

日英機械翻訳のための時間表現の意味と対応関係の解析

池原 悟[†] 村上 仁一^{††} 的場 和幸^{†††}

時間表現が表す概念には、話者と事象の時間的関係を表すテンス、動作や状態の時間的な側面を表すアスペクトがあるが、これらの情報を表現する仕組みは言語によって異なり、機械翻訳における訳文品質低下の一因となっている。本稿では、文末が日本語で代表的な「ル型」、「タ型」、「テイル型」、「テイタ型」、「ル+ダロウ型」の5種類の語尾形式を持つ単文と引用節を含む複文を対象に、言語独立の中間表現を介してテンスとアスペクトを英語に翻訳する方法を提案した。具体的には、まず、言語表現で表される対象と話者の時間関係を事象時 (*E*)、参照時 (*R*)、発話時 (*S*) の3つの時点の客観的な時間関係表現によって表す方法を定めた。次に、与えられた日本語に対して、文末の語尾形式と文中で使用されている動詞、時間副詞の時間的性質に着目して、時間関係表現を導く方法を示し、得られた時間関係表現から、英語表現のテンス、アスペクトを生成する方法を示した。なお、英語表現の生成では、英語文法上のテンス、アスペクトに関する制約条件を考慮し、結合価パターン辞書などを使用して、翻訳に使用された英語動詞の用法チェックする方法を示した。日英対訳標本データ (600文) に適用した結果によれば、単文の場合、一意正解率 73.2%、複数選択正解を含めると正解率は 87.0%、複文の場合の正解率は 79.5%であった。また、日本語教材に収録されたテンスアスペクト用例 (単文 151 文) を用いた実験では、本方式は、市販の翻訳ソフトと比較して 16%程度高い精度を示した。

Meaning Analysis and Correspondence of Time Expressions for Japanese to English Machine Translations

SATORU IKEHARA,[†] JIN'ICHI MURAKAMI^{††}
and KAZUYUKI MATOBA^{†††}

Japanese to English machine translation method for time expressions is proposed for the sentences which have the popular ending forms such as "ru", "ta", "teiru", "teita" and "ru + darou". And similar method is proposed for Japanese complex sentences which have quoted clauses. First, in order to represent the time relations, "Event time (*E*)", "Reference time (*R*)" and "Speech time (*S*)" are introduced and intermediate expressions for time expressions are defined by using these symbols. Next, Japanese time expressions are analyzed using the information about sentence types, verb types and adverbial phrases to obtain the intermediate time expressions. English time expressions are generated from the intermediate time expressions under the constraints of syntactic categories of Tense and Aspect in English by using Valency Pattern dictionary. The result of experimental evaluations showed that 87% of Japanese time expression in simple sentences was correctly translated and the translation accuracy for complex sentences was 79.5%. In the other experiment comparing to a current system, our method showed 16% superior to it.

1. はじめに

自然言語において、話者と対象の時間的な関係は、

通常、テンスとアスペクトによって表現される。テンスは、話者と事象の時間的対応関係を表すのに対して、アスペクトは、動作や状態など対象自身の持つ特性に着目した認識を示す¹⁾²⁾。しかし、実際の言語表現では、これらが文法的に明確に区別されているとは言えず、これらの情報を分離して抽出するのは簡単ではない。特に、異なった言語では、時間概念の認識方法にずれがあり、これらを正しく解析することは、機械翻訳にとって重要な課題の一つとなっている。

日本語研究において、文法的範疇としてのテンスとアスペクトの区別は行われてこなかったと言われてい

[†] 〒 680-8552 鳥取市湖山町南 4-101
鳥取大学工学部 Tottori University
E-mail: ikehara@ike.tottori-u.ac.jp

^{††} 〒 680-8552 鳥取市湖山町南 4-101
鳥取大学工学部 Tottori University
E-mail: murakami@ike.tottori-u.ac.jp

^{†††} 〒 115-0051 東京都北区浮間 3-10-12
ウェルストンハイツ 305 号
E-mail: matoba@ma.kitanet.ne.jp

る³⁾が、他の言語と同様、時の意味を表す様々な表現が存在する。欧米の文法理論が導入されて以来、日本語表現のテンスとアスペクトについて、多くの言語学者によって研究が行われてきた^{4)~9)}。計算機処理の立場からも、時間表現からテンスとアスペクトを解析する技術の研究が行われてきた。

例えば、日本語独自の立場からテンスとアスペクトの基本構造を調べ、文法と意味の形式を判定するアルゴリズムを提案した研究¹⁰⁾や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研究¹¹⁾¹²⁾がある。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、形式的な表現に変換するアルゴリズムを提案した研究¹³⁾などがある。最近では、複文の時の表現を解析した研究¹⁴⁾や論理学的方法によりアスペクト解析する方法として、点と区間の代わりにアロー論理を用いた研究¹⁵⁾も行われているほか、日英両言語のテンス、アスペクトを HPSG で記述する試みもある¹⁶⁾。

また、日英機械翻訳では、日本語のテンス、アスペクト、モダリティの3者に関する表現を分類して英語表現形式に対応付け、時制の一致など英語側の事情を考慮した変形によって適切な英語表現を生成する方法¹⁷⁾が一般的と言えるが、日英両言語での新聞記事の書き方の違いに着目し、計画発表関連の記事では日本語の現在形を英語の未来形に訳し、事実報告の記事では現在形に訳すなどの方法¹⁸⁾も提案されている。また、最近では、用例翻訳を応用した研究¹⁹⁾²⁰⁾や日中機械翻訳のため、テンス、アスペクトの表現を日本語 20 種類、中国語 14 種類に分類し、それらの対応関係を規則化する研究²¹⁾などが行われている。

機械翻訳の研究においては、このように、対象とする言語間での時間表現の対応関係を規則化することが主要な課題とされてきたが、言語によって時間関係をとらえるための仕組みは様々であり、それを表現する方法も様々である。多言語翻訳を考えると、言語表現上の違いを克服できるようなより汎用的な言語処理を実現することが望まれる。具体的には、言語表現から時間関係に関する情報を客観的で言語独立な情報として取り出して定式化し、それを目的言語の枠組みに合わせて生成する技術を実現することが期待される。

そこで、本稿では、日英機械翻訳において、時間表現で表わされた話者と対象との関係を事象時、参照時、発話時の3つの時間からなる言語独立の中間表現(時

間関係表現)を介して翻訳する方法を提案する。具体的には、与えられた日本文の述部の語尾形式、動詞の意味、時間副詞の持つ性質を手がかりに「時間関係表現」を求めたのち「時間関係表現」から英語表現を生成する。また、英語表現の生成では、英語側の事情に合わせ、結合価パターン辞書などを使用して、表現を修正する。

2. 時間関係の表現と翻訳の方法

2.1 日英言語の時間表現について

2.1.1 文法的範疇としてのテンスとアスペクト

言語表現において、時間関係を表現する概念には、テンス(時制)とアスペクト(相)があるが、日本語と英語では、これらの情報の表現する仕組みが異なるだけでなく、時間関係についての認識の仕方も大きく異なる。

英語では、テンスは、文法上の「時間的区分」を表す文法上の「時間区分」を表す文法上のカテゴリとして定義され、主として動詞の語形を変化させることにより、「現在時制」と「過去時制」が区別される²²⁾²³⁾。未来を表す表現も存在する。しかし、アスペクトとしては、未来を表す表現にも動詞の完了形を使用した「完了形」が存在し、さらにそれらのすべてについて「進行形」が存在するから、時間関係表現は、全体で12種類に分類される。

これに対して、日本語では、膠着言語の特徴として、動詞の現在形、過去形は存在せず、時間的関係は助動詞によって表現される。また、文型上、テンスとアスペクトの区別も明確でない。このため、欧米の文法が輸入されるまでは、動詞の語形変化で定義される文法的な範疇としてのテンスやアスペクトは定義されなかった³⁾。最近でも、例えば、日本語のテンスとアスペクトを体系化した文献⁵⁾では、「現在形」の定義はなく、それに相当する表現は「基本形」とされている。(例文1)列車が到着した。The train has arrived.

