

日英機械翻訳における 原文自動書き替え型翻訳方式とその効果

白井 諭 池原 悟 河岡 司

(shirai@nttkb.ntt.jp)

NTT 情報通信網研究所

(神奈川県横須賀市武 1-2356)

あらまし 最近、言語間の発想の違いを克服し、機械翻訳の品質を向上させるための方法として、多段翻訳方式や用例翻訳方式が提案されている。しかし、これらの方法では、表現をカバーするためのルールや用例集が膨大になること、目的に合った用例を発見する方法が容易でないことが問題となる。

本稿では、言語間の発想の違いが、人手による前編集の対象となっている表現や構文に端的に表れていることに着目し、原文自動書き替え機能を加えた原文自動書き替え型翻訳方式を提案する。この方法は、翻訳困難な表現に既存の翻訳機能がそのまま利用できる点に利点があるが、同時に構文解析の曖昧さが減少することによる処理速度の向上などの効果も期待できる。これらの効果についても、この方式を多段翻訳方式に基づく日英翻訳実験システム ALT-J/E の上に実現して行なった走行実験結果に基づいて報告する。

キーワード 機械翻訳, 日英機械翻訳, 原文自動書き替え, 日本語内書き替え, 疑似的日本語表現, 多段翻訳方式

Effects of Automatic Rewriting of Source Language within a Japanese to English Machine Translation System

Satoshi Shirai Satoru Ikehara Tsukasa Kawaoka

(shirai@nttkb.ntt.jp)

NTT Network Information Systems Laboratories

(1-2356 Take, Yokosuka-shi, Kanagawa-ken)

Abstract To improve the quality of machine translation, it is important to develop a translation method that takes into account the conceptual differences between languages that cause difficult problems in translation. These problems typically occur with expressions that must be subjected to manual pre-editing of the source texts.

This paper proposes a translation method that includes an automatic source text rewriting function. This method has the advantage of being able to use existing translation functions for the translation of difficult-to-translate expressions. At the same time, it improves processing efficiency by reducing ambiguities in syntactic and semantic analysis. This method is an extension of the Multi-Level Translation Method, on which the Japanese to English Machine Translation System ALT-J/E is based.

Keywords Machine Translation, Automatic Rewriting, Pseudo Japanese, Multi-Level Machine Translation Method

1 はじめに

従来、機械翻訳において数多くのシステムが研究開発され、翻訳業務にも適用されるようになってきた^[1]。しかし、依然として訳文の品質が問題として残されており、新しい理論や方式の提案が期待されている^[2]。

訳文品質の向上を狙った翻訳方式として、最近、多くの研究が行われている^[3, 4]が、言語が話者の対象に対する見方、捉え方を表現する手段であることを考えると、異なる言語族間の翻訳においては特にこの違いが問題になると考えられる^[5, 6]。

言語における発想の仕方の違いに着目し、原文の意味を失わないように翻訳する方法としては、第1に「原文の表現や構造を分解しすぎないようにしながら、目的言語のなるべく全体の意味に対応する表現を探して置き換える方法」、第2に「原文の意味を変えない範囲で翻訳しやすい表現に書き替えてから翻訳する方法」が考えられる。

第1のアプローチの例としては、言語による話者の認識の違いに着目した多段翻訳方式^[7, 8]などが挙げられる。この方法では、話者の意志や判断を示す主体的表現と対象の姿を示す客体的表現とを分離して翻訳するが、客体的表現に対しては、構造の持つ意味を失わないように、構造の抽象化可能なレベルに応じて段階的に翻訳している。

また最近では経験主義的なアプローチも始められ、知識ベース型の翻訳^[9-11]や用例を利用した翻訳^[12, 13]などの研究が行われ、その効果が期待されている。用例翻訳の方法は原文の表現をダイレクトに目的言語に対応させることを狙っており、やはり第1のアプローチの例と考えられる。

第2のアプローチの例としては、従来から行われている人手による前編集を挙げることができる。最近では、前編集作業を支援するため、翻訳しやすいように言語を制限し、原文をチェックするためのプログラムの開発も行われている^[14, 15]。

