

中国語文解析における文型パターンの作成と その利用について

邵 桂鳳 鎌田 清一郎 河口 英二 安在 弘幸

九州工業大学 工学部

本論文では、動詞と共起する格要素に着目して動詞を分類し、それに基づいた中国語文の文型パターンの設定について述べる。特に、中国語文解析において重要である“是字文”、“存在文”、“把字文”、“被字文”、“連動文”と“兼語文”について詳しく述べる。本研究で設定した文型パターンを利用して中国語文解析実験システムを試作した。システムでは、解析しようとする文を、述語動詞の類別とその文型パターンを用いて語句の構文的、意味的多義の解消を図りながら、対応するモダリティと格要素を抽出して深層構造を生成する。実験を通して、文型パターンが中国語文解析にとって重要な役割を果たすことが確かめられた。

Chinese Sentence Pattern and its Application

Guifeng Shao Seiichiro Kamata Eiji Kawaguchi Hiroyuki Anzai

Kyushu Institute of Technology

In this paper we describe a Chinese verb classification in relation to case elements, and present the Chinese sentence patterns which we settled according to the classification. We study the patterns of shi-sentence, existential sentence, ba-sentence, bei-sentence, liandong-sentence, and jianyu-sentence in detail. Those patterns are very important for Chinese sentence analysis. We implemented an experimental system of Chinese sentence analysis using those sentence patterns. The experiment worked well and we knew that the sentence patterns play an important role in the Chinese sentence analysis.

1 はじめに

我々が全く知らない外国語を勉強する際に取れる一般的な方法は、与えられた文を文型パターンとして捉え、簡単な文型パターンから複雑な文型パターンまでを徐々に覚えることである。同様に、計算機を使って、自然言語文を解析する時、文型パターンは重要な役割を果たす。英語については Hornby の文型パターンがあり [1]、英日機械翻訳では、文型パターンが重要な役割を果たすことが多い [2]。英語の語順と類似する中国語についても同じことが言える。中国語は非形態言語で、品詞の形態変化はなく、文法的な単語機能は語順に依存し、各種のモダリティがマーカーによって示される。このような特徴に基づいて、中国語を解析する時、構文を一定のパターンにまとめておくことは合理的であると思われる。

これまで、中国語文解析において、その前処理としてヒューリスティックな知識の利用 [3] や、複合文の生成 [4] などの手法が提案され、解析文法としては拡張遷移網 (ATN) によって記述される中国語の文法 [5] や、一つの文をいくつかのブロックに細分化して、順序パラメータを用いて語順を制限する中国語の規制文法 [6] などが提案された。しかし、中国語の一般的文型パターンを捉える研究はまだ見あたらない。今後計算機上での中国語処理を考えれば、その文型パターンを整理しておくことは意義深い。

中国語の文は、述語の性質によって、a) 名詞述語文、b) 形容詞述語文、c) 動詞述語文の三つの種類に区分できる。このうち、a) と b) は簡単であり、本研究では、四つの文型パターンにまとめた。一方、動詞述語文は中国語文の大部分を占めている。文の構造はその文の動詞の統語論的特徴によって、著しく限定される。本研究では、中国語の動詞を統語論的特性によって分類し、動詞述語文を 26 種類の文型パターンで捉え、これらの文型パターンを用いて中国語解析の実験システムを試作した。以下、2、3 章では、動詞の分類と、それに基づいて定めた文型パターンを説明する。4

章では、実験システムについて述べる。第 5 章は結びである。

2 動詞の分類

動詞は自然言語文の核心である。文の構造はその文の動詞の統語論的特徴によって、著しく限定される。それゆえ、中国語の動詞は、多くのグループに分かれており、そのグループには、それぞれ共通な文法的特徴があり、それが各グループの特色をなしている。この節では、統語論的特性によって動詞を以下のように分類する。

(1) “是” 動詞

“是” は英語の be 動詞に相当し、中国語の判断文を構成する。中国語は孤立語であるから、英語 be 動詞のような is, are, were などの変形はなく、どんな人称の主語、どんなテンスに対しても、すべて“是”の形である。

(2) 繫合動詞 (v1)

繫合動詞は文の中で二つの名詞句を連結し (3 章の sp5 を参照)、この二つの名詞句の間の関連が一時的な性格のものであることを示す。繫合動詞に含まれるものには、

像 (xiàng) (～ようである、～に似ている)

