

日本語重文・複文を対象とした文法レベル文型パターンの被覆率特性

Coverage of Syntactic Sentence Patterns for Japanese Complex and Compound Sentences

池原 悟* 徳久雅人* 村本奈央** 村上仁一*
SATORU MASATO Tokuhia NAO Muramoto JIN'ICHI Murakami

* 鳥取大学工学部 Faculty of Engineer, Tottori University

〒680-0945鳥取市湖山町南4.101 E-mail: { ikehara, tokuhisa, murakami } @ike.tottori-u.ac.jp

** NTTアドバンステクノロジー株式会社 NTT Advanced Engineer Corporation

〒210-0007 川崎市川崎区駅前本町12-1 川崎駅前タワーリパークビル nao@nlp.ntt-at.co.jp

Abstract: Pattern base MT received attention for long time since it yields good translations for matched sentences. But it was been difficult to build a large-scale pattern dictionary which have semantically independent patterns to obtain a high coverage. This paper experimentally evaluated the coverage of the pattern pair dictionary which has recently been developed for Japanese Complex and Compound sentences and studied possibility of the pattern based MT method. This dictionary contains syntactic sentence patterns of word level (100,000 patterns), phrase level (100,000 patterns) and clause level (10,000 patterns) which are generated from 150,000 example sentence pair for Japanese to English. Evaluation was conducted by using 5 parameters such as "Sentence Recall Ratio", "Sentence Coincide Ratio", "Semantic Precision Ratio", "Recall Pattern Precision Ratio" and "Matched Pattern Precision Ratio". The results are as follows. "Sentence Recall Ratios" are 70%, 90%, ??% for each of Word level, Phrase Level and Clause level sentence patterns and "Semantic Precision Ratios" are only 20% ~ 30%.

1. まえがき

従来の機械翻訳システムは要素合成法を基本としているが、言語表現には意味的に非線形なものが多く、表現を分解して行く過程で全体の意味が失われることが問題であった。この問題を解決するには、文構造とその意味を一体的に扱う仕組みが必要と考えられる。

既に、このような仕組みとしては、古くから文型パターン翻訳(テンプレート翻訳とも言う)の方法が試みられてきた。この方法は、文型パターンに適合した入力文に対して品質の良い訳文が得られることから、トランスファー方式と併用する形で多くの商用システムで実現されている。最近では翻訳メモリとも併用する形式で採用されている例が多い。しかし、使用されている文型パターン数は、いずれも100~300パターン程度で少なく、補助的な仕組みとして使用されるか、もしくは、ある特定の狭い分野の文書に適用されているのが普通である。これは文型パターン作成コストが高いこともあるが、パターン数が増大するとパターン間の意味的な相互干渉が増加して翻訳精度が低下することが主な原因と見られる。

これに対して、文構造と意味の関係を体系化を試みたものとして多段翻訳方式(池原, 宮崎, 白井, 林 1987)の研究がある。この方式では、動詞, 形容詞と及びそれと共起する名詞の関係を1.7万件の結合価パターン(池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林 1997)としてまとめており、単文レベルでの訳文品質は従来に比べて大幅に向上した(金出地, 池原, 村上 2001)。しかし複文(埋め込み節を持つ文)や重文(接続構造を持つ文)では、表現全体に対する意味の非線形性を扱う仕組みがないこと(Ikehara 2001)、また意味の単位と見なされる原言語の表現構造に対して、単一の目的言語表現が対応づけられるため、文脈上不適切な表現への翻訳が防げないことが問題として残されている(池原2001)。

これらの2つの問題を解決することを狙って、最近、「等価的類推思考の原理による機械翻訳方式」(池原2002, 2003)が提案された。この方式は、原言語と目的言語の

表現から非線形な表現構造(池原2004)を取り出してパターン化した後、意味的同等性に着目して類型化(有田1987)すること、また意味的に類型化されたパターン間を「類推思考の原理」(市川1963)によって対応づけることの2つの仕組みから構成される。

この方式を実現するには、原言語、目的言語の表現の「意味類型パターン」を収録した大規模な「意味類型知識ベース」を構築することが必要である。「意味類型パターン」は原言語表現の意味的に非線形な構造を取り出して、それを意味的に分類体系化し、原言語と目的言語に共通する概念(「真理項」と呼ぶ)を介して目的言語の「意味類型パターン」に対応づけるものである。従って、「意味類型パターン辞書」を開発するには、あらかじめ日英対訳文から意味的に非線形な構造を取り出して「対訳文型パターン辞書」を作成することが必要となる。

ところで、文型パターン辞書構築の最大の問題は、文型パターンの網羅性と意味的排他性をいかにして実現するか2点である。

このうち網羅性を実現するには、汎化された大量の文型パターンが必要と考えられるが、汎化が進むにつれて意味的排他性を確保することが難しくなることが予想される。このような網羅性と意味的排他性の問題解決するには、実際にある程度大量の文型パターンを作成しながら、実験的に汎化の可能性を調べていくことが必要と考えられる。

これに対して、新翻訳方式に関する上記の研究では、第1段階として、日英対訳例文100万件の中から重文、複文の例文12.9万件を抽出し、それを文法的なレベルで汎化することによって単語レベル、句レベル、節レベルの文型パターン24.6万件(そのうち異なり文型パターン数は22.1万件)が試作された(池原, 阿部, 徳久, 村上2004)。そこで、本研究では、日本語試験文との照合実験を行って、文型パターンの被覆率と意味的排他性について評価し、今後、文法レベルで定義された文型パターンの問題点と今後の対策について検討する。

*1 慣用表現の例からも分かるように、言語表現の意味が、それを構成する要素の意味の和として表現できないことを言う。言語表現の意味の線形性、非線形性の定義とその判定の方法については、[池原04]を参照。

2. 文型パターン辞書の概要

本章では、文型パターン（以下では単に「文型」と呼ぶ）作成の元となった日英対訳文の内容と作成された文型辞書の概要を示す。

(1) 元となった日英対訳標本文

本稿では、2カ所又は3カ所の述部を持つ¹⁾日本語の重文（文接続のある文）、複文（埋め込み文のある文）を対象に文法レベルの情報を使用して記述された「文型辞書」の被覆率特性について調べる。この辞書は、日本語の基本的な表現が収録されていると見られる辞書や語学教育用の教科書、機械翻訳機能評価用の試験文等、約30種類のドキュメント（日英対訳文100万件）から該当する対訳標本文12.9万件を取り出し、それを汎化することによって作成されたもので、異なり文型22.1万件が収録されている。

