言語処理学会 第 4 回年次大会発表論文集, pp.670–672, 1998

英文語順に準拠した日本語生成

佐田 いち子　九津見 毅　日野 ちなみ　関谷 正明
シャープ株式会社 ソフト事業推進センター
{sata, kutsumi, hino, sekitani}@isl.nara.sharp.co.jp

1　はじめに

現状の英日機械翻訳システムでは、英語と日本語のように語順が大きく異なる言語間で、係り受けの複雑な入力文が入力されると、たとえ正しく構文解析しても、論理的に間違ってはいないが読みにくい日本語文が出力されることが多い。

この改善のため、著者らは以前に、英文読解支援に使われることを想定して、原文を適当な単位で切り分けたうえでその訳を原文に添付して表示する翻訳システムの開発 [1] について報告したが、その後、開発を継続していく上で、翻訳システムの利用形態に拘らず読みやすい訳文を生成することを目指してきた。

原文の英語の語順に近い訳文を生成することで、原文の持つ意味伝達構造を尊重し、スムーズに頭に入る日本語文を生成する [2] という点では、原文を書き換え規則によって適切に分割してから翻訳するといったアプローチ [3] なども行われている。

このたび著者らの採用した具体的方法は、翻訳過程の最終段階である、日本語生成部について、改善効果の大きいポイントに絞って対応することで訳文の改善を図ることである。以下、図１に示す、著者らが採用している翻訳エンジンに適用した例について報告する。

2　間接話法の文の生成の改善

第一の主な対応点は、“The Newspaper said that Clinton did it.” などのような間接話法の文に対し、「新聞によれば、クリントンは、それをした。」といった、報道・レポート文等で用いられる決まり文句を導入することで、読み下しのしやすい訳文を生成することである。

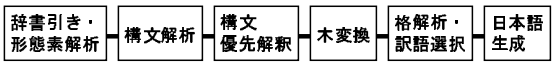

図 1: 翻訳エンジンのブロック図

ID	書き換えパターン	備考
1	(S) によれば (O)	
2	(S) の○○では (O)	元の訳語:サ変
3	(S) が○○するのは (O) ということである	
4	(S) は次のように○○する。(O)	
5	(S) は○○する。(O) と。	
6	(S) の○○は (O) ということである	元の訳語:サ変

図 2: 動詞訳語書き換えパターンの例

通常、このためには、“say” などの動詞の辞書の、THAT 節を目的格にとる型の訳語に、そのための新しい訳語を与える必要があるが、このたび著者らの採った方法は、格解析部にて訳語選択を行った際に、所定の条件を満たしていれば、所定のパターンに従って動詞辞書の訳語を書き換えるというものである。訳語書き換えパターンの例は図２に示すようなものであるが、動詞辞書にもともと記述されている訳語に対して、全く新しい訳語に置き換えたり、元の訳語を利用した言い方に替えたり、あるいは元の訳語の語幹部分を名詞として利用した言い方に替えたりすることが可能となる。(図２で、書き換えパターン中の「○○」は、元の訳語が「○○する」である場合にその語幹または全体を利用することを示す) 各パターンごとに固有の適用条件も定めることができる。

この機構の導入により、辞書本体を大きく書き換えずに新たな訳語を簡便に導入することが可能となり、辞書開発にかかる手間も軽減することができる。

なお、直接話法の文については、４節で述べる訳語再配置処理を使用することで、図４に示すような訳質の改善を図っている。

原文:	The newspaper said that Clinton did it.
従来訳:	新聞は、クリントンがそれをすると述べていた。
改善訳:	新聞によれば、クリントンは、それをした。
原文:	We understand that you're our best partner.
従来訳:	我々は、あなたが我々の最も良いパートナーであると理解する。
改善訳:	我々の理解では、あなたは、我々の最も良いパートナーである。
原文:	We assure that this product helps you.
従来訳:	我々は、この製品があなたに役立つことを保証する。
改善訳:	我々が保証するのは、この製品があなたに役立つということである。

図 3: 動詞訳語書き換え機構による間接話法文の改善例

原文:	"This discovery is very exciting," said Federal arbitrator Nevil Bentley.
従来訳:	米国連邦国家の仲裁人 Nevil Bentley は、「この発見は、非常に興味をそそる」と言った。
改善訳:	「この発見は、非常に興味をそそる」と、米国連邦国家の仲裁人 Nevil Bentley は、言った。

図 4: 直接話法文の改善例

3 従属接続詞を持つ文の生成の改善

従属接続詞を介して主節と従属節から成る英語文には、"Although it was not food, he ate it. " のように従属節が主節の前にある場合と、"He ate it, although it was not food. " のように従属節が主節の後にある場合があるが、著者らの従来の翻訳システムではいずれの場合も、従属節を主節の前に訳出していた。