叫 (jiào) (～と称せられる)

姓 (xìng) (～姓が～である)

などがある。

(3) 存在動詞 (v2)

中国語には二つの存在動詞がある。

在 (zài), 有 (yǒu) (～ある、～いる)

この二つの単語の使い分けは 3 章の文型 sp19, sp20 のところで説明する。

(4) 自動詞 (vi)

・一般自動詞 (vi1)

この種類の動詞は直接主語の後について文の述語となり、一番簡単な動詞述語文を構成する (例文 e11, e12 を参照)。

・運動を表す自動詞 (vi2)

この種類の動詞は名詞句からなる修飾語を取るが、この名詞句は動作の主体が、それに沿って運

動する場所か、運動の終着点か、運動の出発点かの、いずれかを意味する(例文 e13-e15 を参照)。

・自然現象を意味する自動詞(vi3)

下(xià)(雨) (雨が降る)

結(jié)(氷) (氷がはる)

出(chū)(太陽) (太陽が出る)

などの自然現象を意味する動詞はある種のタイプの名詞句 np とそれぞれ特有に共起する(例文 e16, e17 を参照)。

(5) 他動詞(vt)

・一项目的語を取る動詞(vt1)

唱(chàng)(歌う)

写(xiě)(書く)

などの動詞は文の中で主語と目的語の間に保たれる関係を表す。

・与奪動詞(vt2)

中国語には、間接目的語(人物を表す言葉)と直接目的語(物体や物質を表す言葉)という、二つの目的語を同時に要求する動詞がいくつかある。これらの動詞の大多数は、様々なニュアンスを持った(誰かに、何かを)“与える”という概念や、(誰かから、何かを)“とる”という概念を意味する。それゆえ、このグループ全体を、与奪動詞と名付ける。このグループに含まれるものには：

給(gěi)(与える)

送(sòng)(送る、贈る)

借(jiè)(貸す)

などがある。これらに対応する構文は3章の文型パターン sp10, sp11 である。

・強制動詞(vt3)

このグループの動詞はほかの動作を遂行せしめ、それらの動作の実現を助け、または妨げるか、または少なくともこれらの動作の遂行を認め、許すような動作を意味する。強制の意味を有する動詞に含まれるものには：

請(qǐng)(何かをするように頼む、依頼する)

禁止(jìnzhǐ)(禁じる)

派(pài)(何かをするために派遣する)

などがある。

このグループの動詞は、人物を意味する間接目

的語と、動作動詞からなりかつ間接目的語によって示される人物の動作を意味するところの補語を支配する。この間接目的語は、前の強制動詞の対象語であると同時に、後ろの動作の主体でもある。従って、この種の動詞の文型(3章の sp29)を兼語式と言う。

・思想、感覚、思惟動詞(vt4)

このグループに属する動詞は、(a) 各種の精神活動、(b) 感覚器官の活動、(c) 様々な情緒などを意味する。このような動詞には：

知道(zhīdào)(知る)

想(xiǎng)(思う、考える)

觉得(juéde)(感じる)

などがある。構文上の特徴として、この種の動詞は、一つの文全体を目的語として取るか、名詞や代名詞からなる目的語を取るか、さもなければ全く目的語をとらないかのいずれかである。この種の動詞とともに用いられる目的語は、動作の影響のもとに変化する事物ではなく、動作の主体の意識に反応するか、もしくはその意識にある感覚をおこさせる事物や現象を意味するものである。

3 中国語文の文型パターン

3.1 一般の構造文

(1) 名詞述語文の文型は sp1 である。その名詞述語は天候、年月日、時間、または出身地等を表す。

sp1 主語 + 名詞述語

e1 昨天 星期三。(昨日は水曜日だった)

e2 我 二十歳。(私は二十歳です)

(2) 形容詞述語文には次の三つの文型がある。ただし、“状”は状況語の略で、形容詞を修飾する副詞性成分を表す。

sp2 主語 + (状) + 形容詞

e3 張三的太太漂亮。(張の奥さんは美しい)

e4 行李 太多了。(荷物が多すぎる)

sp3 主語 + (状) + 形容詞 + 補語

e5 天氣 晴朗 起来了。(天氣は晴れてきた)

e6 樹葉 綠 得發亮。(木の葉が輝くほど緑だ)

sp4 主語 + (状) + 形容詞句

e7 這声音 非常 微弱, 低沈.