日本語標本文当たりの平均文字数は、23.3字/文（最大148字/文）であり、平均形態素数は、12.9個/文（最大63個/文）である。また、それと対応する英文訳文の平均単語数は、10.3語/文（最大59語/文）である。

(2) 文型の記述レベルと汎化の方法

文型は、対訳標本文中の線形な要素を対象に、変数化、関数化、任意化などを行ったもので、変数化された対象の文法的属性に着目して、単語レベル、句レベル、節レベルの3グループから構成される。表1に各レベルにおける主な汎化の内容を示す²⁾。

表1. 文型の記述レベルと汎化の内容

レベル	主な汎化の内容
単語レベル	線形要素の自立語を変数化した文型で、おおよそ以下の汎化まで行ったもの。 (1)名詞、動詞、形容詞、副詞などの自立語を変数化 (2)線形で文型上不要な要素を任意化し、文型の骨組みとなる要素を抽出 (3)字面要素について可能な範囲でグループ化
句レベル	句の変数は使用されているが、節の変数の使用されていない文型で、単語レベルの文型におおよそ以下の汎化を行ったもの。 (1)適用範囲を品詞から句へ拡大 (2)機能語の適用拡大（格助詞 格助詞相当語、等） (3)英語句生成関数の適用
節レベル	節の変数が一つ以上使用された文型で、句レベルの文型パターンにおおよそ以下の汎化を行ったもの。 (1)名詞節、副詞節、主節、従属節の1変数化による重文、複文の基本構造をパターン化 (2)日本語節から英語句への変換関数の使用 (3)その他英語構造生成関数の使用

これらの3段階は汎化の程度を表すもので、いずれも使用される変数や関数は文法的属性によって変域が指定されるものであり文法レベルで定義された文型である³⁾。

(3) 評価対象とする文型数

前項の方法によって作成された文型の内訳を表2に示す。本稿ではこれを評価対象とする。単語レベル、句レベル、節レベルと汎化が進むにつれて文型は縮退するため文型数は減少している。特に節レベルの文型数は少ないが、これは節レベルまで汎化できるような標本文が少なかったためである。

3. 調査実験の方法

3.1 実験の方法

本実験では、「文型照合プログラム」を使用して、日本語重文、複文を文型辞書と照合し、適合する文型（以下では、単に「適合文型」と言う）を抽出する。その結果から文型辞書の被覆率特性を求める。実験の方法は、断らない限り、下記に示すような「クロスバリデーション法」（オープンテストの一種）を用いる。

<本研究におけるクロスバリデーション法>

文型の作成に使用した対訳標本の日本語を入力文として使用する。その場合、文型は、元となった対訳標本文対毎に作成されているから、各文型はその元となった標本文と必ず適合するはずである⁴⁾。そこで、実験では、テスト文から作成された文型（「自己文型」と呼ぶ）以外の文型にどれだけ適合するかを調べる。この方法は、学習データとテストデータを切り替えて使用する点でクロスバリデーション法（オープンテストの一種）である。

なお、この実験では、文型作成で使用した標本文と同じ母集団から試験文を使用していることから、以下では、この実験で使用する試験文を「母集団試験文」と呼ぶ。

3.2 評価基準

(1) 評価パラメータの種類

被覆率の評価では、情報検索と同様の「再現率」、「適合率」に相当する基準を使用するが、文型検索で注意する点は以下の通りである。

入力日本語の翻訳に使用できる文型は、必ずしも入力文のすべての要素が適合する文型である必要はなく、入力文の主要な構造が適合し意味的に正しい文型であればよい。すなわち、文型照合プログラムは、入力文に対して、当該文型のすべての要素が定義された順に出現する文型を適合文型として抽出するから、適合文型は、以下の2種類に分類できる⁵⁾。

表2 対訳標本文数と作成された異なり文型数

文種別	述部の数	説明	日英対訳標本文数	作成された異なり文型数			
				単語レベル	句レベル	節レベル	合計
文種別1	2	文接続1カ所を持つ文	57,235 (44%)	53,578 (44.0%)	37,356 (42.3%)	5,521 (48.1%)	96,455 (43.7%)
文種別2	3	文接続2カ所を持つ文	6,196 (5%)	6080 (5.0%)	4,952 (5.6%)	417 (3.6%)	11,450 (5.2%)
文種別3	2	埋込み文1つを持つ文	46,907 (36%)	44,008 (36.2%)	30,932 (35.0%)	3,185 (27.7%)	78,125 (35.3%)
文種別4	3	埋込み文2つを持つ文	5,986 (5%)	5,889 (4.8%)	5,084 (5.8%)	811 (7.1%)	11,784 (5.3%)
文種別5	3	文接続、埋込み文各1を持つ文	12,389 (10%)	12,174 (10.0%)	10,025 (11.3%)	1,551 (13.5%)	23,750 (10.7%)
- -	- -	合計	128,713 (100%)	121,729 (100%)	88,349 (100%)	11,485 (100%)	221,563 (100%)

¹⁾ 多数の述部を持つ重文、複文の場合、全体が非線形である場合は少なく、多くの場合は、少数の述部を持つ非線形構造の組み合わせによって構成されていると考えられるため、述部2又は3までの非線形構造を対象に文型パターン化されている。

²⁾ 各レベルの詳細は、文献（池原ほか2004）を参照。また、文型パターン記述方法については、文献（池原ほか2004）を参照。

³⁾ 文型要素が形態解析で得られた文法情報の範囲で記述されている点で「文法レベルの文型」と呼ぶ、これに対して、変数の意味的な制約条件などを付与した文型を「意味レベルの文型」と呼ぶ。これらはいずれもまだ意味的な類型化は行われていない。