例えば、例文1の翻訳例を見ると、日本文では、テンスとアスペクトの情報を表現上分離できない。日本語の動詞「到着した」は、発話時より過去の事象を述べているが、英語訳では、現在完了に訳されている。これは、文脈にもよるが、通常、到着した列車は、ま

言語学におけるテンス、アスペクトの研究については、文献⁴⁾が詳しい。

現代英語学では、助動詞を用いた未来の時を表す表現はテンスとは称されず、テンスは「現在形」と「過去形」の2つとするのが普通とされている。また、最近では、認識論的な立場からテンスの概念そのものを見直す研究も行われている²⁴⁾が、後で述べるように、この論文では、機械翻訳を実現するため、文法的カテゴリとしてのテンスとアスペクトではなく、時の意味を表す言葉としてのテンス、アスペクトを考える。

だ、そこにいるはずだと推測して訳されたもので、英訳では、むしろそれが普通である。

このような日英言語の違いは、時間関係表現の翻訳を困難にしている。

2.1.2 意味的範疇としてのテンスとアスペクト

ところで、テンスとアスペクトを定義する方法としては、これを文法的形式を分類する言葉（文法的範疇）とするか、言語表現で表される意味を分類する言葉（意味的範疇）とするかの2つの立場が考えられる。文法的範疇としてのテンスとアスペクトについては、言語によって定義が困難な場合があるが、意味的範疇としてのテンスとアスペクトは、どの言語でも定義できる可能性がある。

文献²⁵⁾によれば、テンスは、「ある事象や状態を時間の流れの中の1点としてとらえ、それらの発話時との関係を問題とする」とされており、「単純過去」、「単純現在」、「単純未来」の3種類に分類される。表現対象とそれについて発話する人間は、共に時間的な存在である²⁶⁾ことを考えると、「テンス」に分類されるこれら3種類の時間関係は、話者と対象間に客観的に存在している相対的な関係であるから、言語の種類に依存しない概念と言うことができる¹⁾。

これに対して、「アスペクト」は、「ある事象や状態を時間の流れ中で、ある長さを持ったものとしてとらえ、それらが、時間的過程の中でどの位置に属するか属さないかに着目した概念」であり、意味的に「始まりか」、「終わりか」、「継続中か」、「瞬間的な出来事の結果が継続しているか」などを区別するものとされる²⁵⁾。このような「事象や状態の時間的位置」に関する関係も客観的な関係である²⁾から「アスペクト」も言語の種類に依存しない概念と言うことができる。

そこで、本論文では、テンスとアスペクトを言語に共通の意味的な範疇を表す言葉として使用する。

2.2 時間関係の表現形式 — テンスとアスペクトの表現形式

本論文では、話者と対象に関する時間的な関係を言語独立な形式で表現するため、Reichenbach²⁷⁾によ

て定義された「事象時 (E)」、「参照時 (R)」、「発話時 (S)」の3つの時点の概念を導入する³⁾。それぞれは、以下のように定義される。

E (事象時; Event time):

述語 (動詞) によって表現された事象が実際に発生した時点

R (参照時; Reference time):

話者が事象を見た時点

S (発話時; Speech time):

話者が発話している時点

「発話時」に加えて「参照時」を使用することによって、テンスとアスペクトは区別される。すなわち、「単純時制 (単純過去, 単純現在, 単純未来)」は、いずれも、事象時と参照時が同時である⁴⁾のに対して、「完了 (過去完了, 現在完了, 未来完了)」は、いずれも、事象時の後に参照時がくる。すなわち、話者の視点 (S) は、行為や現象そのものにあるのではなく、その結果もしくは、結果としての状態にある。

また、いずれの場合も「過去」、「現在」、「未来」の3つの区別は、「参照時」と「発話時」の関係で決定される。例えば、例文2を考える。

(例文2) 私は家まで走った。

この文では、まず、観念の話者が現れて、「走る」という事象を過去に遡って捉え、その後、観念の話者が消えて現在に戻り、現実の話者が、助動詞「た」によって、事象をとらえたのは過去であることを表現している²⁶⁾。従って、 E (事象時) と R (参照時) は、一致しており、 S (発話時) 以前に位置するから、時間関係は、次の図1のようになる。

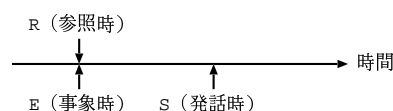


図1 発話時・事象時・参照時の関係

Fig. 1 The relations among Event time, Reference time and Speech time

¹⁾ ところで、与えられた事象や状態を時間軸上の1点としてとらえるか、ある長さを持ったものとしてとらえるかは、言語によって認識の仕方が異なることがあり、文法的範疇としてのテンスとアスペクトは、言語によって一致しない場合があるからである。後で述べるように、本論文では、このような現象については、目的言語生成の段階で、言語毎の特殊性に応じた制約条件を設け、生成された表現を修正する。

²⁾ 対象とする事象を瞬間的なものと見るか、継続的なものと見るか等については、話者のとらえ方に依存するが、その結果として表現される時間関係は解釈に依存しない客観的な関係である。

³⁾ これをテンスとアスペクトの計算に応用した例として、文献²⁸⁾がある。

⁴⁾ 同様の考え方は、日本語でも存在する。三浦つとむ²⁶⁾は、「認識と言語の理論」において、「表現対象となる事物もそれを認識する人間も時間的な存在である以上、人間は、対象と現在の関係でしか対峙することができない」という考えから、日本語の時の情報が助動詞によって表現されることに着目した解釈を示している。すなわち、それは、話者が、過去、未来の事物を表現する場合は、観念的自己分裂によって出現した話者が、過去、未来の時点に移動し、対象と現在の関係で対峙して表現した後、現実の話者が現在に戻ってその結果を表出するという解釈である。

ここで、同時関係を「=」、前後関係を「 \rightarrow 」で表すと、図1は、「 $E = R \rightarrow S$ 」となる。なお、以降では、このように3つの時点点を「=」、「 \rightarrow 」を介して結合した式を「時間関係表現」と呼ぶ。

次に、アスペクトに関する情報を表現する方法を示す。例文3では、「走る」と言う事象が継続的な事象としてとらえられている。

(例文3) 私は彼と走っている。

このように、動作の継続を表す場合には、事象時 E を $E(P)$ で表現することとする。この文では、現在進行中の動作 $E(P)$ を参照時 (R) =発話時 S から捉えているため、「 $E(P) = R = S$ 」となる。

このように、現在、過去、未来の関係では、 $E = R$ が成り立つのに対して、 $E \neq R$ となるのは、完了の場合である。

(例文4) 列車が着いている。

例えば例文4では、話者は「着く」と言う動作と対峙しているのではなく、「着く」動作の結果、すなわち動作の終了後と対峙している。また、それを発話しているのは参照時と同時と考えられるから、3時点の関係は「 $E \rightarrow R = S$ 」となる。

2.3 時間表現の翻訳手順

多言語翻訳への適用を考え、本論文では、日英言語の時間表現を前節で述べたような時間関係表現を介して翻訳する方法を提案する。本論文で提案する日英言語間の時間表現の翻訳手順を図2に示す。