(この声は非常に微かで, 沈んでいる)

e8 晚上的海風 即清新 又涼爽.

(夜の海風はすがすがしくて涼しい)

(3) 動詞述語文

(a) 聚合動詞述語文

sp5 主語 + v1 + np

e9 那座寺院叫白雲觀. (あの寺院は白雲觀という)

e10 他 不姓 王. (彼の姓は王ではない)

(b) 自動詞述語文

sp6 主語 + vi1

e11 我們 呼吸 (私たちは呼吸する).

e12 春天 来了 (春が来た.)

sp7 主語 + vi2 + np(場所/出発点/終着点)

e13 昨天張三回国了. (張三は昨日国に帰った)

e14 王教授去了東京. (王教授は東京へ行った)

e15 他 没来 学校. (彼は学校にきていない)

sp8 vi3 + 從屬主語

e16 出 彩虹了 (虹が出た).

e17 外面下着雨 (外は雨が降っている).

(c) 他動詞述語文

sp9 主語 + vt1 + 目的語

e18 那个人回答了我的問題.

(あの人は私の質問を答えた)

e19 母親在看書. (母親は本を読んでいる)

sp10 主語 + vt2 + (給) + IO + DO

e20 老師 送 給了 小王 二本書.

sp11 主語 + vt2 + DO + 給 + IO

e21 老師 送了 兩本書 給 小王.

訳 e20,e21 先生は王さんに二冊の本を与えた.

注: DO 直接目的語, IO 間接目的語.

sp12 主語 + vt4 + (np/文)

e22 他 知道了 (彼は知っている).

e23 他 知道了 這件事.

(彼はこのことを知っている)

e24 我 知道 他生气了.

(私は彼が怒ったのを知っている)

(4) 二重主語文

二重主語文では, 二つの名詞句 (np) は述語の前に連続して起こる. sp13 はその文型である. 名詞句 np1 と名詞句 np2 は所有者—被所有物の関係を持っている. 文法的には, 文の述語は np2 に呼応する. この種の文の述語になれるものは形容詞と自動詞だけである.

sp13 np1 + np2 + 述語

e25 張三 頭 痛 (張三は頭が痛い).

e26 中国 人口 多 (中国は人口が多い).

3.2 特殊な構造文

1. 是字文 (shi-sentence)

述語が動詞“是”である文は是字文である. 是字文は次の三種に区別されている.

(a) 連結文: 二つの名詞句が“是”によって連結されている文である.

sp14 np1 + 是 + np2

e27 中国的首都是 北京. (中国の首都は北京だ)

e28 魯迅 是 紹興人. (魯迅は紹興人だ)

e29 馬 是 動物. (馬は動物だ)

np1 と np2 との関係は次のいずれかである.

1. np1 と np2 が同一個体を指示する

(np1=np2, 例文 e27).

2. np1 が np2 という類に属する

(np1 ∈ np2, 例文 e28).

3. np1 のすべての成員が np2 の成員に含まれる

(np1 ⊂ np2, 例文 e29).

(b) “的”を含む判断文

sp15 np + 是 + dp + 的

e30 張三 是 昨天来的.

(張三が来たのは昨日だ)

sp16 dp + 的 + 是 + np

e31 昨天来的 是 張三.

(昨日来たのは張三だ)

sp17 是 + np + dp + 的

e32 是 我 請張三来的.

(私が張三に来てもらった)

ここで, dp は動詞成分 (単独の動詞, 各種の動詞構造, 及び動詞を述語とする主述構造を含む) を表す。

(c) “是”を含む存在文

sp18 場所詞 + 是 + np

e33 門前 是 草坪。(門の前は芝生だ)

“是”の後のものは“是”の前の場所に存在する。

この文型の文は一種の描写文である。一方, 次に定義する存在文は叙述文である。

2. 存在文 (existential-sentence)

sp19 場所格 + 有 + np

e34 学校里有圖書館。(学校の中に図書館がある)

e35 客厅里有客人。(応接間にお客がいる)

sp20 主語 + 在 + 場所格

e36 他 在 老師的家里。(彼は先生の家にいる)

e37 圖書館在学校里。(図書館は学校の中にある)

“なにになにがある(いる)”は“有”+“なにになに”となり, “どこそこにある(いる)”は“在”+“どこそこ”となる。“有”の後には場所格は用いられないし, “在”の後には場所格だけしか用いられない。

