The Naming Sharing Structure and its Cognitive Meaning in Chinese and English

Shili Ge, Rou Song
Guangdong Collaborative Innovation Center for Language Research & Service,
Guangdong University of Foreign Studies, Guangzhou, China 510420

Abstract

This research puts forward two concepts, the Naming Sharing Structure (NSS) and the Naming-Telling clause (NT clause). By transforming linear text into a plane structure, we discovered five construction patterns of NT clause for the NSS. The constructions of NT clauses are order-preserving, non-transgressive, and syntactically well-formed. These features of NT clause can be explained from the point of view of cognition.

1 NT Clause and Naming Sharing Structure

Indo-European languages are verb centered in syntactical and semantic analysis, and are analyzed into SVO or verb argument structure (e.g. Pollard & Sag 1994, Dalrymple 2001, and Carnie 2013). This approach is not as suitable to Chinese since many Chinese clauses are subject- or verb-less. Moreover, basically many Chinese verbs can also function as nouns (Shen, 2013). On the other hand, Chinese is a topical language, so it is more convenient to analyze Chinese clauses from the perspective of a topic-comment structure. Topic refers here to the similar concept as used in Systemic Functional Grammar (SFL; Halliday, 1994), but "subject", the traditional term, is also used together with a few adverbial and predicative components, so the concept of topic used here is more general.

We refer to this type of topic as a 'naming'. In combination, a naming (i.e. a reference) and a 'telling', which predicates or explains the naming, comprise an NT clause. When more than one telling shares the same naming, these NT clauses actually form a naming-centered syntactical structure, which referred to as Naming Sharing Structure (NSS). The NSS is distinctively different from verb-centered structure. This study is aimed at discovering the structural patterns of NSS and discusses its cognitive meaning and applicability.

The two-dimensional schema of NSS is illustrated in Section 2, and the shape of the schema can intuitively show the formal properties of NSS. In Section 3 and 4, examples of NSS in Chinese and English texts are illustrated respectively, and their patterns of constructing NT clauses are presented. The cognitive meaning of NSS is discussed in Section 5 and a conclusion offered in section 6.

2 The Newline-Indent Schema and the Naming Sharing Structure

Example 1 (Qian Zhongshu *Fortress Besieged*)

高松年发奋办公，夙夜匪懈，精明得真是睡觉还睁着眼睛，戴着眼镜，做梦都不含糊的。摇篮也挑选得很好，在平成县乡下一个本地财主家的花园里，面溪背山。

Translation:

Kao Sung-nien worked furiously day and night without rest. He was so keen he literally slept with his eyes wide open and his glasses on, so that he was never lazy even in his dreams. The cradle had been very well chosen. It was the garden of a local millionaire, in the countryside of the P'ing-ch'eng district; it faced a stream with mountains in the

13

background. (Translated by Jeanne Kelly and Nathan K. Mao, 2003)

This example includes the eight punctuation clauses c1-c8. A punctuation clause (P-clause) is a text segment which is separated from the surrounding context by a comma, semicolon, period, exclamation mark, or question mark (See Song, 2013). All the P-clauses in this ex-ample, except for the first one, share components from their previous P-clauses to make sense. Following each punctuation mark, a carriage-return and newline mark is inserted, so that every P-clause takes up one line. Then the clause is indented right after its shared components. The left schema in Figure 1 is called Newline-Indent Schema (NIS). It transforms linear text into two dimensional plane structure which shows intuitionally the components in the previous context of a P-clause that should be supplemented at the beginning of the P-clause.
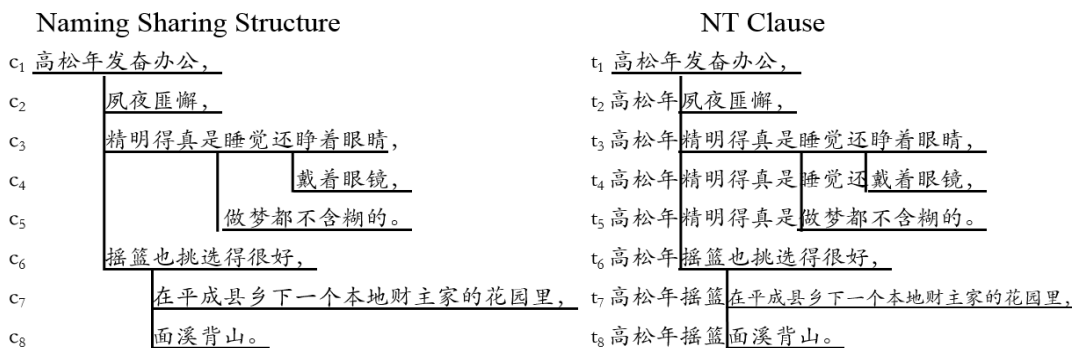