⁴⁾ 適合しないのは文型パターン記述に誤りがある場合であることに着目して文型パターンはデバッグされている。

⁵⁾ このうち、(1)は、対応する英語文型パターンをそのまま使用することで、英訳文が生成できるのに対して、(2)の場合は、文型パターンに含まれない入力文要素に対して、その翻訳結果と文型パターンから得られた翻訳結果を合成することが必要である。

<完全一致文型>：入力文のすべての要素が文型の要素と適合する文型

<部分一致文型>：入力文の一部の要素が文型に定義されない要素となる文型

従って、適合文型だとは言え、必ずしも入力文のすべての要素が解釈を与えるものではないから、適合文型について、それが、入力文の何%をカバーしているかが問題となる。そこで本稿では「入力文単位で見た被覆率」と「入力文の文字単位で見た被覆率」の2つの評価パラメータを使用する。また、検索された文型の意味は、入力文の意味に一致するとは限らないため、意味的な正解率も評価パラメータとする。以下、本稿で使用する評価パラメータを示す。

(2) 被覆率評価のためのパラメータ

本稿で被覆率調査のために使用する評価パラメータは、再現率に相当する2種類と意味適合率に相当する3つの合計5種類である。それぞれを表7に示す。以下、「被覆率特性」と言えば、これら5種類の評価パラメータの値を意味する。

表3. 被覆率評価のパラメータ

評価パラメータ	説明
再現率	R1 文型再現率 テスト入力文の数に対して1つ以上の適合文型があった入力文の数の割合
	R2 文型一致率 テスト入力文全体の文字数に対する「最大文型パターン<注>」でカバーされる入力文全体の文字数の割合
正解率	P1 意味適合率 入力文が「最尤文型<注2>」でカバーされる入力文の文字数の割合
	P2 適合文型意味正解率 適合文型が意味的に正しい確率
	P3 適合文型正解含有率 各入力文に対する適合文型中に意味的に正しいものが1つ以上含まれる確率

<注1>「最大文型パターン」：適合文型のうち、最も広範囲で入力文に適合する文型。<注2>「最尤文型」：意味的に適合した文型の中の最大の文型のこと

4. 日本語文型の被覆率特性

本章では、「文型再現率R1」と「文型一致率R2」を評価する。実験は、クロスヴァリデーション法によることとし、文型作成に使用した日英対訳文12.9万件から、文種別毎の文数の割合を保つこと、それぞれの文種別内での文当たりの文字数の分布が保たれることを条件に合計1万件の日本語を抽出し、テスト用の入力文として使用した。

4.1 被覆率特性と飽和特性

(1) 文型の被覆率

前述の1万件の日本語入力文に対して単語レベル、句レベル、節レベルの文型の「文型再現率」、「文型一致率」を求めた。その結果を表4に示す。

表4. 文型の被覆率

レベル	文型再現率 R1	文型一致率 R2
単語レベル	69.8 %	47.2 %
句レベル	89.0 %	76.2 %
節レベル	78.1 %	63.1 %
混合レベル	91.8 %	78.9 %

表3で「混合レベル」は、単語レベル、句レベル、節レベルの文型全体（異なり22.1万件）に対する照合実験の結果を示す。この表から以下のことが分かる。

- (1) 単語レベル、句レベルの「文型再現率」は、約7割、9割で高い値を示す。

(2) これに比べて最も汎化された節レベルの「文型再現率」は78%で比較的小さい。

(3) また「文型一致率」は、それより13~23%低下する。

(4) 「混合レベル」の「文型再現率」、「文型一致率」は、句レベルの場合より2~3%向上しただけである。

以上、文数から見て、入力文の7割が単語レベルの文型に適合し、9割が句レベルの文型に適合するのに対して、節レベルの文型全体の被覆率は低い。節レベルの文型は、被覆率が高くなることを狙って作成されたものがあるが、実際の対訳例文で節レベルまで汎化できるものは少なく、結果的に句レベルよりも小さい被覆率しか得られなかったものと考えられる。

また、混合レベルの被覆率が他のレベルに比べて余り向上していないが、これは句レベルの文型に比べて節レベルの文型の数が少なく、句レベルの文型以上に広い範囲をカバーできていないことを意味する。なお混合レベルでの被覆率が90%を超えていることから、文型数と被覆率の関係はほぼ飽和状態に達しており、これ以上標本量を増やしても被覆率は簡単には向上しないのではないかとと思われる。

以下、文型の汎用度と文型に定義されない要素について考察する。

a) 各レベルの文型の汎用度

上記(2)で述べたように節レベルの文型の被覆率は比較的小さいが、1文型あたりの汎用性はかなり高いと思われる。そこで、文型の汎用度を「文型再現率/文型数」に比例するとし、単語レベルの汎用度を1とすると、句レベル、節レベルの汎用度は、それぞれ1.8、12.5となる。すなわち、句レベルの文型は単語レベルの文型より約2倍汎用性があり、節レベルの文型はさらにその6倍程度汎用性があることになる。

b) 文型にカバーされない入力文要素

既に述べたように、適合した文型によって入力文のすべての要素が解釈される訳ではない。文型再現率と文型一致率の差から計算すると、平均して、単語レベルでは32%(=1-47.2%/69.8%)、句レベルでは14%(=1-76.2%/89.0%)、節レベルでは、19%(=1-63.1%/78.1%)の文要素が文型では解釈されない要素として残されることになる。この部分の翻訳方法については、英文生成の過程で注意すべきことである。

(2) 被覆率の飽和特性

次に、異なり文型数と「文型一致率R2」の関係を図1に示す。图中、文型数2.5万件以下の波線の部分は外挿したものである。また、参考のため、全体の文型での「文型再現率R1」の値も図示した。

