

# 中国語文内照応現象 解析のための規則ベースの構築

于 素秋<sup>†</sup> 横山 晶一<sup>‡</sup> 西原 典孝<sup>‡</sup>

<sup>†</sup>山形大学大学院工学研究科

<sup>‡</sup>山形大学工学部

## 1 はじめに

文内における照応現象は、名詞や動詞に依存するだけではなく、文構造や、文成分間の意味関係などにも依存する。本論文は話題や情報による文構造上の制約、文と文をつなげる関連語句の役割、文成分間の意味関係などの中国語の特質に基づき、照応規則ベースを構築する。それらの照応規則により、中国語における文内照応現象を解析し、解を求める方法を提案する。

## 2 主題と情報に依存する特質

人間は言語を発するとき、どれを先に言うべきかという発話の順序を、強調対象の異なりにより考慮する。又、他人に何かの情報を伝達しようとする時、情報の重要度を強調するためにも、発話の順序を考慮する。中国語では、文の先頭部に置かれたものを主題部と見做し、その後に置かれたものを述題部と見做す。普通、主題部は談話の出発点であり、既知の、又は分かりやすい情報を表現する。述題部は、伝達しようとする情報を展開し、新しい情報を受話者に伝える部分であり、未知の、又は理解し難い情報を表現する[1]。発話者は、元の話題に対して、引き続き陳述するか、又は陳述した新しい情報を話題にして、更に新しい情報を伝えるかする。従って、述題部の出発点として2つの部分が考えられる。1つは前文の主題、すなわち旧情報である。それは主題を引き継ぎ、文を展開する。もう1つは前文の述題であり、伝達したばかりの情報である。それは新規情報を次の主題にとして、文を広げる。

しかし、実際の文章では、特別に強調する場合以外では、主題の新旧にかかわらず、主題に、なるべく代名詞で代替するか、省略する方が普通である。特に3つ以上の分句からなる複文の場合には、2番目の分句から、主題を省略することが特に多い。

本研究では、複文を「主題チェーン型」と「情報チェーン型」の2種類に分類する。主題チェーン型の複文は、既知情報を利用し、未知情報を伝達する

複文である。それは図1のように示される。

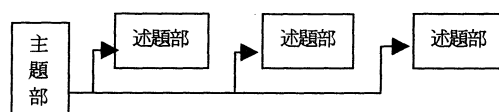


図1 主題チェーン型複文構造

情報チェーン型複文は、新情報を主題にし、更に未知情報を引出して伝達する複文である。それは図2のように示される。

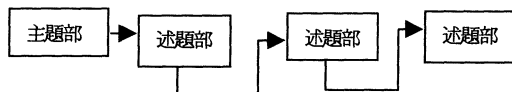


図2 情報チェーン型複文構造

「主題チェーン型」複文は、各分句の述詞が動作や行為を表す動詞文である。「情報チェーン型」複文は、各分句の述詞が非動詞（状態を表す自動詞を含む）文である。実際には、これらが混合する場合が多い。つまり、図3と図4のように、複文には動詞文と非動詞文が共に存在する。

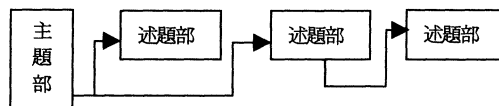


図3 主題型から、情報型に転換する複文

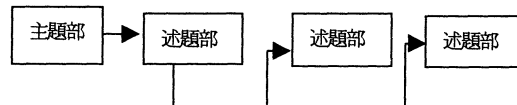


図4 情報型から主題型に転換する複文

どのようなタイプの複文であっても、二つの分句

の間に存在する関係は次の3種類しかない。

1. 二つの分句が同じ主題を共用する。これは主題チェーン型の複文であり、主に主題部にある主語が連続的に動作や行為を行うことを表す。
2. 前の分句は存現や消失などを表す文であり、後の分句は前の分句の動作主体を主題にする。これは情報チェーン型の複文であり、新しい情報を次々に展開する。
3. 前の分句は動作文であり、後の分句は前の分句の目的語を主題にする。つまり、前の分句は、動作や作用、行為などを行う文であり、後の分句は前文の動作や行為の対象や結果などを引き続き表現する。

