Valence alternations and marking structures in a HPSG grammar for Mandarin Chinese

Janna Lipenkova
Freie Universität Berlin
janna.lipenkova@fu-berlin.de

Abstract

The paper discusses HPSG as a framework for the computational analysis of Mandarin Chinese. We point out the main characteristics of the framework and show how they can be exploited to target language-specific issues, describe existing grammar engineering work for Chinese and present our own effort in the implementation of a grammar for Chinese. The grammar is illustrated with two fields of phenomena, namely semantic and syntactic marking and valence alternations. We aim at the integration of work in theoretical linguistics into computational applications in order to complement statistical methods and thus increase their accuracy and scalability.

1 Introduction

This paper presents a grammar fragment of Chinese which is built in the framework of HPSG Pollard and Sag (1994) and implemented in the grammar development system Trale Meurers et al. (2002a). We consider the use of the framework from an NLP perspective; at present, large-scale NLP applications are mostly based on statistical and machine-learning methods with a minimum of theoretical linguistic analysis and information. We believe that the use of a more powerful formal theory in combination with machine learning and induction methods will significantly increase the accuracy of the existing systems and reduce the gap between theoretical linguistics and NLP.

The advantages of the HPSG framework for the computational analysis of Chinese are as follows:

- HPSG provides a range of powerful formal tools for the description of linguistic expressions which are embedded into the model-theoretical framework of *Typed Feature Structure Logic* Carpenter (1992) and allow a seamless implementation in logical programming paradigms.

- HPSG minimizes the use of theory-internal statements about the empiricial properties of linguistic signs. Since Chinese is a language that cannot be straightforwardly explained using the terminology and assumptions of the Western linguistic tradition, HPSG thus provides us with a 'neutral' framework for the formalization of language-specific phenomena based on which more general principles can be derived.

- In contrast to most formal theories, HPSG is not a syntax-first framework; the different levels of linguistic representation – phonology, syntax, semantics, pragmatics – have equal weight. This is especially beneficial for Chinese, which has a poor morphological system and exhibits a high degree of surface ambiguity. The use of a powerful semantic-pragmatic module with fine-grained definitions of semantic types and selectional restrictions and preferences significantly helps disambiguation.

In the following, we first introduce the basic feature architecture and principles of the grammar formalism. Then, we review existing work in HPSG and grammar development for Chinese. The last part contains a synopsis of the covered phenomena; the main analytical choices are exemplified by treatments of two groups of phenomena, namely valence and marking.

*This research is supported by the Deutsche Forschungsgemeinschaft under the grant nr. DFG (MU 2822/5-1).

International Joint Conference on Natural Language Processing, pages 27–35,
Nagoya, Japan, 14-18 October 2013.

2 Framework and implementation

This part provides a very brief overview over the main principles and components of HPSG; for more detailed expositions, the reader is referred to the standard makeup described in Pollard and Sag (1994) as well as Pollard and Sag (1987), Müller (2008) and Sag et al. (2003). The two main systems used for grammar engineering with HPSG are Trale Meurers et al. (2002b); Müller (2007) and LKB Copestake (2002); the implementation presented in this paper uses the Trale system. The semantics follows Minimal Recursion Semantics as described in Copestake et al. (2005).

The main characteristics of the HPSG framework are as follows:

- *Feature-based*: the universal format of representation are typed feature structures Carpenter (1992).
- *Constraint-based*: generalizations on linguistic objects are formulated as declarative constraints; there are no transformations.
- *Lexicalist*: a great part of linguistic information, especially information about syntactic combination, is stored in the lexicon.
- *Monostratal*: multiple levels of linguistic representation (phonology, syntax and morphology, semantic, pragmatics and information structure) are modelled in parallel; no formal priority is given to the structural level.

