

テレビニュース原稿の特徴分析

西脇正通 浦谷則好 畑田のぶ子

NHK 放送技術研究所

1. はじめに

NHKのニュース文は新聞と比較しても長いことが知られている [1]。これは、書き言葉でも話し言葉でもなく、書かれた原稿をアナウンサーが読み、それを視聴者が聞くことを前提にして作られているためと思われる。また構造も、限られた文字数で出来るだけ多くの情報を伝えるために、修飾や引用を交えて複雑なものになりやすい。このような長く複雑な文を対象にして言語処理をする場合、構文解析精度がえられないという問題が起こる。構文解析精度を上げるためには文体の特徴を把握し、単文への分割や文節間の関係の抽出を考える必要がある。このためNHKのニュース文固有の特徴を、複文の出現パターンに着目して分析した。この結果について報告する。

2. 節の分類

NHKのテレビニュースは、通常、記者がワープロで作成した原稿をアナウンサーが読むという形態をとっている。今回は、この原稿のうち、「政治」「経済」「社会」「国際」の各分野の原稿から、30文ずつ計120文を分析した。

複文の出現パターンの分類に関しては、節の概念を利用した。節とは、1つの述語を含み、意味的に1つのまとまりを持つものである[2]。節の種類は、4つに分類されるものとし、以下の順で優先順位をつけた。

(1)副詞節

述語の修飾や文全体の修飾をする節。原因、手段、付帯状況などを表す。

例：彼は、日本に来るため、飛行機に乗った。

(2)並列節

主節との関係が対等で独立している節。

例：式典が行われ、社長が挨拶しました。

今回は、「平均株価は、1万6千5百円で取引を終え、1万6千円台を回復しました。」などの、主節を修飾する並列的なものも、副詞節として分類した。

(3)名詞節

引用節などを含み、体言の性格を帯びる節。

例：私は、彼が言ったことを書きとめた。

例：彼は、「私が言った」と書きとめた。

(4)連体節

体言を修飾する節。

例：彼は、私が作った料理を食べた。

3. ニュース文の分析方法

節の分類では外側にある（木構造上で上位にある）節を優先したが、それだけではパターン化は容易でない。そこで、前節に示した節の優先順位を考慮してパターン化を行なった。ここで、連体節は出現する体言や名詞節の数だけ出現する可能性があり、パターン化を困難にするため、順位を下げた。

また、連体節の修飾先や、名詞節がとる格で、特に重要な格であると考えられる「が格」「を格」「に格」「と格」以外の格は、パターン化する際に関係する節と共に無視している。

無視した節の中に別の節が含まれることもあるが、パターン化の困難さを考慮して今回は無視した。連体節の中に、別の節が含まれる場合も無視した。

節が、入れ子構造になっているものについては、まず、2層目までを分析の対象にした。

4. 分析結果

以上の規則をもとに分類すれば、第1層目では次の4つのパターンとなる。

文型 1. 連体節以外の節が無い文

「が格」「を格」「に格」「と格」をとる体言の前に連体節が来る形をとり、16通りのパターンが考えられる。まったく連体節が無い「単文」もこのパターンに含む。

文型 2. 名詞節を含む文

副詞節、並列節が無く、名詞節を含むもので、連体節も含む場合がある。

文型 3. 並列節を含む文

副詞節は含まないが、その他の節は含む場合がある。

文型 4. 副詞節を含む文

副詞節以外にも、その他の節を含む場合がある。

4. 1. 文型 1

120文中11文が相当した。そのうち4文が、連体節を含まない単文であった。その他のうち4文は、「を格」の体言を修飾する連体節が1つある複文で、例えば次のようなものである。

- 政府・自民党は、
- 金融機関が破綻した（連体節）場合に、借り手を保護するための（連体節）いわゆる「ブリッジバンク」について、
- 破綻直後は金融監督庁が業務を管理した上で、（副詞節）預金保険機構が設立する（連体節）統括機関のもとで公的なブリッジバンクに移行するという（連体節）二段階の方式をとるなどとした（連体節）
- 最終案を
- 固めました。

4文中3文までが上の例のように、政策や発表の内容を伝えるために、このパターンを使用したものであった。

連体節が修飾する格のパターンとその出現数を表1に挙げる。

表1 文型1のパターンと出現数

| 個数 (120文中) | が格 | を格 | に格 | と格 |
|------------|----|----|----|----|
| 4 | × | × | × | × |
| 2 | ○ | × | × | × |
| 4 | × | ○ | × | × |
| 1 | × | × | ○ | × |

