

中国語連体修飾節構文の解析

周振*‡ Alastair Butler* 吉本啓*‡

*東北大学高度教養教育・学生支援機構 †東北大学大学院国際文化研究科 ‡日本学術振興会

要旨

本研究は、中国語連体修飾節構文を対象に、その解析を決めようとするものである。本研究では、連体修飾節構文は、関係節構文および同格節構文に大別された。さらに、関係節構文に対して、主要部終端型、主要部先導型および名詞的關係節という三つの下位カテゴリーも付与した。この分類に基づき、本研究は、各種類の連体修飾節構文について、その統語解析と意味解析をそれぞれ順次考察した。これによって、中国語連体修飾節構文に含まれる統語・意味情報が形式的に捉えられるようになり、各種類の間の相違に関する比較もより直接的に行えるようになった。

1. はじめに

筆者たちは、中国語の無制約のテキストに対して、論理意味表示(述語論理式)を付加した中国語の意味表示コーパスを構築している。その作業は二つの段階に分けられる。すなわち、(1)分析データとして選ばれた中国語の自然テキストに対し統語解析情報を付与すること、および(2)バトラー (Butler 2010) が提唱するスコープ制御理論 (Scope Control Theory; SCT) を実装したシステムで(1)の結果を処理することによる自動的な文の論理意味表示の獲得である。

本研究は、中国語連体修飾節構文を対象に、その解析を決めようとするものである。具体的には、まず先行研究を踏まえて、意味処理からの要請を考慮し中国語連体修飾節構文の定義およびその分類を決定する。さらに、一般的な解析方法を用いて、中国語連体修飾節構文の各パターンに表層的な統語解析情報を各々付与する。これによって、精度の高い意味処理を行い、得られた述語論理式を手掛かりとして中国語連体修飾節構文の各種類の間の相違に関する比較をより直接的に行いたい。

2. 本研究のアノテーション方式

本研究は、統語解析を行う際に、基本的にはペン通時コーパス式の解析規約 (Santorini 2010) に従う。これは、ペンツリーバンク式の解析スキームを修正したもので、後者に比べると、句の内部構成を規則正しく見つけ出せることと節と節の間の統語関係が容易に読み取れるなどの

メリットを持っている。また、中国語の実情および意味処理の要請に配慮し、筆者たちは、前置詞句に関する機能情報をユニークな手法で与えたり、句レベルの要素に対してそのスコープに関する情報をアノテーションしたりするなど、ある程度の修正も行った(その詳細は、Butler et al. 2013 および周他 2015 を参照のこと)。

3. 中国語連体修飾節構文の定義とその分類

名詞の修飾部としては、形容詞句や限定詞句などの句 (phrase) がよく見られるが、それが節 (clause) に任せられることもある。一般的には、名詞の修飾部を節が担当するような構造を持つ文は連体修飾節構文と呼ばれる。中国語の連体修飾節構文の判定については、従来修飾部と被修飾部との相対位置、補文標識 (complementizer) “的 de” の有無および被修飾部の省略可否をポイントにして行われるものが多いが、本研究はより網羅的にデータを考察するために、基本的には以上の三点を全部問わないようにしている。ただし、(1)のような動詞連続構文と区別するには、“的”あるいは修飾部のどちらかが具現化しなければならない(つまり、両方が同時にゼロ形式を取ってはいけない)という制限も加える。

(1)

[IP-MAT [IP-PPL-SBJ 学中文]不难]。
中国語を学ぶことは難しくない。

一方、連体修飾節は、その被修飾部の名詞が修飾部の中で文法的な役割を果たすか否かによって、関係節 (relative clause) と同格節 (appositive clause) に分けられることが一般的である。本研究の場合もその分類法に踏襲するが、中国語の実情に応じ、関係節に対してさらに三つの下位カテゴリーを付与する。従って、本研究における分類は以下のとおりになる。

- i 関係節構文
 - i -1 主要部終端型 (head-final)
 - i -2 主要部先導型 (head-initial)
 - i -3 名詞的關係節 (nominal relative clause)
- ii 同格節構文

