

杜佩芬 國立中正大學語言學研究所

論文題目(中文)：中文關係子句內外部因素對語句處理之影響

論文題目(英文)：Internal and External Factors in the Processing of Chinese  
Relative Clauses

指導教授：戴浩一、麥傑

## 1. Contributions

The multiple factors in this thesis were classified into internal and external factors for Chinese RC processing. This particular classification is more useful to help not only explore multiple factors in the online and offline experiments, but also efficiently distinguish the default syntactic and semantic structures that affect the processing of Chinese RCs. The following are the main contributions in this thesis:

First of all, as indicated above, this classification provides a novel way to explore the factors for researchers in the processing of Chinese RCs. A large number of previous literatures have paid attention on the extraction types, which were defined as an internal factor in this thesis, though it is crucial factor on processing RCs. However, the few studies examined the effects of external factors such as branch direction and the position of *Nage* in the processing of Chinese RCs. Importantly, this classification gives researchers a more holistic picture to investigate Chinese RCs since the internal and external factors can be simultaneously tested. Thus, researchers are in a better position to observe their interaction and underlying relations among these factors and to account for the complexity of Chinese RC parsing.

Secondly, based on the characteristics of language typology, there are different structures of relative clauses between English and Chinese (i.e. the branching direction). This thesis proposes a cross-linguistic comparison of RC processing between both languages. Regardless of branching length (short vs. long RCs), it demonstrates that there is a predominance of processing left-branching RCs in Chinese, whereas there is a predominance of processing right-branching RCs in English (Kemper, 1986; Small et al., 2000). This finding would be the most valuable contribution of this thesis, and also verifies that the factor of branching direction plays an important role in the processing of Chinese RCs.

For the third contribution, the previous corpus-based studies have been taken the initiative in analyzing the asymmetric structures of *Nage* constructions in discourse, and provided the explanations of pragmatic functions. Nevertheless, the factor of *Nage* position was first evaluated in a self-paced reading task. The advantages of this experimental method not only combine the online and offline measures, but also help shed light on interpreting experimental data in which the RC structures show the default of syntactic and semantic linkings between the canonical and non-canonical thematic orders in Chinese. Thus, this novel method may give a straightforward way to investigate how individuals tackle these

different constructions of Chinese RCs, and also provide a reflection for researchers to carefully manipulate this factor when using *Nage* as their experimental materials.

Last but not least, this thesis not only reviewed the corpus-based analyses from the main previous literatures, but also conducted the behavior experiments to investigate the grammatical structures of Chinese RCs. The corpus-based studies provide the empirical evidence how people use language in the real world, whereas psycholinguistic experiments reflect individuals' language process. Interestingly, all experimental results in this thesis were compatible with the previous corpus-based analyses. That is, all factors manipulated in present experiments show the structural preference in the previous corpus-based studies. Thereby, the integration of these two approaches gives us insight into how they interact with each other, as well as brings us a more comprehensive view of Chinese RCs to facilitate linguistic research.

## **2. Abstract**

This thesis aims to examine the multiple factors that affect the processing of Chinese relative clauses (RCs). These factors were classified as internal and external factors according to whether the syntactic items locate inside RCs or not. Two internal factors, which includes extraction types (subject-extracted RCs vs. object-extracted RCs) and branching length (short vs. long RCs), and two external factors, which includes branching direction (left- vs. right-branching) and position of demonstrative + classifier *Nage* (pre- vs. post-RC), were determined in this thesis. A self-paced reading task was conducted in two experiments to observe participants' online sentence processing from the reading time data and offline performance from the accuracy rates of comprehension questions.