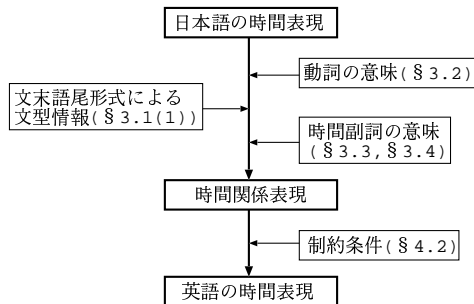


図2 時間表現の翻訳手順

Fig. 2 Translation procedure for Time Expressions

この方法は、以下の2つのステップから構成される。

(1) 日本語表現からの時間関係の抽出

与えられた日本語の述部に使用された動詞、助動詞の種類、および、文中に使用された時間副詞(時間に関する副詞的表現)の意味に着目して、「事象時 (E) 」

観念の話者は「着く」を現在の関係で捉えた後、「いる」時点に移動して、「着く」の結果を見ている。

「参照時 (R) 」、「発話時 (S) 」の3つの時点の関係抽出し、時間関係表現を生成する。

(2) 中間表現からの英語表現の生成

時間関係表現から英語表現を生成する。但し、英語表現のテンス、アスペクトの選択は、訳出に使用された動詞や文型にも依存するので、それらに関する制約条件を設け、条件を満たさない表現形式は修正する。以下、3~4章で、これら2つのステップの詳細を述べる。

3. 日本語時間表現の解析

本章では、日本語動詞の語彙的内容と時間関係副詞に基づいた時間関係決定規則について述べる。

3.1 動詞による時間関係決定規則

(1) 対象とする日本語の語尾形式

日本語は、膠着言語であり、テンス、アスペクトに関する情報の多くは文末の助動詞によって表現されるため、時間表現の解析では、動詞の活用語尾の形式を調べるのが重要である。そこで、本論文では、日本語で、現在、過去、未来の意味を表す基本的な語尾形式として「 \sim ル」(動詞文、形容詞文、ダ文などにおいて基本形(終止形)で終わる文を意味する)、「 \sim タ」、 「 \sim テイル」、「 \sim テイタ」、「 \sim ル+ダロウ」で終わる5種類の文を対象とする。以後、これらの語尾を持つ文をそれぞれ、単に「ル型」、「タ型」、「テイル型」、「テイタ型」、「ル+ダロウ型」と呼ぶ。

(2) 日本語動詞の分類

日本語の述部には、通常、動詞、形容詞が使用される。動詞、形容詞は、それぞれ、実体の持つ動的属性、静的属性を表現する語で、様々な事象を表現するために使用される。また、同一の事象でも話者のとらえ方の違いによって様々な動詞や形容詞が使用される。時間表現の解析では、どの動詞や形容詞によって対象とする動的属性が取り上げられているかを調べるのが重要である。

日本語動詞の意味とアスペクトの関係については、従来、様々な研究が行われてきた^{4)~10)}が、本稿では、文献⁴⁾の分類法に習い、動詞を時間的性質から、「状

例えば、日本語の進行形に相当する「 \sim テイル」の表現から進行形の英語表現が生成されることがあるが、英語の進行形は、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。このような場合、後に述べるように、結合価パターン³⁰⁾を用いて動詞の英語訳語の用法を調べ、文法的に正しいテンス、アスペクトを選択する。

ここでは、形容動詞は「な型形容詞」として形容詞に含める。また、体言述語は「な型形容詞」と同等に扱う。なお、動詞の意味分類は、文献⁵⁾の28分類が参考になる。

状態動詞」「動作動詞」「変化動詞」の3種類に分類する。状態動詞は、存在や静的属性など主体の状態を表す動詞であるので、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」(いわゆる「様相動詞」)と「走る」や「飛ぶ」のように人や物の動的な運動を捉える「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。以上の分類結果を表1に示す。

表1 動詞分類表
Table 1 Classification of Verbs

動詞の種類		例
状態動詞 (形容詞を含む)		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

3.2 動詞と時間関係の検討

前節で示した4種類の語尾形式を持つ日本語について、動詞の種類に着目して、事象時、参照時、発話時の関係を決定する規則を検討する。

(1) 「ル型」の決定規則

状態動詞の「ル型」は、現在の状態、内的動作動詞の「ル型」は、現在の思考、感情を表している。

(例文5) ここに本がたくさんある。<状態動詞>

(例文6) 私は君の成功を祈る。<内的動作動詞>

例文5では、「ある」という現在の状態、例文6は、現在の思考を発話時から捉えている。

(例文7) 彼は50mを泳ぐ。<外的動作動詞>

(例文8) 彼は八時すぎに家を出る。<変化動詞>

これに対して、例文(7)、(8)のように外的動作動詞、変化動詞が「ル型」をとると、現在、または未来の事象を表す。

以上から「ル型」の決定規則は以下の通りとする。

状態動詞, 内的動作動詞: $S = R = E$

外的動作動詞, 変化動詞: $S = R = E$

または, $S \rightarrow R = E$

(2) 「タ型」の決定規則

状態動詞の「タ型」は過去の状態、また動作動詞と

変化動詞の「タ型」は過去の動作、など過去の出来事を事象時から捉えている。

(例文9) 彼らは彼のことを笑った。<動作動詞>

(例文10) ここに本がたくさんあった。<状態動詞>

例文9, 例文10では、いずれも、過去の事象をその発生時点にさかのぼって捉えている。

これより「タ型」の決定規則は「 $E = R \rightarrow S$ 」とする。

(3) 「テイル型」の決定規則

動作動詞の「テイル型」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11) 彼は本を読んでいる。<動作動詞>

(例文12) 窓が開いている。<変化動詞>

例文11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。また、例文12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。

これより「テイル型」の決定規則は、下記の通りとする。

動作動詞: $S = R = E(P)$

変化動詞: $E \rightarrow R = S$

(4) 「テイタ型」の決定規則

動作動詞の「テイタ型」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「テイル型」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13) 彼は本を読んでいた。<動作動詞>

(例文14) 窓が開いていた。<変化動詞>

例文13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。また、例文14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。

これより「テイタ型」の決定規則は、下記の通りとする。

動作動詞: $E(P) = R \rightarrow S$

変化動詞: $E \rightarrow R \rightarrow S$

(5) 「ル+ダロウ型」の決定規則

動詞の「ル型」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15) 彼女は必ず来るだろう。<変化動詞>

例文15は、発話時において未然の「来る」という事象を捉えている。これより「ル+ダロウ型」の決定規則は「 $S \rightarrow R = E$ 」とする。

(6) 動詞による決定規則のまとめ

前項の検討結果から分かるように、単純時制では、

英語動詞では、Vendler²⁹⁾の分類が有名で、それを改良したCarlsonやBachの分類もある。しかし、本研究では、4章で述べるように、英語時間表現生成の段階で、必要に応じて結合パターン辞書を使用して英語動詞の用法を個別に調べることとし、英語動詞の分類は行わない。

事象時 (E) と参照時 (R) は常に同時の関係にあり、両者に前後関係が発生するのは、アスペクトが問題となるときである。

そこで、日本語動詞による時間関係の解析規則を、話者が、事象をどの時点からみているかを意味する「事象時 (E) と参照時 (R) の関係」と、見た事象をどの時点で発話しているかを意味する「参照時 (R) と発話時 (S) の関係」の決定規則に分けて整理すると表 2、表 3 の規則表が得られる。

表 2 は、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である。状態動詞は「テイル (タ) 型」を取らないため空欄となっている。

表 2 事象時・参照時の時間関係決定規則
Table 2 Rules of time relations for Event Time and Reference Time