4. 関係節構文

修飾部に関する文法的な役割の一部が被修飾部の名詞によって担われているような連体修飾節構文は関係節構文と呼ばれる。修飾部と被修飾部が一緒になってさらに名詞句を投射するため、被修飾部の名詞が主要部（ヘッド）になる。中国語の場合、被修飾部の名詞が担当できる文法的役割は、Keenan et al. (1977) が提唱した名詞句接近可能性階層 (NP accessibility hierarchy; 主語>直接目的語>間接目的語>斜格>属格>比較の対象) における最初の「主語」から最後の「比較の対象」まで幅広く観察されている。以下、主要部の位置を手がかりに、関係節構文の各下位カテゴリーについて詳しく見ていくことにする。

4.1 主要部終端型

修飾部が主要部の前に現れる関係節構文は、中国語においてもっとも普通に見られている。(2)はその一例である。

(2) 之前失去的机会又来了。
この前失ったチャンスがまたやってきた。

“机会”は、関係節“之前失去的”によって修飾され、しかもその中で直接目的語の文法的な役割を担っている。これを踏まえて、(2)に対して以下の(3)のような統語解析を与えることが出来る。

(3)

```
(IP-MAT (NP-SBJ (CP-REL (IP-SUB (NP-SBJ *pro*)
                                (NP-OB1 *T*)
                                (ADVP (ADV 之前/この前))
                                (VB 失去/失う))
                                (C 的/補文標識))
    (N 机会/チャンス))
  (ADVP (ADV 又/また))
  (VB 来/やってくる)
  (AS 了/完了))
```

(3)では、関係節は CP-REL というタグが与えられ、その内部に現れるトレース *T* が直接目的語 NP-OB1 として扱われている。そして、(3)を SCT をインプリメントした意味処理システムに入力して自動意味評価することにより、(2)の述語論理式は、(4)のように与えられる。

(4)

```
∃ x4 x1 e2 e3 (
  x4 = pro ∧ 机会(x1) ∧ 之前(e2) ∧
  失去(e2, x4, x1) ∧ 又(e3) ∧ 来_了(e3, x1))
```

二つの述語“机会”と“失去”は連言結合子 ∧ によって結びつけられている。従って、関係節構文の場合、修飾部と被修飾部との意味的關係は、並列関係であるということがわかった。

4.2 主要部先導型

中国語の場合、関係節が常に主要部の名詞に先行しているという印象が強いかもしれないが、(5b)が示すように、実は関係節が主要部の後ろに来るケースも時々見られている。

(5)

a. 我遇到一个会说中国话的外国人。
b. 我遇到一个外国人会说中国话。
私は中国語が話せる一人の外国人と出会った。

(5b)では、“外国人”の様態と状況に関する情報は次に来る修飾部“会说中国话”が提供している。また、主要部終端型である(5a)に比べて、形式上補文標識の“的”が具現化していないほか、両者の意味はほとんど相違がない。とはいっても、被修飾部に関する情報伝達の手法については、両者の間にはやはり微妙な差がある。つまり、主要部先導型においては、その被修飾部にあたる名詞の出現が完全に偶発的なものであるが、主要部終端型においては、関係節により修飾される名詞は、談話の中で既に現れたか、あるいはその出現が予測できるものだと考えられる。

本研究で提唱する主要部先導型は、従来中国語において、関係節として扱われることが少ない。その代わりに、この種の構文は、そのディスコース上の機能としてこれから説明する名詞句の導入が挙げられるので、存在文 (presentative sentences) の一種類とみなされてきた。しかし、以上で見られたように、この種の構文は典型的な関係節構文との類似性が非常に高い。しかも先行する名詞がそれに後続する節の中で(5)で見られた主語をはじめ様々な文法的役割を果たすことが出来る。従って、統語解析の観点から、本研究は、この種の構文を関係節構文の一種類として扱うことにする。(5b)に対して、(6)のような統語解析を付与する。

