

## 日本語構文解析における曖昧さの解消について

3C-7

鈴木 恵美子<sup>1)</sup> 隅田 英一郎<sup>1)</sup> 綾部 祥子<sup>2)</sup> 島本 幸子<sup>2)</sup>

1) 日本アイ・ビー・エム(株) 東京基礎研究所

2) 東京女子大学 文理学部・日本文学科

## 1. はじめに

我々は従来より、広い言語現象に対応できる日本語解析システムの実現をめざしてきた。そして、機械翻訳用に日本語解析システムを作成すればその解析システムは、他のシステムにも利用できる(詳細は文献1参照)と考えて研究を続けてきた。本稿ではその解析システムにおける重要な問題である曖昧さの解消について述べる。

## 2. 曖昧さ

## 2.1. 概要

日本語の構文解析で問題とされる曖昧さには各種のものがあるが、その中の代表的なものに「で」の曖昧さがある。例えば、

・本で殴った。

という場合には「で」は格助詞である。しかし、

・これは本で、あれはペンだ。

という場合の「で」は述態辞(助動詞)となる。このような「で」を「で」の直前にくる体言のフィーチャーで見分ける方法もあるであろうが、「道で(会う)」「収賄の容疑で(あげられる)」「ストライキで(足を奪われる)」等々、フィーチャーも場所、時、主格、状況、理由、道具、など多岐にわたり、その対応関係は複雑とみられる。

本システムでは、1) 浅い知識のみを用いてできる限り曖昧さの数を減し、2) 残った解釈についても優先度を付与することによってより自然な解釈が選択できるようにすることをめざしている(文献1)。そのため格文法でなく結合価文法を導入し、文解析のもっともらしきの目安としている。今回、1)の実現のため、大量のデータをもとに構文解析プログラムに制限を加えることを試みた。

## 2.2. 実現方法

構文解析の対象を限定するわけではないが、入手できた最も多量の日本語が新聞記事であったので、そこから曖昧な語句を前後の文脈と共に切り出し、人手で分類した。

今回は、「で」と「対(する)」について行った。「で」は前述のように、格助詞と述態辞の曖昧さを持ち、「対(する)」は、まとまって1つの格助詞相当となる場合と「対」が本動詞として使われる場合の2通りの解釈ができる。

調査は新聞記事1カ月分について行った。その結果、「で」はごく一部の例外を除く、以下の場合が格助詞、それ以外の場合は述態辞と解釈してよいことが判明した。

「で」が格助詞となる場合

・で+動詞<sup>a)</sup>

・で+助詞<sup>b)</sup>

\*a) 例外は以下のような場合

〈例〉「大変な努力が必要で疲れてしまう」

\*b) 「ではない」など、「ある」または「ある」の変形が後にくる時は述態辞である

〈例〉「初めからあったわけではない」

「対(する)」は、変形(「対して」、「対しても」等)を含めた884例の出現のすべてが格助詞相当連語と認められたので、「対」に動詞の解釈をしないこととする。

## 3. まとめ

以上述べたような方法で、「で」「対(する)」という2種類の語句の解釈における曖昧さを解消することを試みた。今後は更に他の曖昧な語句(「における」「に関する」「として」等の格助詞相当となりうる連語、「て+動詞」等)についても調査、検討の予定である。

## [参考文献]

- (1) 隅田、丸山、1987、「拡張CFGを用いた日本語構文解析」、情報処理学会WGNL63-3。
- (2) 丸山直子、1986、「日本語文法—水谷文法と結合価文法を基盤にして—」、情報処理学会第33回大会予稿集。
- (3) 賀来直子、1985、「日本語のアスペクト」、情報処理学会第31回大会予稿集。

Disambiguation in Japanese Syntactic Analysis

Emiko Suzuki<sup>1)</sup>, Eiichiro Sumita<sup>1)</sup>, Shoko Ayabe<sup>2)</sup>, Yukiko Shimamoto<sup>2)</sup>

1) IBM Research, Tokyo Research Laboratory

2) Department of Japanese Literature, Tokyo Woman's Christian University