

中国語名詞句の内部構造について

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 ‡国立国語研究所

要旨

本論文は、統語・意味情報をタグ付けしたコーパスの構築という視点から、中国語名詞句の内部構造を解明しその解析を行おうとするものである。論文の構成は以下の通りである。第1章では、研究の背景と目的を述べる。第2章では、本研究が使用するペン通時コーパス式の解析規約の特徴を紹介し句の内部構造の一般化を行う。第3章では、名詞句の主要部および主要部の省略について論じる。第4章では、先行研究を踏まえた上で、名詞の修飾部を7つの種類に分け名詞句の内部構造を検討しながらその解析をそれぞれ決定していく。第5章では、複数の修飾部が同時に現れる時の語順について考える。第6章では、まとめを行う。

1. はじめに

名詞句 (NP) は自然言語の中でもっとも重要でしかも頻繁に使われている文法カテゴリーの一つである。動詞と共に文の主題、主語、直接目的語や間接目的語になったり、助詞の前後に現れてその目的語になったりすることが通言語的に見られる。さらに、中国語の場合、一部の名詞が副詞的に使われて直接述語を修飾することや助動詞ぬきでそのまま文の述語になることも可能である。NP は実に豊富多彩な文法的役割を果たしていると言える。

筆者たちは、中国語の無制約のテキストに対して、統語構造 (ツリー) と論理意味表示 (述語論理式) を付加した中国語の統語・意味表示コーパスを構築している (Butler 2015 など) が、名詞句をめぐるアノテーション作業は二つの課題に直面している。すなわち、(1) 名詞句の内部構造を解明しその解析を決めること、および(2) 孤立語である中国語には明示的な名詞の格の変化がないため、様々な状況で現れる名詞句の文法的役割を検討し適切な機能タグを付与することである。(2)を徹底的に論じるためには名詞句を中国語の各構文での用法に即して考察する必要があると思われるが、一つの論文でその全部を論じることは無理がある。そこで、本論文はまず(1)を課題とし、名詞句の内部構造を明らかにした上でその解析を決定しようとする。

2. 本研究のアノテーション方式

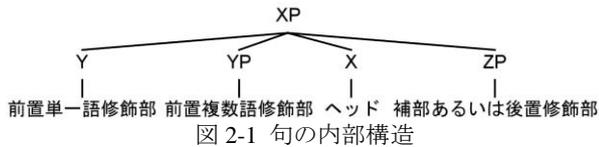
2.1 ペン通時コーパス式の解析規約

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ある。統率・束縛理論 (GB 理論; Chomsky 1981, 1982) に拘った後者に比べて、前者は実用性重視という原則を貫き、特定の言語理論には偏っていない。このことは、当解析規約は文の統語構造 (ツリー) を極力シンプルなままに保とうとする工夫にも見られる。ペン通時コーパス式の解析規約では、中間レベルの構造 (N', ADJ' など) は、いかなる場合も明示的にタグ付けされることはない。また、言語学者たちがこれまで使い続けてきた幾つかの句 (例えば、動詞句や限定詞句) の使用も廃止されている。これらによって、句・節の構造が一般に平坦で、複数の枝別れノードを持つようになる。

2.2 句の内部構造

句のヘッド (head) は各句によって直接支配されるため、中間レベルの構造は示されない。原則として、ヘッド (N, ADJ など) は句レベルのカテゴリー (NP, ADJP) と一致しなければならないが、二つの例外がある。それは、ヘッドが一般的なカテゴリーのラベルよりも特殊なラベルを持っているか、またはヘッドが省略されているか、のどちらかの場合である。

このように、ヘッドは通常句を投射するが、これにも二つの例外がある。まず、動詞 (VB)、助動詞 (AX)、間投詞 (INTJ)、接続詞 (CONJ) などの幾つかの品詞タグは常に句を投射しない。さらに、単一の単語からなる前置修飾部も句を投射しない。これは、単一の単語からなる前置修飾部の投射する句の可能性は予測できることから、枝分かれのない投射を減少することによって句の内部構造を単純化するためである。これに対して、複数の単語から構築される前置修飾部は相当する句を投射する。一方、補部および修飾部が後置する場合は、いずれも相当する句を投射する必要がある。以上を踏まえて、句の内部構造 (IP と CP は節として扱われるため、両者を除く) は、図 2-1 のように一般化することができる。



