CHAPTER IV

DATABASE AND INITIAL OBSERVATION

Previous researchers have mostly emphasized the identification between the missing OBJ and the TOP and have stated that the required NP complement can be satisfied by the NP category realized as TOP. Nevertheless, a fundamental question remains to be answered first: Except for NP category, can any other categories be realized as TOP identified with the missing OBJ apart from the NP category? If so, what are they and why can they do this? According to the corpus data, we find that only VP and CP can be realized as TOP to fulfill the f-structure apart from the NP category. In this thesis, these category types will be examined and formed into a hierarchy according to the frequency of their occurrence in the corpora data.

4.1 The Database

In order to effectively explore which category type of phrase can be realized as TOP identical to OBJ, three online resources are used for the analysis in the main body of this study. Our main resource is the Udndata news database (http://udndata.com/library/). It provides the full text of the latest news of five local newspapers (Lianhebao, Jingjiribao, Minshenbao, Lianhewanbao, Xingbao) in Taiwan. The other is the 5-million-word Sinica Corpus, which is largely composed of both Mandarin written and spoken samples in Taiwan (http://www.

Sinica.edu.tw/SinicaCorpus). Another online archive, the daily-updated online archive, Google (http://www.google.com), contains the latest information that includes various written texts. In addition to the examples found in these three online resources, some of the examples are from my own creations with some modifications by my supervisor.

Acknowledging that the selection of category types is significant to our data analysis (cf. the concept of high-frequent usages in Biber et al. 1998), the total number of occurrences of certain category types in the Udndata, the Sinica Corpus and Google are counted. The category types that can be realized as TOP form a hierarchy: NP > VP > CP > PP. The category type NP with the highest number of tokens is placed highest in the hierarchy on the basis of a high frequency occurrence in the three resources. Moreover, some crucial issues related to these category types arise for further study.

4.1.1 Initial Observation of the Category Types That Could be Realized as TOP

This section starts with an observation of the predicates with movement paradoxes in Mandarin Chinese. The classifications of these predicates in this study are made mainly on the basis of Huang's (1989) and Her's (1999) studies of topicalization in Mandarin Chinese. The number of occurrence of category types

that can be realized as TOP in the data is represented as in Table 1 below:

Table 1: The number of occurrences of category types that can be realized as TOP in

   Mandarin Chinese

	Sources of Data (1000)			
Category types	NP	VP	CP	PP
Number of occurrence	499 (49.9%)	361 (36.1%)	140 (14%)	0 (0%)

The possible category types that can be realized as TOP in Mandarin Chinese in the

corpus include NP, VP, and, CP as shown in examples below.

(51) NP

    a.   [幾何學]$_{NP}$ 他   最   拿手

       [*ji3he2xue2*]$_{NP}$ *ta1*   *zui4*   *na2shou3*

       geometry   he   most   be-good-at

       'He is good at geometry.'

    b.   [那場比賽]$_{NP}$ 我   負責

       [*na4chang3bi3sai4*]$_{NP}$ *wo3*   *fu4ze2*

       that game   I   be-responsible-for

       'I am responsible for that game.'

    c.   [菜單]$_{NP}$ 你   作主

       [*cai4dan1*]$_{NP}$ *ni3*   *zuo4zhu3*

       menu   you   take-charge-of

       'You take charge of the menu.'

d.　[藝術]_{NP}　　我　　　　一竅不通

　　[*yi4shu4*]_{NP}　　*wo3*　　*yi2qiao4bu4tong1*

　　art　　　　I　　know-nothing-about

　'I know nothing about art.'

(52) VP

a.　[修理電視]_{VP}　　他　　最　　拿手

　[*xiu1li3dian4shi4*]_{VP}　　*ta1*　　*zui4*　　*na2shou3*

　repairing a TV　　he　　most　　be-good-at

　'He is good at repairing TVs.'

b.　[買螢幕]_{VP}　　你　　　　負責

　[*mai3ying2mu4*]_{VP}　　*ni3*　　　　*fu4ze2*

　buying monitors　　you　　be-responsible-for

　'You are responsible for buying the monitors.'

c.　[買房子]_{VP}　　你　　　作主

　[*mai3fang2zi*]_{VP}　　*ni3*　　　*zuo4zhu3*

　buying a house　　you　　take-charge-of

　'You take charge of buying the house.'

d.　[修車]_{VP}　　我　　　　一竅不通

　[*xiu1che1*]_{VP}　　*wo3*　　*yi2qiao4bu4tong1*

　repairing a car　　I　　know-nothing-about

　'I know nothing about repairing a car.'

(53) CP

    a. [他   讚美   我]_{CP}   我   心領   了

      [*ta1* *zan4mei3* *wo3*]_{CP} *wo3*   *xin1ling3*   *le*

      he   criticize   I   I   decline-with-thanks   ASP

      'He gave me some praise and I declined its with thanks.'

    b. [你   想   升官]_{CP}   我   無能無力

      [*ni3* *xiang3* *sheng1guan1*]_{CP} *wo3* *wu2neng2we2li4*

      You want   get promotion   I   can-do-nothing-for

      'You want to get promotion but there is nothing I can do about it.'

Table 1 shows that the NP category is the token which occurs with the highest frequency in the three category types. The total number of occurrences of NP is 499. The tokens for the high-frequency of NP here represent the highest status of NPs in the hierarchy and the highest possibility of their being realized as TOP.

4.2 Summary

This chapter introduces the sources of our database selected according to actual use in the Udndata as well as samples retrieved from the Sinica data and Google. The goal of this study is to provide a possible generalization of the category types that can be realized as TOP of Mandarin Chinese predicates with movement paradoxes. It is hoped that this study may provide a plausible hierarchy for the category types that are possible for being realized as TOP on the basis of the findings above.