動詞の種類	ル型, タ型	テイル型
	ル+ダロウ型	テイタ型
状態動詞 (形容詞を含む)	$E = R$	—
動作動詞	内的動作動詞	$E(P) = R$
	外的動作動詞	$E(P) = R$
変化動詞	$E = R$	$E \rightarrow R$

表 3 発話時・参照時の時間関係決定規則
Table 3 Rules of time relations for Speech Time and Reference Time

語尾形式	動詞の種類	時間関係
ル型	状態動詞, 内的動作動詞	$S = R$
	外的動作動詞, 変化動詞	$S = R$ 又は $S \rightarrow R$
テイル型	*	$S = R$
タ (テイタ) 型	*	$R \rightarrow S$
ル+ダロウ型	*	$S \rightarrow R$

また、表 3 は、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である。表中の「*」は、すべての動詞に適用可能である。

3.3 時間副詞の持つ時間特性

3.3.1 時間に関する副詞的表現の役割

前節では、文の語尾形式と動詞の種類に着目した時間関係の決定規則を考えたが、得られた規則には曖昧性があるだけでなく、必ずしも正確な時間関係が決定できない場合がある。

(例文 16) 父は、明日、仕事の関係で東京に行く。

例えば、例文 16 は、解釈に曖昧さが残る。すなわち、動詞「行く」が外的動作動詞の「ル型」であるから前節の規則より、「 $S = R = E$ 」と「 $S \rightarrow R = E$ 」の 2 つの関係が得られる。これは、現在の事実や習慣を表す場合と未来の事象を表す場合に相当する。

しかし、副詞的に使用された名詞「明日」に着目す

れば、いずれかを区別することができる。すなわち、この例では、「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から、時間関係は「 $S \rightarrow R = E$ 」と決定できる。

このように、文の表す時間関係を決定するには、文中で副詞の働きをする語や句（「副詞的表現」と言う）の情報も重要である。本稿では、副詞的表現の情報を用いて時間関係を推定することにより、前節で得られた時間関係表現の絞り込みと修正を行う。具体的には、副詞的表現をその意味によって分類し、時間関係決定のための拘束条件を設ける。

なお、以下では、副詞的表現を単に「副詞」と呼び、上例のような時の意味を持つ副詞的表現を単に「時間副詞」と呼ぶ。

3.3.2 時間副詞の分類

時間副詞をその表す意味に着目して、表 4 に示すように、「基準時特性」、「前後関係特性」、「区間特性」の 3 つの観点から分類する。以下、それぞれの意味について説明する。

表 4 時間副詞の分類法
Table 4 Time related Adverb Classification

副詞の時間特性	特性値
基準時特性	<絶対表現>, <相対表現 (発話時基準, 内容時基準)>, <不明>
前後関係特性	<以前>, <以後>, <同時>
区間特性	<時点>, <開始時>, <終了時>

(1) 基準時特性

時間を表す副詞的表現を単独で位置が決定できる<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」は、<絶対表現>であるのに対して、「明日」、「翌日」などは<相対表現>である。

<相対表現>は、さらに発話時を基準時とする<発話時基準>と文中の事象が発生した時間である内容時を基準とする<内容時基準>に分類する。例えば、「昨日」は、発話時を基準にその1日前を指しているため<発話時基準>であるが、「翌日」は、表現内容の示すある時から見て、その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。

(2) 前後関係特性

基準時特性が<相対表現>の時間副詞を「前後関係特性」によって分類する。「時間関係特性」は、時間副詞の表す時間が、基準時の前か、後か、それとも同時かを表すもので、<以前>, <以後>, <同時>の3

つに分類される。但し、<同時>は、発話時を含むこととする。例えば、「昨日」、「明日」、「今日」は、発話時基準の時間副詞で、「前後関係特性」はそれぞれ、<以前>、<以後>、<同時>に分類される。

(3) 区間特性

時間副詞が「時点を示すか」、「開始時を示すか」、「終了時を示すか」、「間隔を表すか」に着目して、<時点>、<開始時>、<終了時>、<間隔>の4種類に分類する。例えば、「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。

なお、時間副詞には含まれないが、「よく」や「いつも」、「毎日」のような副詞は、テンス、アスペクトの決定に役立つと考えられる。これらの副詞の扱いについては、3.5節で述べる。

例として、表5に「明日」、「翌日から」、「昨日まで」の3つの時間副詞が持つ時間特性を示す。

表5 時間副詞分類例

Table 5 Examples of Time Related Adverb

時間特性	時間副詞		
	「明日」	「翌日から」	「昨日まで」
基準時特性	< 相対表現 > (発話時)	< 相対表現 > (内容時)	< 相対表現 > (発話時)
前後関係特性	< 以後 >	< 以後 >	< 以前 >
区間特性	< 時点 >	< 開始時 >	< 終了時 >

表5で、「明日」は、発話時の一日後を指しており、「区間特性」は<時点>である。「翌日から」は、ある内容時の翌日以降を表すために、基準時は<内容時基準>であり、「区間特性」は<開始時>である。また、「昨日まで」では、発話時の一日前までを表しており、「時間特性」は<終了時>である。

3.4 時間副詞による拘束条件

前項で示した時間副詞の中で、事象時(E)、参照時(R)、発話時(S)の関係の決定に役立つのは、<相対表現>の<発話時基準>の時間副詞である。以下、この種の副詞によって決定可能な時間関係について述べる。

3.4.1 拘束条件の検討

(1) 「前後関係特性」が<以後>の場合

(a) 「区間特性」が<時点>と<開始時>の場合

例えば、「明日」、「明日から」の副詞が使用されている場合、話者は、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、発話時(S)と参照時(R)の関係は「 $S \rightarrow R$ 」とする。

(b) 「区間特性」が<終了時>の場合

この場合は、「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。時間副詞だけでは時間関係は決定できないため、拘束条件は設けない。

(2) 「前後関係特性」が<以前>の場合

(a) 「区間特性」が<時点>と<終了時>の場合

例えば、「昨日」、「昨日まで」の副詞が使用されている場合、話者は、発話時以前の事象を述べている。そのため、事象時(E)と発話時(S)の関係は「 $E \rightarrow S$ 」とする。ここで、時間関係を「 $R \rightarrow S$ 」でなく「 $E \rightarrow S$ 」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。

(b) 「区間特性」が<開始時>の場合

この場合は、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、事象時(E)と参照時(R)の関係は「 $E \rightarrow R$ 」とする。

(3) 「区間特性」が<間隔>の場合

例えば、「毎日」の副詞が使用されている場合は、「前後関係特性」が<以前>であるか<以後>であるかに関わりなく、習慣的な事象をひとまとまりとして捉えているため、事象時(E)と参照時(R)の関係は、「 $E \rightarrow R$ 」とする。「テイル型」であっても事象の継続性が失われるため、事象時がE(P)をとることはないとする。

なお、基準時が<絶対表現>、<内容時基準>、<不明>、および「前後関係特性」が<同時>に分類される時間副詞は時間関係が指定できないため、無視する。以上で得られた時間副詞による時間関係の決定規則を表6に示す。

表6 時間副詞による時間関係の決定規則

Table 6 Decision rules by Time Related Adverbs

区間特性	前後関係特性	
	< 以前 >	< 以後 >
< 時点 >	$E \rightarrow S$	$S \rightarrow R$
< 開始時 >	$E \rightarrow R$ (注)	$S \rightarrow R$
< 終了時 >	$E \rightarrow S$	—
< 間隔 >	$E = R$ (E(P)のときはE)	