Y と YP の順番は問わない。また、中国語の場合、ヘッドが補部 (complement) の後ろに来ることも稀にある。(XP (YP/ZP X))のような極端な例では、構造上から X の前方にくるのは YP かそれとも ZP かを見分けることが理論上不可能だが、このような例はめったにないし、実際に現れたとしても、補部が修飾部 (modifier) よりも優位であることを考えれば両者の区別は困難ではない。中間レベルの構造は存在しないので、従来の所謂指定部 (specifier) および付加部 (adjunct) は修飾部に統合され、しかも修飾部も補部も常にヘッドと同じレベルにあり、みなヘッドの姉妹である。

3. 名詞句の主要部

言うまでもないが、名詞句の主要部は通常普通名詞 (N) である。固有名詞 (NPR) と代名詞 (PRO) は普通名詞ではないが、普通名詞よりも特別な存在だと考えられるため、名詞句の主要部になることもできる。中国語の場合、名詞句の主要部の省略も可能である。

(1)

你是北京大学的學生嗎？
君は北京大學の學生ですか。
是，我是北京大學的。
はい、私は北京大學の（學生）です。

(1)で分かるように、名詞句の主要部の省略は、1) 名詞句の修飾部に助詞“的 de”を含む助詞句が存在することおよび2) 聞き手は文脈から省略された主要部の復元が可能であることを条件とする。(1)のような例のほか、名詞句の主要部の省略は、関係節構文にもよく見られる。

(2)

这里没有我要的。
ここに私のほしい（もの）はない。

(2)における補文標識 (complementizer) の“的 de”と(1)における助詞の“的”とは異なるものである。そして、省略された主要部の指示対象の特定が文脈に頼らなければならない(1)に対して、(2)の場合、文脈がなくても、このような特

定は主要部を修飾している関係節の統語構造からもある程度行える。

4. 名詞句の内部構造

普通名詞、固有名詞および代名詞（以下まとめて名詞）は、そのまま名詞句を投射することができる。(1)における“你 nǐ (君)”、“北京大学 běijīngdàxué (北京大学)”と“我 wǒ (私)”はその例である。一方、名詞は修飾部を取って名詞句を構築することもある。(1)では、“北京大学的學生”という名詞句は、主要部名詞の“學生 xuésheng (學生)”と修飾部助詞句の“北京大学的 (北京大學の)”の両方からなっている。

名詞の修飾語・句については、従来様々な観点から分類、分析されてきた。Li et al. (1981) は、名詞を修飾する要素を数助詞句 (classifier phrase)・度量句 (measure phrase)、連合句 (associative phrase) と限定句 (modifying phrase) に分けて議論している。その他、生成文法の観点から中国語の名詞の修飾部を六つの種類に分けて分析した何 (2011) の研究もある。先行研究に比べて、本研究は中国語のデータを網羅的且つ形式的に扱う必要があるため、より詳しい対応が求められる。そこで、本研究では、本研究のアノテーション方式の枠組みの中で名詞の修飾部を七つの種類に分類する。それは、量化詞 (Q)、指示詞 (D)、数助詞句 (NUMCLP)、助詞句 (PP)、形容詞 (ADJ)、関係節 (CP-REL) および同格句 (NP-RFL/NP-PRN)・節 (CP-THT) である。

4.1 量化詞、指示詞と数助詞句

(3)のような例から、Li et al. (1981) は、数助詞は数詞、指示詞 ((3)の場合は“这 zhè (近称)”)あるいは量化詞のどちらかとの共起が必要で、量化詞も指示詞も数助詞句の一部であると見なしている。

(3)

这条裤子
このズボン
*条裤子

しかし、(4)に示す通り、(3)のような例の場合、実は数詞“一 yī (一)”が省略されたに過ぎないと考えられる。

(4)