今回の分析では、格助詞相当句による格を認めなかったこともあり、第1層目で複数の格が出現するパターンは無かった。

4. 2. 文型 2

120文中20文が相当した。次に挙げる4つのパターンとなる。下線部分が名詞節となっている。

(i) Aが～と（引用）、Bする。（13文）

- 連合会は、
- 「・・・厳しい状態が続きそうだ」と（引用）
- 話しています。

(ii) Aが～ことを、Bする。（3文）

- オルブライト国務長官は、・・・
- ・・・の三か国が協力していくことを（名詞節）
- 確認するものと見られます。

(iii) Aが～のか、Bする。（3文）

- 「環境ホルモン」がどの程度蓄積しているのか（名詞節）
- 厚生省は
- 今年度から本格的な調査を進める事になりました。

(iv) ～がBする。（1文）

- ソニーが始めるのは（名詞節）、・・・
- 自動車損害保険です。

さらに、(i)のパターンを分類すると、2つの並列節を含む引用（2文）や、2つの文を含む直接引用（4文）が、合計で6文出現した。このようなパターンは、次のように長くなりやすいが、3つ以上の並列節や文を含むものは無かった。

- 市場関係者は、
- 「市場では、大規模な減税があれば、不況に陥っている日本経済も回復に向かうのではないか」という思惑から円を買う動きが強まっている。
- 株価の値上がりに加えて、アメリカのルービン財務長官が円安に対する懸念を示したと伝えられたことも円高が進む要因になっている」と（引用）
- 話しています。

その他に、副詞節を含む引用も多く、6文出現した。そのうち3文は、節の主語が共通で、「AがBし、Cすると、Dしました。」のようなパターンであった。

- 総理大臣は、
- 江沢民国家主席に対して、
 - 「今年秋の訪日を重視しており（副詞）、楽しみにしている」と（引用）
- 伝えたということです。

4.3. 文型3

並列節を含む文は120文中31文出現したが、並列節が意味的に独立しているため、並列節内のパターンを考慮して文全体をみた場合は目立った規則性がみられない。

第1層目の並列節の数を見ると、31文すべてにおいて、1つの並列節で構成されていた。ただし、その中には主節の中にさらに並列節を含むものが、2文あった。

- インタビューは、上海市内でおおよそ三十分に行われて、（並列）
 - まず、今回の中国訪問の成果を問われたクリントン大統領は、「江沢民主席が首脳会談後の共同記者会見や北京大学での私の演説を中国全土に向けてテレビで生放送することを認めたことにいささか驚いている」と述べ、（主節に含まれる並列）
- 中国側のこうした対応は中国の民主化につながる出来事だったと評価しました。

また、単文の並列節は、62個中14個で23%しかなく、残りは節をその中に含むものであった。

4.4. 文型4

120文中58文出現した。第1層目のパターンは3つに分けられる。

(a) 副詞節が1つ（32文）

- ノーネクタイもその一つで、（並列）
 - 冷房の設定温度を二十八度以上に行っているかわりに（副詞）
- きょうから二ヵ月間の夏の間は軽装でもよいことにしたものです。

(b) 副詞節が次の副詞節を修飾している（15文）

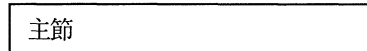
- 今回二日間にわたって行われた協議では、
 - この合意を確認したものの、（副詞）
 - 各国ともこれまでに表明した以上の負担には難色を示したことから、（副詞）
- 結局、費用の分担をめぐる調整はつきませんでした。

(c) 副詞節が並列して主節を修飾する（11文）

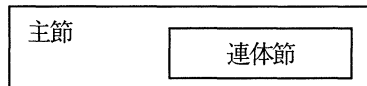
- クリントン大統領は、
 - 中国が貿易上の障壁を緩め、企業や個人に、より自由な経済活動を保証することこそ経済成長の持続には欠かせないとして、（副詞）
 - 世界の金融市場の安定につなげるためにも、（副詞）
- 中国が国際経済の価値観にさらに歩み寄るよう促しました。