言語による発想の違いは機械翻訳しにくいところに端的に現れていると考えられるので、これまで人手による前編集の対象となっている表現をも機械翻訳の対象範囲に取り込むことができるなら訳文品質は向上すると期待できる。

そこで、本稿では、第1のアプローチとして提案されている多段翻訳方式の上に、第2のアプローチの立場から、これまで前編集の対象になっていた機械翻訳困難な表現や構文を自動的に書き替える方法を追加した「原文自動書き替え型翻訳方式」を提案する。具体的には、日英機械翻訳において原文自動書き替えの対象となる表現や構文の種類と性質を調べ、全体を原言語内で書き替える項目と、原言語内書き替えが困難であるため部分的に目的言語の表現に書き替えるべき項目とに分け、書き替え方式と書き替えルール形式を提案する。

この方法は訳文品質の向上を狙ったものであるが、①書き替えられた表現に対しては既存の翻訳機能がそのまま利用できる

ため、新たな翻訳アルゴリズムを作成しなくても良い、②一定の表現構造を固定的に捉えるため、構文意味解析の曖昧さが減少し、処理速度が向上する、などの効果^[16]も期待できる。新聞記事翻訳への適用実験の結果に基づいてこれらの効果も示す。

2 書き替えの対象

2.1 自動書き替えの対象範囲

原文自動書き替えの対象となる表現は、以下の条件を満たす表現と考えられる。

条件1：そのままでは適切な翻訳ができない。

条件2：意味を変えないで別の表現に書き替えられる。

条件3：その書き替えを行えば翻訳可能となる。

条件4：既存の翻訳機能に対して悪い副作用を生じない。

これらのうち、条件1～3は人手による前編集の場合と同様であるが、条件4は異なる。以下、その具体的な内容について検討する。

(1) 機械翻訳不能の原因

まず、条件1について考える。実際のドキュメントで適切な機械翻訳ができない表現を分類すると、ほぼ次のようになる。

(i) 原文が間違っている

① 言語表現の約束違反（誤字、脱字、構文の誤り）

② 表現または内容があいまい

③ 内容が間違っている

(ii) 既存の翻訳技術で翻訳できる範囲であるが、使用しているシステムでは能力が足りない

① システム（辞書、ルール）のバグ

② 該当する表現を翻訳する機能が組み込まれていない

(iii) 高度な意識などが必要なため現状では翻訳困難である

① 原言語の表現に直接対応させられる目的言語の表現がないため、話者の意図を判断して言い直す必要がある

② 慣習の違いなどにより、訳す必要がない、または、補足する必要がある

これらのうち(i)は③を除いてテキスト校正の対象範囲であり、従来から数多くの研究が行われている^{*1}。日英機械翻訳で問題となるのは(ii)と(iii)である。

(2) 意味を変えない書き替え

次に、条件2について考える。人手による前編集の場合は、原言語内に意味を変えない別の表現が存在しなければ書き替えることはできない。これに対して、翻訳システムの内部で書き替える場合は、原言語内に別の表現がなくても、目的言語の適切な表現を直接指示することにより救済することができる。

そこで、原文自動書き替えの対象を以下の2つに分類する。

(A) 着目する表現に対して、当システムで翻訳可能な別の原言語の表現がある場合（原言語内書き替え）

(B) 別の原言語の表現はないが、部分的に対応する目的言

*1 例えば日本語の場合、日本文校正支援システム REVIS E^[17]等によって、おおむね解決が図られている。

語の表現がある場合（疑似的原言語への書き替え）

このうち（A）は、原言語内の書き替えであるため、書き替え後は原言語の文としても意味のわかる文となる*2。これに対して（B）は、目的言語固有の表現に対応づけるための書き替えであるため、書き替え後は必ずしも原言語の文として意味が通じる必要はない。