Formally, an HPSG grammar consists of three parts:

- *Signature*: type hierarchy with feature specifications for the types
- *Lexicon* (constraints on linguistic signs of type *word*):
  - Lexical entries
  - Lexical rules, specifying systematic relationships holding between classes of lexical items
- *Grammar* (constraints on linguistic signs of type *phrase*): constraints on linguistic objects of type *phrase*
  - Small set of broad-range principles holding of large subtypes of *phrase* (Head Feature Principle, Subcategorization Principle, Semantics Principle)
  - Immediate dominance schemata, specifying the constituency of phrases
  - Linear precedence rules, specifying linear constituent order

A linguistic sign is modelled with feature structures built according to a standardized architecture. The feature structures are sets of feature-value pairs; the value of a feature is either atomic or in itself a feature structure. The signature specifies the types of values acceptable for a feature. The following figure shows the basic formal setup of a linguistic sign:

(1)
$$\begin{bmatrix} \textit{sign} \\ \text{PHON } \textit{list(phon-symbol)} \\ \text{SYNSEM} \begin{bmatrix} \text{LOC} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } \textit{head} \\ \text{SPR } \textit{list(synsem)} \\ \text{COMPS } \textit{list(synsem)} \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \text{INDEX } \textit{index} \\ \text{RELS } \textit{list(rels)} \end{bmatrix} \end{bmatrix} \\ \text{NONLOC } \textit{nonloc} \end{bmatrix} \end{bmatrix}$$

At the highest level, there is a separation between the phonological and the syntactic and semantic properties of the sign. This separation is relevant with respect to syntactic selection: heads may only specify the SYNSEM properties of the signs they select. SYNSEM is divided local and nonlocal properties; NONLOC being reserved for the modelling of long-distance dependencies, our following exposition mainly focusses on the LOC feature. LOC contains syntactic and semantic properties (CAT(EGORY) and CONT(ENT), respectively); CAT specifies the part-of-speech specific HEAD features which are propagated by a lexical head to the mother node. It also contains the valence features SPR and COMPS, which specify the valents of the sign. The CONT attribute contains an index variable, which identifies a referential or situational argument, and a set of relations specifying the semantic contributions of the lexical items that compose the sign.

3 Previous work

In this section, we give an overview of the work done so far in HPSG for Chinese. On the one hand, since the 90's, several studies have provided theoretical HPSG analyses of specific phenomena of Chinese. Formal treatments have been proposed for the NP (Gao, 1993; Xue and McFetridge, 1995; Ng, 1999), serial verb constructions (Lipenkova, 2009; Müller

and Lipenkova, 2009) and the well-known *bǎ*-construction (Ding, 2000; Lipenkova, 2011). Besides, two dissertations, namely Gang (1997) and Gao (2000), provide overall sketches of HPSG grammars for Chinese.

On the other hand, there are two ongoing efforts in grammar development for Chinese, presented in Wang et al. (2009); Yu et al. (2010) and Zhang et al. (2011, 2012). Both are oriented towards a large-scale data-driven grammar implementation; they attempt to stay close to the original version of the framework and minimize the use of language-specific postulates. Our grammar aims to complement these efforts and to refine some of the analyses by grounding them on findings from recent descriptive and theoretical research.

3.1 Joint grammar and treebank development

Zhang et al. (2011) and Zhang et al. (2012) use the HPSG framework to combine grammar engineering and treebank compilation. The grammar mainly builds on a part-of-speech hierarchy and the valence specification. The assumed part-of-speech hierarchy is similar to the classification presented in Pollard and Sag (1994). There are two valence features for the arguments of predicates, namely SUBJ for subjects and COMPS for complements. Elements on the COMPS feature have a boolean feature which specifies whether they appear to the right or to the left of the head.

Besides basic clause structures, the grammar covers the structure of NPs and locative phrases, topic constructions, coverbs, resultative verb compounds and simple *bǎ*- and *bèi*-constructions.

3.2 HPSG grammar and treebank conversion

In Wang et al. (2009) and Yu et al. (2010), the authors adopt a data-driven approach with the aim of developing a HPSG parser for Chinese. Starting out with a small set of assumptions about the grammar (sign structure, grammatical principles and schemata), they manually convert a Chinese treebank into an HPSG treebank; the resulting treebank is used for the extraction of a large-scale lexicon of Chinese.

The analyses are based on a rather informal assumption of three levels of sentence structure which is borrowed from traditional Chinese linguistics, namely the predicative part, the simple and the complex sentence. The 'predicative' part contains a verb with its objects and complements. At the simple sentence level, the authors distinguish between sentences with a subject and subject-less sentences, which correspond to a 'standalone' predicative part. Complex sentences subsume coordinated sentences, sentences containing serial verb constructions and topic sentences.

