A Description of Chinese NPs using HPSG

Extended Abstract

The paper proposes a HPSG analysis of Chinese NPs. The analysis is implemented in the LKB system. The authors discuss the structures of Chinese NPs and focus on the relationships between nouns and classifiers, usually known as noun-classifier matching. Previous literatures, concerning Chinese NPs in the framework of HPSG, only simply point out this matching, but the questions about how and why do they match are left open. In this paper, we shed light on this crucial problem. Based on the analysis the paper further establishes a grammar of Chinese NPs and implements the rules and typed feature structures of HPSG in the LKB system. In the concluding section, we give out the advantages and disadvantages of HPSG when implementing in Chinese NPs.

In the first section, the article makes a general description of Chinese NPs, and summarizes the basic structures of Chinese NPs as in Table 1.

Chinese NPs	**Syntactic Structures**	**Examples (Frequent)**
bare nouns	N	wo mai le *shu* 'I bought a book'
modifiers + nouns	A* + N	ci *shu* 'the book' (62)
	Num + **CL** + N	yi ben *shu* 'a book' (56)
	Dem + **CL** + N	zhe ben *shu* 'this book' (143)
	Dem +Num+ **CL** + N	zhe liang ben *shu* 'these two books' (8)

Table 1: The basic structures of Chinese NPs

We also discuss more complex structures, which are often combined by Possessives or relative clauses with the mark of particle 'de'. We can find out these attributes are inserted before 'Dem' or 'N', never between 'Num' and 'CL'. We point out that the matching problem between nouns and classifiers is the crucial factor in Chinese NPs.

Section two focuses on the analysis of the previous studies of Chinese NPs in the framework of HPSG. Three main literatures are Gao (1994), Xue/McFetridge (1995) and Ng (1997). In this part, the authors present the main arguments and figures of each essay and also pick out their good points and pitfalls.

In short, what to note worthy is that, nouns should be the head of NPs, and using a double-specifier hypothesis to analyze 'Dem-CLP-N' in Chinese NP structures is a better one. Also, a deeper insight to the relation between numeral and classifier, which leads to an enlighten idea of the classification of classifiers. Still, many problems of Chinese NPs are not solved, such as the collocation between classifiers and nouns.

Section three treats the relationship between nouns and classifiers. Following the ideas of Aikhenvald (2000), the authors first suggest the basic feature of classifiers, which is known as co-occurring with the noun. Also in Mandarin Chinese, this kind of agreement is determined by lexical selection, rather than matching any inflectional properties, which most Indo-European languages do.

Concerning Chinese classifiers, previous studies since Li (1924/1992) all classify

them into subtypes, as well as nouns. This classification makes sense since the subtypes reveal the matching information. Based on Huang/Ahrens (2003) and Wang (2004), we get this noun-classifier matching as Table 2 shows: (+ refers to matching, - refers to non-matching)

	nouns	Individual	Substance	Group	Abstract	Proper	Event	None
classifiers	codes	ng	nb	nm	na	np	ne	ngq
Individual	qns	+	-	-	-	+	-	-
Measure	qnp	+	+	+	+	-	-	-
Container	qno	+	+	+	-	-	-	-
Group	qnc	+	+	+	-	-	-	-
Kind	qnk	+	+	+	+	-	-	-
Shape	qns	+	+	-	+	-	-	-
Indefinite	qni	+	+	+	+	-	-	-
Time	qt	-	-	-	-	-	+	-
Verbal	qv	-	-	-	-	-	+	-

Table 2: The matching between nouns and classifiers

Further, the author tries to solve the matching problem between individual nouns and classifiers. Considering a simple example of noun phrase ‘yi ben shu’ (a book), we note that, the classifier ‘ben’ could be collocated with the noun ‘shu’ (book), but another classifier ‘tai’ will not be allowed to match with ‘shu’. Based on these kinds of facts, Zhu (1982:49) claims that the matching between individual nouns and classifiers are idiosyncratic, and thus need to be noted in the dictionary. However, nouns are an open set; there is no end to list all the members. Further, it is the classifier that selects the relevant properties of the noun and coerces the appropriate meaning (Huang/ Ahrens, 2003).