中国語複文の以上のような特徴は、文内における有形照応詞や無形照応詞の存在と解析には重要な手がかりと参照点を提供する。本研究は以上のような話題や情報による構文構造の特徴に基づき、それらの関係に依存する照応現象を解析する照応規則A類を作成した。以下はその1つの例である。

[A06] case : PP か ZP の P = V

I : X の P = VT

II : X = MC1. OBJ = N (不定指名詞)

III : MC1. SUB = N (定指名詞)

THEN : X = 先行詞

### 3 文と文を繋げる関連語句に依存する特質

幾つかの文を繋げる手段として、二つの方法が考えられる。それは文の圧縮と文の連結である。圧縮させる手段としては、関係節によって、重複する名詞を省略したり、代名詞化したりする方法がある。文を連結させる手段としては、主題、述題、目的語、時には定語などの関係節で連結したり、等位関係、主従関係などを示す関連語句で連結したりする方法がある[2]。

接続関係を陽に示す関連語句が用いられていなくても、文を理解することができる場合もあるが、その場合は、文と文或は節と節の間の意味的な繋がりによらなければならない。関連語句があれば、聞き手は、言語知識を用いて、前文と後文との間の意味的關係や、文中にある単語と単語の意味関係などに、ある種の予測を立てることができる。普通、継起や累加などの複文は主題をめぐり、主題節や主語節を連結する方法を取り易い。逆に、存現文、非動詞述詞文を含む複文は、前文の述題部分を後文の主題と見做し、直接前文について陳述する述題節や、目的語節を連結する方法を取りやすい。

関連語句は文と文の関係を暗示する役割を持っている[3]。例えば“因为…，所以…（…だから、…）”は因果関係を、“虽然…，可是…（…，けれども…）”は転折関係を示している。又、複文の関連語句の位置は、複文の前後関係をを制約すると同時に、前後文の主語が同一であるかどうかをも制約する。例えば、普通2つの分句の主語が同じであれば、関連語句の先頭部分は、前分句の主語の後に置かれることが多い。2つの分句の主語が異なる時、関連語句の先頭部分は、前分句の主語の前に置かれることが多い。例えば、関連語句「不但…，也…」（だけではなく、…也…）は累加関係を表す。先頭部分の「不但」は、主語の前に置かれた場合、前後文の主語が異なり、主語の後に置かれた場合、前後文の主語が同じであることを制約する代表的な例である。複文は、関連語句の位置によって複文の前後関係を制約する機能を持つので、前後文の主語が同一であるかどうかは、その関連語句先頭部分の所在位置によって判断することもできる。

・不但張三来，李四也来。

（張三が来るだけではなく、李四も来る。）

・張三不但来了，把李四也领来了。

（張三は自分だけではなく、李四を連れて来た。）

本研究は関連語句の機能に基づき、関連語句の有無と、その位置に依存する照応現象を解析する照応規則B類を作成した。以下はその例の1つである。

[B05] case : S = CS (並列)

MC1 【一边 又】…

MC2 【一边 又】…

I : X = MC1・SUB

MC1 = SUB + 【一边 又】 + V…

II : MC2 = 【一边 又】 + V…

MC2 の【…】 = MC1 の【…】

Then : X = 先行詞

### 4 文成分間の関係に依存する特質

中国語の文成分間の関係については、文型から見て、陳述・被陳述、描写・被描写、説明・被説明の関係が存在し、フレーズから見て、修飾・被修飾の関係のみが存在する。陳述・被陳述の関係は、述詞が動詞である文に存在する。描写・被描写の関係は、述詞が形容詞である文に存在する。それらの関係は名詞と述詞との関係として、既に検討した[4]。ここでは説明・被説明関係と、修飾・被修飾関係を検討し、それらの関係に依存する照応現象を解析する規則を作成する。

#### 4. 1 説明関係を用いる規則の構築

説明と被説明関係は、主に動詞「是 (is)」で構成する文に存在する。

「是」は中国語における特別な動詞である。それは動作や行為を表さず、判断や肯定の意味を表す。

「是」の文型は「SV0」であるが、主語のSUBと目的語のOBJは、「是」の支配対象ではなく、「是」の連結対象である。従って、動詞「是」の前後成分は、「是」と直接関係はなく、前のSUBと後のOBJを関係付けるだけである。そして、名詞や代名詞、数詞や数量詞、動詞や形容詞など、どんな単語でも、「是」のSUBと「是」のOBJとなることができる。又、各種のフレーズも、文も、「是」のSUBとOBJとなることができる。一般的に言えば、「是」の前後成分は、名詞であるか、名詞に相当するものである。