In order to provide a formal representation of these sentence structures, the authors use nine phrase structure schemata which determine both constituency and linear order. There are two predicate-argument schemata which are used for the combination of the verb with its arguments. It is assumed that the subject appears before the verb (*specifier-head* schema), whereas the object appears after the verb (*head-object* schema). Adjunction is also analyzed with two schemata, namely the *modifier-head* schema and the *head-modifier* schema, which differ only in linear order. Problemtatically, postverbal elements marked by 得, which have a modifier semantics but syntactically behave on a par with complements, are also analyzed via the *head-modifier* schema; the proposed framework does not provide a schema for the syntactic analysis of these structures as complements.

Further, a *filler-head* structure is proposed for structures with unmarked object preposing. There are two varieties, namely the *pre-object-as-subject* schema, which is used for unmarked passives (2a), and the *pre-object-as-topic* schema which is used for topicalizations (2b):

(2) a. 苹果 吃 了。
Píngguǒ chī le.
apple eat PFV
'The apple was eaten.'

b. 苹果 他 吃 了。
Píngguǒ tā chī le.
apple he eat PFV
'The apple, he ate it.'

However, in HPSG, the *filler-head* schema has been posited for the analysis of nonlocal dependencies (Pollard and Sag, 1994, p. 164). The motivation behind its use for the analysis of unmarked passives is unclear.

A general shortcoming of the proposed implementation is the heavy use of different phrase structure schemata in order to capture alternations in valence and linear order. For example, a

predicate-argument structure consisting of a verb and its object can be analyzed via the *head-object* schema if it occurs in the canonical VO order, a *specifier-head* schema if the object is preposed and marked by *bǎ* and a *filler-head* schema if the object is preposed into the sentence-initial position, giving rise to an unmarked passive. In the original version of the framework, information about constituency, linear order and valence is cleanly distributed between immediate dominance schemata, linear precedence rules and lexicon. The proposed set of schemata mixes up these different types of information and thus obscures the original purpose of immediate dominance schemata; at the same time, it does not exploit the expressive power of the lexicon and of linear precedence rules.

4 The coverage of the grammar

Our grammar is contains a syntactic component which specifies linear order and constituency, a lexicon with about 900 lexical items and a number of lexical rules, as well as a set of macros which are used as 'abbreviations' for recurring descriptions of linguistic objects to ease the work of the grammar writer. The grammar is tested against a testsuite which currently contains 300 phrasal and clausal items which represent different constituent and clause structures of Chinese. At present, we are testing the grammar against a larger corpus of empirical examples of the covered phenomena and extending the lexicon and the grammar as new items and structures arise. The phenomena that can be currently analyzed are:

- NP structure:
  - Internal structure, combination with determiners, numerals and classifiers
  - Prenominal modification: adjectival and possessive modifiers, relative clauses with subject, object and adjunct extraction
- Morphological variation: compounding, reduplication, affixation
- Basic clause structures and valence alternations: transitive, intransitive and ditransitive frames; *bǎ*- and *bèi*-construction; serial verb constructions; topic structures; unmarked passives
- Syntactic marking: nominal *de*-adjunction (的); verbal *de*-adjunction (地); *de*-complementation (得)
- Mood and aspect marking
- Locative and temporal adjuncts
- Resultative constructions

In the following, we outline our analyses of two fields of phenomena, namely valence, incl. argument alternations, and marking. We will see that different formal means provided by the framework are suitable for different types of phenomena. Thus, valence is mostly treated in the lexicon, whereas different types of marking are analyzed via one immediate dominance schema, namely the *head-marker* schema.

5 Example analyses

5.1 Valence alternations and argument realization: a lexicalist analysis

The coverage of the valence 'module' is as follows:

- Unmarked verbal frames in active voice (intransitive, transitive, ditransitive), including differentiation between syntactically obligatory and optional internal arguments; these frames are entirely specified in the lexical hierarchy. Using multiple inheritance, verbs with optional internal arguments can inherit from multiple frames.
- Alternations in argument realization (dative alternation, object preposing with and without subject omission) are captured by lexical rules which act on the valence features of the item. The lexical argument structure is not changed and thus remains accessible to semantic mechanisms such as binding.
- *bǎ*- and *bèi*-constructions: we do not adopt the marker analysis of *bǎ* and *bèi* which is mostly adopted in formal studies. As shown in Sybesma (1999), Bender (2000) and Lipenkova (2011), *i. a.*, the role of these morphemes in sentence formation is more prominent than the role of a normal marker: they can be used with a range of different argument distributions, may select their own arguments and impose semantic constraints

on the predicate and its complement which would be difficult to capture in a marker analysis. In our implementation, *bǎ* and *bèi* are analyzed as clausal heads.