(注)RとEの関係は「テイル型」の変化動詞を除いて $R = E$ であり、この関係が変わるのはこの場合だけである。

3.4.2 時間副詞による時間関係決定の例

表2、表3で求めた時間関係を表6に示した拘束条件を用いて修正する例を示す。

(例文17) 明日運動会がある。

まず、例文17の時間関係は表2、表3の状態動詞の「ル型」の規則より「 $S = R = E$ 」となるが、時間副詞「明日」を見ると、表6より拘束条件は「 $S \rightarrow R$ 」であり、 $R = E$ の関係は保存されるから修正後の時

間関係は「 $S \rightarrow R = E$ 」となる。

(例文 18) 昨日から雨が降っている。

次に、例文 18 の時間関係は、外的動作動詞の「テイル型」の規則から「 $S = R = E(P)$ 」となる。しかし、時間副詞「昨日から」を見ると、「前後関係特性」が「<以前>」、「時間区間」は「<開始時>」であるので、拘束条件は表 6 から「 $E \rightarrow R$ 」となる。よって、時間関係表現は、「 $E(P) \rightarrow R = S$ 」に修正される。

3.5 その他の時間関連副詞の扱い

時間副詞には属さないが、時間副詞と類似した機能を持つ副詞として「習慣」と「完了」を表す副詞がある。これらの副詞から得られる時間関係情報を拘束条件に追加する。

(1) 「習慣」を表す副詞

「いつも」、「よく」、「つねに」、「しょっちゅう」、「このごろ」などの副詞は、表 6 の「時間区間」が「<間隔>」を表す時間副詞と似た性質を持つ。

(例文 19) 彼は、いつも廊下を走っている。

例えば、例文 19 の副詞「いつも」は「テイル型」と共に用いられている。従って、話者は、事象を進行中のものとしてでなく、習慣や反復性を持つ事象として捉えていることが分かる。

そこで、このような副詞のある場合は、拘束条件として、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。すなわち、拘束条件は、「 $E = R$ 」である。但し、「 $E(P)$ 」の場合は、 E に修正する。

(2) 「完了」を表す副詞

完了の意味を表す副詞に「もう」、「かって」、「すでに」、「ちょうど」、「とっくに」などがある。これらの副詞が「タ型」、「テイル型」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ている。

(例文 20) 雨はとっくにやんでいる。

例えば、下記の例文 20 では、過去の事象の完了を発話時から捉えているから、拘束条件は「 $E \rightarrow R = S$ 」となる。

また「テイタ型」の文で用いられたときは、過去の事象の完了をある過去時から捉えているから、拘束条件は「 $E \rightarrow R \rightarrow S$ 」となる。

このように、この種の副詞的表現が使用されている場合は、動詞の種類によらず、発話時、事象時、参照時の関係が決定される。

4. 英語時間表現の生成

4.1 英語の時間表現形式

英語では、意味的範疇としてのテンス、アスペクト

と文法的な表現形式が比較的明確で、文法的形式と事象時 (E)、参照時 (R)、発話時 (S) の関係は、表 7 のように対応づけられる²⁴⁾。空欄は、対応する時間表現がないことを表す。

前章の解析で得られた日本語表現の時間関係表現から、この表に従って英語の時間表現を生成する。なお、日本語の解析結果の事象時が $E(P)$ となつて場合は、表 7 の E を $E(P)$ に読み替えて得られた英語文型を進行形に変形するものとする。

表 7 英語時間表現形式

Table 7 English Time Expressions

	$R = S$	$R \rightarrow S$	$S \rightarrow R$
$E = R$	単純現在	単純過去	単純未来
$E \rightarrow R$	現在完了	過去完了	未来完了
$R \rightarrow E$	単純未来	過去以後	—

<参考>日本語解析結果が $E(P)$ の場合は、該当する英語文型を進行形に変更する。

(例文 21) 彼らは通りを歩いていた。

例えば、例文 21 は、動作動詞の「テイタ型」であるから、3章の規則から時間関係は「 $E(P) = R \rightarrow S$ 」となる。そこで、表 7 で「 $E = R$ 」、「 $R \rightarrow S$ 」の欄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、過去進行形に翻訳される。

4.2 制約条件による英語表現の修正

前節までで、日本語表現から意味的な範疇としてのテンス、アスペクトに関する情報を取り出して、時間関係表現を求め、その結果に基づいて英語側の時間表現を決定する方法を示した。しかし、使用される動詞の種類や時間関係表現から正しい訳文形式が決定できるとは限らない。

第 3 章では、動詞をその意味に着目して分類したが、英語訳文でも同種の動詞が使用されるとは限らないし、名詞や形容詞と組み合わせで意識される場合もある。また、同種の動詞で訳される場合も、使用された動詞の用法によって進行形をとる場合ととらない場合があるなど、文法的範疇としてのテンスとアスペクトを決めるには、個々の動詞の用法にも着目する必要がある。

(例文 22) 娘は高等学校にかよっている。

My daughter goes to the high school.

例えば、例文 22 は、繰り返しの行為であり、表 7 では進行形に相当する。すなわち、3章の結果に従えば、日本語動詞「かよう」は、動作動詞に属し、3章と 4.1 節の規則から英語時間表現は進行形となる。しかし、実際の英語では、単純現在形が使用される。これ

は、一時的な動作も長期的な動作も表す日本語の「テイル型」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ったものと考えられる。

本節では、このような誤りを避けるため、英語側で選択された訳文の文型に合わせて時間表現を修正する方法を示す。

4.2.1 結合価パターン辞書による修正

英語の時間表現の生成では、「日本語語彙大系」³⁰⁾で定義された結合価パターンの用言意味属性を使用し、4.1節で得られた英語表現形式のうち進行形をとるもの(日本語側が「テイル型」)の適切性を確認する。不適切な場合は修正する。

この辞書では、英語文型パターンに対して、テンス、アスペクトに関する制約条件が記載されている。

図3に日本語動詞「疲れる」に対する結合価パターンの一部を示す。図の下線部分がそれぞれのパターンの意味属性、すなわち動詞の意味属性を表しており、()内は受け身、進行形などに関する制約条件である。

(パターン1)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
(パターン3)身体変化(状態, 受け身不可)
N1(#4 人)がN2(*)に疲れる
N1 get tired of N2
(パターン7)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
[注](#nnnXXX): 該当する名詞の意味属性番号と意味属性名を示す。

図3 動詞「疲れる」に対する結合価パターンの例

Fig. 3 Example of Valency patterns for Japanese verb "tsukareru"

例えば、「動作」は進行形可能、「状態」は進行形不可能を意味する。また、特定の助動詞などを伴って英語の文型パターンが変化する場合は、助動詞などが小書きにして後ろに付記されている。そこで、英語の時間表現生成では、これらの情報を制約条件として使用する。

例えば、例文22で示した問題を考えると、この文に適用される結合価パターンは図4のパターンである。

この場合、日本語側から得られた時間関係表現では、事象時が $E(P)$ であるが、結合価パターンの変形情報が「状態」となっているため、進行形はとれない。そこで、 $E(P)$ を E に修正する。

4.2.2 完了形に関する制約条件

日本語では「変化動詞+テイル(タ)型」で変化の結果状態を表すのに対して、英語では、状態として捉え、完了形ではなく、形容詞を使用した「be動詞+adj」

<見出し語>「通う」 <パターン2> 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 <英語パターン> N1 go from N2 to N3

図4 例文22に使用される結合価パターン

Fig. 4 Valency patterns for example # 22

型や状態動詞で表すことがある。以下に例を示す。

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「濡れている」という現在の事象を「濡れる」という事象の結果として見ている。すなわち、原因と結果の関係を内的な視点の移動を伴った表現で表している。これに対して、英語では、be動詞を使うことによって視点を結果の時点に固定し、濡れた状態を表現している。