「是」で構成する文は「A是B」であり、AとBの関係は2種類に分類できる。その1つは、「A=B」であり、同時に「B=A」も成立する同等的な説明関係である。

・北京是中国的首都。=中国的首都是北京。

(北京は中国の首都である=中国の首都は北京である。)

もう1つは、「A<B」の形、すなわちAはBの一部分(或いはBの下位概念)であり、AとBは帰属的な説明関係である。

・北京是大都市。=北京是大都市之一。

(北京は大都市である=北京は大都市の1つである。)

「是」文にあるAとBとの関係が同等関係であるか、それとも帰属関係であるかは、シソーラス辞書により判定できる。AとBが上下関係であるとすれば、「A<B」となり、上下関係でないとすれば、「A=B」となる。

「是」で構成する文は、一般の動詞の文法特徴とは異なる。「是」は述語ではあるが、それ自身が文の意味的中心になることはない。この場合、中心は目的語であり、「是」の後に置かれる名詞である[5]。従って、複文にある照応詞として、省略されたり、代替されたりするのは「是」の前にある名詞であることが多い。「是」の後にある名詞は、新しい情報源となるので、殆ど省略されない。

又、「是」と同じく、説明・被説明関係を結ぶ動詞もある。それは次の自動詞[6]である。「当」；「做」；「成」；「兼」；「属于」；「等于」；「像」；「算」；「分」；「姓」；「叫」。これらの自動詞は、機能が「是」と似ていて、前の主語と後の目的語は名詞や名詞に

相当するものよりなり、それらの間では、同等である「A=B」か、又は帰属である「A<B」の関係をしている。

そこで、「A=B」文である場合、固有名詞である候補名詞を優先し、「A<B」文である場合、下位概念である候補名詞を優先とするという原則に基づき、照応規則C類を作成した。

[C04] case : N1【当, 成, 兼, 像, 叫】N2

I : N1<N2

II : X=N1=MC1.SUB

THEN : X=先行詞

#### 4. 2 「定中フレーズ」を用いる規則の構築

中国語では、修飾・被修飾関係を表すものは文ではなく、フレーズである。各フレーズの中には、修飾・被修飾関係を持つフレーズ「定中フレーズ」がある。「定中フレーズ」とは、修飾語が「定語(連体修飾語)」であり、被修飾語が「中心語」であるフレーズを指す。中心語が必ず名詞であるのが「定中フレーズ」の特徴である[7]。定語は、名詞である中心語を修飾するものなので、いろいろな品詞やフレーズからなり得る。その中で、定語も名詞である「定中フレーズ」は、その他の「定中フレーズ」とは異なり、互いには、それぞれの意味属性により、「譲渡不可能関係」と「譲渡可能関係」が存在する。

譲渡不可能関係：

- a 張三的手(張三の手)
- b 張三父母(張三の父母)
- c 手上的傷痕(手の傷)

譲渡可能関係：

- a 張三的手絹(張三のハンカチ)
- b 汽車輪胎(車のタイヤ)
- c 手上的泥(手の上の泥)

「定中フレーズ」の主な機能は、文中の主語や目的語となることである。それが先行詞として使われるとき、その後の照応詞が、「定中フレーズ」の一部を取るか、それともその全体を取るかは、「定中フレーズ」の譲渡関係により異なる場合がある。又、一部分を取る場合には、前の定語部分を取るか、それとも後の中心語部分を取るかが、名詞と名詞の譲渡関係による場合がある。例えば、

- ・ (張三的) 手 破了, 𐄂 (很) 痛。  
(張三の手がけがをした。痛い)

定中フレーズである「張三的手」は、文の主語として使われている。けがをした「手」は身体の一部なので、「張三」と「手」は譲渡不可能関係である。しかしながら、「痛(痛い)」と感ずるのは「張三」