(6)

```
(IP-MAT (NP-SBJ (PRO 我/私))
  (VB 遇到/出会う)
  (NP-OB1 (NUMCLP (CARD 一/一)
                (NUMCL 个/個))
    (N 外国人/外国人)
    (CP-REL (IP-SUB (NP-SBJ *T*))
```

(AX 会/できる)
(VB 说/話す)
(NP-OB1 (NPR 中国话/中国語))))))

語))))))

(3)と(6)は、関係節 CP-REL の内部における補文標識 C の有無および関係節と修飾される名詞 N との相対位置のほかには特に変わりがなかった。(6)の意味処理結果は(7)のようになる。

(7)

∃ x3 X1 e2 (
x3 = 我 ∧
一 个(x1) ∧ 外国 人(x1) ∧
会(∃ e4 说(e4, x1, 中国话)) ∧ 遇到(e2, x3, x1))

統語処理の段階でトレース NP-SBJ に関するアノテーションを行ったため、“外国人”が“说”の主語項としてその意味役割を果たしているという情報が意味処理の段階で捉えられるようになった。このように、本研究が敢えてこの種の構文を関係節構文として取り扱う理由は、関係節の内部にトレースを追加することにより主要部名詞と修飾部との間の統語関係を容易に明示できることにある。

4.3 名詞的關係節

中国語において、関係節が独立して使われることも可能である。しかし、これはあくまでも関係節の話で、第5章で論じる同格節には通用しない。つまり、中国語では、主要部ぬきで現れる連体修飾節は必ず同格節ではなく関係節である。また、具現化されていない主要部が関係節の内部でどのような文法的な役割を果たしているかについて、Li et al. (1981) は次のようにまとめていた (訳は筆者たちによるものである)。

- ①単独で使用されるために、関係節は必ず動詞を含めしかもその動詞が指定されていない項を少なくとも一つ持たなければならない。
- ②未指定の項が一つしかない場合、具現化されていない主要部の名詞が関係節の中で担う文法的な役割は未指定の項と同じである。
- ③主語と直接目的語が両方とも未指定の場合、具現化されていない主要部の名詞が関係節の中で担う文法的な役割は一般的には直接目的語とされる。
- ④具現化されていない主要部の名詞が関係節の中で間接目的語という文法的な役割を担うことはない。

以上を用いて、具体例を分析してみよう。

(8)

从事服务业的很辛苦。
サービス業に従事する人が辛い。

(9)

这里没有我要的。
ここに私のほしいものはない。

(10)

难民缺少穿的。
難民は着るものに欠ける。

(11)

卖军火的都是恶人。
兵器弾薬を売るのはみんな悪人である。

(12)

卖的比租的贵。
売るものはレンタルのものより高い。

まず、(8)と(9)の関係節において、その動詞“从事(従事する)”および“要(要る)”がみな二項動詞であり、しかも両方の主語と直接目的語がそれぞれ指定されていないため、②により、具現化されていない主要部名詞が(8)と(9)の関係節において、主語および直接目的語の文法的な役割をそれぞれ果たしているということがわかる。次に、(10)の関係節における二項動詞“穿(着る)”はその項が両方とも未指定のままであるので、③に基づいて主要部名詞が担っている文法的役割は主語ではなく直接目的語であるという結論に至れる。最後に、三項動詞が含まれる(11)と(12)についても、②と③に④さえ加えれば、同様に正しい推論が得られる。

①から④までの規則を意味処理システムにインプリメントすれば、今度インデックスと pro のアノテーションをしなくても要素の間の意味的關係が正確的に捉えられるようになる。例えば、(12)に対して、以下のような統語解析を行う。

(13)

(IP-MAT (NP-SBJ (CP-REL (IP-SUB (VB 卖/売る))
(C 的/補文標識)))
(PP (P 比/より)
(NP (CP-REL (IP-SUB (VB 租/レンタルする))
(C 的/補文標識))))
(ADJ 贵/高い))

そして、(13)の意味処理結果が(14)の示すとお

りである。

(14)

∃ x7 x6 x1 x2 e3 e4 e5 (
x6 = pro ∧
卖(e3, x6, x2) ∧
x7 = pro{x2} ∧
租(e4, x7, x1) ∧ 贵(e5, x2) ∧ 比(e5) = x1)