このように、事象時と参照時の時間関係が完了形を表す「 $E \rightarrow R$ 」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「 $E = R$ 」を設ける。

5. 複文(引用節)における時間関係

5.1 日本語複文における時間関係

引用節を持つ日本語複文の時間表現について考える。ここで、引用節とは、「彼が来ると思う」などの文で、引用の助詞「と」に先行する節を示す。引用節を取る動詞には、「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。

主節の時間関係については、すでに述べた単文の方法がそのまま適用できる。そこで、引用節の時間関係を検討することとし、以下の例文を考える。

(例文23) 私は、彼がご飯を食べに来ると思った。

(例文24) 私は、彼がご飯を食べに来たと思った。

(例文25) 私は、彼がご飯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。また、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。

以上から、引用節が表す時間は、主節の時間を基準に考えれば良いことが分かる。そこで、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。

<ステップ1>

3章で示した単文の規則を用いて、主節、引用節の時間関係を個別に決定する。

<ステップ2>

引用節の発話時 (S) を主節の事象時 (E) に一致させる。

<ステップ3>

主節の発話時 (S) を複文全体の発話時 (S) とする。

但し、ステップ1で表3を適用する際、外的動作動詞、変化動詞の場合、発話時と参照時の関係が $S = R$ または $S \rightarrow R$ となっているが、習慣をあらわす副詞が用いられている場合を除いて、引用節については、 $S \rightarrow R$ とする。これは、引用節の動詞の「ル型」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、伝聞したりしているからである。

複文における主節と従属節に対する時間関係表現の例を例文26と例文27に示す。時間関係表現は、それぞれ図5、図6の通りである。

(例文26) 私は、彼が来ると思った。

主節の時間関係 $R = E \rightarrow S$

引用節の時間関係 $S \rightarrow R = E$

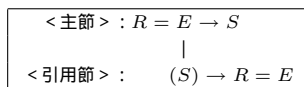


図5 引用節を含む文の時間関係表現 (例文26)

Fig. 5 Example of time relation for Complex Sentences

(例文27) 私は、彼が来たと思った。

主節の時間関係 $R = E \rightarrow S$

引用節の時間関係 $R = E \rightarrow S$

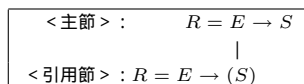


図6 引用節を含む文の時間関係表現 (例文27)

Fig. 6 Example of time relation for Complex Sentences

5.2 複文の英語への翻訳

複文では、主節、従属節のそれぞれについて表7の規則を用いて時間関係表現を求め、英語表現を生成する。しかし、英語では「時制の一致」により、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。

「時制の一致」とは、従属節の事象を、主節の事象と同じ時点で捉えることであるから、英語では、主節と従属節の発話時は同時であり、主節の発話時のこと

である。そこで、時制の一致が要求される場合、従属節の発話時 (S) を参照時 (R) に変更する。すなわち、引用節の参照時 (R) を主節の事象時 (E) に一致させることを制約条件とする。

例文26から抽出した時間関係表現からの翻訳を考える。この場合、主節が過去形であるため、時制の一致が生じる。従属節の発話時 (S) を参照時 (R) に変更する、図5の時間関係は図7のようになる。

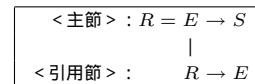


図7 引用節を含む文の時間関係表現 (例文27)

Fig. 7 Example of time relation for Complex Sentences

従って、表7の規則を用いると、主節は単純過去形、引用節は<過去以後> (would) となる。図からは、引用節の「来る」という事象は「発話時以後か」、「以前か」は判断できないが、時制一致の条件に即した翻訳結果が得られる。

また、例文27の翻訳でも時制の一致が生じる。つまり、図6は、図8のように変化する。

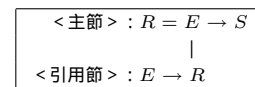


図8 引用節を含む文の時間関係表現 (例文27)

Fig. 8 Example of time relation for Complex Sentences

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。

(例文28) 私は、彼が来たと思うだろう。

主節の時間関係 $S \rightarrow R = E$

引用節の時間関係 $E = R \rightarrow S$

次に、例文28では、主節が未来の事象を表し、引用節の述部が「タ型」である。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図9のようになる。

図9では、2つの解釈が残されている。これは、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

この操作は、日本語と英語における対象認識の違いを反映している。すなわち、日本語では、観念的話者が時間を移動することによって、複文においても述部毎に独立した発話時を持つ表現が使用されるのに対して、英語では、発話時は主節の発話時の位置に固定され、従属節の時は主節の時との相対的關係として表現される。すなわち、複文全体の発話時は主節の発話時であり、それは引用節の発話時と共通であるので、従属節の事象時は、主節の参照時と同時点となる。

節	解釈1	解釈2
<主節>	$S \rightarrow R = E$	$S \rightarrow R = E$
<引用節>	$E = R \rightarrow (S)$	$E = R \rightarrow (S)$

図9 引用節を含む文の時間関係表現 (例文 28)

Fig.9 Example of time relation for Complex Sentences

るため、引用節の事象と発話時の関係が過去か未来かを判断できないためである。

6. 評価実験

評価実験の方法と結果について述べる。以下で示す実験は、実験用の標本選択の方法から言ってオープンテストであるが、各標本に対する動詞の分類(表1)と時間副詞の分類(表4)の情報は、人手で与えたものである。

6.1 日本語単文の時間表現の翻訳

対訳標本データとして、「アンカー和英辞典」³¹⁾から「ル型」「タ型」それぞれ200文「ル+ダロウ型」,「テイル(タ)型」それぞれ100文の合計600文をランダムに取り出し、日本語に対する英訳文の時間表現と比較して正解率を求めた。使用した文はいずれも単文である。

また評価では、単語の訳語の精度は問題とせず、文法的に見てテンス、アスペクトが正しいものを正解とした。なお「ル型」については、複数の候補が得られる場合があるため、以下の3つの基準で評価した。

- ：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

表8に結果を示す。括弧内は該当する文数を示す。

表8から、翻訳正解率は「タ型」の場合92.5%「ル+ダロウ型」では、86.0%である。また、候補が一つに決定できない場合がある「ル型」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も加えると93.0%である。また「テイル(タ)型」については、一意正解率が65.0%にとどまったが、このタイプは、人間でも判定困難な場合が多い。

人手によらず機械的に処理する方法としては以下の方法が考えられる。すなわち、動詞の意味分類については、「日本語語彙大系」³⁰⁾によって、日本語6000動詞に対する意味属性(36種類)が与えられており、あらかじめ、表1とこれらの意味属性との対応関係を決めておく方法が考えられる。また、時間動詞の場合は、日本語の場合はその数も比較的少ないと見られるので、網羅的に抽出しておくことが可能と考えられる。なお、時を表す「名詞」もしくは「名詞+助詞」の形式の副詞的表現については、形態素解析で示される名詞の細分類コードから「時の名詞」を抽出し、副詞的用法の分類表を作成する方法が考えられる。

表8 評価結果(1)

Table 8 Translation results

文のタイプ			×
タ型(200)	92.5%(185)	—	7.5%(15)
ル型(200)	51.5%(103)	41.5%(83)	7.0%(14)
テイル(タ)型(100)	65.0%(65)	—	35.0%(35)
ル+ダロウ型(100)	86.0%(86)	—	14.0%(14)
平均(600)	73.2%(439)	13.8%(83)	13.0%(78)

6.2 引用節を含む複文の時間表現の翻訳

対訳標本データとして「アンカー和英辞典」³¹⁾および「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」²⁵⁾から、引用節を含む日本語文200文を取り出し、前項と同様にして正解率を求めた。結果を以下の表9に示す。