Following Tai (1990:312), we have to seek the basic common features of nouns and classifiers, which are inevitably the semantic features. We could model the four ways of defining the semantic features, such as constitutive, formal, telic and agentive, which are suggested by Pustejovsky (1995). In the case of ‘shu’ (book), the classifier ‘ben’ has the formal role of a bound volume (Huang/ Ahrens, 2003), while the classifier ‘tai’ only modifies machines.

In sum, the classifiers need a classification system based on the matching with nouns. The matching between individual classifiers and nouns reveals the semantic properties of nouns. Considering the multiple matching phenomena between classifiers and nouns, the semantic features described are very complicating.

In section four, we propose an analysis of Chinese NPs in HPSG. We suggest two ways to describe Chinese NPs, (1) Syntactically, we propose a model of the linear sequence of Chinese NPs, including the basic structure, as well as complex one; (2) Semantically, we will define a new features to describe the semantic properties.

Syntactic structures

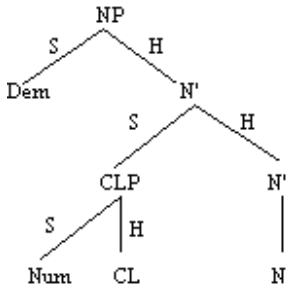

Figure 1: basic structure

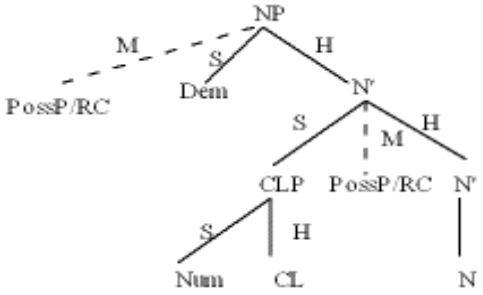

Figure 2: complex structure

In the HPSG framework, we need to modify the head-specifier rule, which also allows a double specifier account. Following Ng (1997), we modify the rule like the following.

$$\begin{bmatrix}\text{phrase}\\ \text{SPR} < >\end{bmatrix} \rightarrow [1]\,,\,[2]\ \text{H}\begin{bmatrix}\text{phrase}\\ \text{HEAD} < \text{SPR}\ [1]\,,\,[2]\end{bmatrix}$$

Figure 3: a modified version of Sag and Wasow (2003: 87)

Type hierarchy of nouns and classifiers

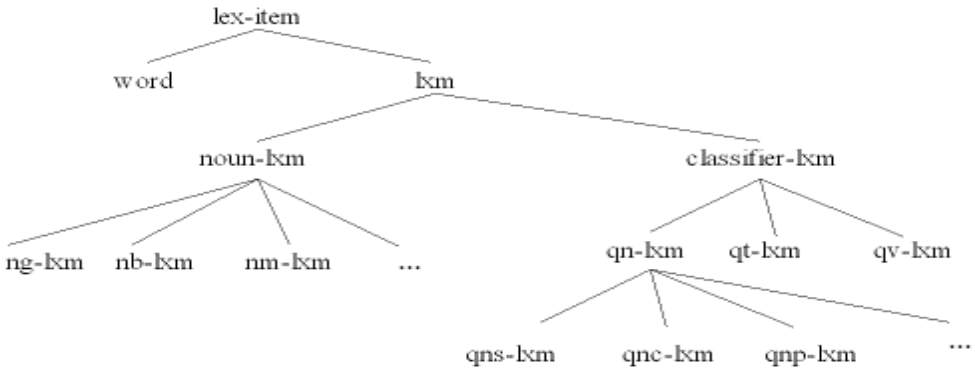

Figure 4: Type hierarchy of nouns and classifiers

As can be seen, the sub types of nouns or classifiers are based on the table 2. Then, similar to AGR, we introduce a feature TYP to entail the constraints of subtypes. And for an individual classifier, the lexical entry is:

[SYN [HEAD [AGR [TYP /individual]]]]

To noun class, we will adopt the properties of classifiers to constraint its property. For example, individual nouns should be described as follows:

[SYN [HEAD [AGR [TYP /individual|measure|container|group|kind|shape|indefinite]]]]

The semantic features of NPs

We introduce another feature CLS to represent the semantic properties of nouns and classifiers. In this way, we can judge the matching by an intersection of the features in CLS.