(13)の IP-SUB の内部に主語の pro と直接目的語のインデックスが見られていない。それにしても、(14)においては、pro の値を持っている項 x6 と x7 が追加され、それに統語の段階で具現化されていない主要部の名詞も、x1 と x2 としてそれぞれ述語“卖”と“租”の arg1 (直接目的語) の位置に現れたのである。

5. 同格節

関係節と異なり、同格節によって修飾される主要部名詞は其中で何の文法的な役割も担当しない。それに、被修飾部の名詞は抽象的な概念を示すものが多い。また、4.3 節で触れたように、主要部名詞の具現化が常に求められ、同格節が単独では使用されることが出来ない。(15)は4.1 節で見られた主要部終端型の関係節構文に対応している同格節構文の例である。

(15)

去美国的.机会又来了。
アメリカに行くチャンスがまたやってきた。

“机会”は抽象名詞で、その内容が修飾部“去美国的”によって補充されるが、“机会”自体が同格節“去美国的”の中で文法的な役割を果たしていない。(16)は(15)の統語解析である。

(16)

(IP-MAT (NP-SBJ (CP-THT (IP-SUB (NP-SBJ *pro*)
(VB 去/行く)
(NP-OB1 (NPR 美国/アメリカ))
(C 的/補文標識)
(N 机会/チャンス)
(ADVP (ADV 又/また))
(VB 来/やってくる)
(AS 了/完了))

関係節は CP-REL というタグが与えられたのに対して、同格節に付与するタグが CP-THT になる。これにより、意味処理システムは両者の識別が可能になり、(16)を入力として、以下の処理結果を返してくれる。

(17)

∃ x4 x3 e1 e2 (
x4 = pro ∧
机会(x3, 去(e1, x4, 美国)) ∧ 又(e2) ∧ 来_了(e2, x3)

(4)と比べると、関係節で見られた並列関係と違って、(17)では、修飾部の意味が被修飾部名詞の意味へ埋め込んでいるということがわかる。

6. まとめ

統語・意味情報をタグ付けした中国語コーパスの構築における連体修飾節構文の解析について論じてきた。本研究は、連体修飾節構文と関係節構文および同格節構文に大別し、そして、関係節構文に対して、主要部終端型、主要部先導型および名詞的關係節という三つの下位カテゴリも付与した。この分類に基づき、本研究は、各種類の連体修飾節構文について、その統語解析と意味解析をそれぞれ順次考察した。これによって、中国語連体修飾節構文に含まれる統語・意味情報が形式的に捉えられるようになり、各種類の間の相違に関する比較もより直接的に行えるようになった。

参考文献

Butler, A. (2010) The Semantics of Grammatical Dependencies. Bingley: Emerald.
Butler, A., Otomo, R., Zhou, Z., and Yoshimoto, K. (2013) Treebank Annotation for Formal Semantics Research. In Y. Motomura, A. Butler & D. Bekki (eds.), JSAI-isAI 2012 Workshops, LENLS, JURISIN, MiMI Miyazaki, Japan, November/December 2012 Revised Selected Papers, New Frontiers in Artificial Intelligence, 25-40. Berlin and Heidelberg: Springer-Verlag.
Keenan, E., and Comrie, B. (1977) Noun Phrase Accessibility Hierarchy and Universal Grammar. Linguistic Inquiry 8, 63-99.
Li, C., and Thompson, S. (1981) Mandarin Chinese: A Functional Reference Grammar. Berkeley: University of California Press.
Santorini, B. (2010) Annotation Manual for the Penn Historical Corpora and the PCEEC (Release 2). Tech. rep., Department of Computer and Information Science, University of Pennsylvania.
周振, Alastair Butler, 吉本啓 (2015) 「中国語意味解析コーパス構築のための句レベルのスコープアノテーション—一文の構成要素の間のコントロール関係の同定および否定の作用域の制御を中心に—」, 『言語処理学会第 21 回年次大会 発表論文集』, 856-859.