Figure 1: The Chinese Naming Sharing Structure and NT Clause illustrated with NIS.

In the right of Figure 1, the supplemented result of every P-clause is listed. They are complete naming-telling structures, referred to as NT clauses.

3 Chinese NSS and patterns constructing NT clauses

3.1 Stack Pattern

In example 1, the pattern in which NT clauses are constructed is called Stack Pattern, since the left end of every line is treated as the stack bottom and the right end as the stack top, and the construction of NT clause is exactly a process of pop and push of a stack as shown in Figure 2.

Figure 2: An iterating process of NT Clause constructed by Stack Pattern.

14

3.2 Back Position Pattern

Example 2 (Zhang Xianliang *Geezer Xing and his dog*)

户口不迁来，再没有个娃娃，女人迟早得回老家，

Translation:

Her Hukou is not official reregistered, and with no children, the woman will go back hometown sooner or later,

Figure 3: Back Position Pattern of Chinese.

In this example, the first two P-clauses are both lack of the subject "女人" (woman), which appears at the beginning of the third P-clause. In order to construct NT clauses of these two P-clauses, "女人" (woman) has to be supplemented based on the later context. Hence, the pattern is named Back Position Pattern.

3.3 New Branch Pattern

Example 3 (Qian Zhongshu *Fortress Besieged*)

他把带到银行里偷空看的书翻开，每个字都认识，没一句有意义。听见外面跑堂招呼客人的声音，心就直提上来。

Translation:

He opened the book he had brought to the bank to read during his spare time but not a single sentence made any sense. When he heard the waiter greeting a customer outside, his heart fluttered. (Translated by Jeanne Kelly and Nathan K. Mao, 2003)

Figure 4: New Branch Pattern of Chinese.

Both P-clause 2 and 3 in the examples lack the naming "带到银行里偷看的书" (the book he had brought to the bank to read during his spare time), instead of "他把带到银行里偷看的书" (He does the book he had brought to the bank to read during his spare time). This is like placing a sluice between "他把" (He does) and "带到银行里偷看的书" (the book he had brought to the bank to read during his spare time) (the sluice is represented with mark ‖). The components on the left of the sluice is not in the naming while on the right is the naming for the following two P-clauses. This naming is like a new branch growing on a tree trunk, hence we refer to it as New Branch, and the pattern New Branch Pattern.

3.4 Influx Pattern

Example 4 (Qian Zhongshu *Fortress Besieged*)

他们都是上了年纪，有孩子的人，

Translation:

They're older and have kids. (Translated by Jeanne Kelly and Nathan K. Mao, 2003)

Figure 5: Influx Pattern of Chinese.

In this example, not only the second P-clause lacks a naming in its beginning, but the first P-clause also lacks an ending "的人" (person) from the tail of the second P-clause should follow it. This is just like that the utterance flow starts from "他们都是" (They all are) into two tributaries "上了年纪" (older) and "有小孩子" (have kids), then the two tributaries flow to "的人" (person). Hence we name the pattern Influx Pattern.

3.5 Enclosed Chunk Pattern

The previous patterns are based on geometrical structure. When mentioning direct quotation, the construction of NT clause needs another kind of pattern to distinguish different layers of utterance.