（3）書き替え後の翻訳の可否

条件3の書き替え後に翻訳可能となるか否かは、人手による前編集の場合と同様であり、実験的に確認する。

（4）副作用のない書き替え

自動書き替えの場合は、人手による前編集の場合と異なり、登録した書き替えルールは該当するすべての表現に適用される。このため、ルールが適切でなければ、書き替えてはならないものまで書き替えてしまう可能性がある。ルール作成時には、訳文品質のデグレードが生じないように、その適用範囲を十分検討して置く必要がある。これも実験的な確認が必要である。

2. 2 書き替え対象表現の分類

日英機械翻訳の場合を例にとり、日英翻訳システムALTE-J/Eの実験結果に基づいて、6種類の日本語自動書き替え項目を提案する。

2. 2. 1 日本語内書き替え

原言語（日本語）内書き替えの対象となる項目を以下の3種類に分類する。ただし、日本語内書き替えが可能であっても、書き替え後の表現が曖昧になるものは、解析が正しく行なわれない可能性があるため、直接英語を意識した疑似的日本語に書き替えることとし、本節の分類から除外し、次節の分類に加えた。

（1）縮約展開型の書き替え

動詞が並列に並べられると、活用語尾が省略され、見かけ上名詞に解釈される現象が発生する。例えば、a-1)では、動詞「追加する」の語尾「する」が省略されているため、名詞「追加」と解釈され、文全体の意味解析に失敗する。

a-1) システムが追加および削除するデータ

a-2) システムが追加しそして削除するデータ

このような失敗を防ぐため、活用語尾を補い、a-2)のように原文を書き替える。

また、動詞を共有する複数の文では、前側の動詞が省略される場合がある。例えば、b-1)では、動詞「担当する」だけでなく、助詞「を」も省略されているため、「米国」と「副社長」が並列に見え、助詞「は」の解釈に支障が生じる。

b-1) 社長は米国、副社長は欧州を担当する

b-2) 社長は米国を担当し、副社長は欧州を担当する

このような場合、格要素の対応関係を見て、省略された述語を補ってあげば、意味解析は容易になる。

（2）冗長除去型の書き替え

通常、接続助詞「ば」は条件を表すが、c-1)のように並列を表す場合がある。

c-1) 男もいれば、女もいる

c-2) 男も女もいる

もともとc-1)の表現のニュアンスは翻訳困難であるから、内容的には同等のc-2)に書き替える。

（3）構文組み替え型の翻訳

日本語の構文に対応する英語の構文がない場合、日本語の文全体の構造を組み替えて英語に対応させるようにするもので、原文からは想像がつかないような英文を生成することができる。例えば、d-1)の文は、主語、目的語の双方が省略されているため、そのままでは英語に翻訳できない。

d-1) 二機種合わせて月五百台生産する

d-2) 二機種の合計月産は五百台だ

これをd-2)のように英語構文に対応する文型に書き替える。

2. 2. 2 疑似的日本語表現への書き替え

疑似的原言語（日本語）への書き替えの対象となる項目を以下の3種類に分類する。

（1）独立句的表現の書き替え

日本語の動詞性の副詞節には、英語側では単純な前置詞句に訳せるにもかかわらず、直訳すれば従属節になり、訳文品質が低下するものが多い。a-1)の「乗る」は動詞であるが、「に乗って」は手段を表す“by”に対応する表現であるので、疑似的日本語として助詞相当語「に乗って」を設け、それに書き替える。手段を表す“by”に相当する日本語の表現は助詞「で」であるが、「で」は多数の多義を発生される可能性があるので使用を避け、疑似的日本語表現へ書き替える。

a-1) バスに乗って学校へ行く

a-2) バス「に乗って」学校へ行く

なお、b-1)の場合も「に乗って」が使用されているが、この場合は本動詞であるので書き替えの対象とはしない。

b-1) 数人がバスに乗って、残りは電車に乗る

（2）様相時制表現の書き替え

様相や時制は、通常、助詞や助動詞などの組み合わせ（主体的表現）によって表現されるが、名詞や動詞などによって客体化されて表現される場合もある。例えば、c-1)では名詞述語「予定だ」によって「計画（意志性）」が示されている。また、e-1)では名詞述語「ところだ」によって完結直後の情態を表出している。このような表現は、それぞれc-2) e-2)のように客体的表現から分離し、疑似的に主体的表現として処理するように書き替える。