这 (一) 条裤子
この (一本の) ズボン

中国語の場合、数助詞の前の数詞“一”はしばしば省略されるのである。さらに、指示詞が文頭に現れる（即ち、文の主語か主題となる）時は数助詞句および主要部名詞の省略が可能であり、量化詞が数助詞句を経由せずにそのまま主要部の名詞を修飾することもできる。従って、指示詞と量化詞は数助詞句の一部ではなく、より高い独立性を持つ名詞修飾部であると考えられる。本研究は、指示詞と量化詞にそれぞれDとQという品詞タグを与え、(4)の統語解析を以下の(5)のように付与する。

(5)
 (FRAG (NP (D 这)
 (NUMCLP (CARD 一)
 (NUMCL 条))
 (N 裤子)))

(5)では、名詞句の主要部名詞“裤子 kùzi(ズボン)”は二つの修飾部を持ち、Dは独立した修飾部として扱われている。そして、前置修飾部が一語からなるもの場合は、句への投射を要求しないので、DはDPを投射していない。一方、もう一つの修飾部の数助詞句(NUMCLP)は、その主要部の数助詞(NUMCL)が数詞(CARD)を取って投射したものである。

4.2 助詞句と形容詞

本研究における名詞の修飾部としての助詞句は、Li et al. (1981) の分類に見られる連合句に比べて、適用の範囲がより広がっている。つまり、本研究は助詞“的”と組み合わせる要素を名詞のみに限定しない。中国語の場合、名詞の他、形容詞、量化詞および前置詞句も助詞“的”と一緒に修飾部助詞句を構築することができる。(6)は“的”と量化詞“所有 suǒyǒu (あらゆる)”とのパターンの例である。

(6)
 所有的学生
 すべての学生

一般的には、助詞と組み合わせられる要素を助詞の補部と見なすため、単一語は助詞“的”に前置しても、その句への投射が必要である。(6)の統語解析は次の(7)のようになる。

(7)
 (FRAG (NP (PP (QP (Q 所有))

(P 的))
 (N 学生)))

(7)では、QはQPを投射してからPの補部となりPPを構築している。なお、本研究は、助詞に異なる品詞タグを付与しないため、助詞の“的”といわゆる格助詞の“的”(所有格)との区別はしないことになる。

一方、“红花 hóng huā (赤い花)”のような主要部名詞をそのまま修飾する形容詞に関しては、関係節を構築する必要がないと本研究は考える。これは生成の経済性を考慮したためである。従って、“红花”の統語解析を(8)のように付与することができる。

(8)
 (FRAG (NP (ADJ 红)
 (N 花)))

形容詞が直接主要部名詞を修飾する時は、ADJがNの前置単一語修飾部になるため、ADJPへの投射を必要としない。

4.3 関係節と同格句・節

Li et al. (1981) は、主要部名詞句を修飾する節のことを名詞化として扱っている。さらに、中国語の名詞化は、助詞“的”が動詞句などに後置する形を取ると述べている。しかし、実は中国語の場合、動詞句などがそのまま主語名詞句や直接目的語名詞句として使用されることが多い。従って、本研究は、主要部名詞を修飾する節のことをその統語的特徴によって関係節および同格節に分けることにする。なお、関係節および同格節の解析に関する詳細は周他 (2016) を参照してほしい。

主要部名詞と同格関係を持つ修飾部は、節の他に、句もある。これは、主に“我自己 wǒzìjǐ (私自身)”のような再帰代名詞“自己(自身)”が代名詞“我(私)”に後置するパターンや“冠军刘翔 guànjūnliúxiáng (チャンピオン劉翔)”のような固有名詞“刘翔(劉翔)”と肩書を示す普通名詞“冠军(チャンピオン)”からなるパターンに見られる。

同格関係を持つ両者のどちらが名詞句の主要部なのかに関する判断はある意味では容易ではない。とはいうものの、先行研究のほとんどは、中国語の名詞句は主要部名詞が常に名詞句の最後の位置に現れると考えている。同格関係も修飾関係の一種と見なせば(同格関係は修飾関係ではないと考える説もある)、“我自己”と“冠