パターン(a)の文の副詞節をさらに分類すると、次のようになる。文の構造の概念図を同時に示す。

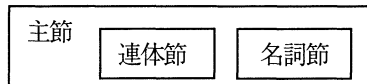
(a-1)単文（16文）



(a-2)連体節のみを含む（10文）

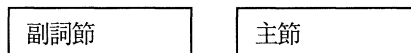


(a-3)連体節と名詞節を含む（6文）

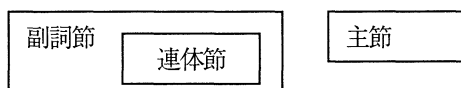


同様にパターン(b)の文の副詞節（合計30個）を分類すると次のようになる。

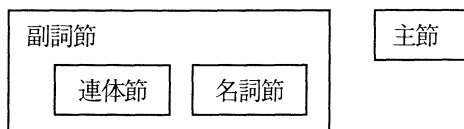
(b-1)単文（15個）



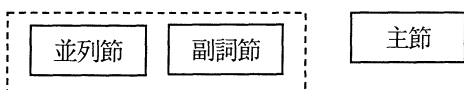
(b-2)連体節のみを含む（6個）



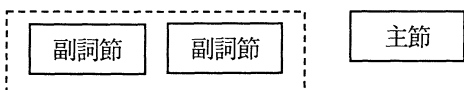
(b-3)連体節と名詞節を含む（1個）



(b-4)並列節を含む（4個）

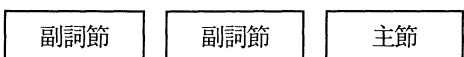


(b-5)副詞節を含む（4個）

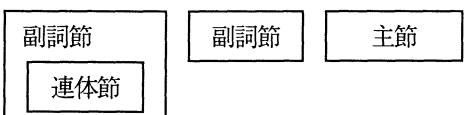


パターン(c)の文で並列した副詞節の数をみると、副詞節が3つある文が1文、2つある文が10文であった。これら、合計23個の副詞節を分類すると次のようになる。

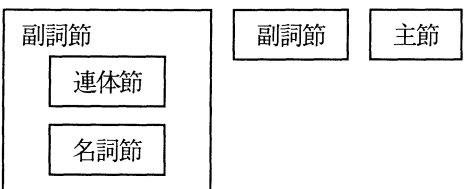
(c-1)単文（11個）



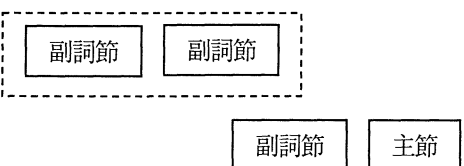
(c-2)連体節のみを含む（2個）



(c-3)連体節と名詞節を含む（6個）



(c-4)副詞節を含む（4個）



このように、副詞節に含まれる節に注目すると、単文である副詞節と単文でない副詞節はどちらも42個で、複雑な構造の副詞節が50%を占めていることがわかる。

4.5. ニュース項目の第1文目のパターン

ニュースの各項目で最初の文となる24文がとるパターンを表2に示す。

第1文目の出現パターンで、文型1、文型2をとるものはそれぞれ1文のみであった。

逆に、24文中の77%が文型4であり、これは、120文中の文型4の出現確率48%に対して、非常に高い割合である。ただし、副詞節が複文となるものは6文（35%）で、120文での副詞節が複文となる割合50%よりも少なかった。

表2 第1文目の出現パターン

| | 政治 | 経済 | 社会 | 国際 | 計 | 全体 |
|-----|----|----|----|----|----|----|
| 文型1 | 1 | | | | 1 | 11 |
| 文型2 | | 1 | | | 1 | 20 |
| 文型3 | 1 | | 1 | 3 | 5 | 31 |
| 文型4 | 4 | 6 | 4 | 3 | 17 | 58 |

5. まとめ

以上のことから、ニュース文には、節を内部に持つ副詞節、並列節が多いことがわかる。また、連体節、名詞節にも節が含まれていることが多く、文を複雑にしている。

逆に、並列構造を持つ場合は、すべてが1つの並列節で構成されていた。

今回は、文が複雑すぎるため上位部分のみしか実施しなかったため、分析が十分ではない。しかし、ニュース文では副詞節をとる文が多く出現すること、複雑な構造が多いことが判明した。連体節や名詞節は、この分析では単純な節に見えるが、格関係や中に含まれる節について、副詞節や並列節と同様に複雑である場合も見うけられた。

今後は、新聞との比較や、パターン分けの手法の精密化などを検討していきたい。

参考文献

- [1] NHK放送文化研究所:放送基本語彙調査, 1989.
- [2] 益岡隆志, 田窪行則:「基礎日本語文法」, くろしお出版, 1989.