c-1) 山谷電気は本社を東京に移す予定だ

c-2) 山谷電気は本社を東京に移す(+計画=plan to変形)

d-1) これは私が出した予定だ

*2 人手による前編集と同様、翻訳システムに合わせた書き替えで、原言語の表現として適切なものになるとは限らない。

- e-1) バスは出発したところだ
- e-2) バスは出発する (+ 完結直後形態 = 完了形変形)
- f-1) 古戦場は武士が戦ったところだ

ただし、d-1) f-1) の例は、ともに全体が「AはBだ」の構文であり、書き替えの対象とはならない。

(3) 接続表現の書き替え

文間の接続を表す語の中には、文字どおりの意味を持たず、そのまま英語にしたのでは意味不明になるような表現がある。例えば g-1) では「のに続き」は行為の順序を示しているだけであるので、内部表現としては、「順接」の接続属性を付加し、原文から削除するように書き替える。

g-1) 高機能を追加するのに続き、改良型を導入する

g-2) 高機能を追加する (+ 順接), 改良型を導入する

また、日本語では文節の順序がかなり自由に記述されるが、接続が関係する場合、文節の順序を変えることで解釈が容易になる場合がある。h-1) では係り受けの交差が発生しているが、h-2) のようにすれば容易に意味解析できるようになる。

h-1) 大統領を 成田に 見送りに 行く

h-2) 大統領を 見送る (+ 目的) 成田に 行く

3 自動書き替え方式

3.1 書き替えルールの形式

書き替えルールの形式を表1に示す。本ルールでは、書き替え対象となる表現は、原文中の単語の表記、品詞、意味属性のほか文節間の係り受け属性を使って記述される。

3.2 書き替え処理の起動のフェーズ

翻訳処理は、形態素解析、構文解析、意味解析などのフェーズから構成されるが、あまり早い段階での書き替えは、解析情報が不足しているため、ルールの適用対象を精密に指定することが困難で、2.1(4)節で述べたような悪い副作用が生じやすい。逆に、解析が深く進行した段階では、後で述べるような解析多義の削減効果が減少する恐れがある。

そこで、ここでは前述のルールの適用条件がチェック可能になる時点、すなわち、構文解析の直後に書き替え処理を起動することとする。図1に書き替え処理の翻訳システム上の位置と書き替え処理の構成を示す。

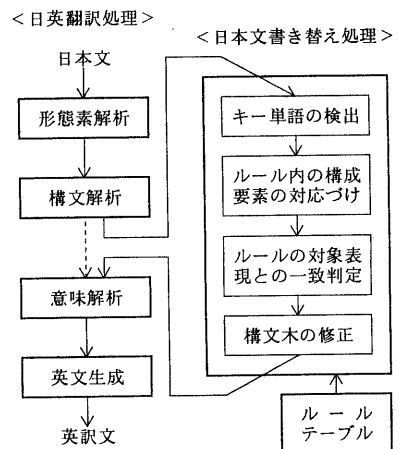


図1 原文自動書き替え型翻訳方式の構成

表1 日本文自動書き替えルールの形式

キー 単語	構文木 内の 位置	書き替えの対象表現			書き替え後の表現		
		単語構成	受け	係り	単語構成	受け	係り
および	X 1	〔サ変名詞〕	<任意>	X 2 (不明)	〔サ変動詞 = “する” を付加〕	<変更なし>	X 2 (接続関係)
	X 2	および(接続詞)	X 1	X 3 (副詞)	そして(接続詞)	<変更なし>	X 3 (接続関係)
	X 3	〔サ変動詞〕 [+*]	X 2	<任意>	<変更なし>	<変更なし>	<変更なし>
乗る	X 1	〔乗り物〕 +に(助詞)	<任意>	X 2 (格関係)	〔乗り物〕 +に乗って(助詞相当)	<変更なし>	X 3 (格関係)
	X 2	乗る(音便形) +て(助詞)	X 1	X 3 (接続関係)	<削除>		
	X 3	行く [+*]	X 2	<任意>	<変更なし>	X 1	<変更なし>

