

# 日本語長文における読点の役割分析

藪 正樹 宮崎 正弘

新潟大学大学院自然科学研究科

## 1 はじめに

複数の文節からなる長い日本語文の係り受けを解析する場合、通常、読点を伴う文節は独立性が強い、すなわち、より離れた文節に係る可能性が高いと考えられる。しかしながら、すべての読点がそのような役割を果たしているわけではなく、並列構造を強調するために読点が用いられている場合も存在する。本稿では、日本経済新聞の記事データ [1] から日本語文における読点の用い方を分析し、表層的な情報から読点の役割を分類する手がかりについて述べる。

## 2 読点の用い方

公にされている句読法で一応の基準とみられる総理庁・文部省「くぎり符号の用い方」(「公文用語の手びき(改訂版)」所収, 昭和24年3月)では、読点の打ち方について、次のような原則を立てている [2]。

- イ 叙述の主題を示す「は」「も」などのあと。
- ロ 対等に並列する同種類の語句の間。
- ハ 文の初めにおく接続詞及び副詞のあと。
- ニ 叙述に対して限定を加え条件をあげる語句のあと。

本稿では上記ロに相当する読点の用い方に注目し、読点前後の品詞および読点と並列構造との関わりについて調査を行う。

## 3 品詞分析

日本経済新聞の記事データ11,741文(読点数17,267)を形態素解析し、読点の直前および直後に出現する品詞を調査した(図1および図2参照)。ここでは、三浦文法[3]に基づく品詞分類[4]を行う形態素解析システムを使用している。

読点直後の品詞としては、名詞が圧倒的に多い。読点直前の品詞として頻度の高い助詞の出現状況を図3に示す。体言およ

び用言については次節以降で検討を加える。

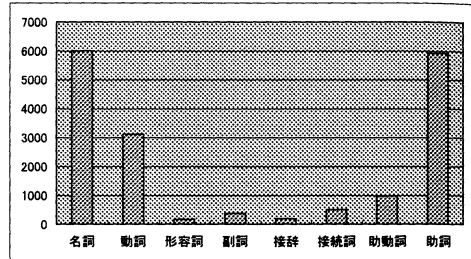


図1 読点直前の品詞

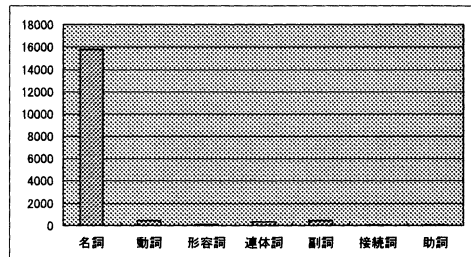


図2 読点直後の品詞

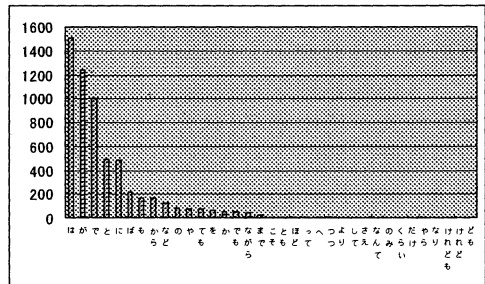


図3 読点直前の助詞

## 4 体言の並列構造

読点直前の名詞が動作名詞、副詞型名詞、および接続属性をもつ形式名詞の場合は、それぞれ用言の連用中止形、副詞、および接続助詞相当の役割をもつ。

動作名詞の例：

南北ウイングとターミナルビル本館を結ぶウイングシャトルにも乗車、車窓からの空港島の眺めを楽しみました。

副詞型名詞の例：

大手人材派遣会社が七月末、大阪市内で開催した就職説明会は当初予定を上回る参加者を集めた。

接続属性をもつ形式名詞の例：

海外への直接投資が拡大した結果、国内の生産活動や雇用が失われることを指す。

読点直前の名詞のうち、動作名詞、副詞型名詞、形式名詞の占める割合は、半数程度である（図4参照）。他の名詞に読点が続く場合、ほとんどは並列構造となるが、一部に同格となるものもある。下記の例では、比喩的な表現を使用しており同格であることを判断するには意味処理が必要となる。

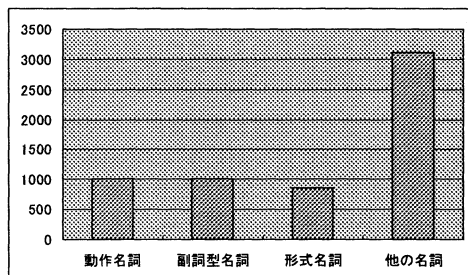


図4 読点直前の名詞

同格となる例：

また少雨傾向を反映して、近畿の水がめ、琵琶湖の水位は三十日に史上最低水位を突破する可能性が強くなった。

さらに、読点後ろの名詞句に「など」が続く場合は、ほぼすべてが体言の並列構造となる。ただし並列内容が省略される場合があり、シソーラスによる類似度などを利用した処理が必要となる。

並列内容が省略される例：

景品には海外、国内の旅行券などを用意する予定だ。

## 5 用言の並列構造

読点の直前にくる用言を活用形で集計すると、大半が連用中止形であった（図5参照）。このうち、読点が用言の並列要素間の区切りを表わすのは、次のような場合である。

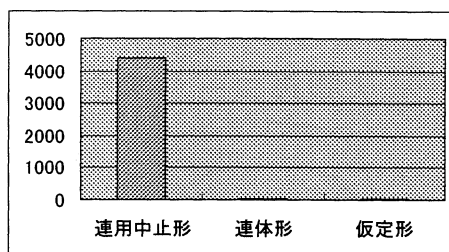


図5 読点直前の活用形

連体形+読点+終止形+「など」：

もっとカネを落としてもらうためにホテルを建てる、店を広げるなど、規模拡大の誘惑はいつも目の前にぶら下がっている。

連用形+読点+連体形+名詞：

だれからともなく握手を求め、互いの労をねぎらう光景があちこちで展開された。

連用形+読点+終止形+「など」：

有機野菜を販売したり、観光客向けにTシャツやステッカーなどのグッズを売り出すなど、あの手この手の資金調達作戦を展開中だ。

連用形+読点+終止形+「と」：

今はアジア自身の市場規模が拡大し、アジアから欧米への輸出も増えているという好循環にあります。

連用形+読点+連体形+名詞+「など」

菱食が持つ販売実績データを基にメーカーが計画生産したり、工場から小売店の店頭までできるだけ短い時間で商品を届ける仕組みを作ることなどを検討する。

「用言+読点」は文節の係り受け解析においてより速くに係ると解釈されるが、読点直前が「たり」となる場合および読点後の用言に「など」が続く場合、その読点が用言の並列構造を表わしていることを示す重要な手がかりとなる。

## 6 おわりに

表層的な情報から日本語長文の並列構造を抽出する手がかりを得るために、日本経済新聞の記事データにおける読点の使われ方を調査した。さらに調査を進め、読点を含む並列構造のパターンを分析していく。

### 参考文献

- [1] 日本経済新聞社：日経 全文記事データベース 日本経済新聞 CD-ROM版 1994年度版 (1995)
- [2] 森岡、永野、宮地：講座正しい日本語 第3巻 表記編、明治書院、pp187-189 (1971)
- [3] 三浦つとむ：日本語はどういう言語か、講談社 (1976)
- [4] 宮崎、白井、池原：言語過程説に基づく日本語品詞の体系化とその効用、自然言語処理、Vol.2, No.3, pp3-25 (1995)