The focus of Experiment 1 was the interaction among branching direction, branching length, and extraction types. Kemper (1986) and Small et al. (2000) propose that English right-branching structure, a RC branches from object noun phrase (i.e. right position) in relation to matrix verb in matrix clause, is easier to process than left-branching structure, a RC branches from subject noun phrase (i.e. left position) in relation to matrix verb in matrix clause. Their findings were attributed to a memory resource capacity theory (Just & Carpenter, 1992; Haarmann, Just, & Carpenter, 1997) that the left-branching RCs interrupt information of their matrix clause, and this resulted in increasing participants' additional memory demands on storage and computation when processing sentences. However, based on the above definition of the branch direction (left vs. right), Chinese left-branching structures, which are sentence-initial RC, do not interrupt the information of their matrix clause, but right-branching structures do. In addition, the corpus-based studies (Pu, 2007; Wu, 2009) also show that the left-branching RCs outnumber the right-branching RCs in Chinese corpora. Thus, we predict that Chinese left-branching RCs are easier to process than right-branching RCs. The results demonstrate a predominance of processing left-branching

RCs in Chinese from online and offline measures. The findings were not only consistent with the previous corpus-based studies, but also showed the opposite views of processing RCs between English and Chinese. That is, a predominance of processing right-branching RCs was found in English (Kemper 1986; Small et al. 2000), whereas a predominance of processing left-branching RCs was found in Chinese. Both findings support the memory resource capacity theory.

On the other hand, Kemper (1986) also proposes that short-branching RCs are easier to process than long-branching RCs (i.e. RC + a temporal adverbial + PPs) in English. Thus, Experiment 1 also manipulated Chinese branching length: short (i.e. a temporal adverbial + RC) vs. long branching (i.e. two temporal adverbials + PPs + RC). However, it shows inconsistent results between on-line and off-line measures. For the online data, it reveals that the participants read long-branching RCs faster than short-branching RCs, and the findings support the anti-locality hypothesis and memory facilitation hypothesis that the long-branching RCs with more specific information facilitate RC processing, particularly at the critical region, head noun. By contrast, for the off-line data, it reveals that the short-branching RCs were recalled more accurately than the long-branching RCs, and it seems to reflect the locality-based hypothesis and the memory resource capacity theory that the long-branching RCs may increase participants' additional memory cost (or integration) when immediately recalling.

The focus of Experiment 2 was the relationship between *Nage* position and the extraction types. This experiment was mainly aimed to observe whether there is difference of RC processing between *Nage* in pre- and post-RCs. Sheng & Wu (2012) claim that *Nage* prefers occurring before subject-extracted relative clauses (SRCs) and after object-extracted relative clauses (ORCs) in their corpus-based analyses and sentence production experiment. Their findings were interpreted within the Audience Design Model that speakers favor to provide an early cue *Nage* for their listeners in pre-SRCs to make the ease of comprehension; by contrast, the speakers favor to use *Nage* in post-ORCs to avoid structural ambiguity where *Nage* occurs in pre-ORCs. However, beyond the above pragmatic interpretations, we propose that there are default syntactic and semantic linkings in Chinese (i.e. the canonicity of thematic order AGENT-ACTION-PATIENT) to directly affect real-time processing (Small et al., 2000; Lin, 2012). Specifically, the structures of *Nage*\_SRCs and ORCs\_*Nage* are compatible with Chinese canonical thematic order, whereas the structures of SRCs\_*Nage* and *Nage*\_ORCs are incompatible with this order. Thus, we predict that *Nage*\_SRCs are easier to process than SRCs\_*Nage*, whereas ORCs\_*Nage* are easier to process than *Nage*\_ORCs in self-paced reading task. Our results demonstrate that there is a predominance of processing *Nage*\_SRCs and ORCs\_*Nage* at the two critical regions, RC and head noun. The findings were also in line with the corpus-based research (Shang & Wu, 2012). It suggests that there was a delayed thematic role assignment in SRCs\_*Nage* and *Nage*\_ORCs, and this delay led

participants to increase additional processing costs when reading sentences.

By integrating the previous corpus-based analyses into the experimental data, the research in this thesis not only discovers the multiple factors that have significant effects on the processing of Chinese RCs, but also provides a more comprehensive view of Chinese RCs for researchers to build linguistic production and/or comprehension models from the empirical evidence.