Example 5 (Jin Yong *The Deer and the Cauldron*)

查伊璜心想："我连吴六奇的名字也没听见过，为何送礼于我？"当下沉吟不语。

Translation:

"I've never even heard this man's name before," Zha thought. "Whatever should he be sending me a present for?" He remained rapt in thought and made no response. （Translated and edited by John Minford, 1999）
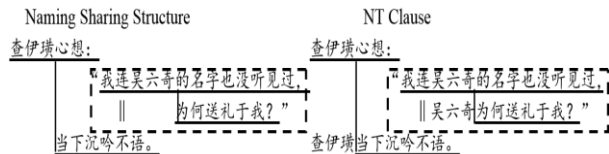
Figure 6: Enclosed Chunk Pattern of Chinese.

This example contains a direct quotation describing psychological activity. Its inside (indicated by the dotted line frame) and outside segments each form a NSS, respectively. And each of NSS constructs NT clauses according to its own pattern. Since their internal structure is enclosed from their outside contexts, we refer to this pattern as the Enclosed Chunk Pattern.

We annotated NSS for texts with different Chinese styles. Each text is annotated from beginning to end, resulting in about 40,000 NT clauses in total. We found that the five patterns of NT clause construction account for 97% of P-clause in Chinese texts.

4 The English NSS and the patterns of constructing NT clauses

4.1 English NT Clauses

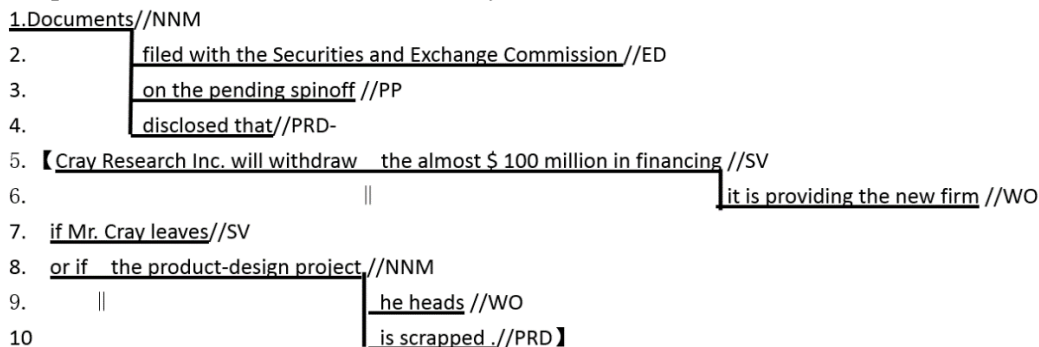Some English sentence (texts separated by period, exclamation mark, question mark) are larger, and of more complex structures than Chinese, but they still can be separated into several NT clauses, based on relationship of naming and telling. The English NT clause can be classified into an SV clause – consisting of subject and predicate –, and a non-SV clause, where the latter may consist of the following:

- antecedent+relative clause
- NP+ post-modifying nonfinite verb phrase (past participle phrase, present participle phrase and infinitive phrase)
- NP+ post-modifying declarative preposition phrase
- NP+ post-modifying adjective phrase
- NP+ post-modifying noun phrase (apposition and other explanatory NP)
- Verbal object + verbal complement

Example 6 (*Wall Street Journal*)

Documents filed with the Securities and Exchange Commission on the pending spinoff disclosed that Cray Research Inc. will withdraw the almost $ 100 million in financing it is providing the new firm if Mr. Cray leaves or if the product-design project he heads is scrapped .

The relationship of naming-telling in this long sentence can be shown in the following NIS:

Figure 7: Naming Sharing Structure in English text.