In the feature architecture, valence is captured by three list-valued features which contain the valents in order of decreasing obliqueness. All signs have the two valence features SPR (external argument) and COMPS (internal arguments):

(3) $\begin{bmatrix} \textit{sign} \\ \text{CAT} \begin{bmatrix} \text{SPR } \textit{list} \\ \text{COMPS } \textit{list} \end{bmatrix} \end{bmatrix}$

These features contain lists of the elements that the sign must combine with in order to grow into a saturated well-formed phrase. The features are dynamic: the Subcategorization Principle (Pollard and Sag, 1994, p. 35) ensures that elements that have already been realized are deleted from the valence lists at the next higher node.

Lexical items have the additional feature ARG-ST. Its value is a fixed list that specifies the dependents lexically selected by the word. In the 'basic' makeup of a lexical item that has not undergone a lexical rule, ARG-ST is the concatenation of SPR and COMPS:

(4) $\begin{bmatrix} \textit{basic-word} \\ \text{SYN} \begin{bmatrix} \text{SPR } \boxed{1}\ \textit{list} \\ \text{COMPS } \boxed{2}\ \textit{list} \\ \text{ARG-ST } \boxed{1} \oplus \boxed{2} \end{bmatrix} \end{bmatrix}$

Thus, to summarize, ARG-ST tells us which dependents are to be realized, whereas SPR and COMPS determine how they are realized.

Marked valence patterns, such as valence alternations and the *bǎ*- and *bèi*-constructions, require additional machinery. Valence alternations which do not come with additional lexical material to which we could tie structural information are analyzed with lexical rules. The formal properties of lexical rules are described in Flickinger (1987), Briscoe and Copestake (1999), Meurers (2000), Meurers (2001) and Müller (2006), *inter alia*. We use lexical rules for valence reduction (e. g. unmarked passives), valence augmentation (e. g. resultative complements) and valence alternation (e. g. dative shift). The following illustrates a simple lexical rule for the use of transitive verbs in the unmarked passive:

(5) Reduced valence in unmarked passive:

a. 苹果　吃 了。
Píngguǒ chī le.
apple　eat PFV
'The apple was eaten.'

b. Lexical rule for valence reduction of *chī*:
$\begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{SPR } \langle \boxed{1} \rangle \\ \text{COMPS } \langle \boxed{2} \rangle \\ \text{ARG-ST } \langle \boxed{1}, \boxed{2} \rangle \end{bmatrix} \rightarrow \begin{bmatrix} \text{SPR } \langle \boxed{2} \rangle \\ \text{COMPS } \langle\rangle \end{bmatrix}$

We see that the output of the lexical entry accommodates the original complement in the specifier position; the original specifier no longer appears on the valence lists and thus cannot be realized.

Besides valence alternations which are not marked by specific morphology, Chinese has two argument structure constructions with additional lexical material, namely the *bèi*- and the *bǎ*-construction. In these constructions, the morphemes *bǎ* and *bèi* determine the argument structure of the clause. *bǎ* preposes an argument of the verb which appears postverbally in the canonical SVO order (6a). *Bèi* either appears alone or marks the external argument of the verb; in any case, *bèi* promotes its internal argument into the subject position (6b):

(6) a. 他 把 苹果　吃 了。
Tā bǎ píngguǒ chī le.
he BA apple　eat PFV
'He ate the apple.'

b. 苹果　被 (他) 吃 了。
Píngguǒ bèi (tā) chī le.
apple　BEI he　eat PFV
'The apple was eaten (by him).'

These constructions are only used with transitive and ditransitive predicates that describe events loosely associated with the semantic concept of 'affectedness'.

Whereas the structual scope of a normal marker is limited to the element it marks, the use of *bǎ* or *bèi* impacts on the overall formation and well-formedness constraints on a sentence. In order to accommodate this information, we analyze *bǎ* and *bèi* as heads which select for an almost saturated verbal complement and 'attract' the yet unrealized argument of that complement in order to realize it in the sentence-initial position. Thus, *bǎ* canonically attracts the external argument, whereas *bèi* attracts the internal argument.