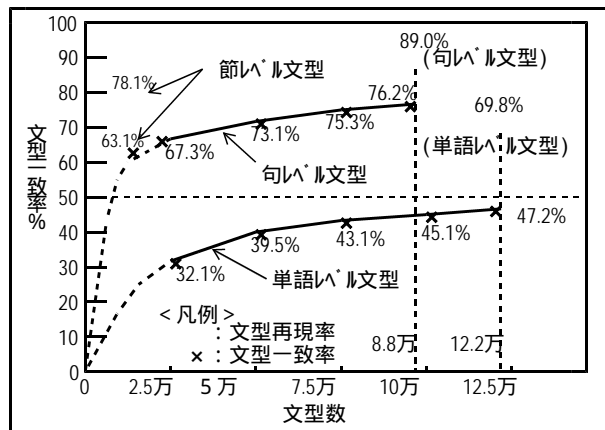


図1. 文型一致率（文字数で見た再現率）

これより、以下のことが分かる。

- (1) 単語レベル、句レベルのいずれの文型も、異なり文型数が数万件になると飽和傾向が現れる。
- (2) いずれの場合も異なり文型数が1万件以下では、有効な被覆率は得られそうにない。
- (3) 文型一致率は、まだまだ向上の余地は残されているが、現在の文型化の方法では、現状(10万件程度)の文型数をこれ以上増加させても、被覆率の向上はあまり期待できそうにない。

以上から、これ以上標本文数を増加させても、その割に被覆率は向上しないとみられる。むしろ、汎化方法について、さらに改良の方法を検討することが重要と判断される。

(3) 文型種別と被覆率特性の関係

表3の結果について、その内訳を調べるため、単文レベルの文型について文種毎の「文型再現率R1」を求めその結果を表5に示す。

表5. 文種別毎に見た単語レベルの文型再現率R1

文種別	1	2	3	4	5	全体
文型再現率R1	70.1%	63.2%	71.8%	65.2%	65.8%	69.8%

この表から、単語レベル、句レベル共に、文型再現率は文種別に依存せずほぼ一定であり、バランスの良い標本集合となっていることが分かる。

4.2 適合の仕方と適合文型数

(1) 完全適合と部分適合の割合

入力日本文に対して、完全一致文型が存在する割合と部分一致文型しか存在しない割合を調べ、表6にその結果を示す。

表6. 完全一致と部分一致の割合

レベル	完全一致文型*	部分一致文型**	合計(R1)
単語レベル	15.0%	54.8%	69.8%
句レベル	54.3%	34.6%	89.0%
節レベル	39.5%	38.6%	78.1%
混合レベル	56.2%	35.6%	91.8%

<凡例>*：入力文に対して完全に一致した文型がある入力文のある割合。 **：入力文に対して部分的に一致した文型しかない入力文の割合

この表から以下のことが分かる。

- (1) 単語レベルの文型では、「完全一致」に比べて「部分一致」となる割合は4倍近い。
- (2) これに対して、句レベルの文型では、この関係が逆転しており、「完全一致」の割合は1.5倍である。
- (3) 節レベルの文型では、「完全一致」と「部分一致」の割合が拮抗している。
- (4) また、混合レベルは句レベルに近い値である。

これらは、各文型の性質をよく表していると思われる。まず、単語レベルの文型に比べて句レベルの文型の完全一致が多いのは、以下の理由によると考えられる。すなわち句レベルの句変数には元の標本文中の句を変数化しただけでなく単語変数の適用範囲を句レベルに拡大したものなどがあり、入力文の中に元の標本文にはないような名詞修飾や格要素が入力文にあっても、それらは名詞句、動詞句などの一部として解釈されるためである。これに対して節レベルの文型の「完全一致率」それほど高くないのは、文型数そのものが少ないためと考えられる。

ところで、節レベルの文型数が少ないのは、節レベルで文型化できない標本文が多かったためである。これは多くの文は節レベルでは非線形であり、複数の節に分離して翻訳しその結果を合成するような要素合成法の方法では良い翻訳ができないことを意味している。また、(4)にあるように混合レベルの被覆率は句レベルと大差ないこと、節レベルの文型では、高品質の翻訳は期待しにくいことから、節レベルの文型には余り期待しない方が良さそうである¹⁾。

(2) 適合文型数

文型との照合では、離散記号の存在によって、同一の文型でも、さまざまな適合の仕方(解釈の多義)が発生する可能性がある。そこで、適合文型を1つ以上持つ入力文について、1文あたり適合文型数を表7に示す。

表7では、適合した異なり文型数の他に、延べ適合文型数と適合文型当たりの解釈の数についても掲載した。なお、表中の混合レベルは、3レベルの和である。

表7. 入力文1文に対する適合文型数

文型種別	異なり文型数の平均	延べ適合文型数の平均	適合文型当たりの解釈の数
単語レベル	9.79件	14.05件	1.44件
句レベル	51.48件	164.72件	3.20件
節レベル	9.85件	41.37件	4.20件
混合レベル	71.09件	220.11件	3.10件

(適合文型を持つ入力文の場合)

この表から以下のことが分かる。

- (1) 単語レベルでは、入力文あたり適合する文型数が少ない。また、文型当たりの解釈の数は最も少ない。
- (2) 句レベルの異なり文型数は、単語レベルの約5倍も多く、文型当たりの解釈の数も2倍以上である。
- (3) 節レベルの場合は、異なり文型数は少ないが、文型当たりの解釈の数は最も多い。

同一の文型に適合した時の複数の解釈の中で、意味的に正しい解釈(適合の仕方)は最大1つと考えられる²⁾。単語レベルで適合文型当たりの解釈多義が1.44件であることは、文法的レベルの文型でも、単語レベルでは、かなり個別性の高い文型となっていることが分かる。句レベル、節レベルになるに従って、文型の解釈の多義が増加していることから、これらのレベルでは、今後、意味的な制約など、より強力な適合条件を指定することが必要と考えられる。

4.3 日本語基本文型に対する被覆率

前節まででは、文型作成に使用した例文を対象に「自己文型」以外の文型にどれだけ適合するかを調べた(クロスヴァリデーション法)が、本節では、文型作成用の標本とは別の日本文を対象に被覆率を調べる。

試験用として使用する例文は、「外国人のための日本語例文・問題シリーズ」全18巻(名柄,加藤,福地 1989)に掲載された日本語例文から抽出したもので、日本語基本文型に対する被覆率特性を知ることができると期待される。具体的には、まず、上記シリーズの例文12,000件に含まれる下記の重文複文(約8,200件)を取りだした。

- 文種別1(4,200文)、文種別2(900文)、
- 文種別3(1,900文)、文種別4(230文)、
- 文種別5(1,000文)