We also try to make a nearly complete account of the semantic features of Chinese NPs. Semantic Compositionality Principle is used to combine all the

information together.

Figure 5 shows a complete analysis of a Chinese NP (yi ben shu, a book)

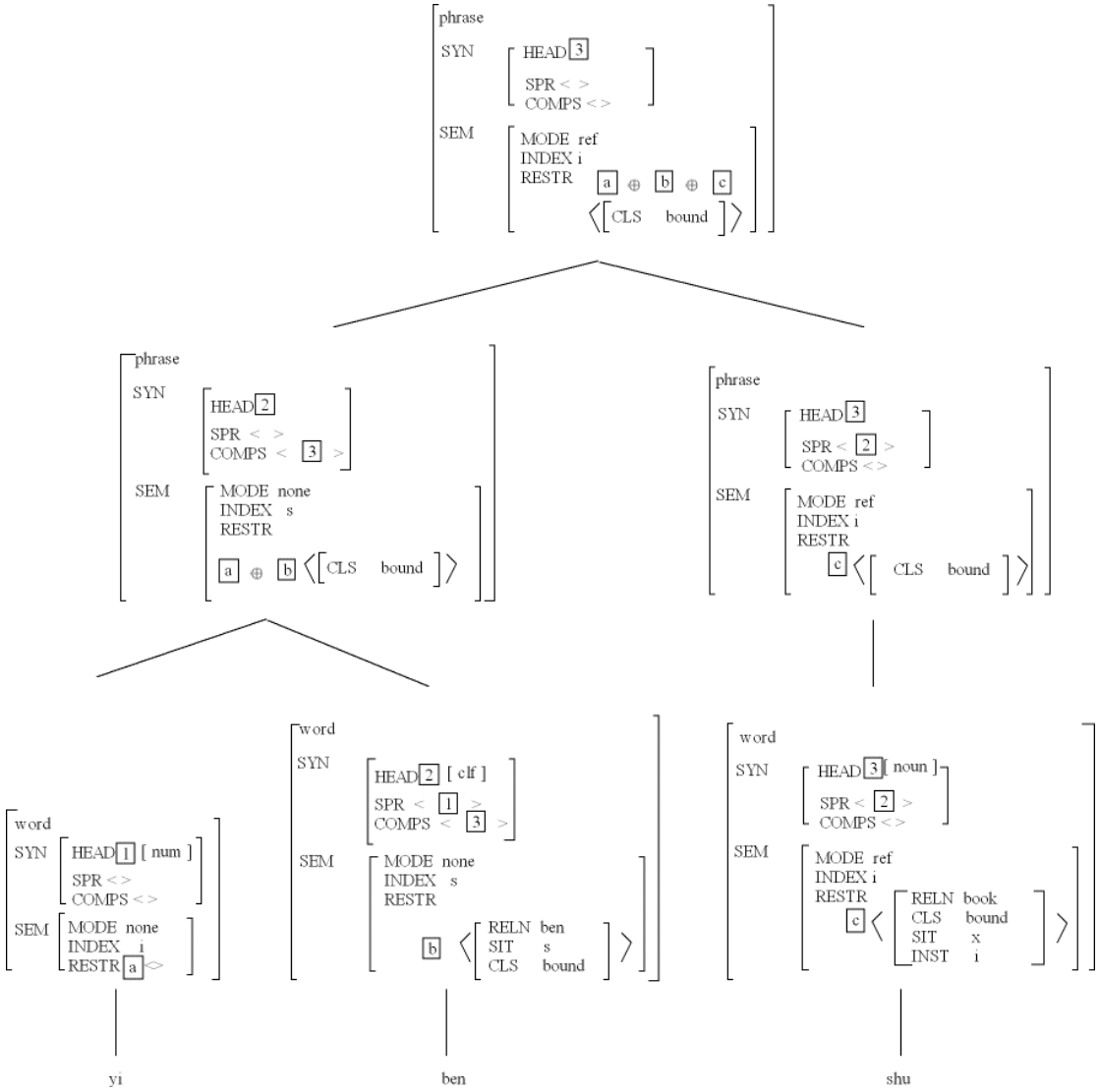

Figure 5: A complete analysis of ‘yi ben shu’

In Section 5, we give a sketch of the implementation of above-mentioned analysis in LKB system. We try to test our proposal through the implementation in LKB system for knowing the pros and cons of these analyses in this paper. The codes of grammar rule, types and lexicon are found in Appendix. As the examples in section three, we will test two phrases: ‘yi ben shu’ and ‘yi tai shu’. Figure 6-8 are output of LKB for the NP: ‘yi ben shu’ (a book).