Note that the output of the lexical rule only specifies features whose value is changed by the rule.

The following structure shows a part of the lexical entry for *bǎ*:

(7) $\left[\begin{array}{l}\text{PHON}\langle bǎ\rangle \\ \text{SYN}\left[\begin{array}{l}\text{SPR}\langle \boxed{1}\rangle \\ \text{COMPS}\left\langle \text{V}\left[\begin{array}{l}\text{SPR}\langle \boxed{1}\rangle \\ \text{SEM}\ \boxed{2}\end{array}\right]\right\rangle\end{array}\right] \\ \text{SEM}\ \boxed{2}\end{array}\right]$

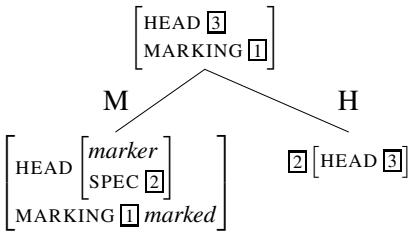
In the canonical case, *bǎ* does not make a semantic contribution; thus, it inherits the content of the verbal complement in order to ensure correct semantic composition at the mother node of the sentence.

Fig. 1 illustrates the syntactic combination for (6a).

5.2 Use of the head-marker structure for different types of marking

As we have said, Chinese has a poor morphology. The lack of expressive force on the morphological level is partially compensated by a rich class of markers. In the following, we distinguish between two kinds of marking; on the one hand, semantic markers mark the aspect of a VP (perfective 了 *le*, durative 着 *zhe*, experiential 过 *guo*) or the mode of a sentence (interrogative 吗 *ma*, imperative 吧 *ba*, change-of-state 了 *le*). On the other hand, syntactic markers make constituents eligible for specific syntactic positions without altering their semantics (three *de*'s: 得, 的, 地).

Structures with markers are analyzed via the *head-marker schema*:

(8)


As part of the HEAD feature, which contains properties specific to a particular part of speech, the marker has a feature SPEC(IFIED) which constrains the marked constituent. Additionally, the feature MARKING ensures that the marker is visible at the top node. This feature takes the value *unmarked* for lexical items that are not markers; for markers, it takes a subtype of *marked*, which subsumes individual values contributed by specific markers.

bǎ also allows for other argument distributions in which it indeed may contribute additional relations and event arguments; cf. **?** and Sybesma (1999), *inter alia*.

5.2.1 Semantic marking

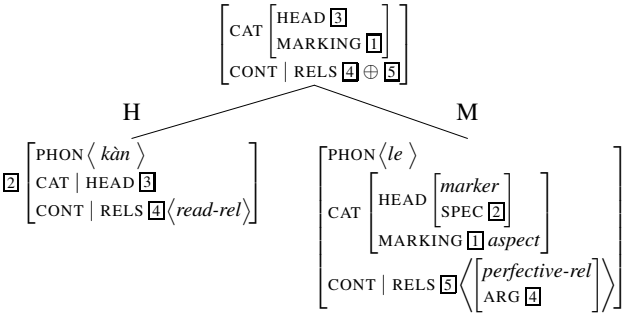
Chinese has three postverbal aspect markers, as illustrated in the following example:

(9) 他 看 了 / 着 / 过 书。
Tā kàn le / zhe / guo shū.
he read PFV / PROG / EXP book
'He read / is reading / once read the book.'

These markers mark the perfective, durative and experiential aspect, respectively. They markers naturally differ in the range of semantic classes of verbs with which they combine; however, the syntactic distribution of aspect markers is identical: they immediately follow the verb. The following AVM shows the supertype constraint for aspect markers:

(10) $\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{HEAD}\left[\begin{array}{l}marker \\ \text{SPEC}\ \boxed{1}\ \text{V}\left[\text{CONT} \mid \text{IND}\ \boxed{2}\right]\end{array}\right] \\ \text{MARKING}\ aspect\end{array}\right] \\ \text{CONT}\left[\begin{array}{l}asp\text{-}rel \\ \text{ARG}\ \boxed{2}\end{array}\right]\end{array}\right]$

With this supertype in place, the entries for the individual markers only specify information about the semantic relation contributed by the marker. Thus, for (9), we get the following combination of the verb with the aspect marker:



Mode markers are analyzed in a similar manner; however, instead of marking the verb, mode markers mark the whole clause. Thus, the lexical entry of a mode marker is as follows:

(11) $\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{HEAD}\left[\begin{array}{l}marker \\ \text{SPEC}\ \boxed{1}\ \text{S}\left[\text{CONT}\ \boxed{2}\right]\end{array}\right] \\ \text{MARKING}\ mode\end{array}\right] \\ \text{CONT} \mid \text{RELS}\left\langle\left[\begin{array}{l}mode\text{-}rel \\ \text{ARG}\ \boxed{2}\end{array}\right]\right\rangle\end{array}\right]$

Figure (**??**) shows the combination for the following example:

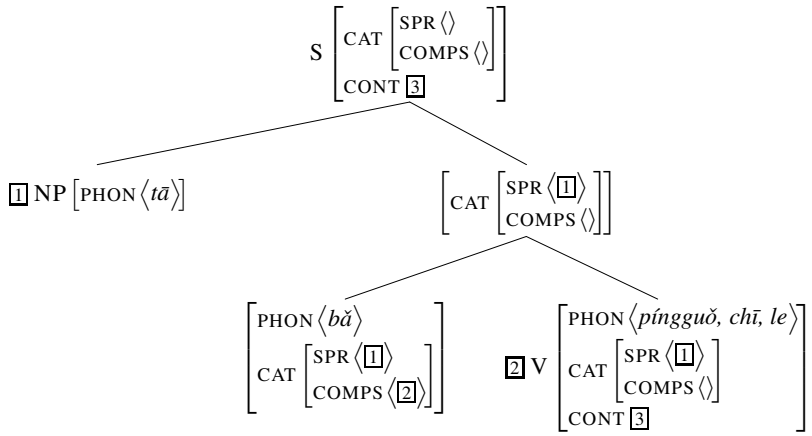


Figure 1: Syntactic combination for (6a): 他把苹果吃了。

(12) 张三 来 了 吗?
Zhāngsān lái le ma?
Zhangsan arrive PFV INTERROG
'Has Zhangsan (already) come?'

5.2.2 Syntactic marking

In this section, we consider the syntactic markers 的 *de*, 得 *de* and 地 *de*. These markers do not carry semantic content; they are used to make constituents eligible for specific syntactic positions. The following examples illustrate:

(13) a. 很 快 的 车
hěn kuài de chē
very fast MK.DE1 car
'a very fast car'

b. 他 很 快 地 跑。
Tā hěn kuài de pǎo.
he very fast MK.DE2 run
'He runs very quickly.'

c. 他 跑 得 很 快。
Tā pǎo de hěn kuài.
he rung MK.DE3 very fast
'He runs very quickly.'

DE1 is used for marking prenominal modifiers (APs, relative clauses, possessives and other NP modifiers). DE2 is used to mark postverbal complements that denote the manner, degree of intensity or result of an action. DE3 is used for the marking of preverbal manner adjuncts.

The constraint on the lexical entry of a syntactic marker is as follows:

(14) $\left[\text{PHON}\left\langle pǎo\right\rangle \; \text{CAT}\left[\text{HEAD}\left[\textit{marker} \; \text{SPEC}\,\boxed{1}\,\text{V}\left[\text{CONT}\,\boxed{2}\right]\right] \; \text{MARKING}\;\textit{syn}\right]\right]$

We can see that the marker does not make a semantic contribution. The *marking* type *syn* has three subtypes which correspond to the three *de*'s. Now, in order to account for the syntactic combination, we posit relational constraints on the resulting *head-marker* structures. These constraints relate the marker with the type of constituent that selects or is modified by the *head-marker* structure:

(15) a. $\left[\textit{head-marker-structure} \; \text{HEAD} \mid \text{MOD N} \; \text{MARKED}\;\textit{de1}\right]$

b. $\left[\textit{head-marker-structure} \; \text{HEAD} \mid \text{MOD V} \; \text{MARKED}\;\textit{de2}\right]$

De3-complements are selected by the verb; the following shows the lexical entry for *pǎo* in (13c):

(16) $\left[\text{CAT}\left[\text{HEAD}\;\textit{verb} \; \text{COMPS}\;\textit{list} \oplus \left\langle\left[\text{MARKING}\;\textit{de} \; \text{CONT} \mid \text{ARG}\,\boxed{1}\right]\right\rangle\right] \; \text{CONT}\,\boxed{1}\right]$