In Example 6, the parts enclosed in the black brackets, the object clause of "disclose", is an enclosed chunk. The token on the right side of double slash, which is at the right end of each line, represents the type of this line, which is naming, telling, or NT clause. They stand for:

- NNM：NP as naming
- ED：past participle phrase as telling
- PP：preposition phrase as telling

- PRD：the predicate in SV structure as telling
- PRD-：the predicate in SV structure as telling, but lacks object clause
- SV：SV structure as NT clause
- WO：relative clause as telling，antecedent as object of the relative clause

All the above categories of telling can be supplemented by naming just as in Chinese, which come exactly from the line right above:

```
1.Documents //NNM
2.Documents filed with the Securities and Exchange Commission //ED
3.Documents on the pending spinoff //PP
4.Documents disclosed that //PRD-
5. 【Cray Research Inc.will withdraw    the almost $ 100 million in financing //SV
6.                                ‖ the almost $ 100 million in financing it is providing the new firm //WO
7.    if Mr. Cray leaves //SV
8.    or if   the product-design project //NNM
9.        ‖ the product-design project he heads //WO
10   or if   the product-design project is scrapped . //PRD 】
```

Figure 8: NT clauses in English text.

4.2 Patterns constructing the English NT clause

Using the texts from the Wall Street Journal, we have built an English corpus containing nearly 2,000 English sentences, which were separated into about 6,000 NT clauses. These NT clauses are constructed through the Stack Pattern, the Back Position Pattern, the New Branch Pattern, the Influx Pattern, and the Enclosed Chunk Pattern. These five patterns account for more than 99% of the NT clauses in the corpus. In Figure 8, each line is an NT clause, except that line 1 and line 8 are only namings. Line 5 and line 7 are SV type of NT clauses; line 2, line 3, line 4, and line 10 are constructed according to the Stack Pattern, whereas line 6 and line 9 to the New Branch Pattern; from line 5 to 10, there is an Enclosed Chunk. In the following examples, NT clauses constructed through the Back Position Pattern and the Influx Pattern are illustrated.

Example 7 (*Wall Street Journal*)

As being so fragile and minute, they will require special robotic handling equipment .

The naming-telling relationship in this text can be illustrated by the NIS:

```
    As being so fragile and minute
they will require special robotic handling equipment .

they as being so fragile and minute
they will require special robotic handling equipment .
```

Figure 9: The Back Position Pattern in English.

In this figure, the preposition phrase in line 1 should take the subject "they" of line 2 as its naming.

Example 8 (*Wall Street Journal*)

Workers dumped large burlap sacks of the imported material into a huge bin, poured in cotton and acetate fibers and mechanically mixed the dry fibers in a process used to make filters.

The naming-telling relationship in this text can be illustrated by NIS:

```
Workers dumped large burlap sacks of the imported material into a huge bin
        poured in cotton and acetate fibers
        and mechanically mixed the dry fibers                        in  a process
                                                                ‖        used to make filters .

Workers dumped large burlap sacks of the imported material into a huge bin in a process used to make filters.
Workers poured in cotton and acetate fibers                                in a process used to make filters
Workers and mechanically mixed the dry fibers                              in  a process
                                                                ‖ a process used to make filters .
```

Figure 10: The Influx Pattern in English.

Line 1 and line 2 need to share the same tail of telling in line 3 "in a process used to make filters", thus the NT clause can be constructed by Influx Pattern, which is to connect the tail of the line 3 to the end of line 1 and line 2. At the same time, the last two lines need a supplementary naming from line 1's subject.

In our English corpus, 99% of NT clauses can be constructed by these five patterns.

5 The nature of the NT clause in Chinese and English texts and the explanation from a cognitive perspective

The following statistics are from:
- Chinese corpus, 187 long texts with 30,963 NT clauses;
- English corpus, 93 short texts with 4,819 NT clauses.

Except for the types of NT clauses listed above, these NT clauses include a few subject ellipsis clause and incompletable clauses. Subject ellipsis clauses can be supplemented with a subject, while for an incompletable clause, it is impossible to supplement missing components.

5.1 Broad Coverage

Within this study, the Chinese and English texts can be covered by NT clauses, i.e. a text can be dissected into sequences of NT clauses, each word belongs to at least one NT clause, and a text is the combination of NT clauses. In the Chinese corpus, there are 2,044 incompletable clauses, 547 subject ellipsis clauses, with a total of 2,591 clauses, which accounts to 8.3% of the NT clauses; in the English corpus, there are 32 incompletable clauses, accounting to 0.66% of NT clauses, without any subject ellipsis clauses.