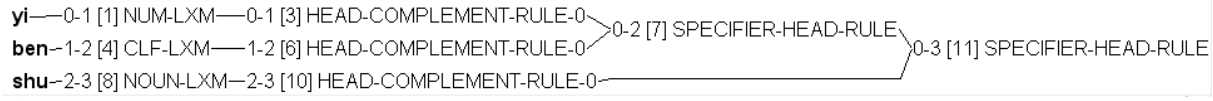

Figure 6: Chart for ‘yi ben shu’

Finally, the result of ‘yi tai shu’ is showed as ‘No parses found’. The matching fails, since ‘tai’ is given a feature of ‘machine’, which is not compatible with that of ‘shu’. In this sense, we successfully implement analyzes above.

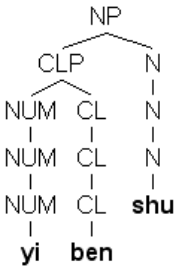

Figure 7: Tree diagram for ‘yi ben shu’

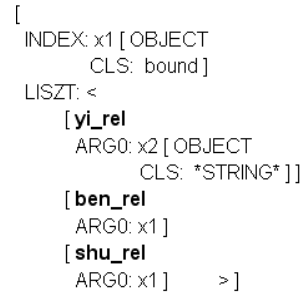

Figure 8: MRS for ‘yi ben shu’

In this paper, we propose a more complete analysis of the constraints between classifiers and nouns and implement it in LKB system.

We also find out several questions: (1) For Chinese HPSG processing, we need a deeper and more complete study of the classification and the semantic features of classifiers as well as nouns. (2) The selection between nouns and classifiers reveals that multiple matching need to be solved, and this matching is not confined to word to word, rather it prefers a selection among semantic features. All present a challenge to HPSG.

References

Aikhenvald, A. 2000. Classifiers: a Typology of Noun Categorization Devices. OUP. 82-84.

Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal Recursion Semantics: an Introduction. Research on Language and Computation 3.4: 281-332.

Dan Flickinger, Emily M. Bender and Stephan Oepen. 2003. MRS in the LinGO Grammar Matrix: A Practical User’s Guide. Technical report, Deep Thought Project Deliverable. 12.

Gao, Q. 1994._Chinese NP Structure. Linguistics 22, 475-510.

Huang, C. R. and Ahrens, K. 2003. Individuals, kinds and events: classifier coercion of nouns. Language Sciences, (25):353–373.

Levy, R. and Oshima, D. 2003. Non-transitive information flow in Japanese noun-classifier matching. In proceedings of HPSG 2003. CSLI. 257-277.

Li, J. X. 1924/1992. Xin Zhu Guo Yu Wen Fa. Beijing: The Commercial Press. 84.

Ng, S. K. 1997. A Double-Specifier Account of Chinese NPs using Head-Driven Phrase Structure Grammar. MSc Thesis. Department of Linguistics, University of Edinburgh.

Pustejovsky, J. 1995. The Generative Lexicon. MIT Press, Cambridge. 61-62.

Sag, Ivan. A. and Wasow, T. 1999. Syntactic Theory: A Formal Introduction. Center for the Study of Language and Information, Stanford, California. 172.

Tai, J.H.-Y. 1990. Variation in classifier systems across Chinese dialects: towards a cognition-based semantic approach. In *Proceedings of the First International Symposium on Chinese Languages and Linguistic*, 308-322.

Wang, H. 2004. A syntagmatic study of noun senses in contemporary Chinese. Beijing: Peking University Press.128.

Xue P. and McFetridge, P. 1995. DP structure, HPSG and the Chinese NP. In *Proceedings of the 14th Annual Conference of Canadian Linguistics Association (CLA '95)*. Montreal, Canada: University of Toronto Press. 10.14.

Zhu, D. X. 1982. Yu Fa Jiang Yi. Beijing: The Commercial Press. 49.151.