また, さらに人なり事物なりが, 如何なる状態において存在しているかを表す場合, 次のパターンがある。

sp21 場所格 + vil/vt1 + 着 + np

e38 院子里 聚集 着 許多蜜蜂。

(庭にたくさん蜜蜂が群がっている)

e39 門前 停 着 一輛車。

(門の前に一台の車が止まっている)

この“着”のはたらきは, 動詞が完成して, 動作の結果が持続していることを示す。このパターンは動作の結果の状態を強調する。もしただ動作に関するイベントを叙述するならば次のパターンを使う。

sp22 np + vil/vt1 + 在 + 場所格

e40 許多蜜蜂 聚集 在 院子里。

(たくさん蜜蜂は庭に群がっている)

e41 一輛車 停 在 門前。

(一台の車は門の前に止まっている)

3. 把字文 (ba-sentence)

把字文の構造は処置式 (disposal construction) とも言う。sp23 はその文型パターンである。文型 sp23 の補語には回数を示す数量詞の補語 (例文 e42), 前置詞句の補語 (例文 e43), 程度の補語 (例文 e44) が含まれる。このタイプの文は, しばしば, 動詞がある要素に修飾されている場合に, “把”によって目的語を前に引き出す構造である。

sp23 主語 + 把 + np + vt + 補語

e42 警察 把 小偷 打了一頓。

(警察はこそ泥を一発殴った)

e43 張三 把 郵票 貼 在信封上了。

(張三は封筒に切手を貼り付けた)

e44 王教授 把 那個問題 解決 得非常好。

(王教授はあの問題を非常によく解決した)

sp24 は sp10, sp11 の別の言い方のパターンである。

sp24 主語 + 把 + DO + vt2 + 給 + IO

e45 老師 把 兩本書 送 給了 小王。

(先生は王君に二冊の本をあげた)

把字文の否定においては, 否定詞“不”あるいは“没有”はいつも“把”の前に現れる。

4. 被字文 (bei-sentence)

被字文の構文は受動態 (passive construction) を表し, sp25, sp26 の形をしている。

sp25 主語 + 被 + np + vt

e46 張三的錢 被 人 偷了。

(張さんのお金は人に盗まれた)

e47 我的心 被 妹妹 哭乱了。

(妹が泣いて, 私の心は乱された)

sp26 主語 + 被 + np + vt + np/補語

e48 張三 被 人 偷了 錢。

(張さんは人にお金を盗まれた)

e49 我 被 人 哭乱了 心。

(妹が泣いたから私は心をかき乱された)

e50 他 被 警察 閔押了 三个小時。

(彼は警察に三時間拘留された)

sp26 での主語は他動詞 vt の後の名詞句との関

係は“所有者”と“被所有物”の関係である。

把字文と被字文に現れる動詞と補語はいずれも極めてよく似ている。“思想，感覚，思惟動詞”(例えば“看見(見かける)，知道(知る)”など)及びその他の若干の動詞(例えば“信任(信用する)，尊敬(尊敬する)”など)が受動態のみに用いられて処置式に用いられないことを除けば，受動態に用いることのできる動詞はいずれも処置式に用いることができる。

5. 連動文 (liandong-sentence)

連動文では二つの動詞句 vp は主語の後に連続して起こり，文型 sp27,sp28 でまとめられている。

sp27 np + vp + vp

e51 小王 吃了早飯 起程了。

(王君は朝ご飯を食べて出発した)

e52 他 乘飛機 回東京。

(彼は飛行機で東京に帰った)

e51 では前の動作が終ってから後の動作は行ふ。e52 では前の動作は後の動作の方式を表す。

sp28 np + (没)有 + np + vp

e53 我們 有 機會 再見面的。

(我々はまた合うチャンスがあるよ)

e54 他 没有 力量 控制自己的感情。

(彼は自分の感情を抑える力がない)

6. 兼語文 (jianyu-sentence)

sp29, sp30 は兼語文の文型である。兼語文は二つの動詞句 vp のリンクとしてはたらく兼語の np1 があるのが特徴である。この兼語 np は，前の動詞の目的語，後ろの動詞の主語として両様にはたらくものである。

sp29 主語 + vt3 + np + vp

e55 我 勸 他 不要喝酒。

(私は彼に酒を飲まないように忠告する)

e56 張三 叫 李四 別難過。

(張三は李四に悲しむなと言った)

sp30 主語 + (没)有 + np + vp

e57 這里 没有 人 關心他的問題。

(ここでは彼の事に関心を持つ人はいない)

