

## 日本語対話文における省略語の補完

3B-5

柳原正秀 飯島正 原田賢一

慶應義塾大学理工学部

## 1 はじめに

日本語の自然言語処理が困難である理由の一つに、省略現象があげられる。特に日本語の場合、主語や動詞（述部）などの解析上必要である語が省略される場合が多くあるため、省略語を補完するためのさまざまな研究が行なわれている。しかし対話文においては、話し手や聞き手が存在するために、これらの方法では様々な問題点が生じてしまう。本研究では、このような問題を解決できるように対話文における省略語補完法を新たに設計し、実際にシステムを作成して評価を行なう。

## 2 従来の省略補完法と問題点

## 2.1 センタリング

ここでは、本研究でも利用しているセンタリングという省略補完法について説明する[2]。これは、直前の文からその文の省略語の照応の対象になり得る語のリストを作り、その中から最も照応の対象になりやすいものを見つける方法である。語のリストは優先順位がつけられており、日本語においてはそれは次のようになっている[3]。

トピック>視点>主語>目的語>その他

日本語においては、トピックとは「は」格、主語は「が」格、目的語は「に」格「を」格である。視点とは授受表現がある場合に決定されるものであり、話し手が自己と最も同一視可できるものと定義されている[1]。ここで、省略語の補完は前文と文法属性が一致しなければならないという制約がある。この方法だと、複数の省略代名詞があってもそれぞれを補完することができる。

## 2.2 センタリングによる例

ここで、実際のセンタリングの例を挙げて見る。

例文1 “太郎が次郎と話していた。” 太郎>次郎  
 “ $\phi_{sub}$  三郎を見かけた。”(太郎)>三郎  
 “ $\phi_{sub}$  三郎を呼び止めた。”(太郎)>三郎

しかし、対話文ではうまくいかないことがある。

例文2 A: “次郎と喧嘩したらしいね。” 次郎  
 B: “してないよ。”(次郎)  
 A: “太郎がだよ。”(太郎)  
 B: “それなら知ってる。”(太郎)

ここでは、最後のBの発話の補完がうまくいかない。なぜならここには、「わたし(話し手)」が補完されなければならないからである。

## 2.3 問題点

ここで、従来の方法に対話文の省略語補完に用いたときの問題点をまとめてみる。

1. 「私(話し手)」「あなた(聞き手)」のように文中に現われない語が省略されていると補完できない
2. 聞き手と話し手の間に誤解があると補完できない
3. 動詞(述部)の省略を補完できない

## 3 対話文省略補完法

## 3.1 対象範囲

ここでは、対象範囲を次のような対話文に限定する。

1. A、B、二人の対話文を扱う
2. 実際の発話文独特の現象(言い直し、言い間違いなど)は扱わない
3. 単文のみ扱う

## 3.2 補完方法

(1) 聞き手と話し手を補完できるようにする。

ここでは、終助詞・助動詞・動詞が対話文においてどのような役割を持つかを検討し、それにより文型と動詞の分類を行なうことにする。

動詞の分類は「知っている」などのような人の心的状態を表す語と、それ以外の語のわける。文型の分類については情報のなわばり[4]を参考に「一般形」「ね形」「らしい形」「らしいね形」四つの文型にわける。

この分類により、拡張センタリングリストを表1のようにまとめる。これを用いることにより、聞き手や話し手を補完することができる。

ここで、Cはセンタリングリスト、話は話し手、聞は聞き手、他は第三者を表す。

文型 \ 動詞	(1) 一般的な語	(2) 心的状態を表す語
一般形	C > 話	話
疑問・ね形	C > 聞	聞
らしい形	C > 話, 他	C
らしいね形	C > 聞, 他	C > 聞

表1: 拡張センタリングリスト

(2) 動詞を補完できるようにする。

動詞が省略されていたら前文の動詞で補完する。そのとき前文の格の情報も補完する。格の情報が重なったら、新しい情報を優先する。

(3) 話し手と聞き手の誤解があっても省略補完できるようにする。

一つの文に対し話し手と聞き手のそれぞれに解析結果とセンタリングリストを持たせる。

補完候補が一つに定まっていなるときには、次の文と動詞が同じである場合に限り、発話者のみ次の文から逆に補完を行う。このようにすることにより、話し手と聞き手の誤解を扱うことができるため、正しく省略補完を行うことができる。

### 3.3 解析例

前記の拡張センタリングリストを用いて作成したプロトタイプにより、実際に例2の文を解析したときの結果を示す。

実際の解析経過はここでは省くが、結果として次のように補完されることになる。

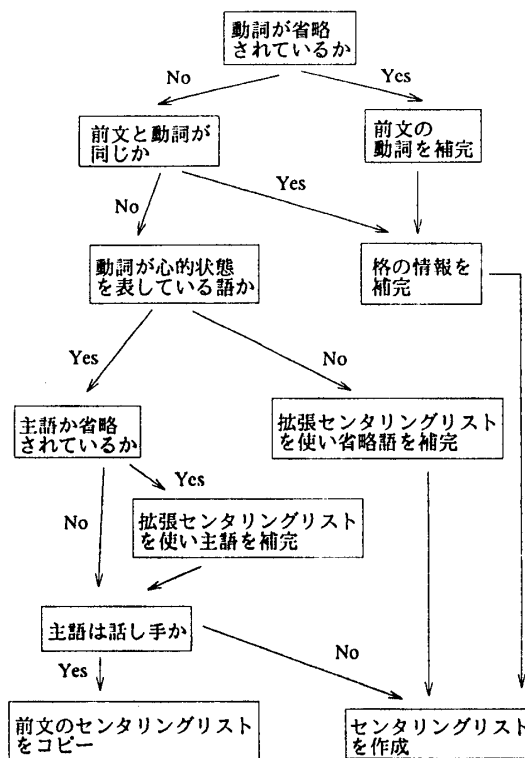
- A: “(太郎が) 次郎と喧嘩したらしいね。”  
 B: “((B、他人) が次郎と喧嘩) してないよ。”  
 A: “太郎が(次郎と喧嘩した) だよ。”  
 B: “それなら(Bが) 知ってる。”

### 3.4 補完の全体図

本研究において省略補完は、構文解析を行いながら行なうことにする。構文解析は格文法を用いる。

入力形態素解析された対話文列、出力は省略語が補完された構文解析結果(格構造)である。

補完方法の流れを示す。



## 4 おわりに

対話文独特の現象を取り上げ検討することにより、対話文における省略語の補完法を提案し、プロトタイプを作成した。ただし、現在のところは文型は四つ、動詞は二つにしか分類していないため、まだまだ改良の余地が残っている。

今後は、終助詞・助動詞・動詞の、対話文における役割をさらに検討し、文型と動詞の分類を細かく行なう必要がある。これにより新たな拡張センタリングリストを作成し、また、実際にシステムを作成し評価を行なわなければならない。また、将来的には複文や重文にも対応できるようにしたいが、そのためには接続詞などの情報も活用できると考えられる。

### 参考文献

- [1] 久野 すすむ: “談話の文法”, 大修館書店.
- [2] Susan E. Brennan, Marilyn W. Friedman, Carl J. Pollard: “A Centering Approach to Pronouns”. ACL Proc. 25th Annual Meeting (1988), pp.155-162.
- [3] Hiroshi Nakagawa: “Zero Pronouns as Experiencer in Japanese Discourse”, COLING 1992 (1992), pp.324-330.
- [4] 神尾 昭雄: “情報のなわ張り理論”, 大修館書店.