その人である。従って、省略された照応詞「 $\phi$ 」は「張三」を同定するはずである。

本研究では、定語も中心語も名詞である「定中フレーズ」に存在する譲渡可能関係と譲渡不可能関係に基づき、照応規則D類を作成した。

[D01] case : NP=N1 {的} N2《譲渡不可》

I : N1=+HUM.all=ATT

II : N2=+HUM.all

THEN : N1+N2=先行詞

名詞間の譲渡関係は、名詞の意味属性により判定する。ここでは、all (全体、完成体) と part (局部、部品) の二つの意味属性も予め与えられているものとする。以下は判定手続きの例である。判定手続きの条件を満たすならば譲渡不可能関係であり、満たさないならば、譲渡可能関係であると判断する。この判定手続きは規則ベースに置き、必要に応じて照応規則より先に動作する。

譲渡不可能関係を判定する手続き：

- 1) NAME+ {的} +N(+HUM.all) (氏名+ {的} +完成体としての人間を表す名詞)。  
例：張三的兒子(張三の息子)。
- 2) personal pronoun+ {的} +N(+HUM.all) (代名詞+人間完成体)。  
例：我母親(私の母親)。
- 3) N1(+ANI.all) {的} +N2(=part(N1)) (有生名詞完成体+同じ名詞の局部)。  
例：張三的手(張三の手)，猫爪子(猫の手)。
- 4) N2がN1から離れない，移せない。  
例：手上的傷痕(手の傷)。

## 5 照応規則の使用

本研究の照応規則ベースは、中国語文内における照応現象を解析するために構築されたものである。使用の前提条件は、先行詞が文中に存在することである。また、照応解析は構文解析が終わった段階からスタートする。それは、入力文がどの照応規則に当てはまるかという判定は、構文解析段階から取り出された情報によるからである。普通、入力文に存在した照応現象は、有形照応詞(代名詞)であれば、代名詞の意味属性を用い、候補名詞とマッチングする方法と、無形照応詞(ゼロ照応)であれば、述詞の意味属性を用い、候補名詞とマッチングする方法で解析する[4]。代名詞でも述詞でも解析できない場合、本研究の照応規則を用い、解を求める。

・馬林又取出一筒牙膏， $\phi$ 也是不大時興的老名牌。(馬林さんは又歯磨きを取り出したが， $\phi$ やはりあまり人気のない古いブランド製品である)。

この代名詞でも述詞でも解析できない複文は、構文解析段階から取り出した情報により、照応規則[A06]に当てはまることが分かる。従って， $\phi$ は、候補名詞「馬林」ではなく，次のような，候補名詞「牙膏(歯磨き)」であるという結果が得られる。

・馬林又取出一筒牙膏，牙膏也是不大時興的老名牌。

(馬林さんは又歯磨きを取り出したが，歯磨きもあまり人気のない古いブランド製品である)。

## 6 結 言

本研究は、文内における照応現象が複文の構文構造や、複文に存在する関連語句、文成分間の意味関係などに依存する特質を用い、照応規則を全部で31個作成した。それらの規則はA, B, C, D四種類に分けられた。本研究の照応規則ベースは、A01からは構文関係，B01からは関連語句，C01からは文成分間の関係(文型)，D01からは文成分間の関係(フレーズ)というようにタイプ別に番号つけられている。又照応規則ベースはオープンにしてあるので、修正、補充は容易である。尚、本研究では、「上下概念」や、「譲渡関係」等の判定に必要なパラメータは、予め辞書項目に与えられていることを前提としている。しかし、実際には、このような判定自体にも、ある程度の文の意味処理に踏み込まないと解決出来ない問題が含まれている。本研究では、先行詞の決定を目的とする為、このような関係の判定問題自体には深く立ち入らない。

本研究で提案した照応規則で解決する方法は、中国語の特徴にあわせて作成したものである。中国語文内照応を解析するには非常に役立つと期待できる。

参考文献：

- [1] 胡状麟，“語篇的銜接與連貫”，上海外語教育出版社，1994。
- [2] 張伯江，方梅，“漢語功能語法研究”，江西教育出版社，1996年。
- [3] 水谷静夫，石綿敏雄，萩野孝野，“文法と意味”，朝倉出版社，1988年。
- [4] 于素秋，横山晶一，西原典孝，“中国語三人称代名詞照応現象解析の一手法”，言語処理学会第4回年次大会発表論文集，言語処理学会，pp39-42，1998年。
- [5] 黄国文，“語篇分析概要”，湖南教育出版社，1988年。
- [6] C.E.ヤーボントフ，“中国語動詞の研究”，白帝社，昭和62年。
- [7] 劉月華，“現代中国語文法総覧”，くろしお出版，1996年。