4 文型パターンに基づく中国語解析の実験システム

文型パターンの応用として中国語解析の実験システムを構築した。システムの入力は単語ごとに分かち書きされた中国語文である。システムは，第3章で述べた文型パターンに基づく構文規則を用いて，左隅統語解析法 [8] で入力文を解析する。その結果に対して，述語に注目して，モダリティと動作主，対象語などの命題情報を抽出して深層構造を生成する。

4.1 構文解析部

構文解析部は入力文を解析してどの文型パターンかを判断する。図1は例文 e58 の構文木である。例文 e58 昨天我們在研究室沒有討論這個問題。(昨日我々は研究室でこの問題を討論しなかった)

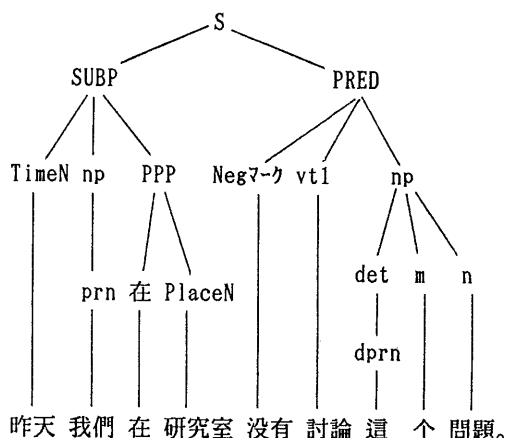


図1. 例文 e58 の構文木

単語の多品詞性に起因する解析の曖昧性を減らすために次の方法を利用した。

(1) 多品詞性単語の役割は構文パターンによって決定する。例えば，“在”は多品詞性単語であり，次の三つの構文パターン

相 標 識: “在” + 動詞 (例文 e19)

存在動詞: “在” + 場所格 + . (例文 e36)

前 置 詞: “在” + 場所格 + …… + . (例文 e58)

注: “. ”: 文の終わる記号

を持っている。例文 e58 では“在”の後ろには場所名詞“研究室”があるから、この“在”は相標識でないと判断できる。また、“研究室”の後ろには単語がまだ続いているから、この“在”は前置詞であると判断される。

(2) 相標識、否定マーク (付録を参照) などによって述語の動詞を決定する。相標識と否定マークの前後に多品詞性単語が来た時、その単語が動詞であると判断できる。例えば、単語“討論”は動詞 (vt1) と名詞の二つの品詞性がある。例文 e58 では否定マーク“没有”により、この語はこの文で動詞であると判断する。

4.2 深層構造の生成

構文木からモダリティと格要素を抽出して、入力文の深層構造を生成する。

4.2.1 モダリティの抽出

いま抽出するモダリティは時制、相、否定、受動である。時制、相と否定はそれぞれ付録の表 1、表 2、表 3 に示した語、相標識、マークによって抽出する。受動は被字文の文型パターンから抽出する。

例文 e58 は文型パターン sp9 である。時間名詞“昨天”から過去時制を抽出する。述語の動詞“討論”の前の“没有”により否定というモダリティを抽出する。表 3 によって文型パターン sp9 の過去時制の否定マークは“没(有)”である。例文 e58 から抽出された否定はこれに合致する。

4.2.2 格要素の抽出

名詞述語文と形容詞述語文は動詞述語文と違って、対応する受動文とか命令文などが存在しない。また、述語には動詞がないので、動作主などは存在しない。これらの種類の文に対して、深層構造

を生成する時、主語部分の名詞句と述語部分の名詞句のみを抽出する。

動詞述語文は文型パターンの違いによって含まれている格要素の種類が違う。

存在文は対象格と場所格が必ずある。その対象がその場所に存在する。動作主と経験格は存在文に含まれない。ほかの文型パターンの文に対して、動作主、時間格と場所格は文の主語部分から抽出され、ほかの格要素は文の述語部分から抽出される。文型 sp9 の例文 e58 は次の格要素を持っている。図 2 にこの文の深層構造を示す

動作主: 我們 時間格: 昨天
場所格: 研究室 対象格: 這個問題

動詞 vt1(討論):