表9 評価結果(2)

Table 9 Translation results(2)

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

表9から、ほぼ80%の正解率が得られることが分かる。なお、一意決定不可能の一文は、例文28と同様の文で、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。

6.3 市販ソフトとの比較

対訳標本データとして「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」²⁵⁾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を行った。

なお、本手法では「ル型」の場合、現在形と未来形の2つの候補が得られる場合が6件存在したが、比較のため、その場合はデフォルトとして単純現在形を選択することにした。その結果、6文中、3文が正解となった。

表10に結果を示す。括弧内は該当する文数を示す。この表より、提案した手法では、市販ソフトに比べて16.6%精度が向上していることが分かる。

7. 考察

6.1 節から6.3 節で示した実験結果について、順に

「アンカー和英辞典」が中心である。現象の網羅性を補強する意味で、日本語教育用のテキスト²⁵⁾の第3章から標本を追加した。なお、この文献の英語模範訳は翻訳家に付与してもらった。例文数は多くないが、日本語教材としての日本語のテンス、アスペクトの典型的な表現が網羅されていると考え、文献²⁵⁾の第1章と第2章に掲載された例文を使用した。なお、英語の模範訳は翻訳家に付与してもらった。

表 10 市販ソフトとの比較
Table 10 Comparison to current system

文のタイプ	提案方式		市販ソフト	
		x		x
タ型 (40)	80% (32)	20% (8)	72.5% (29)	27.5% (11)
ル型 (51)	80.4% (41)	19.6% (10)	52.9% (27)	47.1% (24)
テイル(タ)型 (58)	48.3% (28)	51.7% (30)	34.5% (20)	65.5% (38)
ル+ダロウ型 (2)	100% (2)	0% (0)	100% (2)	0% (0)
平均 (151)	68.2% (103)	31.8% (48)	51.6% (78)	48.3% (73)

考察する。

7.1 単文の場合について

7.1.1 「タ型」について

「タ型」に対する本手法の正解率は 92.5% で、正しく翻訳できなかった文が 15 文存在する。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。その例を例文 29 に示す。

(例文 29) 朝晩めっきり涼しくなった。

It is quite cool these days mornings and evenings.

模範訳： 単純現在形

本方式の訳：単純過去形

この例の模範訳は、「be + 形容詞」の形式となっている。すなわち、日本語では、「涼しくなる」という動的事象にも着目し、その結果としての「涼しくなった」状態を表現しているが、英語は、単に「涼しい」という現在の状況のみが表現されている。日本語側の解析結果が、完了の意味となる場合は、4.2 節で示したようなチェック条件が働くが、この場合は、完了の意味ではないので、チェック条件は働かなかった。

また、単に過去と訳すべきところを時間副詞の情報から完了形に変更して訳した文が 3 文存在した。その例を示す。

(例文 30) かつて国鉄には多くの赤字路線があった。

JNR had many deficit-ridden lines.

模範訳： 単純過去形

選択した訳：現在完了形

この文では、副詞「かつて」から完了形と判断したが、模範訳では、副詞「かつて」は訳出されていない。

7.1.2 「ル型」について

「ル型」については、200 文中、現在の事象かそれとも未来の事象かを判断できない場合(表 8 の に相当)が 86 文存在したが、すべて単純現在形と翻訳すると、そのうち、76 文が正解となり、それを含めた「ル型」200 文の正解率は、89.5% となる。

不正解となった文は、14 文で、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.

模範訳： 単純現在形

本方式の訳：単純未来形

この例では、発話時以後の事象を示す時間副詞「あすから」によって英語時間表現は未来形が選択され、模範訳では、現在形とされている。定期試験は確定的なもので、変更はないだろうと判断される。この判断は、日本語でも同様であるが、未来形としたのは時間副詞が存在することによる副作用である。

このように、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。

7.1.3 「テイル(タ)型」について

「テイル(タ)型」については、一意正解率が 65.0% と、最も低い結果である。日本語の「テイル型」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(タ)型」の実験データ 100 文中で最も頻度の高かった英語時間表現は、単純現在形(「テイト型」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。

7.1.4 「ル+ダロウ型」について

「ル+ダロウ型」の型の翻訳精度は「タ型」の次に良く、86% である。評価に使用した 100 標本文中、翻訳に失敗した例文は 14 文である。いずれも模範訳は単純現在形となっているところを単純未来形に間違えたものである。これらは、以下の 2 種類に分類できる。(例文 32) 彼女は間違いなく舞台上に再起できるだろう。

She is sure to make a comeback on stage.

(例文 33) 先方は我々の案に賛成しないだろう。

I don't think they will agree to our proposal.

一つは、例文 32、例文 33 のように推量の意味を持つ動詞 (be sure to, think など) を使用して訳されるものである。suppose や be afraid of, be certain 等が用いられる場合も含め 12 件がこのタイプに該当する。

(例文 34) このようなチャンスは二度と来ないだろう。

You don't get this sort of chance twice.

残りの 2 件は、例文 34 のように、模範訳では、あえて未来形に訳さなくても良いと考えられたと思われるもので、単純未来形で訳しても誤りとは言えないと考えられる。

7.1.5 結合価パターンの効果について

すでに述べたように「テイル型」の翻訳では、進行形として良いかどうかの判断に結合価パターンを使用している。しかし、この種の例文 100 件中、結合価パターンによって正しい英語時間表現に翻訳できた文が 23 文であった。以下に例を示す。

(例文 35) あいにく店は閉まっていた。

To my disappointment, the shop was closed.

模範訳： 単純過去形

本方式の訳：単純過去形

この日本語に適用される英語文型パターンは、「N1 be closed」となるため、制約条件から単純過去形と翻訳した。

しかし、例文 36 のように、結合価パターンを用いたために、かえって、誤った英語時間表現に翻訳した文も 3 文存在した。

(例文 36) 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.

模範訳： 現在進行形

本方式の訳：単純現在形

例文 36 では、模範訳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、単純現在形を選択した。

これは、日本語側の主語相当語が「大会関係者」であるのに対して、英語の主語は、「連日の雨」となっていることが原因で、日本語の内容を捉え直して表現した英文となっているためである。実際は、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。

7.2 複文の場合について

複文(引用節を含む文)の精度評価の結果では正解率は 79.5%で、不正解となった文が 40 文存在したが、9 文が主節、引用節の動詞ともに「タ型」であるにも拘わらず、英語の引用節の時間表現が単純過去形となった。以下に例を示す。

(例文 37) 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.

模範訳： 主節： 単純過去形

引用節：単純過去形

本方式の訳：主節： 単純過去形

引用節：過去完了形

これは、模範訳が「be out of」を使って訳されているのに対して、動詞の訳語決定に使用した辞書にこの翻訳形式がなかったためである。

7.3 市販ソフトとの比較について

本手法と市販ソフトで正解および不正解となった文を分類した結果を表 11 に示す。

表 11 から、本手法では誤った文の中で、市販ソフトは正解であった文は 13 文であった。これらの文の分析によれば、10 文が「テイル型」を誤った英語時間表現に翻訳したもので、残る 3 文は、動詞の時間的性質からは正しい時間表現が得られたにも拘わらず時間副詞の情報によって誤った時間表現に変更している。

表 11 評価結果の分析

Table 11 Analysis of translation results

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

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 38) 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.