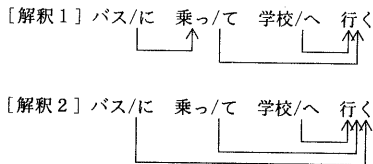
3. 3 構文解析多義の扱い

主として文法情報に基づく構文解析では、構文の解釈の候補が複数個残ることが多い。従って、同一の原文に対する解析結果でも、書き替えルールが適用可能なものと適用不可能なものが生じることがある。両者を比べると、適用する知識内容の違いから、以下の理由で、書き替えルールが適用できる候補の方が相対的に正しい解釈になっていることが推定される。

- ①構文解析（ここでは係り受け解析）では、主として単語の品詞や文節の種類などの文法情報が使用されるのに対して、書き替えルールには3. 1節で述べたように、単語の意味属性などの意味情報が使用される。
- ②構文解析では、文節間の関係は2項関係を基本に解析されるのに対して、書き替えルールには多項関係が記述されるため、表現構造の持つ意味が捉えやすい。

例えば、前節の（a）の例文に対しては、構文解析の結果は図2（a）の2つの解析多義を生じ、〔解釈1〕には表1のルールが適用されて（b）の結果が得られるのに対して、〔解釈2〕には適用できない。この場合、適用できない〔解釈2〕を削除することにより、解釈を一意に定めることができる。

（a）書き替え前の係り受け解析結果（多義数＝2）



（b）書き替え後の係り受け解析結果（多義数＝1）

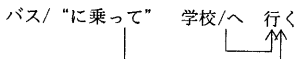


図2 書き替えによる解析多義削減の例

4 実験と評価

4. 1 実験と評価の条件

第2章、第3章で述べた自動書き替え方式を日英翻訳実験システムALT-J/Eの上に組み込み、日本文書き替えを実施した場合としなかった場合について、以下の条件のもとで比較評価を行った。

（1）対象試験文と実装したルール数

日経産業新聞の30記事のリード文を翻訳評価の対象とした。評価対象原文の平均の文字数は44文字／文、単語数は21単語／文である。各記

事のリード文は2～5文から構成され、文脈を持っているため、記事単位で翻訳し^{*3}た後、文単位で評価した。

また、書き替えルールは、上記の評価対象原文を含む新聞記事500文と機能試験文3800文の翻訳実験に基づいて作成した940ルールを実装した。

（2）訳文品質の採点基準

訳文品質の採点基準としては、ALPAC^[19]の9段階採点基準を以下の観点で見直した10点満点法を使用した。

- ①訳文だけで原文の意味が理解できるものを6点以上とし、合格とする
- ②簡単な後修正だけで通用する訳文が得られるものを8点以上とし、秀訳とする

なお、採点は翻訳会社の3名の日英翻訳家が互いに独立に行ない、その平均点を四捨五入した値を訳文の得点とした。

4. 2 実験結果と考察

日本文自動書き替え実験の結果を表2、表3に示す。試験に使用した102文に対して、書き替えルールが適用された文は44文（43％）で、延べの適用箇所は52箇所であった。以下、ルールの適用された文における訳文品質の変化と解析多義の変化について考察する。

表2 書き替え前後の評価点の変化

書き替え 後		書き替え後の評価点分布（平均点＝6.7）											合 計	
		不 合 格					合 格							
書き替え 前	点	0	1	2	3	4	5	6	7	8	9	10		
	0													35 文 （80％）
	1									1			1	
	2								1	1			2	
	3					4	1	2	5		1		13	
	4						1	2	2	3		1	9	
	5							1	3	4	1	1	10	9 文 （20％）
	6								3	1	2		6	
	7									2			2	
	8											1	1	
	9													
10														
合 計						4	2	5	13	11	5	3	1	44文
		11文（25％）					33文（75％）							