新システムでは、これを原文に合わせる形で、従属節が主節の後にある原文の場合は訳文もそのようにする。

このため、主要な従属接続詞に対して、図５のような新たな訳語を導入した。従来は、原文中の従属節の位置に拘らず図５の「従来訳語」を使用してきたが、新システムでは従属節が後置される場合に「新規訳語」を使用する。この場合、訳文では主節と従属節の間に読点を生成し、2 文に分けて訳出する。

従属接続詞	従来訳語	新規訳語
although	が	(S-CL) のだが
even if	としても	たとえ (S-CL) としても

図 5: 従属接続詞の新規訳語の例

原文:	He ate it, although it was not food.
従来訳:	それが食べ物ではなかったが、彼は、それを食べた。
改善訳:	彼は、それを食べた。それは、食べ物ではなかったのだが。
原文:	I have to go, even if it rains.
従来訳:	雨が降るとしても、私は、行かなければならない。
改善訳:	私は、行かなければならない。たとえ、雨が降るとしても。

図 6: 従属節の訳出改善の例

4 関係節の生成の改善

英語文の関係節は、日本語文へと直訳した際に、語順が逆転する、「返り読み」の問題を最も典型的に持つ箇所である。著者らの従来の翻訳エンジンでは、関係節は原則として、通常の英文和訳的に前置修飾として訳出するが、この点の改善を図った。新システムでは、関係節の訳出法は、

- 原文中の関係節の位置
- 関係節の長さ
- 関係詞の種類
- 制限用法か非制限用法か

などの状況によって異なってくる。訳し方の選択肢としては概ね図７のようになり、条件の組み合わせの例としては図８のようになるが、実際には関係節を括弧に入れて訳出する際の生成位置の決定の調整や、補助的な訳語の生成に関し、更に詳細な判断が必要となってくる。

この結果として修正された訳文を図９に示す。

1.	前置修飾として訳す
2.	括弧に入れ、後置する
3.	別文とする
s.	関係節に「その」を補う

図 7: 関係節の訳出法

			訳出法
制限用法	先行詞の修飾が少ない 又は関係節が短い		1.
	上記以外	関係詞が whose	2. s.
		上記以外	2.
非制限用法	関係節の位置が文末且つ 関係詞が who 以外	関係詞が whose	3.
		上記以外	3. s.
	上記以外		2.

図 8: 関係節の訳出規則の例

原文:	An electric plug that will cause little or no arcing when connected or disconnected.
従来訳:	連結している、及び、接続を断たれているとき、電弧を生じることをほとんど引き起こさないであろう電気プラグ。
改善訳:	電気プラグ (連結している、及び、接続を断たれているとき、電弧を生じることをほとんど引き起こさないであろう) 。
原文:	This is Mr. Kirby, whose house faces yours.
従来訳:	これは、家があなたのものに面する Kirby 氏である。
改善訳:	これは、 Kirby 氏である。その家があなたのものに面する。

図 9: 関係節の訳出改善の例

これを実現する手段としては、著者らが以前に開発した [1] 、フレーズ分割の技術を基盤とした、訳語再配置処理を用いる。これは、一旦訳文を生成してから、原文における所定の範囲ごとに訳文単語を配置し直すというものである。ただし、今回は、原文に添えた形で出力するのではなく、通常の翻訳文として利用されることを目的としているので、訳文それだけで、一文あるいは複数の文として不自然でなく成立しているように仕上げる。

5 おわりに

英文和訳特有の読みにくい日本語文について、英日翻訳エンジンの最終段階である日本語生成部での対応を行うことで、相当の訳質改善の効果をあげることを主眼として開発を行ってきた。本稿では便宜上、例文として短い英語文ばかりを掲載したが、節が長く文が複雑になるほど、ここに挙げた新方式による訳文生成の効果は高まる。

日本語生成部は、構文解析部等と比較して、開発・改良の小回りが効く箇所である。そしてまた、この方向での改善の余地は多く残されており、今後の開発においても成果が期待できる。

6 謝辞

本研究にあたり日頃より御指導頂いているソフト事業推進センターの仲川與万所長、福持陽士副参事、並びに有益な示唆を頂いている英日翻訳関連グループの諸氏に深謝する。

参考文献

[1] 九津見, 奥西, 佐田:“英日機械翻訳における速読支援のための日本語生成”. 言語処理学会第 2 回年次大会発表論文集 (1996),pp.221-224.

[2] 今井:“英語「超」速読法”. 時事英語研究 1996 年 5 月号, 研究社,pp.9-23

[3] 佐良木, 黒野:“構文の節結合強度に応じた書換え規則によるテクスト処理およびテクスト処理専用言語 WRAPL の開発”. 言語処理学会第 2 回年次大会発表論文集 (1996),pp.161-164.