模範訳： 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 38 では、時間副詞「十年後」があるため単純未来形と翻訳した。しかし、模範訳は単純現在形となっている。これは、様相の副詞「きっと」に着目して、確信を意味する be certain を使用して訳されたもので、一種の意識である。

本研究では、模範訳を正解としたため、模範訳が意識となっている場合は、不正解となることが多かった。今回、様相については考慮しなかったが、意識は、様相とも関連が深い。今後の検討課題である。

8. おわりに

本稿では、「~ル型」、「~タ型」、「~テイル型」、「テイタ型」、「ル+ダロウ型」の日本語単文、および、引用節を含む複文の時間表現の意味を「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で表現し、それを介して英語表現を生成する方法を提案した。

具体的には、まず、日本語表現の文末の語尾形式、述語動詞の時間的性質、文中で使用された時間関係副詞の情報に着目して「発話時」、「事象時」、「参照時」の

3つの時点の時間関係を言語独立な記号で表現した。また、英語表現の生成では、英語側の特殊事情を考慮し、結合価パターン辞書を使用して、英語側で使われる動詞とその用法から生じる進行形の可否をチェックすると同時に、完了形の意味を、be動詞を使用した表現に訳出する場合についての制約条件を設けた。

提案した手法を和英辞典から抽出した単文600文に適用した結果では、正解率は、文末語尾形式に大きく依存するが、全体では、一意正解率73.2%、複数選択正解を含む正解率87.0%となった。外国人のための日本語テキストから抽出した複文(引用節を持つもの)200文に適用した結果では、正解率は79.5%であった。また、日本語教材のテキストに収録されたテンス、アスペクトの用例151文を用いた実験では、本方式は、市販ソフトに比べて、16%程度高い翻訳精度となった。誤りの多くは、模範訳が意識となっている場合で、そのような場合は、様相情報なども考慮した翻訳方法が必要と考えられる。

また、英語表現でも、現在形、過去形が必ずしも現在、過去の事象を表現するとは限らない。どのようなテンス、アスペクトで表現するかは、話者の対象に対するとらえ方にも依存する。今後は、これらの点を考慮し、機械翻訳として許容される範囲について検討することも必要と考えられる。

参 考 文 献

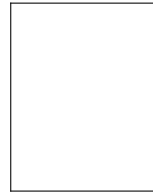
- 1) James Allen: Natural Language Understanding, The Benjamin Cummings Publishing Company, Inc., 1995
- 2) Bob Carpenter: Type-Logical Semantics, The MIT Press, Cambridge Massachusetts, 1997
- 3) 池上嘉彦: 「「する」と「なる」の言語学」大修館, 1981
- 4) 国立国語研究所: 「現代日本語動詞のアスペクトとテンス」, 秀英出版, 1985
- 5) 寺村秀夫: 「日本語のシンタクスと意味 1,2」くろしお出版, 1984
- 6) 金田一春彦編: 「日本語動詞のアスペクト」むぎ書房, 1976
- 7) 工藤真由美: 「アスペクト・テンス体系とテキスト - 現代日本語の時間の表現 - 」ひつじ書房, 1995
- 8) 中村ちどり: 「日本語の時間表現」くろしお出版, 2001
- 9) 三原健一: 「時制解釈と統語現象」くろしお出版, 1992
- 10) 草薙裕: 「テンスとアスペクトの文法と意味」, 朝倉日本語新講座「文法と意味 1」朝倉書店, pp. 166-208, 1983
- 11) 古瀬蔵, 中園薫, 野村浩郷: 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
- 12) 江良浩一, 甲斐郷子, 中村順一, 吉田将: 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
- 13) 溝淵昭二, 住友徹, 泓田正雄, 青江順一: 日本語時間表現の解釈法, 情報処理学会論文誌, Vol. 40, No. 9, pp. 3408-3419, 1999
- 14) 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, Vol. 8, No. 1, pp. 55-64, 1993
- 15) 東条敏: アロー論理によるアスペクトの解析, 自然言語処理, Vol. 7, No. 4, pp. 3-24, 2000
- 16) K. Yoshimoto; Tense and Aspect in Japanese and English, Peter Lang, 2000
- 17) 河合敦夫: 日英翻訳システム ALT-J/E における様相, 時制の処理, 情報処理学会第34回全国大会, 1987
- 18) 白井諭, 横尾昭男, Francis Bond: 新聞記事翻訳における時制の訳出について, 電子情報通信学会秋期全国大会, 1990
- 19) 村田真樹, 馬青, 内元清貴, 井佐原均: 用例ベースによるテンス・アスペクト・モダリティの日英翻訳, 人工知能学会論文誌, Vol. 16, No. 1, pp. 20-28, 2001
- 20) 村田真樹, 馬青, 内元清貴, 井佐原均: サポートベクトルマシンを用いたテンス・アスペクト・モダリティの日英翻訳, 信学技報, 言語理解とコミュニケーション研究会, NLC2000-78, pp. 25-32, 2001
- 21) 謝軍, 池田尚志: テンス・アスペクトの中期会翻訳処理, 言語処理学会第8回年次大会, pp. 77-80, 2002
- 22) Alice G. B. Ter Meulen; Representing Time in Natural Language, MIT Press, Cambridge, Massachusetts, 1995
- 23) Peter Ludlow; Semantics, Tense, and Time, The MIT Press 1999
- 24) 高橋潔, 根岸雅史: 「基礎からの新総合英語」数研出版, 1999
- 25) 加藤泰彦, 福地努: 「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
- 26) 三浦つとむ: 「言語と認識の理論 Vol. 1 ~ Vol. 3」けい草書房, 1967
- 27) Reichenbach, H: Elements of Symbolic Logic, Collier-Macmillan, London, England, 1947
- 28) Bruce B.C.; A Modal for Temporal References and Application in a Question Answering Program, Artificial Intelligence, 3, pp. 1-25, 1972
- 29) Z. Vendler: Linguistics in Philosophy, Ithaca, New York: Cornell University Press, 1967

30) 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩
巳, 小倉健太郎, 大山芳史, 林良彦:「日本語語
彙大系」岩波書店, 1997

31) 山岸勝榮:「アンカー和英辞典」学研, 1991

(平成 14 年 1 月 1 日受付)

(平成 14 年 1 月 1 日採録)



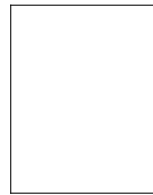
村上 仁一

1984 年筑波大学第 3 学群基礎工
学類卒. 1986 年筑波大学修士課程
理工学研究科理工学専攻修了. 1986
年 NTT に入社. NTT 情報通信処
理研究所に勤務. 1991 年国際通信
基礎研究所 (ATR) 自動翻訳電話研究所に出向. 1997
年鳥取大学工学部知能情報工学科に転職. 現在に至る.
主に音声認識のための言語処理の研究に従事. 電子通信
情報処理学会, 日本音響学会, 言語処理学会, 各会員.



池原 悟 (正会員)

1967 年大阪大学基礎工学部電気
工学科卒業. 1969 年同大学院修士
課程修了. 同年日本電信電話公社に
入社. 数式処理, トラフィック理論,
自然言語処理の研究に従事. 1996 年
スタンフォード大学客員教授. 現在, 鳥取大学工学部
教授. 工学博士. 1982 年情報処理学会論文賞, 1993
年同研究賞, 1995 年日本科学技術情報センター賞 (学
術賞), 同年人工知能学会論文賞, 2002 年電気通信普
及財団賞 (テレコム・システム技術賞) 受賞. 電子情
報通信学会, 人工知能学会, 言語処理学会, 機械翻訳
協会各会員.



的場 和幸

2000 年 3 月 鳥取大学工学部知能
情報学科卒 2000 年 4 月 鳥取大学大
学院工学研究科知能情報工学専攻入
学 2002 年 3 月 鳥取大学工学部工学
研究科知能情報工学専攻修了 2002
年 4 月 ディーアイエステクノサービス株式会社に入社