注1）対象試験文は新聞記事32記事102文（文字数は44字／文）で、本表は書き替え処理の対象になった文のみを集計した

注2）は品質低下領域を示す

^{*3}記事内の文脈から、省略されている主語と目的語を自動的に補完した上で翻訳した^[18]。

表3 書き替えタイプと訳文品質の向上効果

書き替えの種別	番号	書き替えのタイプ	ルールの適用箇所	訳文品質の向上	合格文数の増減	訳文語数の増減
日本語内書き替え	1	縮約展開型	7箇所(7文)	+1.7点	1→5	+4.3語
	2	冗長除去型	2箇所(2文)	+3.5点	0→2	-0.9語
	3	構文組み替え	12箇所(11文)	+1.6点	3→5	-0.1語
疑似的日本語表現への書き替え	1	独立句の表現	21箇所(19文)	+2.3点	3→15	-1.6語
	2	様相時制表現	7箇所(7文)	+2.0点	2→6	-2.3語
	3	接続表現	3箇所(3文)	+1.7点	1→3	0語
合計または平均			52箇所(44文)	+2.0点	9→33	-0.8語

注1) 対象試験文は新聞記事32記事102文(文字数は44字/文)。

注2) 複数の書き替えルールが適用された文が10文あるが、書き替えによる効果は適用されたルールタイプ別に異なり文を集計した。ただし、合計と平均は書き替えが行なわれた44文について集計した。

(1) 訳文品質の向上効果

本方式の適用前後の得点分布を図3に示す。

ルールが適用された44文の内、33文(75%)において訳文品質が向上し、102文全体の合格率は55%から79%に向上した。

なお、102文全体の平均点も適用前の5.71点から適用後は6.59点に向上したが、ルールは訳文品質の悪い例文に適用されたケースが多く、ルールが適用された44文の平均点についてみれば、適用前の4.3点から適用後は6.7点になり、2点以上向上した。

特に、書き替え前の訳文の評価が4~5点の場合、そのほとんど(15/19=79%)が6点以上(合格)となった。もとの評価が3点以下の文では、書き替え対象外の誤りの影響が

大きい、それでも合格まで向上した例がかなり(9/16=56%)あった。

ルール適用によって不合格(5点以下)から合格(6点以上)に変化した例文は24文であるが、その内訳は、日本語内の書き替えによるものが5文、疑似的日本語への書き替えによるものが18文、その両者によるものが1文で、疑似的日本語への書き替えの効果大きい。

疑似的日本語への書き替えは、後の英文生成処理への負担を減少させるとともに、書き替え後の翻訳誤りの発生を防止する効果があるので、今後さらに強化していきたい。

書き替えルールのタイプとその効果の関係をみると、独立句的表現の書き替えルールの適用例が最も多く、訳文品質向上効果も大きい。

(2) 解析多義削減効果

書き替えルールが適用された44文の構文解析の構造多義は、平均5.39から1.31に減少した。この現象は、適切な構造を決定する際の擾乱を防ぎ、訳文品質向上効果を生むと同時に、意味解析過程の高速化にも役立っている。

(3) 訳文語数への影響

縮約展開型の書き替えでは必然的に訳文の単語数が増加する(平均4.3語増)が、それ以外の書き替えでは減少している(平均1.8語減)。全体としてみれば、訳文の単語数は1文平均0.8語の減少にとどまっており、訳文のコンパクト化の効果はあまり期待できない。

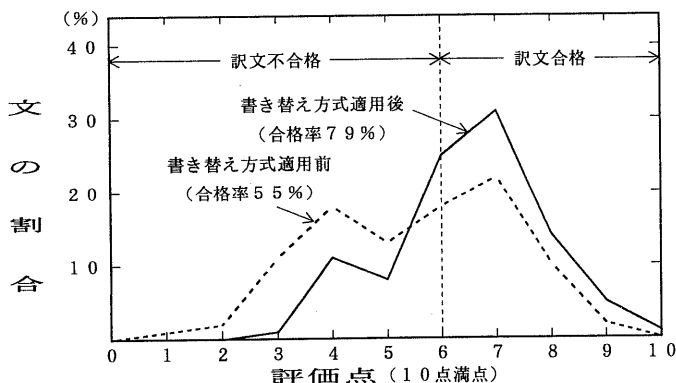


図3 書き替え方式の適用前後の訳文の評価点分布

5 おわりに

言語間の発想法の違いを克服し、機械翻訳の品質を向上させるための1つの方法として、既に提案されている多段翻訳方式（日英翻訳実験システムALTE-J/E）の上に、原文自動書き替え機能を加えた原文自動書き替え型翻訳方式を実現した。

