

# 1C-7 接続助詞に注目した文分割の方式

阿部 ひろみ 奥西 稔幸 三吉 秀夫 小淵 保司

シャープ株式会社

## 1. はじめに

自然言語には多くの曖昧性が存在しており、応用システム開発の上で大きな障害となっている。中でも長文は多くの曖昧性を生じさせる原因の1つである。そこで吉田らが提唱する規格化日本語文法<sup>(1)</sup>のように曖昧性を排除する目的で長文を分割する方法がいくつか提案されている。<sup>(2)</sup>

我々は規格化日本語文法を基礎に、文分割プログラムを開発し、入力解析・修正支援システム<sup>(3)</sup>の1モジュールとして組み込む実験を行った。

本論文では、接続助詞的表現に注目した文の分割方法について述べる。

## 2. 接続助詞による文分割

規格化日本語文法では、接続助詞的表現について文を分割する方法の違いから3群に分類している。本方式では約50個の接続助詞的表現を表1のように分類し直している。

A群	接続助詞的表現の直後で分割可能な表現
B群	接続助詞的表現の直後で分割したあと、書き換えが必要な表現
C群	接続助詞的表現の直後で分割できない表現

表1 分割の分類

本方式では表1の中でA群とB群に属する接続助詞的表現を分割するために表2のような規則を用いる。

接続助詞的表現	意味	接続詞
ので	原因理由+順接確定	したがって
けれども	逆接確定	しかし
てから	列叙接続+継起	それから

表2 分割規則の例(+は付加の意味があることを示す)

文中に接続助詞的表現が複数存在する場合の分割点を決定するため、次の分割点優先規則を用いる。分割点優先規則は、分割可能な接続助詞13種類に関して合計約170種類設けている。表3と表4は、それぞれ1番目、2番目の接続助詞的表現で優先分割する組み合わせの例である。

第1接続助詞的表現	第2接続助詞的表現
ので	あまり、ときに、 のではなく、だけでなく、 かわりに、につれて、 てから、た上で、と、 というように、
ときに	あまり、のではなく、だけ でなく、かわりに、につれ て、と、

表3 1番目を優先分割する例

第1接続助詞的表現	第2接続助詞的表現
てから	ので、けれども、のではな く、かわりに、た上で、と、 というように、
た上で	ので、た結果、けれども、 ときに、のではなく、だけ でなく、につれて、てから、 と、

表4 2番目を優先分割する例

## 3. 分割処理の手順

本方式は形態素解析の結果を入力文として受け取って、以下のような手順で分割処理を行う。

### ①分割条件による判断

あらかじめ指定した分割条件を満たしている場合分割処理を行う。

分割条件とは、文字数、単語数、文節数などである。

### ②接続助詞的表現の代表表現への変換

次に文中の接続助詞的表現を規格化日本語文法で規定している代表表現へ変換し、変換した代表表現が分割可能か否かを調べる。

### ③接続助詞的表現の意味による判断

②のあと、分割可能な接続助詞的表現を更に意味別に分類する。

### ④分割規則の適用

③で意味別に分類された後、それぞれの分割規則に則り、分割処理を行う。

文中に接続助詞的表現が複数存在する場合には、分割優先規則に沿って分割処理を行う。

### ⑤時制変換規則の適用

④の分割規則に沿って文を分割した時、意味的に

不自然さが生じたり、文の意味が分割前と変わってしまう場合がある。これは時制、アスペクト、文体（です／である）等の文末表現の問題であるが、今回は時制についての変換規則のみを構築した。分割後の時制は、分割前の時制と文中にある接続助詞の意味により決定されることから、各文の分割前の時制を次のように定めた。

接続助詞を含む文節（第1文）と最終文節（第2文）において

- ・過去の助動詞「た」が文末に存在する場合は過去とする。
- ・過去の助動詞「た」が文末に存在しない場合は現在とする。

接続助詞的表現の意味	分割前時制		分割後時制	
	第1文	第2文	第1文	第2文
原因理由+順接確定	現在	過去	過去	過去
	過去	過去	過去	過去
	過去	現在	過去	過去
	過去	現在	過去	現在
逆接確定	過去	現在	過去	現在
	過去	現在	過去	現在

表5 時制変換規則の例

表5のように、接続助詞的表現の意味により時制の組み合わせが異なるので、変換規則を個別に設けている。

4. 分割の例

本項では本方式による分割例をいくつか示す。

例1：文中に接続助詞的表現が1つしかない場合の分割の例

接続助詞的表現の意味	分割前の接続助詞的表現	分割後の接続助詞
原因・理由+順接確定	ので	したがって

(分割前) ドアがどうしても開かなかったので、わたしたちは家に入れなかった。



(分割後) ドアがどうしても開かなかった。したがって、わたしたちは家に入れなかった。

例2：接続助詞的表現の直後で分割したあと、書き換えが必要な例

接続助詞的表現の意味	分割前の接続助詞的表現	分割後の書き換え
原因・理由+過度	あまり	非常に～。したがって…。

(分割前) 気が転倒していたあまり、思いもよらぬことをしてしまった。



(分割後) 非常に気が転倒していた。したがって思いもよらぬことをしてしまった。

例3：1番目の接続助詞的表現で分割する例

接続助詞的表現の組み合わせ	結合関係	分割の優先
～た結果～と～。	<	1番目の接続助詞的表現

注) <記号は2番目の接続助詞的表現の方が結合関係の強い事を表す

(分割前) 彼女は医者と相談した結果両親が迎えに来ると病院から家へ戻った。



(分割後) 彼女は医者と相談した。その結果両親が迎えに来ると病院から家へ戻った。

例4：2番目の接続助詞的表現で分割する例

接続助詞的表現の組み合わせ	結合関係	分割の優先
～たうえで～ので～。	>	2番目の接続助詞的表現

注) >記号は1番目の接続助詞的表現の方が結合関係の強い事を表す。

(分割前) よく勉強したうえでテストを受けたので志望校に合格した。



(分割後) よく勉強したうえでテストを受けた。したがって志望校に合格した。

例5：時制変換規則を適用する例

(分割前) 大きな声がするので、目が覚めた。

現在

過去



(分割後) 大きな声がした。したがって目が覚めた。

過去

過去

5. おわりに

長文の曖昧性を軽減する一つ的手段として、文中の接続助詞的表現に注目し、その意味的な役割から文を分割する方法を検討し、プログラムの開発を行った。

その結果、文中に分割可能な接続助詞的表現が存在する場合は、接続助詞的表現の数に関係なく文の意味を保存したままで文を分割することが可能になった。

最後に、今後の課題としては文末表現の処理精度を高めること、そして現方式で処理できない表現の分割方法を検討することなどを考えている。

なお、本研究は第五世代コンピュータプロジェクトの一環としてI C O Tから委託を受けて行っているものである。

◇謝辞

日頃ご指導頂いたI C O T第6研究室の田中室長、並びに有益なご意見を頂いたシャープ(株)情報システム研究所第1開発室の濱田、秋山両氏に感謝致します。

◇文献

1. 吉田 将他, 規格化日本語による文書の作成に関する研究, (1987)
2. 武石 英二他, 複文における分割点の特定について, 第4回人工知能学会全国大会論文集, (1990)
3. 小淵 保司他, 規格化日本文作成支援の一方式, 第36回情報処理学会全国大会論文集, (1988)