军刘翔”の主要部もそれぞれ“自己”と“刘翔”になるべきだと考えられる。しかし、実際のところ、中国語の場合、修飾部が主要部名詞の後ろに来ることもあり得る（その詳細は、周他(2016)を参照のこと）。しかも、本研究のアノテーション方式から見ても、同格関係を内蔵する名詞句の修飾部が前置でなく後置しているという考え方を取るべき理由がある。従来の統語論研究は、“自己”のような再帰代名詞を特別なものとして取り扱うことが多い。そのため、“我自己”と“冠军刘翔”では、主要部名詞と修飾部名詞とがみな同格関係を持ったとしても、両者の区別がツリーバンクの内部ではっきりできるようにしたい。実際のところ、ペン通時コーパス式の解析規約においても、両者の区別がなされている。このような区別は、常に明示的な機能タグの付与によって実現されている。とはいうものの、機能タグが持てるのは句レベルのカテゴリーのみなので、修飾部名詞の名詞句への投射が求められるようになる。さらに、本研究のアノテーション方式では、単一語修飾部の句への投射が認められるのは、それがヘッドの後ろに現れる場合に限られる。従って、“我自己”と“冠军刘翔”に対しては、修飾部を担当する名詞をそれぞれ二番目の“自己”と“刘翔”に設定したほうが適当だと思われる。以上の議論を踏まえて、本研究では、両者の統語構造を次の(9)、(10)のように与える。

(9)
(FRAG (NP (PRO 我)
(NP-RFL (PRO 自己))))

(10)
(FRAG (NP (N 冠军)
(NP-PRN (NPR 刘翔))))

(9)と(10)の示すとおり、“我自己”と“冠军刘翔”の主要部はみな一番目の名詞の“我”と“冠军”になっている。一方、二番目の名詞“自己”と“刘翔”は、どちらも後置修飾部として扱われNPを投射しているが、NPに与える機能タグが異なる。NP-RFLは再帰名詞句を、NP-PRNは同格名詞句を意味する。

5. 名詞修飾部の語順

名詞句にとって、主要部名詞を除いて（主要部が省略される場合もある）、すべての修飾部が不可欠なものではない。本研究は、中国語の名詞句は修飾部後置も可能だという立場を取って

いるが、修飾部後置に比べて修飾部前置の割合がずっと高いというのも事実である。なお、二つ以上の前置修飾部が同時に現れる場合、その順番は決して任意のものではない。先行研究を踏まえて、本研究は、名詞修飾部と主要部名詞との語順は次のような二つの可能性があると考ええる。

- (11)
a. PP (NP + P) + D/Q + NUMCLP + CP-REL + CP-THT + PP (ADJP + P) + ADJ + N + NP-RFL/NP-PRN
b. PP (NP + P) + CP-REL + D/Q + NUMCLP + CP-THT + PP (ADJP + P) + ADJ + N + NP-RFL/NP-PRN

(11)から分かるように、中国語の場合、関係節は数助詞句の後ろにも限定詞・量化詞の前にも現れることができる。

6. まとめ

統語・意味情報をタグ付けしたコーパスの構築という視点から、中国語名詞句の内部構造およびその解析について論じた。実用性を重視した解析規約を採用したため、名詞句の内部構造をシンプルなままに保ちながら、要素間の統語情報を正しく捉えることができた。また、本研究は、名詞を修飾するあらゆる要素を視野に入れて中国語のデータを網羅的に考察した。先行研究に比べて、本研究は、研究の包括性の点で勝っていると考えられる。

参考文献

- 周振, Alastair Butler, 吉本啓 (2016) 「中国語連体修飾節構文の解析」, 『言語処理学会第22回年次大会 発表論文集』, 809-812.
何元建 (2011) 《现代汉语生成语法》, 北京: 北京大学出版社.
Butler, Alastair. (2015) *Linguistic Expressions and Semantic Processing: A Practical Approach*. Switzerland: Springer International Publishing.
Chomsky, Noam. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, M. A.: MIT Press.
Li, Charles N., and Thompson Sandra A. (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
Santorini, Beatrice. (2010) Annotation Manual for the Penn Historical Corpora and the PCEEC (Release 2). Tech. rep., Department of Computer and Information Science, University of Pennsylvania.