具体的には、書き替える対象を①着目する表現に対して、当システムで翻訳可能な別の原言語表現がある場合（原言語内書き替え）と、②別の原言語表現はないが、部分的に対応する目的言語の表現がある場合（疑似的原言語への書き替え）の2つに分け、合計6種類の自動書き替え項目を実現した。

新聞記事リード文を利用した翻訳実験結果によれば、書き替えルールの適用された箇所は102文（32記事）のうち44文（43%）で、延べ52箇所であった。そのうち訳文品質向上効果のあった文は33文（75%）である。また、書き替えルールが適用された文では構文解析による構造多義の数が1文平均5.39から1.31まで減少した。このように実験結果によると、本方式は訳文品質向上効果のみならず構造多義解消効果も大きいことがわかった。

なお、本方式を日英翻訳システムに組み込むと、既存の翻訳機能をそのまま利用して、翻訳困難であった文が翻訳できるようになる利点があり、今後の訳文品質向上策として有望であると判断される。

今後は、本稿で述べた日本文書き替えの考え方を拡張して、適用領域を広げていく方針である。具体的には、日本文書き替え処理を適用するタイミングを、本論文のように係り受け解析直後ではなく、形態素解析の直後に行なうことにより、形態素解析の解析誤りを回復することが可能である。また、本論文で示した疑似的日本語の表現を目的言語の表現に接近させることにより、新たな翻訳のパスを作ることができる。これらについても検討を開始しており、その結果については別の機会に報告する予定である。

<謝辞>

本研究を進めるに当たりご討論頂いた当研究所の横尾昭男氏を始めとする研究グループのメンバー各位と新潟大学の宮崎正弘教授、並びに、書き替え処理の実現とルールの作成にご協力頂いた小見佳恵氏、上田洋美氏を始めとするNTTアドバンステクノロジーの各位に感謝する。

<参考文献>

- [1] J. Carbonell, et. al.: JTEC Panel Report on "Machine Translation in Japan", Coordinated by Loyola College in Maryland, Jan '92
- [2] TMI-92 Proceedings of the Conference, Montreal, Canada, Jun 25-27 '92
- [3] M. Rimon, M. McCord, U. Schwall & P. Martinez: Advances in Machine Translation Research in IBM, Proceedings of MT SUMMIT III, pp.11-18, Jul 1-4 '91
- [4] Proceedings of the COLING '92 France, JUL 23-28 '92
- [5] 池原, 白井: 日英機械翻訳機能試験項目の体系化, 信学技報NLC90-43, pp.17-24, DEC 21 '90
- [6] S. Ikehara: Criteria for Evaluating the Linguistic Quality of Japanese to English MT System, MT Evaluation Workshop, NOV 2 '92 (San Diego)
- [7] 池原, 宮崎, 白井, 林: 言語における話者の認識と多段翻訳方式, 情処論Vol.28, No.12, pp.1269-1279, DEC '87
- [8] S. Ikehara: Multi-Level Machine Translation Method, Future Computer Systems, Vol.2, No.3, pp.261-274, '89
- [9] S. C. Chen, J. N. Wang, J. S. Chang & K. Y. Su: ArchTran: A Corpus-based Statistics-oriented English Chinese Machine Translation System, Proceedings of MT SUMMIT III, pp.11-18, JUL 1-4 '91
- [10] S. Nierenburg: KBMT-89-A Knowledge based MT Project at Carnegie Mellon University, MT SUMMIT II, pp.141-147, AUG 16-18 '89
- [11] 池原, 横尾, 宮崎: 日英機械翻訳のための意味解析辞書, 信学技報NLC91-19, pp.23-30, JUL 19 '91
- [12] O. Furuse & H. Iida: Cooperation between Transfer and Analysis in Example-Based Framework, COLING 92, JUL '92
- [13] M. Nagao: Some Rationales and Methodologies for Example-based Approach, Proceedings of Workshop on Future Generation Natural Language Processing, UMIST, Manchester, JUL 30-31 '92
- [14] M. Nagao: Proposal of Restricted Language, Natural Language Processing Technologies Symposium, Information Processing Society of Japan, '85
- [15] M. Nagao, H. Tanaka & J. Tsujii: Computer Aided Systems for Pre-editing of Sentences Based on Restricted Grammar, IPSJ SIG Report, JUL '84
- [16] 白井: 日本文自動書き替えによる構文多義の解消, 41情処全大4S-6, SEP 4-6 '90
- [17] 池原, 安田, 島崎, 高木: 日本文訂正支援システム (REVERSE), 研究実用化報告Vol.36, No.9, pp.1159-1167, SEP '87
- [18] H. Nakaiwa & S. Ikehara: Zero Pronoun Resolution in a Japanese to English Machine Translation System using Verbal Semantic Attributes, Proceedings of ANLP, MAR 29 '92
- [19] Report by the Automatic Language Processing Advisory Committee, JUN '65

