたが，主節の事象時「言った」において「脱した」という事象を表しているのではなく「be out of」を使って脱した状態を表していると思われる．前もって「be out of」という英語動詞の形がわかれば単純過去形への翻訳は可能であるため，結合価パターンを用いた規則を導入することで精度の改善が望める．しかし，例文 34 では，得られるパターンは「escape from」であり「be out of」というパターンを得ることはできない．

10.3　本手法と市販ソフトとの精度比較実験について

表 11 のように本手法の翻訳精度は 68.2%と，アンカー和英辞典を用いたときより 5%低下した．実験に使用した文が，テンス・アスペクトについて学習するためのテキストに対訳をつけたものであり，進行形や完了形もまんべんなく網羅しているために精度が低下したと思われる．

また，本手法と市販ソフトの正解・不正解の重なりを調べた．以下の表 12 に示す．

表 12: 本手法と市販ソフトとの正解・不正解の重なり

両方正解	本手法のみ正解	市販ソフトのみ正解	両方不正解
65 文	38 文	13 文	35 文

表 12 から本手法では誤ったが，市販ソフトは正解を出力した文は 13 文存在した．13 文中 10 文がテイル形を誤った英語時間表現に翻訳し，残る 3 文は，動詞の時間的性質から正しい時間表現に翻訳できたにも関わらず、時間副詞の情報を使うことで，誤った時間表現に変更している．時間副詞によって誤った時間表現に翻訳した例を以下に示す．

(例文 35) 彼は、きっと十年後に一流のピアニストになる。

He is certain to become a first rate pianist ten years from now.

正解　　: 単純現在形
本手法の訳　: 単純未来形
市販ソフトの訳: 単純現在形

例文 35 においては，時間副詞「十年後」から本手法は単純未来形と翻訳した．しかし，正解は，モダリティーの確信を意味する単純現在形となっている．本研究では，テンス・アスペクトに分類される時間副詞を対象としたが「きっと」のようなモダリティーに分類される副詞についても規則を構築する必要がある．

11　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．本手法の中間表現は言語に依存しない意味表現である．中間表現へ変換する際，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う制約条件を導入した．

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．また，市販ソフトとの精度比較を行ったところ，16%程度高い翻訳精度となった．

今後は対象とする語尾形式を増やし，モダリティーに関する規則を追加することにより，さらにカバー率，正解率を向上させる必要がある．

謝辞

本研究において三年間指導して頂いた鳥取大学知能情報工学科計算機 C 研究室の 池原教授と村上助教授にお礼申し上げます。また本研究で参考にした論文，本の著者の方々にお礼申し上げます。

参考文献

[1] 古瀬，中園，野村：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-63-2，1987

[2] 江良，甲斐，中村，吉田：テンス・アスペクトを考慮した物語文章における事象間の時間関係の抽出，情報処理学会 93-NL-95-6，1993

[3] 溝渕，住友，泓田，青江：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.3408-3419，1999

[4] Reichenbach,H ： *Elements of Symbolic Logic*，Collier-Macmillan，London，England,1947

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

[6] 国立国語研究所：現代日本語動詞のアスペクトとテンス，秀英出版，1985

[7] 桃内 佳雄：日本語複文の時間的構造規則に関する基礎的考察，人工知能学会誌 vol8，No.1，pp-55-64，1993

[8] 加藤，福地：外国人のための日本語例文・問題シリーズ 15 テンス・アスペクト・ムード，荒竹出版，1987

[9] 的場，池原，村上：日英両言語における時間表現の意味的変換規則の検討，第 61 回情報処理学会全国大会 1T-3，2000

[10] 的場，池原，村上：日英時間表現の意味と対応関係の解析，情報処理学会 01-NL-146-9，2001