From a cognitive perspective, naming in an NT clause is a target of cognition, while telling is the content of cognition. Therefore, the NT clause is a basic cognition unit. The fact that NT clauses cover the whole text means that complex cognition forms are made up of basic cognition forms.

5.2 Order-preservation

The original texts in both the Chinese and the English corpus are one-dimensional structures. The NIS is an intermediate form of dissecting original texts into NT clauses, and has a two-dimensional structure. This process keeps the word sequence in the original text intact. To be more specific, for any two words, such as A and B, if A is to the left of B in the original text, then in the NIS, A is either to the left of B or above B, or vice versa.

Both the speech stream and text stream of natural language are linear. The words enter short-term memory of the human brain by their original order. Long-distance adjustment of word order is very hard because the capacity of the short-term memory is very limited. Order-preservation suits the limitations of brain cognition.

5.3 Non-transgressiveness

Naming is topic by and large, and one naming can be explained by several tellings. Under most circumstances, if one new naming-telling relationship occurs, replacing the old naming-telling relationship, then the component in the old naming-telling relationship cannot be shared as a naming by another telling following the new relationship. In another words, the component in the old naming-telling relationship cannot build another naming-telling relationship which trespasses on the new naming-telling relationship.

It is acceptable to say "他衣服笔挺，面料很讲究，鼻梁上架着一副金丝眼镜。" (His clothes are immaculate, and made of fine material, with a pair of gold-framed glasses resting on the ridge of his nose.), but not "他衣服笔挺，鼻梁上架着一副金丝眼镜，面料很讲究。" (His clothes are immaculate, with a pair of gold-framed glasses resting on the ridge of his nose, made of high fine material.） This is because the naming "衣服" (clothes) in the first line cannot trespass on the naming-telling relationship of the "鼻梁上" (on the ridge of his nose) as a naming. Therefore, the naming "衣服" (clothes) cannot be shared by the third P-clause in the second example.
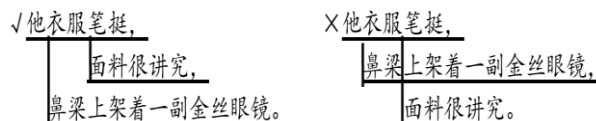

Figure 11: The non-transgressiveness of naming-telling relationship in Chinese.

Naming-telling relationship in English has the same property.

18

Example 9 (by the researcher)

It is acceptable to say "John has a son working in China, and Mary has a daughter studying in the USA.", but it is not acceptable to say "John has a son, and Mary has a daughter studying in the USA, working in China." This is because the naming "a son" cannot trespass the naming-telling relationship of the naming "Mary" and "daughter" to be shared by "working in China".
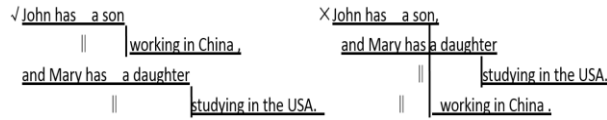

Figure 12: The non-transgressiveness of naming-telling relationship in English.

In our corpus of 30,963 Chinese NT clauses, there are 83 instances of trespassing, which accounts to 0.3%; 4,819 English NT clauses show only 4 instances of trespassing, accounting to 0.1%. 99% of both Chinese and English NT clauses do not trespass. For the few instances of trespassing NT clauses, they offer pragmatic background information. In example 10, the naming in line 1 and the telling in line 5 form a NT clause. This NT clause trespasses on the NT clause in line 4 which expresses the reason of the line 3.
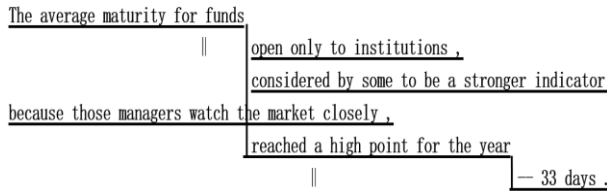
Example 10 (*Wall Street Journal*)

Figure 13: An example of trespassing on the naming-telling relationship.