付表 原文自動書き替え方式による訳文の変化

書き替え のタイプ	番 号	日 本 語 原 文	書き替え部分 と書き替え方	自動書き替えなし翻訳結果	自動書き替え後の翻訳結果
日本語内 書き替え		二階にショールーム、 三階に商談室、会議室、 セミナールームを開設 し、事務室は四階以上 になる。	〔縮約展開型〕 書き替え <u>ショールームを開</u> <u>設し、</u>	C. Itoh Techno-Science Corp. will set up a conference room, a meeting room and a seminar room in the third floor to a show room and the third floor and an office will reach the fourth and higher floors. 《得点 = 3》	C. Itoh Techno-Science Corp. will set up a show room in the second floor and will set up a conference room, a meeting room and a seminar room in the third floor and an office will reach the fourth and higher floors. 《得点 = 6》
		二機種合わせて月四百 台生産する。	〔構文組み替え〕 書き替え <u>二機種の合計月産</u> <u>は四百台だ。</u>	It produces <i>Midori Denki</i> 合わせて in 2 models in 400 units per month. 《得点 = 2》	The monthly output of 2 models is 400 units. 《得点 = 8》
疑 似 的 日 本 語 表現への 書き替え		ソフト会社、N & C ソ フトウェアはシステム ハウスのユニコムオー トメーションと共同で パソコンを使ったカラ ー印刷システムアトリ エ・ビットを開発した。	〔独立句的表現〕 書き替え “と共同で” → jointly with “を使った” → using	N&C Software Corp., a software company, developed Atlier Bit, the system of color printing that used a personal computer by Unicom Automation Corp. and the synergic of a system house. 《得点 = 4》	N&C Software Corp., a software company, developed Atlier Bit, the color printing system using a personal computer, jointly with Unicom Automation Corp. a system house. 《得点 = 9》
		出版取次はもともと利 益率が低いことに加え て、出版物も需要が鈍 化しているため苦しい 経営を余儀なくされて いる。	〔接 続 表 現〕 書き替え “ことに加えて” → not only ~ but also ~	Because it adds a publication agency to that a profit rate is low originally and the demand for publication is slackening, tight management is made to be unavoidable. 《得点 = 4》	Because <u>not only</u> the profit rate of a publication agency is low originally, <u>but also</u> the demand for publication is slackening, tight management is made to be unavoidable. 《得点 = 7》
		富山センターはソフト 開発要員五十人でスタ ート、百五十人に増や す計画。	〔様相時制表現〕 書き替え “計画” → be planning to ~	The Toyama Center is a start in a development staff of 50 and is a plan increased to 150 person. 《得点 = 4》	The Toyama Center starts in a development staff of 50 and <u>is planning to</u> increase the Toyama System Center to 150 people. 《得点 = 7》

____ : 原文自動書き替え処理の着目箇所 ~~~~ : 書き替え箇所 斜体文字 : 文脈処理による前方の文からの補完箇所