Thus, we provide a unified analysis of the three markers which builds on the *head-marker* schema and the MARKING feature. MARKING determines syntactic combination: modifiers relate the MARKING feature to the syntactic type of the

There appears to be a semantic overlap between postverbal manner complements with DE2 and preverbal adjuncts with DE3. The structures are mainly distinguished in terms of syntactic status of the manner constituent (complement for DE, adjunct for DE) and information structure. Thus, under the assumption that the sentence-final position accommodates the focus, the speaker has the choice between two structures that focus either the action or the manner in which it is conducted.

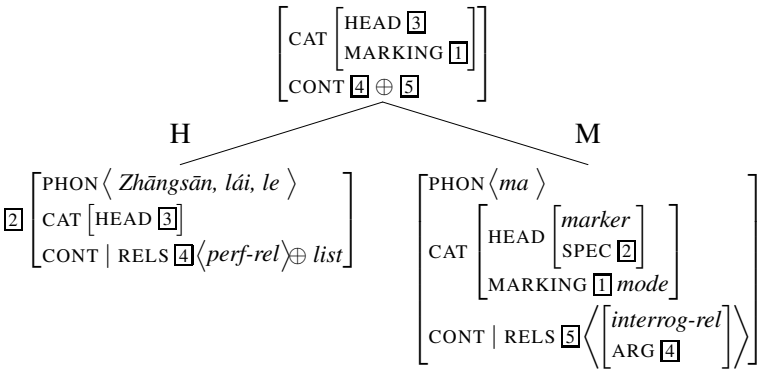


Figure 2: Syntactic combination for (12): 张三来了吗?

modified constituent. Complements can be selected by verbal heads if they satisfy the selectional constraint on MARKING specified by the head.

6 Conclusion

In this paper, we have presented a HPSG implementation of a Chinese grammar fragment; the framework HPSG is well-suited for the analysis of Chinese since it makes a minimal number of empirical assumptions about linguistic objects while providing the grammar writer with a model-theoretically grounded set of descriptive tools. Since the empirical notions and assumptions used in Western linguistics cannot be readily transferred to Chinese, HPSG thus gives us the possibility to formulate theory-neutral analyses which can then be used to derive broader generalizations about the language. In the present paper, we have focussed on two sets of phenomena, marking and valence alternation, and shown how they can be analyzed in a unified manner using the mechanisms provided by the framework.

References

Bender, Emily. 2000. The Syntax of Mandarin *ba*: Reconsidering the Verbal Analysis. *Journal of East Asian Linguistics* 9, 100 – 145.

Briscoe, Ted and Copestake, Ann. 1999. Lexical Rules in Constraint"=Based Grammar. *Computational Linguistics* 25(4), 487–526.

Carpenter, Bob. 1992. *The Logic of Typed Feature Structures*. Tracts in Theoretical Computer Science, Cambridge: Cambridge University Press.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. Stanford: CSLI Publications.

Copestake, Ann, Flickinger, Daniel P., Pollard, Carl J. and Sag, Ivan A. 2005. Minimal Recursion Semantics: an Introduction. *Research on Language and Computation* 4(3), 281 – 332.

Ding, Picus Sizhi. 2000. A computational study of the *ba* resultative construction: parsing Mandarin *ba*-sentences in HPSG. In *Proceedings of PACLIC 14*, Tokyo, Japan.

Flickinger, Daniel P. 1987. *Lexical Rules in the Hierarchical Lexicon*. Ph. D.thesis, Stanford University.

Gang, Liu. 1997. *Eine unifikations-basierte Grammatik für das moderne Chinesisch – dargestellt in der HPSG*. Ph. D.thesis, University Constance, SFB 471, FG Sprachwissenschaft, Universität Konstanz, Germany.

Gao, Qian. 1993. Chinese NP Structure. In Andreas Kathol and Carl J. Pollard (eds.), *Papers in Syntax*, OSU Working Papers in Linguistics, No. 42, pages 88–116, Ohio State University: Department of Linguistics.

Gao, Qian. 2000. *Argument Structure, HPSG and Chinese Grammar*. Ph. D.thesis, Ohio State University.