The cognitive mechanism of the brain needs to process all tellings of a naming before going on to process the next naming and its tellings, so it is not likely for several tellings of different namings be processed all at the same time. The non-transgressiveness of the naming-telling relationship suits such constraint on cognition.

5.4 Syntactic well-formedness

Both in the Chinese or the English corpus, a part of NT clauses constitutes a grammatical SV structure, i.e., they are syntactically well-formed, and, while

another part of the NT clauses is not syntactically well-formed.

In our corpus, 1,963 Chinese NT clauses (out of 30,963) are not syntactically well-formed, which makes up for 6.3%; while 2,463 English NT clauses (out of 4,819) are not syntactically well-formed, which accounts for 51.1%.

A part of the Chinese NT clauses is not syntactically well-formed, because when a naming is connected with its one telling sharing it and constructs an NT clause, the punctuation between them is discarded, causing some semantic relationship markers to be missing or redundant. Adding or deleting such markers can ensure the syntactical well-formedness. For instance, in this sentence "邯郸市在河北省南部，京广铁路穿城而过。" (Handan is in the south of Hebei province, Jing-Guang railway passes through the city), the second P-clause corresponds to the NT clause "邯郸市京广铁路穿城而过" (Handan Jing-Guang railway passes through the city). The semantic relationship between "邯郸市" (Handan) and "京广铁路" (Jing-Guang railway) is not clear, resulting in an awkward clause. We can add "有" (in…there is) to express existential meaning, thus forming the clause "邯郸市有京广铁路穿城而过" (In Handan, there is the Jing-Guang railway passing through the city), which is syntactically correct. Since Chinese grammatical rules are not as strict as English ones, Chinese NT clauses have a lower ratio of ungrammaticality.

In English, the syntactical well-formedness problem in NT clauses is also due to the connection of naming and telling which may result in the missing and redundancy of some connective component. For example, if the telling is a nonrestrictive verb phrase, it usually needs to add an auxiliary "be" to ensure syntacticality. Since English is more rule-governed than Chinese, English has far higher proportion of syntacticality problems.

Even though both Chinese and English NT clauses may exhibit problems in syntactical well-formedness, based on the above analysis, these problems can be solved by adding or deleting some connective component in a finite range. Therefore, all English and Chinese NT clause are syntactical by and large. By this analysis, in human cognition, an NT clause can be processed simply as a com-

mon clause, this attribute meeting the feasibility of NT clause cognition.

5.5 Distribution of NT clause construction patterns

In section 3 and 4, we listed five patterns of the NT clause for both Chinese and English. Among them, the Enclosed Chunk Pattern represents relation between two groups of NT clauses. Therefore, the NT clause can be classified into following six categories: the original SV structure, constructed with a Stack Pattern, Back Position Pattern, New Branch Pattern, or Influx Pattern, and others.

According to statistics in our corpus, the numbers and the proportions of the six categories in Chinese and English are following:

Pattern	Chinese		English	
	No.	Per.(%)	No.	Per.(%)
Original	11970	38.70	2420	50.20
Stack	16449	53.10	515	10.70
Back Position	162	0.50	145	3.00
New Branch	1056	3.40	1723	35.80
Influx	650	2.10	13	0.30
Other	676	2.20	3	0.10
NT clause total	30963	100	4819	100

Table 1: The distribution of NT clause constructing patterns in Chinese and English.

All the NT clauses in the category of "other" are instances of ellipsis which cannot be covered by our patterns. Among the 676 Chinese NT clauses, 547 are subjectless, and 129 are without middle parts; in English, three instances are have no middle parts. All these ellipses are related to specific patterns that are not discussed further in the present study.