¹⁾ 特に節変数が2つ以上使用される文型は、節レベルでの要素合成法が適用でき、従来の要素合成法による翻訳が有効な場合が多いと予想される。

²⁾ 異なり文型の中には意味的に正しいものが複数存在することがある。また、それらに対応する英語文型が異なるときは、微妙な意味の違いによる訳し分けの対象となる。

次に、これらのなかからそれぞれ1/10の割合で、801件の例文をランダムに選択し、試験用の入力文とした。なお以下では、この試験文を「基本文型試験文」と呼ぶ。

(1) 被覆率特性

被覆率実験の結果を表8に示す。

表8. 「基本文型試験文」に対する被覆率

文型の種別	文型再現率 (R1)			文型一致率 (R2)
	完全一致	部分一致	合計	
単語レベル	4.9 %	48.1 %	52.9 %	30.0 %
句レベル	24.7 %	58.2 %	82.9 %	64.5 %
節レベル	22.2 %	52.2 %	74.4 %	54.8 %
混合レベル	28.5 %	57.3 %	85.8 %	67.7 %

この表と表4、表6を比較すると以下のことが分かる。

- (1) 「基本文型試験文」では、単語レベル、句レベル、節レベルの「文型再現率R1」と「文型一致率R2」が共に4~17%低下しており、低下の割合は、単語レベルが最も大きい。
- (2) 完全一致の割合は、3レベルとも10%~30%の範囲で低下している。相対値で見ると、単語レベルの低下は著しく、約1/3となっている。
- (3) しかし、部分一致の割合は、7~24%向上しており、完全一致率の低下が大きいものほど、部分一致の向上の割合が大きい。

このうち、(1)で単語レベルの値が最も大きく低下するのは、このレベルの文型は元となった標本文の個別的な特徴の多くを引きずっているのに対して、句レベル、節レベルではかなり一般化した文型となっていることを意味する。これは、単語レベルの完全一致率が大きく低下していること、また文型数が約1/10の節レベルの文型でR1、R2共にあまり低下しないことから裏付けられる。

また、(2)と(3)の特徴は、テスト文の母集団が標本文の母集団とは異なることから予想されたことである。母集団の異なる入力文では、文型再現率と文型一致率が低下する(特に単語レベル)だけでなく、完全一致の文型が減少し、部分一致の文型が増加することは、文型適用法を検討する上で、この種の文型が無視できないことを意味している。

(2) 入力文1文に適合する文型数

適合する文型のあった入力文の場合、1入力文にどれだけの文型が適合したかを表9に示す。

表9. 入力文1文に対する適合文型数

文型種別	異なり文型数の平均	延べ適合文型数の平均	適合文型当たりの解釈の数
単語レベル	4.53件	7.71件	1.70件
句レベル	33.73件	120.10件	3.56件
節レベル	9.75件	34.85件	3.57件
混合レベル	48.01件	162.66件	3.37件

この表と表6から以下のことが分かる。

- (1) 単語レベルでは、異なり文型数、延べ適合文型数共に、ほぼ半減している。
- (2) それに対して、句レベルの減少は30%程度にとどまっており、節レベルでは、殆ど低下していない。
- (3) 適合文型当たりの解釈の数は、表6の場合と目立った違いはない。

これらは、単語レベルの文型は汎用性が乏しく、母集団が異なる入力文に対して弱いが、句レベル、節レベル

になるにつれて汎用性が増大することを意味している。同一母集団のテスト文では、節レベルの文型の有用性が乏しかったが、異なる母集団の入力文で効果を発揮することも予想される。

5. 文型記述方式と被覆率の関係

4.1節の評価結果から見て、現状の文型の被覆率は実用上まだまだ十分とは言えない。現在の文型は主として対訳例文の線形な要素を変数や関数に置き換えることによって作成されているが、線形な要素でも訳文生成で役立つ要素については任意要素として文型に記述されている。また、助詞、助動詞は表記のレベルで関数化されている。しかし、異なる時制や様相への適合は考えられていないこと、日本語に顕著な語順の自由度も考慮されていないことなど、種々の改良すべき余地を残している。

そこで、本章では、文型記述方法の違いによる被覆率特性の違いについての基礎データを得るため、「任意要素指定が被覆率に与える効果」と「時制、相、様相の情報被覆率に与える影響」についての評価実験を行った。

5.1 任意要素指定機能の効果

文型記述で使用する任意要素は、「原文任意要素」(「離散記号」による)と「文型任意要素」(「省略記号」による)に分けられる。

前者は、任意の文型要素が存在しても良い位置を表すものである。該当する位置に現れた入力文の要素は、別途翻訳し、英文に組み込まなければならないが、組み込むべき位置を知るためには英文解析技術を必要とする。これに対して、後者は、英文中に訳出すべき位置情報が指定されているため、該当する部分の翻訳を組み込むことは容易である。いずれも、単語レベル、句レベルの文型の被覆率向上を目指したものである。

そこで、このような2種類の要素指定の有無と被覆率の関係について「母集団試験文」1万件を対象に調査した。その結果を表10に示す。

表10. 任意要素指定の効果

場合分け		文型再現率 (R1)	
離散記号	文型任意記号	単語レベル	句レベル
あり	あり	69.8 %	89.0%
	なし	47.6 %	83.7 %
なし	あり	15.4 %	55.6 %
	なし	10.7 %	48.7 %

この表から以下のことが分かる。

- (1) 離散記号(「原文任意要素」指定記号)の有無を比較すると、単語レベルでは、この記号の使用によって文型再現率は5倍近く向上しており、句レベルでも1.5~2倍近く向上している。
- (2) 「文型任意記号」も被覆率向上に与える効果は大きい。単語レベルでは1.5倍程度、句レベルでは、1.1倍程度で「離散記号」の場合に比べてその効果は少ない。

これらから、離散記号は被覆率向上で大きな効果を持つが、「文型任意要素」の効果は今ひとつ少ないように思われる¹⁾。現状の文型を見ると字面表記の揺らぎが十分に吸収されているとは言えないので、今後、徹底した異表記のグループ化を行う必要があるであろう。