Lipenkova, Janna. 2009. *Serienverbkonstruktionen im Chinesischen und ihre Analyse im Rahmen von HPSG*. Masters Thesis, Freie Universitt Berlin.

Lipenkova, Janna. 2011. Reanalysis of obligatory modifiers as complements in the Chinese *bă*-construction. In Stefan Müller (ed.), *Proceedings of the 18th International Head-driven Phrase Structure Grammar Conference*, Stanford: CSLI Publications.

Meurers, Detmar. 2000. Lexical Generalizations in the Syntax of German Non-Finite Con-

structions. Arbeitspapiere des SFB 340 145, Eberhard-Karls-Universität, Tübingen.

Meurers, Walt Detmar. 2001. On Expressing Lexical Generalizations in HPSG. *Nordic Journal of Linguistics* 24(2).

Meurers, Walt Detmar, Penn, Gerald and Richter, Frank. 2002a. A Web-Based Instructional Platform for Constraint"=Based Grammar Formalisms and Parsing. In Radev and Brew (2002), pages 18–25.

Meurers, Walt Detmar, Penn, Gerald and Richter, Frank. 2002b. A Web-Based Instructional Platform for Constraint"=Based Grammar Formalisms and Parsing. In Radev and Brew (2002), pages 18 – 25.

Müller, Stefan. 2006. Phrasal or Lexical Constructions? *Language* 82(4), 850 – 883.

Müller, Stefan. 2007. The Grammix CD Rom. A Software Collection for Developing Typed Feature Structure Grammars. In Tracy Holloway King and Emily M. Bender (eds.), *Grammar Engineering across Frameworks 2007*, Studies in Computational Linguistics ONLINE, Stanford: cslip.

Müller, Stefan. 2008. *Head-Driven Phrase Structure Grammar: Eine Einführung*. Stauffenburg Einführungen, No. 17, Tübingen: Stauffenburg Verlag, second edition.

Müller, Stefan and Lipenkova, Janna. 2009. Serial Verb Constructions in Mandarin Chinese. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*.

Ng, Say K. 1999. A Double-specifier Account of Chinese NPs Using Head-driven Phrase Structure Grammar. Master thesis, Department of Linguistics, University of Edinburgh.

Pollard, Carl J. and Sag, Ivan A. 1987. *Information-Based Syntax and Semantics*. CSLI Lecture Notes, No. 13.

Pollard, Carl J. and Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics, Chicago, London: University of Chicago Press.

Radev, Dragomir and Brew, Chris (eds.). 2002.

Sag, Ivan A., Wasow, Thomas and Bender, Emily M. 2003. *Syntactic Theory: A Formal Introduction*. CSLI Lecture Notes, No. 152, second edition.

Sybesma, Rint. 1999. *The Mandarin VP*. Dordrecht: Kluwer.

Wang, Xiangli, Iwasawa, Shun'ya, Miyao, Yusuke, Matsuzaki, Takuya, Yu, Kun and ichi Tsujii, Jun. 2009. Design of Chinese HPSG Framework for Data-Driven Parsing. In Olivia Kwong (ed.), *Proceedings of PACLIC 2009*, pages 835 – 842, City University of Hong Kong Press.

Xue, Ping and McFetridge, Paul. 1995. DP structure, HPSG, and the Chinese NP. In *Proceedings of the 14th Annual Conference of Canadian Linguistics Association*, Montreal, Canada.

Yu, Kun, Miyao, Yusuke, Wang, Xiangli, Matsuzaki, Takuya and ichi Tsujii, Jun. 2010. Semi-automatically Developing Chinese HPSG Grammar from the Penn Chinese Treebank for Deep Parsing. In Chu-Ren Huang and Dan Jurafsky (eds.), *Proceedings of COLING 2010*, pages 1417 – 1425, Chinese Information Processing Society of China.

Zhang, Yi, Wang, Rui and Chen, Yu. 2011. Engineering a Deep HPSG for Mandarin Chinese. In *Proceedings of the 9th Workshop On Asian Language Resources located at IJCNLP, November 12-13, Chiang Mai, Thailand*, ACL.

Zhang, Yi, Wang, Rui and Chen, Yu. 2012. Joint Grammar and Treebank Development for Mandarin Chinese with HPSG. In Nicoletta Calzolari et al. (ed.), *Proceedings of the LREC'12*, Istanbul, Turkey.