For both Chinese and English, there are only a few instances of "other" categories. The Stack Pattern, Back Position Pattern, New Branch Pattern and Influx Pattern cover the majority of non-original SV structure NT clauses. Thus, the four patterns are essential for the NT clause. Section 3 of this study presents the NT clause construction mechanism of the Stack Pattern. This pattern can be assumed as one kind of human cognition mechanism. In Chinese, it is not uncommon that one naming goes with dozens of tellings. A naming can

be connected with a telling which follows after dozens of P-clauses. This phenomenon is surprising, considering that the human brain has only 7 ± 2 units of short term memory. However, if the human brain does follow the Stack Pattern to process, old tellings are discarded constantly, and the naming is always stored in our short term memory to form a connection with the telling, then this phenomenon is no longer strange. Therefore, the Stack Pattern is a reasonable explanation of cognitive brain function for the NT clause. Similarly, the Back Position Pattern, New Branch Pattern and Influx Pattern can be explained in ways which are not covered in detail in this study.

As shown in the above table, both the Back Position Pattern and Influx Pattern have a low frequency in either corpus. Considering this from a cognitive perspective, the Back Position Pattern cannot complete the cognition of the NT clause right after the occurrence of a telling; instead, the telling has to be stored temporarily to await its naming. At the same time, in the Influx Pattern, the front part of the telling occur first, and this part also has to be stored first to await its later part. These two processing patterns of NT clauses both pose high demands on time and space, and therefore their low frequency in the corpora is perhaps unavoidable.

As shown in the table, both the Chinese and English NT clause demonstrate different frequencies for these four patterns. The Chinese NT clause shows higher proportions for the Stack Pattern and the Influx Pattern than English, while the New Branch Pattern and the Back Position Pattern are lower than in English. This frequency difference is attributed to grammatical difference. Taking the Stack Pattern as an example, the Chinese expression usually moves from bigger concepts to smaller concepts, which is suitable to the progressively layered structure of the Stack Pattern. In contrast, the English expression moves from smaller concepts to bigger concepts, which is not suitable for the Stack Pattern. To take another example, for the New Branch Pattern, the English NT clause takes relative clause, preposition phrase or non-restrictive verbal phrase as its telling, and usually they have obvious markers causing less cognitive burden on the brain; however, Chinese does not have similar markers, so in most cases, it has to re-

ly on semantic coordination to distinguish its meaning, and thus exhibits a heavier cognitive load.

6   Conclusion

This study puts forward the concepts of NT clause, dividing linear texts into a two-dimensional plane structure, and drawing up five patterns for constructing the NT clause. Through the annotation of corpus, this study discovered that both Chinese and English texts are formed through combinations of NT clauses according to these five patterns. The construction of the NT clause is order-preserving, non-transgressive, and syntactically well-formed. These features of the NT clause can be explained by appealing to cognitive mechanisms.

In the fields of linguistics and computational linguistics, the theories and methods of clause complex, topic and zero anaphora are related to our study. The kernel of our contribution is the discovery of formal patterns for constructing the NT clause. Based on these patterns, a syntactic platform suiting to both Chinese and English can be built. On the platform the MT of long sentences is decomposed to three steps: NT clause parsing in source language, translation of NT clauses, and assembling the NT clauses in target language. These steps are independent of each other. In this procedure, the first and third steps involve only one language and only the second step involves two languages. Moreover, the NT clauses handled in the second step have smaller scale and simpler structures than long sentence, so that they may be translated easier. This method, expected to ease the difficulties of long sentences in MT, will be discussed in another paper.

Acknowledgments

References

Carnie, Andrew. 2013. *Syntax: A Generative Introduction*, 3rd ed. Blackwell Publishing Ltd., West Sussex.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*, volume 34 of Syntax and Semantics. Academic Press, New York.

Halliday, M.A.K. 1994. *Introduction to Functional Grammar*, 2nd ed. Edward Arnold, London.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. University of Chicago Press, Chicago.

Shen, Jiaxuan. 2013. Can predicates be nominals? *Foreign Studies*, 2013, 1(1): 1-13.

Song, Rou. 2013. Stream model of Generalized Topic Structure in Chinese text. *Studies of the Chinese Language*, 2013(6): 483-494.