¹⁾ 「離散記号」は、むしろ効き目があり過ぎて、意味的に不適切な文型への適合数を大幅に増大させている嫌いがある。その点での使用基準を見直す必要があるようだ。

5.2 時制・様相などの縮退効果

単語レベルの文型化において、時制、相、様相の表現は原則として関数化されているが、文型照合段階での時制変形、相変形、様相変形などの操作を認めることにより、これらの情報を文型から削除することができれば¹⁾、文型の汎用性は大幅に向上することが期待される。そこで現在使用されている時制、相、様相の関数を無視した場合、文型被覆率がどのように変化するかを調査した。

本実験では、文型の作成で使用した12.9万件の日本語（「母集団試験文」）を入力文として、単語レベルの文型辞書12.2万件と照合し、文型再現率を求めた。その結果を表11に示す。また、文型一致率と適合文型数の平均を表12に示す。

表11. 時制・相・様相の縮退効果（単語レベル）

条件	完全一致文型	部分一致文型	文型再現率(R1)	文型一致率(R2)
削除前	14.4%	50.3%	64.7%	42.08%
削除後	6.6%	91.8%	98.4%	64.49%

表12. 文型一致率(R2)と一致文型数平均

条件	一致文型数平均/入力文		適合文型当たりの解釈の数
	異なり文型数	延べ文型数	
削除前	9.45件	22.92件	2.43
削除後	73.5件	255.8件	3.48

表10、表11の「削除前」は現状の文型の場合を示すが、その値は、表3の値と若干異なる値となっている。これは、テストに使用した入力文が異なるためである（いずれも「母集団試験文」であるが、表3は1万文が対象）。

表9から、時制、相、様相の削除によって、完全一致率は大きく減少する²⁾が、部分一致が大幅に増えて、100%近い文型再現率になり、被覆率向上で大きな効果が期待できることが分かる。

また、表10からは、削除後の文型一致率は1.5倍程度に向上するが、その反面、1入力文に対する適合文型数とその解釈の数は10倍程度に増加することが分かる。

これらのことから、今後、汎化により、時制、相、様相の情報を極力縮退させることができれば大きな効果が期待される。しかし、その場合は、意味的に不適切な適合文型が大幅に増大するので、意味的な制約条件を付与するなどにより、意味的に不適切な文型への適合を抑止することが重要と考えられる。

6. 意味的排他性の評価

前章まででは、文法レベルで記述された文型について「文型再現率」と「文型一致率」を用いて再現性を評価

表14. 意味的な正しさの評価

評価パラメータ	意味適合率 P1			適合文型意味正解率 P2			適合文型正解含有率 P3		
	完全一致の場合	部分一致の場合	全体(P1)	完全一致の場合	部分一致の場合	全体(P2)	完全一致の場合	部分一致の場合	全体(P3)
レベル2以上	97.3%	18.5%	34.6%	97.6%	17.2%	24.8%	97.6%	26.7%	44.1%
レベル3以上	80.0%	13.0%	28.2%	77.3%	11.8%	18.4%	81.0%	17.8%	35.6%
レベル4	50.3%	6.6%	16.5%	41.1%	5.4%	9.1%	50.0%	9.6%	20.9%

した。しかし、適合した文型が果たして意味的に適切な文型であるか、また、それに対応する英語文型が訳文の生成に問題なく使用できるかは大変重要である。

そこで、本章では、意味的排他性が最も高いと推定される単語レベルの文型対象に意味的排他性に関する評価を行った³⁾。

6.1 評価の方法

(1) 評価パラメータ

本章で評価したパラメータは、「意味適合率 P1」、「適合文型意味正解率 P2」、「適合文型正解含有率 P3」の3つである。

(2) 評価の手順と判定基準

まず、文型作成に使用した例文12.9万件（「母集団試験文」）に対する文型照合実験結果から、自己文型以外に1件以上の適合文型を持つ入力文を抽出した。次の中から、ランダムに1/400のテスト文（177文）を選び、それらに適合する文型（自己文型を除く1073件）の各々について表13に基づく正誤判定を行った。

表13. 評価の基準

評価レベル	評価区分	説明
レベル1	一見して不正解	基本要素が一致していない
レベル2	一見正しく見える	入力文の骨格要素揃っているが、統語構造が正しくない
レベル3	統語的には正しい	英語文型が使えない
レベル4	意味的にも正しい	英語文型が使える

6.2 評価結果

評価結果を表14に示す。このようから、P1, P2, P3について以下のことが分かる。

(1) 意味適合率(P1)について

- 完全一致文型の場合は、入力文の骨格がほぼ網羅（97%）されており、多くは（80%）は、意味的にも正しく捉えている。また、50%は対応する英語文型をそのまま使用して訳文が作成できる。
- これに対して、部分一致文型の場合は、入力文の構造を捉えているものは少なく（18%）、対応する英語文型パターンがそのまま使えるものは極めて少ない（7%）。
- それらの結果、翻訳に使用できる文型は、適合文型のうち2~3割程度と見られる。

以上の通り、「完全一致文型」と「部分一致文型」の意味適合率が5~10倍も異なることは、適合文型から機械翻訳で使用する文型を選択する方式において、考慮すべき重要な点と思われる。

*1 例えば、過去形の例文や推量の意味の例文から原型に縮退させた文型パターンを作成し、翻訳時には、この文型から過去形、未来の表現、推量や否定などの表現を生成する方法である。この方法は非線形な要素に対しては適用できない。本節の検討は、すべての削除できたときの効果を調べるもので、効果の最大値を知ることが目的としている。

*2 これは、現在形の基本形式のテスト文しか文型と完全一致しないようになるため、当然の結果である。

*3 この評価は、適合した多数の文型の意味判定作業が必要で多大なコストがかかるため、もっとも排他性の高ことが期待される単語レベルに限って評価した。

(2) 適合文型意味正解率 (P2) について

完全一致と部分一致での違いやレベル違いによる「適合文型意味正解率 P2」の相対的な違いは、「意味適合率 P1」の場合とほぼ同様であるが、全体の P2 の値は、P1 に比べて 6 割程度に低下する。