モダリティ: 過去, 否定

格要素: Agent : 我們

Object : 這個問題

Time : 昨天

Place : 在 研究室

図 2. 例文 e58 の深層構造

5 結び

文型パターンは文の構造を形式的に表現したものであり、構文解析の時、単語の多品詞性により得られる不適格な解釈の処理を排除することができる。文の文型パターンが分かったら、意味解析ではモダリティと格要素を正しく抽出することができる。例えば、

例文 e59. 這位技術員把那問題解決得非常好。

e60. 這位技術員非常好地解決了那問題。

(この技術者はあの問題を非常によく解決した)

この二つの文は述語の動詞が同じであるが、文型パターンが違うから、動詞の対象格は、例文 e59 では動詞の前に存在するのに対して、例文 e60 では動詞の後ろに存在する。文型パターンを利用することによって、格要素の存在する位置を知り、格要素を正確に抽出することができる。文型パター

ンが中国語文の構文的意味的多義の解消ならびに格構造の抽出に相当に有効であることが確かめられた。

名詞句の補文の関係節構文、文の従属構文、疑問文、命令文などの文型は本稿で定義した文型には入っていない。また、実験システムもきわめて小さい。今後の課題として、より多くの例文を考察し、言語現象をできるだけ多く整理し、文型パターンを改善する必要がある。そして、改善された文型パターンを用いて実験システムを充実する予定である。

参考文献

- [1] Hornby, A.S.: "Guide to Patterns and Usage in English", Oxford Univ. Press(1975); 伊藤 健三 訳: "英語の型と語法", オックスフォード大学出版局 (昭 52 年)
- [2] 高松 忍, 西田 富士夫: "パターンと格構造に基づく英日機械翻訳", 信学論 (D), J64-D, 9(1981-09).
- [3] 楊 頤明, 堂下 修司, 西田 豊明: "中国語解析システムにおけるヒューリスティックな知識の利用", 情報処理学会論文誌, Vol.25, No.6(1984).
- [4] 范 莉馨, 任 福繼, 宮永 喜一, 柄内 香次: "中国語文中の複合語の生成について", 情報研報 NLC90-1, AI90-29.
- [5] 邨 曉英: "中国語文解析システムのMYLANGによる試作", 九州工業大学研究報告 (工学), 第 49 号, 昭和 59 年.
- [6] 王 丹雲: "規制文法を用いた中国語のCAI実験システム", (人工知能システムの枠組み) シンポジウム, 昭和 62 年 11 月, pp.101-110.
- [7] 野村 浩郷: "自然言語処理の基礎技術", 社会法人 電子情報通信学会編 (1988).
- [8] 田中 穂積 辻井 潤一: "自然言語理解", オーム社 (1988).
- [9] 邵 桂鳳, 蔣 勝平, 安在 弘幸: "中国語の動詞の文法的な特徴とその抽出処理", 九州支部連合大会論文集, pp.665, (1990).
- [10] 土屋 申一: "基礎中国語", 東京/大学書林 (1986).

付録

表 1 : 時制を表す語の例

過去:	已經(すでに), 過去(昔), 原来(もともと), 昨天(昨日), 剛才(先ほど), …了(した), …
現在:	正在(ちょうど…しつつある), 現在(いま), …
将来:	明天(明日), 将来(将来), 過後儿(あとで), 下個月(来月), …

表 2 : 相標識

相	相標識	例
起動前相	快/要…了, 正要…	正要読(読むところ)
起動相	…起来, 開始…	開始読(読み始める)
起動後相	剛/才…(起来), (剛/才)(開始)	剛開始読 読み始めたところ
進行相	正(在)…, …着	正在読(読んでいる)
完了前相	快…(好/完)了	快読完了 (読み終わるところ)
完了相	…(過/好/完)了	読完了(読み終わる)
完了後相	剛/才…過/好/完	剛読完 読み終わったところ
結果相	…着	開着(開けている)

表 3 : 否定のマーク

文型パターン	否定マーク	文に現れる位置
sp2, 5, 13	不	述語の前
sp14-18	不	"是"の前
sp1	不是	主語の前
sp4	不, 不是	主語の後
sp3	没有	主語の後
sp19, 28, 30	没	"有"の前
sp6-12, sp20-27, 29	過去: 没, 没有 現在: 不 未来: 不	把字文: "把"の前 被字文: "被"の前 sp27: vp1の前 sp29: vt3の前 性か: 述語動詞の前