これは、翻訳に使用する文型の選択では、適合した文型からランダムに選択する場合に比べて、一致文字数最大のものを選ぶ方がよいことを意味する。

(3) 適合文型正解含有率 (P3) について

完全一致と部分一致での違いやレベル違いによる文型品質の違いは、「意味適合率 P1」や「適合文型意味正解率 P2」の場合とほぼ同様であるが、品質は、「適合文型意味正解率 P2」に比べて 2 倍、「意味適合率 P1」に比べて 1.3 倍程度向上する。

これは、適切な文型選択の方法があれば、単純に入力文と文型との一致文字数の大きいものを選択する方法に比べて意味的正解率は 1.3 倍まで向上する可能性があることを示している。

6.3 意味的正解率に関するまとめ

以上、単語レベルの適合文型の意味的な正しさに関する 3 種類の検討から以下のことが分かる。すなわち、適合文型の中から翻訳に使用する文型を選択するとき、

- (1) ランダムに文型を選べば、それが意味的にも適切な確率は、約 15%
- (2) 入力文と適合する範囲が最大の文型（「最尤文型」）を選べば、それが意味的にも適切な確率は約 23%
- (3) もし意味的にも最適な文型を選ぶ方法が分かった時は、意味的な正解率は 28% となる。なお、
- (4) 適合文型のうち、「完全一致文型」の適合率は、「部分一致文型」の 5 ~ 10 倍大きい。

上記 (1) ~ (3) 4.1 節の結果を組み合わせれば以下のことが分かる。すなわち適合した文型から最適な文型が選択できる方法が実現できれば、入力文の約 20% (=0.69.8% × 28%) に対して意味的にも適切な文型が発見できること、またその場合、各入力文の 2/3 の部分は対応する英語文型をそのまま使えば翻訳できることである。

しかし、現実の処理では、適合文型の中から意味的にも最適なものを選ぶことが困難である¹と予想され、上記 (2) の方法が現実的であると思われる。

7. 評価結果のまとめと考察

前章までの評価結果をまとめ、今後、文型の適合率と正解率を向上させる方法について考察し、文型を使用した機械翻訳の可能性について考える。

7.1 評価結果のまとめ

(1) 文型の被覆率と正解率の問題

単語レベル、句レベル、節レベルの文型の「文型再現率」は、それぞれ、70%、89%、78% であり、「文型一致率」は、47%、76%、63% である。これらは、母集団（出典）の異なるテスト文では、10% あまり減少する。

これらの値はかなり良い値と思われるが、意味的に見た場合、入力文の翻訳に適用できる文型はこれらの 1/4 程度にとどまっており、実用に耐える品質とは考えられない。これが文法レベルで記述された文型の限度と思われることから、今後は、文型を意味レベルで記述するこ

とが重要と考えられる。

(2) 文型数と被覆率の関係

重文、複文の文型化において、標本文が 1 万件以下の場合には、良い被覆率は得られそうにないが、逆に、標本文数を数万件以上に増大させても飽和現象が発生し、その割には、被覆率向上は上がらない。

(3) 節レベルの文型の被覆率

節レベルの汎化ができる標本文は少ないため、全体から見ると、その被覆率は、単語レベル、句レベルの文型に比べて比較的小さい。従って、節レベルの文型の役割は保険として捉え、被覆率向上の点では、句レベルの文型に重点を置いて今後の汎化を進めるのが適切と判断される。

(4) 適合文型を使用した機械翻訳

単語レベルの文型では、「部分一致文型」に比べて「完全一致文型」の割合は低い。また句レベルの文型でも適合する文型のかなりのものが「部分一致」である。

(5) 意味的にも適切な文型の選択

適合文型を持つ入力文について見ると、適合する文型 (3 ~ 60 件) は大変多い。また、適合する文型一件当たりの適合の仕方 (平均 10 通り位) も多い。この中から意味的にも適切な文型と解釈を決定する問題は今後重要な問題になる可能性があり、今後の文型の汎化と改良はこの点を重視して進めるべきである。

なお、文型選択では、意味的正解率は、「完全一致文型」の意味的正解率は、「部分一致文型」に比べて 5 ~ 10 倍大きいので、それを考慮する必要がある。

7.2 今後の対策について

既に述べたように文型方式の成否の鍵は、どれだけの被覆率と意味的排他性が実現できるかであるが、被覆率について明確な成否の分岐点を設けるのは難しい。

本稿では、文型は、非線形な構文構造を対象に適用される技術としている。そこで、日本語重文、複文において非線形な構文構造の占める割合を 50% と仮定すると、文型は、これらの非線形な構文構造がカバーできればよいから、意味適合率の達成目標は 50% となる²。

しかし、この目標に比べて、現状の意味的な正解率は極めて低く、適合した文型から意味的にも正しいものを選択することも容易でないと推定される。また、意味的排他性も実現されていない。従って、今後の問題は、「文型一致率」と「意味適合率」の両者を上げること、また同時に、「文型の意味的排他性を向上させ意味的にも不適な文型に適合しないようにすること」の 2 点であると考えられる。

そこで、以下では、これらの 2 つの課題の可能性について考える。

(1) 意味的にも排他性向上のための課題について

前節の (5) から、適合文型の中で、意味的にも正しい文型を選択することはかなり困難な仕事だと思われる。この問題を解決する方法としては、文型で使用された変数に意味的制約条件を付与することが有効と見られる。そこで、どの程度の分解能で意味属性を付与すると良いかであるが、本検討結果から見て、文型では、平均数十通りの意味が縮退していると思われるから、平均で数十通り、重なり大きいところを考えると 1000 通りくらいの分解能（「日本語語彙大系」(池原ほか 97) の意

¹ 意味的にも最適な文型パターンを選択するには、別途、入力文の意味解析を行うことが必要となり、意味解析の実現を目指した文型パターン方式の目的に反する。

² これは、日本文中の重文、複文のうち、50% は直訳が困難な文型であることに相当する。現実に訳しにくいものの上位 50% が適切に意識できるようになると全体の翻訳品質は格段に向上すると思われるから、ここでは、これを目標として考えることにする。

味属性の分解能のレベルに相当)があれば、ほぼ意味的な多義は解消できそうである。

(2) 被覆率向上のための課題について

変数に意味的制約条件を付与した場合は、意味的適合率は大幅に向上すると期待できるが、逆に文型再現率と文型一致率は、低下することが予想される。従って、これらの値を大幅に向上させる方法が必要である。

その対策であるが、「日本語語彙大系」の開発例と今回の被覆率調査から見て(1)「格要素の語順の自由指定や副詞の位置変更可能指定を行うこと」、(2)「時制、相、様相の汎化を行うこと」、また、(3)「字面表記の自立語、付属語の表記の揺らぎを徹底的にグループ化すること」が重要と考えられる¹⁾。

そこでまず(1)について考えると、格要素の語順変更指定の効果は、単語レベルの文型から句レベルの文型に汎化した場合と同程度の被覆率向上が見込めそうで、適合文型は何倍かに増加する可能性がある。

次に、(2)の効果を予測するため、現在使用されている時制、相、様相の情報を汎化した場合どれだけの効果が期待できるかに関する今回の調査結果を見ると、単語レベルの文型一致率が、54%から98%に向上する可能性が示されている。このことから、汎化不能な情報は残したとしても、適合文型はかなり大幅に増大することが予想される。

第3に(3)の効果を推定するため、現在の文型で使用されている「任意要素」指定の機能の効果を見ると、「文型一致率」は、離散記号の有無によって、単語レベルで4~5倍、句レベルで1.5倍~2倍の差が生じている。文型任意記号についても単語レベルでは、数十%の違いがあり、適合文型の大幅な増加が期待できる。

(3) 文型を使用した機械翻訳の可能性

以上、本節の考察は、かなり大胆なものとならざるを得なかったが、文法レベルで文型を記述した現段階での評価では、日本語重文、複文の非線形構造が50%と見た場合、それをカバーするような文型が十万件程度の文型として記述できる可能性があると推定される。今後は、前項の3つの課題を早急に行い、結果を確認することが必要である。

8. あとがき

日本語重文、複文を対象に文法レベルで記述された文型それぞれ12.2万件、8.8万件、1.1万件を対象に被覆率特性(「文型再現率」、「文型一致率」、「意味適合率」、「適合文型意味正解率」、「適合文型正解含有率」の5種)を調べ、今後の課題と文型方式の可能性について検討した。

その結果によれば、文法レベルで定義された現状の文型の文型再現率は、単語レベル、句レベル、節レベルの順に70%、89%、78%で、かなり高い値を示すが、完全一致する文型は少なく、文型を使用した翻訳では部分一致した文型をうまく使用することが大切である。

また、単語レベル文型の場合、適合文型の中に意味的に正しいものが含まれている確率が、20%弱にとどまっていること、また、入力文は多数の意味的に適合しない文型と適合してしまうことは、文法レベルの情報で記述された文型の限界を示しているものと見られるため、早急に意味レベルの文型化を試みる必要がある。

今後は、汎化レベルと意味的排他性をさらに向上

させる方法について検討することが必要である。このうち、の対策としては、格要素の語順の自由指定や副詞の位置変更可能指定を行うこと、時制、相、様相の汎化を進めること、また、字面表記の自立語、付属語の表記の揺らぎを徹底的にグループ化すること等が挙げられ、これらによって意味的に適切な文型の割合は数倍程度以上に向上することが期待される。

また、では、文型内の変数の変域に対して意味的制約条件を付けることが必要と考えられるが、現在の多義の発生程度から考えると、その意味的分解能としては1000種類の分類があれば良さそうである。

今後は、意味的な適合率50%以上の実現を狙ってこれらの課題に取り組んで行く予定である。

謝辞

この研究は、科学技術振興機構(JST)の戦略的基礎研究事業(CREST)で行っているものである。ご議論を頂いた宮崎正弘先生(新潟大)、佐良木昌先生(長崎純心大)、池田尚志先生(岐阜大)、新田義彦先生(日本大)、山本和英先生(長岡技科大)、横尾昭男様(NTT)に感謝する。また、文型化作業を担当して頂いたNTTアドバンステクノロジ株式会社の皆様に感謝する。

参考文献

- 有田潤(1987):「ドイツ語講座」南江堂, pp.48-56
池原悟、宮崎正弘、白井諭、林良彦(1987):言語における話者の認識と多段翻訳方式,情報処理学会論文誌,Vol.28, No.12, pp.1269-1279
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(1997):「日本語語彙大系」岩波書店
池原悟(2001):自然言語処理の基本問題への挑戦,人工知能学会誌, Vol.16, No.3, pp.422-430
Satoru Ikehara(2001): Meaning Comprehension Using Semantic Patterns in a Large Scale Knowledge-Base, Proceedings of the PACLING'01, pp.26-35
池原悟(2002):究極の翻訳方式の実現に向けて = 類推思考の原理に基づく翻訳方式 = , AAMT Journal, アジア太平洋機械翻訳協会, No.32, pp.1-7
池原悟(2003):言語で表現される概念と翻訳の原理,電子情報通信学会技術研究報告「思考と言語」研究会TL2003, pp.7-12
池原悟(2004):非線形な言語表現と文型パターンによる意味の記述,情報処理学会自然言語処理研究会,(2004.1/13)
池原悟、宮崎正弘、佐良木昌、池田尚志、白井諭、村上仁一、徳久雅人:機械翻訳のための日英文型パターン記述言語、電子情報通信学会、思考と言語研究会TL2002-48、pp.1-6(2003.3)
池原悟、阿部さつき、徳久雅人、村上仁一(2004):非線形な表現構造に着目した日英文型パターン化,情報処理学会第160回自然言語処理研究会(2004.3/4.5)
市川亀久彌(1963):「創造的研究の方法論」(増補版),三和書房
金出地真人、池原悟、村上仁一(2001):結合価文法による動詞の訳語選択能力の評価,情報処理学会第63回全国大会, 6Y-04, 2-267-268
名柄、加藤康彦、福地努(1989):「外国人のための日本語例文・問題シリーズ」全18巻,荒竹出版,1989

¹⁾「日本語語彙大系」の結合価パターンの開発作業では、これらの作業に加えて変数意味属性自身の汎化作業を徹底して行っており、大きな効果のあることが確認されている。しかし、意味属性汎化作業は膨大な人手コストがかかるため、今回の文型パターン化では、それは行わない方法が望まれる。